THE MISSING ANTIMATTER • WHAT THE MILKY WAY HIDES • THE SLIME EEL

SCIENTIFIC AMERICAN

OCTOBER 1998 $4.95

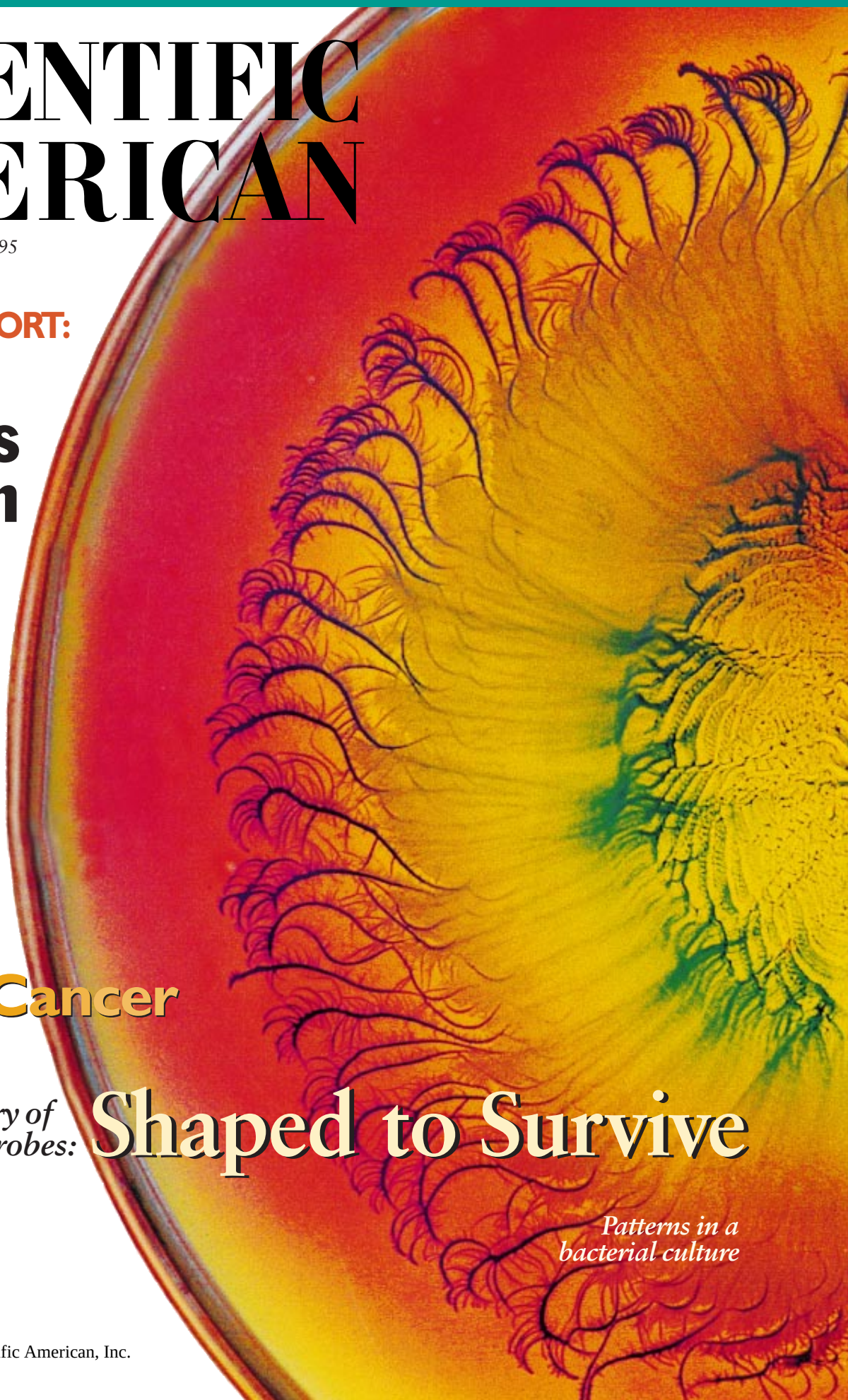
SPECIAL REPORT:

How Hackers Break In

Keep networks and data safe from Internet spies

Drugs that Prevent Breast Cancer

The Artistry of Microbes: **Shaped to Survive**



Patterns in a bacterial culture

SCIENTIFIC AMERICAN

October 1998 Volume 279 Number 4

About the Cover
Feathery shapes in a culture of *Bacillus subtilis* result from the microbe's swift but off-center movements in this soft medium. Photograph by Eshel Ben-Jacob.

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111.

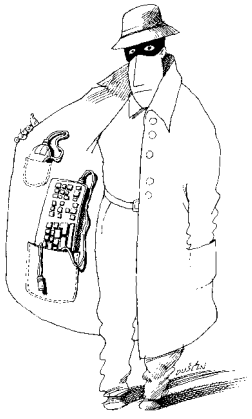
Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 242764. Canadian BN No. 127387652RT; QST No. Q1015332537. Subscription rates: one year $34.97 (outside U.S. $49). Institutional price: one year $39.95 (outside U.S. $50.95). Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. Reprints available: write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; fax: (212) 355-0408 or send e-mail to sacust@sciam.com **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.**

Learning from the Hackers

You've found a vulnerability in certain widely used e-mail programs that could let vandals wreck computers by remote control. Announcing your discovery will allow programmers to fix this problem, but it will also tip off would-be saboteurs. To publish or not to publish?

This past summer some experts faced that dilemma. They wisely chose to tell the world. Fixes were written and distributed quickly, and as of this writing, no one seems to have exploited the weakness.

We faced a similar decision over Carolyn P. Meinel's article on page 98, "How Hackers Break In … and How They Are Caught." Meinel describes how a fictional hacker might penetrate a corporation's computer system. Is publishing it irresponsible? Obviously, we think not. Improving the security of networked computers is crucial. We can best inform readers about how to defend themselves by explaining what attacks to expect.



HIGH-TECH BURGLARS don't need to be programmers.

Serious hackers already know these secrets. Anyone who wants to know how to crack a system can get all the advice he (or, rarely, she) needs on Web sites and bulletin boards. The software equivalents of crowbars and lockpicks are available on-line. Hackers don't need to be programmers these days any more than burglars need to be architects.

And cracking a system doesn't take a criminal mastermind when the authorized users are locking the front door with masking tape and string. Every person on a network who chooses an obvious password or, worse, patches in an unguarded phone line is shaving years off the life of some poor system administrator.

Here is how mainstream hacking has become: thousands of hackers gather in Las Vegas every summer for a meeting called Def Con. (Reality check: subversive groups don't hold annual conventions in Vegas.) Luckily, most hackers are more curious and adventurous than malicious and so are willing to share their knowledge of the Internet's soft underbelly. Smart corporations, law enforcers and the military are listening.

We all should be. Vulnerability to hacking is not a passing phase. No matter how strong the firewalls around systems, some people will always try to break in—and administrators will retaliate with stronger walls.

Vigilance and prudence can keep malicious hacking in check. Reading our special report on computer security and the Internet is a good way to start. Then think about changing your passwords—but for heaven's sake, stay away from birthdays, J.R.R. Tolkien characters and *Star Trek* references!

JOHN RENNIE, *Editor in Chief*
editors@sciam.com

DUSAN PETRICIC

SCIENTIFIC AMERICAN®
Established 1845

John Rennie, EDITOR IN CHIEF

Board of Editors
Michelle Press, MANAGING EDITOR
Philip M. Yam, NEWS EDITOR
Ricki L. Rusting, ASSOCIATE EDITOR
Timothy M. Beardsley, ASSOCIATE EDITOR
Gary Stix, ASSOCIATE EDITOR
W. Wayt Gibbs, SENIOR WRITER
Kristin Leutwyler, ON-LINE EDITOR
Mark Alpert; Carol Ezzell; Alden M. Hayashi; Madhusree Mukerjee; George Musser; Sasha Nemecek; David A. Schneider; Glenn Zorpette
CONTRIBUTING EDITORS: Marguerite Holloway, Steve Mirsky, Paul Wallich

Art
Edward Bell, ART DIRECTOR
Jana Brenning, SENIOR ASSOCIATE ART DIRECTOR
Johnny Johnson, ASSISTANT ART DIRECTOR
Bryan Christie, ASSISTANT ART DIRECTOR
Dmitry Krasny, ASSISTANT ART DIRECTOR
Bridget Gerety, PHOTOGRAPHY EDITOR
Lisa Burnett, PRODUCTION EDITOR

Copy
Maria-Christina Keller, COPY CHIEF
Molly K. Frances; Daniel C. Schlenoff; Katherine A. Wong; Stephanie J. Arthur; Eugene Raikhel; Myles McDonnell

Administration
Rob Gaines, EDITORIAL ADMINISTRATOR
David Wildermuth

Production
Richard Sasso, ASSOCIATE PUBLISHER/ VICE PRESIDENT, PRODUCTION
William Sherman, DIRECTOR, PRODUCTION
Janet Cermak, MANUFACTURING MANAGER
Tanya Goetz, DIGITAL IMAGING MANAGER
Silvia Di Placido, PREPRESS AND QUALITY MANAGER
Madelyn Keyes, CUSTOM PUBLISHING MANAGER
Norma Jones, ASSISTANT PROJECT MANAGER
Carl Cherebin, AD TRAFFIC

Circulation
Lorraine Leib Terlecki, ASSOCIATE PUBLISHER/ CIRCULATION DIRECTOR
Katherine Robold, CIRCULATION MANAGER
Joanne Guralnick, CIRCULATION PROMOTION MANAGER
Rosa Davis, FULFILLMENT MANAGER

Business Administration
Marie M. Beaumonte, GENERAL MANAGER
Alyson M. Lane, BUSINESS MANAGER
Constance Holmes, MANAGER, ADVERTISING ACCOUNTING AND COORDINATION

Electronic Publishing
Martin O. K. Paul, DIRECTOR

Ancillary Products
Diane McGarvey, DIRECTOR

Chairman and Chief Executive Officer
John J. Hanley

Co-Chairman
Rolf Grisebach

President
Joachim P. Rosler

Vice President
Frances Newburg

SCIENTIFIC AMERICAN, INC.
415 Madison Avenue
New York, NY 10017-1111
(212) 754-0550

PRINTED IN U.S.A.

LETTERS TO THE EDITORS

One story—really, one sentence—in the June issue irritated many readers. In "Millennium Bug Zapper" in the News and Analysis section, staff writer Alden Hayashi remarked that there is "only a year and a half left until the new millennium"—meaning, of course, that 2000, not 2001, marks the historic turning point. John Rabold of Oakland, Calif., wrote in exasperation, "Has even SCIENTIFIC AMERICAN capitulated to mass opinion about the year 2000?" In reply, we might cite Stephen Jay Gould's recent book *Questioning the Millennium*. Gould predicts that, at least this time, the popular view (the new millennium begins in 2000) will win out over the purist view (it starts in 2001). Gould's book offers several rationales in support of marking the transition in 2000: that the first decade of the first century had only nine years, or that the year 1 B.C. is equivalent to A.D. 0, so there really was a year 0. For people who find these explanations flippant, Gould writes, "Arbitrary problems without conceivable final answers require consistent but arbitrary solutions." The rest of the issue also prompted interesting, though not quite so heated, comments from readers.

DEALING WITH DEPRESSION

Charles B. Nemeroff provided an interesting and thought-provoking article in "The Neurobiology of Depression" [June]. Without question, our understanding is advanced by research into the biology of this tragic condition, and better treatments are almost sure to follow. But Nemeroff's statement that "often psychotherapy is needed [as a treatment] as well, but it usually is not sufficient by itself, especially if the depression is fairly severe" may lead readers to conclude that psychotherapy is less effective than medication and that their combination provides a better outcome than either treatment alone.

This space is clearly not the forum to debate the relative merits of therapies, but the bulk of the evidence does suggest three important conclusions. Psychotherapy and medication are both helpful in the treatment of depression; there is no clear evidence that one is superior to the other. And despite what common sense would dictate, a combination of psychotherapy and medication does not appear to be more effective than either alone. Finally, neither treatment is effective enough—a large percentage of patients receive little or no benefit from either method, resulting in unacceptably high relapse rates. It is, of course, the third conclusion that really matters. It is through the research efforts of Nemeroff and his colleagues that widely effective, durable treatments for depression will be found.

DIRK ELTING
Nebraska Health and Human Services System
Norfolk, Neb.

Nemeroff replies:

I appreciate Elting's comments on my article. I agree that space constraints preclude a comprehensive discussion of the use of antidepressants and psychotherapy alone or in combination. I would like to make a few points, however. First, there is increasing evidence that both psychotherapy and pharmacotherapy are biological interventions in the sense that they produce changes in brain function. Second, my reading of the limited literature suggests that the most severe forms of depression—psychotic depression and melancholia—respond best to pharmacotherapy. Finally, there is evidence that the combination of pharmacotherapy and psychotherapy is more effective than either method alone. Elting's point that many patients do not respond to any treatment is indeed true and will be the subject of a large study, sponsored by the National Institute of Mental Health, currently being planned.

NAJLAH FEANNY SABA

TREATING DEPRESSION can involve either psychotherapy or medication, or a combination of both.

EFFECTS OF AQUACULTURE

Because my name was associated with the June article "Shrimp Aquaculture and the Environment," by Claude E. Boyd and Jason W. Clay (I took several of the photographs that appeared in the article), I want to respond to some of the authors' points. The article concentrates mainly on the technical shortcomings of shrimp aquaculture as well as a few environmental consequences. All the technical fixes in the world won't amount to much if the rights of local populations are compromised. Thousands of people have been forced from their land by the encroachment of shrimp aquaculture in Asia and Latin America, in part because of salinated drinking water, failing agricultural croplands, declining fisheries and mounting environmental degradation. Certainly, technological fixes are needed. But issues of social justice are more complicated and yet most important to solve now.

ALFREDO QUARTO
Co-director, Mangrove Action Project
Port Angeles, Wash.

ALCOHOL THROUGH THE AGES

While reading the June article by Bert L. Vallee entitled "Alcohol in the Western World," I came across the statement that "Western civilization has wine and beer to thank for nourishment and hydration during most of the past 10,000 years." But in my high school class this year on human anatomy and physiology, I learned that alcohol dehydrates you. Thus, people who are planning to drink a lot of alcohol also need to drink a lot of water.

BLAKE GOUD
West Bethesda, Md.

Although it is certainly true that there is considerable biblical evidence for consumption of wine, Vallee hasn't looked as closely as he should for evidence of consumption of water. He states that "both the Old and New Testaments are virtually devoid of references to water as a common human beverage." Yet I found 15 references to consumption of water in less than an hour—including Genesis 21:14, "Abraham took some food and a skin of water and gave them to Hagar." And John 4:7 reads, "When a Samaritan woman came to draw water, Jesus said to her, 'Will you give me a drink?'" (Both passages are quoted from the New International Version.)

MARTIN LABAR
Southern Wesleyan University

The Editors reply:

Goud is getting a good education—drinking alcohol certainly dehydrates you. As noted in the article, however, "alcoholic drinks were diluted with the sullied water supply." The amount of water mixed with alcohol more than made up for any dehydrating effect. As for the biblical references, although there are passages concerning water fit for drinking, in many of them clean water is held in high esteem, suggesting that it was hard to find. For instance, Revelation 22:1 reads, "Then the angel showed me the river of the water of life, as clear as crystal, flowing from the throne of God and of the Lamb."

LOSS OF POWER

Regarding the "Cracking a Combination" caption on page 71 in "Quantum Computing with Molecules," by Neil Gershenfeld and Isaac L. Chuang [June], shouldn't it read, "On average, an n-bit lock requires $(2^n)/2$ tries" before you stumble on the correct combination (rather than just $^n/_2$)?

MATT FANTE
Annapolis, Md.

Editors' note:

Fante is correct. We apologize for the confusion.

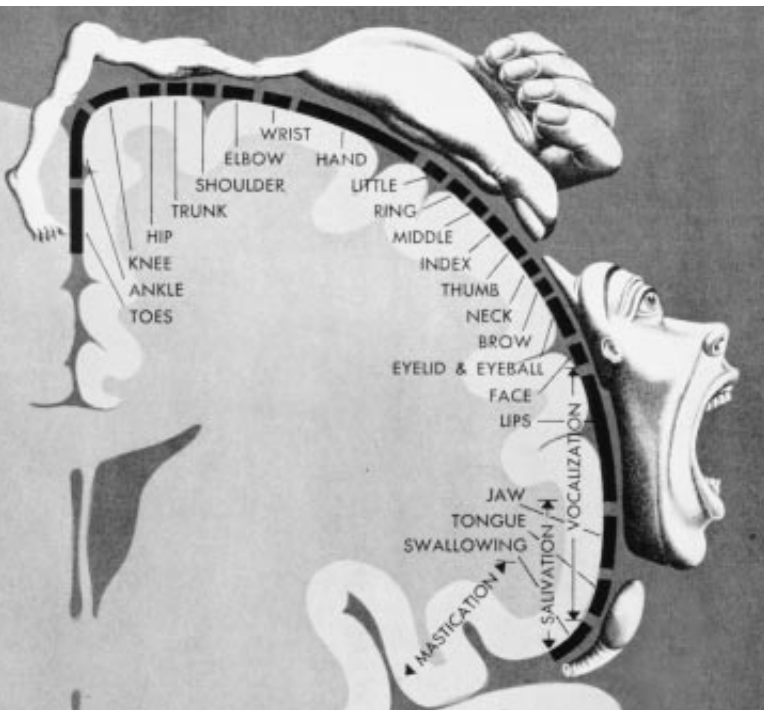
Letters to the editors should be sent by e-mail to editors@sciam.com or by post to Scientific American, 415 Madison Ave., New York, NY 10017. Letters may be edited for length and clarity.

50, 100 AND 150 YEARS AGO

SCIENTIFIC AMERICAN

OCTOBER 1948

RED SCARE—"During the past year, the federal government has found scientists and technical personnel increasingly unwilling to accept or continue in government jobs. For instance, all but 15 of the 150 outstanding scientists engaged in wartime atomic research have quit the government since the end of the war. This situation is due (in the words of President Truman) to 'attacks on scientists in the ostensible name of security.' The president denounced 'the creation of an atmosphere in which no man feels safe against the public airing of unfounded rumors, gossip and vilification' as 'the most un-American thing we have to contend with today.'" [*Editors' note: Truman himself was partly to blame for this atmosphere by ordering loyalty investigations of all government employees in March 1947.*]



Brain control of motor functions

PSYCHOSURGERY—"Based on the initial work of Egaz Moniz and Almeida Lima in Portugal, 2,000 persons in North America have been operated on by leucotomy [now termed lobotomy] or related techniques. This severing of the connecting fibers between the prefrontal region and the thalamic center in the brain stem apparently releases the 'new' brain of the prefrontal region from the emotional dominance of the 'old' brain of the cerebral stem. Whatever the mechanisms involved, there have been amazing transformations of violently insane persons into seemingly normal ones. Adverse changes can also result, however, which can cause a deterioration in personality."

MOTOR HOMUNCULUS—"Our illustration shows a schematic interpretation of the brain's motor area, projected on a cross section of a hemisphere. Each of the areas outlined by this grotesque manikin is devoted to sending impulses to the corresponding part of the body. Parts of the homunculus are enlarged or diminished in proportion to how much the related part of the body is used. The drawings are from Dr. Wilder Penfield's monograph 'The Cerebral Cortex of Man.'"

OCTOBER 1898

ETHEREAL CHEMISTRY—"Mr. Charles F. Brush read a very important paper before the American Association for the Advancement of Science, in which he describes extracting from the atmosphere a gas which is lighter than hydrogen. The new substance has been called 'etherion.' Mr. Brush says that the ability of etherion to conduct heat is fully a hundred times as great as that of hydrogen. He also considers that the gas reaches out indefinitely into space."

DIRIGIBLE BALLOON—"M. De Santos-Dumont, a well-known Parisian sportsman, made a highly interesting experiment with an aerostat. It is a cylinder tapered at both ends, is 82 feet long and is made of extra-light Japan silk rendered waterproof. The weight of the balloon, car, engines and rudder is 114 pounds. The motor is of the kind usually found on automobile tricycles, provided with two superposed cylinders. This is said to be the first time that motors of this type have been used on aerostats. The aeronaut followed a course toward the Bois de Boulogne at an altitude of 650 feet before the aerostat began losing its rigid form and he was forced to land."

OCTOBER 1848

MODERN BAROMETER—"A new barometer without the use of alcohol or mercury has lately been exhibited in London and which is said to be a simple, beautiful, and accurate indicator of atmospheric changes on an entirely novel principle [using the action of air pressure on a diaphragm covering an evacuated chamber]. It is termed by the inventor, a French gentleman, the Aneroid Barometer."

FLORIDA WRECKS—"An average of a million dollars' value is annually wrecked on the Florida Reefs and Keys, for the want of an accurate chart of that coast. Although Florida has been held by the United States for twenty-seven years, no original American chart has ever been made of its dangerous coast. Navigators have to depend upon old Spanish charts, and those made by the British from 1763 to 1784."

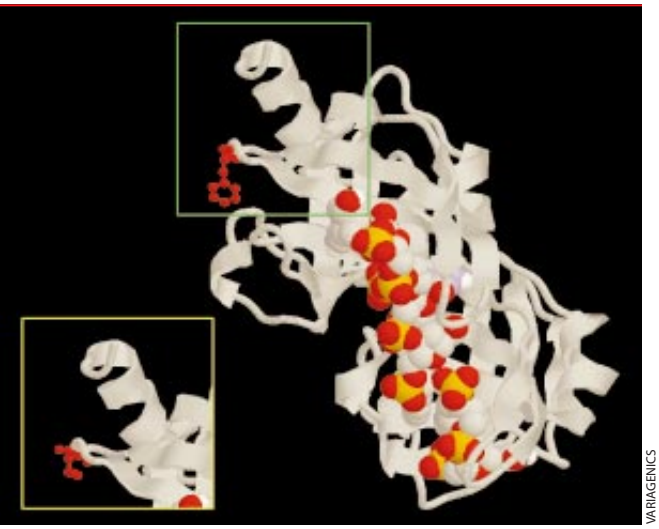
FOR MEDICINAL PURPOSES—"Dr. Rennes, of Bergerac, advises that leeches should be put for an instant into weak wine-and-water, the better for being a little warm, just before applying them; no sooner are they laid on than even the most sluggish pierce the skin instantly; those even that had been for a short time before used immediately attach themselves."

NEWS AND ANALYSIS

IN FOCUS

PERSONAL PILLS

Genetic differences may dictate how drugs are prescribed



VARIAGENICS

SINGLE NUCLEOTIDE CHANGE *alters a protein's amino acid from phenylalanine* (outlined in green) *to serine* (yellow outline), *an example of the type of genetic variation that can affect drug action.*

Even before the human genome is fully decoded, academic and industry researchers have begun to take the next step: comparing how genetic information varies from individual to individual. The databases compiled from these endeavors will provide a record of human migrations and will show how multiple genes contribute to common diseases. But biotechnology and pharmaceutical companies also want to use this knowledge to tailor drugs to certain groups of patients. A customized pharmaceutical might eliminate life-threatening adverse reactions. And knowing how genetically distinct individuals react differently to a certain compound may reduce the cost of clinical trials by targeting only those patients capable of responding to a drug.

Pharmacogenomics is the term that has evolved to describe the use of advanced genetic tools to elucidate how variations in patients' DNA may diminish or amplify drug effects or render a pharmaceutical toxic. Earlier this year an article in the *Journal of the American Medical Association* estimated that adverse drug reactions accounted for more than two million hospitalizations and more than 100,000 deaths in 1994, making them a leading cause of mortality in the U.S.

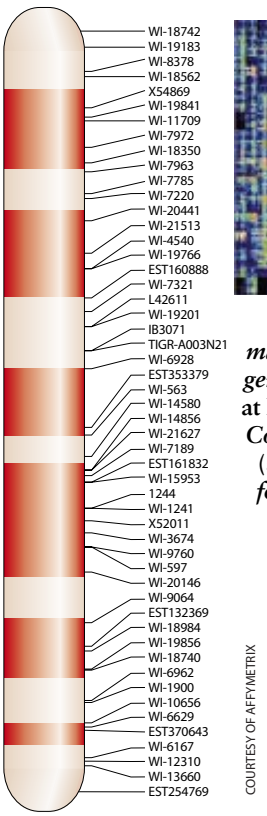
Many of the ideas that underlie pharmacogenomics are not new. It has been understood for decades that genes affect the way patients respond to drugs. For instance, pharmaceutical researchers sometimes look at how differences in the genes for liver enzymes called cytochrome P450 affect how patients metabolize a new drug candidate. But until now the genes one could study for such variations were few in number. The tools for rapidly compiling large compendiums of the minute variations in nucleotides (DNA bases) are of recent vintage.

Indeed, a race is under way to catalogue genetic variations among these single DNA bases, known as single nucleotide polymorphisms (SNPs, pronounced "snips"), which can be used in characterizing drug responses. The National Institutes of Health has launched a $36-million, three-year program to collect data on 50,000 to 100,000 SNPs, a new goal for its Human Genome Project. The information would be used not only to gauge drug responses but also to study disease susceptibility and to conduct basic research on population genetics. In midsummer a group of pharmaceutical companies discussed forming a consortium with the NIH that would supply

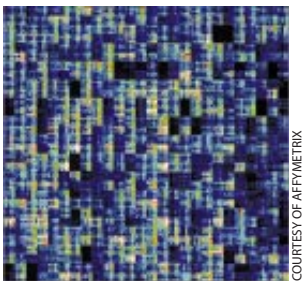
additional funding and research resources to create an even larger public database. One impetus to establish a tie with industry has been a concern that private attempts to patent SNPs could choke off access to data for basic research. "These research tools are far upstream of any particular product," notes Francis S. Collins, who oversees the Human Genome Project at the NIH. "The public is best served by having them accessible to any researcher who wants to use them."

A pharmaceutical industry collaboration with the NIH would promote public access to SNPs. Still, some biotechnology companies have rushed to embrace pharmacogenomics by creating private databases. A French company, Genset, is testing the DNA of more than 100 people to develop a map of the entire human genome. The Genset map will contain 60,000 SNPs that are within or near genes that cause disease or differing drug reactions. Genset's chief genomics officer, Daniel Cohen, devised the first rough physical map of the human genome in 1993.

Abbott Laboratories, a major U.S. pharmaceutical manufacturer, has invested $20 million in Genset. The companies will market SNP map data to drug companies that wish to pinpoint during clinical trials a common set of variant nucleotides shared by people who do not respond to a drug. This information could then be used to create diagnostic tests to filter out unresponsive patients. Abbott, in fact, is paying Genset an additional $22.5 million to help it develop a diagnostic test to screen patients for zileuton, its own asthma drug, which can induce liver toxicity in 3 percent of patients. Genset is not the only one putting together SNP databases. In August, Incyte Pharmaceuticals announced plans to purchase Hexagen in Cambridge, England, as part of its effort to detect genetic variation.



COURTESY OF AFFYMETRIX


COURTESY OF AFFYMETRIX

DNA CHIP ***made by Affymetrix finds genetic variations*** **(labeled at left)** ***on chromosome 12. Colored areas on the chip*** **(above)** ***represent SNPs found throughout a full set of chromosomes.***

The application of rapid tools for screening SNPs may eventually make it possible to look for the unique signature of an individual's DNA in a matter of hours. Traditional gene-sequencing technology might take two weeks and $20,000 to screen a single patient for variations in 100,000 SNPs. "That's going to make this prohibitive to put into a clinical-trial kind of system," noted Robert Lipshutz of DNA chipmaker Affymetrix at the annual meeting of BIO, a biotechnology industry trade group. Affymetrix is testing a chip that can detect 3,000 SNPs in less than 10 minutes. As the technology progresses, Affymetrix expects to be able to mill through 100,000 SNPs dispersed through a patient's genome in several hours, for as little as a few hundred dollars.

Not everyone wants to assess patient drug response by scanning the entire genome. Variagenics in Cambridge, Mass., selects a few target genes thought to be associated with drug responses for a given disease, a more established approach intended to speed assessment of drug safety and refinement of diagnostic tests. "Genes involved in drug action are over-represented among the genes whose sequences are already known," says Fred D. Ledley, the company's chief executive. To locate SNPs, an enzyme called resolvase scans the selected genes. It cuts the DNA when it finds a nucleotide that differs from a reference sequence. Using these data allows investigators to glean the genetic profile of patients who experience ill effects from a drug. One of Variagenics's goals is to improve the prescription of existing drugs. It is fashioning a test that will let physicians adjust the dosage of a widely prescribed cancer drug, 5-fluorouracil, that produces severe gastrointestinal side effects in some patients.

Before genetic profiling for drug prescriptions becomes routine, pharmacogenomics must overcome other obstacles. Individualizing pharmaceuticals may not necessarily sit well with big pharmaceutical companies, which are constantly in search of blockbuster drugs to offset multimillion-dollar development costs. A drug tailored to a specific subpopulation may fragment and diminish markets. Several drugs may be needed for a given condition, one for each genetic subtype. This strategy might still work if a manufacturer can charge enough for each drug. The real push toward pharmacogenomics may be driven by managed health care. A diagnostic test, even if it does add cost, could avoid the expense of today's trial-and-error methods of making multiple doctor's visits to have a prescription adjusted.

The hazards of placing patients in subgroups has not gone unnoticed. Without safeguards, health insurance providers might deny coverage to those with a certain genetic profile—patients for whom a drug is too expensive or for whom there is no treatment. The Human Genome Project's Ethical, Legal and Social Implications program will make the use of information about genetic variation its "number one priority" during the next five years, Collins says. "When you're cataloguing large numbers of SNPs on large numbers of people, it greatly accelerates the potential for this information to be misused in discriminatory ways," he remarks.

And according to one biotechnology industry leader, pharmacogenomics may simply be an ill-chosen approach to designing new drugs. William A. Haseltine, chairman and chief executive of Human Genome Sciences, asserts that pharmaceutical companies should be using genetic technologies to find the safest possible drug, not trying to save failed candidates by targeting them to selected patients. Diagnostic tests can be unreliable, he notes, and some patients could still sustain life-threatening reactions. Moreover, the multiple genes involved in a drug reaction can be hard to decipher. Environmental factors—food, other drugs ingested, a patient's gender and overall state of health—may account for much of how someone responds to a drug. "You've got to consider the whole person when using a drug," Haseltine says. "The pharmacogenomic argument is very similar to the sociobiology argument that everything is in the genes, when it is not." Debate may never fully settle. Technology that can identify a patient's distinctive genetic profile—and thus alter the way drugs are prescribed—may always prove contentious. *—Gary Stix*

CLIMATE CHANGE

IN THE HEAT OF THE NIGHT

Warmer nights may be slowing tropical forest growth and raising carbon dioxide levels

Researchers working in Costa Rica have discovered disturbing evidence that increasing temperatures have markedly slowed the growth of tropical trees over the past decade. The slowdown may explain calculations suggesting that tropical forests, which are usually considered to take up carbon dioxide, have actually added billions of tons of the greenhouse gas to the atmosphere each year during the 1990s, making them a huge net source, comparable in size to the combustion of fossil fuels. The trend could exacerbate global warming: as the mercury rises, tropical forests may dump yet more carbon dioxide into the atmosphere, causing still more warming.

In 1984 researchers Deborah A. Clark and David B. Clark of the University of Missouri, collaborating with Charles D. Keeling and Stephen C. Piper of the Scripps Institution of Oceanography in La Jolla, Calif., began measuring the growth rates of scores of adult tropical rain-forest trees at La Selva Biological Station in central Costa Rica. The sample includes six different tree species, with both fast- and slow-growing types represented. Using special measuring collars, the scientists obtain reliable data on aboveground growth each year. Deborah Clark presented the team's findings in August at a meeting of the Ecological Society of America in Baltimore.

The group found that growth of all the trees fluctuated considerably from year to year. Moreover, the year-to-year changes correspond strikingly with the results of separate calculations of the size each year of a colossal unexplained tropical terrestrial source of carbon dioxide. In years when this theoretical source was large, the trees had grown slowly; in years when it was small or negative, the trees had grown faster.

The apparent lesson is that the varying annual growth rate of trees in tropical forests could account, in large part, for a calculated increase in carbon dioxide released from land in the tropical zone in the 1980s and 1990s (although other sources, such as soil microbes, probably also contribute). Although trees take in carbon dioxide and release oxygen during photosynthesis, they also release some carbon dioxide as a by-product of respiration, as most organisms do. When growing vigorously, plants take up more than they produce. But if growth slows, the balance shifts.

The annual excess of carbon coming from tropical forests, according to a preliminary calculation by Keeling and his associates, has been more than four billion tons in some recent years. Many researchers regard such estimates as provocative but not ironclad. The new data on tree growth "increase confidence in Keeling's work," Clark says. For comparison, worldwide carbon release into the atmosphere from the combustion of fossil fuels is estimated to be about 6.5 billion tons each year.

In an effort to understand what was causing the year-to-year variations in the rate of tree growth in Costa Rica, Clark and her colleagues evaluated climatic factors. They found that rate of growth was strongly linked to average temperature, slowing down in warmer years. The negative link was even stronger between growth rate and daily minimum temperature. "Tropical trees are being increasingly stressed through higher nighttime temperatures," Clark concludes. She thinks higher nighttime temperatures force the trees to respire more, thus promoting release of carbon dioxide. Yet warming does not increase photosynthesis, leading to a growing imbalance.

The new information from Costa Rica has not yet been published in a peer-reviewed journal, so it remains to be seen whether the scientific community will accept it. Globally, daily minimum temperatures have been increasing faster than average temperatures, so the data suggest that tropical forests might become an even bigger net source of carbon dioxide in coming years. On the other hand, studies of trees in temperate regions indicate that artificially increased levels of carbon dioxide cause trees there to grow faster, which in principle might counter the heat-induced suppression of tree growth. But Clark's observations seem to suggest that the growth-slowing effect of increased temperatures in tropical regions is now stronger than any beneficial fertilizing effect from rising carbon dioxide.

Lest anyone get the mistaken idea that destroying tropical forests would help, James T. Randerson of the Carnegie Institution of Washington notes that clearing a forest adds much more of the gas to the atmosphere than does leaving it be. Researchers believe that tropical forests account for about one third of all carbon dioxide taken out of the atmosphere by photosynthesis on land, making them a crucial part of the global atmospheric equation. The newly detected slowing effect of temperature on tropical forest growth "could be a positive feedback" that will speed global warming, Clark warns.

—Tim Beardsley in Washington, D.C.

STEVE KAUFMAN *Corbis*

***TREES IN COSTA RICA** and other tropical regions may be feeling the effects of global warming.*

ASTRONOMY

POLARIZED LIFE

Astronomers probe Orion to answer one of life's mysteries

For Europe, 1848 was a year of revolutions—one of which was scientific. It was then that the young Louis Pasteur showed that certain organic molecules come in two mirror-image forms, one that rotates polarized light to the right, the other to the left. Such molecules are said to have a definite handedness, or chirality. And it has long been a great mystery why organisms show "homochirality," or, more specifically, an overwhelming preference to build their cells with right-handed sugars and left-handed amino acids. But now a team of astronomers has stumbled on what may be an important piece of the puzzle.

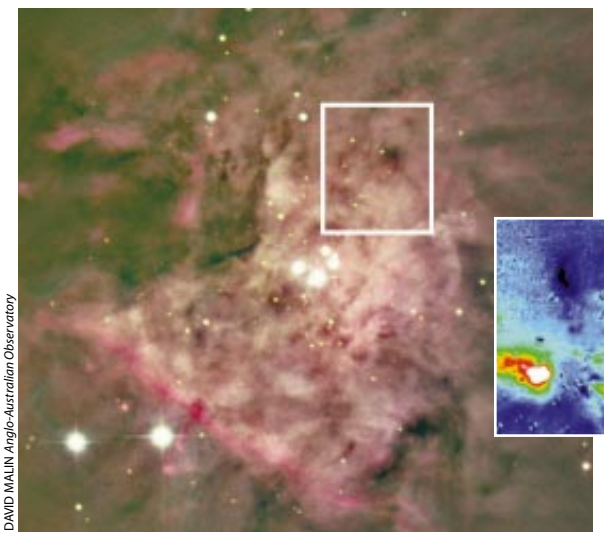
This past July in *Science*, Jeremy Bailey of the Anglo-Australian Observatory and seven colleagues reported that they had discovered large areas of circularly polarized light coming from a region of star formation in the constellation Orion. (The circular polarization of a light wave refers to the orientation of its oscillating electric field, which rotates 360 degrees clockwise or counterclockwise during each cycle.) Some of these immense patches emit circularly polarized light that is predominantly right-handed, some left-handed.

The astronomers were measuring the polarization that can come about when celestial dust grains scatter light from nearby stars. By doing so, they hoped to learn more about the makeup of these particles. For a while, they detected 1 or 2 percent circular polarization at most. Then, according to team member James H. Hough of the University of Hertfordshire in England, one night their primary targets were obscured, and the researchers said to themselves, "Let's look at Orion; it's always interesting." They were stunned to find as much as 17 percent circular polarization.

Bailey realized that such high percentages might have relevance to the enigma of biological homochirality. Even if lifeless, such dusty regions probably contain organic molecules, including amino acids, a supposition based in part on the discovery of extraterrestrial amino acids within the meteorite that fell on Murchison, Australia, in 1969. The handedness of life could be explained if circularly polarized ultraviolet light bathed the dusty cloud that condensed into our own solar system and preferentially destroyed the right-handed amino acids. (Laboratory experiments show that such selectivity readily occurs, but whether the right- or left-handed form breaks down depends on the spectrum of the light.) When the first life-forms eventually emerged, they used the more numerous left-handed amino acids to build proteins, which were shaped in a way that naturally favored right-handed sugars.

One objection to this hypothesis is that the astronomers observed only circularly polarized infrared light (a wavelength that can pierce dusty regions), whereas ultraviolet light is needed to weed out chiral molecules. But the researchers' computations showed that the scattering of light from elongated grains aligned by a magnetic field should generate circularly polarized ultraviolet along with infrared. Another objection is that perhaps life needed no external influence beyond chance to choose its handedness. Perhaps so, yet last year's discovery that even the nonbiological amino acids in the Murchison meteorite tend to be left-handed argues that some extraterrestrial mechanism must have operated to create this imbalance.

Scientists have invoked many other ideas to explain the chirality bias, such as tiny asymmetries in fundamental physics, light from exotic neutron stars and spontaneous chemical reactions. Though still possible, these explanations are only speculations, whereas the astronomers' new work, in the words of Dilip K. Kondepudi, a physical chemist at Wake Forest University who studies homochirality, "gives us some hard facts." —*David Schneider*



DAVID MALIN *Anglo-Australian Observatory*

***ORION'S OMC-1 STAR-FORMING REGION** has dust clouds that circularly polarize light. Where present to a large degree* **(red and white areas in inset),** *this phenomenon may favor certain organic molecules.*

PHYSICS

STRING INSTRUMENTS

String theory may soon be testable

The theory of strings, which attributes the infinite variety of the cosmos to the harmonies of subatomic membranes, has emerged over the past two decades as the leading contender for the "theory of everything." It would explain the four forces of nature—gravity, electromagnetism, and the weak and strong nuclear forces—as a single force with different manifestations. But how could such a theory ever be proved? The last time the four forces acted as one was at the big bang; to re-create those conditions, physicists would need a particle accelerator larger than the solar system, which Congress might be reluctant to fund. Despairing of the task, some scientists call theories of everything an exercise in theology. "For the first time since the Dark Ages," physicists Paul Ginsparg and Sheldon L. Glashow wrote 12 years ago, "we can see how our noble search may end, with faith replacing science once again."

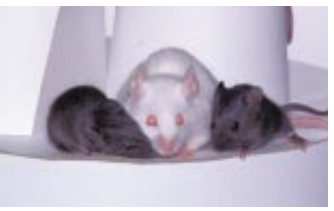
That proclamation now seems premature. Researchers have devised the first astronomical probe of theories of everything and have also discovered that the four forces may unite under conditions short of the big bang. "Unification, the theory of everything, might actually be accessible experimentally," says Nima Arkani-Hamed of the Stanford Linear Accelerator Center.

The probe was conceived by Giovanni Amelino-Camelia of the University

IN BRIEF

Cloned Clones

This time the creators of Dolly the cloned sheep have truly been outdone. An international team of scientists led by Ryuzo Yanagimachi of the University of Hawaii has created multiple clones and clones of those clones using a new technique. A paper describing the success appeared in the journal *Nature* on July 23. The group is the first to duplicate mammals in a reproducible fashion, making more than 50 mice that genetically match their sister/parent, sister/grandparent and sister/great-grandparent. Some of the animals now reside at the Liberty Science Center in Jersey City, N.J. For more information, see http://www.sciam.com/explorations/1998/072798clone/index.html at the *Scientific American* Web site.



Mouse clones with white surrogate mother

PROBIO AMERICA, INC.

Good News Blues

A new study has found that it may not pay to look on the bright side in your golden years. Derek M. Isaacowitz and Martin E. P. Seligman of the University of Pennsylvania followed 71 adults, aged 64 to 94, for a year to see how they responded to life's disappointments. They expected to find that generally pessimistic individuals would be more prone to depression, as is true in younger people. Instead they found the reverse: elderly pessimists appeared to experience less depression. Being optimistic may not always be realistic later in life, the researchers suggest.

Camp Toxic

People often blame urban runoff for lake pollution, but in one Sierra Nevada spot, scientists have identified another culprit. John E. Reuter and Brant C. Allen of the Tahoe Research Group at the University of California at Davis and their colleagues found that fully 86 percent of a suspected carcinogen, methyl *tert*-butyl ether, in Donner Lake resulted from summer boating. The group presumes that two-cycle engines, which have exhaust ports at or below the water's surface, are responsible.

More "In Brief" on page 30

of Oxford and the Institute of Physics in Neuchâtel, Switzerland, and his colleagues. They propose using gamma-ray bursts to check whether the speed of light in a vacuum depends on its wavelength. According to special relativity, light has the same speed in a vacuum regardless of wavelength. Therefore, the detection of a wavelength-dependent speed would unearth a level of physical law more fundamental than relativity.

Variations in the speed of light are familiar to anyone who has looked at a prism. Because glass, water and other substances allow red light to go faster than blue, the prism splays white light into a rainbow.

Empty space, too, is a substance of sorts. By the laws of quantum mechanics, particles burble in and out of existence as the void fluctuates around complete emptiness. Present quantum theory, which incorporates special relativity but not gravity, says that these fluctuations affect all wavelengths of light equally. But theories of everything also allow for fluctuations in gravity, which might act as subatomic lenses that bend light. The shorter the wavelength of light, the more it might induce such lensing and the slower it would travel.

Although the retardation is predicted to be small, it might show up in gamma-ray bursts. Whatever their mysterious origins, these intense flashes travel billions of light-years and flicker frenetically. The blinking gives astronomers a handle on any dispersion: at shorter wavelengths, a flicker would register a moment after it appeared at longer wavelengths. Across a typical range of gamma rays, the time difference would be around 10 microseconds—not much, considering that the radiation has traveled for 10 billion years. But it may be just enough for current instruments to detect. And the Gamma-ray Large Area Space Telescope, scheduled to begin operation in 2004, will certainly have the requisite resolution.

Meanwhile there is another way that predictions of string theory could be detectable sooner: namely, if the forces of nature unite under unexpectedly mild conditions. Two years ago Edward Witten of the Institute for Advanced Study in Princeton, N.J., and Joseph D. Lykken of Fermi National Accelerator Laboratory in Batavia, Ill., realized that strings could come into play at lesser energies than previously assumed. In other words, maybe strings aren't so tiny.

The standard argument that strings should appear at high energies is based on theoretical extrapolations from the measured strength of the four forces. Electromagnetism and the two nuclear forces should become equally strong at the so-called Grand Unification scale. At a slightly higher energy, the Planck scale, gravity is supposed to join in. Both scales are trillions or quadrillions of times beyond the reach of today's accelerators.

But these extrapolations don't take into account a key prediction of string theory: the presence of extra dimensions, on top of the four familiar ones—three for space, one for time. New dimensions could lower both the Grand Unification scale (as shown recently by Keith R. Dienes, Emilian Dudas and Tony Gherghetta of CERN near Geneva) and the Planck scale (according to Arkani-Hamed, Savas Dimopoulos of Stanford University and Gia Dvali of the Abdus Salam International Center for Theoretical Physics in Trieste).

Specifically, string theory adds six minuscule dimensions, which Dienes compares to hairline cracks in the pavement. Each crack adds an extra (third) dimension to the two-dimensional road, but if it is small, your car rolls right over it.

If the crack is large enough and if your tire is small enough, however, your car rattles. Similarly, if the extra dimensions of space are large enough and a particle is small enough, the particle could begin to vibrate in those dimensions. New harmonics would develop, generating new particles—and altering the way the electromagnetic and the two nuclear forces are transmitted. Gravity might shift in a telltale way, too: for simple geometric reasons, extra dimensions would cause gravity to weaken more rapidly with distance. Experimenters are starting to look for such an effect.

Lower unification scales would allow the Large Hadron Collider, now being built at CERN, to make strings. To be sure, that prospect is still speculative. "All these proposals are in the spirit of 'unlikely to be right, but so extremely interesting if they are that they are well worth thinking about,'" says Sean M. Carroll of the University of California at Santa Barbara. But along with other hints of new physics—the neutrino mass, the cosmological constant, the odd behavior of meson particles [see "The Asymmetry between Matter and Antimatter," on page 76]—they suggest that we won't need to take a theory of everything on faith after all. —*George Musser*

In Brief, *continued from page 28*

Quitting Aids

It's a familiar argument: even if Congress hikes cigarette prices, addicts will continue to buy them. But in fact, a new study shows that this may not always be the case. Researchers at the Centers for Disease Control and Prevention reviewed 14 years' worth of data from national health surveys and found that lower-income, minority and younger smokers were all particularly responsive to price increases. A 10 percent jump, they estimate, would inspire 25 percent of Hispanic smokers, 10 percent of African-American smokers and nearly 1 percent of non-Hispanic white smokers to kick the habit.

Future Time

At the moment, the most accurate way to keep time is with an atomic clock—a device in which millions of cesium atoms flip between two configurations and count out split seconds. One limit to their precision, though, is that these masses of atoms interact and trip up one another's regular beats. Now, however, a group led by Carl E. Wieman of the University of Colorado at Boulder has solved the problem by creating the world's first—albeit still primitive—superatom, or Bose-Einstein condensate, clock. Because all the atoms in such a condensate are in the same quantum state, they tick off billionths of seconds perfectly in step.

Stick Up

Some drugs, such as protein-based molecules, can't be taken orally, but they are needed often enough to make injections inconvenient. Now, engineers at the Georgia Institute of Technology have come up with a solution: microneedles. These tiny needles, made using ion-etching microfabrication techniques developed for integrated circuits, are far thinner than a human hair and leave prick marks about one micron in diameter. Also, because they penetrate only the outer, nerve-free layer of skin, they cause no pain.

MARK PRAUSNITZ/MARK ALLEN *Georgia Institute of Technology*

Microneedles

More "In Brief" on page 32

ECOLOGY

AFTER THE DELUGE

A hurricane's effects on species are starkly revealed

A natural disaster in a study site might seem like bad news for field ecologists, but when Hurricane Lili swept over the Exuma Islands in the Bahamas on October 19, 1996, the havoc presented two researchers with a unique opportunity.

David A. Spiller of the University of California at Davis and Jonathan B. Losos of Washington University had just surveyed lizard and spider populations on 19 of the small islands around Great Exuma. After recovering their boat from a clump of trees, they were able to recensus the populations over the next few days and so chart the hurricane's violence. Most such surveys, in contrast, are conducted weeks or months later and so can leave key questions unanswered. When animals vanish, for instance, it could be either because habitat was destroyed or because individuals were literally blown away.

Eleven of the Exuma Islands that Spiller and Losos studied were exposed to Lili's full force and were hit by a storm surge of almost five meters (16 feet). These islands lost all their previous lizard and spider populations, although a few spiders of a species that had not previously been seen were noted, presumably blown in from other areas. A year later many spiders had returned, although very few lizards had made it back.

Eight islands were partly sheltered from the storm surge. Here animals fared better: despite a 34 percent drop in the number of lizards counted, no lizard species became extinct, although almost half the spider species vanished. A year later the number of species of spiders had rebounded, but lizard populations had not recovered significantly.

The serendipitous findings, reported in *Science*, boost several ideas about extreme events that ecologists have had difficulty demonstrating. Most biologists assume that large animals are better able to weather a disturbance than small ones, but until now there have been "precious few examples," says Stuart L. Pimm of the University of Tennessee. Small organisms seem to have a different talent: if they can endure through a calamity, their populations recover more quickly.

Lili's long-term effect on the species mix of the Exumas lends weight to an idea put forward in 1983 by Spiller and Losos's co-author, Thomas W. Schoener (also at Davis): that hurricanes might explain why some species are absent from islands that seem to offer a good habitat. "That's something people didn't consider in their theories 20 or 30 years ago," Spiller says. Losos and Spiller had deliberately introduced lizards to some of the Exumas in 1993 and 1994 to test Schoener's idea.

What happened to the creatures that vanished in the hurricane's force? Losos thinks they were blown or washed out to sea, where presumably most perished. But maybe not all. Several researchers have observed that storms can transport live animals over great distances. Ellen J. Censky, a biologist at the Carnegie Museum of Natural History in Pittsburgh, has described how after a storm iguanas rafted ashore on vegetation to a Caribbean island where they were not previously known and established a permanent colony. Hurricanes, it seems, play a role in evolution as creators as well as destroyers.

—Tim Beardsley in Washington, D.C.

PHOTOGRAPHS BY JONATHAN B. LOSOS

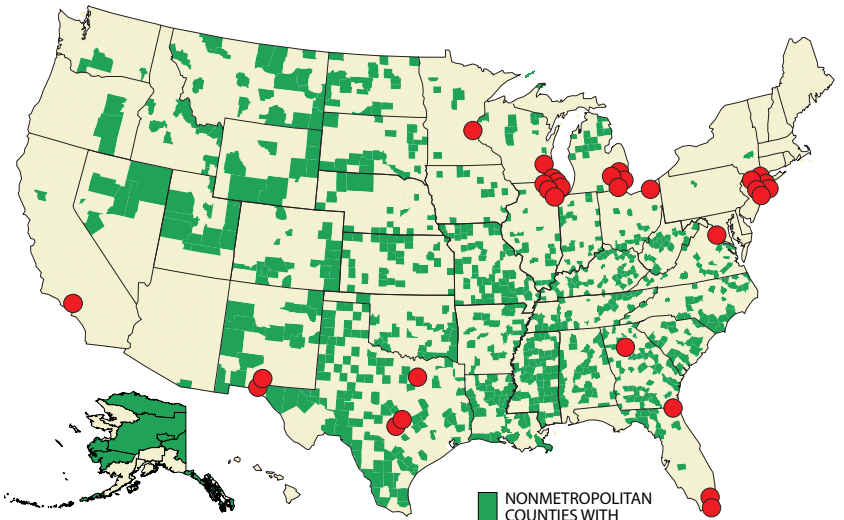
BEFORE AND AFTER:
Hurricane Lili swamped the Exumas.

BY THE NUMBERS

Where the Doctors Aren't

The U.S. has more than 640,000 physicians involved in patient care—one for every 420 Americans. This ratio would seem to be adequate, but because the distribution of doctors is not even, one of every 10 Americans is medically underserved solely by reason of geography. Both urban and rural areas are affected, the latter in part because, in an age of growing specialization and technological advances, young physicians are drawn increasingly to urban hospitals. Moreover, medical schools discourage development of primary-care physicians, who are far more apt to go to rural areas than specialists are. Another factor is doctors' spouses, who are pursuing professional careers in greater numbers than before and thus are unlikely to find jobs in the countryside. Minority communities are more likely to suffer shortages because white physicians are reluctant to work there. Because minority physicians are more apt than their white counterparts to practice in minority neighborhoods, the dismantling of affirmative action programs in higher education systems, as has happened in California, could worsen the shortages.

The Health Resources and Services Administration of the U.S. Department of Health and Human Services (HHS) has identified more than 2,750 regions where the number of primary-care physicians is inadequate. These areas, which were chosen on the basis of low physician-to-population ratio, high infant mortality and traveling time to physician offices, are home to 27 million medically underserved people, of whom 57 percent are in metropolitan areas. According to the HHS, more than 12,000 additional physicians are needed to bring these places up to the standard of one primary-care physician per 2,000 people (the minimum needed to serve a population adequately). Shortages may occur in towns, neighborhoods, correctional facilities, school districts or whole counties.



SOURCE: Health Resources and Services Administration, U.S. Department of Health and Human Services, as tabulated by Quality Resource Systems in Fairfax, Va. County designations are as of September 1996, and urban designations are as of March 1998.

In rural and small-town America, 814 whole counties are designated as shortage areas (*dark green areas on map*). Over 120 of these sparsely populated counties have no physician whatsoever, not even a physician's assistant. (Information on other health professionals, such as nurses or certified midwives, is lacking for these zones.) The map also indicates the 30 urban areas with the greatest need for primary-care physicians (*red circles*). Most are inner-city neighborhoods, such as the Bedford-Stuyvesant area in Brooklyn, N.Y., where there is a shortage of 59 physicians, and the Logan Square neighborhood in Chicago, where there is a shortage of 18.

Because financial incentives have been largely ineffective in luring physicians to shortage areas, the HHS has been exploring the possibilities in telemedicine, which allows remotely situated physicians to evaluate patients. Practicing telemedicine can be as simple as using a fax machine or as sophisticated as transmitting two-way audio and video, with zoom cameras for dermatological exams and signal-transmitting equipment, including electronic stethoscopes, endoscopes and electrocardiograms, for internal diagnoses.

Currently, the primary obstacles to telemedicine are regulatory and procedural. Barriers to the interstate practice of medicine would have to be removed, for instance, and insurance coverage for telemedicine would have to be arranged. Ideally, telemedicine could, within the next decade, bring the same quality of service to currently underserved people as is now enjoyed by those patients with easy access to centers of medical research in cities such as New York City, Boston and Cleveland.

—Rodger Doyle (rdoyle2@aol.com)

RODGER DOYLE

In Brief, *continued from page 30*

Wet Weekends

The rain in White Plains (New York State, that is) falls mainly on Saturdays, according to a recent analysis. Randall S. Cerveny and Robert C. Balling, Jr., of Arizona State University reviewed weekly climate fluctuations recorded at Sable Island—just east of Nova Scotia and downwind of the U.S. eastern seaboard—from July 1991 to January 1995. They found that concentrations of ozone and carbon monoxide peaked late in the week—a cycle that strongly correlated with observed rainfall patterns. The conclusion? Aerosol pollution probably ensures that coastal areas are wetter on the weekends, making Saturdays 22 percent rainier than Mondays.

J. CRAWFORD *The Image Works*

Murphy's Law in action

Dilbert's Corollary

You're not imagining it. Your boss *is* wasting your time. Ohio State University researcher Paul C. Nutt studied decision-making tactics used by 376 business managers. His paper, which appeared in the *Journal of Management Studies,* reported that only 8 percent of the decisions he examined were implemented using the most successful means, a strategy he calls intervention: establishing standards and then measuring performance against them. Another highly effective technique—having subordinates take part in the decision making—was used no more than 16 percent of the time. In contrast, nearly 40 percent of the decisions were implemented using the least successful tactic, issuing edicts. And overall, 37 percent of the decisions were never put to use.

Breakthrough Brewing?

Fuel cells, which combine hydrogen and oxygen to generate electricity, have yet to achieve widespread use in industry. But they seem to have found a foothold in brewing. Fuel cells made by Mitsubishi Electric Corporation are providing power to Asahi and Sapporo breweries in Japan, according to Nikkei America, and the Kirin Brewery Company's Tochigi brewery will be installing a fuel cell next year, utilizing hydrogen made from methane produced during the brewing process. *—Kristin Leutwyler*

ANTI GRAVITY

Whale Weight Watching

Call me a schlemiel. Some weeks ago—it was in July actually—having few if any pressing assignments, and nothing particularly interesting for sure, I hooked up with an outfit called the Institutes for Journalism and Natural Resources, which schlepped me and a bunch of other reporters around Maine for a week and a day so we could acquire expertise in environmental issues. During this excursion, I found myself at the town of Lubec, home to salmon pens and a sardine cannery. Despite the revelation that each sardine was scissors-decapitated and gently placed in the familiar round-edged can by hand, and despite the computation that one worker, slicing for six decades, has beheaded some 300 million herring, this story has bigger fish to fry. Mammals, actually.

I contend that I am a schlemiel because it was not until a week later, as I wandered the streets of Woods Hole, on Cape Cod in Massachusetts, contemplating knocking Red Sox fans' hats off, that I by chance saw a flyer for a lecture on blubber. Thus did I discover Michael J. Moore. Captain Moore seeks the right whale. From a base back in Lubec.

