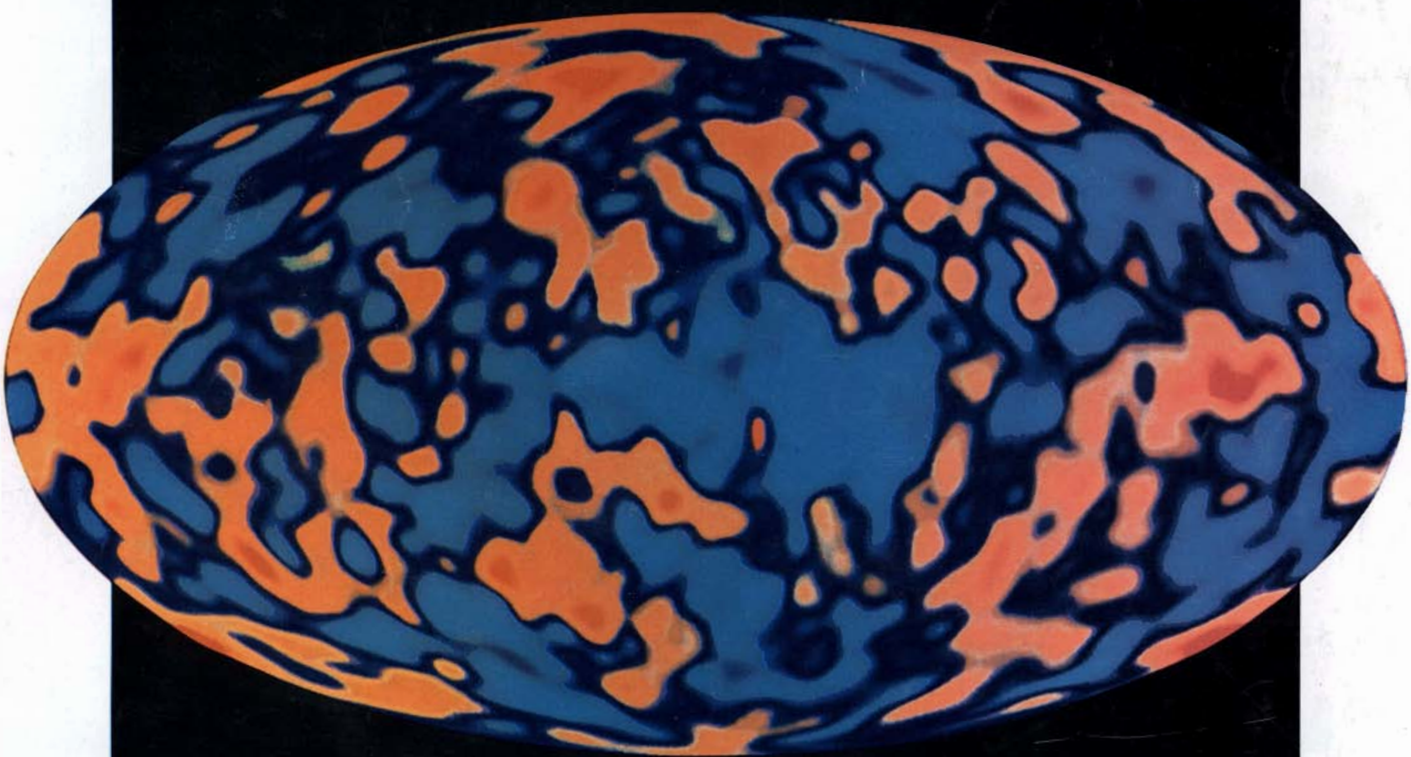
SCIENTIFIC AMERICAN

JULY 1992
$3.95

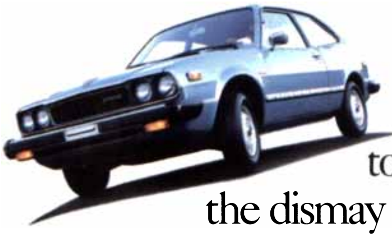
The strange world of quantum physics.
Computer programs that evolve.
Living Stone Age artisans of New Guinea.

MORE PROOF FOR THE BIG BANG



COBE *sky map shows telltale fluctuations in the temperature of the early cosmos.*

When's the last car chase that



It was as if a starter's gun had been fired.

The year was 1976, and the Honda Accord arrived much to the delight of America's drivers and critics. And much to the dismay of other automotive designers and engineers.

After all, this was a car with a totally different interpretation of efficiency, comfort and performance. An automobile so thoughtful, practical and complete in its design, the competition had no other choice than to follow in its tracks.

Now, more than a decade and a half later, the Accord continues to perform an amazing feat. Each year, it somehow improves.

For instance, the interior. It's surprisingly large, and amazingly quiet. Which

makes it more comfortable than you ever imagined. You'll notice it when you sit in the firm, contoured seats. When you slide open the power moonroof. And when you ride silently over bumps and through dips.

The refinements underneath the hood have been equally dramatic. A fuel-injected, 140-horsepower engine gives the Accord more power than ever before. And the 4-wheel double wishbone suspension translates into tremendously agile

time you saw a lasted 17 years?

handling. The result is truly amazing. You can respond confidently to virtually any driving situation you might encounter.

This car also excels in safety. A driver's side airbag is now standard. Along with a sophisticated anti-lock braking system to help you control the car on rain-slick surfaces. You can actually steer the car while braking.

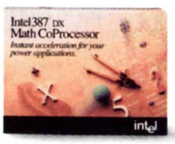
The body is also quite enviable. The lines are clean and smooth. The fit and finish are consistently excellent.

Not surprisingly, workmanship like this has produced a bounty of rewards. The Accord has been the best-selling car in the U.S. for the past three years. And it consistently appears on all the "top 10" lists for cars in its category.

Considering all these facts, one comes to a rather obvious conclusion. The Accord has always been, and continues to be, a true industry standard. The car by which all others in its class are measured and judged.

And while the competition remains relentless in their quest to catch us, we remain dedicated to building a car that's a few years ahead of them. Which means this chase might be just getting started, after all. The Accord EX. HONDA

It takes lots of screaming fans



Eleven million, to be exact. And in the math coprocessor business, that's definitely a record.

It just goes to show that Intel is the de facto standard in math coprocessors. And has been for the last ten years.

Not surprisingly, a whopping nine out of ten math coprocessors in use today carry the Intel name. And you can bet almost every one of them is sitting right next to an Intel CPU.

Get the connection? It's as perfect a match as you can get. Intel Math CoProcessors are 100% compatible with your Intel CPU, not to mention more than 2,100 software applications.

to go platinum eleven times.

It's all part of Intel's commitment to providing powerful solutions today, like the world's fastest microprocessors, plus the enhancements for even greater performance. Today and down the road.

WE MAKE COMPUTERS WORK HARDER.
Call and ask for packet BZ
1-800-538-3373

To receive a detailed information packet about the complete family of Intel Math CoProcessors, call 1-800-538-3373.

You'll see why Intel Math CoProcessors have topped the charts for ten years running.

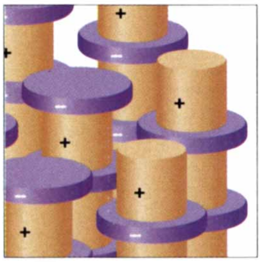
intel®

SCIENTIFIC AMERICAN
July 1992 Volume 267 Number 1

42

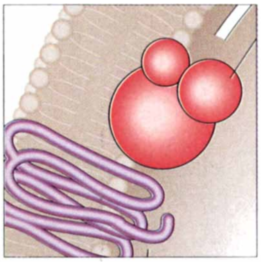

Origins of Western Environmentalism
Richard H. Grove

During the Age of Discovery, tropical islands became a powerful metaphor for the European idea of an untouched Eden. Those utopian images were soon shattered. In the 17th and 18th centuries, scientists employed by the companies exploiting colonial resources began to voice alarm over large-scale ecological changes. In doing so, they laid the foundation of modern environmentalism.

48


Building Molecular Crystals
Paul J. Fagan and Michael D. Ward

From snowflakes to semiconductors, the orderly lattice of crystals is governed by the size and shape of constituent molecules and the forces between them. Researchers have now begun to understand the conditions that influence the structure of crystals as they form. The achievement makes feasible the design of materials that have specific electronic, optical and magnetic properties.

56


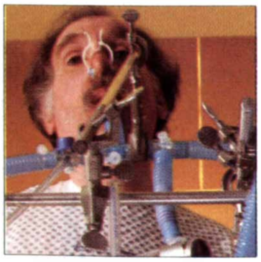
G Proteins
Maurine E. Linder and Alfred G. Gilman

Like messages in a game of "telephone," the signals between living cells are passed through a series of intermediaries. Critical among them is a class of substances attached to the inner surface of the cell membrane, called G proteins. They play a central role in many cellular activities, from vision to cognition. Malfunctioning G proteins have been implicated in diseases such as cholera and cancer.

66


Genetic Algorithms
John H. Holland

The consummate ability of species to adapt to an environment arose through natural selection. A group of computer programmers are emulating that process in the design of software. Programs based on genetic algorithms can evolve solutions to complex problems. They have demonstrated their practicability in designing jet turbines and controlling the flow in gas pipeline systems.

74


Breath Tests in Medicine
Michael Phillips

Since the time of Hippocrates, physicians have known that the odors of breath can convey vital information. With modern analysis and understanding of metabolic processes, breath tests can diagnose diseases of the stomach, intestine and pancreas. They can also monitor exposure to industrial chemicals. Standardized apparatus is needed before such tests take a place beside the x-ray for routine screening.

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017. Second-class postage paid at New York, N.Y., and at additional mailing offices. Authorized as second-class mail by the Post Office Department, Ottawa, Canada, and for payment of postage in cash. Canadian GST No. R 127387652. Subscription rates: one year $36 (outside U.S. and possessions add $11 per year for postage). Subscription inquiries: U.S. and Canada 800-333-1199; other 515-247-7631. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537.

80

Reproductive Strategies of Frogs
William E. Duellman

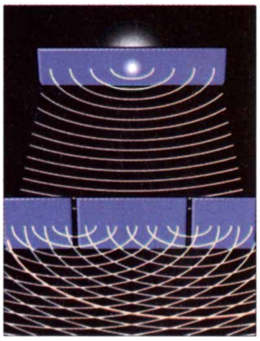
Frogs reproduce by laying eggs resembling those of fish. Yet frogs have managed to colonize niches throughout the terrestrial environment. To do so, they have developed a diversity of strategies that range from the improbable to the bizarre for ensuring that their eggs stay moist and that their young are nourished.

88

The Last Stone Ax Makers
Nicholas Toth, Desmond Clark and Giancarlo Ligabue

The village of Langda on the cloud-shrouded slopes of New Guinea's central cordillera provides a priceless glimpse of Stone Age technology. There, skilled craftsmen, who lived in complete isolation from the modern world until 1984, fashion stone axes that resemble those first made 20,000 years ago.

94



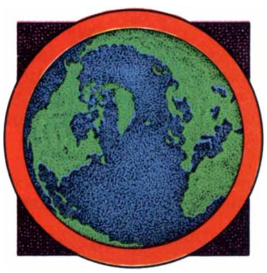
TRENDS IN PHYSICS

Quantum Philosophy
John Horgan, senior writer

The deeper physicists inquire into the mysterious world of quantum theory, the stranger it gets. New experiments continue to challenge the common notion of reality. Photons, neutrons, even objects large enough to be seen, lack form until they are observed. Observation can alter the outcome of experiments that have already occurred; measuring one entity can influence another far away.