Moore, a researcher at the Woods Hole Oceanographic Institution (WHOI), studies the whales in search of clues about the overall health of their ecosystem. "They are filter-feeding, plankton-feeding animals," he noted between bites of lobster roll in a local bistro. "And we really want to get a sense of what the base of the food chain looks like toxicologically." Moore and his colleagues already knew that the 300 or so right whales left in the northwest Atlantic reproduce at a lower rate than their southern cousins. Then, four years ago, one of his colleagues, Amy R. Knowlton of the New England Aquarium, offhandedly remarked that the whales looked a bit skinny.

Moore, a veterinarian before getting a doctorate from WHOI, thought about this observation. "And we decided that it was appropriate to try to measure their body-fat thickness in the field." If the whales simply weren't finding enough to eat, that would certainly reflect on the overall quality of their habitat. But how do you measure wild whale fat? Moore pointed a french fry at me and said, "It's not easy."

Ultrasound devices have long been used to gauge fat thickness in pigs. Moore embarked on an odyssey of instrument development by simply lashing a pig-fat meter to a reject 20-foot-long, carbon-fiber America's Cup yacht batten. Then, looking like a man combining fly-fishing with metal detecting, he snuck up on some right whales. "We successfully got measurements of the back-fat thickness of two animals," he recalled, "so we knew it could be done." But a 20-foot pole meant getting so close that the whales became skittish, in effect saying fat chance.

Experiments with 40-foot lengths of sailboat spars and four-by-four lumber revealed the structural weaknesses in spars and lumber. Moore finally turned to a WHOI engineer, Richard Arthur, who shored up the longer pole with clamps, struts and wires and designed a nifty turret for bracing the rod on the boat. In a test last summer, Moore was able to get multiple readings from more than 30 tranquil animals a day. Unfortunately, the blubberometer itself was doing less than a whale of a job.

THE BRIDGEMAN ART LIBRARY INTERNATIONAL

Moore taking readings (artist's conception)

So Moore recently updated to a sophisticated ultrasound device ordinarily used to detect flaws and cracks in steel. He combined that with a computer program specifically designed to suck the ultrasound images into an onboard laptop. When we talked, he was just days from Lubec and the whales whose backs he sought to scratch in the Bay of Fundy. And I was reminded, as he rushed from the restaurant, that men obsessed with whales tend not to take no for an answer. *—Steve Mirsky*

PROFILE

Starving Tumors of Their Lifeblood

No, Judah Folkman probably won't cure cancer in two years. He says he simply hopes to render it a manageable, chronic disease

At 9 A.M. on a hot, sticky Friday in the middle of June, the tiny conference room on the 10th floor of Children's Hospital in Boston is packed with the eager faces of an ethnically diverse cast of young people munching bagels. Among the Birkenstocks, T-shirts and jeans of his 30-odd graduate students and postdoctoral fellows, Judah Folkman stands out because of three things: his age (65 years), his necktie and white lab coat, and his courtly but authoritative manner.

At this weekly laboratory meeting, several of the lab members stand to describe their most recent results studying the link between angiogenesis—the growth of new blood vessels—and cancer. Folkman offers everything from detailed remarks on the methods of a particular experiment to advice on how to make the best use of an overhead projector. He's the consummate manager and mentor: one minute he's upbraiding a cocky postdoc for not taking criticism as easily as he dishes it out; the next he's commending the same young man for "good progress" and joking with him that it's not yet time for him to give up and go to business school. The postdoc wraps up his presentation and sits down with a smile.

Folkman and his prolific laboratory hit the news in a major way this past May, when an overenthusiastic, front-page story in the *New York Times* trumpeted results by Folkman's group using naturally derived angiogenesis inhibitors to cure cancer in mice by preventing the growing tumors from attaining a blood supply. The focus of the story was a scientific paper published in November 1997 that had already been the subject of a *Times* news story, though not on page one.

One of the most provocative aspects of the *Times* article was a quote attributed to Nobel laureate—and biology legend—James D. Watson: "Judah is going to cure cancer in two years." Although four days later the newspaper published a letter from Watson saying his "recollection of the conversation" with the *Times*'s reporter was "quite different," the damage had been done. Hordes of people with cancer were already rushing their physicians' offices, demanding access to the impending "cure"—despite the fact that it has yet to be tested in a single human. Folkman's office alone logged more than 1,000 calls a day from cancer patients and their loved ones the week after the *Times* ran the story.

Folkman says he is puzzled over why the *Times* decided to publish such a belated, breathless article on his group's work. "Our published results have all been in mice," he emphasizes. "Many different substances have been shown to inhibit cancer in mice over the years, but unfortunately, so far not all of them have worked as well in people." Most of all, he says, he is concerned that the story might have instilled false hopes in so many of those desperately ill with cancer.

Folkman is a leading pediatric surgeon but shows none of the ego of the stereotypical topflight surgeon. Quite the opposite. He dislikes giving interviews (especially for television) to the point that this summer he even turned down a request by NBC morning anchor Katie Couric—who had recently lost her husband to colon cancer—to appear on the *Today* show. He also hates having his photograph published—not because he is vain about his looks, he says, but because he doesn't want to seem to be taking sole credit for the dogged work of the many scientists who make up his laboratory. In addition, he says, he wants to avoid being thought the leader of the only laboratory in the world devoted to angiogenesis, because many other labs contribute to the field.

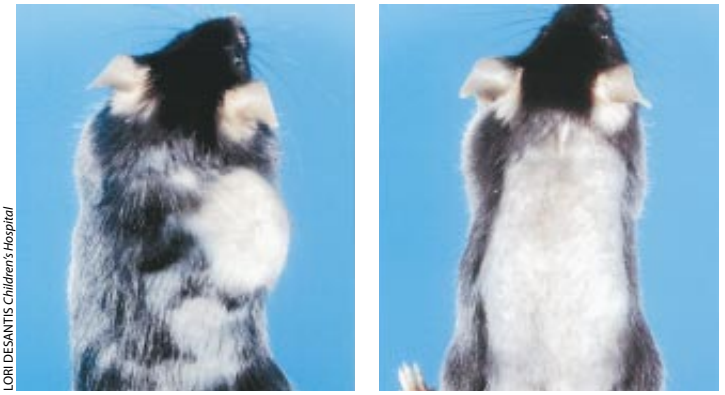
A cavernous, elaborate workspace is not for Folkman: his office, which he rarely uses, is small and furnished with tattered, 1970s-era furniture. Every horizontal surface is stacked with books, journals, files and papers, so that the room more closely resembles an attic.

LORI DESANTIS *Children's Hospital*

***FINDING NEW ANGIOGENESIS INHIBITORS** is the goal of Judah Folkman* **(far right)** *and his colleagues* **(from left)** *Robert J. D'Amato, Michael S. O'Reilly and Donald Ingber. Folkman dislikes having his photograph taken for publication without others from his laboratory.*

Although his office has a computer, Folkman's secretary says that when he needs to write something he usually pulls a chair up to a spare computer next to her desk, in a cramped corner in front of a mini-refrigerator, because there is more room there than in his own space.

Folkman's lifework on cancer and angiogenesis began in circumstances not of his own making: he was drafted into the U.S. Navy in 1960. Although he had just finished his assistant residency in surgery at Massachusetts General Hospital in Boston, the navy set him up with a small lab at the National Naval Medical Center in Bethesda, Md., to help in the military's drive to create blood substitutes for use on aircraft carriers, which often spend months at sea.



LORI DESANTIS *Children's Hospital*

DISAPPEARANCE OF A TUMOR ***implanted in a mouse took place in 12 days when the mouse was treated daily with the angiogenesis inhibitor endostatin.***

There Folkman conducted the pivotal experiments that focused him on angiogenesis. While studying the ability of a cell-free blood substitute to keep a rabbit thyroid gland alive in culture, Folkman and navy colleague Frederick Becker placed a few rabbit melanoma cells on the gland's surface. To their surprise, the cells grew but stopped once they formed tumors the size of peas. "Why did the tumors stop growing?" Folkman asks. "That question kept me going for years."

After leaving the navy in 1962, Folkman returned to Mass General, where he became chief surgical resident two years later. As one of Harvard Medical School's brightest young surgeons, by 1967 Folkman had attained tenure, going directly from associate (instructor) to full professor and chairman of the department of surgery at Children's Hospital in just one year. Folkman had distinguished himself as a surgeon through his technical skill and his ability to train others. He had also participated in the early development of implantable drug-delivery devices, which eventually led to the commercialization of products such as the contraceptive Norplant.

Along the way, Folkman kept a small research lab going on the side to pursue his interests in angiogenesis. But when he tried to publish his animal results, he was turned down by dozens of journals. Many scientists scoffed at his idea that devising a way to block angiogenesis might keep growing tumors in check.

It was only through giving a lecture in 1971 that Folkman got his ideas into an important journal for the first time. That year he was asked to give a special seminar at Beth Israel Hospital in Boston that has often been invited for publication in the *New England Journal of Medicine*. Finally, Folkman had a well-read platform for describing his conclusion from the rabbit thyroid gland experiments: that tumors are incapable of growing beyond a certain size unless they have a dedicated blood supply and that finding a way to block the process of angiogenesis might nip emerging cancers in the bud.

But the *NEJM* article simply egged on Folkman's critics. In 1973, for example, when Folkman and his co-workers reported that injecting human tumor cells into the eyes of rabbits prompted angiogenesis, some scientists argued that the observed blood vessel growth was simply part of an inflammatory reaction to foreign cells. One researcher subsequently showed that implanting a chemical irritant, a crystal of uric acid, in rabbits' eyes also spurred angiogenesis. It took years for Folkman and his colleagues to explain this finding by demonstrating that immune system cells called macrophages had entered the rabbits' eyes to destroy the uric acid and had secreted substances that promote angiogenesis.

Folkman's struggles for credibility affected all the factors crucial to the success of a biomedical researcher: his ability to obtain grants from the National Institutes of Health, his chances of publishing his ideas in leading journals and his capacity to attract scientists in training to work for him in his laboratory.

"In the 1970s professors dissuaded their best students from coming to work in my lab," Folkman says matter-of-factly. The only way he could convince outstanding young scientists to join him, he says, was by reminding them that they were so good that even if things didn't work out and they left after a year, their careers wouldn't be harmed.

Throughout the 1980s, Folkman and the other scientists in his laboratory kept adding pieces to the puzzle of angiogenesis and slowly gaining adherents to the idea that inhibiting angiogenesis might be a key to keeping cancer in check. A significant break came in 1994, when Michael S. O'Reilly in Folkman's lab isolated one of the most potent natural inhibitors of angiogenesis, which they named angiostatin [see "Fighting Cancer by Attacking Its Blood Supply," by Judah Folkman; SCIENTIFIC AMERICAN, September 1996]. Folkman, O'Reilly and their co-workers isolated a second natural inhibitor, endostatin, in 1996.

Folkman and his colleagues have now published articles in all the most prestigious research journals, and the list of awards and honors Folkman has received takes up two full pages of his curriculum vitae. Although researchers are not yet clear exactly how angiogenesis inhibitors work, angiostatin is expected to be tested in humans beginning late this year. (Several synthetic angiogenesis inhibitors are already in clinical trials.)

When asked how he persevered despite his early critics, Folkman credits his wife of 38 years, Paula, an alto who sings as a full-time member of the chorus with the Boston Symphony. "I would come home at night so disheartened," he says, "and she would ask, 'Why do you care what they think?' She has always been very supportive."

Does Folkman believe that he will eventually cure cancer? "No, I don't think angiogenesis inhibitors will be the cure for cancer," he answers. "But I do think that they will make cancer more survivable and controllable, especially in conjunction with radiation, chemotherapy and other treatments. I'm very excited to see how they will work in people." —*Carol Ezzell*

MEDICAL TECHNOLOGY

BLOODLESS TESTING

Noninvasive glucose monitors for diabetics are on the way, but cost could be a problem

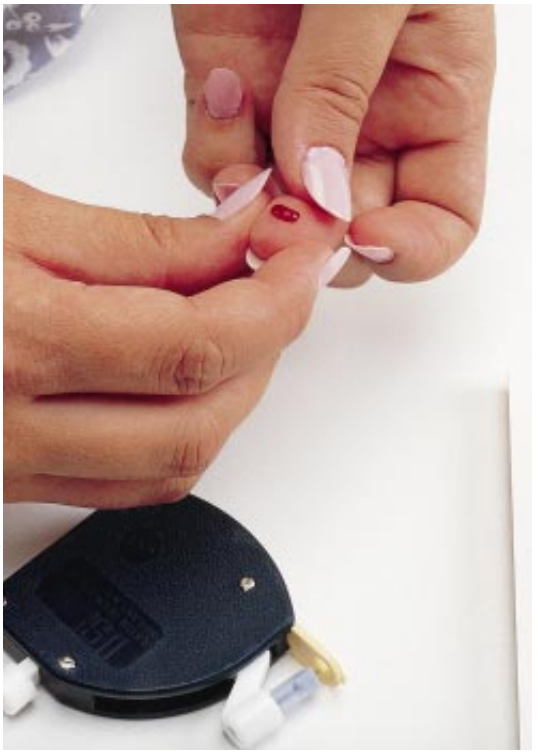
Lancing the finger to check blood glucose levels is an all too common ritual for diabetics, whose bodies cannot produce enough insulin to metabolize the sugar. Unfortunately, given the discomfort and inconvenience of finger pricking, many diabetics do not monitor themselves as regularly as they should, putting themselves at risk for kidney failure, blindness and stroke. Several new instruments that do not require drawing blood, however, are on the way, which could help the estimated 15.7 million diabetic Americans and reduce the $92-billion toll that the American Diabetes Association says the disease takes after medical costs, disability and lost earnings are totaled. Still, the new devices are expensive, and given the reluctance of insurance companies to cover glucose monitoring, they may reach only a few diabetics.

The world's first noninvasive glucose sensor, the Diasensor 1000, has recently received market approval from the European Union. It relies on a fiber-optic probe that emits infrared light, which passes through the skin into the blood. The light is reflected back into the sensor and then is analyzed in the Diasensor's computer. The machine's creator, Pittsburgh-based Biocontrol Technology, has been working with the Food and Drug Administration since 1994 to get the Diasensor approved for domestic use, but in 1996 an FDA review panel delayed making a final recommendation and requested more information, which Biocontrol has yet to present.

Other noninvasive glucose sensors are in various stages of development. Cygnus in Redwood City, Calif., for example, is finishing up clinical trials for its GlucoWatch, which is worn like a wristwatch. The monitor uses low-level electric current to extract glucose painlessly and move it into a transdermal pad. Cygnus hopes to submit its data to the FDA before the end of this year.

A different approach is taken by the SalivaSac, being developed by Pacific Biometrics, based in Lake Forest, Calif. It was designed to improve the diagnostic use of saliva by "ultrafiltrating" it—that is, separating out enzymes, food and other contaminants. Pacific Biometrics is currently in an early phase of testing its device and does not expect to market it for at least two years.

Looking further into the future, George S. Wilson of the University of Kansas is developing an implant that would continuously monitor blood sugar levels and set off an alarm when insulin is needed. The sensor, about three times the thickness of a human hair, can be easily implanted underneath the skin with a needle. National Applied Science in Portland, Ore., holds the rights to market the system, which is about five years from FDA approval.



GRAPES/MICHAUD Photo Researchers, Inc.

***MEASURING GLUCOSE LEVELS** may soon no longer require finger pricking.*

The worldwide market for glucose monitoring has reached $1.5 billion and is growing 15 percent a year. With some 800,000 new diabetes cases appearing in the U.S. annually, any company that develops an accurate, convenient bloodless monitor seemingly stands to make a huge sum of money. That is, if the monitoring systems are competitively priced and if health insurance will pick up the tab. Michael Mawby, national vice president for advocacy at the American Diabetes Association, believes insurance companies may be less than enthusiastic. "I'd be very surprised if any insurer rushed out to cover them," he says. "FDA approval does not translate into automatic coverage."

Mawby's caution is strongly supported by past performance records. Both private insurers and Medicare have been reluctant to provide comprehensive insurance coverage for the needs of diabetic patients. In August 1997 LifeScan in Milpitas, Calif., the leading manufacturer of blood glucose–monitoring systems, commissioned a Gallup survey of 252 of America's largest companies. The survey showed overall coverage for diabetes management tools to be limited: 38 percent provided no coverage for glucose-monitoring meters and strips. That number, Mawby believes, is fairly accurate, although aggressive lobbying on behalf of diabetics has convinced 30 state legislatures to pass laws requiring insurance companies to offer or include coverage. The federal government also expanded Medicare coverage for diabetes-related expenses beginning this past July 1.

Existing glucose meters are not themselves expensive. LifeScan's new compact FastTake system, for example, retails at around $65, but with a rebate and trade-in allowance, it can end up costing as little as $5. The real expense lies in the test strips, which are used once and average between $0.50 and $1 apiece. A diabetic can conceivably spend $150 a month on the strips.

The new technology offers little relief on the financial end. Biocontrol plans to sell the Diasensor 1000 in Europe for $9,000. But unlike the standard meter, the Diasensor does not require any additional supplies and should last 10 years. The price of the GlucoWatch will range between $225 and $250, with an additional cost of approximately $4 for each transdermal autosensor pad, which

needs to be changed every 12 hours.

Without insurance, the devices may be out of reach for many diabetics. Andrea Sobel of Cygnus says the firm has already discussed its GlucoWatch with several managed health care providers and will work with them to get at least partial coverage after the product is launched. And Pacific Biometrics intends to price the SalivaSac competitively with existing products. "The components are not exceptionally expensive," says Sayed M. Badrawi, vice president of marketing. "Our object is not to make this an elite product but a product for everyone who wants to use it." He and the other developers hope insurers will see it that way, too. —*Roxanne Nelson*

ROXANNE NELSON, based in the San Francisco area, described the return of potato late blight in the June issue.

DEFENSE POLICY

INNER-CITY VIOLENCE

The U.S. military tries to prepare for urban warfare

For much of the cold war, American military planners devoted most of their energy and budget to preparing for all-out battle with the Soviet Union—a global war that Dwight D. Eisenhower predicted before his presidency would make targets of "the cities mankind has built." With the cold war over, the U.S. is again focusing on cities, but the threat of nuclear attack is no longer the main worry. Today the military expects to be called on most often not for major wars but for a variety of smaller conflicts, many of which will have enemy soldiers lurking around every corner, behind every door and under every manhole cover. The Pentagon calls these conflicts military operations in urban terrain (MOUT), and planners are scrambling to ensure that U.S. forces are ready for them.

The problem is that vastly outnumbered and outgunned forces can use cities as cover for guerrilla warfare, putting at risk troops relying on training and equipment designed for different scenarios. For instance, in 1993, 18 U.S. soldiers from an elite fighting unit died in the streets of Mogadishu, Somalia, during a peacekeeping operation that turned violent. Similarly, thousands of Russian troops died during the following two years in horrific urban battles with separatists in Grozny, the capital of Chechnya.

Traditionally, the U.S. military has instructed its leaders to stay away from cities and has developed technologies for other types of battles. "Our military doctrine has been one of avoiding cities, and we can't avoid them any longer," says Carol Fitzgerald, the director of an army-led program designed to test new concepts and technologies for fighting in urban areas.

Russell W. Glenn, a defense analyst with Rand Corporation, agrees "there hasn't been much done on the subject," but he believes that during the past two years the armed forces "have begun to recognize the challenges that MOUT offers far more than they did in the past."

Fitzgerald's demonstration is one of several new MOUT studies aimed at finding ways to keep soldiers alive while simultaneously keeping cities standing and civilians unharmed. It's not an easy task. "Urban areas make combat more difficult," says Timothy Jones of the Marine Corps Warfighting Laboratory in Quantico, Va. "Rubbled streets and buildings limit mobility, [and] large numbers of noncombatants limit our use of fire support. At the same time, dense infrastructures affect our ability to communicate."

With urban battles increasingly likely (about 70 percent of the global population is expected to reside in cities by 2020), something clearly has to change. The army, Glenn notes, has not published a MOUT field manual since 1979, and even then its tactics dated to World War II: "Throw the grenade through the window, bust through the door and storm the room," he sums up.

For future operations, then, the goal is to ensure first that civilians or friendly troops are not on the other side of the door. Accordingly, demonstrations such as Fitzgerald's program are tinkering with sensors that can "see" around corners and through walls. Night-vision equipment and other sensors and cameras, some mounted on robots and remotely controlled aircraft, will help clear up confusion in close quarters. Digital map systems will allow improved so-called situational awareness. Technologists are also testing specially made explosives that can blow human-size holes

SARAH UNDERHILL *U.S. Army*

MARINES TRAIN FOR URBAN WARFARE
in a mock city constructed at Camp Lejeune, N.C.

in walls without knocking them down; guided parachutes, or parafoils, that can be used to maneuver from one building top to another; portable water-purification systems; and sophisticated hands-free communications equipment.

But it's not all gee-whiz technology. Fitzgerald's demonstration, for example, has shown that soldiers in city combat can benefit mightily from simple things like elbow and knee pads developed for in-line skaters.

Eventually, Glenn predicts, technology may produce weapons such as missiles that can pinpoint enemies within a building "without endangering civilians or friendlies" in a nearby room. But he and other experts caution that the military first must change the way it trains—and there are signs that this is happening. Both the army and the Marine Corps have constructed mock cities for training and have begun studying the psychological forces at play in unconventional warfare, where soldiers may be cornered, targeted from all sides or unable to distinguish friend from foe from noncombatant. Marines have also traveled to places such as Chicago to get a better feel for what battle could be like in a real city. The Russians know. In Chechnya they were taught a lesson the U.S. learned in the jungles of Vietnam: dedicated forces, no matter how small, can be extraordinarily effective when fighting on their own turf. Future enemies no doubt will try the same tactics, which is why the U.S. military is seeing the city for what, in wartime, it really is—a concrete jungle. *—Daniel G. Dupont*

DANIEL G. DUPONT edits Inside the Army, *an independent weekly newsletter based in Washington, D.C. He wrote about missile defense failures in the June issue.*

ACOUSTICS

IN THE AUDIO SPOTLIGHT

A sonar technique allows loudspeakers to deliver focused sound beams

Imagine being able to project sound in a narrow beam. A public address system could then target a message to its intended recipient, calling a worker, say, from the shop floor without disturbing everyone in the factory. Museum curators could also play a description of an exhibit just to the person standing in front of it. And special effects would be that much more convincing with a highly directional loudspeaker: sounds might then come from whatever part of the room is illuminated by the audio spotlight. As eerie as it seems, these thrills may not be far off.

The long-standing difficulty in projecting sounds in this way arises from basic physics. Typically the angular size of a beam is determined by the ratio between the wavelength of the emanation and the aperture of the source. Sending out a collimated beam of audible sound, which can have wavelengths of several meters, would normally require a loudspeaker the size of a small building.

But acoustical engineers have long known a clever way to sidestep this difficulty. Since the early 1960s they have been able to project (or detect) low-frequency sound in one direction using special transducer arrays that are smaller across than the wavelength of interest. Such parametric arrays, which were developed for underwater sonar systems, take advantage of subtle nonlinearities in the fluid carrying the sound. Just as a loudspeaker distorts music if driven too hard (out of its range of linear response), water will also distort high-intensity sound waves. Rather than being a problem, this nonlinearity offered sonar engineers a way to make the water itself generate low-frequency sound waves from high-frequency ones. (In essence, the water mimics the detection circuit in a radio receiver, which uses nonlinearity to transform a modulated carrier wave of high frequency into an audible signal.) And because of its small wavelength, an ultrasonic carrier can be sent from a physically small source in a tight beam.

For a decade acoustics experts debated whether such parametric arrays would also work in air, where the nonlinearity is much less pronounced. Then, in 1975, Mary Beth Bennett and David T. Blackstock, both at the University of Texas at Austin, put an end to the controversy. They created an audible tone in air using ultrasonic waves. And in the early 1980s Masahide Yoneyama and several Japanese colleagues at Ricoh and Nippon Columbia used the same technique to direct audio in a tight beam using a compact array of ultrasonic transducers.

The Japanese group grappled with the difficulties involved in generating sufficiently powerful ultrasonic sounds and in modulating the carrier properly so that the resulting audible signal would not be too distorted. Yet the application of parametric arrays to send sound through air languished until about two years ago, when F. Joseph Pompei, then an engineering student at Northwestern University, and Elwood G. Norris, an inventor at American Technology Corporation in San Diego, began independently tinkering on the problem.

Pompei had worked earlier for Bose (a loudspeaker manufacturer) and was playing flute and bass guitar at various Chicago clubs. His interests in music and acoustical engineering converged, and after moving to the Media Lab at the Massachusetts Institute of Technology last year, he constructed a parametric array for air. Unlike earlier arrays built with piezoelectric elements, his ap-

paratus employs wide-bandwidth electrostatic transducers, which minimize distortion. Pompei recently demonstrated this device by beaming a John Coltrane saxophone solo around the room. "You'd probably notice that it's not as good as a loudspeaker," he admits. Still, he says, he is able to project a three-degree-wide beam for some 200 meters.

Norris and American Technology Corporation have also been pursuing the prospect of using parametric arrays in air. Norris's system, like Pompei's, has had difficulty with distortion, particularly at low frequencies. But Norris claims that custom piezoelectric transducers that he and his colleagues have started using in the past few weeks have sufficient bandwidth to have "licked" the problem: "All of a sudden now we can play music." So the competition between his company and the Media Lab for beaming sax players around large rooms is sure to heat up. —*David Schneider*

FIELD NOTES

PHANTOM TOUCH

Imbuing a prosthesis with manual dexterity

"Whee!" exclaims Melissa Del Pozzo, a vivacious 10-year-old who is watching an electrical trace on a computer screen undulate up and down like ocean swells. Born without a left hand and lower forearm, Melissa makes the signal oscillate by moving either a tendon or muscle in the arm that would have been used to flex her thumb. A sensor attached to the skin just below the elbow detects the slight movement and relays it to a window on the screen.

"Whee!" Melissa repeats. This time a line wiggles in another window. This one corresponds to the tendon or muscle that initiates pinkie motion. Melissa's ability to trigger separate waves marks the promise of a new technology that may allow her to achieve her dream of playing piano with both hands. The electric signals represented by the fluctuating lines can be used to move independent fingers in what may be the first dexterous prosthesis.

Melissa found out about the testing taking place at Rutgers University's department of biomedical engineering in Piscataway, N.J., just this past June. Her

Scientific American **Resource Center**

Welcome to Scientific American! For your convenience, we have summarized the answers to often-asked customer service questions and how to get in touch with us.

Scientific American Presents
Our new quarterly publication. Each 1998 issue is dedicated to a single theme: March, Magnificent Cosmos; June, Women's Health; September, Oceans; December, Intelligence. Annual subscriptions: In U.S.: $19.80. Elsewhere: $23.80 surface; $43.80 air service.

www.sciam.com
Our award-winning Internet resource, updated weekly and thoroughly linked to related sites. Here you will find timely and interesting current news, recent illustrated scientific articles, current issue highlights, ask the editors, our on-line store, subscriptions, e-mail, and other features not found in Scientific American, and much more.

Scientific American subscription rates
In U.S., Scientific American is: $34.97. Elsewhere: $49. Your 12-issue subscription includes all regular issues with the in-depth special reports. Our 1998 report subjects: April, Wireless Technologies; July, AIDS update; and October, Computer Security.

SCIENTIFIC AMERICAN

Subscription inquiries
Call us: In the U.S. and Canada: (800) 333-1199.
In other countries: (515) 247-7631.
e-mail: subscriptions@sciam.com
or write: Subscription Manager; Scientific American; PO Box 3187; Harlan, IA 51537.
The date of the last issue of your subscription appears on the mailing label. For change of address, please notify us at least four weeks in advance, and include both old and new addresses.

Back issues
In U.S. $9.95 each. Elsewhere: $12.95 each. Many issues are available. Fax your order with your Visa, MasterCard or AMEX information to: (212) 355-0408.

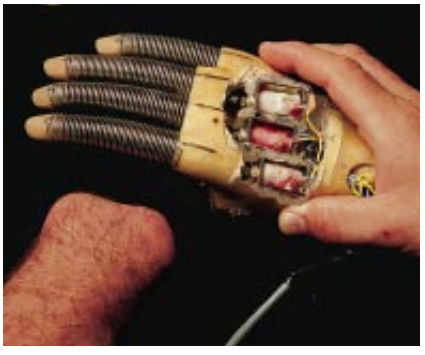
Reprints
$4 each, minimum order 10 copies, prepaid. Limited availability. Write to: Reprint Department; Scientific American; 415 Madison Avenue; New York, NY 10017-1111; fax: (212) 355-0408 or tel. (212) 451-8877.

NAJLAH FEANNY *SABA*

CONTROLLING ELECTRIC SIGNALS BY TENDON AND MUSCLE ***is Melissa Del Pozzo*** **(right)**, ***who might one day similarly control a prosthetic limb.***

parents showed her an article in the *Asbury Park Press* about Keith St. John, a 35-year-old amputee who was testing a hand with three-fingered movement. That was a definite improvement over her own state-of-the-art hand, which consists of a claw covered with plastic that can execute only a simple open-and-shut grabbing motion when activated by an electric potential from muscles in the forearm. After reading the article, Melissa implored her parents to make a call to the researchers, William Craelius, an associate professor of biomedical engineering, and his doctoral student Rochel Lieber Abboudi.

Some half a dozen amputees had made the mechanical fingers wiggle or the lines on the screen oscillate. Melissa so far is the only one to have been born without a hand who could manipulate the signals; two others missing hands from birth could not make the lines jump. Melissa, like some other subjects, reports that she can feel control over missing hands and fingers—a phenomenon known as phantom limbs.



NAJLAH FEANNY *SABA*

PROSTHETIC HAND ***allows flexing of single fingers when actuated by tendons or muscles.***

Her visits to Rutgers are a prelude to fitting her with a hand and the requisite control apparatus. Researchers will fashion a silicon sleeve equipped inside with pressure sensors. On top of the sleeve will sit a hard plastic socket that serves as an exoskeleton on which to anchor the hand. The hand itself is a commercial wooden product used on other prostheses. It is fitted with electromagnets that move each of three fingers separately. When a tendon moves, it causes the sensor—a small diaphragm filled with air—to emit a puff that travels through a tube to a transducer that senses the pressure and transmits an electric signal to the artificial hand. Craelius decided to focus on a tendon-actuated system because of the difficulty in isolating specific muscle groups that can move individual fingers. In the case of Melissa and others missing a lower forearm, where tendons are most accessible, muscles from the upper arm can sometimes be utilized.

The tapping motion that Craelius and his colleagues have demonstrated may suffice to play the piano or saxophone, type on a keyboard or let a court reporter work a stenotype machine. One subject, Jay Schiller, played "Mary Had a Little Lamb" at one-quarter speed. He made only two mistakes, one with his still intact hand. The entire project, though, demonstrates the difficulties inherent in designing the bionic human. Activating a tendon or muscle for each finger may eventually enable Melissa to play the piano, but it will remain a daunting challenge to achieve the full 24 degrees of freedom—that is, the 24 distinct movements—that the human hand can produce.

Triggering finger movement by retraction and extension of a tendon is what engineers call biomimesis, a replication of the body's own control mechanisms. But the tendon's simple back-and-forth motion will not suffice to reproduce a full range of motion. If it ever becomes possible to flash the "V" sign with a prosthetic hand, it will require some novel stratagem. More sophisticated sensors and control programs might anticipate and act on the prosthetic user's demands. Alternatively, the user might initiate different finger movements by flexing sensor-fitted toes.

Going beyond mere taps will require additional engineering. If the signal is held constant for a long time, the hand's electromagnets burn out. "What if you want to hold a cup of coffee for more than a few seconds? That's pounds of pressure and amperes of current," Craelius points out. The Rutgers team is interested in a hand that contains more than one magnet to manipulate each finger. An electromagnet could move a finger, and an accompanying permanent magnet could hold it in place.

The mechanical hand may prove unnecessary for some tasks. In fact, Craelius, Abboudi and their co-workers received a patent not for the prosthetic appendage but for the method of tendon-based control. The importance of the control system is underlined by an upcoming project. The prosthesis—replete with controller and mechanical hand—may serve as a backup for the first hand-transplant operation, which is scheduled to take place at the University of Louisville in Kentucky before year's end. If the patient's immune system rejects the hand, the Rutgers prosthesis could serve as either a permanent replacement or a device that would permit the tendons to be exercised until another transplant can be found. But Craelius emphasizes that the hand is not essential. The Louisville transplant patient, a computer programmer, could attach the tendon sensors directly to the computer for writing software. As such, the Rutgers project may revise the very definition of manual control.

—Gary Stix in Piscataway, N.J.

Y2K: The End of the World as We Know It

Every religion has its doomsday prophecy, and it turns out that computing is no exception. (If you doubt that computing is a religion, just try mentioning Windows to a Mac owner.) The Year 2000 problem, or Y2K as it is called, is definitely real—businesses, governments and individuals are all working on retooling software and embedded systems written in the days when it was safe to call years familiarly by their last two digits. But predictions as to what will actually happen vary wildly, encompassing everything from minimal disruption to "The End of the World as We Know It" (shortened to TEOTWAWKI on newsgroups).

At the denial extreme, we have astronomer Clifford Stoll, author of the high-tech critique *Silicon Snake Oil,* who in a January 1, 1997, debate with Y2K consciousness raiser Peter de Jager claimed that the Year 2000 problem can be fixed in a long weekend, and chief information officer David Starr of Reader's Digest Association, who told *Computerworld* in mid-1997 that Y2K is a fraud. Even more stringently, Jim Wilson, the science editor of *Popular Mechanics,* has dismissed Y2K as an urban legend, apparently on the grounds that the computer industry couldn't possibly be that stupid.

On the Usenet newsgroup comp.software.year-2000, which was set up in 1996 to allow technical people to exchange useful information, these guys are known as Pollyannas. So am I, unfortunately: someone read and posted a newspaper report of a talk I gave at the recent skeptics conference in Germany (sponsored by the Committee for Scientific Investigation of Claims of the Paranormal), in which I admitted that I am a moderate who believes human adaptability and the fact that most people want civilization to survive will help us muddle through. This puts me firmly on the doomed list, and the advice of one particularly rabid TEOTWAWKI type was that I unsubscribe from the newsgroup: "You have learned nothing."

He's wrong, of course. I've learned from reading the newsgroup that I ought to be stocking up with 300 pounds of grain, 60 pounds of legumes, 60 pounds of sugar or honey, five pounds of salt and 20 pounds of fat or oil for the first year, along with a gallon of water per person per day; that I should be buying candles, fuel, medical supplies, a generator, canned vegetables and fruits, garden seeds, blankets, sleeping bags, hand tools, lots and lots of batteries, and even more guns and ammunition to protect the stockpile from the starving and desperate hordes who will flee the burning cities in search of sustenance; and that gold is a poor choice for storing currency because the government can seize it at any time during a national emergency. I should also be buying any books that might tell me how to make things I need when civilization falls. And I should work out, so that I'm physically fit enough to survive whatever humanity and nature throw at me. Except for

DAVID SUTER

the guns (illegal where I live), none of this advice is necessarily bad.

Aside from sad postings about how most of the world's population is going to die—four fifths, according to some postings—there's an element of satisfaction among these Cassandras. They make up the in-group that is going to survive because they're smarter and tougher than the rest of us. Computing gurus are at the mercy of the political and financial decisions of others, just like the rest of us (*Wired* magazine recently featured a few software programmers who were stocking up and taking to the hills). People who have rigorously refused to have computers still rely on the ready availability of electric power, food, telecommunications and, most important, a clean supply of water. About the only people in the U.S. who might escape all effects are the Amish.

On the newsgroup, you can watch at work what one skeptic in another context called the "ratchet effect." Anything—the doubling of the federal government's estimate of the cost of remedying its systems, for example—that depicts Y2K as a catastrophe is carefully reported and believed. Any news suggesting that a remediation effort might succeed is dismissed as lies, stupidity or denial. Off the newsgroup, a computer science researcher of my acquaintance tells me he figures the chances of catastrophe are about 5 percent, and that's enough for him to have sold out of the stock market and filled his country home with supplies, just to be safe.

Over the centuries, of course, there have been many doomsday prophecies: a list published in James Randi's *The Mask of Nostradamus* gives many historical dates on which the world was to end: 1524, when a deluge was supposed to flood London; 1719, when mathematician Jakob Bernoulli expected the earth to be hit by a comet; and 1947, when "America's greatest prophet" John Ballou Newbrough thought (in 1889) that all governments and rich monopolies would cease. After that, the cold war made it completely rational to believe "they" might blow up the world. And the 1980s were a popular time for postholocaust science fiction, from David Brin's 1985 novel *The Postman* (recently a Kevin Costner film) to the Mad Max movies. Yet here we still are, to face Y2K, the perfect mechanism for bringing the world down on our heads.

We have to remember, of course, that they might be right. The only problem is, if everyone runs for the hills, they'll get kind of crowded. People who live in the country may be surprised to learn that urban folk think the country is a safer and easier place to live in the event of a catastrophe, because rural areas are typically the last places to get assistance. In the meantime, it should be interesting to watch the effect that all this stockpiling and withdrawal of cash in advance "just to be safe" will have on the economy. The secondary effects may just turn out to be the bigger problem.

—Wendy M. Grossman

WENDY M. GROSSMAN, a London-based writer and author of net.wars *(New York University Press, 1997), has stockpiled several dozen bags of chocolate chips.*

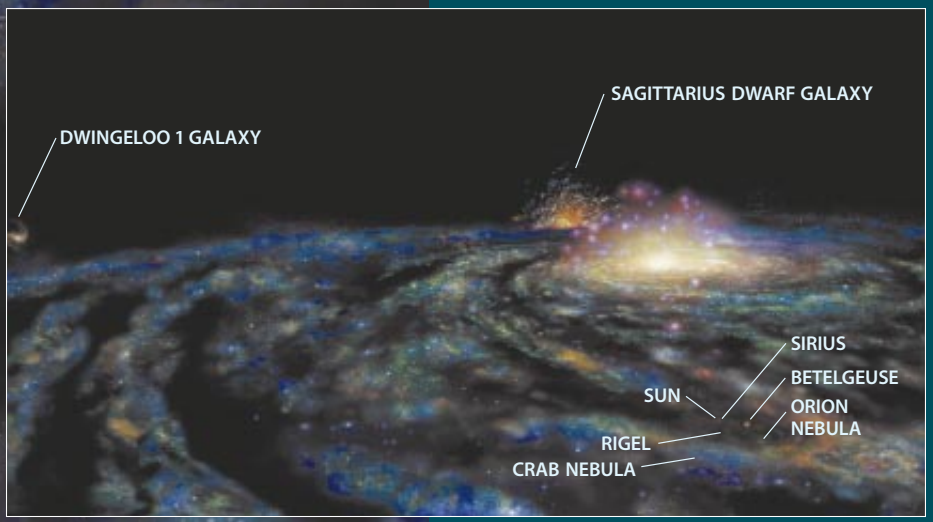
Galaxies behind the Milky Way

Over a fifth of the universe is hidden from view, blocked by dust and stars in the disk of our galaxy. But over the past few years, astronomers have found ways to peek through the murk

DISK OF MILKY WAY GALAXY, a cosmic crepe with one trillion suns' worth of stars, dust and gas, prevents us from viewing a fifth of the universe. Among the hidden objects is the Sagittarius dwarf spheroidal galaxy, apparent in these artist's impressions of the view from below (*main illustration*) and above (*inset*) the plane of the Milky Way. Our sight lines to the dwarf are almost completely blocked by the bulge of stars surrounding the center of our galaxy. Although Sagittarius is the closest galaxy to our own, it was discovered only four years ago. Another hidden galaxy, Dwingeloo 1, is shown in the inset.

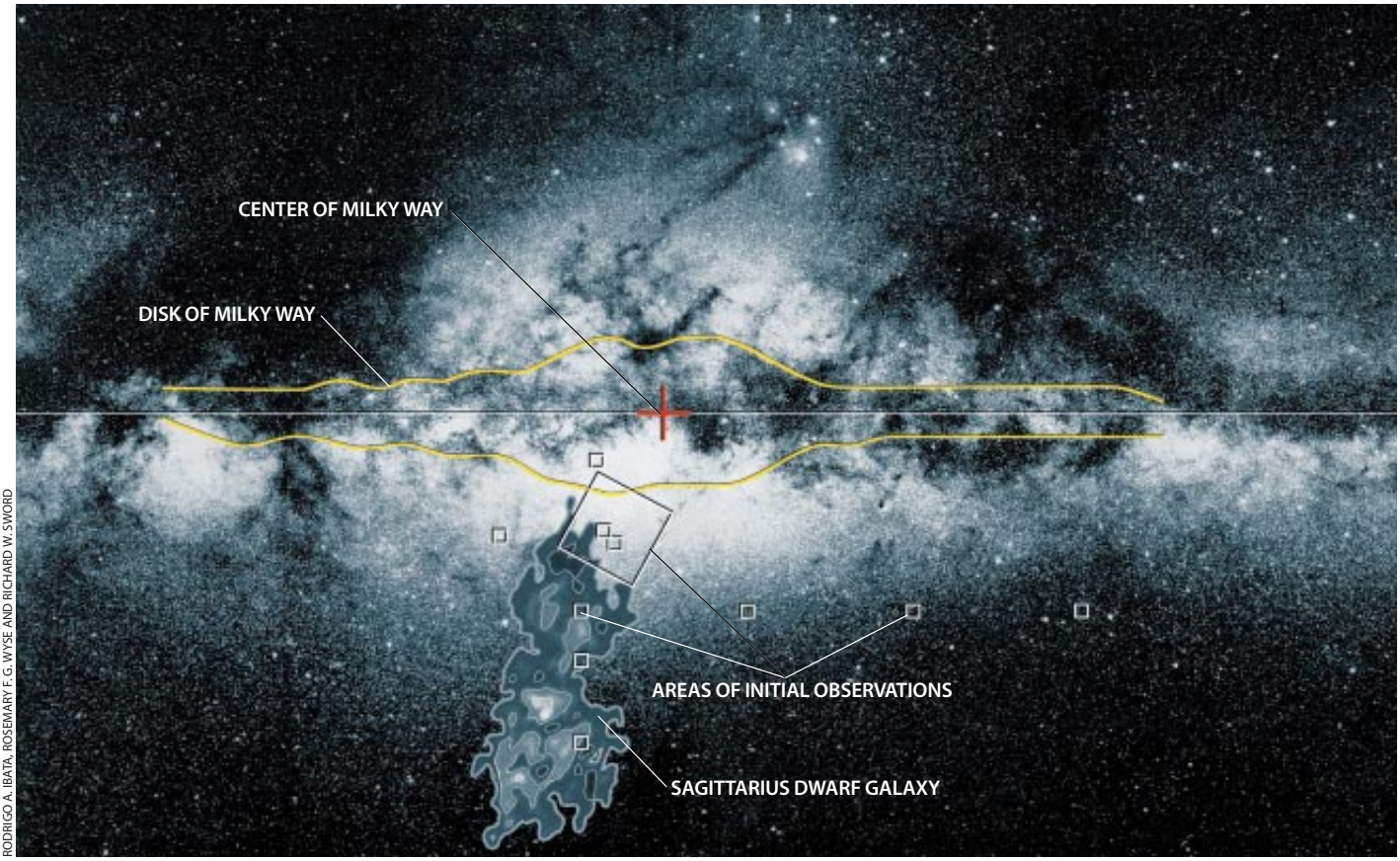
DON DIXON

VIEW FROM 500 LIGHT-YEARS ABOVE GALACTIC PLANE



by Renée C. Kraan-Korteweg and Ofer Lahav

On a dark night, far from city lights, we can clearly see the disk of our galaxy shimmering as a broad band across the sky. This diffuse glow is the direct light emitted by hundreds of billions of stars as well as the indirect starlight scattered by dust grains in interstellar space. We are located about 28,000 light-years from the center of the galaxy in the midst of this disk. But although the Milky Way may be a glorious sight, it is a constant source of frustration for astronomers who study the universe beyond our galaxy. The disk blocks light from a full 20 percent of the cosmos, and it seems to be a very exciting 20 percent.

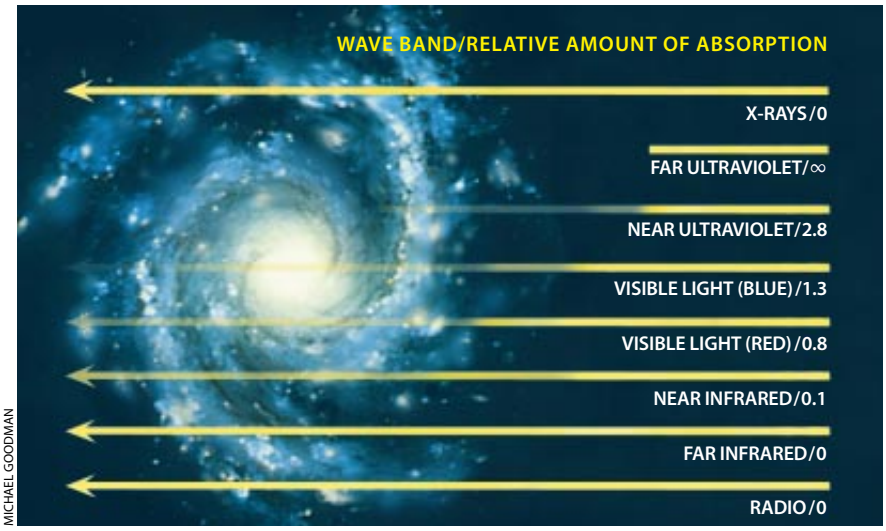


RODRIGO A. IBATA, ROSEMARY F. G. WYSE AND RICHARD W. SWORD

Somewhere behind the disk, for example, are crucial parts of the two biggest structures in the nearby universe: the Perseus-Pisces supercluster of galaxies and the "Great Attractor," a gargantuan agglomeration of matter whose existence has been inferred from the motions of thousands of galaxies through space. Observations also show a tantalizing number of bright and nearby galaxies in the general direction of the disk, suggesting there are many others that go unseen. Without knowing what lies in our blind spot, researchers cannot fully map the matter in our corner of the cosmos. This in turn prevents them from settling some of the most important questions in cosmology: How large are cosmic structures? How did they form? What is the total density of matter in the universe?

Only in recent years have astronomers developed the techniques to peer through the disk and to reconstruct the veiled universe from its effects on those parts that can be seen. Although observers are far from completing this tedious task, some spectacular discoveries have already proved that it is worth the effort. Among other things, astronomers have found a new galaxy so close that it would dominate our skies were it not obscured by the disk. They have found colossal galaxy clusters never before seen and have even taken a first peek at the core of the elusive Great Attractor.

The obscuration of galaxies by the Milky Way was first perceived when astronomers began distinguishing external galaxies from internal nebulae, both seen simply as faint, extended objects. Because galaxies appeared everywhere except in the region of the Milky Way,



MICHAEL GOODMAN

LIGHT FROM OTHER GALAXIES penetrates the Milky Way to varying degrees, depending on its wavelength. The longest wavelengths, which correspond to radio and far-infrared radiation, are hardly affected, but shorter wavelengths (such as near-infrared, visible and ultraviolet light) are blocked by the dust and gas clouds within our galaxy. For very short wavelengths, such as the most powerful x-rays, the gas becomes transparent again.

DWINGELOO OBSCURED GALAXY SURVEY TEAM, SHAUN HUGHES AND STEVE J. MADDOX *Isaac Newton Telescope*

HIDDEN GALAXIES emerge in careful astronomical observations. Dwingeloo 1, first detected by a Dutch radio telescope in 1994, is a spiral galaxy visible as a faint background to the constellation Cassiopeia (*near left*). The Sagittarius dwarf spheroidal galaxy, in contrast, cannot be seen directly even with hindsight; its stars are jumbled with those of the Milky Way and must be identified one by one (*far left*). The Sagittarius image is a mosaic of pictures mainly from the southern sky, projected so that the Milky Way runs horizontally across the center.

this region was named the "zone of avoidance" [*see illustration on page 56*]. Scientists now know that external galaxies consist of billions of stars as well as countless clouds of dust and gas. In the zone of avoidance the light of the galaxies is usually swamped by the huge number of foreground stars or is absorbed by the dust in our own galaxy.

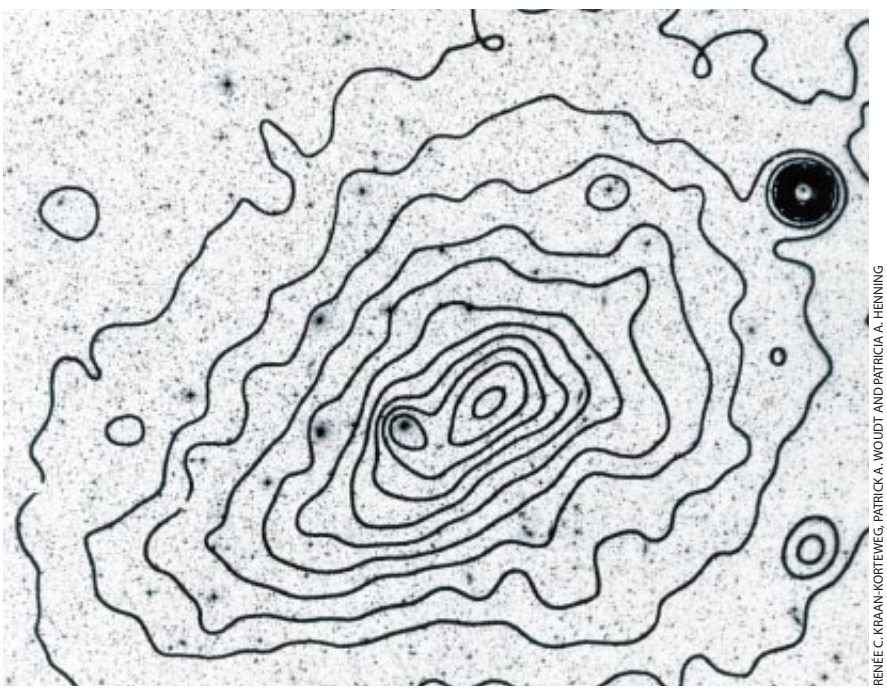
Extragalactic astronomers have generally avoided this zone, too, concentrating instead on unobscured regions of the sky. But 20 years ago a crucial observation hinted at what they might be missing. Crude measurements of the cosmic microwave background radiation, a relic of the big bang, showed a 180-degree asymmetry, known as a dipole. It is about 0.1 percent hotter than average at one location in the sky and equally colder in the catercornered site.

These measurements, confirmed by the Cosmic Background Explorer satellite in 1989 and 1990, suggest that our galaxy and its neighbors, the so-called Local Group, are moving at 600 kilometers per second (1.34 million miles per hour) in the direction of the constellation Hydra. This vector is derived after correcting for known motions, such as the revolution of the sun around the galactic center and the motion of our galaxy toward its neighbor spiral galaxy, Andromeda.

Where does this motion, which is a small deviation from the otherwise uniform expansion of the universe, come from? Galaxies are clumped into groups and clusters, and these themselves agglomerate into superclusters, leaving other regions devoid of galaxies. The clumpy mass distribution surrounding the Local Group may exert an unbalanced gravitational attraction, pulling it in one direction. At first glance, it might seem hard to believe that galaxies could influence one another over the vast distances that separate them. But relative to their masses, galaxies are closer to one another than individual stars within our galaxy are.

The expected velocity of the Local Group can be calculated by adding up the gravitational forces caused by known galaxies. Although the resulting vector is within 20 degrees of the observed cosmic background dipole, the calculations remain highly uncertain, partly because they do not take into account the galaxies behind the zone of avoidance.

The lingering discrepancy between the dipole direction and the expected velocity vector has led astronomers to postulate "attractors." One research group, later referred to as the Seven Samurai, used the motions of hundreds of galaxies to deduce the existence of the Great Attractor about 200 million light-years away [see "The Large-Scale Streaming of Galaxies," by Alan Dressler; Scientific American, September 1987]. The Local Group seems to be caught in



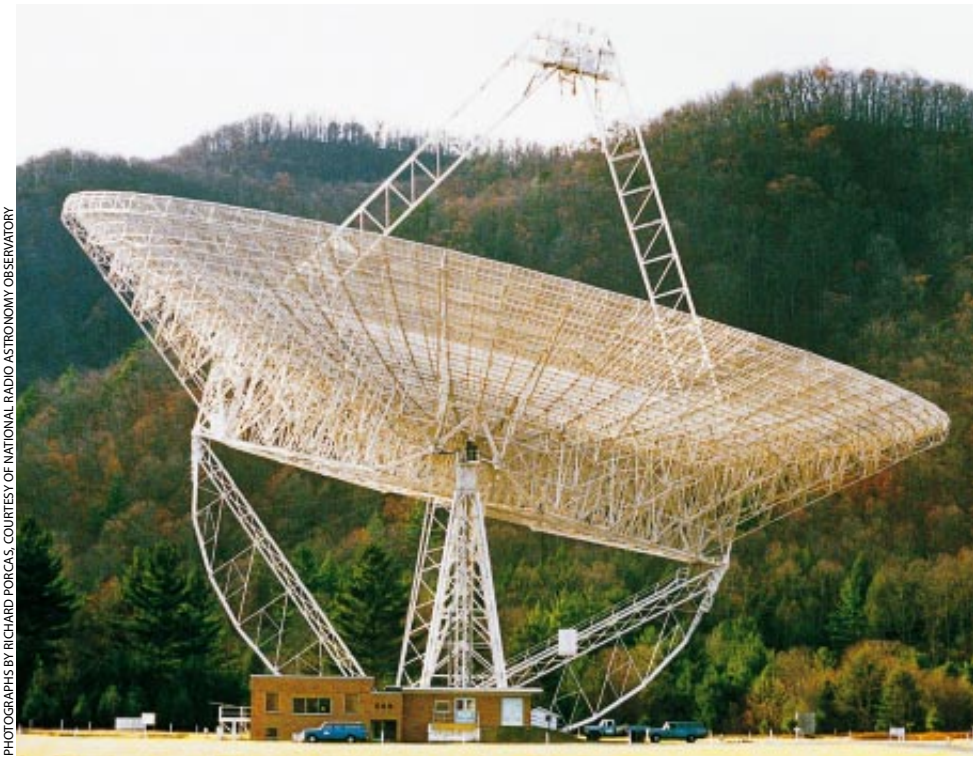
RENÉE C. KRAAN-KORTEWEG, PATRICK A. WOUDT AND PATRICIA A. HENNING

CORE OF "GREAT ATTRACTOR" has been identified as galaxy cluster Abell 3627. It appears both in a visible-light image (*background*) and in x-ray observations (*contours*). Over 100 galaxies show up in this negative image; most of the dots are stars in our own galaxy. The tight concentric contours (*top right*) mark a bright galaxy in the cluster.

a cosmic tug of war between the Great Attractor and the equally distant Perseus-Pisces supercluster, which is on the opposite side of the sky. To know which will win the war, astronomers need to know the mass of the hidden parts of these structures.

Both are components of a long chain of galaxies known as the Supergalactic Plane. The formation of such a megastructure is thought to depend on the nature of the invisible dark matter that makes up the bulk of the universe. Chains of galaxies should be more likely in a universe dominated by particles of so-called hot dark matter (such as massive neutrinos) rather than by cold dark matter (such as axions or other hypothetical particles). But astronomers cannot distinguish between these two possibilities until they map the structures fully.

Nearby galaxies are not to be ignored in the bulk motion of the Local Group. Because gravity is strongest at small distances, a significant force is generated by the nearest galaxies, even if they are not massive. And it is intriguing that five of the eight apparently brightest galaxies lie in the zone of avoidance; they are so close and bright that they shine through the murk. These galaxies belong to the galaxy groups Centaurus A and IC342, close neighbors to our Local Group. For each member of these groups that astronomers manage to see, there are probably many others whose light is entirely blocked.



PHOTOGRAPHS BY RICHARD PORCAS, COURTESY OF NATIONAL RADIO ASTRONOMY OBSERVATORY

RADIO TELESCOPE at Green Bank, W.Va., 91 meters (300 feet) in diameter, was busy conducting observations on Nov. 15, 1988 (*left*). The next day it was a heap of rubble (*right*). Scientists at the National Radio Astronomy Observatory say the dish collapsed

Lifting the Fog

Our vantage point, to be sure, could be worse. If we lived in the nearby Andromeda galaxy, the obscured part of the sky would not be much different, yet we would also lose our clear view of the nearest galaxy cluster in Virgo. But even a habitual optimist would admit that we are somewhat unlucky. Because the orbit of the sun about the galactic center is inclined to the galactic plane, the solar system partakes in an epicyclic motion above and below the plane. Currently we are elevated only 40 light-years from the plane. If we had been born 15 million years from now, we would be located nearly 300 light-years above the plane—beyond the thickest layer of obscuration—and could view one side of the current zone of avoidance. It will take another 35 million years to cross the disk of the Milky Way to the other side.

Most astronomers do not want to wait that long to learn about the extragalactic sky behind the zone of avoidance. What can they do in the meantime? A first step is careful review of existing visible-light images. The dust in the zone does not completely blot out every galaxy; some poke through, although they seem dimmer and smaller the closer they are to the middle of the galactic plane. The odd appearance of these galaxies, in combination with the high density of foreground stars, can confuse the computer software used to analyze images and recognize galaxies. So various groups of astronomers have gone back to the old-fashioned way of examining images—by eye. Photographic plates from the Palomar Observatory sky survey and its Southern Hemisphere counterpart, conducted in the 1950s, have been painstakingly searched over the past 10 years. Researchers have covered a major fraction of the zone of avoidance, identifying 50,000 previously uncatalogued galaxies.

In areas where the extinction of light by dust is too severe, however, galaxies are fully obscured, and other methods are required. The leading option is to observe at longer wavelengths; the longer the wavelength, the less the radiation interacts with microscopic dust particles. The 21-centimeter spectral line emitted by electrically neutral hydrogen gas is ideal in this respect. It traces gas-rich spiral galaxies, intrinsically dim galaxies and dwarf galaxies—that is, most galaxies except gas-poor elliptical galaxies.

In 1987 a pioneering 21-centimeter project was launched by Patricia A. Henning of the University of New Mexico and Frank J. Kerr of the University of Maryland. They pointed the 91-meter radio telescope at Green Bank, W.Va., toward random spots in the zone of avoidance and detected 18 previously unknown galaxies. Unfortunately, the telescope collapsed spectacularly before they could finish their project. (Its replacement is due to be completed next year.) A more systematic survey was initiated by an international team that includes us. Conducted at the 25-meter Dwingeloo radio telescope in the Netherlands, this longer-term project is mapping all the spiral galaxies in the northern part of the zone of avoidance out to a distance of 175 million light-years. So far it has discovered 40 galaxies.

Last year another international col-

when a metal plate in the main support gave way. The $75-million replacement, now under construction, will be a paraboloid 110 by 100 meters; unlike traditional circular dishes, the new telescope will not have any support struts that block its view of the sky.

laboration, led by Lister Staveley-Smith of the Australia Telescope National Facility in Marsfield and one of us (Kraan-Korteweg), began an even more sensitive survey of the southern Milky Way. This survey, which maps galaxies out to 500 million light-years, uses a custom-built instrument at the 64-meter radio telescope at Parkes, Australia. More than 100 galaxies have already been detected, and thousands more are expected when the survey reaches its full depth.

The radio-wave bands are not the only possible peepholes through the zone of avoidance. Infrared light, too, is less affected by dust than visible light is. In the early 1980s the Infrared Astronomical Satellite (IRAS) surveyed the whole sky in far-infrared wavelengths (those closer to radio wavelengths). It tentatively identified infrared-bright galaxies, particularly spirals and starburst galaxies, in which stars are forming rapidly and plentifully. IRAS-selected galaxy candidates near the zone of avoidance are now being reexamined with images taken in the near-infrared wavelengths (those closer to visible light).

Two systematic near-infrared surveys, due to be finished in 2000, are also under way: the Two Micron All-Sky Survey, an American project, and DENIS, a European project that focuses on the Southern Hemisphere. Both surveys take digital images in three wave bands that probe the older stellar population in galaxies. The surveys easily trace the elliptical galaxies found at the center of dense galaxy concentrations; they therefore complement the far-infrared and 21-centimeter bands, which predominantly find spiral galaxies. A pilot study has shown that the near-infrared surveys do indeed uncover galaxies that fail to register on visible-light photographs. Unfortunately, neither visible nor infrared light can pick out galaxies in the thickest parts of the galactic plane.

Another possible way to overcome the obscuration is to observe at very short wavelengths, such as x-rays. Highly populated galaxy clusters emit copious x-rays, which pass through the Milky Way almost unhindered. But an x-ray investigation, which could draw on existing data from ROSAT and other satellites, has not been done yet.

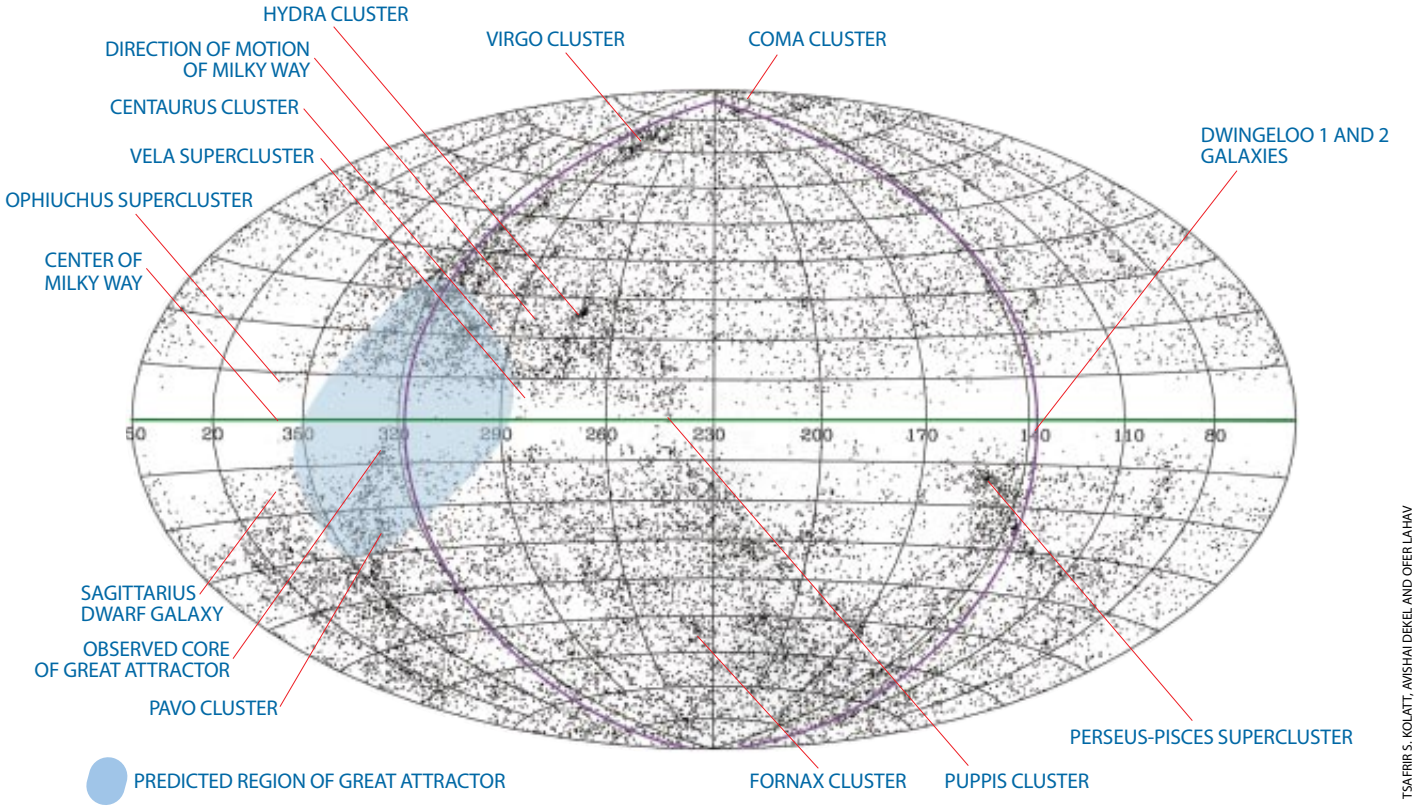
In addition to direct observations, astronomers are exploring the zone of avoidance by indirect means. Signal-processing techniques, commonly applied by engineers to noisy and incomplete data, have been used successfully by researchers at the Hebrew University and one of us (Lahav) to predict the existence of clusters such as Puppis and Vela, as well as the continuity of the Supergalactic Plane across the zone. The galaxy velocities can also be used on both sides of the zone to predict the mass distribution in between. With this method the center of the Great Attractor was predicted to lie on a line connecting the constellations Centaurus and Pavo. These reconstruction methods, however, deduce only the largest-scale features across the zone; they miss individual galaxies and smaller clusters.

Prey of the Milky Way

Such methods are slowly opening up the hidden fifth of the universe to astronomical investigation. A most surprising discovery came in 1994, when Rodrigo A. Ibata, then at the University of British Columbia, Gerard F. Gilmore of the University of Cambridge and Michael J. Irwin of the Royal Greenwich Observatory in Cambridge, England, who were studying stars in our Milky Way, accidentally found a galaxy right on our doorstep. Named the Sagittarius dwarf, it is now the closest known galaxy—just 80,000 light-years away from the solar system, less than half the distance of the next closest, the Large Magellanic Cloud. In fact, it is located well inside our galaxy, on the far side of the galactic center.

Because the Sagittarius dwarf lies directly behind the central bulge of the Milky Way, it cannot be seen in direct images. Its serendipitous detection was based on velocity measurements of stars: the researchers spotted a set of stars moving differently from those in our galaxy. By pinpointing the stars with this velocity, looking for others at the same distance and compensating for the light of known foreground stars, they mapped out the dwarf [*see top illustration on page 52*]. It extends at least 20 degrees from end to end, making it the largest apparent structure in the sky after the Milky Way itself. Its angular size corresponds to a diameter of at least 28,000 light-years, about a fifth of the size of our galaxy, even though the dwarf is only a thousandth as massive.

Many popular models of galaxy formation postulate that large galaxies are formed by a long process of aggregation



TSAFRIR S. KOLATT, AVISHAI DEKEL AND OFER LAHAV

30,000 GALAXIES, culled from three standard astronomical catalogues, are shown as dots on this map. The galaxies appear all over the sky except in the so-called zone of avoidance, which corresponds to the plane of our Milky Way galaxy (*green horizontal center line*). Outside the zone, the galaxies tend to clump near a line that traces out the Supergalactic Plane (*purple line*).

of many smaller galaxies. Such a process should still be common today, yet has been observed only rarely. Sagittarius appears to have undergone some disruption from the tidal forces exerted by the Milky Way, but the disruption of the core of Sagittarius is unexpectedly minor. The dwarf may have orbited our galaxy 10 times or more yet remains largely intact, indicating that it is held together by large amounts of dark matter (as opposed to luminous matter such as stars or gaseous clouds). Even so, its demise is just a matter of time; some studies suggest that Sagittarius may have only another billion years to go before being swallowed by our galaxy. Its discovery has demonstrated that mergers do happen, that they happen today and that they do not necessarily wreck the disk of the larger galaxy.

Sagittarius is one of many surprises to have surfaced from the zone of avoidance. In August 1994 we and the rest of the Dwingeloo Obscured Galaxy Survey team examined our first 21-centimeter spectra. We selected a region where many filaments are lost in the zone and where the nearby galaxy group IC342 resides. Quite soon we came across an intriguing radio spectrum in the direction of the constellation Cassiopeia. Radio observations are prone to interference, which can mimic extragalactic radio profiles; moreover, the feature blended with the emission from galactic gas. Yet various tests confirmed the signal, marking the discovery of another previously unknown nearby galaxy.

George K. T. Hau of the University of Cambridge identified an extremely dim visible-light object that matched the location of this radio signal. Before long, deeper images were obtained at various telescopes, which fully revealed the shape of the galaxy: a bar with spiral arms protruding at its ends [*see top illustration on page 53*]. If it were not lying behind the plane of the Milky Way, the galaxy—named Dwingeloo 1—would be one of the 10 brightest in the sky. Judging from its rate of rotation it has about one third the mass of the Milky Way, making it comparable to M33, the third heaviest galaxy of the Local Group after the Milky Way and Andromeda.

While conducting follow-up observations of Dwingeloo 1, the Westerbork Synthesis Radio Telescope in the Netherlands discovered a second galaxy just one third of a degree away: Dwingeloo 2, a dwarf galaxy with half the diameter and a tenth of the mass of Dwingeloo 1. Located at a distance of 10 million light-years, the pair of galaxies is close to, but just beyond, the Local Group. They seem to be associated with IC342. Two other galaxies in this assemblage were later discovered on sensitive optical images.

Although astronomers have yet to explore the entire zone of avoidance, they can now rule out other Andromeda-size galaxies in our backyard. The Milky Way and Andromeda are indeed the dominant galaxies of the Local Group. Disappointing though the lack of another major discovery may be, it removes the uncertainties in the kinematics of our immediate neighborhood.

Clusters and Superclusters

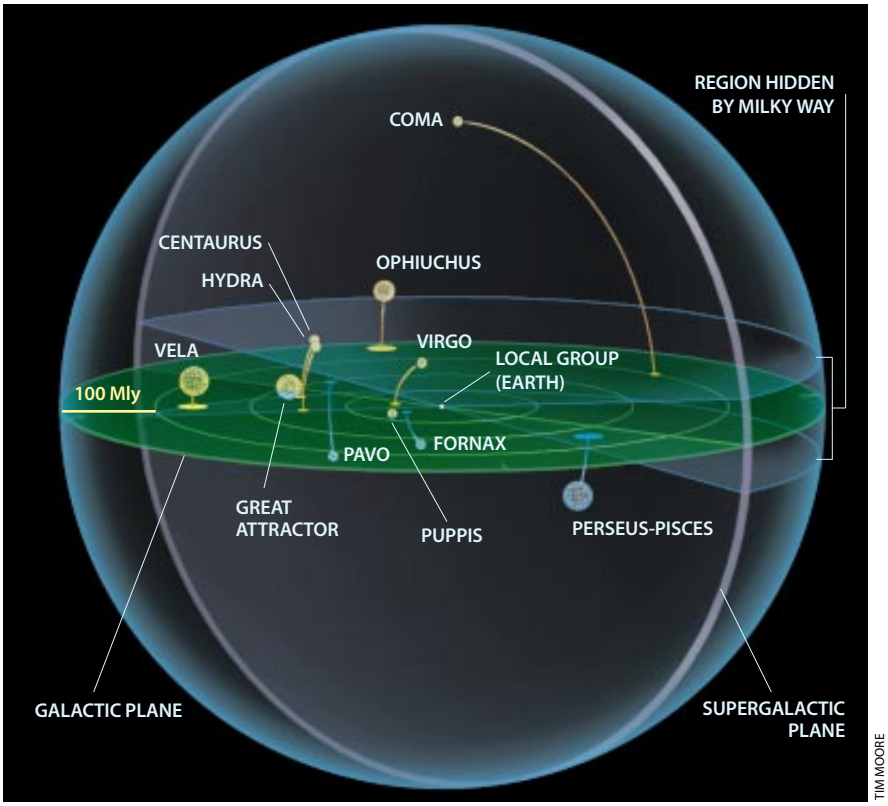
Studies in the zone of avoidance have also upset astronomers' ideas of the more distant universe. Using the 100-meter radio telescope near Effelsberg, Germany, astronomers discovered a new cluster 65 million light-years away in the constellation Puppis. Several other lines of evidence, including an analysis of galaxies discovered by IRAS, have converged on the same conclusion: the

inclusion of Puppis brings the expected motion of the Local Group into better agreement with the observed cosmic background dipole.

Could these searches demystify the Great Attractor? Although the density of visible galaxies does increase in the attractor's presumed direction, the core of this amorphous mass has eluded researchers. A cluster was identified in roughly the right location by George O. Abell in the 1980s, at which time it was the only known cluster in the zone of avoidance. But with a mere 50 galaxies, it could hardly amount to an attractor, let alone a great one.

The true richness and significance of this cluster has become clear in the recent searches. Kraan-Korteweg, with Patrick A. Woudt of the European Southern Observatories in Garching, Germany, has discovered another 600 galaxies in the cluster. With colleagues in France and South Africa, we obtained spectral observations at various telescopes in the Southern Hemisphere. The observed velocities of the galaxies suggest that the cluster is very massive indeed—on par with the well-known Coma cluster, an agglomeration 10,000 times as massive as our galaxy. At long last, astronomers have seen the center of the Great Attractor. Along with surrounding clusters, this discovery could fully explain the observed galaxy motions in the nearby universe.

The hierarchy of cosmic structure does not end there. Searches in the zone of avoidance have identified still larger clumpings. One supercluster 370 million light-years away in the constellation Ophiuchus was identified by Kenichi Wakamatsu of Gifu University in Japan. Although this supercluster lies behind the galactic center, a region extremely crowded with stars, Wakamatsu identified thousands of its galaxies on sky-survey plates. The Ophiuchus supercluster might be connected to another supercluster in the constellation Hercules, suggesting coherent structures on scales that are mind-boggling even to astronomers.



THREE-DIMENSIONAL VIEW of the local universe reveals the uneven distribution of galaxy clusters. The blue sphere represents a distance of 400 million light-years (Mly) from the Milky Way, the green plane is the galactic plane extended into intergalactic space, the small knots of dots are galaxy clusters, and the circles are their projections onto the galactic plane. Many galaxy clusters lie on or near the Supergalactic Plane (*purple*). Some are hidden in the zone of avoidance (*gray wedge*).

For generations of astronomers, the zone of avoidance has been an obstacle in investigating fundamental issues such as the formation of the Milky Way, the origin of the Local Group motion, the connectivity of chains of galaxies and the true number of galaxies in the universe. The efforts over the past decade to lift this thick screen have turned the former zone of avoidance into one of the most exciting regions in the extragalactic sky. The mysterious Great Attractor is now well mapped; the discovery of the Sagittarius dwarf has shown how the Milky Way formed; and the vast cosmic filaments challenge theories of dark matter and structure formation. More surprises in this *caelum incognitum* may await astronomers. Step by step, the missing pieces of the extragalactic sky are being filled in. SA

The Authors

RENÉE C. KRAAN-KORTEWEG and OFER LAHAV joined forces in 1990, after they met in Durham, England, at a conference on cosmology; independently, they both had discovered a previously unknown cluster behind the Milky Way in the constellation Puppis. Kraan-Korteweg is a professor in the department of astronomy of the University of Guanajuato in Mexico. Lahav is a faculty member of the Institute of Astronomy at the University of Cambridge and a Fellow of St. Catharine's College. Kraan-Korteweg explores the zone of avoidance by direct observation, whereas Lahav utilizes theoretical and computational techniques.

Further Reading

PRINCIPLES OF PHYSICAL COSMOLOGY. P.J.E. Peebles. Princeton University Press, 1993.

UNVEILING LARGE-SCALE STRUCTURES BEHIND THE MILKY WAY. Edited by Chantal Balkowski and R. C. Kraan-Korteweg. Astronomical Society of the Pacific Conference Series, Vol. 67; January 1994.

A DWARF SATELLITE GALAXY IN SAGITTARIUS. R. A. Ibata, G. Gilmore and M. J. Irwin in *Nature,* Vol. 370, pages 194–196; July 21, 1994.

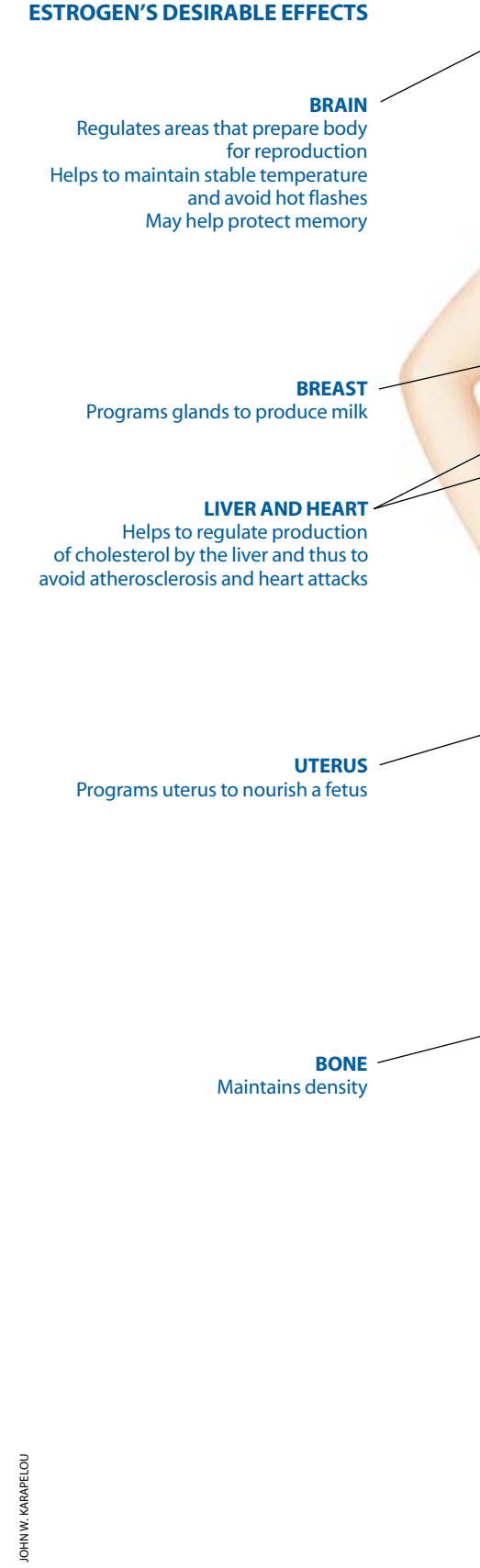
DYNAMICS OF COSMIC FLOWS. Avishai Dekel in *Annual Review of Astronomy and Astrophysics,* Vol. 32, pages 371–418; 1994.

A NEARBY MASSIVE CLUSTER BEHIND THE MILKY WAY. R. C. Kraan-Korteweg et al. in *Nature,* Vol. 379, pages 519–521; February 8, 1996.

Designer Estrogens

These compounds—also called SERMs—have evolved from mere laboratory curiosities into drugs that hold promise for preventing several major disorders in women

by V. Craig Jordan



JOHN W. KARAPELOU

Dramatic research findings have recently focused new attention on so-called designer estrogens, the medicines known more technically as selective estrogen receptor modulators (SERMs). These agents behave like estrogen in some tissues but block its action in others.

This past spring a large study demonstrated that one such compound—the cancer treatment tamoxifen—can prevent breast cancer in women as young as 35 who are at high risk for the disease. No drug had ever before been found to avert primary breast tumors. Soon after, a preliminary report revealed that a related agent, raloxifene, may similarly protect against breast cancer. This time the drug was tested in postmenopausal women who, aside from advanced age, had no particular signs of elevated risk.

Further, various lines of evidence have begun to suggest that raloxifene (now prescribed for maintaining bone density in older women) or other SERMs in development may be able to shield women single-handedly from a suite of serious disorders that become increasingly prevalent after menopause. Those conditions include breast cancer and osteoporosis (bone thinning that weakens resistance to fractures) as well as endometrial cancer (which affects the uterine lining) and coronary artery disease, the single leading cause of death in women and men alike.

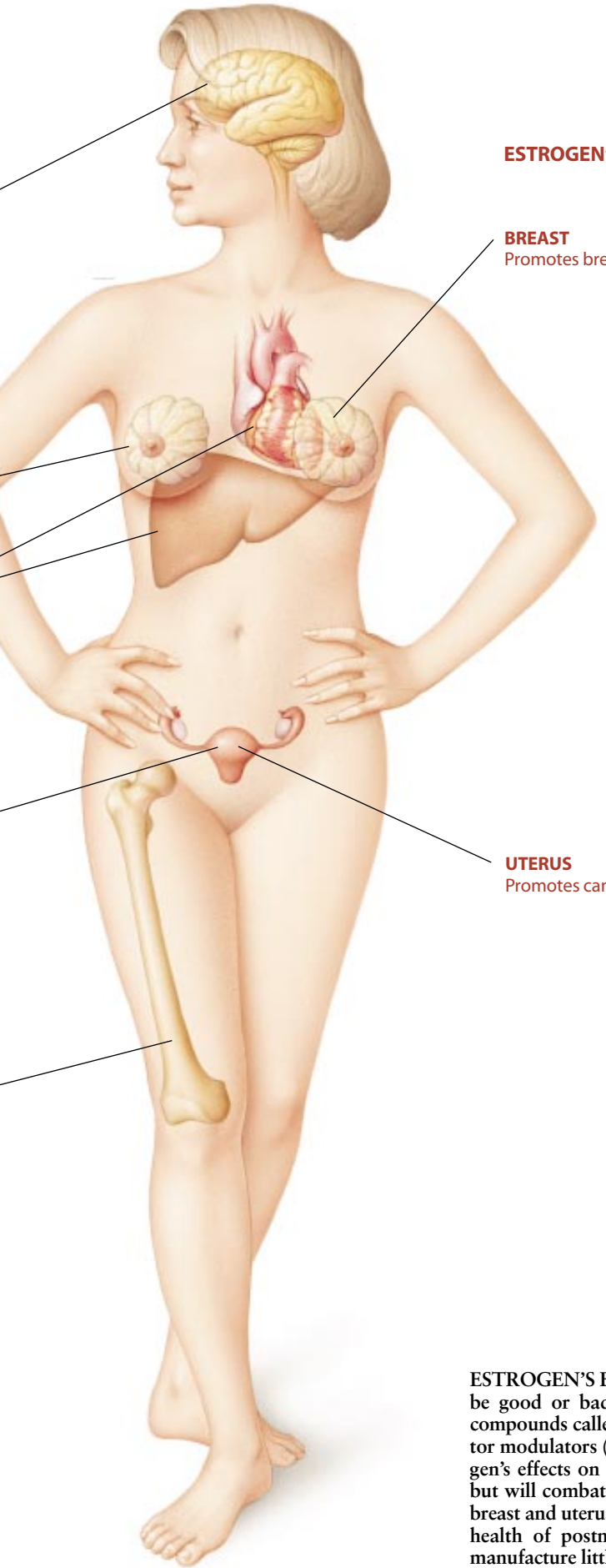
If SERMs fulfill their promise, they should greatly improve women's health in the 21st century. I have spent almost 30 years examining the actions and side effects of these agents, providing clinical investigators with information needed for designing human trials. I am therefore well situated to tell the often surprising tale of how these remarkable compounds evolved from laboratory oddities to cancer therapies to medicines that may preserve the vitality of healthy women as they age.

The story properly begins nearly four decades ago, with tamoxifen's discovery. The narrative will be most comprehensible, however, if I briefly review estrogen's role in the body and explain why drugs that selectively mimic or impede this hormone are so intriguing.

Why Regulate Estrogen?

Estrogen turns out to be a contradictory molecule, both crucial and harmful to women. The most obvious benefit is the programming of the female body (especially the breast, uterus and brain) for reproduction, a function essential to perpetuating the human race.

During the past 20 years, the hormone has also been found to act on other organs in ways that optimize overall health. For instance, it influences brain centers that maintain body temperature, and it enables the vaginal lining to stay thick and lubricated. More important, it safeguards the heart, largely by limiting the buildup of atherosclerotic plaque in the coronary arteries. The arteries are protected in part by estrogen's ability to modulate the manufacture of cholesterol in the liver; it restrains the output of atherosclerosis-promoting LDL (low-density lipoprotein) cholesterol, and it elevates that of plaque-fighting HDL (high-density lipoprotein) cho-



ESTROGEN'S EFFECTS on a woman can be good or bad. Drugmakers hope that compounds called selective estrogen receptor modulators (SERMs) will mimic estrogen's effects on the liver, heart and bones but will combat its harmful effects on the breast and uterus. The aim is to protect the health of postmenopausal women, who manufacture little estrogen of their own.

lesterol. Estrogen also preserves bone, helping to keep a balance between its creation and degradation. And recent findings hint that it may support memory and delay or ease Alzheimer's disease.

Many women best appreciate estrogen's value when they confront its loss during and after menopause—the phase, usually beginning around age 50, when the ovaries exhaust their supply of eggs and stop making the hormone. Hot flashes and night sweats may occur for several years. Meanwhile rises in LDL cholesterol and other cardiovascular changes progressively increase the odds of coronary disease and heart attacks. In parallel, bone thinning, which begins slowly in a woman's late 30s, accelerates, often resulting in crippling fragility and deformity in her 70s.

Estrogen's more malevolent aspect—an ability to promote cancer of the breast and of the uterine lining—also becomes most evident in the years following menopause. The hormone probably does not initiate the cellular changes that result in these cancers, but it does stimulate the growth of cells that have taken the first steps toward uncontrolled proliferation. Decades of exposure to her own estrogen greatly elevate a woman's chances of having a detectable tumor, which makes advancing age a major risk for both cancers. Breast cancer, though, is much more common, affecting one woman in every 15 between the ages of 60 and 79.

After women stop ovulating, they frequently try to compensate for the loss of estrogen's desired effects by taking a synthetic or natural replacement. Such replacement therapy eases hot flashes and, with long-term use, guards against osteoporosis. In addition, several retrospective studies suggest that it reduces heart attacks, although a definitive, prospective study remains to be done. Yet replacement therapy presumably promotes breast and endometrial cancer, and in some women it triggers formation of potentially dangerous blood clots in the veins. Addition of progestin, a synthetic form of the hormone progesterone, to the regimen protects against endometrial cancer, by inducing monthly shedding of the uterine lining (including any malignant cells). But many older women dislike having "periods" again, and progestins do not alleviate the threat of breast cancer or of blood clots.

A major hope for SERMs is that some will replicate estrogen's critical benefits for the bones and heart but will act as antiestrogens (estrogen blockers) in the

breast and uterus, thus avoiding estrogen's key dangers. Because SERMs can worsen hot flashes, however, any that meet those demands will probably be used primarily by women whose acute menopausal symptoms have subsided.

I should note that exercise and careful attention to diet and to calcium intake can help preserve the bones and heart without drugs; some scientists contend these approaches may even fend off breast and endometrial cancer. But how well they stack up against drug therapy has yet to be examined rigorously.

Tamoxifen: The Pioneer

Although SERMs are suddenly in the spotlight for their disease-preventing potential, that capacity was never considered when tamoxifen, the best-known member of the class, was uncovered some 40 years ago. In fact, no one had yet realized that it, or any chemical, could be estrogenic in some tissues and antiestrogenic in others. What is more, tamoxifen, then classified solely as an antiestrogen, was not expected to forestall or treat any disease at all. It was going to be a contraceptive. Its failure in that role was the first of many disappointments that would, ironically, become valuable revelations—pointing

How SERMs Work

Investigators do not yet know how SERMs, such as tamoxifen and raloxifene, can be antiestrogenic in some cells and estrogenic in others. But speculation abounds.

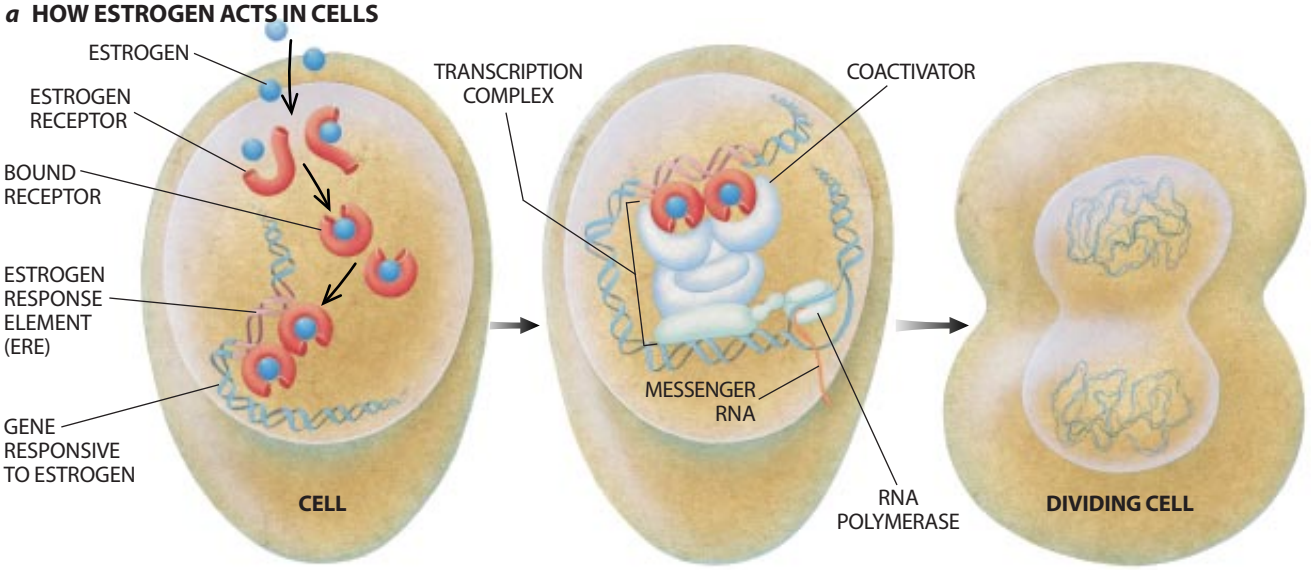
All our ideas build on an understanding of how estrogen works (*sequence in a, below*). The hormone affects only cells that make nuclear proteins called estrogen receptors. When these nuclear proteins encounter estrogen (*1*), they bind to it, change shape, form pairs and attach to docking sites, termed estrogen response elements (EREs), in certain genes. This attachment, in turn, triggers formation of a transcription complex (*2*), a cluster of so-called coactivators and other proteins that fit around the receptors like pieces in a jigsaw puzzle. The completed complex then activates the bound genes. It induces an enzyme, RNA polymerase, to transcribe the genes into molecules of messenger RNA—mobile RNA templates from which new proteins are made. These proteins then induce the cell to divide or change in other ways (*3*).

When SERMs block estrogen action, they do so by plugging the estrogen binding site on the receptor (*b*), thus preventing estrogen from gaining access. SERM binding in those cells also probably keeps the receptor from adopting the shape needed for interacting with coactivator proteins that cooperate with the receptor in forming a normal transcription complex. Evidence for that possibility comes from the laboratory of Roderick E. Hubbard at the University of York in England. In comparing the structure of a receptor bound by estrogen with one bound by raloxifene, the researchers found that estrogen causes a helical region of the receptor—helix 12—to swing across the binding site, much as a restraining bar on an amusement park ride might swing across a seat to lock passengers into place. This displacement of helix 12 positions certain of its constituent amino acids so that they can combine with particular coactivators. Binding by raloxifene, in contrast, jams the "hinge" of the bar, so that helix 12 cannot rotate properly [*see illustration on page 64*].

My laboratory has found that raloxifene and tamoxifen jam the hinge by binding to a specific amino acid—number 351—in the string of amino acids that compose the receptor. This discovery helps to explain some instances of lost responsiveness to tamoxifen in breast cancer patients originally treated successfully with the drug. In these cases, the tumors alter the receptor so that the proper amino acid in position 351 is replaced by a different one. This substitute amino acid reacts to tamoxifen by allowing helix 12 to rotate normally. In consequence, the receptor behaves as if it were bound by estrogen, not tamoxifen, and the tumor grows.

But how would SERMs that are antiestrogens in some healthy cells end up mimicking estrogen in others? Presumably, the drugs are affecting the receptor identically in all cell types, so the variation

a **HOW ESTROGEN ACTS IN CELLS**



1 Estrogen binds to its receptor, which then binds to certain genes

2 A transcription complex forms and activates gene transcription

3 Cell behavior changes

to new, more profound applications for tamoxifen and its relatives.

In the early 1960s Dora N. Richardson, Michael J. K. Harper and Arthur L. Walpole of ICI Pharmaceuticals (now Zeneca) in England identified tamoxifen as an antiestrogen, after seeing that it kept estrogen from increasing the size of the rat uterus. In the era of "make love, not war," sales of birth-control pills were soaring. The team realized that tamoxifen, by inhibiting estrogen in the uterus, might block pregnancy and serve as a convenient "morning after" pill. Yet in subsequent studies, the compound fell flat. In fact, instead of interfering with pregnancy, it enhanced the odds of conception in subfertile women.

Fortunately for breast cancer victims, Walpole had reason to suspect that ICI's apparently useless contraceptive might offer a new way to treat breast tumors. Charged with producing birth-control agents, he could not pursue the notion himself. Nevertheless, he vigorously encouraged others to do so. One of those people was me. I first met Walpole in 1967, when I was a student intern at ICI. Later, in 1972, he served as the chief examiner of my doctoral dissertation, on the action of antiestrogens.

Walpole based his anticancer hopes for tamoxifen on emerging understand-

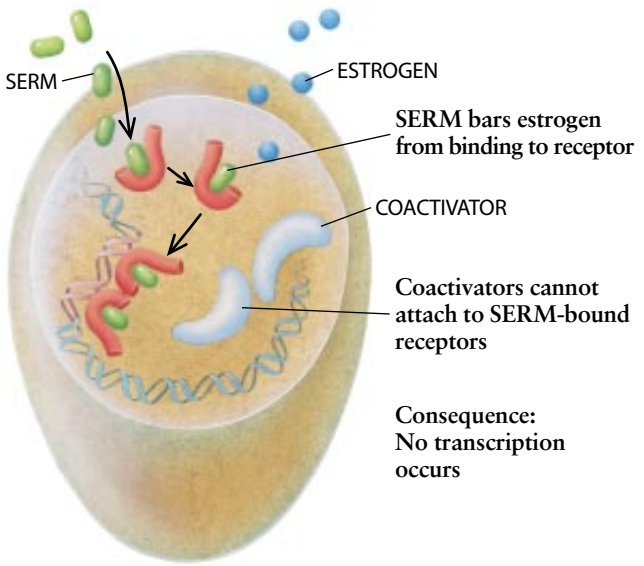
in response has to lie with some features that differ between cells.

One hypothesis focuses on the coactivator proteins that interact with receptors on genes (*c, top*). Cells that show no estrogenlike responses to SERMs may produce no coactivators able to interact with SERM-bound receptors, whereas cells that do respond may, by chance, produce coactivators that can cope with the abnormally shaped receptors and initiate formation of a working transcription complex.

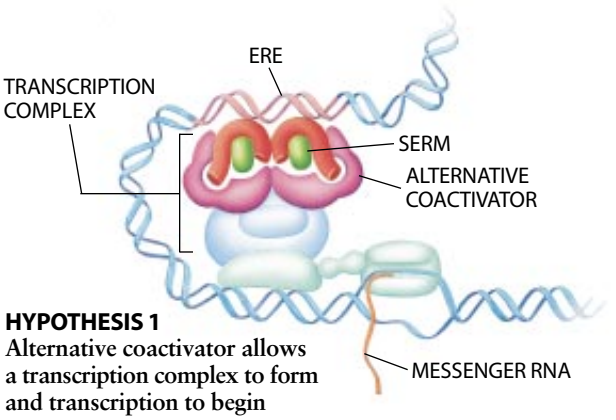
Another idea holds that SERM-bound receptors are unable to bind to standard estrogen response elements in genes. In some cells, however, certain genes reveal alternative docking sites (*c, middle*). When the SERM-bearing receptors land on those sites, transcription complexes form and transcription ensues much as if estrogen were driving it.

Finally, Jan-Åke Gustafsson and his colleagues at the Karolinska Institute in Stockholm have recently discovered that the estrogen receptors come in at least two forms. The one researchers have long known about is now termed the alpha receptor; the new variety is the beta receptor. It is conceivable that in cells that make only alpha receptors, SERMs keep the estrogen receptor from stimulating transcription. But in cells that make beta receptors (*c, bottom*), binding by SERMs may induce those receptors to activate transcription. In this case, some findings suggest, the receptors may abet formation of a transcription complex by binding not to genes directly but to the DNA-binding proteins Fos and Jun. Other explanations for estrogenic activity are under consideration as well. *—V.C.J.*

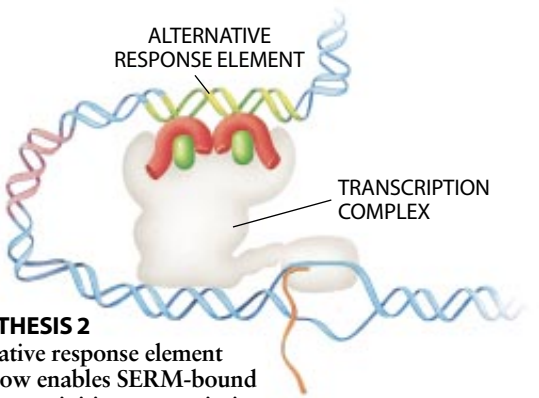
b HOW SERMS BLOCK ESTROGEN ACTION



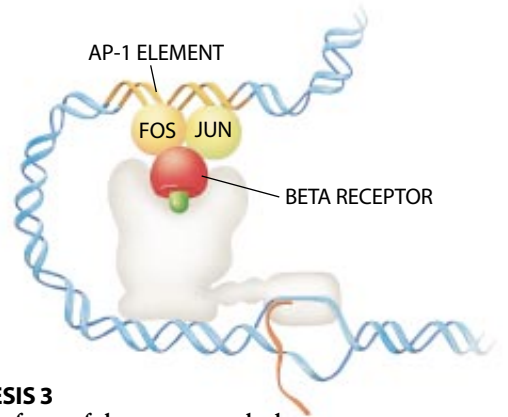
c HOW SERMS MIMIC ESTROGEN



HYPOTHESIS 1
Alternative coactivator allows a transcription complex to form and transcription to begin



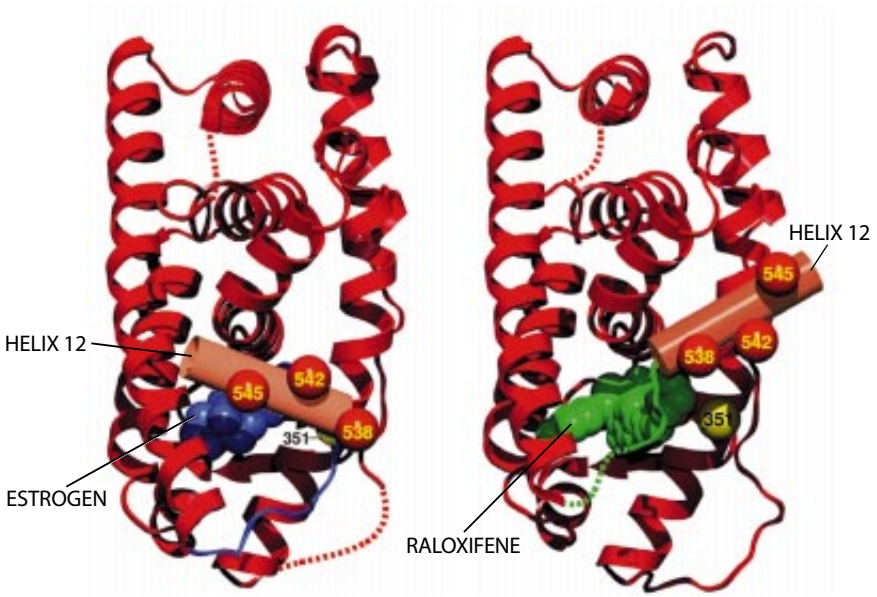
HYPOTHESIS 2
Alternative response element somehow enables SERM-bound receptors to initiate transcription



HYPOTHESIS 3
Alternative form of the receptor—the beta receptor—initiates transcription when the standard, alpha type is unable or is absent

TOMO NARASHIMA

ASHLEY PIKE *University of York*; ADAPTED FROM ANDRZEJ M. BRZOZOWSKI ET AL. (*Nature*, October 16, 1997)



STRUCTURAL FINDINGS have helped to explain the antiestrogenic activity of raloxifene. When estrogen binds to the estrogen receptor (*red ribbon*), helix 12 in the receptor folds across the estrogen molecule (*left*). This rotation positions certain amino acids (*red balls*) so that they can mesh with other molecules needed to activate estrogen-responsive genes. But the helix does not rotate when raloxifene binds (*right*). Lack of rotation may prevent the receptor from interacting properly with the activating molecules. Raloxifene appears to impede rotation by binding to amino acid 351.

ings of how estrogen contributes to breast cancer. Researchers had known for decades that the hormone somehow promoted the growth of breast tumors. By the turn of the 20th century, for example, physicians had learned that removing the ovaries, the main source of estrogen in females, could induce a degree of tumor regression in about a third of those with advanced breast cancer (disease that has spread beyond the lymph nodes under the arms).

A distinguished French scientist, Antoine Lacassagne, subsequently postulated, in 1936, that a drug able to counter estrogen action could prevent breast cancer. In the 1930s, though, there were no antiestrogenic drugs and no molecular target to aim at. No one knew the molecular interactions through which estrogen stimulated tumor growth in the breast or, for that matter, induced maturation of the female reproductive tract.

The target turned up in the late 1950s, when studies by Elwood V. Jensen and Herbert I. Jacobson of the University of Chicago established that estrogen altered the behavior of responsive tissues by linking to a binding molecule, or receptor, produced only in those tissues. About five years later Jack Gorski and his colleagues at the University of Illinois confirmed the receptor's existence by isolating it from cells of the rat uterus. Both groups also predicted correctly that bound receptors in cells then activated genes, which in turn altered the cells' behavior. For instance, in the breast and endometrium, estrogen-stimulated gene activation results in cell division.

Jensen also made the crucial conceptual leap that would eventually catalyze development of tamoxifen as a breast cancer therapy. He reasoned that breast cancer patients who responded to removal of the ovaries had tumors containing abundant estrogen receptors. To grow, these tumors required estrogen to bind to the receptors; in the absence of such contact, the tumors frequently stopped expanding and sometimes disappeared. In contrast, tumors that did not respond to ovarian removal presumably lacked estrogen receptors and thrived even in the absence of estrogen. Hence, women with receptor-rich tumors would very likely benefit from endocrine (estrogen-removing) therapy, whereas women with receptor-poor tumors would not.

Jensen's views led to the tailoring of therapy based on the presence or absence of estrogen receptors in breast tumors. They also suggested by the 1970s that tamoxifen, as an estrogen antagonist, might halt the growth of receptor-bearing breast tumors and even eliminate them without forcing women to undergo surgery to remove estrogen-producing organs. Tamoxifen would bind to estrogen receptors in receptor-rich breast tumors, thereby keeping estrogen from gaining a foothold and from issuing growth signals.

In those days, though, most clinicians were dubious. They instead placed their hope in chemotherapy, the administration of toxic chemicals that travel through the blood, killing stray cancer cells but destroying many healthy cells in the process.

In spite of the prevailing skepticism, after I earned my Ph.D. I set about evaluating tamoxifen's selective ability to decrease the growth of breast tumors. My experiments showed that it blocked estrogen action by binding to estrogen receptors in place of estrogen. In rats, it also shrank mammary tumors bearing estrogen receptors and prevented known carcinogens from producing new tumors.

When those studies were completed, in 1974, at the Worcester Foundation for Experimental Biology in Massachusetts, the research establishment had yet to turn its attention to breast cancer prevention. Hence, for at least a decade, no one attempted prevention trials in humans. Nevertheless, with colleagues at various universities, I continued to conduct laboratory experiments that provided the scientific rationale for a range of human studies of tamoxifen. In the course of this work, we learned that the body converts tamoxifen into several different derivatives, or metabolites. Modified versions of some of the metabolites are now being tested as drugs; one variation became raloxifene.

Tamoxifen Proves Therapeutic

While our investigations proceeded, other teams conducted trials of tamoxifen in women with breast cancer. In the early 1970s two small tests had shown the drug to reduce tumors temporarily in about a third of patients with advanced, widely disseminated disease. Further, the side effects were much milder than those of chemotherapy. On that basis, the British government allowed tamoxifen onto the market in 1973. It became available for late-stage breast cancer in the U.S. in 1978.

Advanced cancer is difficult to cure, but the tumor regression in those early experiments implied that tamoxifen might yield more lasting effects if it were given to women who harbored less cancer. Indeed, I began to wonder, as did others, if patients with more limited

disease might be cured by receiving tamoxifen as adjuvant, or additional, therapy after surgery removed any identifiable tumor. The drug, we hoped, would destroy micrometastases—undetectable tumor cells that had already spread around a woman's body and that, left unopposed, could evolve into fatal masses.

Clinical researchers decided to evaluate the approach by giving patients a year of tamoxifen as adjuvant therapy. My group, meanwhile, worked with animals. We induced formation of microscopic malignancies and then tested the ability of tamoxifen to "cure" the disease, keeping the tiny abnormalities from growing into tumors.

In the human trials and in our tests alike, short-term treatment failed. On the other hand, the animal experiments showed that unbroken, long-term delivery (equivalent to more than five years of tamoxifen delivery in human subjects) blocked tumor development. In the initial phase of therapy the drug acted like a fire blanket: it snuffed out malignancies only as long as it was in the bloodstream. Later, however, it apparently had more enduring effects. These results convinced investigators to try more extended courses in human subjects.

Tamoxifen has now been studied in patients for more than 20 years. This experience has confirmed that longer is better: five years of adjuvant therapy with tamoxifen is superior to one or two years. Investigators at the University of Oxford regularly review clinical research into tamoxifen. Their latest report includes data on 30,000 breast cancer patients whose tumors had not spread beyond the lymph nodes before surgery or radiation therapy was done. Of these women, 18,000 had tumors rich in estrogen receptors, and 12,000 had cancers of unknown receptor status (8,000 of which were probably receptor-bearing). Five years of adjuvant tamoxifen therapy reduced cancer recurrence near the original site or as metastases by close to 50 percent, and the effect persisted for at least five years after the subjects finished their course of tamoxifen. Similarly, women receiving the adjuvant therapy for five years had about a 50 percent reduction in the appearance of tumors in the second breast, and the protection again lasted for at least five years after the drug therapy was completed.

A Troubling Effect Emerges

By 1986 tamoxifen was poised to be prescribed widely as adjuvant therapy for all stages of breast cancer, and at least two major trials of its value for preventing the disease were being planned. With its use about to soar, I became increasingly fearful that long-term delivery would cause unanticipated side effects. My laboratory at the University of Wisconsin–Madison therefore initiated a series of animal experiments to define potential toxicities.

At first we were worried that tamoxifen, as an estrogen antagonist, would inactivate estrogen in tissues that needed it. What would be the advantage of controlling breast cancer with tamoxifen only to increase disability from osteoporosis and death by heart attack?

We began addressing our concerns by evaluating the drug's effects on bone. We looked not only at tamoxifen but also at raloxifene, then an experimental compound that Eli Lilly was pursuing as an alternative antiestrogenic therapy for breast cancer. To our amazement, both ostensible antiestrogens maintained bone density in estrogen-deprived rats, much as estrogen would.

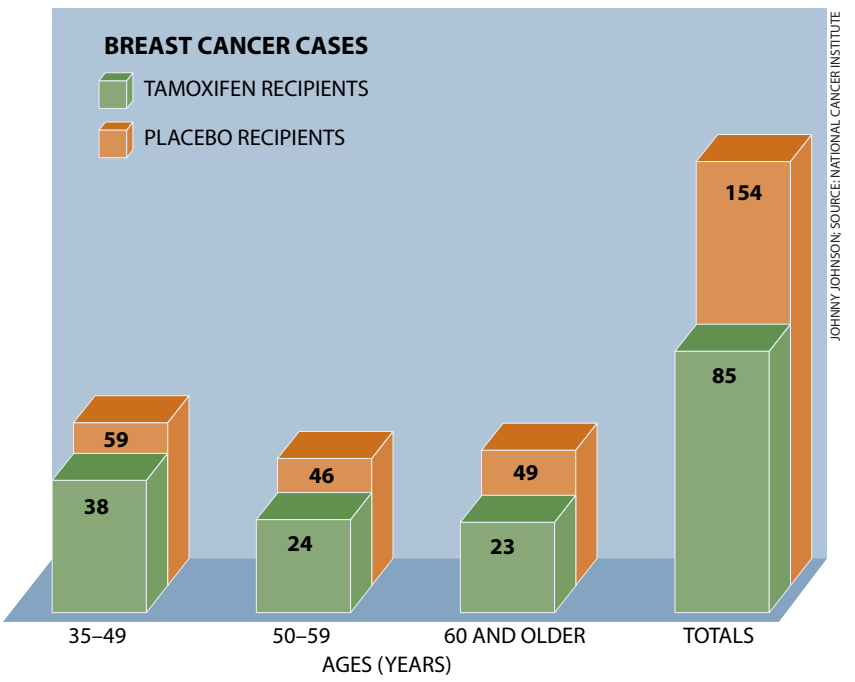
Clearly, their actions were target-site-specific. The chemicals always bound to the estrogen receptor in place of the real hormone, barring estrogen from attaching. And in some cell types, such as the breast, this binding ensured (as expected) that the receptors remained quiescent. At least in bone cells, though, the drugs induced the receptors to behave almost as if they were bound by estrogen.

These agents, then, were not pure antiestrogens after all. They were selective estrogen receptor modulators. Naturally, we were curious about the explanation for this tissue-selective activity—a variability that remains poorly understood [*see box on pages 62 and 63*]. At the time, though, we were mainly delighted that tamoxifen did not erode bone.

Our joy was short-lived, because an

TAMOXIFEN STUDY evaluated the drug's ability to prevent breast cancer in 13,388 women at high risk for the disease. Subjects, three of whom are shown below, were eligible if they were at least 60 years old or if their medical or family history indicated a high propensity for the disorder. The four-year study ended this past spring. Compared with subjects in the placebo-receiving group, 45 percent fewer of the treated patients were diagnosed with breast cancer (*graph*). The drug was effective in all age groups.

JEAN-MARC GIBOUX *Gamma Liaison*



JOHNNY JOHNSON; SOURCE: NATIONAL CANCER INSTITUTE

important side effect of tamoxifen soon came to the fore. We knew from the scientific literature that the agent could weakly enhance the growth of endometrial cancer in laboratory animals. Following up on this observation, Marco Gottardis in my group performed a now classic experiment demonstrating that the chemical behaved selectively like estrogen in the uterus. When he delivered tamoxifen and estrogen to mice that had been implanted with human breast and uterine tumors, he saw that tamoxifen impeded estrogen's ability to stimulate growth of breast cancer but was unable to check the growth of endometrial cancer. In 1989 Tommy Fornander and Lars Erik Rutqvist of the Karolinska Institute in Stockholm observed a similar phenomenon in women: in breast cancer patients who were postmenopausal, tamoxifen reduced the incidence of tumors in the second breast, but it increased the rate of endometrial cancer.

Since then, the public and physicians alike have expressed great anxiety over this effect. Vigilance is certainly warranted, but the extent of fear may be excessive. Recently the World Health Organization's International Agency for Research on Cancer concluded that women taking tamoxifen for breast cancer therapy should not stop because of concern about endometrial cancer: the survival benefits strongly outweigh the risks. In numerical terms, about 30 times more lives are saved by tamoxifen than are lost to its side effects. Further, with monitoring, most cases of endometrial cancer in tamoxifen users can be found early, at a stage when they are readily cured.

Another relatively uncommon but potentially lethal side effect has turned up as well. Like estrogen, tamoxifen can give rise to venous blood clots, mostly in women older than 50.

Tamoxifen Prevents Breast Cancer

As research into tamoxifen's side effects progressed, so did studies of its value for preventing breast cancer. One, a small endeavor headed by Trevor J. Powles of Royal Marsden Hospital in London, is ongoing. The other—the Breast Cancer Prevention Trial—produced the result mentioned at the start of this article.

The latter project, a joint American and Canadian effort led by Bernard Fisher, now at the Allegheny University of the Health Sciences, enrolled 13,388 women 35 years or older who were considered to be at high risk by virtue of their family or medical history or age (older than 60). Half received tamoxifen, half a placebo. Tamoxifen's benefit was so strong—the treated group had 45 percent fewer cases of breast cancer—that the study leaders halted the trial after four years, instead of the projected five, and offered tamoxifen to the placebo group.

Because physicians can prescribe an already marketed drug for any purpose they deem reasonable, doctors are starting to offer tamoxifen to certain patients whose risk profile matches the actual situation of women in the study. The subjects turned out to have at least four times the risk as that of their age groups in the general population.

Whether or not unease over tamoxifen's effect on the uterus is overblown, the fact that the drug can increase susceptibility to endometrial cancer means it cannot be used as a breast cancer preventive in women having no obvious susceptibility to that disorder (except, perhaps, in women who have already had their uterus removed for other reasons). Even the relatively low likelihood of acquiring endometrial cancer would be indefensible in that group. And yet some fraction of them will certainly go on to acquire breast cancer.

While pondering whether any other substance might be safer for broad distribution, I realized that a less risky compound suspected of being able to prevent breast cancer could be made available to the general population fairly quickly if it also had value for some other medical application. Experimenters could test the agent's efficacy in the other application without having to restrict subjects to those with a documented propensity for breast cancer and without having to wait many years for cancer to develop in the treated and placebo-receiving groups. Then, over time, researchers could gather data on its ability to prevent breast cancer.

By 1990 Gottardis had demonstrated that raloxifene—already known to preserve bone in rats—could prevent both mammary and endometrial cancer in rodents. Meanwhile clinical investigators had shown that tamoxifen, which seemed to behave very much like raloxifene, could lower LDL cholesterol in women. Together these findings implied that raloxifene or a similar SERM might provide the long-term benefits of estrogen replacement therapy (protection of the bones and heart) without posing the carcinogenic dangers.