DEPARTMENTS

17



Science and the Citizen

COVER STORY: **Ripples from the big bang....** The debate over dyslexia.... Subsea sonograms.... More trouble for the SSC.... Selfish genes.... Getting artificial intelligence up to speed.... PROFILE: *COBE* team leader George Smoot.

106

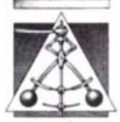
Science and Business

Are natural genes patentable?... Virtual ichthyology.... Clues to metastasis.... Improving electrorheological fluids.... Morphing the movies.... Biosensors.... THE ANALYTICAL ECONOMIST: What's the return on education?

10

Letters to the Editors
Picking bones.... The largest volcanic eruption.

14



50 and 100 Years Ago
1942: Physicists are the fair-haired boys of World War II.

114

The Amateur Scientist
A program that allows the best answer to evolve.

118

Book Review: *Stephen Jay Gould*
Retrying the monkey trial in a kangaroo court.

122

Essay: *Anne Eisenberg*
Sexism still lurks in the language of science.

THE COVER image shows the sky through the microwave-sensitive eyes of the *Cosmic Background Explorer* satellite. Orange indicates warmer than average regions; blue shows cooler than average ones. Temperature variations are thought to trace the uneven distribution of matter shortly after the formation of the universe (see "The Golden Age of Cosmology," by Corey S. Powell, page 17). Most of the features seen here are instrument artifacts, however. Primordial lumps could have seeded the complex tangle of galaxy clusters now observed.

THE ILLUSTRATIONS

Cover image courtesy of the National Aeronautics and Space Administration/Goddard Space Flight Center

Page	Source
43	*The Voyage of John Huygen Van Linschoten to the East Indies,* London, 1589
44	Johnny Johnson (*top*), courtesy G.H.C. Sprot (*bottom*)
45	Johnny Johnson (*top*), photograph by Nicolas Braunt/Alain François (*bottom*)
46	Public Records Office, London
47	Nicolas Braunt/Alain François
49	Joseph C. Calabrese and John J. Christy, Du Pont
50	Joseph C. Calabrese and John J. Christy (*left*), Jared Schneidman (*center* and *right*)
51	Layne Kennedy
52	Jared Schneidman; Joseph C. Calabrese and John J. Christy (*photo insets*)
53	Layne Kennedy
54	Joseph C. Calabrese and John J. Christy (*left*), Jared Schneidman (*right*)
56-59	Ian Worpole
60	Gabor Kiss
61	Ian Worpole
64	Johnny Johnson
65	Ian Worpole
67	Gordon Lansburg, Bruce Coleman, Inc.
68	Laurie Grace (*left*), David Dressler and Huntington Potter, Harvard Medical School (*right*)
69	Laurie Grace; Patricia J. Wynne (*inset*)

Page	Source
70	Laurie Grace
71	Laurie Grace (*top*), Johnny Johnson (*bottom*)
72	General Electric Corporation
75	Dan Wagner
76	Bettmann Archive (*top*), Laurie Grace (*bottom*)
77-78	Laurie Grace
79	Johnny Johnson
80-81	Juan M. Reijh, Animals/Animals
82-83	Patricia J. Wynne
84	Carol Hughes, Bruce Coleman, Inc. (*top left*), Michael Fogden, Animals/Animals (*top right*), Michael Fogden, Bruce Coleman, Inc. (*bottom*)
85	Michael Fogden, Bruce Coleman, Inc. (*top* and *bottom*)
86-87	Patricia J. Wynne
89	Christopher Burke, Quesada/Burke
90	Giancarlo Ligabue (*top*), Patricia J. Wynne (*bottom*)
91	Nicholas Toth (*top*), Patricia J. Wynne (*bottom*)
92	Viviano Domenici
93	Kathy Schick, University of Indiana (*left*), Rose A. Sevcik (*right*)
94-95	Andrew Christie
96-97	Boris Starosta
98	Robert Prochnow
99-100	Boris Starosta
101	John Pinderhughes
114-115	Johnny Johnson

SCIENTIFIC AMERICAN®

Established 1845

SCIENTIFIC AMERICAN, INC.

415 Madison Avenue
New York, NY 10017
(212) 754-0550

Protecting The Earth Doesn't Mean Going Without. It Means Making Better Use Of What We Have.

We're all concerned with finding ways to protect the environment. In the electric utility industry, we believe that the efficient use of energy should be at the top of the list. And we've focused our efforts on electric technologies that will help us achieve it.

Consider what has already been accomplished. Today, a microwave oven can cook a meal using 90% less energy than a gas oven. Refrigerators can now consume 80% to 90% less energy than those built 10 years ago. And fluorescent lights are 75% more efficient than common incandescents. All are examples of how electric technology enables us to get more out of each kilowatt hour, while reducing both fuel consumption and emissions.

Automated controls make it possible to optimize heating, ventilating and air conditioning systems.

Each day, further advances are being made that increase our efficient use of energy. For more information, call the Edison Electric Institute at 1-800-438-8334 or, in Washington, at (202) 508-5002. A cleaner environment is on everyone's agenda. Translating clean electric technologies into environmental solutions is at the top of ours.

ELECTRICITY
Taking you into the future.

ANNOUNCING THE RETURN OF THE PURE SPORTS CAR. This is the kind of sports car manufacturers stopped making years ago. The all-new 1993 Mazda RX-7 is specially designed for a handful of uncompromising individuals. Those who believe in the simple thrill that only comes from driving a lightweight car with a lot of power. What makes the RX-7 lightweight is also what makes it powerful. The world's only sequential twin-turbo rotary engine. More compact, and hundreds of pounds lighter than a comparable piston engine, it powers the RX-7 from 0-60 in 4.9 seconds and offers unique design

advantages including perfect weight distribution. It even inspires the car's fluid shape. No gimmicks. No gadgets. No compromises. The all-new Mazda RX-7. Unless you race, you've never felt anything like it.

THE 1993 MAZDA RX-7

Sequential twin-turbo rotary engine. 255 horsepower. Independent double-wishbone suspension. 11.6-inch, ventilated 4-wheel disc brakes with ABS. Torsen® limited-slip differential. 36-month/50,000-mile, no-deductible, "bumper-to-bumper" warranty. Free roadside assistance. See dealer for limited-warranty details. For a free RX-7 brochure call 1-800-639-1000.

LETTERS TO THE EDITORS

Big Bang Booster

Geoffrey Burbidge's essay "Why Only One Big Bang?" [SCIENTIFIC AMERICAN, February] argues that alternative theories are being suppressed and that strong evidence suggests the standard hot big bang model is wrong. Unfortunately, his methods resemble the attacks of creationists on the conspiracy of the "evolutionist establishment."

He says the most favored model of the big bang yields a universe between seven and 13 billion years old, then points to data suggesting an age of 15 billion years. His age choice assumes a high mass density. In fact, we do not know the mass density; if it is low, the age could be 20 billion years.

Burbidge maintains that a conflict between the smoothness of the microwave background and the clumpiness of galaxies is a challenge to the theory. This section of his essay is analogous to arguing that the theory of evolution is in serious doubt if we cannot explain in detail the death of the dinosaurs. The only cosmological models in question are the very simplest ones, which we naturally chose to examine first. Since his essay appeared, the *COBE* satellite has discovered fluctuations in the background radiation of one part in 100,000, confirming a key prediction of those simplest, most elegant models. Such essays no doubt stimulate interest, but they mislead the public.

ADRIAN L. MELOTT
Department of Physics and Astronomy
University of Kansas

Some Bones to Pick

Thank you for the fine article about "The Mammals of Island Europe," by Gerhard Storch [SCIENTIFIC AMERICAN, February]. It had one problem: it used the term "amateur paleontologists" in connection with the plunder of the site. Amateur paleontologists contribute collections and funds to museums, notify the professionals about where bones are exposed and continue to advance the science. "Commercial collectors" might have better fit the people whom Storch described.

RICHARD S. COHEN
Boulder, Colo.

I was amused by Storch's unkind slap at amateur paleontologists. His comment indicates what he thinks of free markets. Just recently in Los Angeles, a complete cave bear skeleton was sold. A beautiful, almost perfect saber-toothed skull was also sold for about $7,000. It is far better to have a free market for fossils than to keep them locked in some museum basement drawer. Besides, in the long run, good pieces usually end up in museums.

GEORGE H. WILLIAMS
Universal City, Calif.

Storch replies:

Amateur paleontologists do make essential contributions: I have named new mammalian species after amateurs to honor their work. I agree that "commercial collectors" would have been a better description. Commercial collectors at Messel damaged excavation sites, removed fossils from their stratigraphic settings and sold specimens at unreasonably high prices. Those acts meant bidding farewell forever to most of these fossils, including rare or even undescribed taxa.

Please consider, Mr. Williams, that it may be the 100th specimen of a particular mammal that reveals its embryos or its complete inner ears or other informative features for the first time. The Senckenberg Research Institute and Museum of Natural History is in charge of the public display of Messel fossils, which now belong to the state of Hesse. They are not buried "in some museum drawer" but accessible to anyone who is seriously interested.

Earthshaking Myths

Perhaps the June 15, 1991, eruption of Mount Pinatubo was near first in sulfur release for this century, but it was far from being the largest eruption, as mentioned in "Volcanic Disruption," by John Horgan ["Science and the Citizen," SCIENTIFIC AMERICAN, March]. Beginning on June 6, 1912, the Novarupta Vent near Mount Katmai on the Alaska Peninsula produced about 12 cubic kilometers of magma. It formed the still desolate Valley of Ten Thousand Smokes in what is now Katmai National Park. Mount Pinatubo vented only three to five cubic kilometers of magma, which makes it only the third or fourth largest eruption of the century.