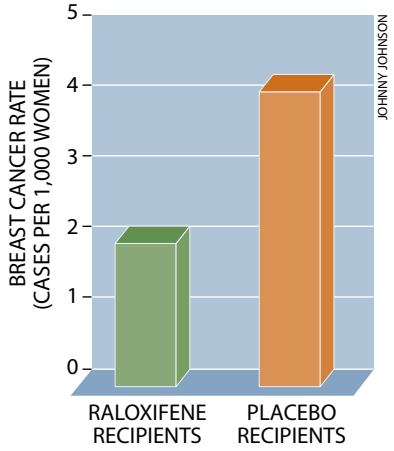
Focus on Health after Menopause

I therefore challenged pharmaceutical companies to develop drugs that aimed to prevent osteoporosis or heart attacks, or both, in postmenopausal women and that, as a beneficial side effect, prevented breast and endometrial cancer. I call this therapeutic approach my Trojan horse strategy, because the drug goes into the body ostensibly to do a single job—say, combating bone loss—but then does other tasks surreptitiously.

Many companies have responded. Such efforts led, less than a year ago, to the approval of raloxifene for maintaining bone density in postmenopausal

PHOTOGRAPHS BY P. MOTTA *University of La Sapienza, Rome/SPL/Photo Researchers, Inc.*

BONE from a healthy individual (*left*) is significantly sturdier than that from someone who has osteoporosis (*center*), the bone thinning that accelerates after menopause. One SERM, raloxifene, has been approved for maintaining bone density. Preliminary findings suggest that in postmenopausal women, raloxifene might also reduce risk for breast cancer (*graph*),



endometrial cancer and heart attacks, but further investigation is needed. The breast cancer data come from an analysis of more than 10,000 women followed for almost three years.

women. And signs from human trials favor the likelihood that it can meet the other criteria we put forward in 1990.

This past spring, for example, raloxifene was shown to lower "bad," LDL cholesterol in postmenopausal users without decreasing "good," HDL cholesterol. Perhaps, then, it will manage to lower the incidence of heart attacks. Moreover, preliminary data from studies that initially focused on osteoporosis imply that the drug can probably prevent breast cancer in that same population, reducing its incidence by about 50 percent. Also, small clinical studies have so far seen no disturbing estrogenlike activity in the uterus, but this effect must be evaluated more thoroughly.

Still, raloxifene is not perfect. Like estrogen replacement therapy and tamoxifen, it increases the incidence of blood clots in the veins. And many questions persist. Clinicians have yet to determine whether raloxifene remains effective and safe indefinitely and whether it is as successful as estrogen at preventing osteoporosis. Also uncertain is whether it might be antiestrogenic in areas of the brain involved in memory, although nothing indicates that it—or tamoxifen—impairs memory.

Further, even though animal and some human findings support the notion that raloxifene may reduce heart attacks and endometrial cancer in women who have stopped ovulating, doctors lack definitive data on these points. The heart question is currently being confronted in the RUTH (Raloxifene Use for the Heart) trial, which is testing raloxifene against a placebo in 10,000 women at high risk for coronary disease. That trial, which should begin reporting results in six or seven years, will additionally collect data on the incidence of breast cancer in the subjects.

Not surprisingly, beyond assessing raloxifene's value for women past menopause in general, scientists are eager to learn whether the drug is an effective but safer choice than tamoxifen for specifically preventing malignancy in females known to have a high likelihood of acquiring breast cancer. After all, raloxifene's effects in most tissues are similar to those of tamoxifen, except for raloxifene's apparent lack of unwanted stimulation of the uterine lining. A direct comparison of the drugs in postmenopausal women at risk for breast cancer should be under way by the time this article is published. The trial, named STAR (Study of Tamoxifen and Raloxifene), is sponsored by the National Cancer Institute. It will include 22,000 subjects and should last five to 10 years.

New SERMs are in the pipeline. One day women and their doctors may be able to select substitutes for estrogen replacement therapy and cancer preventives on the basis of each patient's unique set of risks and worries.

As studies of SERMs for those purposes proceed, efforts to improve SERM-based therapies for cancer are advancing as well. Breast cancer sometimes becomes resistant to tamoxifen and stops responding to it. Cancer researchers would like to explain how the resistance arises so they can find ways to overcome it. In the past year, my team at Northwestern University has uncovered one, albeit rare, mechanism of resistance: substitution of a tiny component—a single amino acid—in the estrogen receptor. That change causes the receptor to behave in breast cancer cells as if it were bound by stimulatory estrogen even when it is bound by the normally inhibitory tamoxifen [*see box on pages 62 and 63*].

In some cases, tamoxifen-resistant tumors respond to pure antiestrogens, chemicals that display no estrogenic activity at all. Those drugs are used as a last resort, however, because (as was originally feared for tamoxifen) they block estrogen at sites where it is wanted, such as the bones and liver. Just as multiple SERMs are being evaluated as substitutes for estrogen replacement therapy, pure antiestrogens are being developed as new cancer treatments.

A revolution in women's health has occurred in the past 20 years. Tamoxifen, the first known SERM, has extended millions of lives by acting as an antiestrogen in breast cancer cells. After it was introduced as an antiestrogen, recognition of its estrogenic aspects led to the study of other SERMs that are antiestrogenic in some tissues but estrogenic in others. Evaluations of those agents, and especially of raloxifene, then led to the prediction that certain SERMs might spare many women from osteoporosis, heart disease, breast cancer and endometrial cancer. Today that prediction shows strong signs of coming true. SA

The Author

V. CRAIG JORDAN, often said to be the driving force behind the development of tamoxifen, earned his doctorate in pharmacology from the University of Leeds in England. He is now director of the Lynn Sage Breast Cancer Research Program at the Robert H. Lurie Comprehensive Cancer Center of Northwestern University in Chicago and scientific chairman of the STAR trial, comparing the ability of tamoxifen and raloxifene to prevent breast cancer. He and his colleagues also proposed the idea that selective estrogen receptor modulation could protect women's health. Jordan has won many prizes for his work on tamoxifen—most notably the Cameron Prize, which has been identifying top medical scientists since 1879.

Further Reading

CONTRASTING ACTIONS OF TAMOXIFEN ON ENDOMETRIAL AND BREAST TUMOR GROWTH IN THE ATHYMIC MOUSE. M. M. Gottardis, S. P. Robinson, P. G. Satyaswaroop and V. Craig Jordan in *Cancer Research,* Vol. 48, No. 4, pages 812–815; February 15, 1988.

TAMOXIFEN: A GUIDE FOR CLINICIANS AND PATIENTS. V. Craig Jordan. PRR, Huntington, N.Y., 1996.

TAMOXIFEN: THE HERALD OF A NEW ERA OF PREVENTIVE THERAPEUTICS. V. Craig Jordan in *Journal of the National Cancer Institute,* Vol. 89, No. 11, pages 747–749; June 4, 1997.

BASIC GUIDE TO THE MECHANISMS OF ANTIESTROGEN ACTION. Jennifer I. MacGregor and V. Craig Jordan in *Pharmacological Reviews,* Vol. 50, No. 2, pages 151–196; June 1998.

Secrets of the Slime Hag

by Frederic H. Martini

Loathsome though they may seem, hagfishes may also resemble the earliest animals to have a braincase—making them even older than the first animals to develop a backbone

Thump. After an hour of descending through near-total darkness in the research submarine *Alvin,* we slide into the silty ocean bottom roughly 1,700 meters (just over one mile) below the surface of the Pacific, off the coast of southern California. The pilot switches on the floodlights, illuminating a dense cloud of sediment kicked up by our arrival. Several minutes pass as we wait for the debris to settle and activate the sub's sonar system, which shows a large target roughly 240 meters away. As we move closer, we see through *Alvin*'s portholes the ghostly white carcass of a 32-metric-ton gray whale. The whale's watery grave is anything but peaceful: it is swarming with hundreds of half-meter-long hagfishes, which are methodically gnawing away at the whale's chalky blubber, bite by bite.

Scenes like this are eerie enough to keep some people up at night—and they change forever one's concept of burial at sea. But for my colleagues and me, who study the biology of hagfishes, they provide a fascinating glimpse into the lives of these strange and slimy animals. For years, the habits of hagfishes—which are sometimes called slime hags—and their place on the evolutionary tree of life have been a matter of conjecture. But recent studies indicate that hagfishes, which appear to have changed little over the past 330 million years, in many ways resemble the first craniates (animals with a braincase). For instance, the evolutionary path leading toward humans—and all other vertebrates (animals with a backbone)—probably diverged from that of hagfishes 530 million years ago. New research also shows that hag-

DEAD WHALE on the deep-sea floor off the coast of California is aswarm with hagfishes. Scientists are finding that hagfishes—which have changed little over 330 million years of evolution—play important roles in the ecosystem of the ocean bottom, as both predators and scavengers.

fishes are much more abundant—and probably play a much more important role in the ecology of the ocean-bottom community—than anyone would have guessed a decade ago.

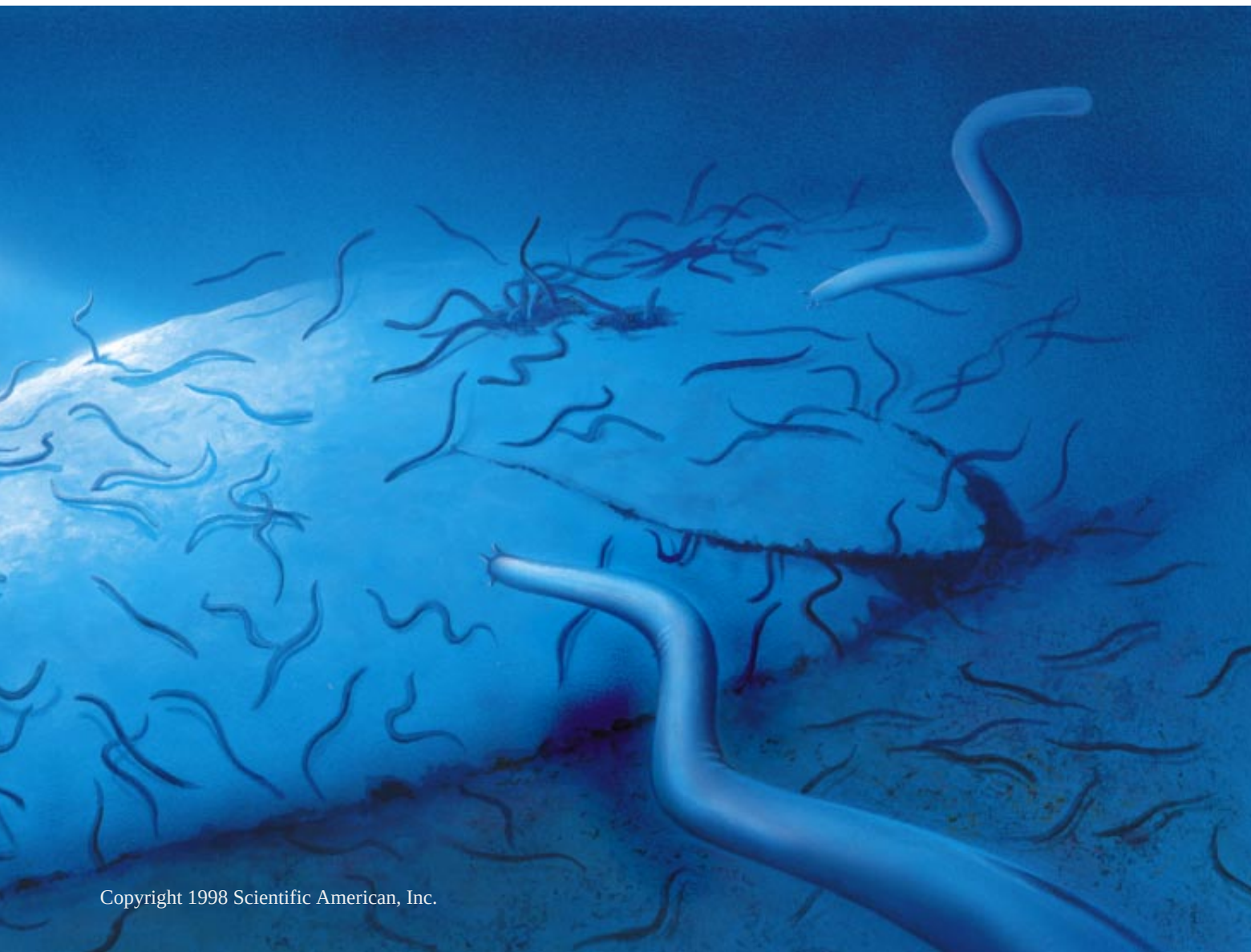
Slime Balls of the Sea

The word "slimy" can only begin to describe the average hagfish: one good-size adult can secrete enough slime from its roughly 200 slime glands to turn a seven-liter bucket of water into a gelatinous mess within minutes. Hagfishes release slime in varying amounts, depending on the circumstances. They tend to produce slime in small amounts while feeding on a carcass, a behavior that might be designed to ward off other scavengers. But when attacked or seized, a hagfish can ooze gobs of goo, from all its slime glands at once. Although the slime is initially secreted as a small quantity of viscous, white fluid, it expands several hundred times as it absorbs seawater to form a slime ball that can coat the gills of predatory fish and either suffocate them or distress them enough to make them swim away. But for all its utility, the slime appears to be equally distressing to the hagfish. To rid its body of the sticky mucus, a hagfish literally ties its tail in a knot and sweeps the knot toward its head to scrape itself clean.

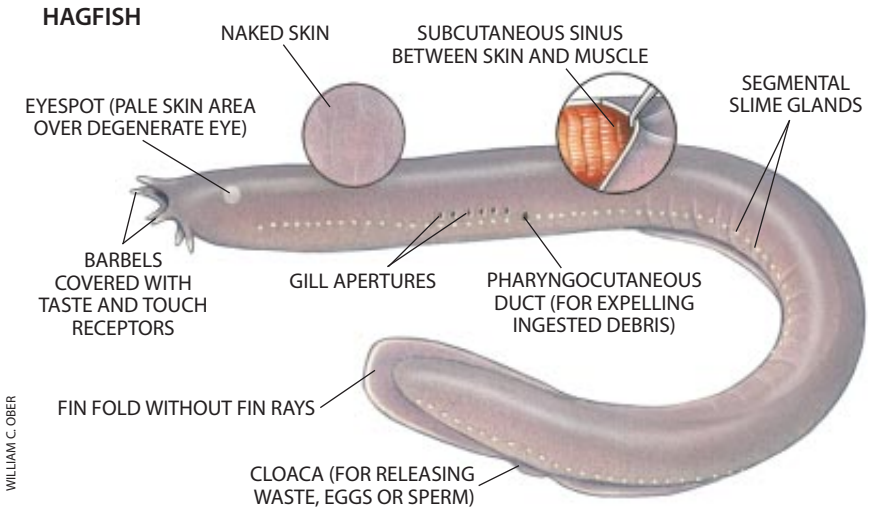
People often mistake the hagfish for an eel because both animals are long and cylindrical. The common names of several species of hagfish even include the term "eel," usually accompanied by a descriptive adjective ("slime eel," for example). As is so often the case, however, such common names are misleading. Hagfishes are not eels at all: true eels are bony fish with the requisite prominent eyes, paired pectoral and pelvic fins, a hard skeleton, bony scales and strong jaws. Like other bony fish, eels rely for respiration on gills that are attached to bones called gill arches and covered by a bony flap called an operculum.

In contrast, hagfishes are much simpler in form and function [*see illustrations on pages 72 and 73*]. They lack true eyes and paired fins, and their rudimentary skeleton consists only of a longitudinal stiffening rod made of cartilage, called the notochord, and several smaller cartilaginous elements, including a rudimentary braincase, or cranium. Hagfishes do not have scales; instead they have a thick, slippery skin and large, complex slime glands. In addition, they lack jaws, and their gills are a series of pouches that are different from the gills of any other living fish.

Hagfishes can be found in marine waters throughout the world, with the apparent exception of the Arctic and Antarctic seas. Although the animals always live near the ocean bottom, they can survive at a variety of depths. Water tempera-



ROB WOOD



EXTERNAL VIEWS of a Pacific hagfish (*left*) and an eel of the genus *Anguilla* (*right*) reveal that despite their similar shapes, the eel is a

ture is the primary factor that limits the habitat of hagfishes: they appear to prefer waters cooler than 22 degrees Celsius (71 degrees Fahrenheit). In the cold coastal waters off South Africa, Chile and New Zealand, the animals sometimes enter the intertidal zone, where they have been collected in tide pools as shallow as five meters. In tropical seas, though, hagfishes are seldom seen at depths shallower than 600 meters.

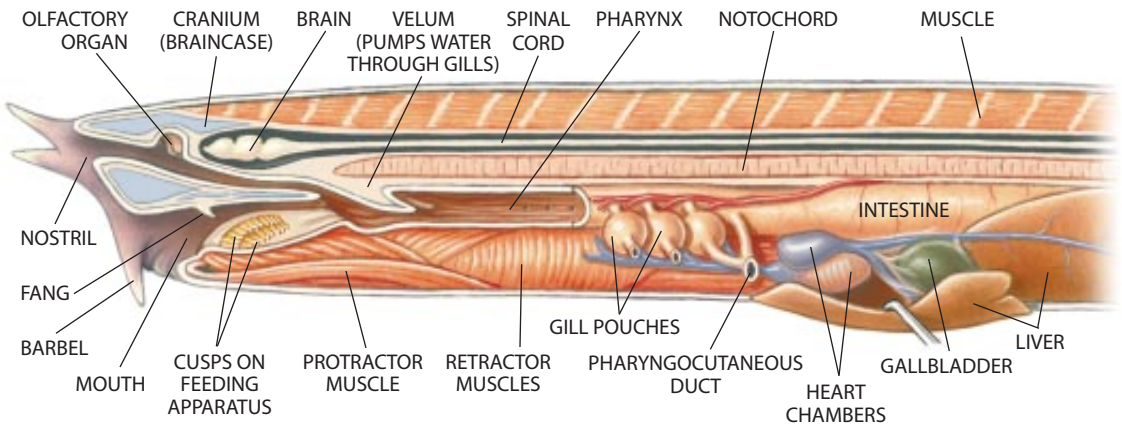
There are roughly 60 species of hagfishes, most of which are members of two major genera: *Eptatretus* or *Myxine*. (Many species in these genera, however, are known only from single specimens.) The genus *Eptatretus,* with roughly 37 species, includes the largest hagfish known, *E. carlhubbsi,* which can reach a length of 1.4 meters and can weigh several kilograms. Underneath their skin, *Eptatretus* species have the evolutionary remnants of eyes that are covered by translucent eyespots. Their heads also bear traces of lateral lines, sensory structures that extend down the sides of bony fish. Individual *Eptatretus* live in long-term burrows in the ocean floor but may roam widely among rocks or other hard substrates.

Members of the genus *Myxine,* which includes roughly 18 species, are more specialized than *Eptatretus* for living in burrows. *Myxine* are generally more slender, have even more degenerate eyes that lack eyespots, and show no traces of lateral lines. Typical *Myxine* live in transitory burrows and are always found in or near soft, muddy sediments.

The feeding habits of hagfishes—which can eat small, live prey and act as scavengers—are particularly distinctive. As a hagfish feeds, it protrudes a very effective feeding apparatus consisting of two dental plates, each supporting two curved rows of sharp, horny cusps. The dental plates are hinged along the midline, allowing them to open and close like a book. To take a bite, a hagfish extends its feeding apparatus, causing the "book" to open, and presses the dental plates against a fleshy surface—whether it be the body of a sea worm, a dead fish or your hand. When the hagfish withdraws the apparatus, the book closes and the opposing cusps grasp and tear the flesh, carrying it into the mouth. (The fang situated above the dental plates keeps live prey from wriggling away between bites.)

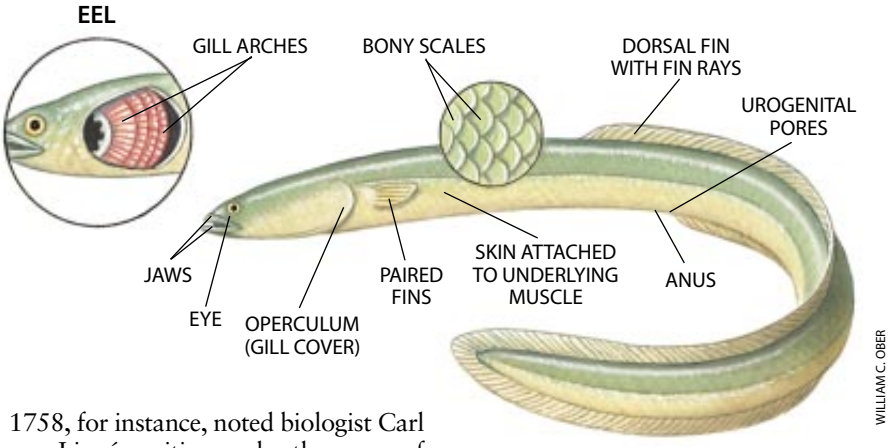
This feeding method works quite well when a hagfish preys on thin-skinned, soft-bodied sea worms, but the cusps cannot pierce the scales of fish or the skin of whales. Unless other scavengers have already opened the way, when feeding on a large carcass a hagfish usually takes the easy route, entering the body through the mouth, gills or anus. It then consumes the soft tissues from within, until only the bones and skin remain. More than one disappointed fisher has hauled in a prize fish that turned out to be a hollow shell full of hagfishes.

Only a few details are known about hagfish reproduction. Hagfish gonads form in a fold of tissue on the right side of the abdominal cavity. In a female an ovary forms in the anterior two thirds of the fold; in a male, a testis forms in the posterior one third. Curiously, individuals with both types of gonads are found occasionally. Females, which in some species outnumber males more than 100 to one, produce between 20 and 30 yolky, shelled eggs at a time. There are no oviducts; mature eggs are released into the abdominal cavity. The eggs—which vary in size from 20 to 70 millimeters, depending on the species—usually have hooked filaments at either end that enable them to lock together and be ejected in a chain. In males the testis produces sperm in follicles that release sperm into the abdominal cavity. Eggs or sperm then leave the abdominal cavity through a large pore into the



ANATOMICAL VIEWS of the anterior (*left*) and posterior (*right*) parts of a Pacific hagfish highlight both the animal's unique specializations and other, more general characteristics—such as a cranium—that persist in more evolved animals. (The middle—roughly one third of the animal's length—has been omitted.) Like a small proportion of most species of hagfishes, this specimen has both an ovary and a testis.

more highly evolved, bony fish. Unlike the eel, the hagfish lacks jaws, paired fins, eyes, scales, dorsal fin rays, gill arches and gill covers.



cloaca, an excretory chamber that also receives and expels urinary and digestive wastes.

Apart from these anatomical details, however, the sex lives of hagfishes remain almost a complete mystery. We assume that hagfish females lay their eggs for subsequent fertilization by males, but we have no idea where, when or how this occurs. We also do not know why the sex ratio of the hagfish is biased toward females or how often the females produce eggs.

The embryonic development of the hagfish is also still a black box. Despite more than 100 years of searching, only three fertilized eggs of the genus *Myxine* have ever been found, and those were damaged. The situation is only slightly better for other genera of hagfish: roughly 200 fertilized eggs of the genus *Eptatretus* were collected in California's Monterey Bay between 1896 and 1942, but none have been recovered since.

Many other aspects of hagfish life are equally mysterious. For instance, juvenile *Myxine*—those under 170 millimeters in length—have never been collected. Where are they, and what do they eat? How fast do they grow? At what age do they mature sexually? As yet, we have no answers.

Living Fossils

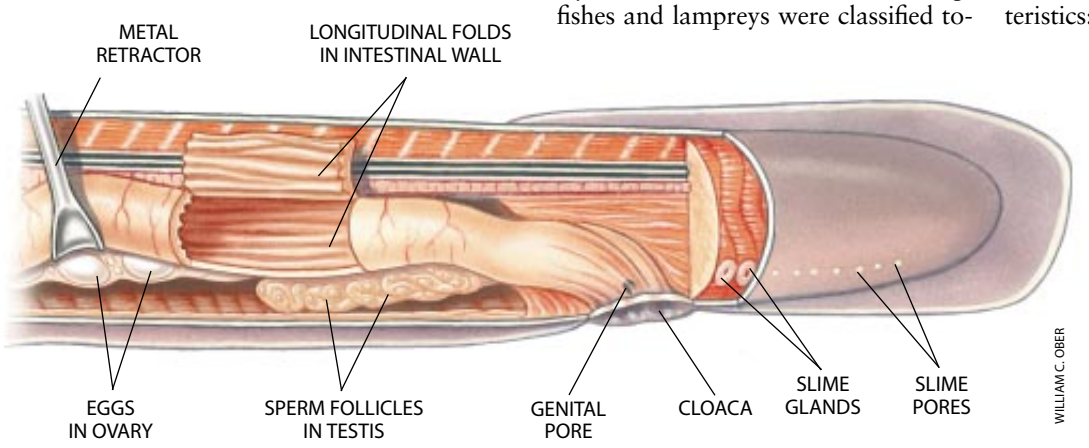
Given the bizarre and mysterious biology of hagfishes, it is no wonder that these blind, jawless, scaleless, finless, bottom-dwelling creatures were not immediately recognized (or acknowledged) as distant cousins of humans. In 1758, for instance, noted biologist Carl von Linné, writing under the name of Linnaeus, classified hagfishes as Vermes, or worms (rather than fishes), although we now know that hagfishes and worms are only distantly related.

Today, however, scientists recognize that hagfishes are virtual biological time machines. The term "living fossil" is often used to refer to the coelacanth, a rare, deepwater fish with fleshy, lobed fins that was first caught in 1938 in waters off the east coast of South Africa, among the Comoro Islands. But hagfishes make coelacanths look like evolutionary newborns: coelacanths may have changed little since they first appeared in the fossil record 60 million years ago, but a fossilized hagfish, *Myxinikela,* was found in sediments deposited roughly 330 million years ago. Aside from *Myxinikela*'s large eyes, if it were alive today it could easily pass for a modern hagfish.

Biologists neglected hagfishes for much of the past century primarily because of the way they classified animals. Until relatively recently, they relied on common features, such as the presence or absence of eyes or jaws, to establish relatedness between creatures. Under this scheme, hagfishes were lumped together with lampreys in a group called either Agnatha (literally, "no jaws") or Cyclostomata ("round mouths"). Hagfishes and lampreys were classified together because both lack jaws, paired fins, a bony skeleton and scales. Because the habitats of hagfishes make them relatively hard to come by, biologists concentrated on lampreys, which spend part of their lives in freshwater streams and rivers and therefore are much easier to catch.

In more recent years, the acceptance of phylogenetic systematics, or cladistics—classifying animals according to shared, specialized characteristics—has forced a reevaluation of the old methods for deciding what is related to what. Biologists now recognize that it is impossible to tell whether the ancestors of a given organism never had a particular feature or whether they had the feature but their descendants simply lost it sometime during evolution. Neither hagfishes nor snakes have legs, for example, but that does not mean that they are related. Hagfishes have never had paired fins—let alone limbs—but the ancestors of snakes had both forelimbs and hindlimbs.

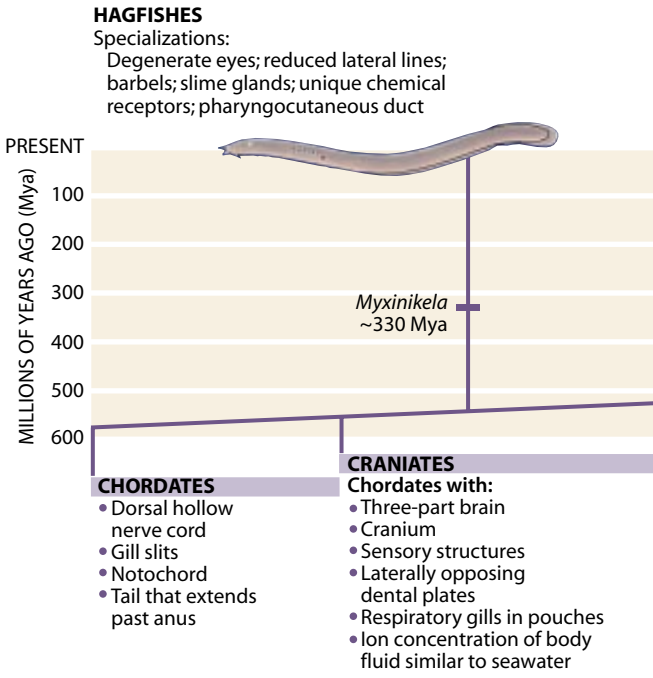
According to cladistics, hagfishes and lampreys are separate and distinct groups within the chordates (Chordata) [*see top illustration on pages 74 and 75*]. At some point in their lives, all chordates display the following characteristics: a hollow, dorsal nerve cord; a notochord, situated immediately below the nerve cord; gill slits; and a segmentally muscled tail that extends past the anus. Hagfishes are considered the most primitive living craniates. Lampreys also have a cranium, but unlike hagfishes, they also have segments of cartilage to protect their nerve cord. These cartilage segments are the first evolutionary rudiments of a backbone, or vertebral column. Lampreys, there-



fore, are considered the most primitive living vertebrates.

By comparing fossils with living creatures, biologists can create diagrams called cladograms that display the evolutionary relations among organisms. A cladogram of the chordates suggests that the hagfish diverged from the vertebrate evolutionary line around 530 million years ago. It also reveals that the predecessors of the hagfish never had a bony skeleton but that those of the lamprey did. What is more, the cladogram suggests that all early craniates had a complex protrusible feeding apparatus comparable to that of hagfishes. Early vertebrates, including the distant ancestors of humans, probably shared many other anatomical and physiological characteristics with modern hagfishes. But hagfishes have evolved many unique specializations: their eyes and lateral lines regressed, and they developed slime glands.

Besides their key position on the tree of life, hagfishes are also gaining new respect as members of the complex ecosystem of the ocean bottom. Scientists now know that the animals are more abundant than was once thought. Based on trapping surveys done between 1987 and 1992, my colleagues and I estimated that the inner Gulf of Maine contains population densities of up to 500,000 *M. glutinosa* per square kilometer. W. Waldo Wakefield of Rutgers University, who was then at the Scripps Institution of Oceanography, found comparable densities of *E. deani* off the California coast at depths of between 600 and 800 meters.



EVOLUTIONARY TREE, or cladogram, shows that hagfishes are the oldest surviving craniates (animals with a braincase)—predating lampreys, cartilaginous fishes and ray-finned fishes by millions of years. (Fossils are indicated in italics.)

We also now recognize the degree of hagfish predation on the populations of other animals that live near the bottom of the ocean. Although individual hagfishes have extremely low metabolic rates, their energy needs add up. The average number of *M. glutinosa* that inhabit one square kilometer of seafloor (59,700 animals) must consume the caloric equivalent of 18.25 metric tons of shrimp, 11.7 metric tons of sea worms or 9.9 metric tons of fish every year. And that amount would be sufficient only to keep the animals alive at rest; when they swim or burrow, their energy demands increase between four and five times.

Hagfishes also consume discarded so-called bycatch from commercial trawl-fishing fleets and play a central role in recycling the carcasses of dead marine vertebrates, including whales. Craig Smith of the University of Hawaii has found that hagfishes can remove roughly 90 percent of the energy content of small packages of bait sunk to the seafloor at depths of 1,200 meters. But hagfishes are not just important ecologically for their roles as predators or scavengers, they also serve as important prey for a surprising number of marine animals, including codfish, dogfish sharks, octopuses, cormorants, harbor porpoises, harbor seals, elephant seals and some species of dolphins.

The Slime Hag Trade

In many locales around the world, hagfishes have become the focus of a large and flourishing commercial fishery. Since the 1960s there has been a booming trade in leather goods produced from tanned hagfish skin. These products, which are manufactured primarily in South Korea, are sold as "eelskin" (presumably because consumers would be less likely to pay high prices

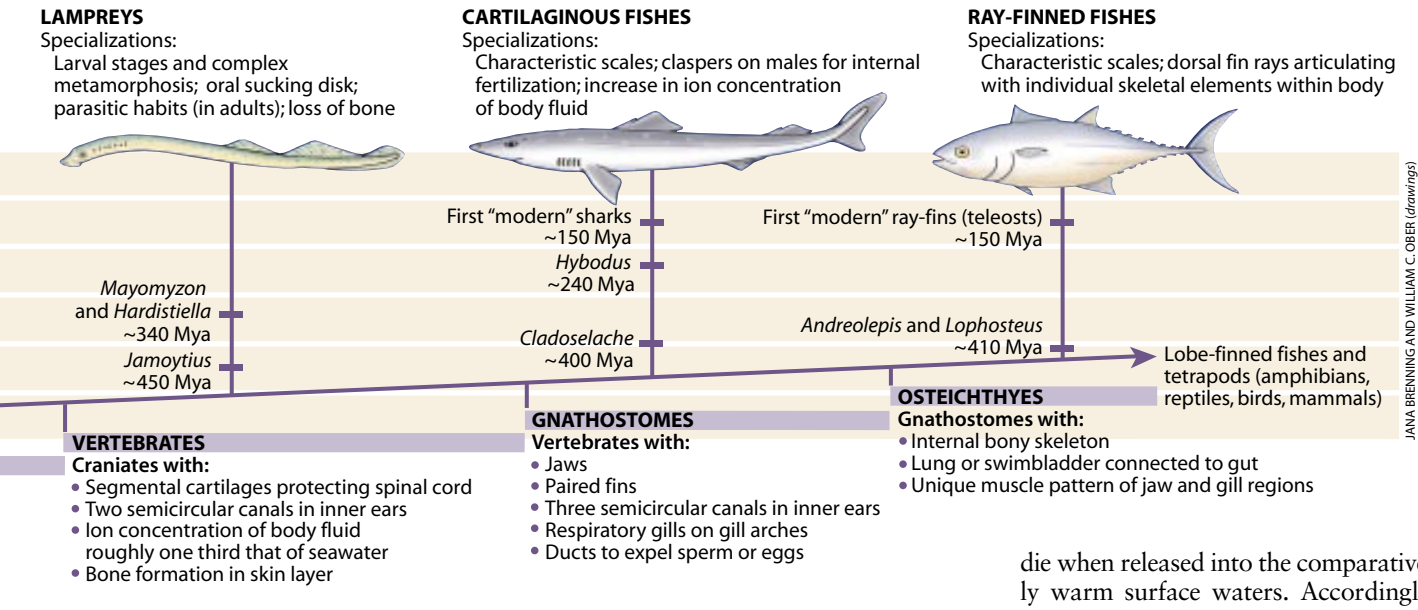
JASON GROW *SABA*

JASON GROW *SABA*

JASON GROW *SABA*

JOHN B. HEISER *Cornell University*

PACIFIC HAGFISHES are on display at the Steinhart Aquarium in San Francisco (*top* and *bottom left*); the dissected dental plate of a preserved hagfish (*bottom center*) reveals its toothlike cusps. A photograph from a fishing boat (*bottom right*) shows a characteristic ball of slime.



for goods labeled "slime-hag hide").

Hagfish skin, which is smooth and slick to the touch, consists of a superficial layer of epidermis overlying a dermis containing multiple, dense layers of collagen fibers. In the leather preparation process, the epidermis is removed, and the treated dermis is used to produce designer handbags, shoes, wallets, purses, briefcases and so forth. Removing the skin is relatively easy because it is attached to underlying muscles only along the dorsal midline and along the ventral surface at the level of the slime glands. Thus, the leather produced by one hagfish is a long strip, with a wrinkled band down the midline marking the site of the dorsal attachment to muscle.

The demand for suitable skins has supported commercial hagfish ventures around the Pacific Rim and in the western North Atlantic. The strip of skin must be above a minimum width (roughly five centimeters) but must not be too thick. This combination eliminates many species from the fishery: some are too small or too slender, others too large and too thick-skinned. The collection method is very low tech: multiple traps baited with anything from herring to kitchen scraps are set along a line on the sea bottom and left overnight. The traps can be 19-liter pails with lids or 190-liter barrels with small holes in the sides; once inside, most hagfishes become trapped in the bait and their own slime. In previously unfished areas, more than 100 hagfishes have been found to enter a given trap during its first hour on the bottom.

Unfortunately, the demand for hagfish skin has depleted the populations of many species because the trapping rate has far exceeded their rate of reproduction. And as each species becomes less abundant, fishers target others. Over the past three decades, fishers have exploited *Paramyxine atami, E. burgeri, E. okinoseanus* and *M. garmani* in the western North Pacific, *E. stouti* and *E. deani* off the Pacific coast of North America, and *M. glutinosa* in the Gulf of Maine.

In the region of New England, annual landings of hagfishes went from virtually zero in 1991 to roughly 1,950 metric tons in 1996. Over that five-year span, roughly 50 million hagfishes were processed and shipped overseas. Fishers discard hagfishes that are shorter than 500 millimeters—the minimum length suitable for leather—and these usually die when released into the comparatively warm surface waters. Accordingly, the actual impact of fishing on hagfish populations is far greater than indicated by the landings alone. By 1996 there were signs that hagfish fisheries were in trouble; recent declines in landings, average size and catch per trap suggest that the trouble is serious.

This state of affairs is not likely to improve, because almost everywhere hagfishes are classified as "underutilized species," and their exploitation is usually permitted without efforts to regulate the effect on hagfish populations. Hagfish trapping was considered a growth industry in New England, for example, when other fisheries were nearing a state of collapse.

Although hagfishes are more plentiful than once thought, we do not yet know enough about them to manage a sustainable hagfish fishery. In the meantime, we should take simple steps—such as requiring holes in commercial traps through which small, young hagfishes could escape—to reduce the impact on hagfish populations. When we drastically reduce the number of any species—even the lowly (and to some, loathsome) hagfish—we are performing an ecosystem experiment on a grand scale. As usual, we cannot yet begin to predict the eventual results. SA

The Author

FREDERIC H. MARTINI is an adjunct research associate in the School of Ocean and Earth Science and Technology of the University of Hawaii at Manoa but spends most of his time developing textbooks, teaching supplements and multimedia learning tools from his home in Maui. He is the author of five texts and reference books, including *Fundamentals of Anatomy and Physiology,* now in a fourth edition. Martini received a Ph.D. in comparative and functional anatomy from Cornell University in 1974.

Further Reading

THE BIOLOGY OF *MYXINE*. Edited by Alf Brodal and Ragnar Fänge. Universitetsforlaget, Oslo, 1963.
THE HAGFISH. David Jensen in *Scientific American,* Vol. 214, No. 2, pages 82–90; February 1966.
THE HAGFISHERY OF JAPAN. Aubrey Gorbman et al. in *Fisheries,* Vol. 15, No. 4, pages 12–18; July 1990.
THE BIOLOGY OF HAGFISHES. Edited by J. M. Jørgensen, J. P. Lomholt, R. E. Weber and H. Malte. Chapman & Hall, 1998.

The Asymmetry between Matter and Antimatter

In 1999 new accelerators will start searching for violations in a fundamental symmetry of nature, throwing open a window to physics beyond the known

by Helen R. Quinn and Michael S. Witherell

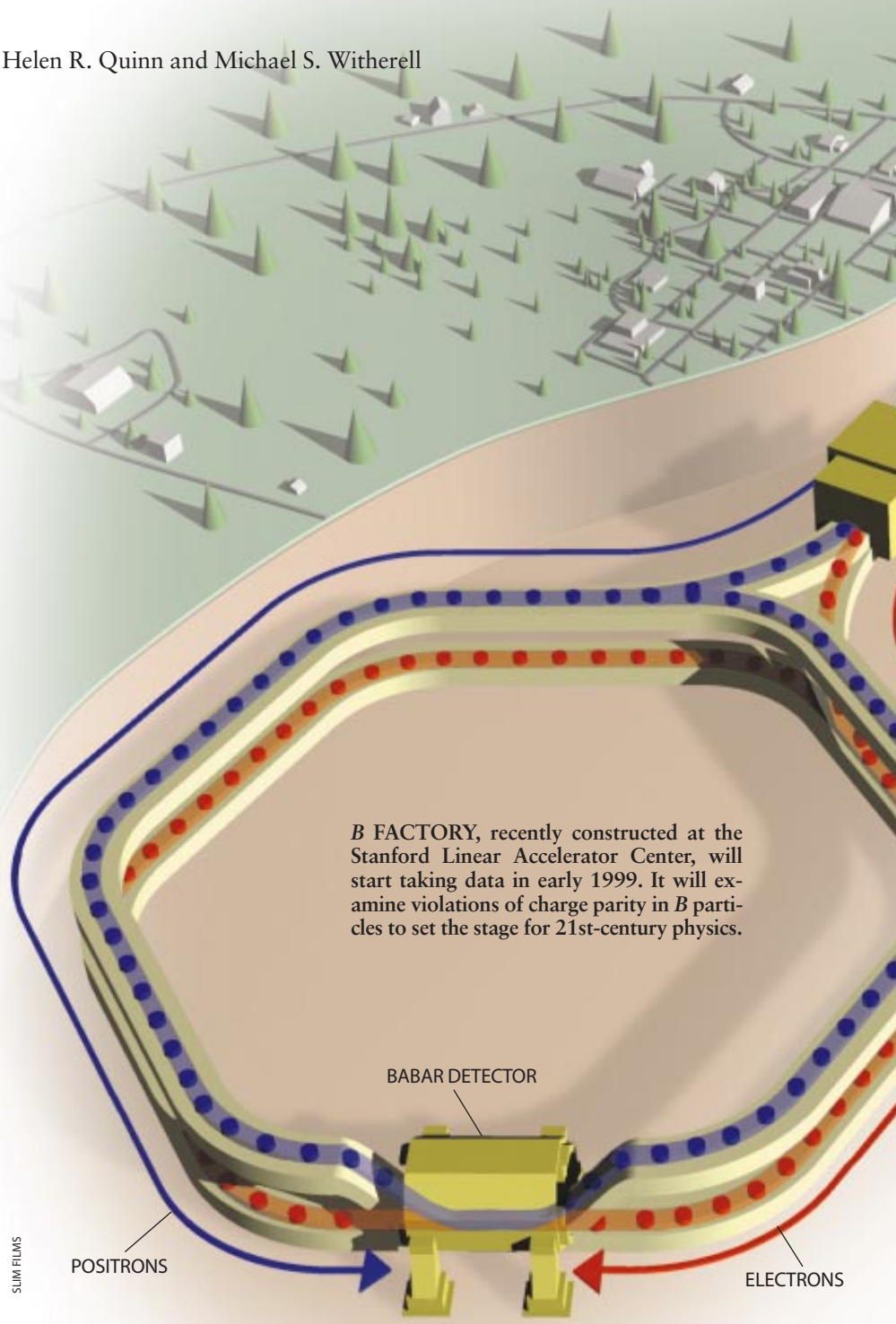
As far as humans can see into the universe, an essential imbalance strikes the eye. Stars, planets, asteroids, rocks—everything is made of matter. Essentially no antimatter is evident.

Is this imbalance the result of an accident, a chance occurrence during the birth of the universe? Or is it an inevitable outcome of some asymmetry in the laws of nature? Theorists believe that the excess of matter comes from fundamental disparities in how matter and antimatter behave. These differences amount to violations of a symmetry called charge-parity reversal, or CP.

After years of effort, experimental and theoretical physicists have found a natural way for CP symmetry to be broken within the prevailing theory of particle physics, called the Standard Model. Curiously, the amount of CP violation the model predicts is too small to explain the matter excess in the universe.

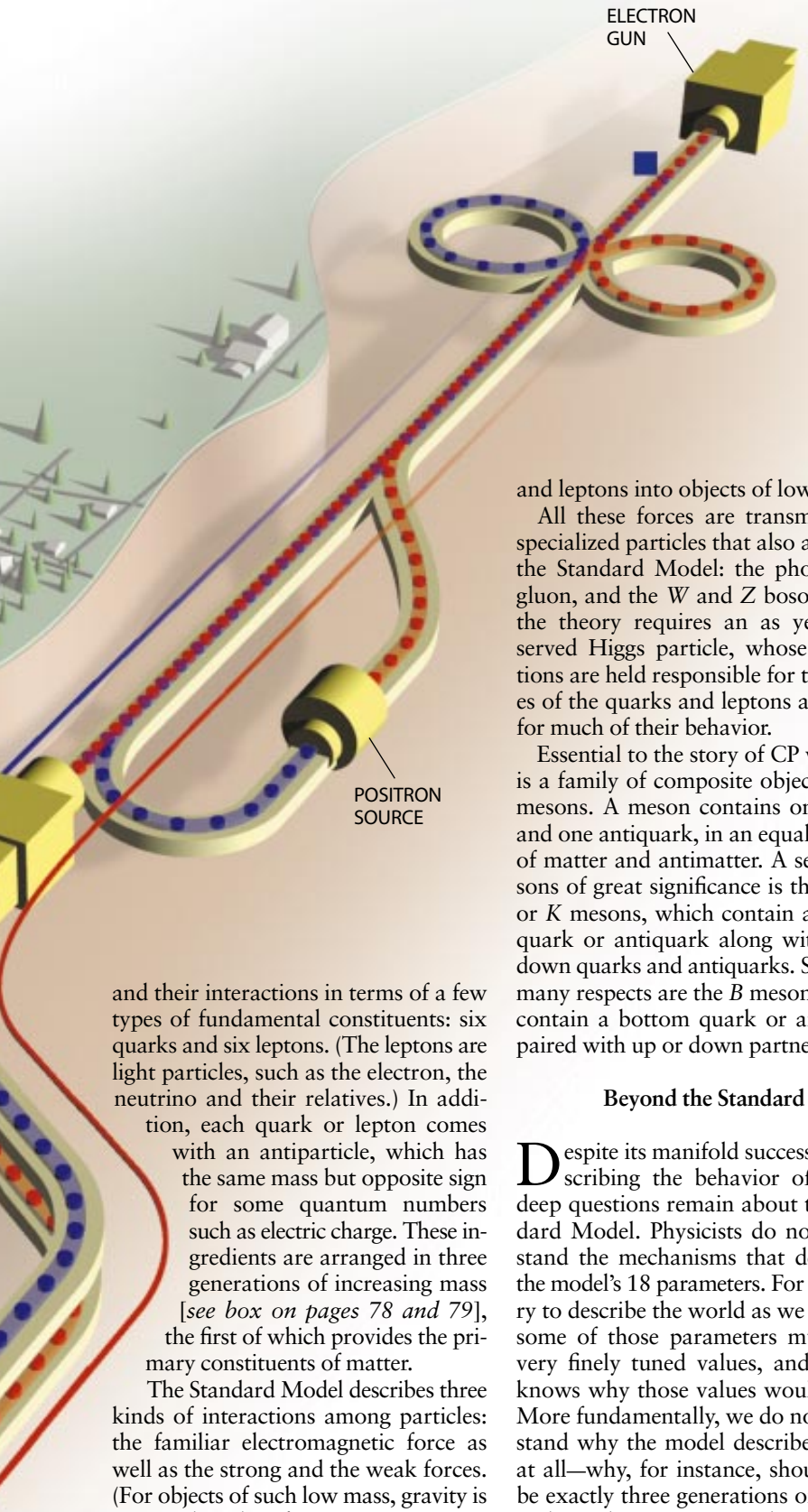
This finding is a vital clue that not all is well with the Standard Model: unknown factors are very likely at play. Two new accelerators, just now being completed in California and in Japan, will soon begin to probe violations of CP, with the aim of understanding whether the Standard Model needs to be revamped or replaced. These accelerators, which will produce enormous numbers of particles called *B* mesons, are known as asymmetric *B* factories. They are the latest tool in the search for physics beyond the Standard Model.

Everything known about the elementary properties of matter is encapsulated in the Standard Model. It describes all the hundreds of observed particles



SLIM FILMS

***B* FACTORY**, recently constructed at the Stanford Linear Accelerator Center, will start taking data in early 1999. It will examine violations of charge parity in *B* particles to set the stage for 21st-century physics.



and their interactions in terms of a few types of fundamental constituents: six quarks and six leptons. (The leptons are light particles, such as the electron, the neutrino and their relatives.) In addition, each quark or lepton comes with an antiparticle, which has the same mass but opposite sign for some quantum numbers such as electric charge. These ingredients are arranged in three generations of increasing mass [*see box on pages 78 and 79*], the first of which provides the primary constituents of matter.

The Standard Model describes three kinds of interactions among particles: the familiar electromagnetic force as well as the strong and the weak forces. (For objects of such low mass, gravity is too weak to be of interest.) Strong interactions confine quarks, which are never seen alone, within composite objects such as protons. Weak interactions cause instability—in particular, the slow decays of all the more massive quarks and leptons into objects of lower mass.

All these forces are transmitted by specialized particles that also appear in the Standard Model: the photon, the gluon, and the *W* and *Z* bosons. Last, the theory requires an as yet unobserved Higgs particle, whose interactions are held responsible for the masses of the quarks and leptons as well as for much of their behavior.

Essential to the story of CP violation is a family of composite objects called mesons. A meson contains one quark and one antiquark, in an equal mixture of matter and antimatter. A set of mesons of great significance is the kaons, or *K* mesons, which contain a strange quark or antiquark along with up or down quarks and antiquarks. Similar in many respects are the *B* mesons, which contain a bottom quark or antiquark paired with up or down partners.

Beyond the Standard

Despite its manifold successes in describing the behavior of matter, deep questions remain about the Standard Model. Physicists do not understand the mechanisms that determine the model's 18 parameters. For the theory to describe the world as we know it, some of those parameters must have very finely tuned values, and no one knows why those values would apply. More fundamentally, we do not understand why the model describes nature at all—why, for instance, should there be exactly three generations of leptons and quarks, no more or less? Finally, aspects of the theory that involve the Higgs particle are all untested. The Large Hadron Collider, now under construction at CERN, the European laboratory for particle physics near Geneva, will, however, allow the Higgs to be observed if its properties are as predicted by the Standard Model. The Higgs is believed to lie behind most of the mysteries of the Standard Model, including the violation of CP symmetry.

A theory of physics is said to have a symmetry if its laws apply equally well even after some operation, such as reflection, transforms parts of the physical system. An important example is the operation called parity reversal, denoted by P. This operation turns an object into its mirror reflection and rotates it 180 degrees about the axis perpendicular to the mirror [*see box on page 80*]. In mathematical terms, parity reverses the vectors associated with the object.

A theory has P symmetry if the laws of physics are the same in the parity-reversed world as in the real world. Particles such as leptons and quarks can be classified as right- or left-handed depending on the sense of their internal rotation, or spin, around their direction of motion. If P symmetry holds, right-handed particles behave exactly the same as left-handed ones.

The laws of electrodynamics and the strong interactions are the same in a parity-reflected universe. But in a famous experiment in 1957 Chien-Shiung Wu of Columbia University and her collaborators found that the weak interactions are very different for particles of different handedness. Peculiarly, only left-handed particles can decay by means of the weak interaction, not right-handed ones. Moreover, so far as we know there are no left-handed neutrinos: these particles are always right-handed. Because neutrinos have only weak interactions with the rest of the universe, this asymmetry is attributed to the weak force. So the weak force violates P.

Another basic symmetry of nature is charge conjugation, or C. This operation changes the quantum numbers of every particle into those of its antiparticle. Charge symmetry is also violated in weak interactions: antineutrinos are not left-handed, only right-handed.

Theorists combine C and P to get the operation CP, which turns all particles into their antiparticles and also reverses the direction of all vectors. When subjected to CP, the left-handed neutrino becomes a right-handed antineutrino. Not only does the right-handed antineutrino exist, but its interactions with other particles are the same as they are for left-handed neutrinos. So although charge and parity symmetry are individually

broken by neutrinos, in combination their dictates would seem to be obeyed.

Much to the surprise of physicists, the story of CP turned out to be far from simple. A mathematical theorem proved in 1917 by German mathematician Emmy Noether states that every symmetry implies the existence of a related quantity that is conserved, or immutable. For instance, the fact that spacetime is the same in all directions—that is, has rotational symmetry—leads to the conservation of angular momentum. Noether's theorem implies that if charge parity were an exact symmetry of nature, then a quantity called CP number would be conserved.

CP Violated

A particle and its antiparticle moving in opposite directions with equal energies form a pair with charge-parity symmetry: the CP operation does not change the system (taken as a whole), except that its mathematical representation acquires an overall factor. This factor is the CP number.

Either C or P, if acting twice on a system, returns it to the original state. This property is expressed as $C^2 = P^2 = 1$ (where 1, the identity operation, imparts no change at all). As a result, the CP number can be only +1 or –1. If nature has perfect charge-parity symmetry, Noether's theorem rules that no physical state with CP number –1 can transform into a state with CP number +1.

Consider the electrically neutral kaons. The K^0 consists of a down quark and an antistrange quark, whereas the anti-K^0 consists of an antidown quark and a strange quark. Because CP transposes quarks and antiquarks, it would turn each kaon into the other instead of leaving it unchanged. Hence, neither of these kaons has a definite CP number. Theorists can, however, construct a pair of kaons with definite CP numbers by superposing the wave functions for K^0 and anti-K^0. By the rules of quantum mechanics, these mixtures correspond to real particles and have definite mass and lifetime.

The conservation of CP number would explain an odd detail: the two "combination" kaons, though apparently similar, differ in their life spans by a factor of about 500 [*see bottom illustration on opposite page*]. The kaon with CP number +1 can change to two pions, a state that has the same CP number. This decay proceeds rapidly, because the kaon is massive enough to yield two pions readily. But the kaon with CP number –1 can decay only to another state with CP number –1: three pions. This latter breakdown takes time, because the kaon has barely enough mass to generate three pions. So when physicists found a long-lived kaon in addition to a short-lived one, they acquired strong evidence that the combination kaons obeyed CP symmetry.

This tidy picture was shattered in 1964, when in a groundbreaking experiment at Brookhaven National Laboratory on Long Island, James Christenson, James Cronin, Val Fitch and René Turlay observed that about one out of every 500 of the long-lived kaons (those with CP number –1) decays into two pions. If CP were an exact symmetry of nature, it would forbid such a decay. Few experiments in particle physics have produced a result as surprising as this one. Theorists found it hard to see why CP symmetry should be broken at all and even harder to understand why any imperfection should be so small.

PARTICLE			ANTIPARTICLE		
NAME	SYMBOL	CONTENT	NAME	SYMBOL	CONTENT
PROTON	p	u u d	ANTIPROTON	$\bar{p}$	$\bar{u}\,\bar{u}\,\bar{d}$
NEUTRON	n	u d d	ANTINEUTRON	$\bar{n}$	$\bar{u}\,\bar{d}\,\bar{d}$
PI-PLUS	π^+	u $\bar{d}$	PI-MINUS	π^-	$\bar{u}$ d
PI-ZERO	π^0	u $\bar{u}$ + d $\bar{d}$	ANTI-PI-ZERO	$\bar{\pi}^0$	u $\bar{u}$ + d $\bar{d}$
K-ZERO	K^0	d $\bar{s}$	ANTI-K-ZERO	$\bar{K}^0$	$\bar{d}$ s
B-ZERO	B^0	d b	ANTI-B-ZERO	$\bar{B}^0$	$\bar{d}$ b

SLIM FILMS

COMPOSITE PARTICLES are either baryons (such as the proton and the neutron) made up of three quarks, or mesons, which contain one quark and one antiquark. The most common meson is a pion, containing up and down quarks and antiquarks. *K* mesons and *B* mesons, important to the study of charge-parity violation, contain strange and bottom quarks (or antiquarks), respectively.

Particles of the Standard Model

The primary constituents of matter, quarks and leptons, are divided into generations. The first generation contains up and down quarks and antiquarks as well as the electron, a neutrino and their antiparticles. Ordinary matter is made almost exclusively of first-generation particles: an atom's nucleus contains protons and neutrons, themselves made of up and down quarks. The other generations occurred in the early universe, may still exist in hot environments such as neutron stars and are routinely observed in accelerators.

In addition, the Standard Model contains several particles that transmit force as well as a mysterious and unobserved particle called the Higgs. In the Standard Model the Higgs is responsible for the masses of all particles and for violations in charge-parity symmetry. *—H.R.Q. and M.S.W.*

TRANSMITTERS OF FORCE

WEAK BOSONS	PHOTON	GLUON	HIGGS
W	γ	g	H

In 1972 Makoto Kobayashi and Toshihide Maskawa of Nagoya University showed that charge parity could be violated within the Standard Model if three or more generations of quarks exist. As it happened, only two generations of quarks—the first, containing the up and down, and the second, with the strange and charm—were known at the time. So this explanation began to gain currency only when Martin L. Perl and others at the Stanford Linear Accelerator Center (SLAC) spied τ (tau) leptons, the first particles of the third generation, in 1975. Two years later experimenters at Fermi National Accelerator Laboratory in Batavia, Ill., found the bottom quark. But only recently, with the top quark being nailed down, also at Fermilab, has the third generation been completed.

Skewing the Universe

It is imaginable that the universe was born skewed—that is, having unequal numbers of particles and antiparticles to begin with. Such an initial imbalance, however, would be quickly eliminated

CONSTITUENTS OF MATTER

	PARTICLE	SYMBOL	CHARGE	MASS (GeV/c^2)
FIRST GENERATION				
QUARKS	UP	u	+2/3	0.03
	DOWN	d	−1/3	0.06
LEPTONS	ELECTRON	e^-	−1	0.0005
	ELECTRON NEUTRINO	ν_e	0	0?
SECOND GENERATION				
QUARKS	CHARM	c	+2/3	1.3
	STRANGE	s	−1/3	0.14
LEPTONS	MUON	μ^-	−1	0.106
	MUON NEUTRINO	ν_μ	0	0?
THIRD GENERATION				
QUARKS	TOP	t	+2/3	174
	BOTTOM	b	−1/3	4.3
LEPTONS	TAU	τ^-	−1	1.7
	TAU NEUTRINO	ν_τ	0	0?

SLIM FILMS

if the early universe contained any processes that could change baryon number—the number of matter particles minus the number of antimatter particles. (In extensions of the Standard Model called Grand Unified Theories, such processes would have been very common soon after the big bang.) Theorists prefer the alternative scenario, in which particles and antiparticles were equally numerous in the early universe, but the former came to dominate as the universe expanded and cooled.

Soviet physicist (and dissident) Andrei Sakharov pointed out three conditions necessary for this asymmetry to develop. First, fundamental processes that do not conserve baryon number must exist. Second, during the expansion the universe must not attain thermal equilibrium. (When in thermal equilibrium, all states of equal energy contain equal populations of particles, and because particles and antiparticles have equal mass or energy, they would be generated at the same rate.) Third, CP symmetry—essentially, the symmetry between matter and antimatter—must be violated. Otherwise any process that changes the amount of matter would be balanced by a similar effect for antimatter.

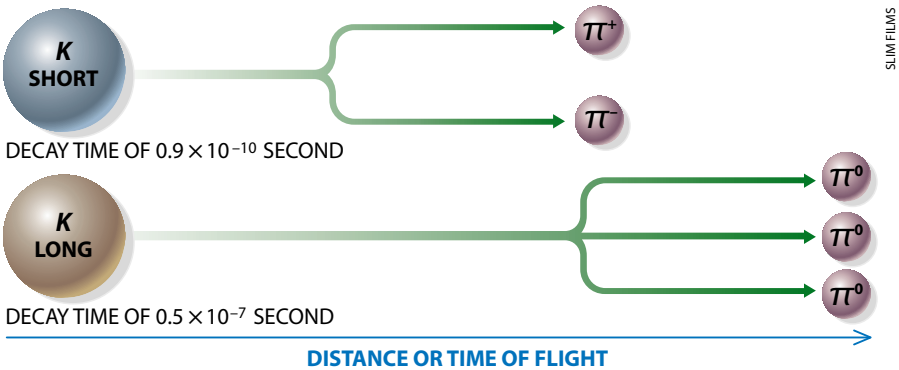
The prevailing theory holds that when the universe was born, the quantum field associated with the Higgs particle was everywhere zero. Then, somewhere in the universe, a bubble developed, inside which the Higgs field assumed its present nonzero value. Outside the bubble, particles and antiparticles had no mass; once inside, however, they interacted with the Higgs field to acquire mass. But as the bubble grew, particles and antiparticles were swept through its surface at unequal rates because of CP violation. Any imbalances between matter and antimatter thus created outside the bubble were quickly corrected by processes that change baryon number.

Such processes were extremely rare inside the bubble, however, so the imbalance was frozen in. By the time the bubble had expanded to occupy the entire universe, it contained more particles than antiparticles. Eventually the universe cooled to a point where particles and antiparticles could no longer be generated in collisions but would annihilate when they found one another.

Unfortunately, when theorists calculate how much of an imbalance between matter and antimatter this mechanism can create, it comes out too small—by many orders of magnitude. This failure suggests that there must be other ways in which CP symmetry breaks down and hence that the Standard Model may be incomplete.

A fruitful place to search for more violations is most likely among the B mesons. The Standard Model predicts the various decays of the B^0 and the anti-B^0 to be highly asymmetric. A B^0 contains a down quark bound to an antibottom quark, whereas the anti-B^0 consists of an antidown quark and a bottom quark. The B mesons behave much like the kaons discussed earlier: the observed B mesons consist of mixtures of the B^0 and anti-B^0.

Consider the evolution of a B^0 meson produced at a certain instant. Some time



SLIM FILMS

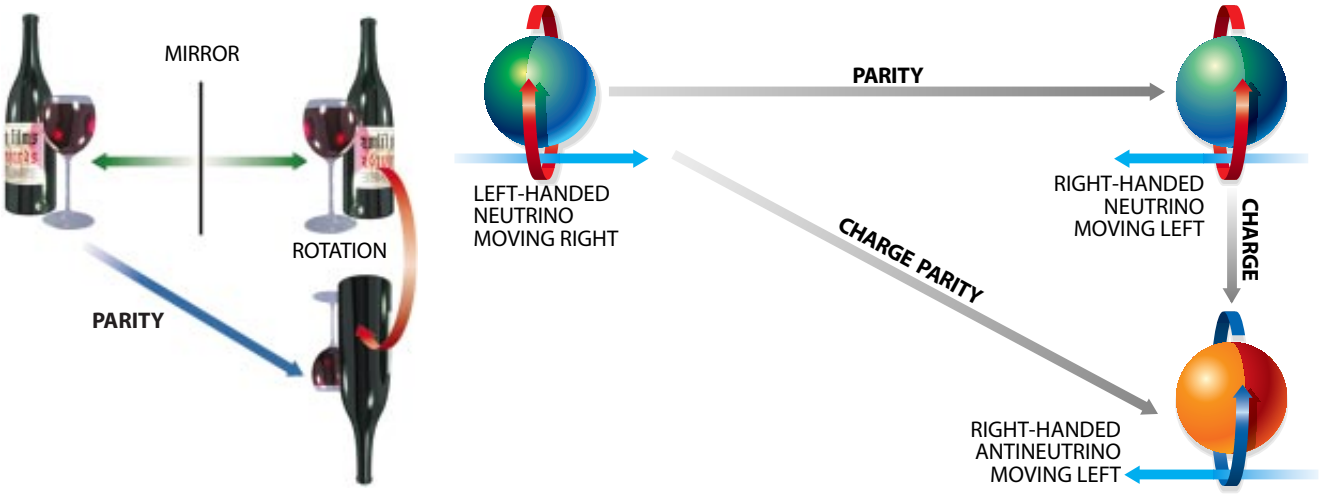
NEUTRAL KAONS, or K mesons, are observed to have two very different life spans. One type of kaon decays quickly into two pions, whereas the other decays slowly into three pions. The different behavior comes from the two kaons having opposite charge-parity symmetry. On rare occasions, however, the second type of kaon also decays to two pions, proving that charge parity can be violated.

Reversal of Charge and Parity

Symmetries are vital to the study of physics, and few symmetries are more intriguing than the combination of charge and parity. Charge reversal gives the opposite sign to quantum numbers such as electric charge, changing a particle to its antiparticle. Parity reversal reflects an object and also rotates it by 180 degrees (equivalent to changing the arrow on all vectors associated with the object).

The laws of classical mechanics and electromagnetism are invariant under either of these operations, as are the strong interactions of the Standard Model. The weak interactions, however, are changed by the reversal of either charge or parity.

For many years, it appeared that parity and charge flipped in succession ("charge parity") was invariant even for weak interactions. Experiments in 1964 shattered this illusion, posing the puzzle of why nature looks different when reflected in the charge-parity mirror. *—H.R.Q. and M.S.W.*



SLIM FILMS

later an observer has a certain probability of finding the same particle and also some probability of finding its antiparticle, the anti-B^0. This peculiar meson state, oscillating between a given quark-antiquark combination and its antiparticle, is a remarkable illustration of quantum mechanics at work.

The Bottom Line

To study CP violation, experimenters need to study decays of B^0 into those final states that have a definite CP number. Such decays should proceed at a different rate for a particle that is initially B^0 compared with one that is initially anti-B^0. This difference will indicate the extent of CP violation in the system. But rather than resulting in the one-in-1,000 effect seen in K^0 decays, the predicted asymmetry for B^0 decays grows so large that one decay rate can become several times larger than the other.

Models other than the Standard often have additional sources of CP violation—sometimes involving extra Higgs particles—in general offering any value for imbalances in B^0 decays. Thus, measuring the pattern of asymmetries will provide a clear test of the predictions.

When the bottom quark was discovered, its mass was measured at around five giga-electron volts (GeV), or about five times the mass of a proton. Consequently, theorists calculated that it would take a little more than 10 GeV of energy to produce two B mesons (because the added down or antidown quarks are very light). In the early 1980s at Cornell University, operators of an electron-positron collider—a machine that accelerates electrons and positrons into head-on crashes—tuned it so that an electron-positron pair would release an energy of 10.58 GeV on annihilating. As predicted, this burst of energy preferentially converts to B mesons, providing a very rich source of the particles. About one in four annihilations results in a B meson and its antiparticle, leaving behind no other particles at all.

At SLAC in 1983 experimenters found an unexpectedly long lifetime of about 1.5 picoseconds for the B meson. The extended life improved the chances that a B^0 would turn into an anti-B^0 before decaying, making CP-violating asymmetries easier to observe. Furthermore, in 1987 experimenters at the Electron Synchrotron Laboratory (DESY) in Hamburg, Germany, measured this "mixing" probability at 16 percent, making it likely that the asymmetries would be far larger than those for the K^0. Still, these large asymmetries occur in relatively rare decays of the B mesons. For a true study of CP violation, a great number of B mesons would be needed.

In 1988 at a workshop in Snowmass, Colo., the major topic of interest was the Higgs particle. A group of participants also discussed CP violation, especially in B mesons. It determined that a favorable way to study the B mesons would be with an electron-positron collider tuned to 10.58 GeV in which the electron and positron beams had different energies. This rather unusual feature would facilitate the measurement of a B meson's life span. Experimenters identify the point of birth and the point of death (that is, decay) of a B meson from traces of particles in the detector. Dividing the distance between these two points by the calculated velocity of the meson yields its life span. But an ordinary electron-positron collider at 10.58 GeV produces two B mesons that are almost at rest; the small distances they move are hard to measure.

Pier Oddone of Lawrence Berkeley National Laboratory had pointed out that if the electrons and positrons have different energies, the B^0 mesons that are produced move faster. For instance,

if the electron beam has an energy of 9.0 GeV and the positron beam an energy of 3.1 GeV, the B^0 mesons move at half the speed of light, traveling about 250 microns (about one hundredth of an inch) before they decay. Such a distance can yield a reasonably accurate measure of the lifetime.

An accelerator facility with two separate rings delivering different energies to the electrons and positrons would fit the task. Each ring would have to deliver very intense beams of particles, obtaining a high rate of collisions. Such a machine came to be called an asymmetric *B* factory: asymmetric because of the different beam energies, and *B* factory because of the large numbers of *B* mesons it would produce.

Teams at several laboratories developed designs that could generate about 30 million pairs of *B* mesons a year. In 1993 the U.S. Department of Energy and the Japanese agency Monbusho approved two proposals for construction: one at SLAC in California and the other at KEK, the High Energy Accelerator Research Organization in Tsukuba, Japan. The SLAC project is utilizing the existing linear tunnel to accelerate the positrons and electrons. These will then be circulated in separate rings newly constructed in a 20-year-old tunnel and set to collide at a point of crossing. The accelerator construction cost $177 million. The Japanese project is also employing extant tunnels—those that previously housed the Tristan collider.

Physicists and engineers are busy setting up a large experiment that can identify the rare decays of a *B* meson and measure their positions to within the requisite 80 microns. This accuracy is obtained by using the silicon microstrip technology that helped to unearth the top quark [see "The Discovery of the Top Quark," by Tony M. Liss and Paul L. Tipton; SCIENTIFIC AMERICAN, September 1997]. Experimenters aim to identify almost every particle that emerges from the decays of the *B* mesons in order to isolate the rare events that shed light on charge-parity questions.

In the BABAR detector that is being built for SLAC, the silicon microstrip will be the innermost layer, forming a cylinder roughly 30 centimeters in diameter and 60 centimeters long. Outer layers will measure energy, velocity and penetration power for each particle created, allowing physicists to reconstruct the original events. More than 500 participants—including both of us—from 70 institutions in nine nations are building the detector and also sharing its cost of $85 million. (It was, in fact, to facilitate international collaborations of this kind that the World Wide Web was invented at CERN.) The BELLE collaboration that is building the Japanese experiment is also international in scope, with members from 10 countries. Both *B* factories are scheduled for completion later this year, with the first data arriving in early 1999.

STANFORD LINEAR ACCELERATOR CENTER

BABAR DETECTOR at Stanford University, being built by a collaboration from nine nations, will capture the decays of *B* mesons in charge-parity studies.

Other kinds of violations of charge parity, less predictable than the quantum-mechanical mixing, should also occur in *B* decays. The Cornell collider and detector are being upgraded to search for such effects. A number of experiments on *B* physics are also planned at proton accelerators around the world. Both types of colliders will provide crucial, and complementary, pieces of evidence on CP violation.

The *B* factories could definitively tell researchers that the Standard Model concept works and then help to determine its remaining parameters. Alternatively, they could show that the model's predictions cannot fit the data no matter what the choice of parameters. Indeed, the results could rule out entire classes of models beyond the Standard Model, thus helping theorists to zero in on a successor. And if all goes well, we may even come to understand why our world is made exclusively of matter. SA

The Authors

HELEN R. QUINN and MICHAEL S. WITHERELL bring complementary skills to the study of CP violation. Quinn is a theorist, whose main contributions include a demonstration of how the strong, weak and electromagnetic forces can become unified at high energies, as well as an explanation of why strong interactions conserve CP symmetry. She is currently involved in designing tests of the Standard Model to be carried out at the *B* factories and has been a staff scientist at the Stanford Linear Accelerator Center since 1976. Witherell is an experimenter who led a Fermilab effort to study the decays of charm mesons, for which he obtained an award from the American Physical Society in 1990. For the past four years, he has been helping build the silicon vertex detector for the BABAR experiment. He is a professor at the University of California, Santa Barbara, and was elected to the National Academy of Sciences this year.

Further Reading

THE CHARACTER OF PHYSICAL LAW. Richard Feynman. MIT Press, 1965.

FEARFUL SYMMETRY: THE SEARCH FOR BEAUTY IN MODERN PHYSICS. A. Zee. Macmillan Publishing, 1986.

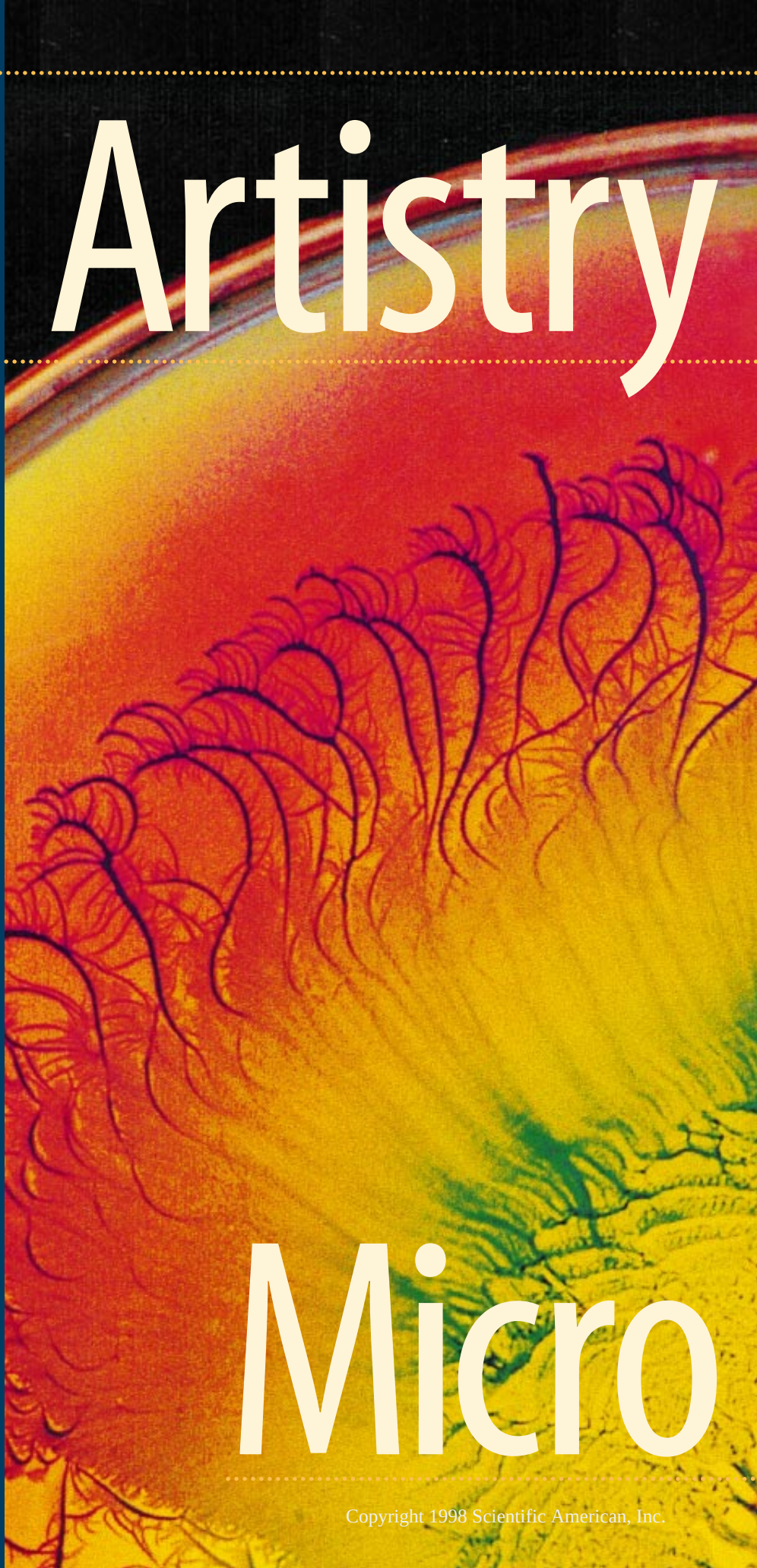
THE PHYSICS OF TIME REVERSAL. Robert G. Sachs. University of Chicago Press, 1987.

THE NEW AMBIDEXTROUS UNIVERSE: SYMMETRY AND ASYMMETRY FROM MIRROR REFLECTIONS TO SUPERSTRINGS. Martin Gardner. W. H. Freeman and Company, 1990.

The Artistry of Microbes

Colonies of bacteria or amoebas form complex patterns that blur the boundary between life and nonlife

by Eshel Ben-Jacob and Herbert Levine



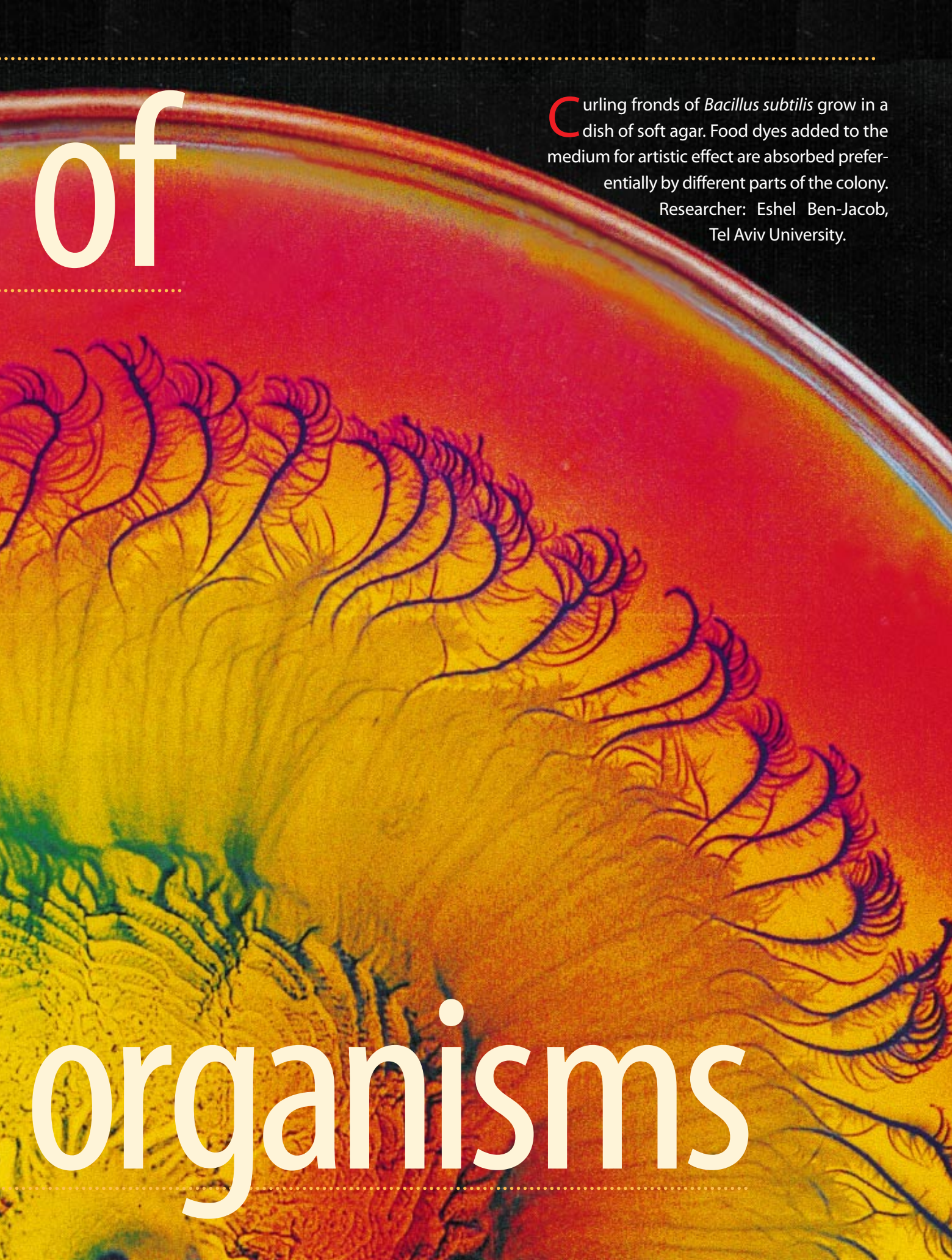
In the past few decades physicists have come to understand how some patterns, such as the shape of a snowflake or of a flame, form. Think of how a snow "fern" grows on a cold window. Randomly migrating molecules of water on the glass pane occasionally collide with the fern, getting stuck. Because the molecules are more likely to hit the protruding parts of the fern and to attach there, any region that sticks out tends to grow longer. This process leads to a series of branches. In addition, the molecules are more likely to stick in certain positions depending on the orientation of ice molecules already in the fern. As a result, an intricate, crystalline structure spontaneously grows.

Colonies of microorganisms can take up even more complex forms because of an extra ingredient—life. Physicists were surprised to find that their pat-

Continued on page 87

COURTESY OF ESHEL BEN-JACOB

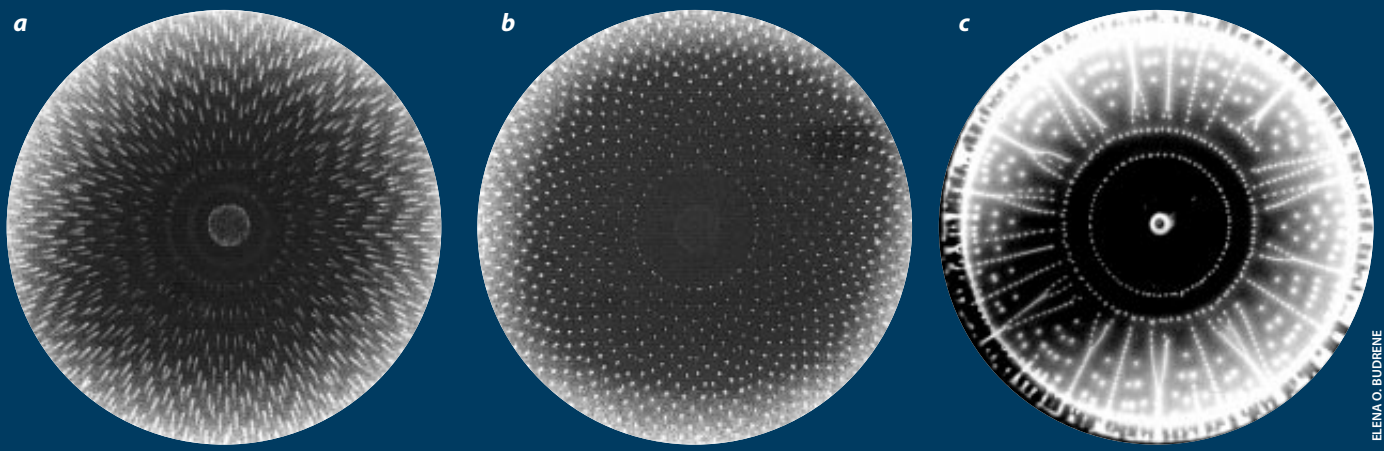
# of organisms



Curling fronds of *Bacillus subtilis* grow in a dish of soft agar. Food dyes added to the medium for artistic effect are absorbed preferentially by different parts of the colony. Researcher: Eshel Ben-Jacob, Tel Aviv University.

COURTESY OF ESHEL BEN-JACOB

Branches give way to curls when *B. subtilis* from a colony in hard agar are placed instead in soft agar. The organisms mutate to a longer form that moves fast but tends to veer toward one side, giving rise to the curls. The asymmetry is most likely caused by the spiral tail that propels each bacterium. In extremely hard agar, however, *B. subtilis* congregate into tight, rotating vortices that cut through the medium like a circular saw. The cooperative behavior allows the colony to grow outward (*inset*); the "saw" is the dark dot at the end of each line. Researcher: Eshel Ben-Jacob.



ELENA O. BUDRENE

Geometric lattices may result when commonplace *Escherichia coli* (*a, b*) or *Salmonella typhimurium* (*c, d*) organize themselves. When exposed to certain nutrients or agents that cause stress, the cells secrete chemical attractants and aggregate in response to them. Depending on conditions, the groups can arrange themselves into stable periodic arrays of spots or stripes. Researchers: Elena O. Budrene and Howard C. Berg, Harvard University.

Patterns also emerge in computer simulations (*right*) in which the bacteria migrate toward higher densities of the signaling chemical. Researcher: Eshel Ben-Jacob.

ESHEL BEN-JACOB

JAMES A. SHAPIRO

Concentric circles form within a colony of *E. coli* grown on hard agar. Under these conditions the bacteria cannot move but propagate outward only by cell division. The rings show periodic chemical activity: the blue color represents the metabolized products of an enzyme. Researcher: James A. Shapiro, University of Chicago.

FELICE FRENKEL *Massachusetts Institute of Technology*

Terraces result when some bacteria in colonies of *Proteus mirabilis* temporarily transform into long "swarmer" cells that move in concert over hard agar to colonize new regions. These pioneers establish a habitable swath that is filled in by normal "swimming" cells. Then the process begins again, perhaps triggered by an internal clock. *P. mirabilis* is responsible for many urinary tract infections suffered by hospital patients. Researcher: James A. Shapiro, University of Chicago.

Continued from page 82

terns often resemble those of inanimate objects. But the bacteria clearly choose certain designs over others to facilitate their survival.

Take, for example, *Bacillus subtilis* bacteria. In the late 1980s Mitsugu Matsushita of Chou University in Tokyo, along with his collaborators, showed that these microorganisms could create branching patterns when cultured on agar surfaces with limited nutrients. They grow by much the same principles as a snow fern does. Imagine the extreme case where the agar is too hard for the bacteria to move. Molecules of food can, however, bounce randomly through the gel, tending to reach those bacteria that are protruding from the colony. The bacteria then eat, grow twice as long and split into two. In this way, the protrusion slowly grows into a branch. Usually, of course, the bacteria also move. They push toward regions where more food is available—that is, just beyond the branch tips—thus causing the branches to grow even faster.

Inspired by this finding, one of us (Ben-Jacob) decided to seek the point at which the pattern would betray the living nature of its constituents. He found that if the bacteria are grown on softer agar, the initially branched pattern changes spontaneously to a curling one that spreads much faster. Under a microscope, each bacterium is seen to become significantly longer when the pattern begins to curl. Although the transformation obviously helps to make the bacteria more mobile—and thus to improve their access to food—no one knows how it is triggered.

Bacteria not only sense their environment but also influence it, as well as one another: colonies of *Escherichia coli* can aggregate by secreting attractant chemicals. Amoebas, which have nuclei and other structures that make them more complex than bacteria, are more robust. Whereas bacteria often respond to stresses by generating new variations of themselves—and thereby new patterns—amoebas are quite resistant to change. But they show collective behavior of astonishing sophistication.

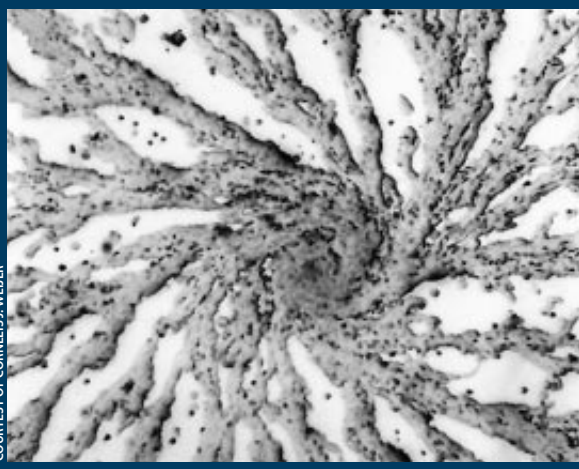
One such amoeba, *Dictyostelium discoideum*, is a favorite of developmental biologists who wish to study how cells can join together to form multicellular organisms. If these cells are spread on a surface that has no food at all, after about four hours they set up a chemical signaling system (perhaps an alarm call?). Each cell has receptors for a specific chemical, cAMP (cyclic adenosine monophosphate), emitted by the other cells; when one detects the chemical, the cell itself emits more. The net effect is that spiral waves of cAMP pulse through the system.

At some point, each amoeba begins to move toward the center of the nearest spiral. On the way it meets aggregates of other cells and merges with them into a pattern resembling a river system. Along with David A. Kessler of Bar-Ilan University in Israel, one of us (Levine) showed that these streams can form without any specialized genetic apparatus, as a result of simple principles of physics. Ultimately, the amoebas aggregate into mounds, each of which gives rise to tens of thousands of spores that await transport to a more generous environment.

At present, biologists rely mostly on biochemical and genetic techniques to analyze how the motion of cells is controlled by external signals. Their task involves unraveling the microscopic rules of the game. But physicists have learned that getting from such microscopic laws to macroscopic patterns is also an exceedingly challenging problem. We trust the science and art of microorganisms will continue to fascinate scientists for a long time to come. SA

The Authors

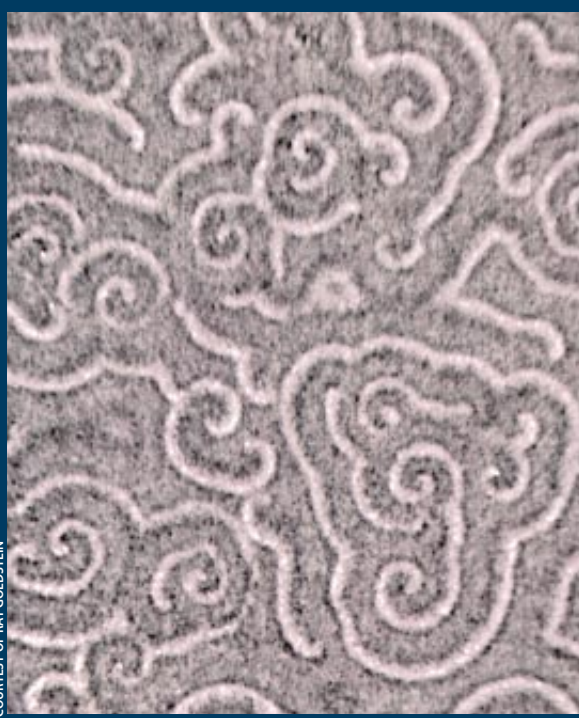
ESHEL BEN-JACOB and HERBERT LEVINE are condensed-matter physicists whose interest has diverged from snowflakes to microorganisms. Ben-Jacob is a professor at Tel Aviv University in Israel, and Levine is a professor of physics at the University of California, San Diego. The collaborators were among the first to demonstrate the role of anisotropy in patterns formed during solidification.



COURTESY OF CORNELIS J. WEIJER

River network (*above*) forms as starving cells of *Dictyostelium discoideum* aggregate into a mound, eventually to give rise to hardy spores. Researcher: Cornelis J. Weijer, University of Dundee.

The process begins with the amoebas setting up a chemical signaling system. Spiral waves of a chemical pulse over the colony, causing each amoeba to cringe: to pull in all its pseudopods (essentially, feet), changing shape. The alteration is visualized by dark-field microscopy, providing an image of the chemical waves (*below*). Researcher: Ray Goldstein, University of Arizona.



COURTESY OF RAY GOLDSTEIN

Simon Newcomb: Astronomer with an Attitude

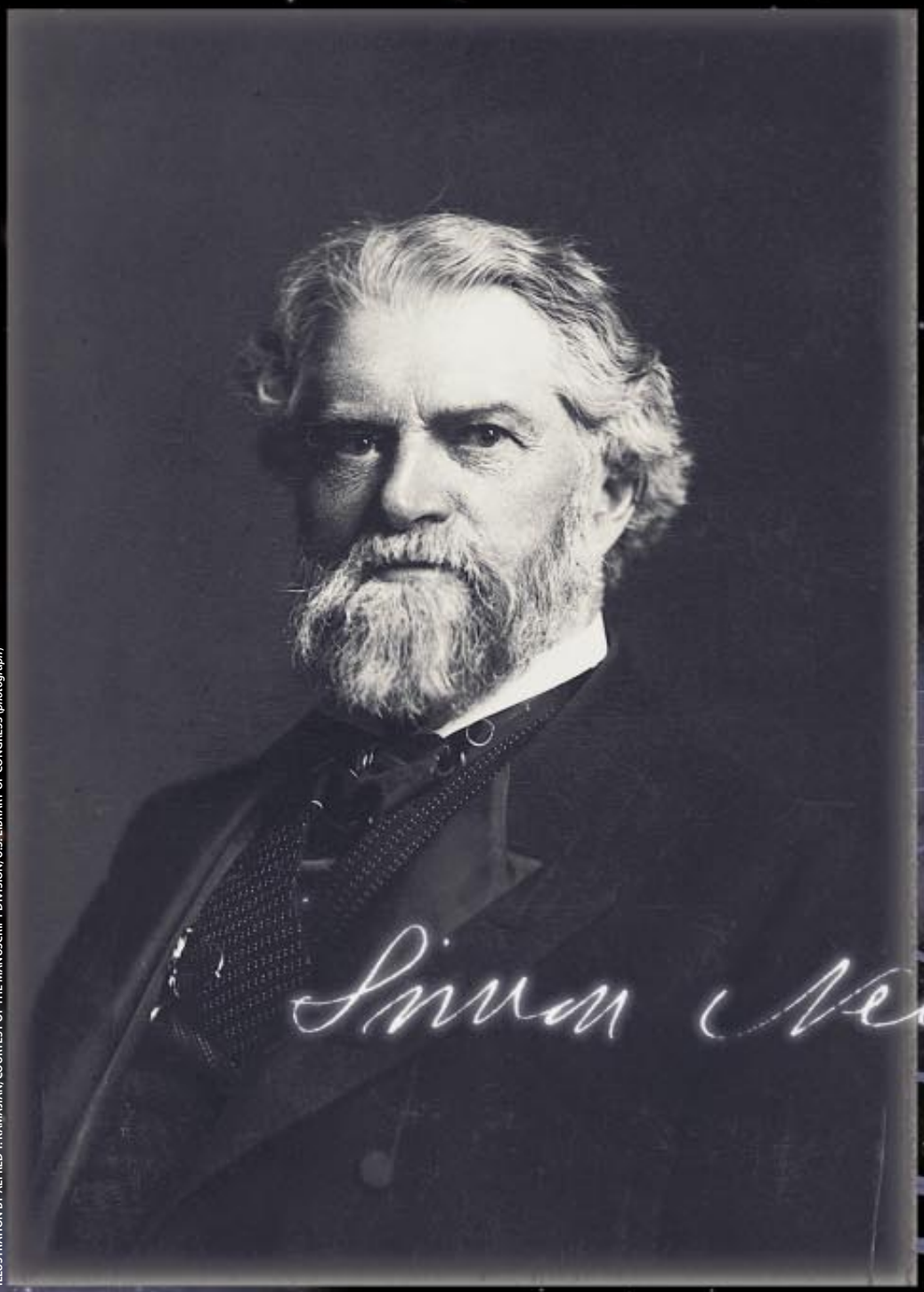


ILLUSTRATION BY ALFRED T. KAMAJIAN; COURTESY OF THE MANUSCRIPT DIVISION, U.S. LIBRARY OF CONGRESS (*photograph*)

SIMON NEWCOMB'S investigations of the stars and planets lifted classical, positional astronomy to a new level of refinement and helped to make him one of the most well-known astronomers of the late 19th century. But developments in the field of physics and astronomy during the 20th century—namely, new spectroscopic observations and the theory of relativity—overshadowed his contributions. Nevertheless, Albert Einstein would judge Newcomb's lifework to be "of monumental importance to astronomy," calling him "the last of the great masters" of classical astronomy.

The most celebrated American astronomer of the late 19th century advocated broad social and cultural reforms based on the use of scientific method

by Albert E. Moyer

The next time you hear someone appeal to the authority of the "scientific method" to prove a point, think of Simon Newcomb. Although this 19th-century astronomer is not well known today, in the decades following the Civil War, he became the era's most acclaimed American scientist. He first achieved recognition through his comprehensive studies of the motions and positions of the sun, moon, planets and stars, which helped to resolve many lingering problems in classical astronomy.

But, convinced of the value to human progress of both science and its seemingly unfaltering method, Newcomb also gained prominence in the U.S. as a commentator bent on wider social and cultural reforms. Outspoken and often openly partisan, he eagerly discussed science's place in American life and its relation to politics, economics, religion and even studies of the paranormal. In the slang of our era, Simon Newcomb had an attitude. He doggedly refused to accept the status quo in any investigative arena.

By all accounts, Newcomb was a congenial, energetic and inquisitive person. American social reformer Frederic Howe described him as "a big, lusty, joyous man." After meeting the astronomer in 1876, prominent British physicist William Thomson (later Lord Kelvin) commented that Newcomb was a "first-rate man—full of *go*." Another colleague, William Alvord, president of the Astronomical Society of the Pacific, found him to be "ruggedly independent in thought and in speech," adding that the "essential quality of his mind is that of a philosopher, rather than that of a mathematician or an astronomer merely.... In his treatment of all questions, it is the philosophical habit of his mind which is the most remarkable and the most valuable."

Newcomb was born in 1835 in Wallace, Nova Scotia, the oldest of seven siblings. His father was a teacher who traveled from school to school, while his mother—a woman of strong Calvinistic convictions—tended the home. At the age of 18, young Newcomb left his native Maritime Provinces for Maryland to work as a teacher. In his spare time, he visited the Smithsonian Institution's library in Washington, D.C., where he gradually gained the respect of its distinguished director, noted American physicist Joseph Henry. In late 1856

one of Henry's contacts found a position for Newcomb at the Nautical Almanac Office in Cambridge, Mass., an agency of the U.S. Navy responsible for creating astronomical tables used in navigation. While there, Newcomb enrolled at Harvard University's Lawrence Scientific School, where he studied under Benjamin Peirce—at the time, the nation's leading mathematician and mathematical astronomer. After two years, Newcomb graduated summa cum laude with a bachelor of science degree.

Newcomb first captured the attention of the international scientific community with a convincing mathematical demonstration, presented in 1860, showing that the asteroid belt did not result from the breakup of a former planet between Mars and Jupiter—then a commonly held belief. In 1861, having made rapid strides in astronomy, mathematics and physics, he accepted the position of professor at the U.S. Naval Observatory in Washington, D.C., continuing what would evolve into a lifelong affiliation with the navy. Three years later he became a naturalized U.S. citizen.

Although he had been hired at the Naval Observatory as an observational astronomer, Newcomb managed to squeeze in further studies in mathematical astronomy. In 1877 he parlayed this research into a position of authority, winning an appointment as superintendent of the more mathematically inclined Nautical Almanac Office, by then located in Washington, D.C. The same year, thanks to a burgeoning reputation as a researcher and administrator, Newcomb served as president of the American Association for the Advancement of Science (AAAS). He followed this with a six-year stint as vice president of the National Academy of Sciences. Also, beginning in the 1870s, he taught various courses at nearby Columbian (later George Washington) University and at the new Johns Hopkins University.

To be sure, Newcomb performed credibly as an observational astronomer. He raised the quality of work at the Naval Observatory, initially by eliminating systematic errors that tainted past values of stars' right ascensions (longitudinal positions on an imaginary celestial sphere encasing Earth) and later by taking charge of the observatory's new telescope, then the nation's largest.

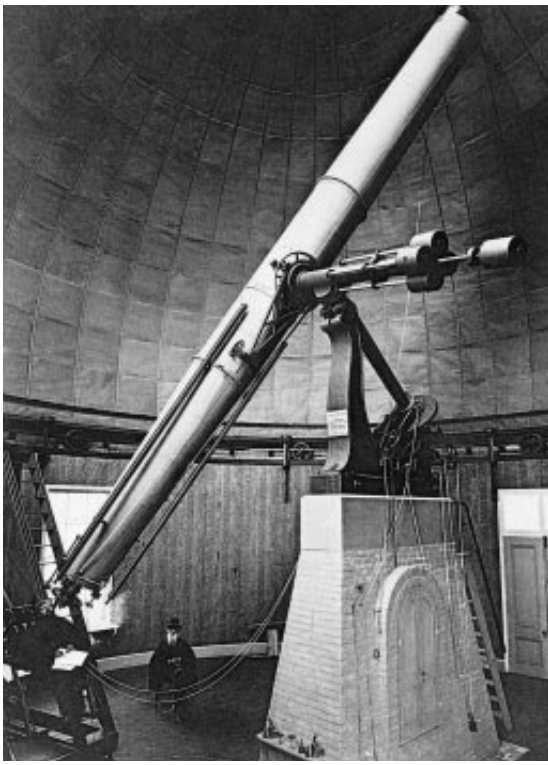
Newcomb's true calling, however, was theoretical, mathematical analysis of the orbital motions of the planets and the moon in relation to each other and to the sun. In 1874 his painstaking planetary and lunar studies earned the 39-year-old Newcomb the prestigious gold medal of London's Royal Astronomical Society. As head of the Almanac Office, he charted an even more challenging course of reform: a multidecade reevaluation of commonly accepted positions of the planets, moon and sun, coupled with a recasting of the corresponding mathematical formulas and the construction of associated tables. With support from talented assistants, he completed the bulk of this reevaluation by the mid-1890s and published his "preliminary results" in the monograph *The Elements of the Four Inner Planets and the Fundamental Constants of Astronomy*.

As the 20th century dawned, Newcomb found his work becoming the standard in positional astronomy—a status it would hold for decades to come. Furthermore, he found himself in possession of top honors from around the nation and the world, including the award of the Copley Medal of the Royal Society in London and election as one of eight Foreign Associates of the Academy of Sciences in Paris. William W. Campbell, a fellow astronomer from Lick Observatory, later described Newcomb's research program to be "of herculean and monumental proportions."

Commentators from the 20th century would look back, for example, at his exhaustive treatment of Mercury's orbit, noticing that he had pinpointed the modern value of a slight orbital anomaly (known as precession of the perihelion and first detected earlier in the 19th century). This anomaly, which Newcomb suspected defied conventional Newtonian gravitational explanations, would become intelligible only through Albert Einstein's general theory of relativity. Indeed, Einstein would judge Newcomb's lifework to be "of monumental importance to astronomy," labeling him "the last of the great masters" of classical, positional astronomy.

A Popular Advocate of Science

In addition to his prominence among astronomers, Newcomb also had a large following among the general public. His writings, including the widely reprinted and translated book *Popular Astronomy* and a successful series of mathematical textbooks, made him a household name among educated people. He also gained public attention by taking part in an intensifying national dialogue concerning science's place in American culture.



A particular image of science colored Newcomb's view of its relation to other areas of life and inquiry. Influenced by European scientific commentators such as British political philosopher John Stuart Mill (whom he met in 1870), Newcomb felt that the power of science derived mainly from the scientific method. That is, the successes of science occurred because scientists carried out their studies using well-established procedural rules. The foremost rule, in Newcomb's opinion, was to employ only those concepts that could be defined by reference to concrete experiences—an emphasis on precise meanings he called "the scientific use of language."

Newcomb expected that comparable successes awaited anyone who chose to apply the same readily accessible rules to investigations outside natural science. Scientific method was as useful in the halls of Congress as in the laboratory. As Newcomb promoted this idea in his speeches and magazine articles, he not only sparked popular support for science in the U.S. but also promoted an idea of science in the public mind as reliable, practical, even democratic. Indeed, his vision of the scientific method as dependable and accessible helped to catalyze the intellectual movement later known as American pragmatism—a cluster of cultural aspirations and philosophical predilections that developed at

PHOTOGRAPHS COURTESY OF THE LIBRARY, U.S. NAVAL OBSERVATORY

U.S. NAVAL OBSERVATORY served as Newcomb's base of operations from 1861 to 1877. While there, Newcomb had authority over what was then the nation's largest telescope; this photograph (*left*) from 1873 shows Newcomb standing at the eyepiece of the 35-foot-long telescope. In 1874 astronomers in the U.S. and Europe mounted expeditions to watch Venus pass in front of the sun as a way to measure the distance between Earth and the sun more accurately. In the picture above, Newcomb is shown speaking to the American party at the Naval Observatory; he is standing toward the left in the foreground.

the end of the 19th century (and in many ways continues to this day).

Pronouncements about scientific method provided Newcomb and the pragmatists with powerful tools of persuasion. For instance, in the 1870s Newcomb's rhetorical dependence on method quickly came to the fore when he began to call attention to the deficiencies of physical science and mathematics in the U.S. In various public forums he complained about inadequate educational programs, technical journals and professional societies. If this institutional framework were to gain strength in a nation dedicated to democratic doctrines, there needed to be fuller public support of basic research. Such support, he felt, would be forthcoming if educated citizens could be convinced of the social value of the scientific method. "I make bold to say," he asserted in an 1880 speech, "that the greatest want of the day, from a purely practical point of view, is the more general introduction of the scientific method and the scientific spirit into the discussion of those political and social problems which we encounter on our road to a higher plane of public well-being."

Newcomb did more than merely extol the use of scientific method to attack social problems. He himself developed into a political economist of wide repute, publishing technical and popular expositions on finance, trade, taxation, currency and labor. Through a steady stream of books, articles and speeches, he sought to provide a dispassionate analysis of political and economic issues, thus demonstrating—rather than merely touting—the social utility of the scientific method. Indeed, his stated goal was to reform the discipline of economics by giving it a more logical, mathematical and scientific structure. But it appears that Newcomb's unstated and perhaps unconscious goal was not only to enhance science's image but also to use its method to promote a particular political and economic platform.

His personal agenda stands out most explicitly in an article written in 1870 and incorporated 15 years later as a chapter in his massive textbook *Principles of Political Economy.* (John Maynard Keynes would look back at *Principles* and term it "one of those original works which a fresh scientific mind, not perverted by having read too much of the orthodox stuff, is able to produce from time to time in a half-formed subject like Economics.") The article's title encapsulates Newcomb's ideology; he tagged it "The Let-Alone Principle," a phrase that harks back to the hands-off policy of Adam Smith. Reflecting the British and American tradition of political and economic liberalism fostered by Smith, Mill and David Ricardo, Newcomb stated that the government should not interfere with the freedom of individuals to follow their own economic interests. (Of course, the meaning of the term "liberalism" is quite different today than in the 19th century.)

Newcomb used scientific method to justify controversial positions, such as promoting free trade and protecting the gold standard. He implied that his arguments favoring these positions were credible because he had framed them using the proper method. These contentions, of course, were not universally accepted. In the mid-1880s he found himself defending his positions (in the pages of the journal *Science,* among other outlets) from attacks by political economists who supported more aggressive government interventions. Invoking his basic rule of scientific method, Newcomb faulted his critics for their unscientific use of language. He complained that instead of restricting themselves to concrete discussions of "ploughs, fields, fences, cattle, factories, cloth, railways, locomotives" and other "really practical things," popular commentators and political economists were inappropriately pontificating about "capital," "labor," "capitalism" and other equally abstract, ill-defined terms. These pretenders were thus spawning confusion of thought.

Science and Religion

Newcomb applied the scientific method to other human issues as well. In his 1878 address as retiring president of the AAAS, he discussed the relation between natural science and the Christian religion, calling for the separation of scientific reasoning and theological arguments. The speech was reprinted in several magazines and distributed widely, resulting in a nationwide discussion on the topic. Christians who still endorsed the central claims of "natural theology" were particularly upset. They believed that science augmented religion and that the study of "design" in nature could reveal God's existence.

In his arguments, Newcomb used the scientific method as a criterion by which to differentiate scientific knowledge from religious belief, portraying the two as distinct but complementary—as alternative approaches to understanding the world. He placed the scientific use of language at the center of his methodological argument, insisting that true science restricts itself to terms that "have exact literal meanings, and refer only to things which admit of being perceived by the senses, or, at least, of being conceived as thus perceptible." Traditional natural theology failed to meet this methodological test and therefore should

be divorced from modern science, he declared. This argument enabled Newcomb to claim a province for science in American culture without, he hoped, alienating the Christians who generally dominated the culture.

In the public debate that followed, Newcomb was careful to avoid direct references to his personal religious beliefs. In private, however, he held that Christianity was an untenable, dying religion. In anonymous articles written during the time, he expressed his low esteem for Christianity and argued for the adoption of a new, humanist religion founded on personal virtue and duty to others. "Such a faith," he suggested, "fears no false teaching, sets no limit on the freedom of human thought, and views with perfect calm the subversion of any and every form of doctrinal belief, confident that the ultimate result will tend to the elevation of the human soul and the unceasing progress of spiritual development."

REPRODUCED FROM *McCLURE'S* MAGAZINE (OCTOBER 1910)

NEWLYWEDS Simon and Mary Newcomb were photographed in 1863. The couple eventually had three daughters, Anita, Emily and Anna. The oldest, Anita Newcomb McGee, went on to become a well-known physician and founder of the Army Nurse Corps.

Newcomb displayed a similarly deliberative but iconoclastic outlook when, in the mid-1880s, he became the first president of the newly organized American Society for Psychical Research (ASPR). Though skeptical of extrasensory perception, mental telepathy and other alleged paranormal phenomena, Newcomb felt that scientists should at least review the evidence for the purported effects.

The ASPR took shape in late 1884, when a group of distinguished British scientists who had recently founded their own psychical society visited North America. They convinced an equally respected group of American scientists and educators to organize their own group. Harvard psychologist William James, a prime mover of this unorthodox but fashionable society, felt that Newcomb gave the organization credibility, or as he expressed it to a colleague, "I think Newcomb, for President, was an uncommon hit."

James had a practical reason for wanting Newcomb as president, but why did Newcomb accept? Certainly, the enthusiasm of the British visitors, the flattery of being asked and the challenge of the research problem all swayed him. In spite of being an avowed skeptic, he found intriguing the possibility of mental telepathy. But he had a deeper motive. He felt a moral imperative to weigh the evidence concerning psychic effects. Specifically, his objective was to sort out what, if anything, was scientifically defensible and what must be debunked as socially dangerous fraud. Indeed, an editorial in *Science* portrayed the scientists in the ASPR as possessing the proper moral qualities to deliver the nation from the evils of raw spiritualism and substitute in its place a systematic study of phenomena such as telepathy.

Never one to investigate a topic halfheartedly, Newcomb threw himself into the thick of research on paranormal events, poring over the literature and attempting to witness occurrences firsthand. With colleagues from Johns Hopkins, he observed mediums and other masters of the paranormal in Washington, D.C., Philadelphia and New York City. By the end of his first term as president, however, Newcomb's skepticism had hardened. This became evident in his formal presidential address to the ASPR. Throughout this forceful discourse, he used elementary canons of scientific method to question claims about paranormal incidents. He targeted the British researchers, accusing them of inferring new laws of mental action without being able to replicate the relevant phenomena. James, who inclined toward belief in psychic effects, objected to Newcomb's speech and initiated a friendly skirmish with Newcomb through an exchange of published letters in *Science* and in personal notes.

Newcomb's doubt was not enough to remove him from the presidency of the ASPR, many members of which shared his skepticism. Although he offered to step down after his first term, he won reelection. And when he declined reelection at the end of a second term, he continued on the governing council until the society began to fade and the British organization absorbed it in 1889. Through these final years, he remained a dissenter. He was convinced and hoped to convince others that, on methodological grounds, psychical research was a scientific dead end.

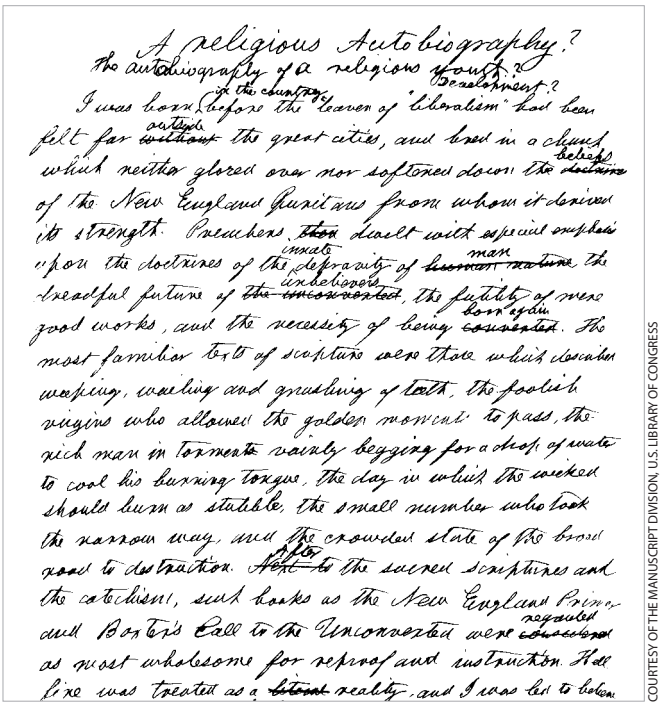
Retirement, then Back to Work

Forced by law to leave his naval post at age 62, Newcomb stepped down from the Nautical Almanac Office on his birthday—March 12, 1897. Retirement simply meant a realignment of work, not an end or even slackening of research, writing and public speaking. A modest congressional appropriation and then, beginning in 1903, generous grants from the new Carnegie Institution in Washington, D.C., enabled the distinguished retiree to maintain his intense schedule of research and professional chores. Building on his previous contributions to positional astronomy, he assumed the lead in a major international campaign to bring further order to astronomical computations and tables through the adoption of uniform constants and consistent data. Newcomb did his job so well that many of his numerical values remained in official use until the arrival of electronic computers and artificial satellites in the 1950s.

Continuing to display great drive, he also helped to organize and, in 1899, became the first president of what is now the American Astronomical Society. The creation of this society exemplifies the professional gains that the scientific community had made in the U.S. during the quarter of a century since Newcomb's first article calling attention to the nation's deficiencies in physical science.

Having a bit more leisure time, the new retiree also indulged in autobiographical reflections and even tried his hand at fiction, publishing two short stories and a science-fiction novel. Set in the distant year of 1941, his novel traces the successful quest of a Harvard physics professor to disarm the armies and navies of the world by building airships propelled by a new substance that nullifies gravity. Tongue in cheek, Newcomb included a British character named "W. K. Constant," who was the fictional counterpart of Lord Kelvin. (At least one character based on Newcomb may have appeared in fiction as well: evidence suggests that Arthur Conan Doyle picked him as the model for particular biographical features of Professor James Moriarty, the canny arch rival of Sherlock Holmes.)

Newcomb's final projects included the closing parts of his long-standing study of the moon's motion—conclusions he dictated to stenographers in 1909 as he faced imminent death from cancer of the bladder. Around the same time, he drafted a chapter on his religious views that he intended to add to a new edition of his autobiography. He had remained convinced that scientific inquiry must be kept distinct from traditional natural theology. But toward the end of the manuscript, this dying 74-year-old religious skeptic acknowledged that he had always allowed for the possibility of a "Great First Cause."