JOHN C. EICHELBERGER
Geophysical Institute
University of Alaska at Fairbanks

Showdown on Body Armor

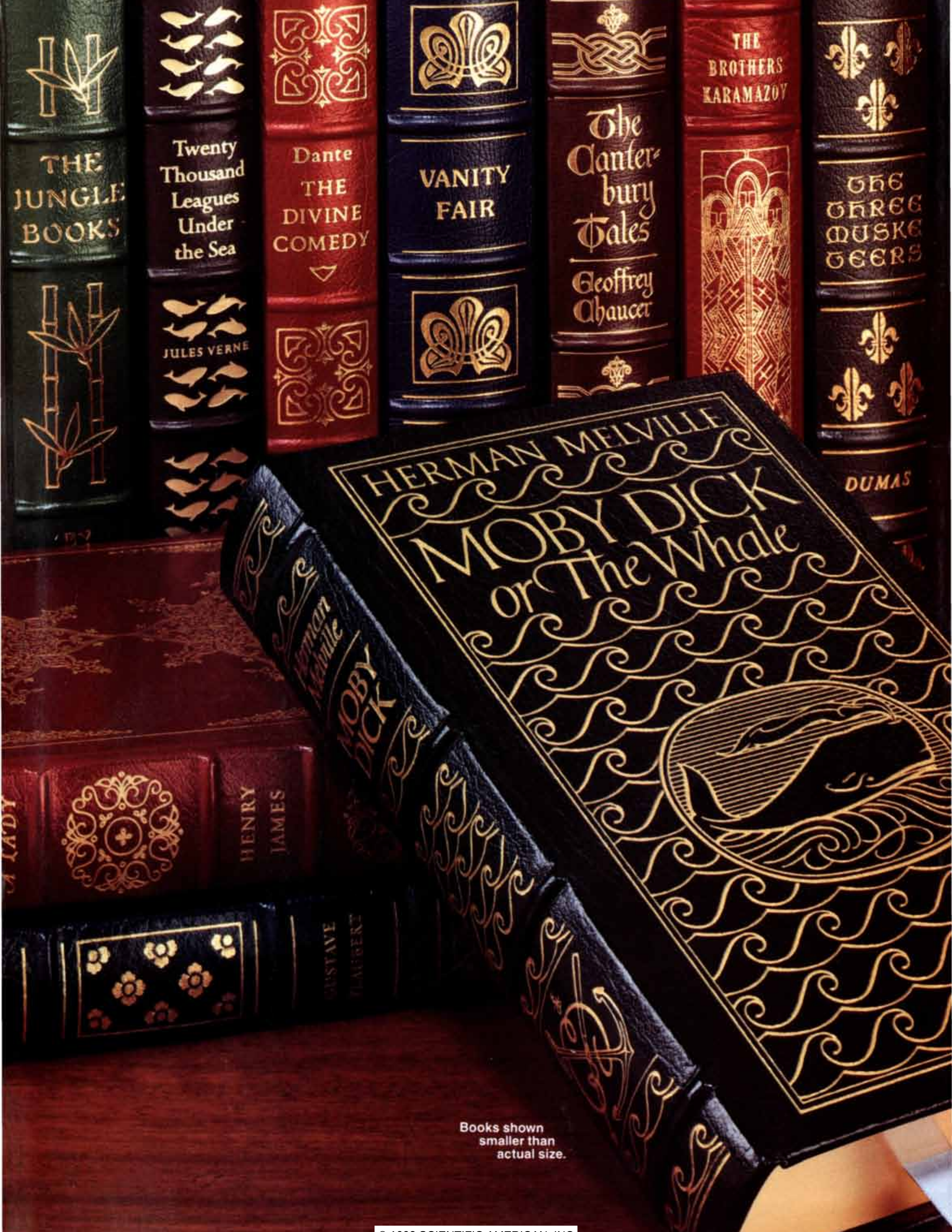
Tim Beardsley badly misinterpreted the issues in "Vested Interests" ["Science and Business," SCIENTIFIC AMERICAN, March]. He presented the fight against the flawed National Institute of Justice (NIJ) soft body-armor test as an attempt by Du Pont to protect the Kevlar market. If this were so, Du Pont should favor the NIJ test since any vest can be made to pass it simply by adding more Kevlar. Du Pont, however, wants the vests to be worn—which is less likely for vests made unnecessarily uncomfortable by excess Kevlar.

During the past two decades, about 600 bullets striking police officers have been stopped by their protective vests. Nearly 40 percent of the officers killed by gunfire in the past decade would have been saved if they had been wearing their vests when they were shot. No vest has ever failed against a bullet it was rated to stop. The NIJ test therefore proves itself invalid by failing 50 percent of the vests that have a perfect safety record.

MARTIN L. FACKLER
President, International Wound Ballistics Association
Pinole, Calif.

Beardsley is to be congratulated for his clear and succinct description of the status of this "controversy." Before retiring as director of science and technology of the NIJ, I was responsible for the development of Kevlar into modern soft body armor. I was also responsible for managing the Law Enforcement Standards Laboratory at the National Institute of Standards and Technology (formerly the National Bureau of Standards), which developed all the NIJ standards for police equipment. I did not find any errors of fact in Beardsley's article, and I wish to commend him for fairly stating the current situation.

LESTER D. SHUBIN
Fairfax, Va.


THE JUNGLE BOOKS
Twenty Thousand Leagues Under the Sea
JULES VERNE
Dante
THE DIVINE COMEDY
VANITY FAIR
The Canterbury Tales
Geoffrey Chaucer
THE BROTHERS KARAMAZOV
THE THREE MUSKETEERS
DUMAS
HERMAN MELVILLE
MOBY DICK or The Whale
Herman Melville
MOBY DICK
HENRY JAMES
GUSTAVE FLAUBERT
Books shown smaller than actual size.

50 AND 100 YEARS AGO

JULY 1942

"The great significance of polymerization is not so much that we have at last learned how to make these synthentic elastomers, important as that may be in these grave days; it is that we have mastered the secrets of finding the proper introducers, of building endless molecules to *our own specifications*. We have just made a start, and at present we are still trying largely to imitate the qualities of natural rubber."

"As most middle-aged and older readers will recall, the chemists were the fair-haired boys in World War I. Today, in World War II, the fair-haired are the physicists—exalted to the level of a kind of war nobility. In the systematic search for the ablest physicist workers, the personnel of physics has been combed over again and again, and now the nation is to be fine-tooth-combed for physicists of lesser degree clear down to the rank of high-school graduates and even below that. Officers and faculty members of high schools are being officially urged to single out youths who have shown innate aptitude for physics and its working tool, mathematics, and to direct their interests more actively into these channels."

"There is at all times a negative electric charge distributed over the surface of the earth and a positive charge in the outer regions of the atmosphere. These two sets of charges produce an electric field in the air directed vertically toward the earth. During fair weather this field is about 30 volts per foot. The air has a low conductivity and the presence of the electric field produces a leakage current, tending to discharge the earth's negative charge. If it were not for some restoring agent the earth would be completely discharged in five or six minutes. Lightning plays the fundamental role of returning this charge to the earth and maintaining it in its normal charged condition."

SCIENTIFIC AMERICAN

JULY 1892

"Professor Loeffler, the originator of the system of destroying field mice by typhus bacillus infection, has returned to Germany from Greece, where he had gone to put his system to a practical test. The professor reports that his mission has been a complete success, and that within eight or nine days the swarms of field mice which infested the parts of the country visited by him, and destroyed the crops, were absolutely annihilated. The peasants were asked to meet at a given point with baskets of odd pieces of bread broken small. This bread was soaked in the solution containing typhus bacilli, and returned to the owners with instructions to spread it in the fields. Pieces of bread saturated with the bacillus were eaten by Dr. Loeffler and his assistants to demonstrate its harmlessness."

"At a recent meeting of the Academy of Sciences, Paris, much interest was created by Dr. Brown-Sequard's paper on his system of treatment with injections of a solution of sperm fluid. The *savant*'s address bristled with proofs of its efficacy. He instanced a patient of 80, living at Mauritius, who was restored from a paralytic and atonic state to health and vigor. Another somewhat younger man, bed-ridden, and regarded as moribund, was re-established sufficiently in a fortnight to take severe horse exercise, 'and now,' drily added the doctor, 'his health improves so rapidly that the family have forbidden the medical man to continue the treatment.' Dr. Brown-Sequard claims that the 20,000 injections made by him during the last three years have been invariably successful."

Wellington's chair

"Professor Dewar's friend, Mr. McKendrick, had tried the effect of low temperatures upon the spores of microbe organisms, by submitting putrefied blood, milk, and such like substances for one hour to a temperature of –182° C.; they afterward went on putrefying. Seeds, also, withstood the action of a similar amount of cold. He thought, therefore, that the experiments had proved that the idea of Lord Kelvin uttered some years ago was possibly true, when he suggested that the first life might have been brought to the earth upon a seed-bearing meteorite."

"A pound of coal used to make steam for a fairly efficient refrigerating machine can produce an actual cooling effect equal to that produced by the melting of 16 to 46 pounds of ice. These figures are sufficient to prove the practicability of artificial cooling for office buildings, hospitals, theaters, hotels, and even for the best class of private houses. It is a curious example of the slowness with which people take advantage of modern inventions that thousands of men sit sweltering in hot offices in the midsummer days; business lags and the efficiency of workers is greatly reduced. At the same time not more than three or four blocks away are great provision warehouses where the temperature is kept at freezing the year round. If it pays to keep dead ducks and turkeys cool on Greenwich Street, why would it not pay to keep live business men cool on Broadway?"

"In the chair shown in the illustration, the occupant is supposed to sit crosswise of the seat, as one would sit in a saddle. Such a chair in a library or study, affording convenient opportunity for such changes of position from the usual posture as are often sought, cannot fail in many cases to contribute materially to one's comfort. The picture is very nearly a representation of a chair used for many years by the Duke of Wellington, at Walmer Castle, England."

WIZARD®

ORGANIZATIONAL POWER: PURE AND SIMPLE.

The Wizard® OZ-8000 has all the power you need to stay on time, on track and in control. Eleven built-in functions and a raised typewriter keyboard make it easy to store, retrieve and sort phone numbers, important dates, your schedule, memos and more. You don't need to enter DOS commands. Just touch any function key and the Wizard displays the information you want.

EASY TO EXPAND.

Insert any optional Wizard software card for easy access to more power. The Wizard's patented touch-screen turns each card into a customized keypad. Choose from reference, business, entertainment, telecommunications and memory expansion cards that are as easy to use as the Wizard itself.

EASY TO EXCHANGE DATA.

With the optional Organizer-Link and cable, you can load files from your PC into your Wizard, or download work done on your Wizard to your PC.

EASY TO COMMUNICATE.

An Organizer Fax/Modem gives you the power to send a fax directly from your Wizard. Or, connect with on-line services to send and receive data over direct phone lines and even from a cellular phone. To learn how easy it is to organize your business and personal life, call 1-800-321-8877. In Canada, call 1-416-890-2100. And ask for the dealer nearest you.

SHARP®

FROM SHARP MINDS COME SHARP PRODUCTS™

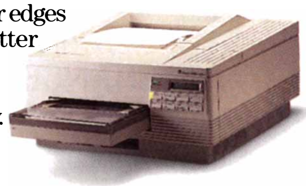
"I haven't taken a LaserJet in for repairs once in seven years."

—Clyde Alpert, ARL Services Inc., Chicago, IL

"I repair other printers once, twice, maybe three times a year," says Clyde Alpert. "But I never have to take the LaserJets in. We've pumped thousands and thousands of copies through them, and they've never failed us."

In addition to great reliability, HP also offers exceptional customer support. For example, the LaserJet Customer Assist Line is staffed with experts to advise you after the sale. They'll help you take full advantage of all the HP LaserJet printers' unique features and answer your hardware and software questions.

With five award-winning models, ranging in price from $1,249 to $5,495,* HP offers an unbeatable selection. You can also choose genuine PCL5 printer language, with scalable typefaces. As well as HP's exclusive Resolution Enhancement technology, which creates sharper edges and better overall print quality.



To hear more about what other customers have to say about HP LaserJet printers, remember to call **1-800-752-0900, Ext. 2986** for a free video. Then forget about repairs.

HP LaserJet Printers.

HEWLETT PACKARD

*Suggested U.S. list prices.