COURTESY OF THE MANUSCRIPT DIVISION, U.S. LIBRARY OF CONGRESS

RELIGIOUS AUTOBIOGRAPHY was written by Newcomb between 1879 and 1880. It was his attempt to sort out his ideas on Christianity following a controversial speech he gave in 1878 on science and religion at the American Association for the Advancement of Science. The unpublished manuscript begins, "I was born in the country before the leaven of 'liberalism' had been felt far outside the great cities, and bred in a church which neither glazed over nor softened down the beliefs of the New England Puritans from whom it derived its strength."

True to form, Newcomb clarified this position by returning to a critique of language. "But if I am asked whether I regard this cause as an intelligent one I am unable to answer until the word 'intelligent' is defined. This term implies a certain mental quality belonging to the human race and it seems to me a belittling of a great universal cause to apply any such term to it."

"All we can say," he added in his final words on the subject, "is that the cause exists and must be capable of preceding the result, which is the universe as we find it to be. But I have never indulged in vain speculation on subjects which I found it impossible to form a clear and definite conception, and so shall not pursue the subject further." Newcomb died a few weeks later, on July 11, 1909.

Throughout the U.S. and beyond, newspapers and professional journals announced the eminent astronomer's death. President William H. Taft joined other national and international dignitaries at his elaborate state funeral. Holding the relative rank of rear admiral in the navy, Newcomb was buried with full military honors in Arlington National Cemetery. The U.S. Marine Band, three companies of marines and one company of U.S. Navy bluejackets led the funeral procession from the family church to the graveside. A detail from an artillery regiment drove Newcomb's flag-draped, black caisson. A long file of carriages followed.

Although the mourners remembered Newcomb as a reform-minded astronomer and to a lesser extent as a political economist, they overlooked his broader contribution to American thought, culture and society: promoting a popular faith in science and the scientific method. For better or for worse, that faith took hold and persists within American culture to this day. To promote preferred positions, to justify pet points or to mark advantageous boundaries, people still appeal to the supposed authority of the scientific method. SA

The Author

ALBERT E. MOYER taught physics before turning to the history of science. For the past two decades, he has been at Virginia Polytechnic Institute and State University, where he is a professor and chair of the department of history. He also holds appointments as an adjunct member of the university's Center for Science and Technology Studies and as a research associate at the Smithsonian Institution. Moyer is the author of three books and a variety of articles on the development of American physical science in the 19th and 20th centuries. He remains interested in teaching physics and has worked to strengthen U.S. ties to educational programs in China.

Further Reading

THE REMINISCENCES OF AN ASTRONOMER. Simon Newcomb. Houghton Mifflin and Company, 1903.

THE PROBLEM OF PRAGMATISM IN AMERICAN HISTORY. David A. Hollinger in *Journal of American History*, Vol. 67, No. 1, pages 88–107; 1980.

SIMON NEWCOMB'S ROLE IN THE ASTRONOMICAL REVOLUTION OF THE EARLY NINETEEN HUNDREDS. Arthur L. Norberg in *Sky with Ocean Joined: Proceedings of the Sesquicentennial Symposia of the U.S. Naval Observatory*. Edited by Steven J. Dick and LeRoy E. Doggett. U.S. Naval Observatory, 1983.

THE POLITICS AND RHETORIC OF SCIENTIFIC METHOD: HISTORICAL STUDIES. Edited by John A. Schuster and Richard R. Yeo. Kluwer Academic Publishers, 1986.

A SCIENTIST'S VOICE IN AMERICAN CULTURE: SIMON NEWCOMB AND THE RHETORIC OF SCIENTIFIC METHOD. Albert E. Moyer. University of California Press, 1992. (Material from this book, the preceding article's main source, is used with permission of University of California Press.)

SPECIAL REPORT

Computer Security and the Internet

How Hackers Break In... and How They Are Caught

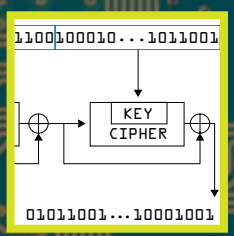
by Carolyn P. Meinel

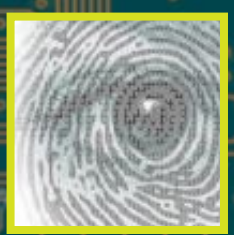
How Computer Security Works

1 Firewalls by William Cheswick & Steven M. Bellovin
2 Digital Certificates by Warwick Ford
3 The Java Sandbox by James Gosling

Cryptography for the Internet

by Philip R. Zimmermann





The Case against Regulating Encryption Technology

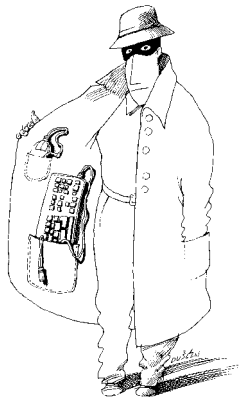
by Ronald L. Rivest

This past February hackers reached through the Internet to break into the computer networks at various U.S. Air Force and Navy sites. The intruders, allegedly two northern Californian teenagers, were trying to gain access to systems that contained sensitive shipping, personnel and payroll information. At one point, the Pentagon thought that Saddam Hussein might be responsible for the attacks—a suspicion that could have had disastrous consequences. Such electronic transgressions are hardly rare: a recent study co-authored by the Federal Bureau of Investigation found that nearly two thirds of the organizations and companies surveyed were the victims of a cyberviolation within the past year.

In this special report, experts describe advanced technologies for thwarting cybertrespassers, thieves and eavesdroppers. Defensive software and hardware, such as firewall servers, can detect and block intruders; advanced encryption techniques ensure the privacy of data should a security breach occur. Cryptography also enables confidential messages to be dispatched freely around the world over the open Internet. Such approaches for securing computers—and the electronic data they store, transmit and receive—will help preserve the Net as a shared, invaluable tool.

—The Editors

How Hackers Break In...



Port scanners, core dumps and buffer overflows are but a few of the many weapons in every sophisticated hacker's arsenal. Still, no hacker is invincible

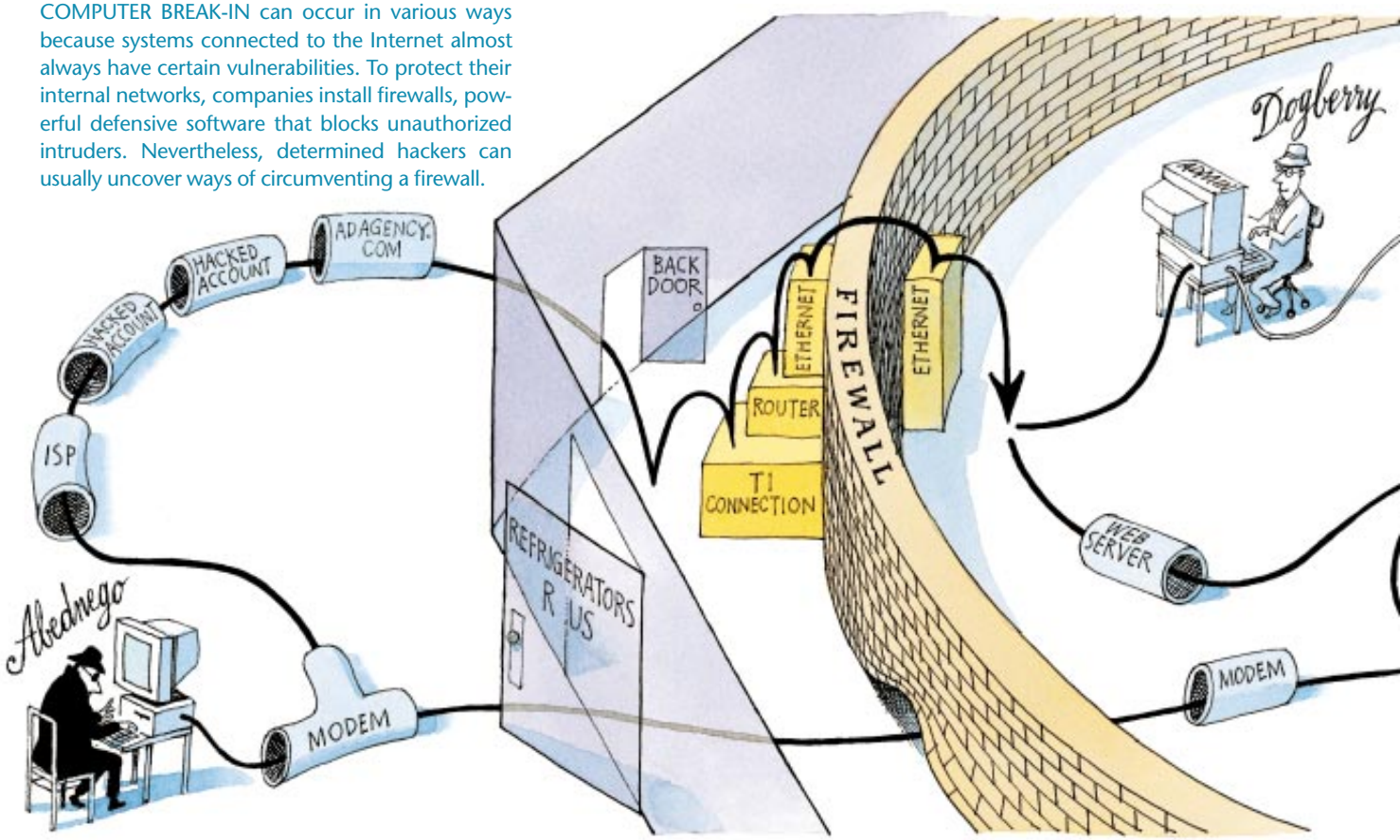
by Carolyn P. Meinel

Editors' note: *This fictionalized account is a composite of many incidents that have occurred, at one time or another, somewhere in cyberspace. The names of people and other details have been changed, but the technologies and software do exist. Some of the events reported here are drawn from the firsthand experiences of the author, who is known both in the computer underground and among security experts for her hacking skills and for her countless battles against hackers.* SCIENTIFIC AMERICAN *thanks Rt66 Internet, an Internet service provider in Albuquerque, N.M., which tested much of the software and hardware described in this article to verify the technologies involved.*

Sitting at his home computer one night, Abednego logs on to the Internet Relay Chat, the cyberspace equivalent of CB radio. After connecting to a channel devoted to the powerful Unix operating system, he watches as the on-line habitués meet to make contacts, build alliances and exchange knowledge. The scene is reminiscent of the cantina in *Star Wars*.

Eager to interject himself into the conversation—and to impress others—Abednego waits for someone to ask a simple-minded question so that he can incite a "flame war," in which the participants begin hurling venomous insults at

COMPUTER BREAK-IN can occur in various ways because systems connected to the Internet almost always have certain vulnerabilities. To protect their internal networks, companies install firewalls, powerful defensive software that blocks unauthorized intruders. Nevertheless, determined hackers can usually uncover ways of circumventing a firewall.



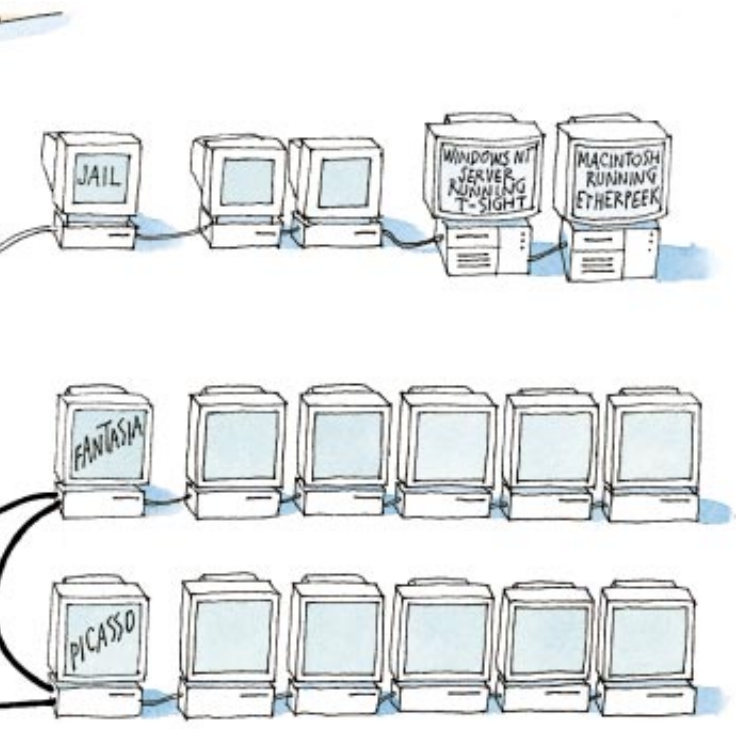
and How They Are Caught

one another. Just then, someone with the handle "Dogberry" asks about writing a device driver for a home weather station. Abednego seizes his chance. "RTFM" is his response. It stands for "read the f——g manual."

Others begin launching nasty insults, but not at Dogberry. Apparently, the question was far more complex than Abednego had realized. Dogberry's terse put-down—"Newbie!"—fans the flames. Humiliated, Abednego vows revenge.

Using the "finger" command on Internet Relay Chat, Abednego obtains the e-mail address "Dogberry@refrigerus.com." Abednego figures that if Dogberry is such a Unix whiz, he might be manager of the computers at refrigerus.com. To confirm his hunch, Abednego uses "telnet" to connect to the mail server of that computer. He then issues the command "expn root@refrigerus.com" and learns that Dogberry is indeed the head system administrator there.

His interest sufficiently piqued, Abednego runs Strobe, a program that attempts to connect with each of the thousands of virtual ports on refrigerus.com. The scanner will meticulously record responses from any daemons, which are automatic utility programs, such as those that handle e-mail. Abednego knows that each port might be an open door—or a door that he might be able to break down—if he can take advantage of some flaw in its daemon.



ALL ILLUSTRATIONS BY DUSAN PETRICIC

But Strobe hits a wall—Dogberry's firewall, to be exact. That powerful defensive software intercepts each incoming packet of data, reads its TCP/IP (transmission control protocol/Internet Protocol) header and determines with which port it seeks to connect. The firewall compares this request with its own strict rules of access. In this case, refrigerus.com has decreed that there should be only one response to Abednego's scanner.

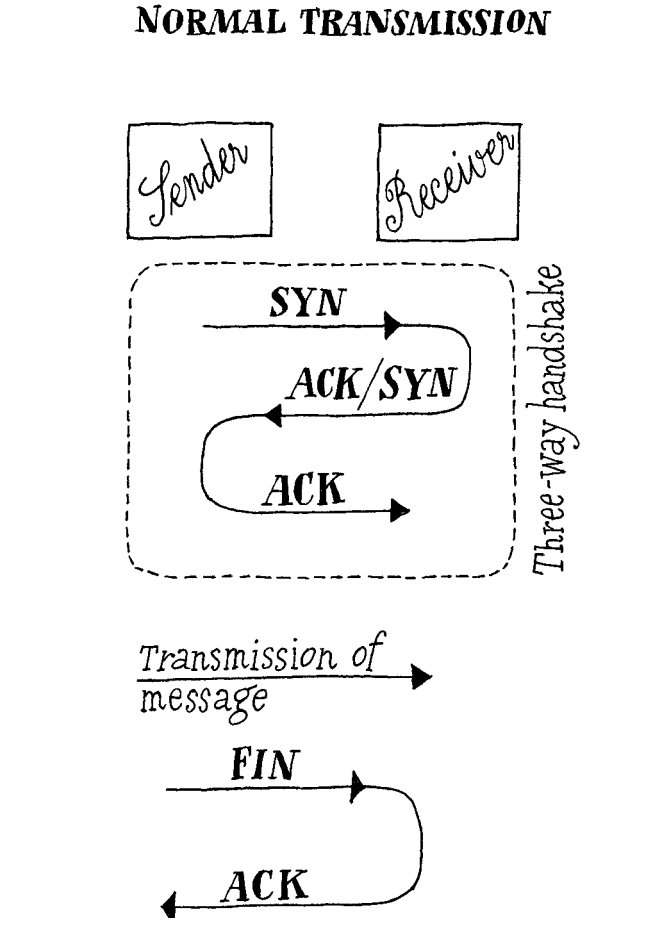
From that instant on, a program on refrigerus.com sends a blitzkrieg of meaningless data, including random alphanumeric characters, back to Abednego, overwhelming his home PC. Meanwhile another daemon sends e-mail to Abednego's Internet service provider (ISP), complaining that someone is attempting to break into refrigerus.com. Within minutes, the ISP closes Abednego's account for suspicion of computer crime.

Although Abednego is caught off guard—many ISPs would not have taken such a strong measure so quickly—the setback is minor. The closed account was only one of several he had created after breaking into that ISP. But the termination of the account at that particular moment causes him to be dumped from Internet Relay Chat in the midst of the flames against him. To the others on-line, it looks as if Abednego has been unceremoniously booted or, worse, that he has fled for cover.

Abednego burns for retaliation. His next step is to try a stealth port scanner. Such programs exploit the way in which IP transmissions work. When one computer wishes to talk to another, it must first transmit a short message packet containing a SYN (synchronize) flag. The header of the packet also contains other important information, such as the IP address of both the source and destination. In response, the recipient daemon sends back a packet that contains an ACK (to acknowledge the received packet), a SYN and a sequence number that is used to coordinate the upcoming transmission. When the first computer gets the return ACK/SYN, it issues an ACK of its own to confirm that all is ready, thus completing a three-way handshake. Then, and only then, can the sender computer begin transmitting its message using the sequence number provided. At the end of the communication, the sender transmits a packet with a FIN (finish) flag, and the receiver returns an ACK to signal that it is aware the transmission has ended.

Abednego knows that a stealth port scanner can take advantage of this process by sending just premature FIN packets to each port on a computer. Typically if a port is open, the recipient daemon will not send any response. If a port is closed, however, the computer will return an RST (reset) packet. But because this computer does not truly recognize a connection until it has completed the opening three-way handshake, it does not record the transmission in its logs. Thus, a FIN scanner can probe a computer in relative secrecy, without ever having opened any official connections. (Yet, as Abednego will soon learn, there is enough information in even one FIN packet to establish a sender's identity.)

Abednego surfs the Internet to search for an advanced



INTERNET TRANSMISSIONS follow certain rigid protocols. Normally, the sender first transmits an introductory message packet containing a SYN flag to synchronize the upcoming communication (*top*). The receiver then returns an ACK, which acknowledges the request, and a SYN. After obtaining this information, the sender transmits an ACK, which completes the necessary three-way handshake. Only then can the sender dispatch the message itself. When finished, he issues a FIN flag, and the receiver returns an ACK, which officially closes the correspondence. A hacker can circumvent the process by sending just a premature FIN, from which the hapless receiver might return an RST, or reset, packet (*bottom*). The response—or lack of one—reveals certain information about the receiver, but because no three-way handshake was ever completed, the transmission is not recorded in the receiver's logs. The hacker can thus probe an unwitting computer in relative secrecy.

stealth port scanner and finds one at an underground Web site. The program, like most other hacker tools, is written in the C computer language. Abednego struggles to compile, or convert, the scanner from C into a form that can be executed on his home PC, which runs on Linux, one of the many variants of Unix.

Abednego's difficulty in converting the software is not unusual because of the many peculiarities of the different flavors of Unix. And Abednego, like many hackers, did not formally study computer science. In fact, also like most hackers, Abednego never learned to program because he never had to: almost any software a computer criminal might ever want is available on the Internet, already written and free for the taking—as long as the hacker knows how to compile it (or has cohorts who do).

The young Dogberry had taken a different path. After befriending a technician at a local ISP, he learned how to administer a network. Before long, Dogberry and the technician were playing computer break-in and defense games. The payoff came when they used the results to help the ISP improve its security. With that success, Dogberry was hired by the ISP to work part-time while he pursued his computer science degrees.

Thus, when Abednego decided to take on Dogberry, he had already made his first mistake. Dogberry is a white-hat (or nonmalicious) hacker and a veteran of many cyberbattles.

Casing the Joint

As dawn breaks, Abednego has finally finished compiling the code and is ready to deploy it. Within minutes, the FIN scanner has given him a snapshot of the services that refrigerus.com offers to those coming only from an approved IP address. Two that draw his attention are a secure-shell daemon, which is a way to make encrypted Internet connections, and a Web server.

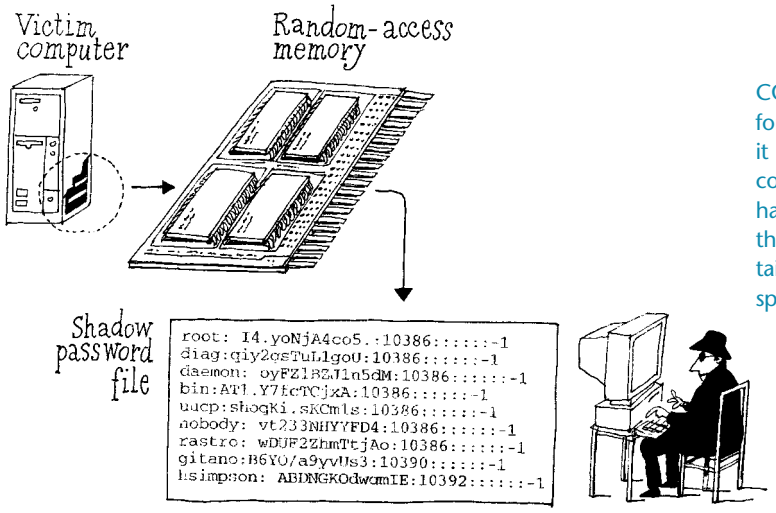
Then Abednego's heart skips a beat. An unusual port number, 31,659, has also turned up on his FIN scan. Could another intruder have preceded him and left a back door, a secret passage to enter the system undetected?

The beeping of a pager jolts Dogberry out of a deep sleep. EtherPeek, a sniffer program installed on the refrigerus.com network, has detected the port scan. Dogberry rushes into the office to watch for more attacks from the console of his administrative computer. His best defensive programs run only from that machine and only for someone who is physically there, so that they cannot be tampered with remotely by an attacker.

Meanwhile, despite the powerful temptation of that 31,659 daemon, Abednego leaves the chase for now. Something—his hacker intuition—tells him that he should return on another night. So by the time Dogberry arrives at work, he sees no more activity.

Curious about the unusual attack, though, Dogberry begins analyzing his computer logs and is able to retrieve the source address from the hacker's FIN packets. With this information, he sends an e-mail to Abednego's ISP, advising the firm of the break-in attempt and asking for details about Abednego's account. But the system administrator at the ISP rejects Dogberry's request, citing a confidentiality policy, because merely running a scanner breaks no law.

Three evenings later Abednego resumes the hunt. But when his computer dials into his account, he finds out his



CORE DUMP can be used by hackers to obtain secret information. When a program running on a computer fails, it sometimes causes the machine to dump, or flush, the contents of a part of its random-access memory (RAM). A hacker can force such an incident to occur so that he can then sift through the discarded data, which might contain important information, such as the passwords for specific accounts on the network system.

password is no longer good. Upset, he phones the ISP and learns that his account has been shut down because of the FIN scan. Yet this turn of events does little to discourage him. In fact, he is now even more determined.

With his credit-card number and a telephone call to a different ISP, he is back on-line within minutes. This time, though, Abednego is more cautious. Through this new account, he logs on to one of his hacked accounts at yet another ISP. Once there, he gives the simple command "whois refrigerus.com." The response tells him the domain name belongs to Refrigerators R Us, a national retail chain.

Next, Abednego tries to log on to refrigerus.com through the 31,659 port by issuing the command "telnet refrigerus.com 31,659." The response is, "You lamer! Did you really think this was a back door?!" Then the 31,659 daemon attempts to crash his PC by sending corrupt packets, while e-mailing the system administrator at Abednego's hacked ISP that someone has attempted to commit a computer crime. Within minutes, Abednego's connection dies.

More determined, Abednego now tries to tiptoe around the firewall instead of forcing his way through it. Using yet another of his many hacked accounts, he begins by attempting to catalogue the computers that belong to refrigerus.com. To obtain this information, he tries "nslookup," which initiates a search throughout the Internet for master databases containing directories of IP addresses.

But "nslookup" is unable to retrieve anything useful. Dogberry must have set up the refrigerus.com network so that all packets destined for any of its internal addresses are sent first to a name-server program, which then directs them to the appropriate computers within the network. This process hinders anyone on the outside from learning details about the computers inside the firewall.

Abednego's next attempt is through an IP address scanner. First, he converts refrigerus.com to a numerical address, using "nslookup." With that number as a starting place, he scans the IP addresses above and below it. He discovers some 50 Internet host computers. Although there is no guarantee that these belong to refrigerus.com, Abednego knows it is a good bet they do.

Next, he uses "whois" to ask whether any other domain names are registered to Refrigerators R Us. The response reveals another: refrigeratorz.com, with an address that is numerically distant from that of refrigerus.com. The IP address scanner soon reveals five additional Internet hosts on numbers nearby refrigeratorz.com.

HACKER LEXICON

Abednego—A biblical Israelite held in Babylonian captivity who walked through a wall of fire and survived.

ACK—See illustration on page 100.

Back door—A secret way to enter a computer that bypasses normal security procedures.

Buffer overflow—See illustration on page 102.

Core dump—See illustration on this page.

Daemon—An automatic utility program that runs in the background of a computer.

Dogberry—The constable in William Shakespeare's *Much Ado about Nothing.*

FIN—See illustration on page 100.

Firewall—Defensive software that protects a computer system from unauthorized intruders.

FTP—File transfer protocol, a common protocol and program used to transfer files over the Internet.

IP—Internet Protocol, a low-level convention that allows computers to move packets of data across the Internet.

Internet Relay Chat—An on-line chat service.

ISP—Internet service provider.

Keystroke logger—A program that records everything a user types at a keyboard.

Port—A connection, or channel, into a computer.

RAM—Random-access memory.

Root—The highest level of access to a Unix computer.

Root kit—A program that hackers implant in a victim computer to hide their nefarious activities.

RST—See illustration on page 100.

Scanner—A program that attempts to learn about the weaknesses of a victim computer by repeatedly probing it with requests for information.

Sequence number—A number used to coordinate an upcoming IP transmission.

Shell—A software layer that provides the interface between a user and the operating system of a computer.

Sniffer—A program that records computer and network activity.

Spoof—See illustration on page 104.

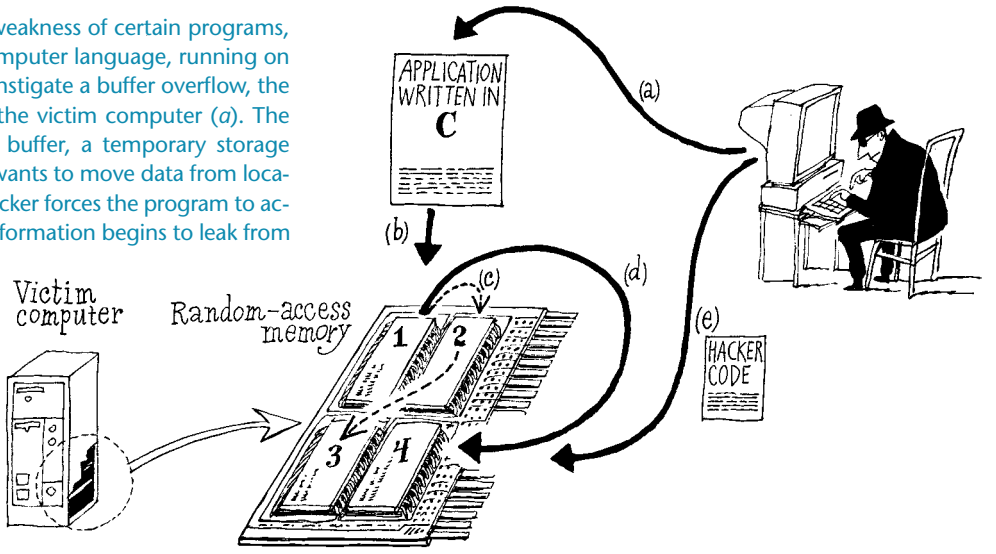
SYN—See illustration on page 100.

TCP—Transmission control protocol, the set of communications conventions that enable the sending and receiving of data over the Internet.

Telnet—A Unix command that enables a user to log on to a computer from a remote location.

Unix—A powerful operating system.

War dialer—A program that will automatically dial a range of telephone numbers.



BUFFER OVERFLOW is an exploitable weakness of certain programs, for example, those written in the C computer language, running on an operating system such as Unix. To instigate a buffer overflow, the hacker might run a C application on the victim computer (*a*). The program begins to write data into a buffer, a temporary storage space in memory (*b*). The application wants to move data from location 1 into 2, then into 3 (*c*). But the hacker forces the program to accept excess data so that some of the information begins to leak from location 1 into 4 (*d*). The hacker can take advantage of the overflow to insert his own code (*e*), which has been written to help him gain high-level privileges to the victim computer.

As a safety precaution, Abednego telnets from his current hacked account into another of his pirated accounts. He then telnets from that location to yet another account that he has hacked, remotely logging on to it in preparation to run more FIN port scans. The extra steps will force anyone in law enforcement to obtain search warrants for three companies, encumbering the process.

He also decides to hide on the third hacked computer under the protection of a root kit, a Trojan horse program that, despite its harmless appearance, will automatically delete any evidence of his actions from the logs used to detect abnormal activities. The software also defeats other programs that seek to detect alterations to system files on that computer. A root kit will even prevent people from determining that he is logged on and running programs.

From this safe perch, Abednego scans one after another of the Internet host computers at refrigerus.com and refrigeratorz.com. The FIN scanner slips straight through the firewall to every one of them. The activity, though, is detected by the EtherPeek sniffer, which again sets off Dogberry's beeper.

A haggard Dogberry, after rushing to work, soon identifies the origin of the FIN scans and alerts the system administrator at Abednego's third hacked account. But the root kit has done its job, hiding Abednego from mystified computer operators there. Abednego boldly continues, switching from the stealth scanner to Strobe in hopes of finding an IP address that the firewall does not protect.

He succeeds only in having the refrigerus.com firewall unleash a flood of meaningless data. The sudden load finally convinces the system administrator at Abednego's hacked account that there really must be an attacker at work. She takes the drastic step of cutting the entire system off from the Internet. As his connection fizzles, Abednego realizes there is no elegant way around the firewall.

Finding a Workaholic

For each of the several dozen Internet hosts at Refrigerators R Us, Abednego guesses that there are probably many other desktop computers sitting quietly in employees' cubicles and offices. What are the chances, he muses a few nights later, that somewhere among those hundreds of users are workaholics who circumvent the company firewall by phoning into their computers from their homes to perform late-night tasks? It's simple, really, for someone to buy a modem, connect it to a computer at work and plug a phone line into it before leaving for the day.

Knowing that almost every large corporation has at least one unauthorized modem on its network, Abednego sets up ShokDial, a war-dialer program that will call each of the extensions to the phone system at Refrigerators R Us as well as other numbers within that exchange. At the headquarters building of the company, the night watchman hears the ringing of one office phone after another but thinks nothing of it.

Then, at 2:57 A.M., the war dialer pulls up a modem, and Abednego is greeted with the log-on screen of a Silicon Graphics computer: "Refrigerators R Us Marketing Department. Irix 6.3." Great, Abednego thinks, because Irix is a type of Unix, which means he has found a potent portal into Dogberry's world.

Abednego's next strategy is to try brute force, using a program that will repeatedly dial the Irix box and guess passwords for root, a top-level account (usually reserved for system administrators) from which he can run any command and access all information on that particular computer. He is hoping that the owner of the Irix machine, like many harried workaholics, has negligently allowed remote access to a root account.

The password guesser starts with common words and names and from there tries less obvious choices. The slow, painstaking process can take months, even years, as the program exhausts every word in an unabridged dictionary, all names in an encyclopedia and each entry from a local phone book. But Abednego gets lucky. Around 5 A.M. he learns that the password is simply "nancy."

"Yes!" Abednego shouts as he logs on to a root shell, from which he can then issue other commands to run on that machine. Next, he secures his beachhead, using FTP (file transfer protocol) to plant a root kit and sniffer onto his latest victim. He sets the program to capture and record everything typed in at the console (a process known as keystroke logging), as well as any log-on sessions from the network. The sniffer will hide this information in an innocuously named file right there on the unwitting host. Within min-

utes, Abednego's root kit has even set up an additional way to log on: user name "revenge," password "DiEd0gB."

Abednego's last deed that morning is simple. To find the Internet address of the hijacked box, he types the "who" command, and his computer shows user "revenge" logged on to picasso.refrigeratorz.com. Later that morning the rightful owner of picasso logs on and sees no indication that someone has usurped control of her computer. Abednego's root kit is doing its job.

For Dogberry's part, all that his log reveals is an early-morning attempt to enter refrigeratorz.com from the Internet. Remembering the recent FIN scans, Dogberry is troubled by this latest incident, but he has too little information to take action.

Two nights later Abednego dials in and connects with picasso to view his logs. To his dismay, he sees that information on the internal network traffic has been encrypted. But the keystroke logger of his sniffer has recorded that someone on picasso had logged on to another computer named fantasia. Abednego now owns a user name and password for fantasia. Open sesame!

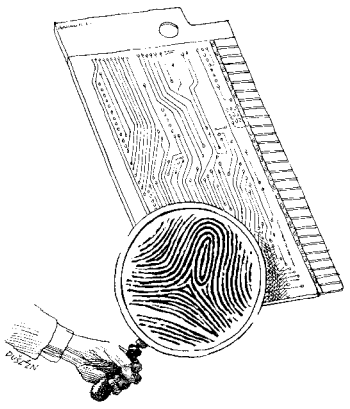
Abednego discovers that the computer is a SPARC workstation used for rendering animated sequences, perhaps for television ads. Because the box is probably a server used by many other computers, Abednego begins hunting for a password file, hoping that some of the passwords he finds will also work on other machines inside the company network.

He locates the file but discovers only "x" characters where the encrypted passwords should have been. Apparently, the information he seeks is hidden elsewhere in a shadowed file. Smiling to himself, Abednego runs the FTP program and tricks it into crashing. Bingo, core dump!

Fantasia is forced to flush a part of its random-access memory (RAM). Fortunately for Abednego, the discarded information—a record of what was being held in that RAM sector at that moment—ends up in the user directory.

The legitimate purpose of a core dump is to enable programmers to perform an autopsy on the digital remains in search of clues to a program's failure. But, as Abednego well knows, a core dump has other uses. A shadowed password system sometimes places encrypted passwords in RAM. When a person logs on, the computer does a one-way encryption of the password the user attempts and compares that with the encrypted password from the shadowed file. If the two match, the person gets in.

The shadowed password file that Abednego is able to retrieve from the core dump on fantasia is encrypted, so he starts running his password cracker. The program could be busy for the next few days, maybe even weeks.



Too impatient to wait, Abednego is already working on his next maneuver—exploiting a common vulnerability of Unix. When a program running on that operating system pours excessive data into a buffer (a temporary storage space in memory), the information will leak, infiltrating other areas of the computer's memory.

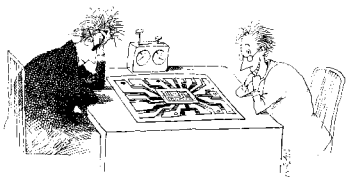
Abednego takes advantage of the buffer overflow by using it to insinuate his own code into the SPARC. The added software helps him create a root shell, from which he can then run other commands and programs. Pleased with his latest effort, Abednego next installs a root kit and sniffer. Because the kit will hide evidence of his activities only from the time when the program was activated, Abednego must mop up by deleting previous actions of his busy night.

One task remains. Is there anyone who is allowed to log on to fantasia from the Internet? Abednego types the "last" command to display records of connections people may have made to fantasia. He perks up as he sees that user names vangogh and nancy have recently entered fantasia from the Internet through the domain "adagency.com," which lies outside the Refrigerators R Us firewall.

Abednego can hardly fall asleep that morning. His adrenaline flowing, he buzzes with the knowledge that he will soon "own" Refrigerators R Us.

Closing in for the Kill

The next evening Abednego makes short work of breaking into adagency.com. At first he uses IP spoofing to trick that computer into recording a false IP address for his location. By probing adagency.com with SYN packets to elicit ACK/SYN responses with an assortment of sequence numbers, Abednego's program is able to tease out a pattern from which he can then guess the next sequence numbers and use that knowledge to fake his origin. Abednego quickly installs a sniffer on adagency.com and uses a secure-shell program to create an encrypted connection for logging on to fantasia.



From that computer, he types the "netstat" command to view tables of active connections within the network. He discovers a computer that he had missed in his earlier search. Its name, "admin.refrigerus.com," is promising. Could that be from where Dogberry oversees the system?

Meanwhile every time Abednego's PC cracks yet another combination of user name and password, he tries it on various refrigerus.com computers. But none of them works anywhere except on fantasia, which he already "owns."

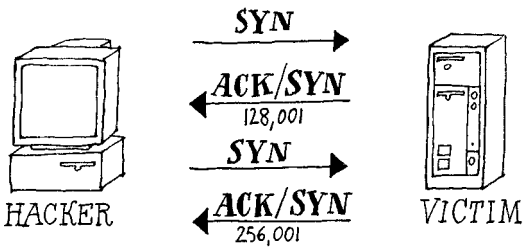
Then Abednego hits the jackpot. Twice.

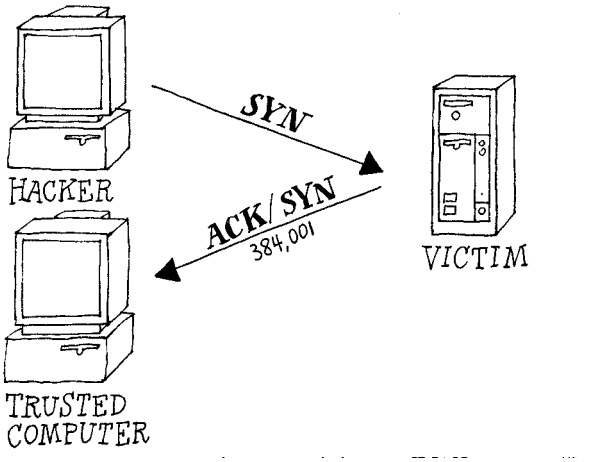
On fantasia he captures keystrokes made by vangogh as that user updated the company's Web server. Now Abednego has the password he needs to hack the Refrigerators R Us Web site. In addition, his sniffer on picasso reveals that someone, Nancy, has dialed into that computer and from there used a back door to log on to a root account, hidden by her root kit, at admin.refrigerus.com.

He slips right behind Nancy into admin.refrigerus.com. Using the root account there, he tries logging on to one Refrigerators R Us computer after another. Dogberry, however, has been exceedingly careful. On the Refrigerators R Us network, even root privileges do not allow someone to enter other computers without providing new passwords.

Only briefly distracted, he turns his attention back to the Web server and logs on to it using his recently acquired password. From his home PC, he then uploads a new version of the Refrigerators R Us home page that he had put together in anticipation of this day.

Back at Refrigerators R Us, Dogberry is working late, poring over his logs. It seems the marketing people have been





IP SPOOFING enables a hacker to fake his identity. The hacker first probes his victim by sending multiple SYN packets [*see illustration on page 100*] to obtain ACK/SYN messages with sequence numbers (*left*). From these responses, the hacker is able to uncover a pattern. In this example, he notices that the numbers increase by an increment of 128,000. Next, the hacker sends a SYN that impersonates another computer that the victim trusts. The victim then transmits an ACK/SYN to this authorized host (*center*). Although the hacker does not receive this particular response, he can nonetheless continue the correspondence as if he had: he is-

getting an unusual number of connections from adagency.com. Tomorrow he will ask those folks exactly what is going on. He will also call the system administrator at adagency.com, a colleague whom he once helped to install some new system software.

Just as Dogberry is about to head home for the night, the phone in his office rings. An angry customer complains that Refrigerators R Us's Web site features a pornographic movie with a refrigerator as a prop. After bringing up and viewing the defaced Web page, Dogberry moves quickly to sever the umbilical Ethernet cable that connects the company network to the Internet.

Abednego is enraged when his obscene masterpiece is taken down so quickly. But he is also worried that he has left too much evidence behind, so he returns using the dial-up line to picasso—an entryway that is still unknown to Dogberry. He buys time by reformatting completely the administrative computer's hard disk, which shuts down the company network, temporarily thwarting Dogberry's efforts to gather details of the attack.

Dogberry rushes to the administrative computer with hopes to reboot it from the console, but he is too late. Dogberry must now rebuild the software on that computer from scratch. (Unbeknownst to Abednego, though, the EtherPeek sniffer running on a nearby Macintosh has also been making logs.)

Abednego, still peeved about the Web site, has one final act that night: he unleashes a flood of data packets against refrigerus.com. Soon Dogberry gets a frantic call from a company salesperson who, using her laptop PC and a phone line in her hotel room, wants to retrieve her important e-mail but has been unable to connect to the mail server at Refrigerators R Us.

The next morning an exhausted Dogberry begs the vice president of technology at Refrigerators R Us for an okay to wipe clean every computer in the network, reinstall every program and change all passwords. But the extensive—though prudent—measure would require shutting the system down for days, and the vice president denies the request.

At this point, Abednego's malicious and destructive exploits have gone well past the legal bounds for hacking. But the FBI, which is severely understaffed, has been busy investigating some recent break-ins at several army and navy computer systems around the U.S. Dogberry will have to gather more evidence himself.

Because the attacker remained on the system even after it had been physically disconnected from the Internet, Dogberry suspects there must be a contraband modem somewhere in the building. He runs his own war dialer and discovers the culprit. He will soon have words with the marketing department!

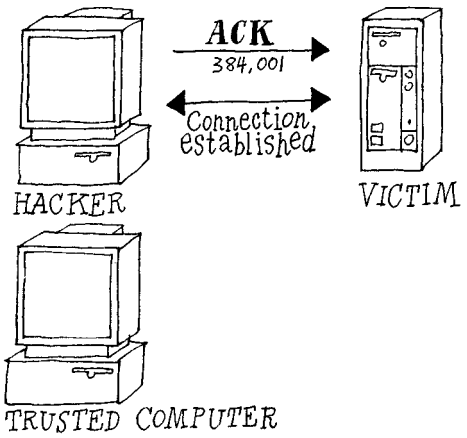
Dogberry then reloads a clean version of his main administrative computer. Next, on a Windows NT server that Dogberry knows has not been tampered with, he deploys T-sight, an advanced antihacker program that can monitor every machine on the company network.

Last, Dogberry sets his trap. T-sight will watch for the attacker's next connection to admin.refrigerus.com and will redirect the intruder into a "jail" computer. Once there, the culprit can be monitored and traced. To keep the unsuspecting person distracted, Dogberry enlists a team of programmers to make the jail look like an accounting system, complete with the tempting bait of fake financial data.

Pride Goeth Before…

Just two nights later Dogberry is standing watch at 8:17 P.M. when he discovers someone once again entering admin.refrigerus.com. It is Abednego. Why has he returned so soon? Abednego was exhilarated when he learned that his pornographic Web site had become the talk of the hacker underground. He had even rated a brief mention on CNN. The publicity and his hubris were a potent combination, making Abednego feel invincible.

In fact, tonight he has brazenly reentered Refrigerators R Us without his customary caution. After dialing into a guest account on an ISP, he telnetted directly to adagency.com to gain faster access to fantasia's back door.



sues an ACK message with the correct predicted sequence number, thus establishing a connection between his computer and the victim (*right*). The hacker can then transmit information that the victim will assume is benign because of the mistaken belief that it is coming from the trusted host computer.

From admin.refrigerus.com, Abednego is lured to the jail by T-sight. He can hardly control his excitement as he begins sifting through what he believes are sensitive financial records.

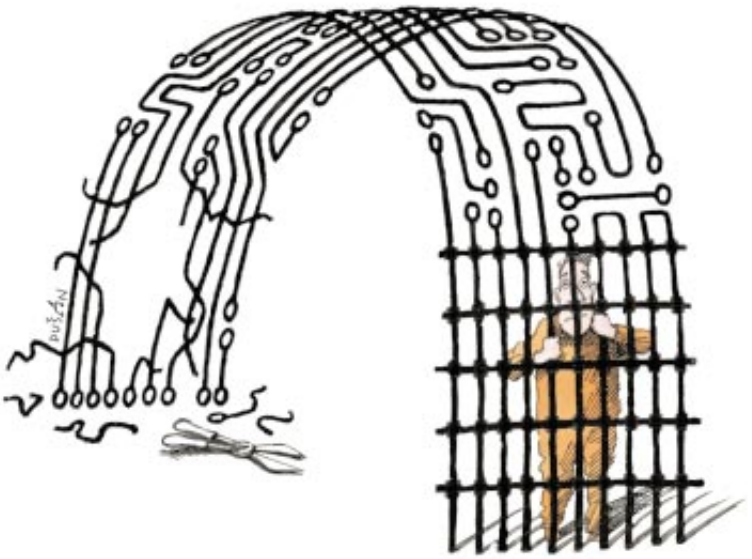
Dogberry, too, is busy. Quickly analyzing data from T-sight, he obtains Abednego's root password on fantasia—DiEd0gB—and is able to trace the intruder back to adagency.com. Dogberry calls the pager of the system administrator there. She has already left work, but she phones Dogberry from a restaurant to help him continue tracking Abednego.

So while Abednego is retrieving a huge file containing bogus credit-card numbers, Dogberry installs a sniffer on adagency.com. He is even able to sneak unnoticed into Abednego's account on that computer by typing DiEd0gB, because Abednego has lazily used the same password for all his root kits. Then, just minutes before Abednego finishes his download and logs off, Dogberry succeeds in tracking the trail of the plundered credit-card file back to Abednego's dial-up account at the ISP.

The information Dogberry has obtained is enough to bring in the FBI, which contacts the ISP the next day to obtain Abednego's identity from the company's phone logs. With enough evidence in hand, including the Macintosh's high-quality EtherPeek logs, the U.S. Attorney's office approves a search warrant.

Soon after, FBI agents raid Abednego's apartment and confiscate his PC. The hard disk of the computer will reveal all. Abednego had taken the precaution of erasing incriminating files from his PC after each night's escapade. He is chagrined to learn that the FBI can extract that information from his hard drive even after it had been erased and overwritten several times. Soon a laboratory has recovered details of his past trespasses, including the time he romped through the computer system at a major banking institution in the Northeast.

The megabytes of incriminating data provide the smoking gun necessary to indict Abednego on multiple counts of computer fraud. Unfortunately for him, the trial judge assigned to his case is known for her tough stance on cybercrime. Taking his attorney's advice, Abednego wisely accepts a plea bargain even though, like many hackers who have crossed the line, he insists that his activities—which, for Refrigerators R Us alone, resulted in thousands of dollars in damages—were just playful pranks. Abednego is currently serving a two-year sentence in a federal prison. SA



The Author

CAROLYN P. MEINEL is addicted to the frontiers of technology. She is the author of *The Happy Hacker: A Guide to (Mostly) Harmless Computer Hacking* (American Eagle Publications, 1998) and the president of Happy Hacker, Inc., a nonprofit organization devoted to teaching people how to hack responsibly and legally. The group runs a hacker wargame on its World Wide Web site (http://www.happyhacker.org). Meinel holds an M.S. in industrial engineering from the University of Arizona. She is currently working on a book with Jason Chapman, recounting her many adventures battling hackers. Meinel wishes to acknowledge the help of Michael and Diana Neuman, Damian Bates, Emilio Gomez, Mahboud Zabetian and Mark Schmitz in researching this manuscript.

Further Reading

ESSENTIAL SYSTEM ADMINISTRATION. Second edition. Æleen Frisch. O'Reilly & Associates, Sebastopol, Calif., 1995.

INTERNET FIREWALLS AND NETWORK SECURITY. Second edition. Chris Hare and Karanjit Siyan. New Riders Publishing, Indianapolis, 1996.

MAXIMUM SECURITY: A HACKER'S GUIDE TO PROTECTING YOUR INTERNET SITE AND NETWORK. Anonymous. Sams Publishing, Indianapolis, 1997.

THE GIANT BLACK BOOK OF COMPUTER VIRUSES. Second edition. Mark Ludwig. American Eagle Publications, Show Low, Ariz., 1998.

Additional information can be obtained at http://www.geek-girl.com/bugtraq, http://ntbugtraq.ntadvice.com, http://rootshell.com, http://www.infowar.com, http://www.antionline.com and http://www.happyhacker.org on the World Wide Web.

Details of EtherPeek and T-sight can be obtained at http://www.aggroup.com and http://www.engarde.com, respectively.

How Computer Security

Three types of safeguards offer a formidable defense against Internet intruders

1 Firewalls

by William Cheswick, *Bell Labs* (*division of Lucent Technologies*) and Steven M. Bellovin, *AT&T Research*

Computer networks will always be vulnerable to attack. As long as companies use the Internet—for transferring files, sending e-mail, downloading programs and so on—there will always be the chance that some malicious outsider will find a way to wreak havoc with their computer systems. But there are ways to make a network much more resistant to attack. The first line of defense is the firewall, a software program that acts as a gatekeeper between the Internet and a company's "intranet"—the network of computers used by the company's employees.

The two most common kinds of firewalls are packet filters and application-level firewalls. A packet filter, which typically runs on a machine called a router, examines the source address and destination address of every packet of data going in or out of the company's network. The filter can block packets from certain addresses from entering the network—and prevent other packets from leaving. An application-level firewall examines the content of the Internet traffic as well as the addresses; it is slower than a packet filter, but it allows the company to implement a more detailed security policy.

In the illustration, the maze of offices represents a computer network protected by firewalls. The green figures symbolize authorized packets of data. The red figures are potentially harmful packets that should not be allowed into the network.

A

Outside the firewall (*outlined in orange*), a company may set up a network that is accessible to the general public. Such a network—often called a demilitarized zone, or DMZ—allows customers to send e-mail to the company or browse through the company's site on the World Wide Web. The DMZ is analogous to the public lobby of an office building. It is a good place for customers to get information on the company's products, but because the DMZ is open to anyone, sensitive data should not be placed there.

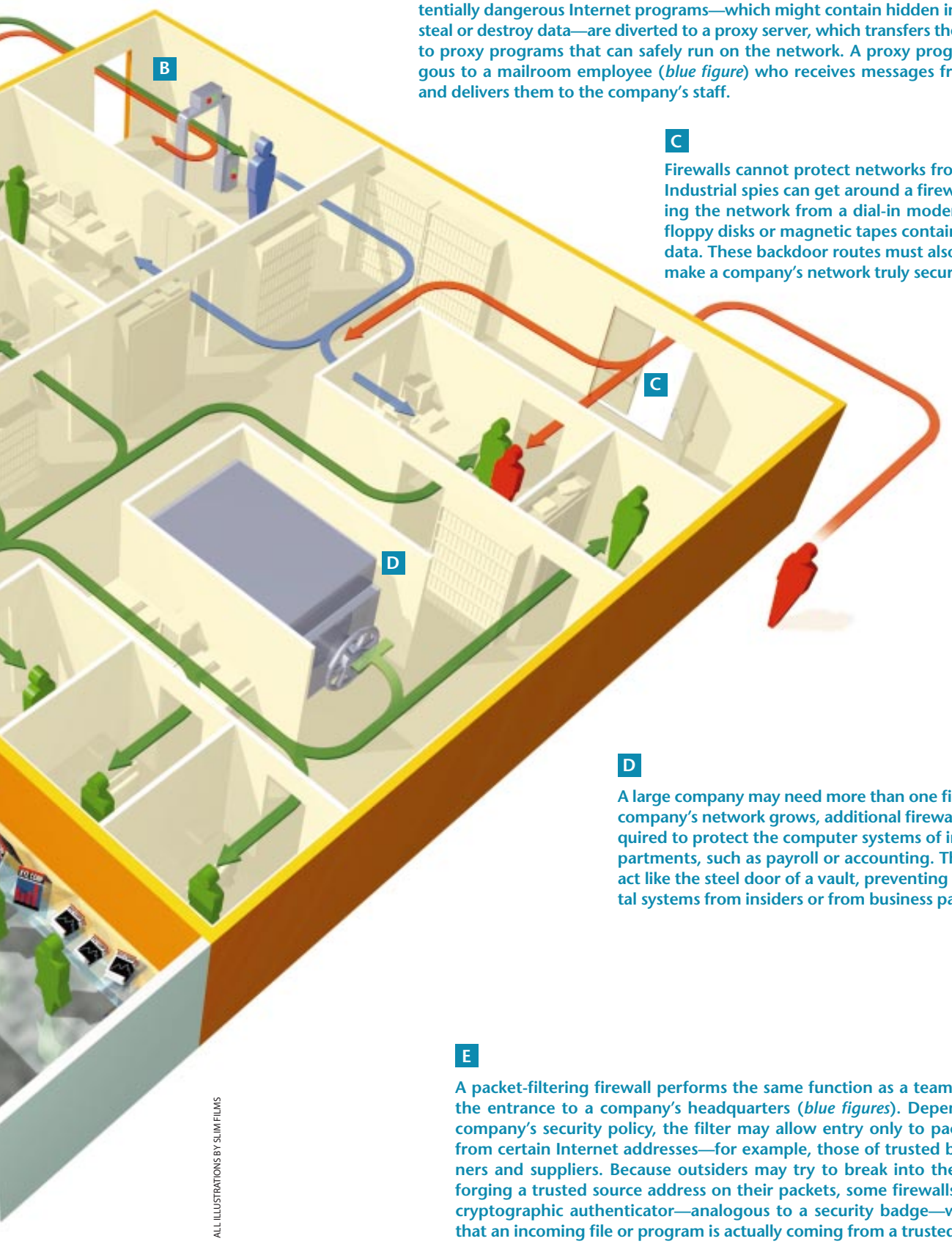
Works

B

An application-level firewall is like the company's mailroom: it can scan incoming e-mail for computer viruses just as mailroom employees can x-ray bulky packages. Potentially dangerous Internet programs—which might contain hidden instructions to steal or destroy data—are diverted to a proxy server, which transfers the information to proxy programs that can safely run on the network. A proxy program is analogous to a mailroom employee (*blue figure*) who receives messages from outsiders and delivers them to the company's staff.

C

Firewalls cannot protect networks from all attacks. Industrial spies can get around a firewall by accessing the network from a dial-in modem or stealing floppy disks or magnetic tapes containing sensitive data. These backdoor routes must also be closed to make a company's network truly secure.



D

A large company may need more than one firewall. As the company's network grows, additional firewalls may be required to protect the computer systems of important departments, such as payroll or accounting. These firewalls act like the steel door of a vault, preventing attacks on vital systems from insiders or from business partners.

E

A packet-filtering firewall performs the same function as a team of guards at the entrance to a company's headquarters (*blue figures*). Depending on the company's security policy, the filter may allow entry only to packets coming from certain Internet addresses—for example, those of trusted business partners and suppliers. Because outsiders may try to break into the network by forging a trusted source address on their packets, some firewalls search for a cryptographic authenticator—analogous to a security badge—which verifies that an incoming file or program is actually coming from a trusted address.

ALL ILLUSTRATIONS BY SLIM FILMS

2 Digital Certificates

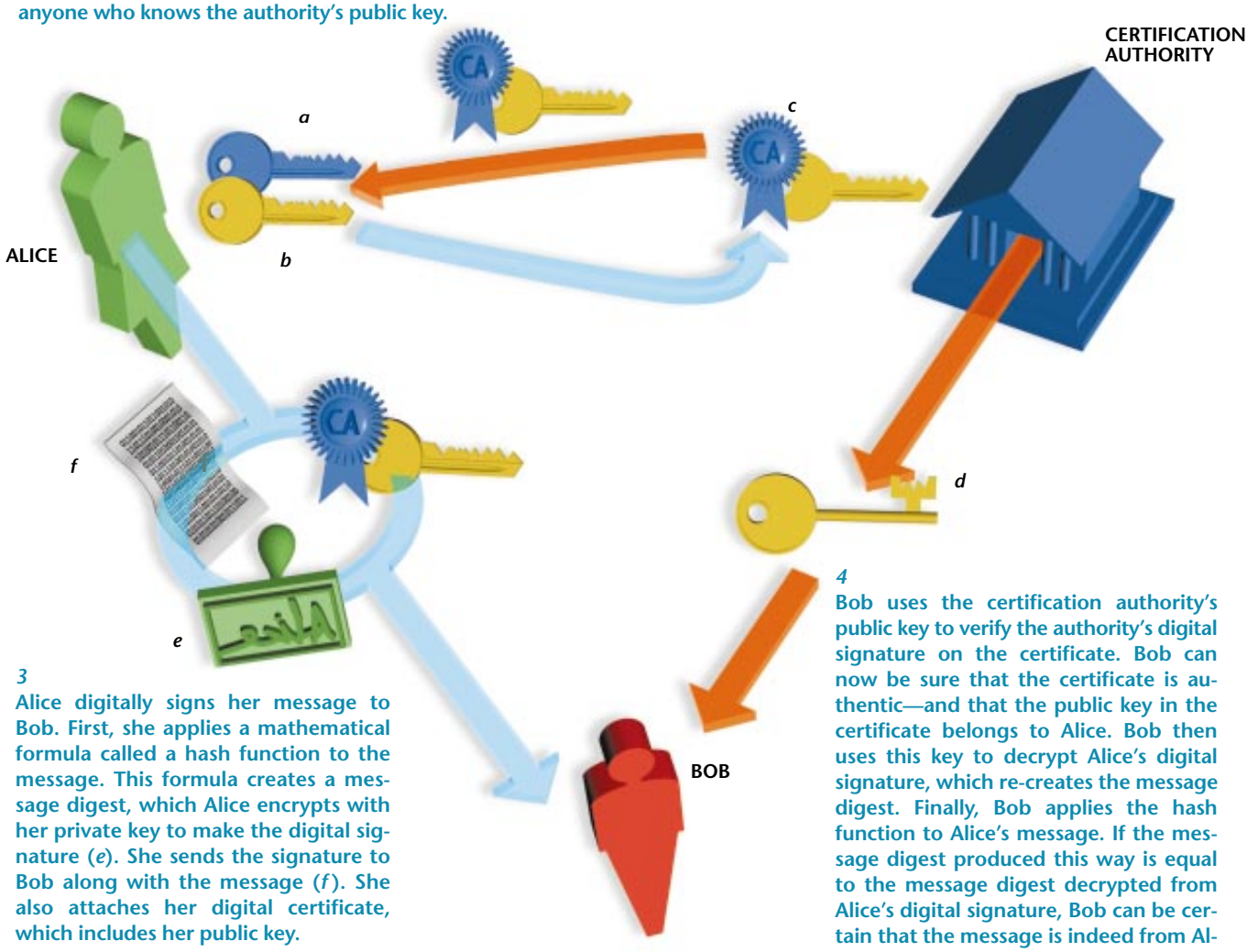
by Warwick Ford, *VeriSign*

Digital certificates play an essential role in public-key cryptography, a method widely used on the Internet to keep communications secure. To send and receive messages with this method, a computer user must have a pair of cryptographic keys—a private key and a public key—which are long strings of data, usually containing 500 to 1,000 bits. The user keeps the private key in a safe place—encrypted on a computer's hard drive, for example—but makes the public key known to the people with whom he or she wants to communicate.

Let's say that Alice wants to send a message to Bob. Because she wants Bob to be sure that the message is really coming from her, Alice uses her private key to create a digital signature, which accompanies the message. Bob uses Alice's public key to verify the signature. But how can Bob be sure that the public key actually belongs to Alice? An impostor could create her own key pair and send the public key to Bob, claiming that it belongs to Alice. To prevent such a possibility, Alice must obtain a digital certificate, a data item issued by a widely trusted certification authority, such as VeriSign or GTE CyberTrust or an authority set up by Alice's company. The digital certificate can be thought of as the cyberspace equivalent of a driver's license. It confirms that a particular public key belongs to a particular person or entity.

1
Alice uses cryptographic software to generate a private key (*a*) and a public key (*b*). She sends the public key to a certification authority and asks for a digital certificate. The authority needs to authenticate Alice's identity; depending on the type of certificate, this may involve verifying private information that Alice supplies. If her credentials check out, the authority issues a digital certificate (*c*) affirming that the public key belongs to Alice. Attached to the certificate is the authority's digital signature, which can be verified by anyone who knows the authority's public key.

2
The certification authority's public key (*d*) is distributed to anyone who needs it, including Bob. The key is typically embedded in Web browsers and other application software used for secure computer communications.



3
Alice digitally signs her message to Bob. First, she applies a mathematical formula called a hash function to the message. This formula creates a message digest, which Alice encrypts with her private key to make the digital signature (*e*). She sends the signature to Bob along with the message (*f*). She also attaches her digital certificate, which includes her public key.

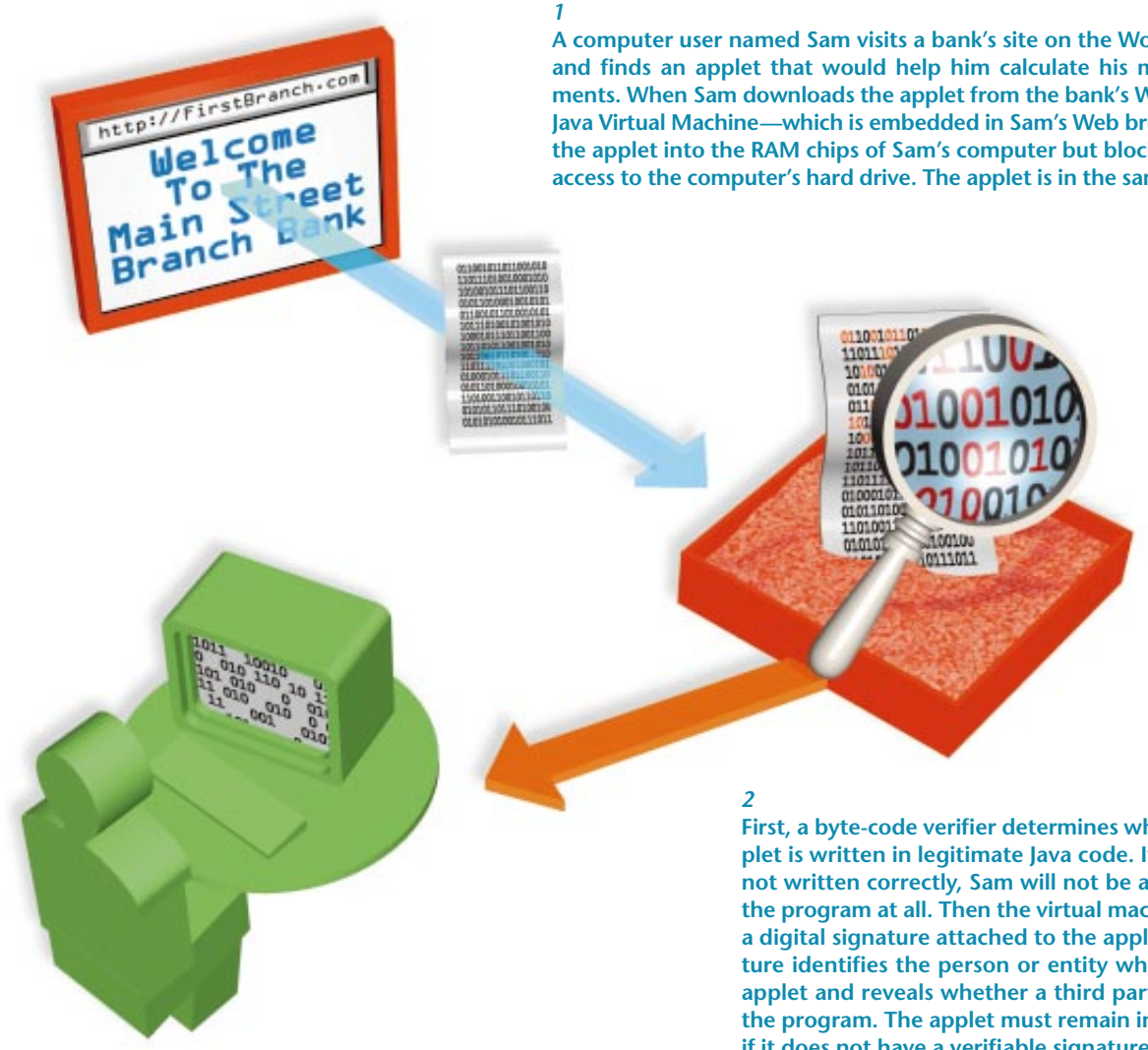
4
Bob uses the certification authority's public key to verify the authority's digital signature on the certificate. Bob can now be sure that the certificate is authentic—and that the public key in the certificate belongs to Alice. Bob then uses this key to decrypt Alice's digital signature, which re-creates the message digest. Finally, Bob applies the hash function to Alice's message. If the message digest produced this way is equal to the message digest decrypted from Alice's digital signature, Bob can be certain that the message is indeed from Alice and that it has not been tampered with by anyone else on the network.

3 The Java Sandbox

by James Gosling, *Sun Microsystems*

The Java programming language can be used by software developers to write small applications—called applets—that can be downloaded from the Internet or other computer networks. The danger is that an unscrupulous person might create an applet that would tamper with a user's computer system by erasing files, stealing data or passing along a computer virus. But the Java language has been designed to prevent such breaches.

The key to Java's security is a layer of software called the Java Virtual Machine, which is needed to execute any applet written in the programming language. When a computer user downloads an applet, the virtual machine initially prevents the program from gaining access to the computer's hard drive, network connections and other vital system resources. At this stage the applet can be imagined as sitting in a child's sandbox, a place where it can do no damage. An applet can get out of the sandbox only if the virtual machine verifies that the program comes from a trusted source.



1
A computer user named Sam visits a bank's site on the World Wide Web and finds an applet that would help him calculate his mortgage payments. When Sam downloads the applet from the bank's Web server, the Java Virtual Machine—which is embedded in Sam's Web browser—allows the applet into the RAM chips of Sam's computer but blocks the applet's access to the computer's hard drive. The applet is in the sandbox.

2
First, a byte-code verifier determines whether the applet is written in legitimate Java code. If the applet is not written correctly, Sam will not be allowed to use the program at all. Then the virtual machine looks for a digital signature attached to the applet. The signature identifies the person or entity who created the applet and reveals whether a third party has altered the program. The applet must remain in the sandbox if it does not have a verifiable signature. Sam can use the program to perform calculations but not to read or write files on his hard drive.

3
If the applet's signature is verified, the virtual machine determines how much access to give the program. Sam can adjust the amount of access according to his security needs. For example, he can allow the applet to read any file on his hard drive or bar it from parts of the drive that contain confidential data.

Cryptography for the Internet

E-mail and other information sent electronically are like digital postcards—they afford little privacy. Well-designed cryptography systems can ensure the secrecy of such transmissions

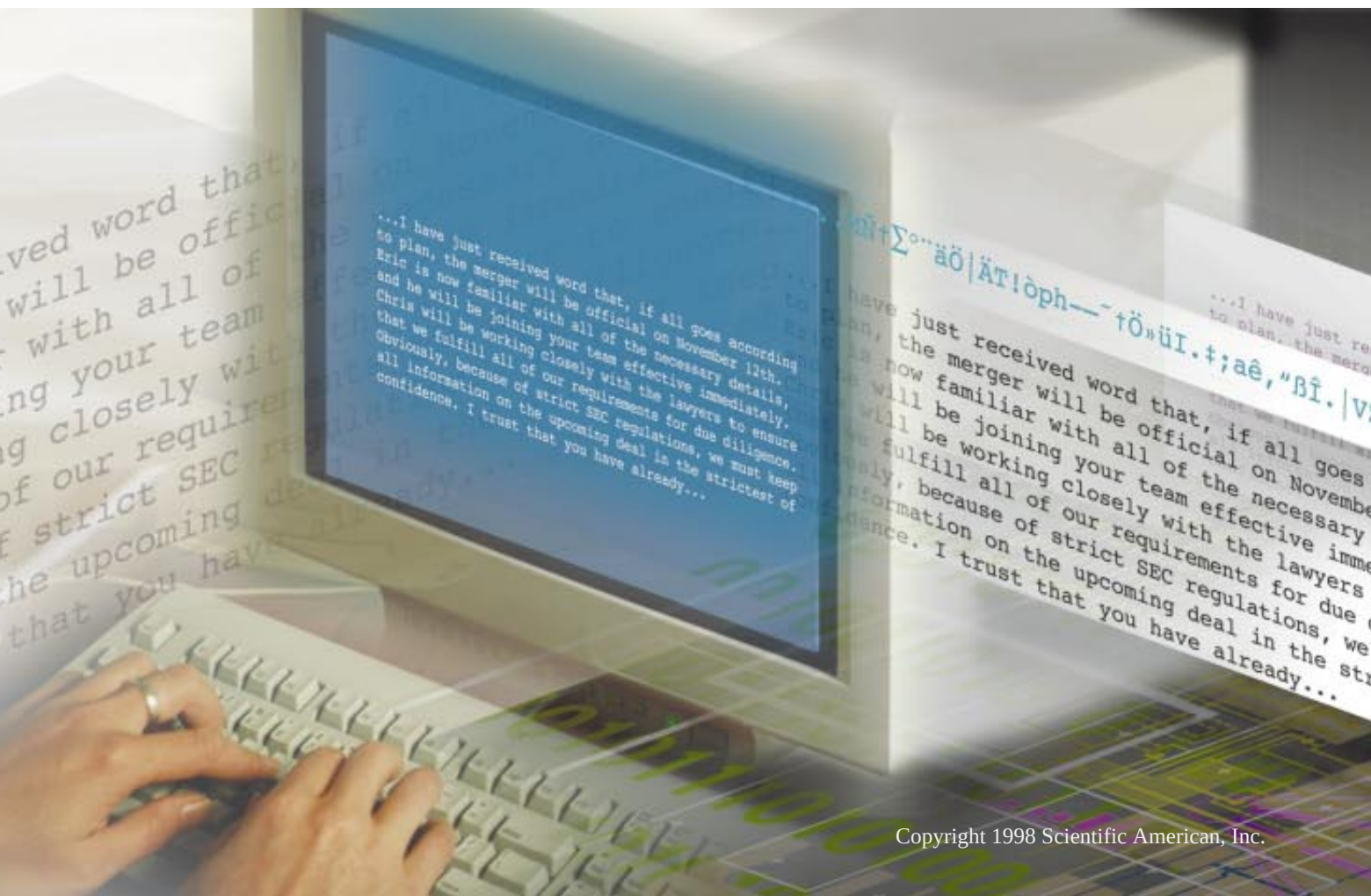
by Philip R. Zimmermann

Sending letters through the post office might take days, but at least the correspondence is guaranteed some degree of privacy. E-mail delivered over the Internet, on the other hand, can be blindingly fast but is highly susceptible to electronic eavesdroppers. One way to increase the privacy of such transmissions is to encrypt them, scrambling the information in complex ways to render it unintelligible to anyone but the intended recipient.

Since the 1980s the development of sophisticated algorithms and fast but affordable computer hardware have made powerful, military-grade cryptographic systems available to millions of people with ordinary personal computers. Recent technological improvements promise to make such systems increasingly resistant to even the most advanced cipher-cracking techniques.

Out from the Shadows

Four decades ago the Pentagon's requirements for tiny custom circuits to fit into missiles and spacecraft were the driving force behind the U.S. electronics industry. Today civilian demands dominate, and the military currently



satisfies most of its needs with off-the-shelf products designed for the much larger consumer market. The same thing is happening with cryptography.

Until the mid-1970s the National Security Agency (NSA) had a virtual monopoly on U.S. encryption technology, a field that was kept shrouded in secrecy. Then, in 1976, the seminal article "New Directions in Cryptography," in which Whitfield Diffie and Martin E. Hellman of Stanford University first described "public-key cryptography" in the open literature, forever changed the landscape. In the years since that publication, an energetic cryptographic community in academia and industry has emerged, publishing an ever increasing number of papers and building a mature discipline. The growing popularity of the Internet—and people's concerns about the privacy of that medium—has only intensified the trend. Today some of the best ciphers and systems are being developed by cryptographers at universities and in the private sector all over the world. In fact, the NSA is now beginning to buy commercial products for a portion of its cryptographic needs.

Why was Diffie and Hellman's introduction of public-key cryptography so crucial? In conventional cryptosystems, a single key is used for both encryption and decryption. Such systems, called symmetric, require the key to be transmitted over a secure channel—a process that is often inconvenient. After all, if a secure channel exists, why is encryption needed in the first place? This limitation hobbled cryptography.

Diffie and Hellman removed that constraint. Public-key cryptography allows the participants to communicate without requiring a secret means of delivering the keys. Such asymmetric systems rely on a pair of keys that are different but complementary. Each key unlocks the message that the other key encrypts, but the process is not reversible: the key used to encrypt a message cannot be used to decrypt it. Thus, one of the complementary keys (public) can be disseminated widely, whereas the other key (private) is held only by its owner. When Bob wants to send a message to Alice, he can use her public key to encrypt the information, which she will then use her private key to decrypt.

Public-key cryptosystems are based on mathematical problems that are easy to compute in one direction but painfully slow to solve in the reverse. The two main public-key algorithms are the Diffie-Hellman (and its variants, such as the Digital Signature Standard from the National Institute of Standards and Technology, ElGamal and elliptic curve approaches) and RSA, developed at the Massachusetts Institute of Technology by computer scientists Ronald L. Rivest, Adi Shamir and Leonard M. Adleman.

ENCRYPTING A PRIVATE MESSAGE that Bob will send to Alice over the Internet requires several steps. In this conceptual schematic, Bob first computes a hash of the text [*see diagram on page 113*]. He then encrypts the hash using his private key [*see box on next page*]. The resulting information (*blue, below*) serves as Bob's "signature." Bob compresses the signature and his message electronically (*purple*) and enciphers the file (*green*) using a particular session key. Bob encrypts this key using Alice's public key, and the result (*orange*) is added to the message. Finally, the file is converted into alphanumeric characters (*red*) for transmission over the Internet. At the receiving end, the steps are essentially reversed, with Alice using her private key to decrypt the session key, which she can then use to decipher the rest of the message.

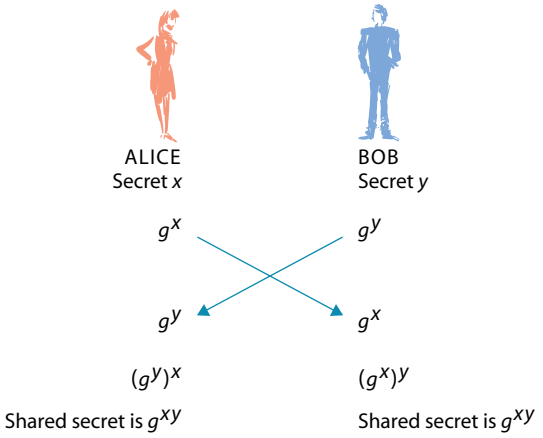
The former approach uses discrete logarithms. It is simple to compute g^x modulo p: just raise g to the x power, divide that quantity by a large prime number p, and then take the remainder of that operation. But given g, p and the value of g^x modulo p, it is infeasible to recover x [see "The Mathematics of Public-Key Cryptography," by Martin E. Hellman; SCIENTIFIC AMERICAN, August 1979].

The RSA system is based on the difficulty of factoring. It is straightforward to multiply two large prime numbers together, but it is extremely difficult to factor that huge composite back into its two primes [see "Mathematical Games," by Martin Gardner; SCIENTIFIC AMERICAN, August 1977].

Another beauty of public-key cryptography is that it can be used for message authentication: a recipient can verify the identity of the sender. When Bob transmits a message to Alice, he first encrypts it with his private key, then reencrypts the encrypted message with Alice's public key. Alice, after receiving the transmission, reverses the steps. She first decrypts the message with her own private key, then decrypts it again with Bob's public key. If the final text is legible, Alice can be confident that Bob actually wrote the message.

Of course, all this encrypting and decrypting requires myriad mathematical calculations. But software applications, such as PGP, running on PCs can automate the process. Using one of those packages, Bob and Alice need only press the "encrypt" and "decrypt" buttons on their computers, and the number crunching is performed behind the scenes.

For all its innovation, public-key cryptography has two severe limitations. First, because of its relatively slow speed, the technology is impractical for encrypting large messages. Second, and perhaps more important, public-key cryptography sometimes allows patterns in a message to survive the encryption process. The patterns are thus detectable in the enciphered text, making the technology vulnerable to cryptanalysis. (Cryptography is the science of making ciphers, cryptanalysis is the study of breaking them, and cryptology is both disciplines.)



Public-Key Cryptography

For centuries, cryptography was hampered by the so-called key-exchange problem. Specifically, if Bob wanted to send Alice an enciphered message, he also somehow had to transmit to her the secret encryption key that he had used. Public-key cryptosystems overcame this limitation by relying on clever mathematics.

In the Diffie-Hellman algorithm, which helped to spawn the field of public-key cryptography, Alice uses her secret number x to calculate g^x and sends that quantity to Bob. On his end, Bob uses his secret number y to compute g^y and sends that to Alice. (Note that the value of g is publicly known.) After Alice receives this information, she can then compute $(g^y)^x$, which is equal to $(g^x)^y$, the value that Bob calculates. This quantity can become their shared, secret encryption key.

But someone who has intercepted Alice's g^x and Bob's g^y would be able to derive the secret x and y. So to thwart any eavesdroppers, Alice and Bob insert the modulo function, which calls for the remainder from a division operation. (For example, 14 modulo 4 = 2 because the remainder of 14 divided by 4 is 2.) This added twist ensures secrecy—instead of sending g^x to Bob, Alice transmits the value of g^x modulo p, from which eavesdroppers would have great difficulty in recovering x, even if they know g and p.

With additional mathematics, the Diffie-Hellman algorithm has evolved into cryptosystems that generate two complementary keys, one private (for Alice, x) and the other public (consisting of g, p and the value of g^x modulo p). Ingeniously, the private key deciphers the message that was enciphered by the public key, but the key used to encrypt a message cannot be used to decrypt it. Thus, Bob can use Alice's public key (which she has disseminated to everyone) to encrypt a message to her, which she—and only she—can decrypt using her private key. —*P.R.Z.*

Symmetric Workhorses

Consequently, the bulk of encryption is usually performed by faster and more secure symmetric ciphers, with public-key cryptography limited to the small—but essential—function of exchanging the symmetric keys. Specifically, Bob encrypts his message with a quick and strong symmetric cipher. He then needs to send Alice the symmetric key that he used, so he enciphers it with her public key and attaches the result to his encrypted message. Alice will decrypt the symmetric key with her private key so that she can use that information to decrypt the rest of Bob's message.

For authentication, Bob again does not use public-key cryptography to "sign" his transmission directly. Instead he computes a hash, or fingerprint, of his message. Such mathematical procedures can be used to condense an input of any size into a digest of fixed length, typically 160 bits long. (A bit is the most basic unit of computer data. It stores one of two possible states, represented by 0 or 1.) Cryptographically strong hash functions, such as SHA-1, RIPEMD-160 and MD5, are designed so that a forger would find it computationally infeasible to devise a different message that would yield the same hash. In other words, the fingerprints generated are virtually unique: two different messages will almost certainly yield distinct digests.

After computing a hash of his message, Bob encrypts that information with his private key. He sends this "signature" with the rest of his encrypted transmission. Alice receives the encrypted hash and decrypts it with Bob's public key. She can then compare the result with the hash she computes herself after decrypting the message. A match proves both that the transmission has not been tampered with and that Bob is the sender.

For encrypting such information to be sent over the Internet, the most common method is to break the data into fixed-size blocks, each usually 64 or 128 bits long, so that the encryption can be performed a chunk at a time. So-

ILLUSTRATIONS BY DMITRY KRASNY

called block ciphers usually encrypt each chunk using multiple rounds (the exact number is dictated by the particular algorithm) of mathematical operations, with the output of one iteration fed as input to the next. Each round often involves both permutation (shuffling "xtv" to "tvx") and substitution (changing "tvx" to "cb2"). A section of the key helps to transform the data during the iterations.

Feeding identical chunks of text to a block cipher will lead to encrypted blocks that are identical to each other. To suppress any such block-aligned patterns from forming (which would make the cipher easier to crack), block algorithms typically use some kind of chaining. Blocks that have already been encrypted are looped back to help encrypt subsequent chunks. In effect, the encryption of a block of text depends on *all* the previous blocks.

Block ciphers have symmetric keys that are usually 56, 128 or 256 bits long. Well-known examples are the Data Encryption Standard (DES), triple-DES, CAST, IDEA and Skipjack. The workhorses of cryptography, block algorithms have become the focus of much recent research.

The Key Is the Key

The most sensitive operation in cryptography is the generation of keys. For a system to be as secure as possible, the keys should be numbers that are truly random, unpredictable by an attacker. Such numbers are different from the deterministic pseudorandom sequences that computers generate algorithmically for games and simulations. Truly random numbers can be derived only from the environmental "noise" of the physical world, such as the process of radioactive decay.

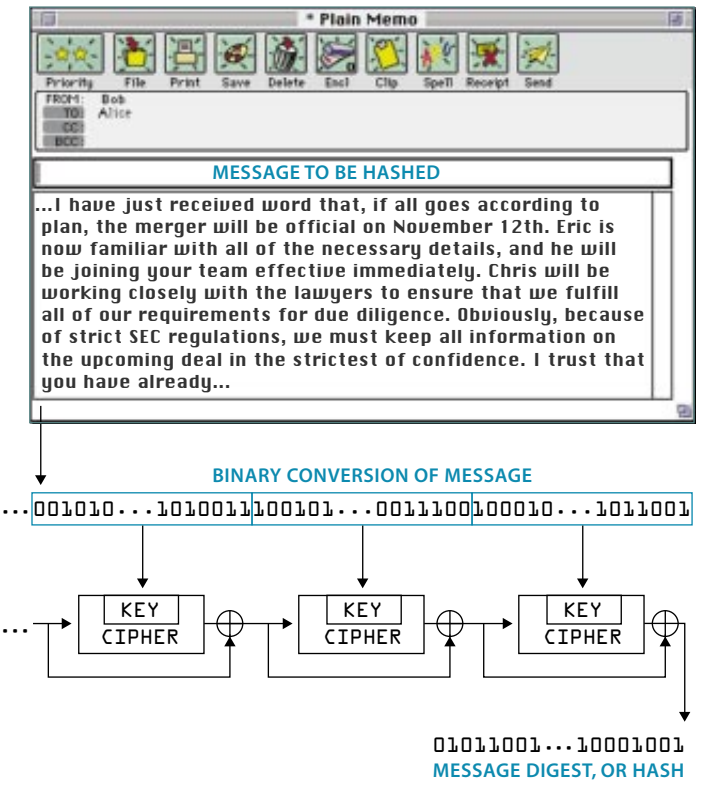
Such high-quality randomness is difficult to generate in a computer. One method is to measure the time, in microseconds, between each human-supplied keystroke, which is impossible to predict. Data gathered in this way are not quite random enough for generating keys directly, but the information can be passed through a hash function to distill the disorder.

Interestingly, the only cipher that cryptologists have ever proved to be perfectly secure is the one-time pad (OTP), in which the key is as long as the message itself. In an OTP, a random sequence is used to encipher a message bit for bit—that is, the 34th bit of the key is used to alter the 34th bit of the message. The key must be truly random. It cannot be a pseudorandom sequence produced by a deterministic algorithm; otherwise the cipher may be crackable. OTPs are rarely used because of their impracticality: the key must be as long as the message, and it must be sent to the receiver over a secure channel. Moreover, it can be used only once, or an attacker could break the messages.

Although many people think key size is the determining factor in cryptographic strength, an equally important criterion is the quality of the cipher's design. Consider a simple substitution cipher in which all *A*s are changed to *W*s, all *B*s turned into *K*s, all *C*s transformed to *Q*s, and so on. The number of different ways to rearrange the alphabet is given by 26 factorial (that is, $26 \times 25 \times 24 \times \ldots 3 \times 2 \times 1$). That quantity is roughly equivalent to 2^{88}, a "key space" of different combinations that is regarded as fairly respectable, requiring enormous computing resources to break if every possible key must be tried. Yet when I was a kid I would crack this type of cryptogram all the time with no more than a pencil and paper. I simply looked for the most common letter and assumed it was probably *E* and then found the second most common letter and assigned *T* to it, and so on. Clearly, despite its vast key space, this type of cipher is very weak.

For a well-designed cryptography system, though, the key size does relate directly to the effort required to crack it. For block ciphers, the relation is usually exponential. Adding just one bit to the key length doubles the work the attacker must do to try all the keys. And doubling the key size squares the amount of effort. On average, a 128-bit key requires about 2^{127} (in decimal, 1.7×10^{38}) operations to break.

Public-key algorithms are less sensitive. Typically, they have subexponential but superpolynomial key spaces, which means that doubling the length of the key increases the work substantially, but the amount is less than a squaring of the work effort. To use RSA as an example, modern factoring algorithms can do much better than simply trying all the possible smaller prime numbers to factor a large composite. Diffie-Hellman is also subexponential. For com-



HASH ALGORITHM condenses a message into a digest, a digital fingerprint that can be used to detect forgeries. The text of the message is first converted into binary form. (The letter *A* might be represented by 00000, the letter *B* by 00001, the letter *C* by 00010, and so on.) The resulting string of 0s and 1s is then separated into equal-size blocks. Next, the chunks are fed in sequence as key material into a cipher. The final output is the digest, or hash, of the original message. Note that a message of any length will always yield a digest of fixed size. The operation is called "one way" because it is virtually impossible to recover a message from its hash. Also, the algorithm is designed so that any given two messages will almost certainly yield distinct hashes, and it is computationally infeasible to find another message that produces the same hash as a given message. Thus, a digest can serve as a "fingerprint" for its corresponding message.

parison's sake, a 3,000-bit RSA or Diffie-Hellman key requires about the same amount of work to crack as a 128-bit key for a block cipher.

Still, block ciphers are hardly invincible. This year a special-purpose massively parallel machine built for less than $250,000 by the Electronic Frontier Foundation, headquartered in San Francisco, broke a DES message by exhausting its 56-bit key space in less than a week.

Brute force is not the only way to crack a cipher. Cryptanalysts can apply powerful mathematical and statistical tools to find any shortcuts, perhaps by uncovering patterns in the encrypted text. Attempts to break ciphers can be grouped into three categories, depending on how much is known about the original message (called plaintext) and the corresponding enciphered transmission (called ciphertext).

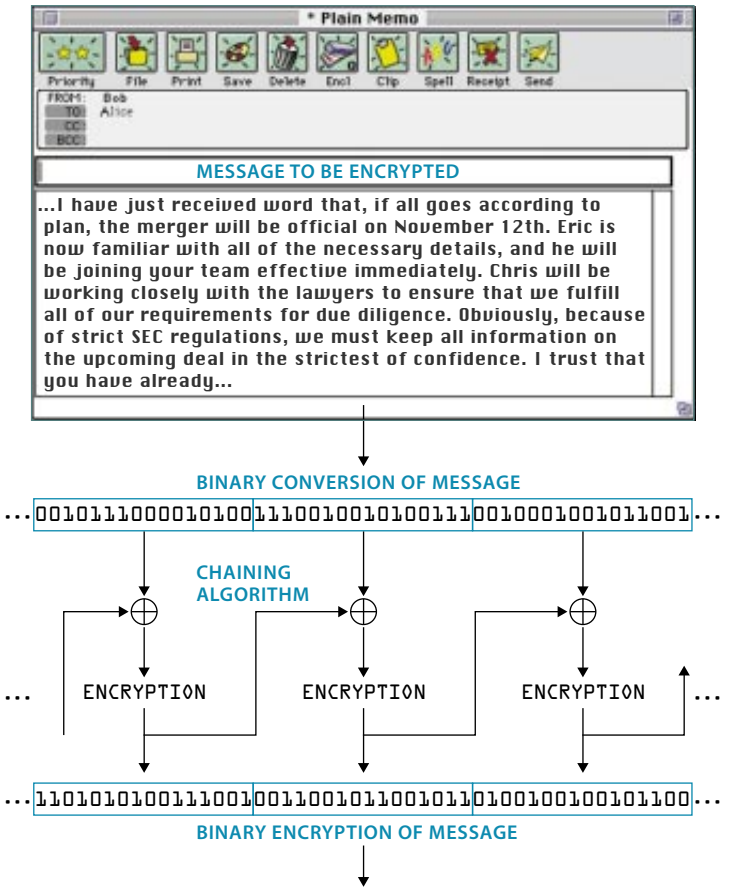
In some cases, all that the attackers have to work with is the ciphertext, so they have little to guide their efforts in guessing the key. Even a poorly designed cipher might be able to withstand such ciphertext-only attacks.

But if the attackers know at least a part of the message—for instance, that the text begins with "Dear Mr. Jones"—the opportunities for success increase significantly. At a minimum, they can try different keys until they find one that decrypts the "Dear Mr. Jones" part of the plaintext. Even if the attacker knows only the language (Russian or French or COBOL) of the plaintext, that information can be exploited. If the message is in English, for example, the most common word is probably "the." To thwart such known-plaintext attacks, some cryptography systems electronically compress the message, squeezing out easily predictable patterns in the plaintext, before encrypting it.

Often an attacker knows much more. If a person steals a "smart" card containing crypto hardware, the thief can present perhaps billions of carefully chosen messages to the card and study the ciphertext output. Such chosen-plaintext attacks will crack a poorly designed cipher easily. Another example is public-key systems. An attacker can write a message, encrypt it with the public key (which is, after all, public) and then analyze the resulting ciphertext.

Two very effective methods of cryptanalysis, differential and linear, have recently been developed. Both approaches have been used to crack a number of well-known block ciphers and to show that DES can be broken hundreds or thousands of times faster than by key exhaustion.

In differential cryptanalysis, introduced by Shamir and Eli Biham of Technion Israel Institute of Technology, many pairs of plaintext messages with carefully chosen differences are encrypted to find a corresponding pair of ciphertexts that have a certain dissimilarity. When such a pair is found, it reveals information about the key. Linear cryptanalysis, developed by Mitsuru Matsui of Mitsubishi Electric Corporation, searches for correlations between plaintext, ciphertext and key that are true slightly more often than not. The method then gathers statistics on large numbers of known plaintext-ciphertext pairs, looking for biases that will disclose clues about the key.



...hkdWDecU2xh/+q5NZjCLapEcdqZermz/8bErhfzqI92ECID4iHRf
pvR0+j3RZPwXIIEjLciMqFMZyExEeHSh8R030QukMED/CPYYME4d
Nj+EgKcp2zhUP8nwhp00iE1TncYOo8Cw3U5zRm9X5s2oF7G60gn
NWJkid8nUDi7OApRQiJZdOY64FgHnz6X/oXFHDf6wHypNOGde6e8
UCJcU/W4+MSaHYh+dT5PBajn94BiaIOt/NdrIJ+A6Oq1nafEDj/I/o
G+cZhf+8ZpMy/w4cvy3VE3z0BqdOO8V6hItYtXz8y6httzACCZdPI
R75J6wyEGp2GQgYtIg/0o/MX7Q3kUJ5/GST2NkM+cNybm8kArJN
+iSMJDCzO1nYiEHeawaEhqEz4ASI2xGBoxs0LMSEGuYeZRXZkWZa
Od3QCMte5vqji5bYLFbQU3+wc3TIth1B+hcP8jE6i2BMHOkengROp
UsCkMTq6+EHCXjEqUI+8Naik/9hePz45vx//2WPJXyj4cxfWz...

CHAINING ALGORITHM increases the security of block ciphers. A message is converted into a string of 0s and 1s, and the long sequence is then broken into blocks of equal size. Before each of these chunks is encrypted, it is first mathematically combined with the enciphered previous block. Thus, the encryption of the 23rd chunk depends on the enciphered 22nd block, which itself was affected by the encryption of the 21st block, and so on. Because of this feedback chain, an encrypted block depends on all the previous blocks, making the cipher more difficult for cryptanalysts to crack.

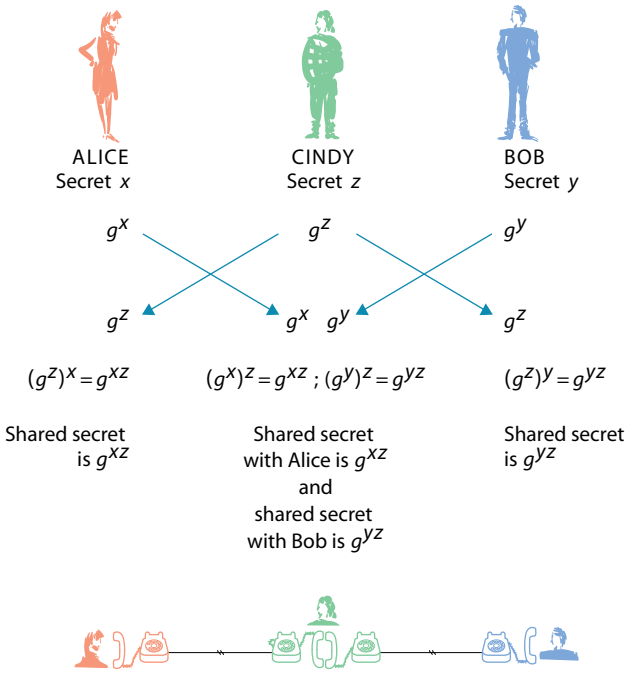
Beware of Middlemen

Though powerful, cryptanalysis techniques usually require a backbreaking number of computations. Often instead of trying to crack a cipher, it is easier to attack the protocol, or implementation, of that cipher.

One potential threat is man-in-the-middle attacks, which are the biggest vulnerability of public-key cryptosystems. When Bob wants to send a message to Alice, he may be unaware that Cindy is attempting to impersonate Alice. If Cindy can trick Bob into using her public key instead of Alice's, she will be able to decrypt Bob's message.

The only way to prevent this type of attack is for Bob to confirm somehow that Alice's public key is really Alice's. Most of the complexity of well-designed implementations of public-key cryptosystems is devoted to this one particular vulnerability. One solution is to have a trusted third party verify and sign the keys. This approach, however, begs the

ILLUSTRATIONS BY DMITRY KRASNY



MIDDLEMAN ATTACK is the greatest vulnerability of public-key cryptosystems. If Cindy, an eavesdropper, can intercept transmissions between Alice and Bob, she can trick Bob into using her g^z instead of Alice's g^x and similarly deceive Alice into using g^z instead of Bob's g^y [*see box on page 112*]. Cindy would then be able to decrypt and reencrypt Alice's and Bob's messages to each other—all unbeknownst to the couple. The process is analogous to Alice and Bob talking on special, encrypted telephones while Cindy listens in by using a pair of such phones to decrypt, then reencrypt, the transmission.

major—and politically controversial—question: Should the keys be certified in a top-down manner by government authorities or in a decentralized grassroots method by different entities, including various private companies and individuals, allowing people to choose for themselves which key signers to trust? In fact, this issue is so crucial that I could have written this entire article on it.

As cipher-breaking techniques have improved, so have the algorithms for stronger cryptography. Recently the National Institute of Standards and Technology solicited designs for the Advanced Encryption Standard (AES), a new block cipher to replace the DES, which has reached the end of its useful life, mainly because of its short 56-bit key and 64-bit block size. The AES, which has been generating considerable excitement in the cryptography field, will use a key size of 128, 192 or 256 bits to encrypt data in 128-bit blocks.

Good AES designs will meet several criteria. They will offer flexibility in various key and block sizes; they will be efficient in setting up keys and in encrypting and decrypting, particularly when implemented on 32-bit processors as well as on eight-bit microprocessors, such as in "smart" cards, and on other hardware; and they will perform well in a wide range of applications, from satellite communications to high-definition television.

Several of the AES candidates appear to be extremely well designed. The better proposals have capitalized on the experience of cryptographers who have studied block ciphers for the past 20 years, including their knowledge of how to defend against linear and differential cryptanalyses.

Of the 15 submissions, I believe more than a few would make credible encryption standards. MARS, which draws on the experience of IBM's original DES team, uses two very different structures for the encryption rounds. The mixed approach, the IBM cryptographers claim, will result in better security than that achieved with a homogeneous cipher. CAST-256 extends the earlier CAST architecture to a 256-bit key and 128-bit block size. Twofish is more mathematically rigorous than its predecessor, Blowfish. Serpent deploys an unusual parallel design to make it as fast as DES, with a short time for key setup, which should enable the cipher to be used efficiently as a hash function.

Deciphering the Future

Whichever candidate is selected, the AES promises to tip the balance further in favor of cryptographers in their ongoing arms race against cryptanalysts. Today the very best cryptosystems are beyond the reach of the best cryptanalytic methods known. Still, it is conceivable that powerful, new cipher-breaking techniques will be developed in the coming years. Even so, many cryptologists contend that the gap between cipher makers and cipher breakers will only widen.

I agree with that assertion, in part because of the active community of cryptographers in academia and the private sector, which has grown and matured to reach parity with military expertise in the field. Evidence of this was supplied by the recent declassification of the Skipjack cipher, which the NSA had developed in secrecy for the Clipper chip. A review by Technion's Biham, an academic cryptologist, revealed the algorithm to be less conservative, with a smaller margin of safety, than the best designs from academia. It appears that cryptography—like the Internet itself—has stepped from the dark shadows of the military into the bright sunshine of the free market. SA

The Author

PHILIP R. ZIMMERMANN is the author of Pretty Good Privacy (PGP) encryption software, for which he received the Chrysler Award for Innovation in Design. He is a software engineer with more than 20 years of experience in cryptography, data communications and real-time embedded systems. He has testified before Congress, urging the U.S. government to loosen its control of encryption technology. Selected in 1995 by *Time* magazine as one of the 50 most influential people on the Internet, Zimmermann is currently a senior fellow with Network Associates.

Further Reading

THE OFFICIAL PGP USER'S GUIDE. Philip R. Zimmermann. MIT Press, 1995.

APPLIED CRYPTOGRAPHY. Second edition. Bruce Schneier. John Wiley & Sons, 1996.

Additional information can be found at http://www.pgpi.com, http://www.pgpinternational.com, http://www.pgp.com/phil, http://csrc.nist.gov/encryption/, http://www.epic.org, http://www.eff.org and http://www.cdt.org on the World Wide Web.

The Case against Regulating Encryption Technology

One of the pioneers of computer security says the U.S. government should keep its hands off cryptography

by Ronald L. Rivest

The widespread use of cryptography is a necessary consequence of the information revolution. With the coming of electronic communications on computer networks, people need a way to ensure that conversations and transactions remain confidential. Cryptography provides a solution to this problem, but it has spawned a heated policy debate. U.S. government agencies want to restrict the use of data encryption because they fear that criminals and spies may use the technology to their own advantage.

Before the 1970s, cryptography was too complicated and too expensive for everyday use. Two inventions changed this picture dramatically: public-key cryptography and the microprocessor. The idea of using public and private encryption keys—first proposed in 1976 by electrical engineers and computer scientists Whitfield Diffie, Martin E. Hellman and Ralph C. Merkle—paved the way for the general use of strong cryptography, which scrambles messages so effectively that it would take many years of computer time to break the code. And the growing availability of fast microprocessors gave more and more computer users the ability to make the calculations necessary for this kind of encryption.

As strong cryptography became easily accessible in the late 1980s and early 1990s, two government agencies grew concerned about its widespread deployment. The National Security Agency (NSA), which monitors electronic communications around the globe, worried that it would be unable to decipher the encrypted messages of potential spies and terrorists. Similarly, the Federal Bureau of Investigation feared that criminals in the U.S. would use the encryption software to thwart surveillance of their voice or data communications. Over the past decade these agencies have pushed for government regulation of encryption technology and have favored the continuation of current restrictions on the export of strong encryption software.

The government's concern is that the "bad guys" will benefit from the new cryptographic technology. This is certainly possible—the sun shines on the evil as well as the good. But it is poor policy to clamp down indiscriminately on a technology merely because some criminals might be able to use it to their advantage. For example, any U.S. citizen can freely buy a pair of gloves, even though a burglar might use them to ransack a house without leaving fingerprints.

I rather like the glove analogy; let me expand on it a bit. Cryptography is a data-protection technology just as gloves are a hand-protection technology. Cryptography protects data from hackers, corporate spies and con artists, whereas gloves protect hands from cuts, scrapes, heat, cold and infection. The former can frustrate FBI wiretapping, and the latter can thwart FBI fingerprint analysis. Cryptography and gloves are both dirt-cheap and widely available. In fact, you can download good cryptographic software from the Internet for less than the price of a good pair of gloves.

Should the use of cryptography be restricted to satisfy the concerns of the NSA and the FBI? It is true that these two agencies may find their jobs more difficult as cryptographic technology spreads. But we should also consider cryptography's benefits to society as a whole. Most people use cryptography to prevent crime rather than to hide it, just as most people wear gloves to protect their hands rather than to hide their fingerprints. By ensuring the confidentiality and authenticity of electronic banking and Internet commerce, cryptography prevents theft and credit-card fraud. The vigorous application of cryptography may also improve national security: the encryption of communications, for example, protects U.S. businesses from industrial espionage. Paradoxically, we may create a safer society by promoting a technology that somewhat hampers law enforcement.

Some have hoped for compromise solutions that would allow strong cryptography to be widely used while still enabling the NSA and the FBI to decrypt messages when lawfully authorized to do so. For example, there have been key-escrow proposals that would require users to register their software encryption keys with law-enforcement agencies, and key-recovery proposals that would give government agencies backdoor access to the keys. In a typical key-recovery scheme, an encrypted version of the message encryption key is sent along with each message. An FBI-authorized key-recovery center can use a master backdoor key to decrypt the message key, which is then used to decrypt the message itself.

In my opinion, these systems would satisfy no one. They are very easy to circumvent: spies and criminals could modify the encryption software to disable the key-recovery features, or they could simply download alternative software

from the Internet. Key recovery would be very expensive, too. Someone would have to pay for creating, staffing and maintaining the key-recovery centers. But the most subtle and serious cost in the long run would be the erosion of confidence in the government resulting from an increased sense of "Big Brotherism." To get an idea of the intrusiveness and impracticality of key recovery, imagine that whenever you bought a pair of gloves you were legally required to sew latex copies of your fingerprints onto the gloves' fingertips!

SLIM FILMS

Key-recovery systems would also create substantial security risks. The system's most serious flaw is that the same back doors used by the FBI to decipher encrypted messages would become targets for criminals, hackers, spies and even disgruntled employees of the FBI itself. If criminals or hackers managed to penetrate a key-recovery center and steal a master backdoor encryption key, they would be able to decrypt Internet communications at will. Millions of corporate, personal and government secrets would suddenly become vulnerable to theft and tampering.

In 1993 Congress asked the National Research Council to study U.S. cryptographic policy. The council then convened a blue-ribbon committee of 16 members. Its superb 1996 report, the result of two years' work, offered the following conclusions and recommendations:

- "On balance, the advantages of more widespread use of cryptography outweigh the disadvantages."
- "No law should bar the manufacture, sale or use of any form of encryption within the United States."
- "Export controls on cryptography should be progressively relaxed but not eliminated."

The committee members concluded that a ban on unregulated encryption would be "largely unenforceable." But the FBI and the NSA continue to push for key recovery and to oppose the relaxation of export controls unless key recovery is incorporated into the exported software.

Strong cryptography only gets easier to implement—and harder to regulate—over time. Professional societies are adopting public cryptographic standards that even a high school student can convert into programs. And new techniques such as "chaffing and winnowing"—which does not encrypt a message but achieves confidentiality by hiding pieces of the message in a welter of random data, or chaff—illustrate the enormous technical difficulties involved in trying to control cryptography.

The economic consequences of our current policy are also becoming clearer. A recent study conducted by the Economic Strategy Institute, a think tank in Washington, D.C., concluded that continuing the export controls on cryptographic products will cost the U.S. economy more than $35 billion over the next five years. My personal opinion is that the U.S. risks losing its leadership position in the software industry because of its restrictive export policy.

Finally, the ability to have private conversations is in my view an essential democratic right. Democracy depends on the ability of citizens to share their ideas freely, without fear of monitoring or reprisal; this principle should be upheld as much in cyberspace as it is in the real world. For the U.S. to restrict the right to use cryptography would be a setback for democracy—and a victory for Big Brother. SA

The Author

RONALD L. RIVEST is the co-inventor of RSA encryption, the most widely used public-key cryptosystem. He is Edwin S. Webster Professor of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology and a founder of RSA Data Security (a subsidiary of Security Dynamics Technologies).

Further Reading

CRYPTOGRAPHY'S ROLE IN SECURING THE INFORMATION SOCIETY. Edited by Kenneth W. Dam and Herbert S. Lin, National Research Council. National Academy Press, 1996. The report can be found at http://www.nap.edu/readingroom/books/crisis on the World Wide Web.

THE ELECTRONIC PRIVACY PAPERS: DOCUMENTS ON THE BATTLE FOR PRIVACY IN THE AGE OF SURVEILLANCE. Bruce Schneier and David Banisar. John Wiley & Sons, 1997.

PRIVACY ON THE LINE: THE POLITICS OF WIRETAPPING AND ENCRYPTION. Whitfield Diffie and Susan Landau. MIT Press, 1998.

by Shawn Carlson

Home Movies of an Invisible World

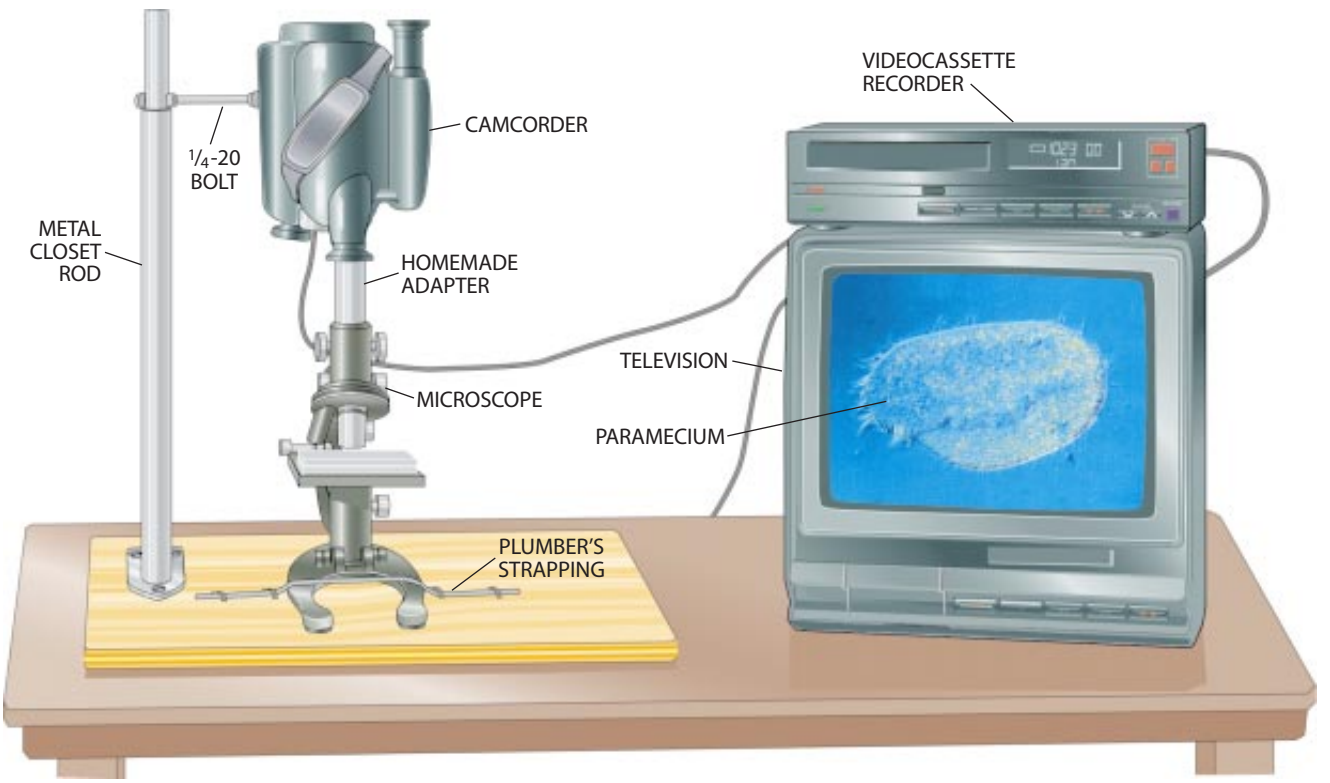
A few years ago I decided to explore the microscopic menagerie living in a bit of rainwater that had collected in an open barrel. It proved to be a rich find. As the infinitesimal neighborhood of a single droplet came into focus under my microscope, I discovered many organisms I had never seen before. One in particular intrigued me. At first it looked like a cylindrical creature with a moving gut. But after a few moments I realized that I was peering at a tiny, tubular home being constructed by an even more tiny and highly industrious architect. The little worker trundled back and forth along the tube, snatching floating bits of organic refuse, which it then used to extend its domicile. I watched transfixed, all the while wishing I had some way to document this activity. Still photographs would have been wholly inadequate; the situation clearly required a video recording. But I didn't have a way to attach my camcorder to my microscope and thus was unable to share the odd antics of this aquatic charmer with others.

So you can understand why I was thrilled to receive a package from Charles Carter, a talented amateur scientist in London, Ontario, addressing just this problem. One can, of course, buy a commercial video camera and adapter specially built for this task, but these units are expensive. Besides, using a home camcorder has certain advantages. For example, it makes it easy for you to include a running commentary about your procedures and observations. And with the software and hardware now widely available, you can easily capture individual video frames on your computer for additional analyses.

Carter's invention consists of two parts: an adapter that optically links the camcorder with the microscope and a stand that holds the camcorder in place. Both pieces can be made in an afternoon for very little money.

If you lack a microscope, worry not. Microscopes often turn up where secondhand items are sold. For instance, you might scour your local thrift stores and pawnshops for a bargain. Carter found a simple monocular instrument at a garage sale for $10. I bought a research-quality binocular microscope from a friend for $100. And some brand-new microscopes are well within a typical amateur's budget. Small instruments can be purchased at many shopping malls for less than I paid. You can also consult the forum hosted by the Society for Amateur Scientists on the World Wide Web (www.thesphere.com/SAS/WebX.cgi) and check the on-line swap meet there for bargain equipment.

Carter's adapter couldn't be simpler. It uses a hood that screws around the lens of the camcorder to shade it from glare. Many camcorders just recess the lens in the housing for shading, but they still

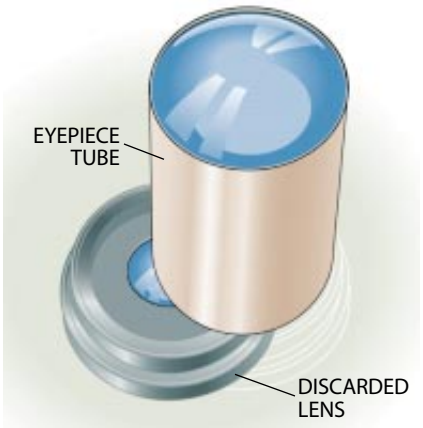


IAN WORPOLE

INEXPENSIVE VIDEO MICROSCOPE SYSTEM,
which allows several people to share the same view, can be built with a few dollars and an afternoon's work.

have threads for filters in front. So if your camcorder did not come with a separate lens hood, or if you don't want to sacrifice it for this project, you can probably find a lens hood of the proper diameter by rummaging through the junk box of your local camera shop. Alternatively, you can always buy a new one for a few dollars.

The eyepiece for most microscopes consists of two lenses situated at either end of a short metal tube. The top lens, the one you hold your eye near, is normally blocked off except for a hole at the center about the size of your pupil. This opening is too small for your camcorder to see through. The bottom lens, however, is larger and virtually unobstructed. And your camcorder will focus just fine with only this one lens in the eyepiece.



IAN WORPOLE

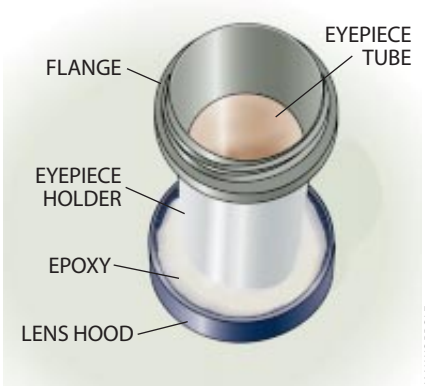
MODIFIED EYEPIECE contains only one lens.

Because eyepieces sometimes need to be cleaned, the top lens is normally designed to unscrew from the tube. The other lens of the eyepiece may be attached more permanently, usually recessed slightly from the bottom end of the tube. Unscrew the top lens and discard it. You will also need to remove the outer housing that holds the eyepiece in place by unscrewing it from the main body of the microscope. Turn the eyepiece upside down and insert it into the eyepiece holder so that the lens projects above the top of the holder by about 1/4 inch (about half a centimeter). Use a few drops of Krazy Glue to hold it in place.

Set the camcorder lens hood thread side down onto a wide strip of masking tape that is positioned sticky side up on a flat surface. Then place the eyepiece holder neck down in the center of the hood, pressing it firmly against the tape. Now mix a batch of epoxy and pour it between the neck and the lens hood. Take care not to allow any epoxy to ooze onto the threads of the lens hood. After the epoxy sets, lift the assembly and remove any tape sticking to the lens hood or covering the eyepiece.

Although you could now just screw the adapter to the microscope and the video camera to the adapter, that top-heavy arrangement would be quite precarious. Moreover, some high-end microscopes have their eyepieces canted to the side for the comfort of the viewer, and these instruments would not be able to support the weight of a camcorder attached at an angle. But Carter devised a sturdy stand that helps to hold his camcorder in position yet lets it freely slide up and down as he focuses the microscope. The stand functions best when the eyepiece is vertical, but it should also work in situations where the eyepiece must remain at a slight angle.