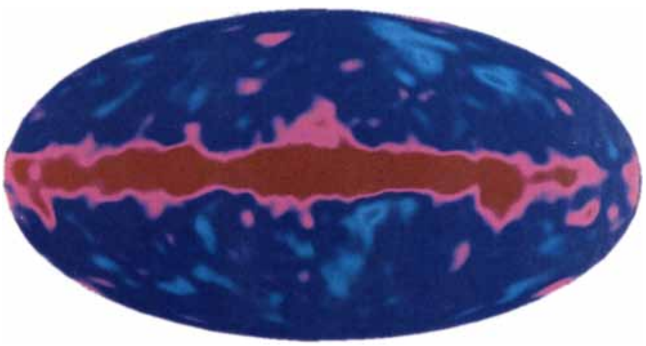
The Golden Age of Cosmology

On April 23, George F. Smoot of Lawrence Berkeley Laboratory and his collaborators added another chapter to the scientific version of the story of creation. As camera lights glared and shutters whirred, they announced that an instrument on the National Aeronautics and Space Administration's *Cosmic Background Explorer* (*COBE*) satellite had detected relics of lumpy structures that existed 300,000 years after the universe was born in a vast explosion, the big bang, some 10 to 15 billion years ago. In essence, *COBE* may have found the primordial blueprint that determined the structure of the modern universe. "When fossils were found in the rocks, it made the origin of species absolutely clear-cut," observes Jeremiah P. Ostriker of Princeton University. "Well, *COBE* found its fossils."

The fossils unearthed by *COBE* consist of variations in the faint microwave glow that permeates the universe. Those variations are thought to trace ancient ripples that could have seeded the formation of galaxies and clusters of galaxies. In a single swoop the *COBE* findings gave important support to the big bang but simultaneously relegated most of cosmologists' specific models for the formation of the universe to the trash bin. The toll would have been higher if *COBE* had failed to find the telltale fluctuations. "A null result would kill off the whole present crop of theories," says Philip Lubin of the University of California at Santa Barbara, who has collaborated on *COBE*.

The satellite, which was launched late in 1989, was conceived nearly two decades ago by scientists at NASA's Goddard Space Flight Center to provide precise measurements of the background radiation. One set of instruments on *COBE*, known as the Differential Microwave Radiometers, has spent the past two years searching for any deviations from uniformity in the microwave sky. Rumors of the results had been circulating through the astronomical community for months before Smoot and his co-workers summarized their findings at the annual meeting of the American Physical Society. "Lots of people on the *COBE* team were walking around grinning like Cheshire cats," reports John C. Mather of the Goddard Space Flight Center, project scientist for the satellite. The formal announcement of *COBE*'s findings touched off a new round of scientific debate. Are the results real? What do they mean for cosmology?

The microwave background has been a linchpin of cosmological theory since it was discovered by Arno A. Penzias and Robert W. Wilson of Bell Telephone Laboratories in 1964. At first they were puzzled by the mysterious microwave noise that seemed to enter their antenna from all directions. Only when they spoke to Robert H. Dicke of Princeton did they learn that the big bang theory predicts just such global distribution of microwaves.



COBE *MAP shows that our galaxy* (**horizontal band**) *dominates the microwave sky. Such local signals were subtracted from the map on the cover. Image: Lawrence Berkeley Laboratory.*

Theorists believe this radiation is a relic from the time of the big bang, when the universe consisted of an exceedingly hot, opaque soup of charged particles and radiation. When the temperature of the universe dropped to 3,000 kelvins (300,000 years post–big bang), protons and electrons combined to form transparent, neutral hydrogen. At that time, matter and radiation went their separate ways. In the 15 billion years of expansion since, the radiation has cooled to an effective temperature of 2.74 kelvins, so cold that every element except helium would freeze solid.

"I was not immediately sold on the cosmology," Wilson recalls. But a fruitless year of searching for other sources of the background noise finally convinced him. The microwave background remains one of the prime pieces of evidence in favor of the theory that the universe began as an unimaginably hot, compact mass.

One of *COBE*'s top goals was to test the theory by measuring the spectrum of the microwave background. The simplest version of the big bang predicts that radiation from the hot early universe should describe a smooth spectral curve, known as a blackbody spectrum. In January 1990 Mather and his fellow team members released *COBE*'s microwave spectrum at a meeting of the American Astronomical Society. The chart that Mather tossed on the overhead projector looked so much like a perfect blackbody curve that it could have been lifted from a physics textbook. The astronomers in attendance broke out in spontaneous applause.

The recent *COBE* data are helping solve an even knottier problem. For the past 28 years, cosmologists have been looking, without success, for variations in the radiation produced by irregularities in the big bang. At the same time, astronomers have developed an increased appreciation of the scale and complexity of cosmic structure. Surveys of vast swaths of space reveal galaxies organized into vast sheets and filaments. The largest formations stretch about 300 million light-years across, roughly 3 percent of the radius of the visible universe. Observers have also turned up galaxies and quasars so distant that they must have formed only about one billion years after the big bang.

To produce such large and distant objects, clumpiness in the early cosmos need not have been terribly pronounced. Gravity amplifies any variations: regions that are slightly denser than average tend to pull together, becoming increasingly concentrated at the expense of regions that are comparatively rarefied. Nevertheless, some structure must have been present right at the start for gravitational amplification to get under way.

In places where matter piled up, the local tug of gravity would steal energy from outward-bound photons (individual particles of electromagnetic radiation), so background radiation that originated in that region would appear relatively cool. Low-density voids should look warm for the same reason. Nascent structure therefore should show up as patches of varying temperature in the microwave sky.

Much to the embarrassment of big bang boosters, increasingly sensitive studies of the microwave background continued to show a completely uniform glow of radiation. Theorists obligingly adjusted their models to accommodate ever smaller initial density fluctuations. But they did so knowing that, at some point, the absence of detectable temperature variations would doom the big bang cosmology. *COBE*'s precision instruments seem to have come to the rescue. The detected fluctuations differ from the average temperature of the sky by only 30 millionths of a kelvin—near the limit of *COBE*'s sensitivity. When he was asked about his confidence in the results, Smoot smiled and said that he would stake his personal reputation on the fact that they are real.

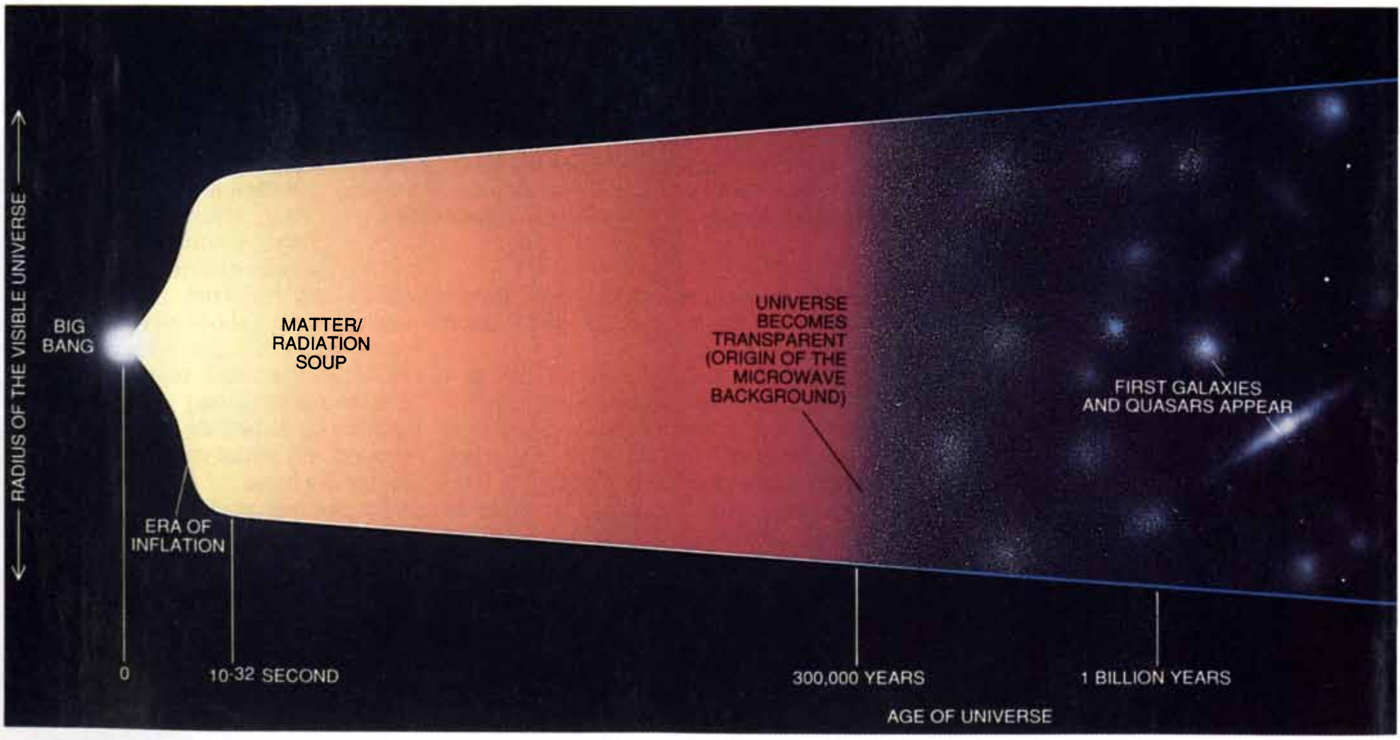
In this case, "real" is a somewhat blurry term. *COBE*'s map of the microwave sky is dominated by instrument noise; roughly two thirds of the data shown on the map originated in *COBE* or in unaccounted-for nearby sources and not in the infant universe. Some of the fluctuations shown on the map represent genuine signals, but *COBE* scientists cannot as yet say which ones. "I can't emphasize strongly enough that you cannot look at any one point and say, 'That's a cosmic fluctuation,' " cautions Charles L. Bennett of the Goddard Space Flight Center, the deputy principal investigator of *COBE*'s microwave radiometers. Only by applying mathematical analysis techniques, such as statistical averaging, can one prove that some of the patches are not instrument artifacts, he explains.

The reason for the ambiguity lies in the Herculean task of accounting for every source of microwave emission other than the cosmic background. Our galaxy, the Milky Way, radiates copious microwaves that Bennett and his coworkers subtracted to reveal the background signal. Other galaxies also emit microwaves, and hot gas clouds in clusters of galaxies introduce their own, minuscule temperature fluctuations.

Such noncosmological signals should show up on small angular scales, but *COBE* looks at large swaths of sky, seven degrees wide (200 times the apparent area of the moon). Individual galaxies should average out in the *COBE* observations. Edward L. Wright of the University of California at Los Angeles checked to see if the signals seen by *COBE* matched up with the locations of known, comparatively nearby galaxy clusters and found no correspondence. Even so, some astronomers remain slightly skeptical. John P. Huchra of the Harvard-Smithsonian Center for Astrophysics, playing devil's advocate, suggests that the microwave fluctuations could be produced by a previously unknown class of nearby astronomical objects and not by density variations shortly after the big bang.

For now, though, most cosmologists are taking the *COBE* results at face value and buckling down to understand their implications. The most obvious losers are theories in which dense, energetic field defects (such as domain walls or textures) served as the nuclei around which galaxies formed. Such theories predicted the existence of a few particularly intense temperature fluctuations, which *COBE* does not see. David N. Spergel of Princeton, who has worked on texture-based cosmologies, took the news rather well. "Oh well, textures are dead," he said cheerfully. "Cosmology is like that."

The biggest beneficiary so far from the *COBE* results has been inflationary cosmology, a highly popular elaboration of the big bang developed in the early 1980s by Alan H. Guth of the Massachusetts Institute of Technology and refined by Andrei D. Linde of Stanford University and by Paul Steinhardt of the University of Pennsylvania. In this model the universe passed through a brief phase very soon after its birth during which the natural forces behaved very differently from the way they do now, so that gravity effectively became a re-



COBE *INVESTIGATOR George F. Smoot tells the world that "the big bang is alive and well... very well." Photo: AP/World Wide Photos.*

pulsive force. As a consequence, in a tiny fraction of a second (10^{-32} second or so) the universe expanded by a mind-boggling factor of 10^{30} or more.

Inflation appeals to many cosmologists because it explains two difficult cosmic puzzles: why the microwave background has virtually the same temperature in all directions and why the density of the universe is close to the value needed eventually to halt its expansion. During the era of inflation, local temperature variations expanded and smoothed out; at the same time, the geometry of space became flat (a characteristic of critical density), regardless of its initial shape. Inflation also provides a mechanism for creating cosmic structure. In those first moments, tiny quantum energy oscillations would have expanded along with the rest of the universe to become macroscopic regions of comparatively high density. These overdense regions could have evolved into clusters and superclusters of galaxies.

Inflation models predict that the overdense regions should have a scale-invariant (Gaussian) distribution—that is, the pattern of variation should look qualitatively identical on all different angular scales. *COBE* sees just such a distribution of hot and cold spots on the sky. Guth is quick to note, however, that although *COBE*'s results are consistent with inflation, they do not prove it.

Theorists have widely diverging views on the viability of other cosmological theories. In an upcoming paper in *Astrophysical Journal Letters,* Wright and his co-authors single out a class of theories known as isocurvature models as being eliminated by *COBE*. "Oh, not at all!" responds P. James E. Peebles of Princeton when asked if his work on such models has been rendered obsolete. Several cosmologists comment that theories producing structure via cosmic strings, ultradense serpentine field defects related to textures, could match the *COBE* observations. Ostriker strongly doubts that strings have fared any better than textures. "David Spergel was just more honest," he says.

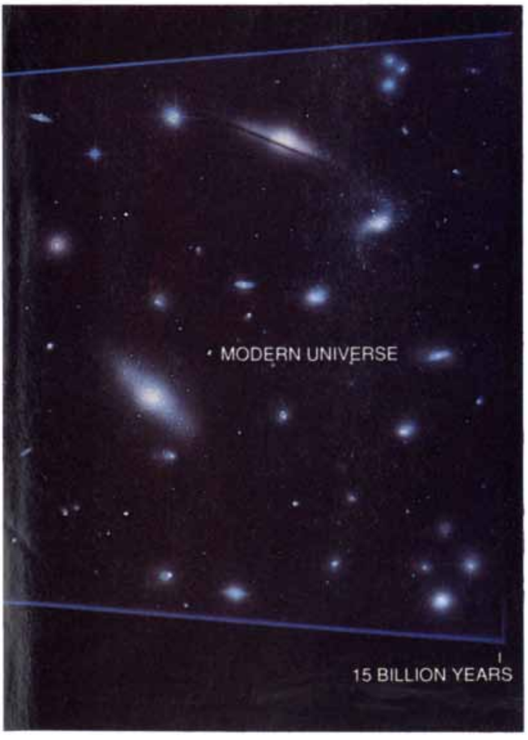
Regardless of the particular cosmological model one chooses, the *COBE* maps raise a perplexing problem. The temperature fluctuations are minuscule, only about one part in 100,000. Other experiments have shown that the fluctuations on smaller angular scales are also extremely slight. Such slight variations could not easily have produced dense, highly organized galaxies within a billion years or two after the big bang. Ordinary matter could not begin accumulating into condensed structures until the universe became transparent to radiation. Even then, the radiation would have kept matter too hot to coalesce quickly into galaxies.

In recent years, cosmologists have proposed that the universe contains a second, invisible component—cold dark matter—that clumps together far more readily. The existence of dark matter sits happily with inflationary cosmology because such models predict that the universe is exactly dense enough to halt the present expansion. Luminous material visible to astronomers provides only about 1 percent of that critical density.

In order for the dark matter to help form galaxies, it must consist of particles that do not interact with electromagnetic radiation, referred to as nonbaryonic dark matter. Protons, neutrons and electrons are ruled out. Theorists have proposed a seemingly endless roster of candidate particles. One by one, high-energy physics experiments have been ruling them out; none has yet been confirmed.

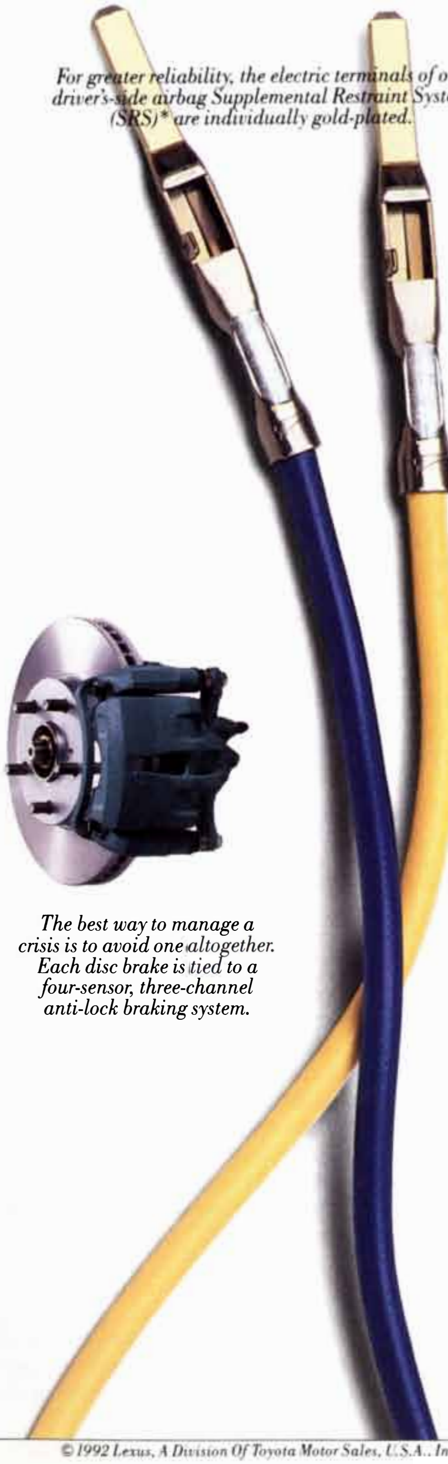
Cold dark matter has other problems as well. Some cosmological models incorporating cold dark matter can account for the existence of large clusters and superclusters of galaxies. Others can explain the formation of individual galaxies. None can do both. Some researchers therefore suspect that cold dark matter may turn out to be a chimera. "The physicists have been enormously reluctant to accept what astronomy shows them," says Penzias, who argues in favor of a less dense universe containing only ordinary matter. "Cold dark matter is dead," Peebles agrees.

On the other hand, the inflationary cosmologies that answer several of the most fundamental questions about the



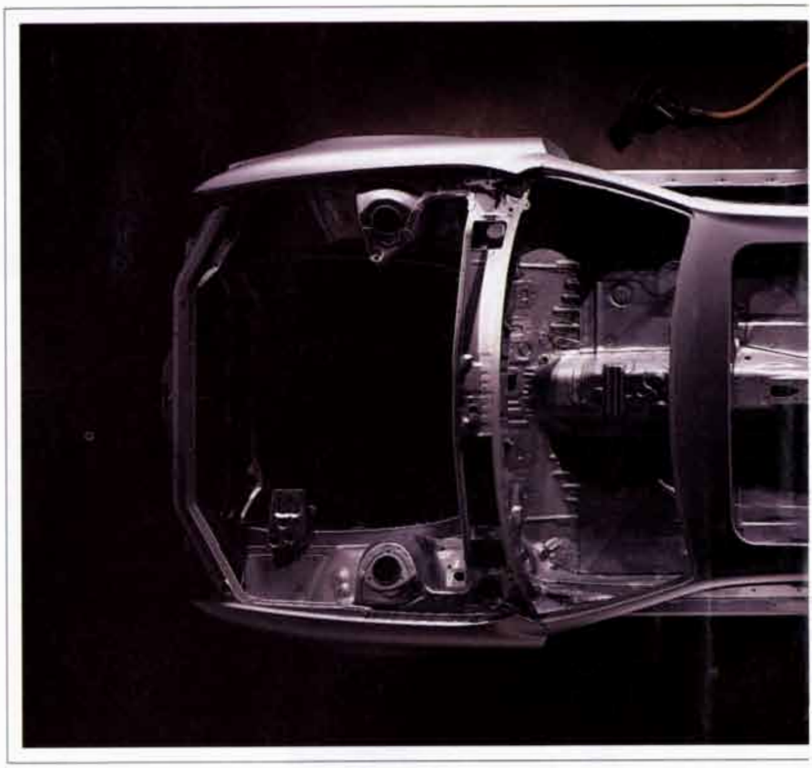
COSMIC TIMELINE depicts a stylized outline of the inflationary big bang model. The microwave background offers the earliest possible glimpse at the primeval structures that eventually evolved into all the objects visible in the night sky.

Our Idea Of Cri

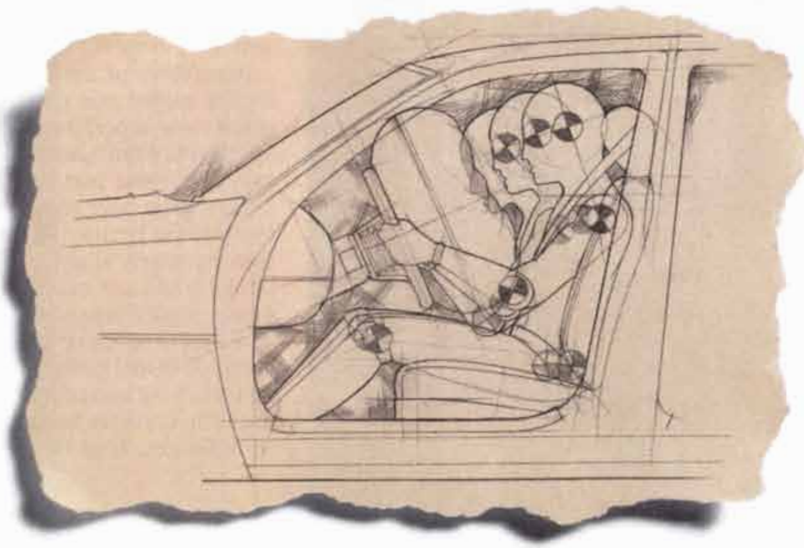


For greater reliability, the electric terminals of our driver's-side airbag Supplemental Restraint System (SRS) are individually gold-plated.*

The best way to manage a crisis is to avoid one altogether. Each disc brake is tied to a four-sensor, three-channel anti-lock braking system.



Major chassis elements are laser-welded before assembly. For even greater strength, they're



Every Lexus comes with a fifth shock absorber: a driver's-side airbag (SRS) activated by five impact sensors.*

sis Management.

made with high-tensile-strength steel.

The front and rear crumple zones are designed to progressively give upon impact, rather than directly pass the energy on to the driver and passengers.