Carter built the base of his stand from a piece of scrap 3/4-inch plywood to



IAN WORPOLE

VIDEO ADAPTER
is made by attaching the microscope eyepiece holder to a lens hood using quick-setting epoxy.

which he attached an adjustable closet rod. First affix the end of the closet rod to the base. Then place your microscope on the base, screw on both the adapter and your camcorder and adjust the microscope's focusing knob so that the camcorder is as low as it can go. Take care and make sure things don't topple over at this point.

To attach the camcorder to the rod, Carter cleverly exploits the threaded sleeve on the base of the unit (where a tripod would normally screw in). While holding the microscope-camera combo upright with one hand, mark a line on the outer tube of the closet rod about one inch below the tripod mount. Then cut the outer tube off at that point with a pipe cutter or hacksaw. Next, cut the inner tube of the closet rod so that when it is inserted all the way in, it sticks out about four inches. If you have a stage-focusing microscope, you can dispense entirely with the inner rod and attach the camera directly to the primary support.

Next, drill a 1/4-inch hole into the inner rod so that the camera can be secured using a long 1/4-20 bolt, as shown in the illustration on page 118. Slip a washer on the bolt, push the bolt through the rod, and add another washer and two nuts. Now screw the bolt into the camcorder and tighten the nuts, one against the closet rod and the other against the camera.

Fix the microscope to the plywood stand by securing it with a length of plumber's strapping across the base. Now you're ready to spy on an invisible realm. To use your new apparatus, focus the microscope until a clear image appears on your monitor and then adjust the zoom on the camcorder until the image fills the screen.

Although this system will enable you to take leisurely safaris through microscopic jungles, it may be difficult to gauge the size of your minute prey. The best way to address this problem is to measure on your monitor the dimensions of something of known size. For example, you can view the tiny, gridded dots in the large, gray *A* at the beginning of this article; they are spaced at intervals of 169 microns. I invite professional and amateur microscopists to share other suggestions for calibrating this instrument with the Society for Amateur Scientists by joining the discussion on the society's Web site. SA

For more information about projects described in this column, check the forum conducted by the Society for Amateur Scientists on the World Wide Web (www.thesphere.com/SAS/WebX.cgi). You can also write the society at 4735 Clairemont Square, Suite 179, San Diego, CA 92117, or call 619-239-8807.

by Ian Stewart

Playing with Chocolate

Just because a game has simple rules doesn't necessarily mean there is a simple strategy for winning. Sometimes there is—ticktacktoe is a good example. But sometimes there isn't. For example, there is no winning strategy for dots and squares (also known as boxes), the childhood game in which players draw lines in a grid of dots and capture any squares they complete. I call games with winning strategies "dream games" and the others "nightmare games." Games with very similar rules can be surprisingly different when it comes to their dream or nightmare status. Nightmare games are often more interesting because you can play them without knowing in advance who ought to win. And in some nightmare games, you may know who ought to win—the player who takes the first turn, for example—but not know how.

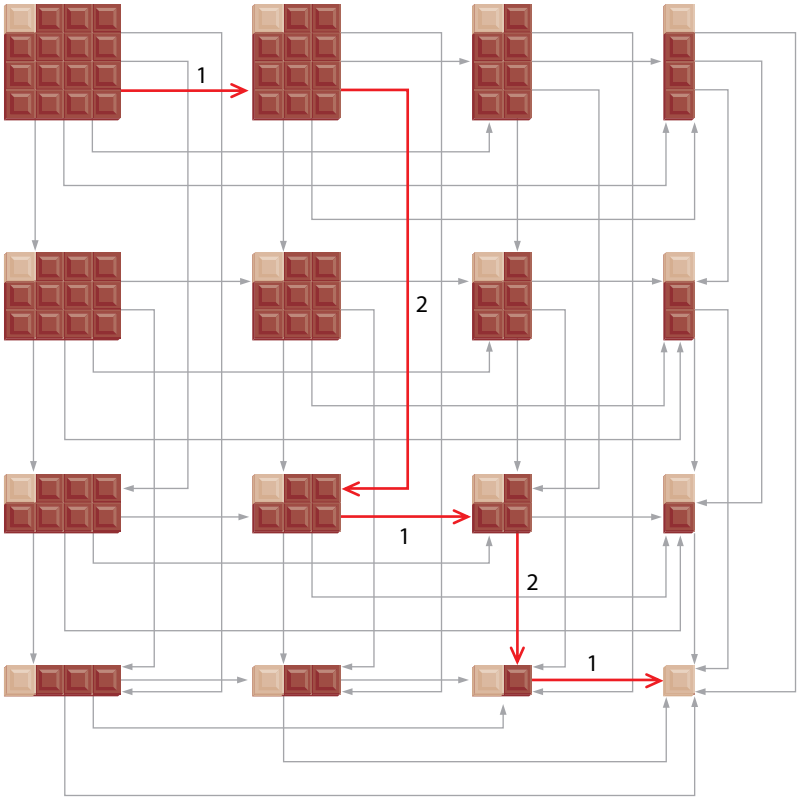
Let's consider two simple games involving chocolate bars. One, called Yucky Chocolate, is a dream game. The other, Chomp, has very similar rules, but it's a nightmare game: with optimal play the first player should always win, but nobody knows the winning strategy.

I have no idea who invented Yucky Chocolate. It was explained to me by Keith Austin, a British mathematician at the University of Sheffield. It takes place on an idealized chocolate bar: a rectangle divided into smaller squares. Two players—I'll name them Wun and Too after the order in which they play—take turns breaking off pieces of chocolate, which they must then eat. Each break must be a single straight line cutting all the way across the rectangle along the boundaries between the squares. The square in the upper left corner of the rectangle contains a lump of soap; the player who is forced to eat this square loses. The red arrows in the illustration at the right show the moves in a hypothetical game played by Wun and Too with a 4-by-4 chocolate bar (a square divided into 16 smaller squares). In this game, Too makes a bad mistake on his first move and loses a game that he should have won. The gray arrows show all the other moves that could have been made instead. This diagram is the game tree for 4-by-4 Yucky Chocolate.

A winning strategy is a sequence of moves that results in victory no matter what moves the opponent makes. Strategy theory for "finite" games—ones that can't continue forever and in which draws are impossible—relies on two simple principles: a position is a winning one if you can make some move that places your opponent in a losing position, and a position is a losing one if every move you can make places your opponent in a winning position. The logic here may seem circular, but it's not. I'll use these two principles to find a winning strategy for 4-by-4 Yucky Chocolate. The trick is to start from the end of the game and work backward—a process called "pruning the game tree."

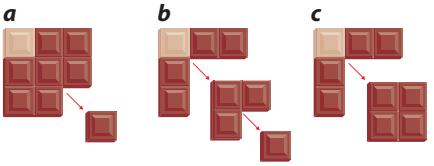
Yucky Chocolate ends when one of the players is forced to eat the upper left square. If, after several rounds of breaking, a player winds up with a rectangular piece of chocolate consisting of the upper left corner and one or more squares in the top row, the player is in a winning position: he can break off the other squares and leave the soap to his opponent. Similarly, a player is in a winning position if left with a rectangular piece consisting of the upper left corner and one or more squares in the column below it. In the game tree, each of these positions has a gray arrow leading directly to the final position.

What if a player is left with a square piece of chocolate containing the upper left corner and the three adjacent squares? As the game tree shows, this position is a losing one because the only



ALL ILLUSTRATIONS BY BRYAN CHRISTIE

GAME TREE
for 4-by-4 Yucky Chocolate shows a hypothetical game between two players (red arrows) *and all the other possible moves.*



IN THE GAME OF CHOMP the second player's winning strategy (a *and* b) *can always be accomplished in the first move* (c).

moves the player can make at this point will leave the opposing player in a winning position. This in turn implies that all the positions leading directly to the square piece of chocolate are winning positions. Working backward in this manner, we can deduce a pattern: a player is in a losing position when he is left with a square piece of any size but is in a winning position when left with a rectangular—but not square—piece of chocolate. The game of Yucky Chocolate has this pattern because any rectangular piece can be converted to a square piece in a single break—and breaking a square piece will always leave a rectangular piece for one's opponent.

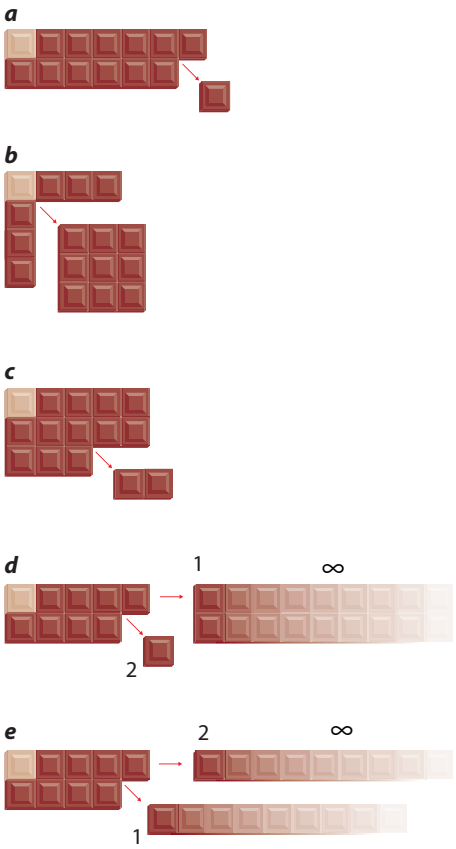
Thus, given optimal strategy, the player who takes the first turn in 4-by-4 Yucky Chocolate will always lose. The second player will win by repeatedly breaking the rectangular pieces of chocolate into ever smaller squares. (In the hypothetical game, Too lost because he didn't follow this strategy.) But in the 4-by-5 game or the 3-by-5 game, or in any game that starts with a piece of chocolate that is not square, the first player should always win. Yucky Chocolate is a dream game no matter how big the bar of chocolate is.

In principle, this procedure can be applied to any finite game to determine the winning strategy. The "root" of the game tree is the opening position, and the "twigs" are the final positions. Because we know the win/lose status of the final positions, we can work backward along the branches of the game tree, determining whether any given position is a winning or losing one. By assumption, the game tree is finite, so eventually we reach the root of the tree—the opening position. If this position is a winning one, the first player can always win by repeatedly maneuvering the second player into losing positions. But if the opening position is a losing one, the second player can always win. Unfortunately, pruning the game tree can become quite difficult if the tree is large. And because the game tree includes all possible positions and lines of play, even a simple game can have a huge tree.

The game of Chomp has almost the same rules as Yucky Chocolate, but pruning Chomp's game tree rapidly becomes impossible—and where pruning is possible, it does not reveal any pattern that could lead to a simple winning strategy. Chomp was invented by David Gale, a mathematician at the University of California at Berkeley, and is described in his new book, *Automatic Ant: And Other Mathematical Explorations* (Springer-Verlag, 1998). Gale describes Chomp using a rectangular array of cookies, but I'll stick with chocolate. As with Yucky Chocolate, the goal of Chomp is to force one's opponent to eat the soapy piece in the upper left corner of the chocolate bar. But in Chomp a player can break the bar in a greater variety of ways. Specifically, a player chooses one of the component squares and breaks off the chunk that includes it and all the squares to the right of and below it [*see illustration at above left*].

There is a neat proof that for any size of bar other than 1-by-1, Chomp is a win for Wun, the first player. Suppose, to the contrary, that Too has a winning strategy. Wun then begins the game by removing the square in the lower right corner. This cannot leave Too in a losing position, because we are assuming that the opening position is a loser for Wun. So Too can play a winning move, such as the move shown in the illustration, to leave Wun in a losing position. But Wun could have played the same move at the start of the game! This contradicts the assumption that Too has a winning strategy, so that assumption must be false. Therefore, Wun has a winning strategy. Proofs of this kind are called strategy stealing. The irony of this proof, when it works, is that it offers no clue to what Wun's winning strategy should be!

For Chomp, winning strategies are known only in a few simple cases. In the 2-by-*n* (or *n*-by-2) case, Wun can always put Too in a losing position by leaving him with a 2-by-*n* rectangle minus a single corner square [*see illustration at above right*]. In the *n*-by-*n* case—any square bar of chocolate—Wun can win by removing everything except an L-shaped edge. In succeeding turns, Wun



*FIRST PLAYER WINS in 2-by-*n (a), n*-by-*n (b) *and 3-by-5* (c) *Chomp—but the second player should always win 2-by-∞ Chomp* (d *and* e).

copies whatever moves Too makes, but reflected in the diagonal. A few other cases are known: for example, in 3-by-5 Chomp the winning move for Wun is removing the two squares at the right end of the bottom row. The winning move need not be unique: in the 6-by-13 game there are two different winning moves.

Chomp can also be played with an infinitely large chocolate bar—in which case, paradoxically, it remains a finite game, because after a finite number of moves only a finite portion of the bar remains. But with an infinite bar, Too can sometimes win. This happens, for example, with the 2-by-∞ bar. (In this case, ∞ means the entire set of positive integers.) Whatever Wun does in the first move, in the second move Too can leave Wun with a 2-by-*n* rectangle minus a single corner square—which we already know is a losing position. Chomp can also be played on a doubly infinite array, or in three or more dimensions. On the whole, little is known about winning strategies for these generalizations. SA

THE PARADOX OF GENDER

Review by Carol Tavris

The Two Sexes: Growing up Apart, Coming Together
BY ELEANOR E. MACCOBY
Belknap Press of Harvard University Press, Cambridge, Mass., 1998 ($39.95)

Talking about sex differences is America's second favorite indoor sport. (The first is practicing them.) Women wonder why little boys love guns, dump trucks and robots, why men hog the remote and why their husbands don't talk about their feelings. Men wonder why women talk so much about feelings and don't just get on with it. Maybe, we privately think, scientists really will one day discover a techno-gizmo gene on the Y chromosome and a recessive verbo-blather gene on the X.

ELIZABETH CREWS/*Image Works*

IN CHILDHOOD, BOYS AND GIRLS often establish two separate "cultures" that reflect their different play preferences…

Everyone is fascinated by sex differences, and that's the problem for researchers who study them. More than any other topic of inquiry, we *live* this one—in our beds, boardrooms, playgrounds, kitchens—so we all have our favorite theories that fit our experiences and prejudices. Scientists, though, must confront what they call "the paradox of gender": the fact that while they are rummaging around in their laboratories trying to find sex differences and locate their origins, in the outside world sex differences rise and fall (so to speak) as rapidly as hemlines and stock prices.

A mere 40 years ago, for example, who could have predicted the blurring of gender rules and roles we see now? What would Ozzie and Harriet have made of basketball star Dennis Rodman, sashaying around in outfits appropriate for several different genders; the growing political activism of transsexuals, intersexuals and bisexuals; sex chat rooms on the Internet; hard-muscled women running marathons; and soft-hearted men changing their babies' diapers?

In *The Two Sexes,* Eleanor E. Maccoby, professor emerita of psychology at Stanford University and a member of the National Academy of Sciences, has taken a terrific stab at the paradox of gender. The most important theme of her book is that the behavior we attribute to gender is not a matter of individual personality; it is an emergent property of relationships and groups. What people say, what they do and how they speak with members of their own sex differ considerably from how they behave when the other sex is around. Maccoby threw a hand grenade into her field of developmental psychology years ago when she showed that gender differences in children couldn't be accounted for by personality traits but rather by the gender composition of a group. Little girls aren't "passive" as some ingrained quality, for example; they are passive only when boys are present.

This approach shows why traditional efforts to measure sex differences in terms of individual traits or abilities (empathy, vanity, submissiveness, intelligence, math abilities and so forth) are fruitless and become quickly dated. Sex differences that show up in any study tend to be artifacts of education, power, the immediate social context and the historical moment, which is why they wax and wane with the times. For example, "female intuition" about other people is actually subordinate's intuition: both sexes are equally intuitive when they have to read a superior's mood, nonverbal signals or intentions—and equally thick-headed, when they are the bosses, about their subordinates' feelings.

Maccoby sets out to explain the great mystery of gender development: the virtually universal existence of gender segregation among children, which remains impervious to the best efforts of egalitarian-minded parents and teachers. Boys and girls will play together if adults require them to, although it's often "side-by-side" play, in which each does his or her own thing, but given their druthers, children self-segregate. The result, Maccoby argues, is the emergence of a "girls' culture" and a "boys' culture" that are strikingly different in play styles, toy preferences and ways of interacting. Before long, as with any two nations, schools or ethnic groups, boys and girls identify with their own in-group, they stereotype and disparage members of the out-group, and they misunderstand or feel uncomfortable with the other group's ways of doing things.

The most puzzling fact about the two cultures of gender, however, is their asymmetry. Boys' groups, Maccoby shows, are "more cohesive than girls' groups: more sexist, more exclusionary, more vigilant about gender-boundary violations by their members, and more separate from adult culture." Throughout childhood, as throughout life, there are fewer penalties for girls who encroach on boys' turf and who like to do boy things than for boys who venture onto girls' territory.

And so the great question is: Why? Why are children, in the words of sociologist John H. Gagnon, the Gender Police, enforcing rigid stereotypes that many of their parents have long discarded? Why do they behave differently with their own sex than with the other? And what, if anything, is the link between childhood and adulthood, considering how many members of the Gender Police eventually become gender criminals, breaking as many gender rules as they can, or gender revolutionaries, trying to rewrite the rules altogether? Maccoby's answers are both timely and old-fashioned, falling squarely between two antithetical trends in the current study of gender.

Opposite or Other?

One, the oldest empirical tradition, takes an essentialist approach. Essentialists regard a gender-related attitude, trait or behavior as being something embedded in the person—internal, persistent, consistent across situations and time—and thus they tend to regard the sexes as "opposites": men are aggressive, women pacifistic; men are rational, women emotional. The most extreme version of essentialism is represented by pop-psychologist John Gray, who thinks men are from Mars and women are from Venus. But here on Earth all kinds of other notions of inherent sexual opposition are widespread. For Jungians and psychoanalysts, men and women are guided by opposite archetypes and unconscious dynamics. For some feminist psychologists, men and women have inherently different ways of knowing, ways of speaking, ways of moral reasoning and the like. For neuroscientists, men's and women's brains operate differently. For sociobiologists, male promiscuity and female monogamy are opposite, hard-wired reproductive strategies. (When sociobiologists learned that the males of many species are nurturant and monogamous and the females of most species are promiscuous, they reconnoitered and decided that these reproductive strategies too are adaptive.)

In contrast, researchers who take a social constructionist approach vigorously dispute all forms of essentialism. Social constructionists hold that there is no "essence" of masculinity and femininity, for these concepts and labels are endlessly changing, constructed from the eye of the observer and from the historical and economic conditions of our lives. "Opposition," for example, is a social construction, not an empirical reality; it is a stereotype that blinds us to the greater evidence of gender similarity. Are men rational? Sure, except in love, war and sporting events. Are women unaggressive? Sure, unless you define "aggressiveness" as the intention to harm another, in which case they don't differ from men. Constructionists regard gender as a performance, not an attribute. People don't *have* a gender, they *do* gender, which is why their behavior changes so much depending on the situation. A teenage boy may "do" masculine when he's with a pack of his male friends but "do" feminine by tenderly caring for his baby brother (if his friends aren't watching).

For the constructionists, therefore, the really interesting news about gender lies not in the traditional oppositional categories but in the increasingly diverse and growing numbers of people who aren't conforming to the categories at all. Even to the fundamental categories of male and female: biologists such as Anne Fausto-Sterling and others have shown that human dimorphism is neither as obvious nor as universal as most people believe. The number of "intersexed" infants born with anatomical, hormonal or genotypic ambiguities is about 2 percent of all live births—a small percentage, but many thousands of individuals. Recent books in this genre include Suzanne J. Kessler's *Lessons from the Intersexed,* Marianne van den Wijngaard's *Reinventing the Sexes* and Alice D. Dreger's *Hermaphrodites and the Medical Invention of Sex*.

Maccoby, calling her book "The Two Sexes," is not remotely interested in the "transgender" research that is revolutionizing gender studies; she finds the whole subject tangential to the question of male-female differences. Yet she also rejects biological reductionism and other essentialist ideas of opposition. She refers always to the "other" sex, never the "opposite" sex; she never assumes that biology is the whole story, emphasizing repeatedly that it interacts with experience and culture. For example, childhood sex segregation may be universal, but it differs in form and degree depending on culture. Societies in which men clearly have higher status than women, Maccoby reports, are those in which boys make "the earliest and strongest efforts to distance themselves from women and girls—from their own mothers, as well as from other females."

In the second part of her book, Maccoby reviews the voluminous research on biological factors, socialization practices and cognitive processes that might explain the mystery of children's self-segregation. Many of the findings here are fascinating. For instance, sex segregation does not originate because boys have a greater "activity level," as commonly believed. In fact, boys aren't more physically active than girls when children are playing on their own. But when boys play with other boys, they become more excited and aroused than girls do and by different things: threats, challenges and competition. High rates of male activity are a *consequence* of male-male play, not a cause. Besides, activity levels decline from ages four to six, when gender segregation steadily increases.

Maccoby, a scrupulous scientist, gives us a state-of-the-art review of the re-

BILL TRUSLOW *Liaison International*

…BUT IN ADULTHOOD, most men and women become "bilingual."

search, not a cohesive argument designed to support a thesis. In this age of simplistic pop-psych overgeneralizations, her caution and scholarly rigor are refreshing. Yet readers may occasionally get lost in the dense thickets of evidence for and against each line of explanation. I felt I was eating many delicious raisins while being denied the satisfaction of a whole piece of cake.

The third section of the book, in which Maccoby examines the links between childhood and adulthood, is the weakest, perhaps because of her own ambivalence. On the one hand, she argues that the gender segregation established in childhood and the asymmetrical cultural differences that result from it persist in many adult contexts, including the workforce and men's and women's habits, preferences and disputes. Many men don't listen to their wives, Maccoby suggests, for the same reason that little boys refuse to be influenced by little girls.

On the other hand, she subtitles her book "Growing up Apart, Coming Together," which accurately reflects the fact that vast changes in men's and women's relationships have occurred "in spite of, rather than because of, the way boys and girls are socialized by their parents"—and in spite of, I might add, sex differences in hormones or alleged brain function. Among adults, circumstances and experiences supersede the maturational pull of genes and hormones and even the socializing pull of parental instruction. That is why a random group of 50-year-olds is more diverse than a group of five-year-olds and why adults today find themselves doing things they once would never have imagined for themselves. And it's why Maccoby's generalizations about adults seem flat and stereotypic (although certainly they have an element of truth), in contrast to her brilliant portrayal of children.

The war between essentialists and constructionists is bound to continue, and this book will provide ammunition for both sides. But perhaps the war between men and women will find a lasting truce if, as Maccoby hopes, we understand that men and women don't have to be the same in order to be equal—in opportunities, income or love.

CAROL TAVRIS is a social psychologist who writes frequently on the behavioral sciences. She is author of The Mismeasure of Woman *and co-author of two textbooks,* Psychology *and* Psychology in Perspective.

THE EDITORS RECOMMEND

MAXWELL'S DEMON: WHY WARMTH DISPERSES AND TIME PASSES. Hans Christian von Baeyer. Random House, New York, 1998 ($25).

Simple questions, von Baeyer points out, often lead to profound insights. "What is warmth?" is one such question, and the efforts of many curious people to answer it have built the science of thermodynamics. Those efforts comprise von Baeyer's story. He describes felicitously and lucidly the laws of thermodynamics, entropy, the dissipation of energy and time's arrow, providing along the way little-known details about the lives of the famous and not so famous pioneers of the field: Benjamin Thompson (Count Rumford), Sadi Carnot, Robert Mayer, James Joule, Hermann von Helmholtz, Rudolf Clausius, William Thomson (Lord Kelvin), James Clerk Maxwell and Ludwig Boltzmann. Their work moves von Baeyer to put in his own words a concept stated by physicist Gerald Holton: "A few simple themes—unspoken assumptions and intuitively held prejudices that originate outside science—underlie all scientific thought."

GEONS, BLACK HOLES AND QUANTUM FOAM: A LIFE IN PHYSICS. John Archibald Wheeler with Kenneth Ford. W. W. Norton and Company, New York, 1998 ($27.95).

Wheeler, an eminent Princeton physicist, is the originator of the first and third terms in this book's title and the man who gave the second its push into the language. They reflect the breadth of his interests and accomplishments in physics. "Geon" incorporates g for "gravity," e for "electromagnetism" and -on as the word root for "particle" and identifies a "hypothetical entity, a gravitating body made entirely of electromagnetic fields." A voice from the audience at a talk Wheeler gave in 1967 suggested "black hole" as a name for what he had discontentedly called a "gravitationally completely collapsed object." He immediately recognized the term's felicity and adopted it. At another time, pondering gravitation and general relativity, he found himself "forced to invent the idea of 'quantum foam,' made up not merely of particles popping into and out of existence without limit, but of spacetime itself churned into a lather of distorted geometry." Wheeler describes these and many other concepts in physics with characteristic clarity and salts his tale with many fine anecdotes about his encounters with other famous physicists, Lyndon Johnson and American railroads.

LOOKING FOR EARTHS: THE RACE TO FIND NEW SOLAR SYSTEMS. Alan Boss. John Wiley & Sons, New York, 1998 ($27.95).

It is a riveting question: Are there other Earths, bearing life in some form? Earlier the question was, Are there other planets outside the solar system? A flurry of discoveries over the past three years has provided the answer to that one: yes. Boss traces the story chronologically, telling it from the viewpoint of an astrophysicist and incidentally providing a rewarding account of how astronomers and astrophysicists do their work. Now, he says, we are in a new era, "in which we will discover many planetary systems circling stars in our neighborhood of the galaxy, systems containing Earth-like planets capable of supporting life."

THE THERMAL WARRIORS: STRATEGIES OF INSECT SURVIVAL. Bernd Heinrich. Harvard University Press, Cambridge, Mass., 1996 ($27).

BERND HEINRICH

"For insects … the struggle to keep body temperature within an acceptable range is constant," Heinrich writes, "and often it is a matter of life or death. Each insect is a 'thermal warrior' in a contest with its predators and competitors in the context of its physical environment." Heinrich tells of this struggle as it affects insects, from a glacier-dwelling midge to a variety of bees, ants, moths and termites. He writes with an unflagging sense of wonder at what insects can accomplish.

THE LAST RECREATIONS: HYDRAS, EGGS AND OTHER MATHEMATICAL MYSTIFICATIONS. Martin Gardner. Copernicus, New York, 1997 ($25).

Gardner, who conducted the immensely popular Mathematical Games department

of this magazine from 1956 to 1981, presents here his 15th and final (he says) collection of those columns. There are 23 of them, culled from his last seven years of writing for the magazine. They deal with such engaging topics as "The Wonders of a Planiverse," "Bulgarian Solitaire and Other Seemingly Endless Tasks," "*M*-Pire Maps," "The Monster and Other Sporadic Groups" and "Taxicab Geometry." As in previous collections, Gardner brings his topics up-to-date and includes some of the letters from readers that his beguiling problems brought forth.

THE MAN WHO LOVED ONLY NUMBERS: THE STORY OF PAUL ERDÖS AND THE SEARCH FOR MATHEMATICAL TRUTH. Paul Hoffman, Hyperion, New York, 1998 ($22.95).

The peripatetic Hungarian mathematician Paul Erdös (1913–1996) was renowned for his almost total concentration on his work. Hoffman describes him as "a mathematical monk" who renounced physical pleasure and material possessions for an ascetic, contemplative life, a life devoted to uncovering mathematical truth. This he did in 1,475 papers that he wrote or co-authored with 485 collaborators—more than any other mathematician has produced and a landmark that has given rise to the cherished "Erdös number." An Erdös co-author's number is 1; a mathematician who has published with someone who was an Erdös co-author is a 2, and so on in widening circles to infinity for everyone who has never written a mathematical paper. Hoffman is among those at infinity, but he describes Erdös's life and eccentricities engagingly and deals comprehensively with the great man's mathematical work.

SILENT THUNDER: IN THE PRESENCE OF ELEPHANTS. Katy Payne. Simon & Schuster, New York, 1998 ($25).

Acoustic biology is Payne's field. In earlier studies, she learned that great fin and blue whales communicate by infrasound. One day, standing outside a cage in the elephant house at Washington Park Zoo in Portland, Ore., she felt a silent thunder—a throbbing that seemed to be associated with what the elephants were doing. A question struck her: "Were the elephants calling to each other in infrasound?" With tape recordings at the zoo and from field studies in Africa, she established that infrasound is a component of elephant communication, particularly over long distances. Payne's chatty tale embraces not only the science of her work but also her emotional involvement with her subjects, which leads her to vent her distress at the culling and poaching of elephants in Africa for the ivory trade. She pleads for a different approach: to "harvest the ivory without harvesting the elephants."

PATHS OF FIRE: AN ANTHROPOLOGIST'S INQUIRY INTO WESTERN TECHNOLOGY. Robert McC. Adams. Princeton University Press, Princeton, N.J., 1996 ($29.95).

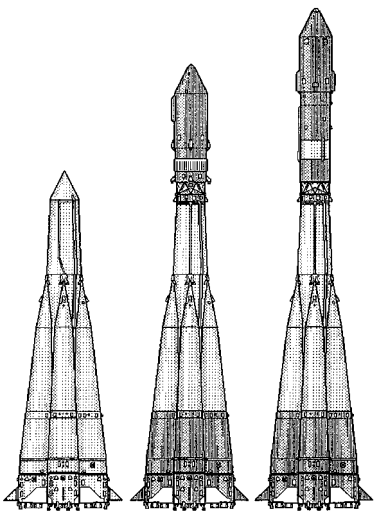
Anthropology is the study of human activity, and technology is an advanced embodiment of that activity. Adams, former secretary of the Smithsonian Institution, presents a scholarly study of the advance of Western technology, with emphasis on developments in Britain and the U.S. Beyond tracing the history of Western technology and its great rewards, he also points out that technological advance has entailed consequences such as pollution and the destruction of environmental resources and amenities, leading to the need to answer a difficult question: "What is an acceptable role for government monitoring and regulation, which necessarily constrains, and sometimes can distort, the range of techno-economic choices?"

LIFE: A NATURAL HISTORY OF THE FIRST FOUR BILLION YEARS OF LIFE ON EARTH. Richard Fortey. Alfred A. Knopf, New York, 1997 ($30).

"The narrative of life requires a scale of thousands to millions of years, acting over a drama of more than 3,000 million years." It is a grand narrative, told grandly by Fortey, a senior paleontologist at the Natural History Museum in London. Drawing on a great breadth of knowledge, he flavors the narrative with illuminating and often surprising analogies and quotations from the likes of Pope, Swift and Yeats. His story takes life from the first single-celled organisms to prehistoric humans—over "the vast tract of time after the Sun blazed into heat … and before humans started making pots, building ceremonial centres, and recording the details of their daily transactions on pottery slabs."

THE CAMEL'S NOSE: MEMOIRS OF A CURIOUS SCIENTIST. Knut Schmidt-Nielsen. Island Press/Shearwater Books, Washington, D.C., 1998 ($24.95).

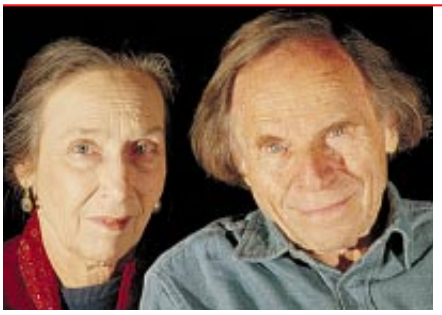
"It has been said," Schmidt-Nielsen writes, "that the primary function of schools is to impart enough facts to make children stop asking questions. Some, with whom the schools do not succeed, become scientists." Schmidt-Nielsen did and has had a prominent career in animal physiology. "The questions I have tried to answer have been very straightforward, perhaps even simple: Do marine birds drink sea water? How do camels in hot deserts manage for days without drinking?… How can snails find water and food in the most barren deserts? Can crab-eating frogs really survive in sea water?" Even that summary does not fully indicate the liveliness of his curiosity. In his travels to almost everywhere, he made a point of eating the local delicacies, and he reports that scorpion meat "had no pronounced flavor," that sampling boiled locusts "was like eating tasteless shrimp without peeling them" and that durian, a foul-smelling tropical fruit that he tried in Bangkok, was "wonderful" for him but merely "tolerated" by his traveling companion. Discussing the avalanche of papers and books faced by today's workers in science, he says: "I maintain that word of mouth is the most important aspect of scientific communication today, and in this sense we have returned to medieval conditions."



COURTESY OF PETER GORIN

EYE IN THE SKY: THE STORY OF THE CORONA SPY SATELLITES. Edited by Dwayne A. Day, John M. Logsdon and Brian Latell. Smithsonian Institution Press, Washington, D.C., 1998 ($29.95).

The cold war propelled the U.S. and the Soviet Union to stunning technological heights—both figuratively and literally. Among the most impressive—and most deeply concealed—of those achievements is satellite reconnaissance, on which the U.S. continues to spend billions of dollars every year. This book chronicles in satisfying detail the origins of U.S. satellite reconnaissance by focusing on the pioneering Corona program, under which some 800,000 satellite images were made between 1960 and 1972. SA



COMMENTARY

WONDERS

by Philip and Phylis Morrison

Field Guide to the New Biology Lab

The guinea pig, still a metaphor for any subject of experiment, is no longer a laboratory fixture. Molecular biology, built on the rock of DNA and its double helix, was realized largely out of the study of a humble, single-celled microorganism: a human symbiont, the colon bacterium *Escherichia coli*. Bacterial DNA is the purest quill of all, one molecular loop of code, all tape and no cassette. Five decades of attention have yielded a rich understanding of single-celled life.

During most of those years, many biology research labs hardly looked the part; they were austerely chemical, with plenty of glassware and circuitry but no animals or plants at all. The investigators themselves were the main visible life-form, eagerly transferring precise droplets among bubbling tubes and dripping columns. Outsiders joked that soon M.I.T. biology Ph.D.'s. wouldn't remember how to raise a white mouse or grow a green shoot—or even to tell the two apart.

Single test tubes nurtured bacterial cells by the 100 million. In that crowd, such rarities as the nonlethal mutants, key to most genetic analysis, were there to be found. For study of rare forms, the creatures screened must be numerous, yet if a lab is to provide them they must be small in size and swift to multiply. Before World War I, classical chromosomal genetics had set that standard firmly on fruit flies. No more! At M.I.T. and elsewhere, we have had the opportunity to enjoy enthusiastically guided visits to the grazing areas, greenhouses and ponds—over rooms, not acres, to be sure—where today's biologists are raising new populations of mutant organisms. Their intensive study seeks to carry the triumph of molecular biology out to the world of multicellular life.

The first of the novel organisms is a threadworm, gracile, crystal-clear, a millimeter long; the second is a small weed; and the latest is a little striped freshwater aquarium fish. These amazing choices display the development of an individual organism from egg to egg, as well as the past billion years of evolution—ontogeny and phylogeny both—right on the benches of a couple of hundred exciting labs worldwide.

The threadworm is a tiny nematode; its manifold kin include human parasites such as the hookworm, although the lab species is no parasite but an abundant, free-living predator of soil bacteria. Its diet in captivity is a satisfying strain of *E. coli,* supplied as a living lawn on dishes of nutrient agar. Thereon the transparent worms sinuously bend and sway, grazing as you watch; the species is well named *Caenorhabdytis elegans*. The microscope at high power shows every individual cell, as well as the store of eggs and sperm. Under low power you can watch many worms; in a few days most *elegans* mature after four moltings into hermaphrodite adults. Forming both eggs and sperm, they each will lay and fertilize about 300 eggs during three days at (English) room temperature and continue living for weeks if fed, but with no more cell divisions.

In 1963 the far-seeing Sydney Brenner selected *C. elegans* as the creature to lead the way from microbes to multicellular development: "We propose to identify every cell in the worm." It has no blood cells, no immune cells, no cells that circulate at all. Each cell and the fate of every lineage are now mapped, a total of 959, plus the germ cells. Another 131 cells are programmed to die along the way. For instance, when two cells develop, either is able to form part of the tail of a male worm, but only the one better placed to function survives.

Cells can be extirpated one by one with a laser "scalpel," the ultimate in dissecting tools. The worm has been mapped through 20,000 gossamer slices viewed under the electron microscope. Every neuron path was traced to complete the writing diagram of the nervous system, just 302 cells made up into 118 types of elaborately linked neurons, each with its proper chemical neurotransmitters, mostly the same as those of humans. The larvae are easily bathed in mutagen, bred, sometimes even crossed,

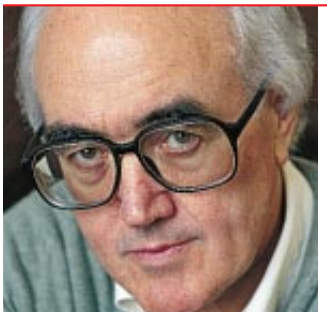
We saw mature zebra fish transformed—their dark stripes transverse instead of lengthwise!

and the desired odd progeny stored indefinitely in cryogenic refrigerators.

For decades, it has been the genes that attract most workers: 150 feet of paper stretch out of the lab and down the hall, the growing map of the *C. elegans* genome, six chromosomes spelled out in base pairs. Although gaps still exist, this year the work will be complete; 100 million base pairs should encode the entire worm. To turn this coded manuscript into a functional map of genes is the urgent task. Right now the physical map of base pairs is being studied worldwide, as much over the Internet as in any single lab. Clues to its meaning reside in plausible matches to DNA sequences found in more sporadic work on other worms but also to sequences in yeasts, fruit flies and mice; human analogues are not rare. DNA is taken from mutant worms whose "defect" has been cured by insertions of trial alien genes, and matching fragments are then tested, even synthe-

Continued on page 133

DUSAN PETRICIC



CONNECTIONS

by James Burke

Does This Take You Back?

DUSAN PETRICIC

Disease: nostalgia. Symptom: an irrepressible desire to go home." That plaintive entry caught my eye recently while I was noodling through a dusty history of quackery that was otherwise so dull it would cure your insomnia. The line occurred in the section devoted to nosology, the classification of disease by symptoms, a rave craze in the 18th-century medical world.

Turns out that the top nosology maven of the late 1760s lived in Scotland. Name of William Cullen, he was professor of the theory of medicine at the recently opened Edinburgh University Medical School and author, throughout a long and distinguished career, of one single underwhelming research paper on evaporating fluids. Back then, you didn't perish if you didn't publish.

William Cullen's star pupil (and the guy who would succeed him as prof) was Joseph Black, who was famous for at least three things. He always carried a green silk umbrella, he discovered latent heat and so was able to tell James Watt how to make his steam engine work, and he founded a dining club in Edinburgh known as the Oyster. This select nosherie was the regular elite-meet for most of the luminaries of the Scottish Renaissance (the *Dictionary of Scientific Biography*'s phrase, not mine). Among those to be heard enlightening the world over the seafood each week were economist Adam Smith and a now half-forgotten expert on the circulatory system (and friend of Black), James Hutton.

I suppose Hutton was a great example of the Scottish Renaissance man the *DSB* had in mind. He studied humanities, physics, geography, law, medicine and chemistry and qualified as a doctor. Then, in the manner of such eclectic people, he became a farmer. Why not? It may have been his consequent landowner's interest in rocks and soil that got him into geology. In 1764 he began a series of trips to stonier parts of the British Isles, tapping and chipping away. The chief object of his attentions tended to be basalt, because of the sundry explanations available at the time regarding the formation of the earth. Hutton was most attached to the molten-interior, liquid-granite hypothesis.

Well, all the hammering must have been really productive, because in 1785 Hutton penned the outline of a modest work, eventually published under the modest title of *Theory of the Earth,* and blew everybody away with his description of a great, cyclical process: land

Back then, you didn't perish if you didn't publish.

degradation by erosion, resultant deposits washed into the sea, sedimentary layers settling over millions of years, ultimately to be thrown up again, to be eroded once more and so on. And, as Hutton said, if this process had taken as long in the past as it seemed to take in the modern world, then the planet was humongously ancient, never mind the Bible. It would be this particular bit of Hutton's geologic uniformitarianism (the fancy name for his theory) that would in time inspire Darwin.

Hutton was accompanied on some of his peregrinations by another Edinburgh pal, John Clerk, also an Oyster member and amateur rockhound. Clerk goes down in history for his "Essay on Naval Tactics," reputed to have inspired Admiral Horatio Nelson to victory on the *Victory.* Clerk's other notable feat was to marry the sister of Robert Adam, one of the hottest architects in 18th-century Britain. You wanted a scrambled hash of oh-so-chic neoclassical bits and pieces all over your interiors, you called Robert, who would transform your crumbling pile into instant pseudo-Greco-Roman for only an arm and a leg. One of his more fastidious imitators was a cabinetmaker and furniture fiend named George Hepplewhite, who added simplicity and elegance to some of Adam's more extravagant chairs and in 1788 came out with his own best-selling *Cabinet-Maker and Upholsterer's Guide,* whose designs were to be often copied (in America) but rarely acknowledged (in England).

Hepplewhite's guide included instructions for japanning mahogany. Japan lacquer (so called because in the 17th century, when lacquer arrived in the West from China, nobody knew the difference between one place and the other) had become so popular among European royals that you needed a king's ransom to buy some. The Chinese wouldn't reveal the secret of lacquer manufacture, so buyers took a real shellacking. Till 1732, when Thomas Allgood, in the Welsh town of Pontypool, came up with a new kind of lacquer, which became known as Pontypool Japan. If you possess one of your great-great-grandmother's tea caddies with Chinese figures on it, you know what I mean. The big advantage of Allgood's mix of linseed oil, umber, litharge, pitch and turpentine was that it was cheaper and more available than the real thing. It could also be put onto tin, which was

cheaper and more available than wood. If, that is, you knew how to make something out of tin. For which you had to know how to roll sheets of iron.

So guess where one of the best iron-rolling mills around at the time was? Pontypool. At one point, one of the Allgood family worked for the mill's owner, John Hanbury, who had developed a technique for processing red-hot iron through several sets of rollers to make it extremely thin, whereupon it could be dipped in molten tin and shaped into utensils ready for Pontypool japanning. When Hanbury originally set up his mill, the cutting edge in sliced sheet-iron manufacture was at Stjärnsund, Sweden, where an unsung genius named Christopher Polhem designed and built incredible water-powered machines that would do anything you wanted to hot metal. Polhem had started out as a mining engineer, and his work would end up giving Sweden the reputation for expertise in metallurgy it still has today. Not much is known about Polhem, who got written up by a young admirer whose reputation then totally eclipsed his, because after the two of them had worked together on the Swedish Royal Board of Mines, the guy in question went on to what I suppose might be described as higher things.

Emanuel Swedenborg (you're probably there before me) was yet *another* polymath. He did humanities, geology, metallurgy, paleontology, flying machines and submarines, started the first Swedish science journal and dabbled in astronomy. In 1745 he had an epiphanous vision, in which God ordered him to dump science and technology in favor of the Bible and in doing so changed Swedenborg's life from that of propeller-head to that of prophet of his Church of the New Jerusalem.

One American commentator on Swedenborg was a journalist and businessman named John Bigelow, who in 1849 became managing editor and part-owner of the *New York Evening Post*. His co-owner at the *Post* was one of America's best-known Romantic poets, name of Bryant, who wrote about the woods and streams of his native Berkshires. All his life he wanted to return there. Not surprisingly, for such a nostalgic type, his first two names were William Cullen. Does that take you back … to the beginning of this column? SA

Wonders, *continued from page 131*

sized, by recombinant DNA techniques.

The stiff little weed *Arabidopsis thaliana* has been greening biology at M.I.T. and in many another lab, especially during the past five years. Its small size, facile growth, self-fertility and tens of thousands of tiny seeds per plant all commend it. But what has been wonderful is the discovery of how to transform its genome. Conventionally, one must prepare and transform a single plant cell, then let it mature in tissue culture, a tricky, months-long process. In contrast, an entire *A. thaliana* can merely be soaked in a culture of a celebrated plant DNA vector, the crown gall bacterium, augmented by stretches of specifically cloned DNA that it will insert at random into host nuclei. High school students can do it.

The plants grow under greenhouse lighting, hundreds in a small tray of soil, and the seeds are harvested. The transformed progeny—each of them billions of cells—are screened by eye. It is exciting to see among 1,000 weeds a couple of dozen novelties, some dwarfed, some scrawny, some whose tiny flowers have strange petal counts, even a few that grow a pulpy ring of green tissue rather resembling their cabbage cousins. Genes can be roughly located on the physical map and new plants transformed to help fine-tune for the code particulars.

The zebra fish, popular among aquarium enthusiasts, is the latest of the new guinea pigs at M.I.T.: we saw mature, silvery, centimeter-long fish transformed from the wild type, their dark stripes transverse instead of lengthwise! These are animals like ourselves, with blood, bone, brain, eyes and complex actions. The zebra fish is small and fast-developing, but its eggs and embryos, all transparent, are open to microscopic study cell by cell, almost like a giant *elegans*. The M.I.T. lab is preparing for 4,000 new tanks in three roomfuls of fish stocks and their mutants. The powers of recombinant DNA can now extend scrutiny to a class of vertebrates.

The vista here is wide: the universality of life on the earth. The dances of genes become plants and animals through a choreography we are beginning to grasp. In its intricacy lies the combinatorial wealth of the living world, its modular patterns open to investigators of a species at once part of it and yet distinct. SA

MOTION-PICTURE PROJECTORS

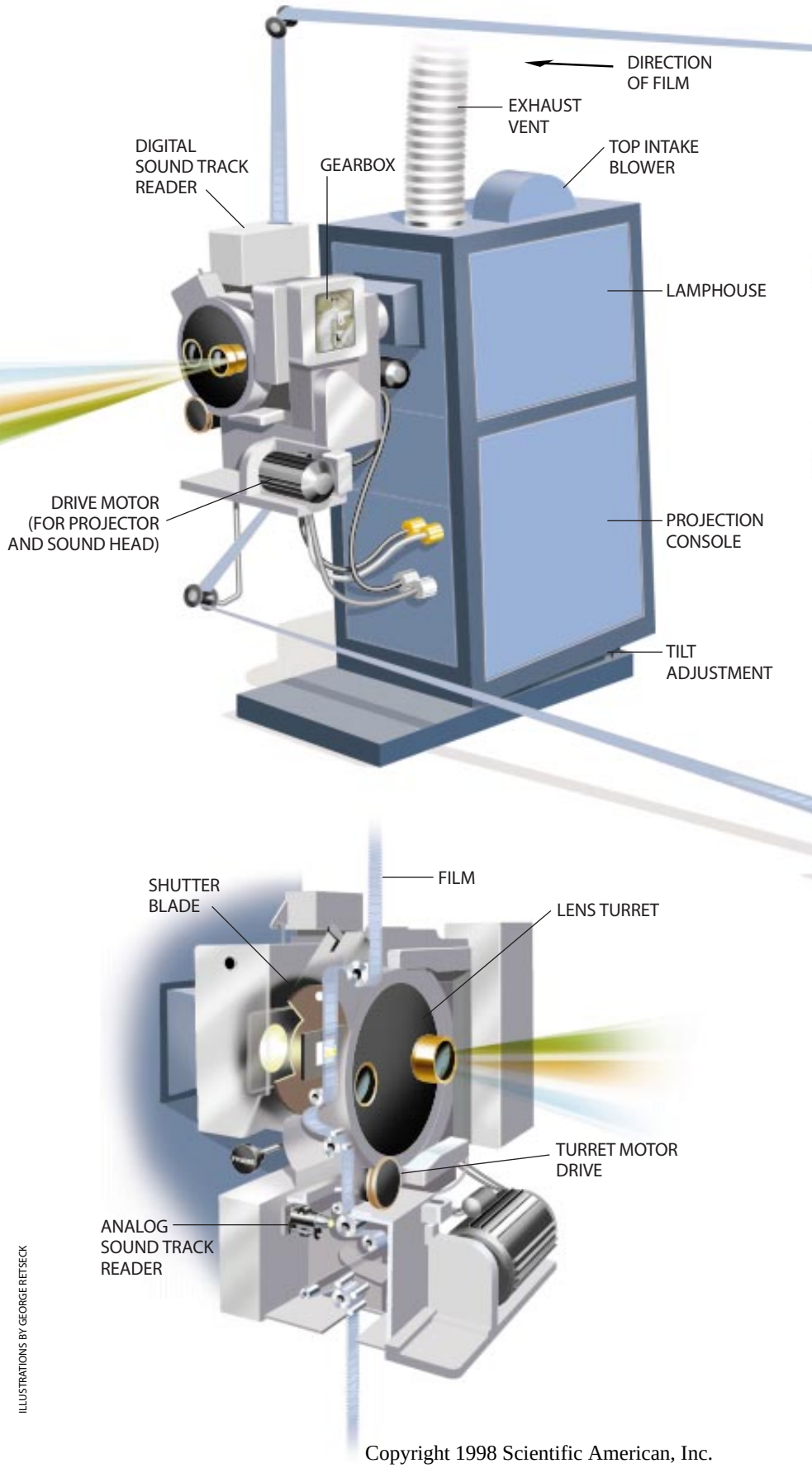
by Ray F. Boegner
Senior Vice President, Ballantyne of Omaha

If you are like most people, the term "movie projector" probably conjures up an image of a clattering piece of equipment in which film spools from one big reel to another. It is an enduring image, kept alive by scenes on television and in motion pictures themselves—even though it is at least two decades out of date.

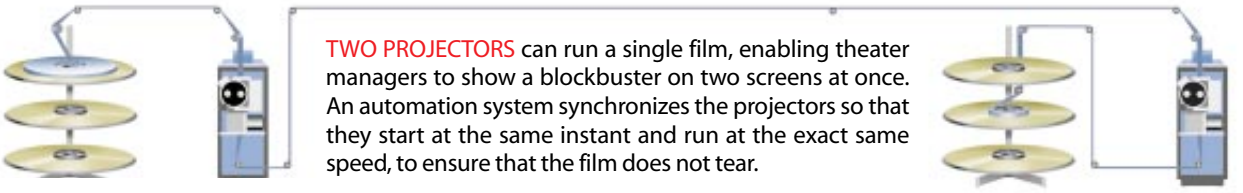
Since the mid-1970s, almost all the projectors in theaters in developed countries have used platters rather than reels to hold the film being run through the projector. These platter systems, invented around 1970, have several major advantages. One is that the film never needs to be rewound and is therefore subjected to much less stress. Film feeds from the center of one platter through the projector and onto the other platter, where it coils around itself from the center of the platter toward its outer edge. When the movie is over, the projectionist need only switch platters before the next showing.

Another advantage is that a platter can hold up to four and a half hours' worth of film—7.4 kilometers (4.6 miles)—so essentially all films can be set up and then left to run until the closing credits. In contrast, reels can hold only a little over one hour's worth of film (about 1.8 kilometers). So the showing of any film longer than an hour requires two projectors and a projectionist to stand by and switch on the second projector as the film is running out in the first one.

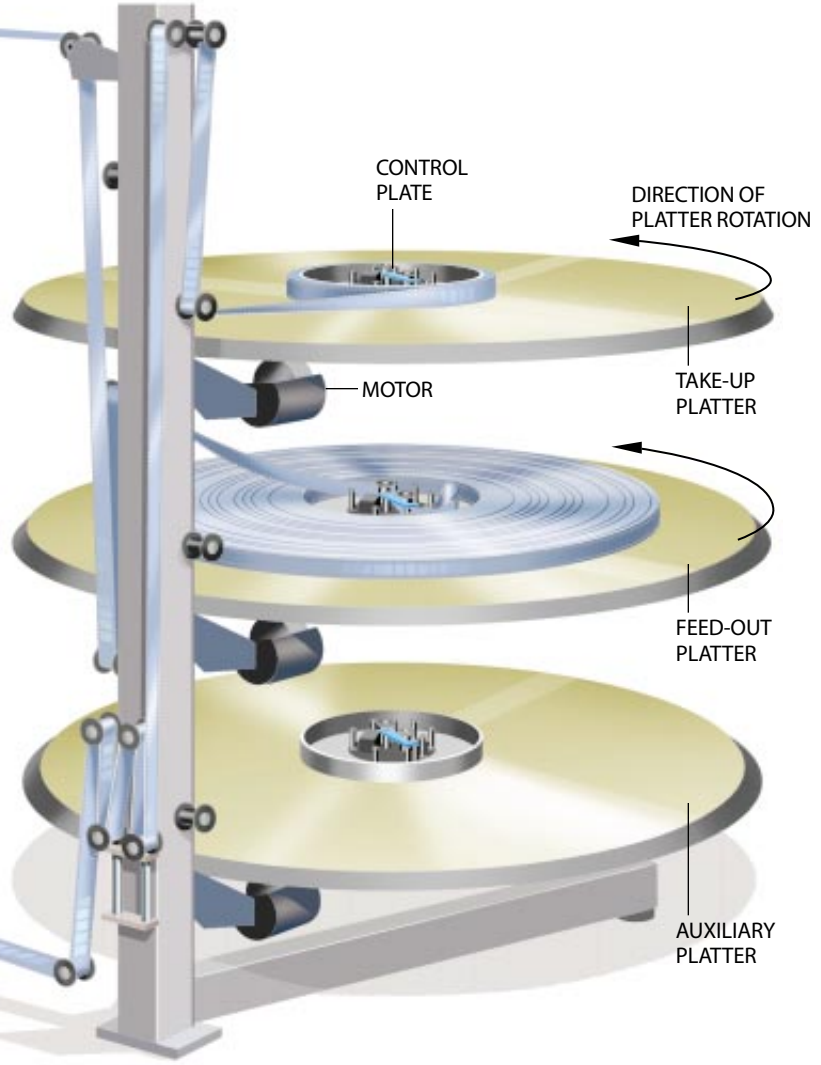
What made the platter system possible was the invention of the xenon lightbulb by German engineers in the late 1940s and its subsequent adaptation, in 1954, as a light source for use in movie projectors. Before xenon bulbs, movie-theater projectors used carbon-arc lamps. The carbon rods in these lamps were vaporized after about an hour and were changed when the projectionist switched to the second projector during a screening. Thus, before xenon bulbs, which can last for 2,000 hours or more, there was little incentive to have a projector that could run for more than an hour.



ILLUSTRATIONS BY GEORGE RETSECK

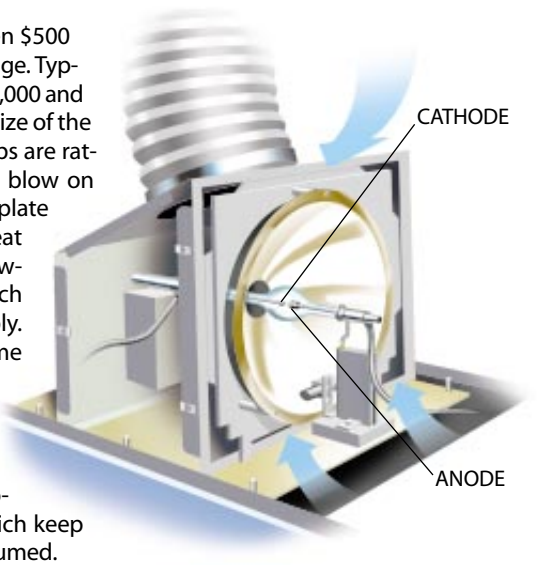


TWO PROJECTORS can run a single film, enabling theater managers to show a blockbuster on two screens at once. An automation system synchronizes the projectors so that they start at the same instant and run at the exact same speed, to ensure that the film does not tear.



MODERN MOTION PICTURE FILM is polyester-based and 35 millimeters wide. In addition to the picture seen, the film has printed on it one or more audio formats—only one of which is used at a time, depending on the type of sound head the motion-picture projector is equipped with and the type of audio-processing equipment the theater has installed.

PROJECTOR BULB costs between $500 and $2,000, depending on wattage. Typical power ratings are between 2,000 and 4,000 watts, depending on the size of the theater; the most powerful bulbs are rated at 7,000 watts. Several fans blow on the bulb and nearby aperture plate to dissipate the tremendous heat and prevent the bulb from blowing out inside the projector, which would damage it considerably. Bulbs work on basically the same principle as arc welders: direct current produces an arc of electricity up to a half a centimeter long. Surrounding the arc are several atmospheres of inert xenon gas, which keep the electrodes from being consumed.



POPULAR SOUND FORMATS are Dolby SR (analog), where dual tracks of analog audio information are printed on the film right next to the picture; Dolby Digital Sound, where the digital sound information is printed between the sprocket holes on the film; Sony SDDS Digital Sound, where the digital sound information is printed near both edges of the film and can produce up to eight discrete channels of sound; and DTS Digital Theater Sound, in which a proprietary time code (*not shown*) printed next to the analog track synchronizes with the film's digital sound track, which is stored on a CD-ROM.