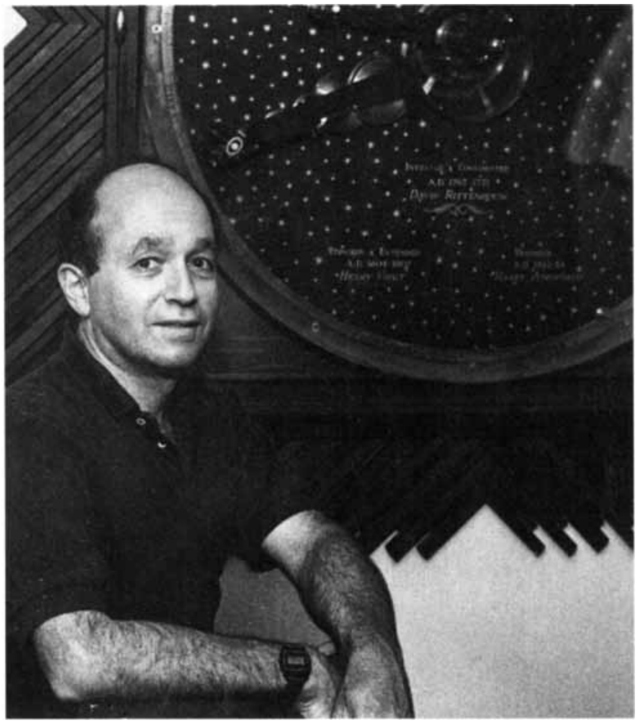
Any collision powerful enough to activate the airbag may deactivate the battery it's connected to. The answer? A back-up power source.

What if the driver in front of your car suddenly brakes hard? The outcome of this hypothetical scenario may depend on the caliber of car you're driving. If you're behind the wheel of a Lexus ES300, for example, you can be confident that the car will handle superbly; between your skill, the car's prowess and a little bit of luck, hopefully you'll avoid this accident. But Lexus has spent years preparing for those occasions when luck is not with you. In the event of an accident, the ES300 is equipped with several safety features, all aimed at reducing the risk of injury.

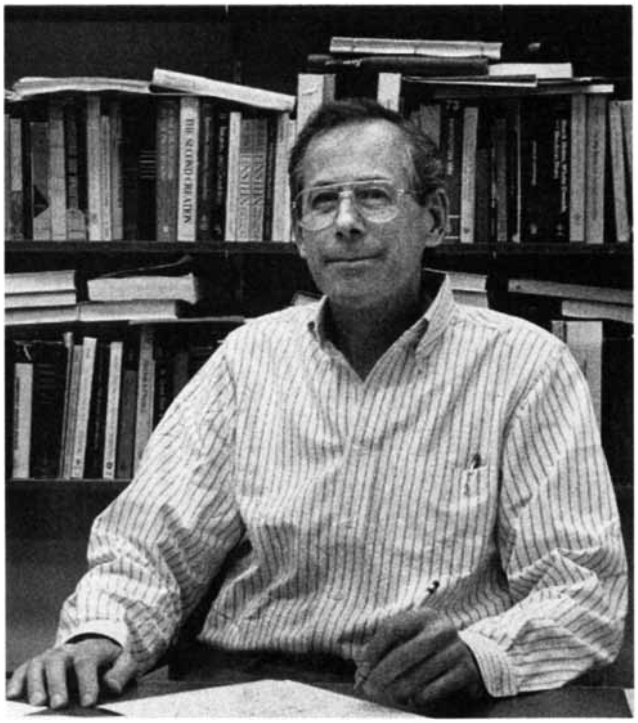
"Reducing," however, is not the same as "eliminating." The potential for a crisis is and always will be there. But while you may not think about the possibility of getting into an accident every time you're in a Lexus automobile, you can be certain of one thing: we have.

LEXUS
The Relentless Pursuit Of Perfection.

driver's-side Supplemental Restraint System (SRS) will inflate only in a severefrontal impact. In a moderate collision, the three point seat belt provides primary protection; the airbag will not inflate.



JEREMIAH P. OSTRIKER hails* COBE*'s findings but warns against mistaking cosmology for theology. Photo: J. Levinson.



P. JAMES E. PEEBLES perceives a need for fresh ideas on the origin of cosmic structure. Photo: R. Matthews.

big bang absolutely require that much of the dark matter be nonbaryonic. (A dense universe consisting only of ordinary matter would have acquired a composition quite unlike the one observed.) Studies of the large-scale flows of galaxies also point toward a universe dominated by nonbaryonic dark matter. And if theorists abandon their mystery particles, they must conjure up another mechanism that would cause galaxies to form rapidly.

Huchra offers "two words that cause theorists to quake in their boots: magnetic fields." Present cosmological models do not take magnetic fields into account. Although few researchers believe magnetism had much to do with galaxy formation, a vocal minority, led by Swedish Nobel laureate Hannes Alfvén, continues to argue otherwise. J. Anthony Tyson of AT&T Bell Laboratories, who has extensively mapped dark matter around galaxy clusters, suspects that cosmic structure originated via "some other pathological process, perhaps black holes." And there is always the possibility that some completely new mechanism remains to be discovered.

Listening to the present round of disputes, one may find it hard to recall just how far cosmological theories have advanced since the discovery of the microwave background less than three decades ago. Penzias recollects that when he started out, there were no data at all to test the theories. Smoot struck a similar tone at the April press conference as he gloated again and again, "This is the golden age of cosmology."

The fate of the current crop of cosmological notions will hinge on more detailed measurements of the microwave background. Because of its huge field of view, the smallest features seen by *COBE* in the early universe correspond to features that exceed in size the largest amalgamations of galaxies seen in astronomers' deep-sky surveys. Scrutinizing the microwave sky for smaller (arc-minute size) features will yield more direct information on how the present galaxies and galaxy clusters originated. In fact, such studies have been under way for a number of years now, but astronomers have yet to see any unevenness in the temperature of the sky. "It's not a real comfortable situation," Lubin admits, "but it's not yet a head-to-head conflict with *COBE*."

Now that *COBE* has shown the magnitude of the temperature variations, other researchers will have a better sense of how sensitive their experiments need to be. Lubin and his colleagues are analyzing data from a balloon-launched microwave telescope and preparing for another flight this fall. "We really ought to be seeing something soon," he says, although "it's a no-lose situation: if *COBE* is wrong and there are no fluctuations, that would be even more exciting."

Meanwhile the *COBE* scientists are not resting on their laurels. Wright promises that by the time *COBE* completes its mission in 1994 the microwave measurements will contain far less noise, so that it will finally be possible to point to a feature on the map and say, with certainty, "That is a cosmic fluctuation."

How cosmology will look after the impending shakedown is anyone's guess. "Every generation thinks it has the answers, and every generation is humbled by nature," Lubin reflects. Even as *COBE*'s measurements seemed to reinforce the reigning inflation-plus-cold-dark-matter scenario, cracks have been forming in the intellectual edifice. Recent observations of the rate at which the universe expands hint that inflationary models may paradoxically require the universe to be younger than the objects it contains. Cosmologists are considering adding a second dark matter component (called, logically enough, hot dark matter) to shore up their models of structure formation.

Almost nobody doubts the validity of the big bang itself. On the other hand, no one knows how to make a direct connection between the conditions prevailing at the time of cosmic origin and those prevailing today. Ostriker chastises those who "treat this [cosmology] like theology that you have to believe. This is science!" The universe has spent billions of years writing the story of creation. Humans will no doubt be trying to read it for quite some time to come. *—Corey S. Powell*

Compulsive Canines

Dogs that can't stop grooming provide clues to obsessions

Jonathan Swift conceived such a horror of human ordure that he could no longer tolerate the presence of people, not even his own family. Howard Hughes, overwhelmed by the increasing demands of his own purification rituals, finally gave up grooming altogether and became a helpless recluse. Such mania for cleanliness constitutes the most common expression of obsessive-compulsive disorder (OCD), an often incapacitating problem that afflicts an estimated four to six million people in the U.S. alone.

In this affliction, man has company. Some dogs also can't stop washing themselves. Moreover, they respond to the drugs that help OCD patients, and they apparently inherit the disorder. Judith L. Rapoport, a neurologist at the National Institute of Mental Health, says the canine syndrome—called acral lick—should make a good model for OCD. Such a model might lead to new treatments and illuminate the evolution of grooming and other presumably primitive behaviors.

Rapoport unwittingly summoned her pooch patients through her popular book on OCD, *The Boy Who Couldn't Stop Washing.* Owners of dogs that had licked their paws to the point of ulceration came to her in desperation, after other cures had failed. "The vet told the first one that the acral lick came from loneliness," Rapoport says. "So the owner brought her dog to the office with her every day. Now she had an ecstatically happy dog—but it went on licking."

Rapoport prescribed clomipramine, one of the three known antiobsessional drugs, and the dog stopped licking. To prove that clomipramine had worked purely against the compulsion, without any coincident antidepressant effect, Rapoport recruited 42 dogs with acral lick and put them on either antiobsessional drugs or conventional antidepressants. The antiobsessionals worked; the antidepressants did not.

Antiobsessional drugs work by preventing nerve cells from reabsorbing serotonin, a neurotransmitter that prevents them from firing again. No one knows how this process ameliorates OCD, although imaging techniques locate the action in the brain's basal ganglia, which hardly differs between humans and other mammals.

A dog model might provide some answers. First, it should help workers find any genes that predispose to acral lick. "The animal model, like the human syndromes, seems a straightforward inheritance," Rapoport says. "All you have to do is look for dog families of three generations with the syndrome and compare DNA of dogs with and without the syndrome, and the geneticists will tell you which sequences they have in common—even if you don't know which chromosome they're on."

Compulsions may also lie at the root of animal behavior. One theory, which Rapoport champions, treats them as hard-wired subroutines: little computer programs left over from the primitive common ancestor of dogs and humans. "These behaviors seem phylogenetically old," she says. In her view, the compulsion to wash may be a grooming program gone wild. The compulsion to check that doors are locked, pictures straightened and ovens turned off may stem from an ancestral need to check the cave for predators and other dangers.

Galling the Bears

Pity the bears in Asia. Not only are their paws valued as food, but their gallbladders are one of the most valuable medicines in the traditional pharmacopoeia. According to Judy A. Mills and Christopher Servheen of the University of Montana, bear gallbladders sell in South Korea for several thousand dollars. As a result, they say, Asian bears are declining and "cannot maintain viable wild populations," and evidence is mounting that North American bears are being hunted to satisfy Asian demand.

The gallbladders and the bile are used in many Asian countries to treat liver disease. Unlike rhinoceros horn, a traditional medicine with no known pharmacological value, the gallbladders do—unfortunately for bears—contain an active substance, ursodiol. Under the name of Actigall, a synthetic version is sold in the U.S. by Ciba-Geigy as a treatment for gallstones. And a study published last year in the *New England Journal of Medicine* established ursodiol as an effective therapy for primary biliary cirrhosis. In Asia, the substance is widely used as a prophylactic against hangovers.

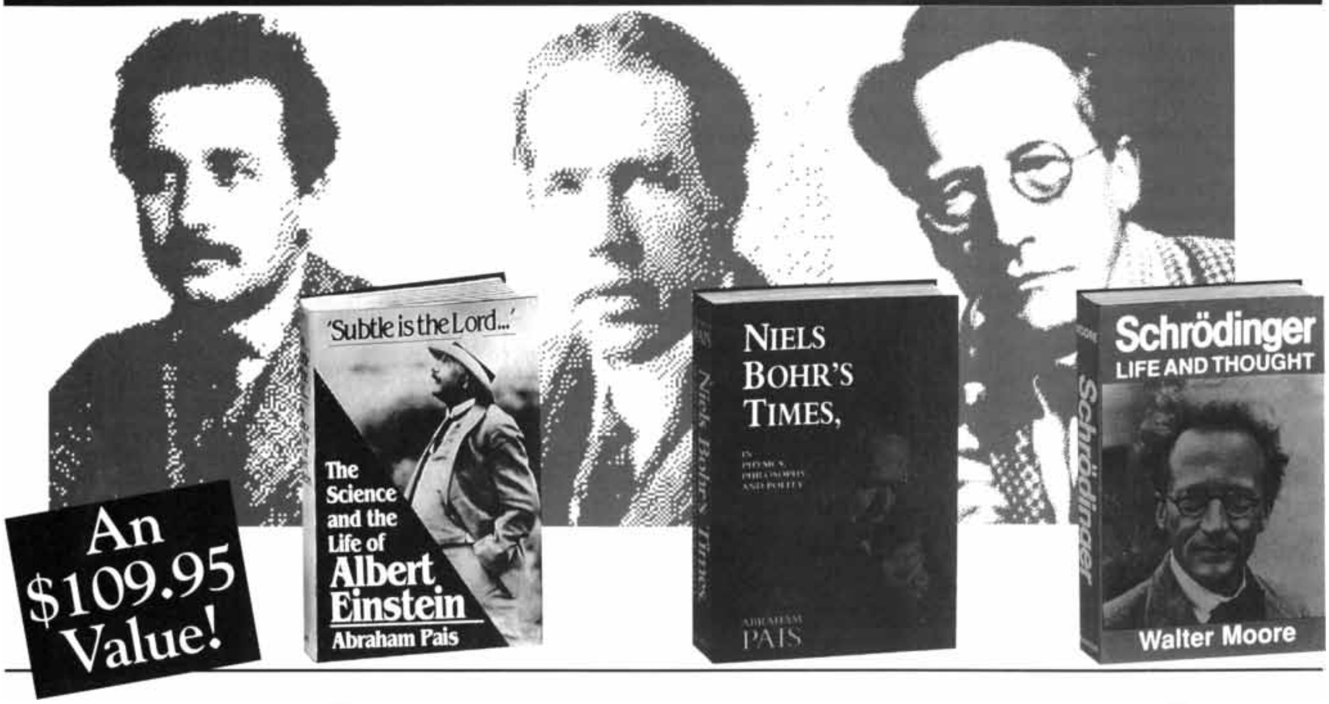
AMERICAN BLACK BEARS and their Asian cousins should gain when trading in bear parts becomes harder. Photo: Erwin and Peggy Bauer/Bruce Coleman, Inc.

But the availability of inexpensive, synthetic ursodiol has not dented the demand for bear gallbladders, Mills says. Wealthy Asians favor the natural product because the synthesized chemical "is so cheap people do not believe it is effective," one Japanese official told Mills and Servheen.

In Canada, a number of black bears have been found dead with their gallbladders cut out. Hard data are impossible to come by, but in the U.S., "the number of investigations is increasing," says G. Adam O'Hara of the U.S. Fish and Wildlife Service. Trafficking in bear parts is legal in seven states. Trade in Asian bears and their parts is banned under the Convention on International Trade in Endangered Species (CITES), but the ban has been impossible to enforce because traders claim that their gallbladders came from American bears killed legally.

There is hope on the horizon. This past March both the American black bear and the Russian brown bear were placed on a special appendix of CITES that lists "look-alike" species—ones that are difficult to distinguish from a controlled species. It will now be harder to pass off Asian bear parts as American, because gallbladders from American bears will need a CITES permit. —*Tim Beardsley*

A fascinating look at the three most influential men of 20th Century physics...



THE GREAT MEN OF SCIENCE COLLECTION

only **$2.95** *as your introduction to* **The Library of Science**

You simply agree to buy three more books — at significant savings — within the next 12 months.

'Subtle is the Lord' The Science and the Life of Albert Einstein

by Abraham Pais

"An outstanding biography"
— *Physics Today*

Author Abraham Pais, an award-winning physicist who knew Einstein personally, presents an intimate guide to the life and thought of the most famous scientist of our century. Using previously unpublished papers and personal recollections, this narrative illuminates the man through his work with liveliness and precision, making this the authoritative biography of Einstein.
Publisher's Price: $35.00

Neils Bohr's Times, in Physics, Philosophy, and Polity

by Abraham Pais

"...illuminates the probing, deep exertion within Bohr that was midwife to the birth of twentieth century physics. "
— Mitchell Feigenbaum

Niels Bohr is one of the three men without whom the development of quantum physics is unthinkable. The author, a colleague of Bohr's at the Institute for Theoretical Physics (now the Niels Bohr Institute), gives a first-hand account of Bohr as a great scientist and teacher as well as philosopher, humanitarian, administrator, and devoted family man.
Publisher's Price: $35.00

All Nobel Prize Winners...

All friends ...

All fiercely competitive colleagues ...

All three biographies for only $2.95...

Schrödinger Life and Thought

by Walter Moore

A few days before Christmas in 1925, Erwin Schrödinger left for a two-and-a-half week vacation in the Swiss Alps. When he returned, Schrödinger brought with him the fundamental formulations of wave mechanics. This, the first comprehensive biography of Schrödinger, is an attempt to explain this astounding burst of creativity and to describe one of the most fascinating and enigmatic figures of modern physics.
Publisher's Price: $39.95

THE LIBRARY OF SCIENCE is the nation's oldest and most respected science book club. Dedicated to serving both professionals and serious amateurs, it surveys the entire field of published sources on scientific subjects to make the very finest works available to club members at substantial savings, with prices starting as low as $9.95.

MEMBERSHIP BENEFITS- In addition to getting the Great Men of Science Collection for only $2.95 when you join, you'll also receive discounts on books you choose to purchase. Discounts generally range from 15% to 30% off the publishers' prices, and occasionally even more. • Also, you will immediately become eligible to participate in our Bonus Book Plan, with savings of more than 50% off the publishers' prices. • Moreover, our books are always identical to the publisher's editions. You'll never receive an "economy edition" with inferior paper and bindings from us. • At 3-4 week intervals (15 times per year), you will receive the Library of Science News, describing the coming Main Selection and Alternate Selections, together with a dated reply card. • In addition, up to three times a year, you may receive offers of Special Selections which will be made available to a group of select members. • If you want the Main Selection, do nothing, and it will be sent to you automatically. • If you prefer another selection, or no book at all, simply indicate your choice on the card and return it by the date specified. • You will have at least 10 days to decide. If, because of late mail delivery of the News, you should receive a book you do not want, we guarantee return postage.

If reply card is missing, please write to the Library of Science, Dept. 2-DL4 (00166), 3000 Cindel Drive, Delran, NJ 08370-0001, for membership information and an application.

Scientific American 7/92

Irrational fears suggest a similar evolutionary mechanism. "What are people phobic about, even in big cities?" Rapoport asks. "Not guns or cars, but snakes and spiders. You can condition a fear stimulus in monkeys after one exposure to a snake, but never to a flower." The same applies to the common human disgust for insects. "In one study the researcher put a sterilized cockroach in a glass, took it out and offered $5 to any student who would drink from the glass. Many of them wouldn't."

Yet not all animal compulsions can be compared with OCD, cautions Nicholas H. Dodman, a veterinarian at Tufts University. Stereotyped behaviors sometimes reflect the stress of confinement, which can stimulate the release of endorphins. These hormones, like the opiates they resemble, excite rather than relax certain animals. To work off their angst, stir-happy animals often engage in mindless activity, a solution ethologists call displacement. "Horses will grab their feeding trough in their mouths and pull back on it, or kick the door, or dig holes in the ground," Dodman says. Bears, on the other hand, will pace their cages; parrots pick their feathers out.

Even tail chasing, a common displacement in many dogs, can take on a pathological color. A remarkable case was reported recently in Canada, where police dogs are trained to attack on command but to desist when the suspected criminal raises his hands above his head. "A suspect was threatening a policeman by holding a chair above his head, and when the policeman told the dog to attack, the animal just blew a fuse," Dodman says. "It started chasing its tail and wouldn't stop. Nothing could cure it, so they finally had to put the dog down."

People, too, blow fuses when placed in unbearable situations. Whether or not such breakdowns themselves reflect a process of natural selection remains a mystery. *—Philip E. Ross*

DNA Fingerprinting Reconsidered (Again)

When the National Research Council published its study on DNA technology in forensic science in April, the *New York Times* reported that the panel had urged a moratorium on the use of DNA fingerprinting as evidence in court. Victor A. McKusick, who headed the panel, was quick to deny the report. The Johns Hopkins University human geneticist insisted that his group still endorsed the technique, although it did call for stricter controls on its use.

Even so, defense lawyers are using some of the research council's less well-publicized conclusions to try to overturn convictions won on the basis of DNA fingerprinting evidence. "We're going to attempt to get lawyers in every state to look at every relevant case," vows Barry C. Scheck, a prominent defense lawyer and professor at the Cardozo School of Law in New York City.

One of the recommendations of the panel that has attracted the attention of attorneys is that fingerprinting tests showing a phenomenon called band shifting be declared "inconclusive" until laboratories have performed adequate studies on the effect. Band shifting can make two forensic samples from different biological sources appear the same; private companies that perform DNA fingerprinting and the Federal Bureau of Investigation have already made attempts to correct for it. But the McKusick panel was not convinced that such attempts protect those accused: a suspect could be wrongly convicted if the "correction" was itself in error.

Scheck, who is co-chair of a committee that is reviewing old convictions in the light of the research council's report, believes that many people have been convicted on the strength of DNA fingerprinting tests that showed band shifting. One company offering DNA fingerprinting, Lifecodes, concludes that 30 percent of the forensic tests it has performed showed the effect.

The research council study also recommended that in the future a new, safer mathematical technique be used to interpret the results of DNA tests, and it outlined how the technique should be performed. That recommendation is likely to be grounds for reopening many cases in which the old technique was used to secure convictions, Scheck says.

A brief arguing against the admissibility of DNA evidence that cites the research council's report has now been filed in a "consolidated" case in the District of Columbia, *U.S.* v. *Porter*. The case is being appealed by the government because DNA fingerprinting tests carried out by the FBI were excluded. Similar briefs in California and Chicago are likely to follow quickly, Scheck predicts. *—Tim Beardsley*

Mother's Little Favorite

A spiteful gene ensures, "Like mother, like child"

In the words of one hymn to motherhood, "M is for the many things she gave me." If Mom is *Tribolium castaneum,* the flour beetle, one of her gifts had better be a gene called *Medea*. As three Kansas entomologists have learned, if a female beetle carries *Medea,* any of her offspring that lack the gene will die in the cradle.

Its discoverers puckishly declare that *Medea* is an acronym for "maternal effect dominant embryonic arrest," but don't be fooled: their real inspiration was Medea of Greek mythology, who killed her children in a jealous rage. *Medea* seems to represent the first member of a previously unnoticed class of so-called selfish genes. Such genes may play an important part in evolution and the origin of new species.

Selfish genes are parasitic sequences of DNA that increase their own frequency in a population, usually to the detriment of the organisms that carry them. Although most genes can prosper only by contributing to the well-being of their bearers, "there are certain times in the life cycle of an organism when genetic elements can potentially cheat and gain a transmission advantage," explains Jack H. Werren, who studies selfish genes at the University of Rochester.

Some genes cheat by inserting extra copies of themselves into cells; some cripple or obliterate their genetic counterparts on other chromosomes. A few can even change the sex of their host so that it produces only eggs or sperm favorable to the gene's distribution. Virtually all species, including humans, harbor at least some selfish genes. "It's a jungle in there," Werren quips.

Medea is remarkable because it is the first maternally linked selfish gene that acts on embryos rather than ova or sperm. Richard W. Beeman of the Agricultural Research Service in Manhattan, Kan., and his colleagues Kenlee S. Friesen and Rob E. Denell of Kansas State University became aware of *Medea* while testing whether flour beetles from different parts of the world had any problems breeding with one another. When they tried to cross beetles from a rice warehouse in Singapore with ones from a farm in Georgia, they noticed that some of the hybrid combinations had consistently small broods: many of the resulting larvae died before or during hatching.

Further experiments, which Beeman,

Friesen and Denell described in *Science* in April, convinced them that the deaths could be attributed to the activity of a selfish maternal gene killing noncarriers in the next generation. The lethality is linked strictly to the maternal line, Beeman observes. "If the father carries it and the mother doesn't, everything is normal," he says. "You wouldn't know the gene was there."

Destroying offspring may seem counterproductive from the standpoint of the beetles, but it suits the gene's purpose by eliminating the competition—in this case, individuals that cannot transmit it further. When *Medea* is carried by a female, the gene raises its frequency in the next generation by eliminating larvae that lack it. Males carrying *Medea* can meanwhile mate indiscriminately and infect other pedigrees with the gene.

As *Medea* spreads, many larvae die, but the gene can eventually overtake an entire population. "Once it becomes fixed in a population, it becomes invisible, which may be why no one discovered it before," Beeman says.

No one yet knows how *Medea* exerts its murderous influence. One possibility is that the gene has two products, one a poison, the other its antidote. If the poison persists in the mother's eggs, *Medea*-free offspring developing from those eggs might die because they are unable to counteract it. (Some selfish genetic elements in primitive bacteria are known to operate in a similar manner.) Beeman believes that if the *Medea* gene can be specifically located and cloned, its killing mechanism may eventually become clear.

The existence of *Medea* invites speculation about the possible role of selfish genes in evolution. New species can arise when reproductive barriers, such as distance or mutations, split up populations. If factors resembling *Medea* appear commonly in nature, they could contribute to that isolation by making it difficult for strangers to join a protected population. Both Beeman and Werren caution, however, that the actual isolating effects of *Medea* and other selfish genes remain to be demonstrated.

There is reason to suspect that *Medea*-like selfish genes could be widespread among animals and plants. Once they knew what to look for, Beeman and his colleagues quickly found factors resembling *Medea* in other populations of flour beetle and in a completely different beetle species. Crossbreeding experiments in other species could uncover new analogues of *Medea*.

Notwithstanding its selfishness, *Medea* could sometimes turn out to be surprisingly advantageous for its hosts after all. Werren argues that for many mammals, insects, plants and other organisms, the chief competitors of any individual are actually its siblings, who share the same environment and consume the same resources. By weeding out some siblings for its own purposes, *Medea* may indirectly enhance the fitness of its carriers—thus boosting its own success still more.

As biologists continue to investigate the growing ranks of selfish genes, even more crafty proliferation strategies are likely to emerge. If someone can work out an appropriate acronym, perhaps the next selfish gene will be called *Machiavelli.* —*John Rennie*

RED FLOUR BEETLE, **Tribolium castaneum,** *stands atop a kernel of corn* **(left).** *Many females of this species bear a recently discovered selfish gene that discourages crossbreeding by killing offspring that fail to carry it* **(right).** *Photo: Richard W. Beeman.*

Colliding Physicists

Will the SSC lose out in the battle of big versus small?

"The tide is running against big science right now," observes William Happer, director of the office of energy research at the Department of Energy (DOE). He ought to know—his job is to coordinate plans for the accelerators that physicists use to study the behavior of fundamental particles, including the huge Superconducting Super Collider (SSC). And that project is under determined attack by those who argue it will strip funds from other science, including "small science," which is arguably more likely to lead to technologies that will produce economic returns.

Opponents of the SSC got some unexpected confirmation of their fears in a recent report by a panel chaired by Michael S. Witherell, a physicist at the University of California at Santa Barbara. The report gives some of the first clues about the impact the $8.25-billion SSC is likely to have on the rest of particle physics. DOE officials insist the behemoth, which is now under construction in Waxahachie, Tex., and will be 54 miles in circumference, will not come at the expense of other science. But the department told Witherell's panel to assume that once the SSC goes into operation (under present plans, in 1999), the budget for the rest of high-energy physics will be reduced by more than $150 million a year, or close to 25 percent of next year's amount.

In turn, Witherell's panel warned that if budgets fail to keep pace with inflation, one of the former crown jewels of particle physics, the Stanford Linear Accelerator Center (SLAC), should be closed permanently. Even if budgets keep pace with inflation, the panel recommended that SLAC, the site of several breakthrough discoveries over the past 25 years, stop doing particle physics until there are sufficient funds to start building a new accelerator at Stanford, known as the B-factory, in 1996. Happer says he is now planning on a budget that will just match inflation.

The proposed $184-million B-factory, like the existing linear collider at SLAC, would collide electrons and positrons together. But the Witherell panel declined to give the SLAC B-factory top priority after the SSC, giving the nod instead to a $204-million upgrade of the Tevatron, currently the world's biggest proton-antiproton collider, which is at Fermi National Accelerator Laboratory in Batavia, Ill.

Systems Integration magazine says IRIS Indigo has, "the best price/performance ratio we've seen on a workstation – or any product for that matter." They've also been given the BYTE Award of Distinction and the Digital Review Editor's Choice.

Desktop discovery: IRIS Indigo RISC/PC.

We've built a system that hits the cool price point of Macs and PCs, but delivers the hot performance of workstations – the IRIS Indigo™ family.

These are the machines that can take your work to new heights of discovery.

Run leading productivity software and thousands of technical and graphics programs for General Science and Biology, Physics and Medicine. Plus, use IRIS Explorer™ to build your own visual applications. And with Showcase,™ our integrated drawing and presentation tool, turn concepts into funded projects.

Imagine the power: interactive 2D and lifelike 3D – in a range of affordable configurations – put together by the leader in visual computing.

Discover them for yourself. Get our free 20-page full-color capabilities brochure, and find out where you can check out the IRIS Indigo family, as well as our full line of servers and supercomputers.

Call 1 (800) 800-7441. Ext. 24

SiliconGraphics®
Computer Systems

That recommendation has not been well received at SLAC, where electron-positron researchers see themselves as victims of the proton physicists, who dominate the field. The Tevatron upgrade, known as the main injector, will extend the life of that machine. But the same experiments that the main injector will make possible could be done at the SSC a decade later. Happer says it is all a matter of timing. "If you're willing to wait for the SSC, what do you do in the meantime?" he asks.

Burton Richter, the director of SLAC, wants to know what his accelerator will do in the meantime if building the B-factory is put off until 1996. The existing SLAC collider has been superseded by the Large Electron Positron collider at CERN, the European laboratory for particle physics near Geneva. And Richter says the Witherell report's recommendation that SLAC should do just research on accelerator technology for two years is unrealistic. As for the report's contingency plan to abandon SLAC if budgets fall behind inflation, Richter complains that it would leave U.S. high-energy physics "monolithically proton oriented."

In Congress the budget-cutting knives are out, and the General Accounting Office has accused the DOE of playing fast and loose with its bookkeeping for the SSC. Happer denies there are serious problems. "We're keeping a tight rein on it," he insists. A computerized management system is in place, and contracts have been renegotiated to eliminate budget and schedule overruns that were forecast. But Victor S. Rezendes of the General Accounting Office points out that if a foreign contractor agrees to supply components for less than the budgeted amount, as some have, then the DOE chalks up the saving as a foreign contribution. Yet if a U.S. firm does the same thing, the saving is recorded as a reduction in the project's total cost.

The DOE has promised that nonfederal funds will account for one third of the cost of the SSC. India has pledged $50 million, Korea has offered $30 million and negotiations are in progress with the former Soviet Union, Happer says. But even after Texas's contribution of $875 million has been taken into account, that still leaves the agency a daunting $1.7 billion short.

All hopes rest with Japan. Happer says he expects the Japanese will agree to pay as much as $1.5 billion over the lifetime of the SSC, in addition to any contribution to the two $500-million detectors, only one of which is included in the SSC's budget. The Japanese government has established a task force to consider the matter, although it will not appropriate any funds before the end of the year. But Happer is counting on a commitment by "late summer or early fall." One congressional staff member tracking the project, however, believes Japanese funds are unlikely to arrive before 1994. If so, the SSC would be running $100 million behind budget during 1993.

If the proposed cuts at SLAC prompt California's representatives in Congress to reconsider their support for the SSC, the project's political viability could be jeopardized. Doubts will be reinforced if CERN proceeds with plans to build quickly a new accelerator, the Large Hadron Collider, which could bag some discoveries the SSC was designed to make. That prospect should, according to Happer, be no source of joy for other physicists. "If the SSC were to go down," he says grimly, "there'd be no chance of getting a nickel of that money back for science." *—Tim Beardsley*

Sound System

Using the ocean's noise to image undersea objects

To peer through the murky depths of the sea, oceanographers generally transmit sonar signals or listen to the sound an object might emit. Now Michael J. Buckingham and Broderick V. Berkhout of the Scripps Institution of Oceanography and Stewart A. L. Glegg of Florida Atlantic University have come up with a new way of forming underwater images. They examine how objects modify the naturally occurring random noise in the ocean.

In effect, Buckingham and his colleagues are imaging objects with the acoustic version of daylight. Objects are visible because they scatter the atmospheric light. An approach using acoustic daylight makes the ears work as the eyes do, says Ralph R. Goodman, an ocean acoustics scientist at Pennsylvania State University. The technique exploits the fact that "the sea is filled with sound."

Most of the ambient noise comes from breaking waves. As a surface wave rolls over and breaks, it traps air, forcing it down a few meters into the water. The air forms tiny bubbles, which resonate, or ring, for a few tens of milliseconds. "The net result," Buckingham says, "is that the ocean is quite noisy."

Seeing with acoustic daylight may offer a view of objects located anywhere in the ocean because sound travels with remarkable efficiency in the sea, essentially to the ocean bottom. Furthermore, sound from such sources as boat engines or dolphins does not interfere with acoustic daylight. Those passing sounds, which produce problems for sonar, would actually contribute to the overall ambient noise, much as stadium lights add to the illumination of a late-afternoon baseball game.

In their experiment the researchers submerged a piece of plywood covered with neoprene rubber off the Scripps pier. Then they placed a parabolic reflector and a hydrophone—the "acoustic lens"—seven meters away. The arrangement was similar to taking a photograph with the sun behind the camera. They were able to "see" the acoustic signature of the target at sound frequencies between five and 50 kilohertz; there was a four-decibel difference in the scattered noise and the ambient noise field.

Moreover, the targets absorbed certain sound frequencies and reflected others. That finding implies that underwater objects can be rendered in false color, Buckingham says. Visually, objects appear to have color because they absorb certain frequencies of light.

The researchers are still a long way from actual pictures. The data collected in that initial experiment "correspond to one point of an image," Buckingham says. A complete portrait would require an array of hydrophones, each pointing in a different direction. Buckingham speculates that within a couple of years the technology may be able to generate an image once every 0.1 second and thus act as an acoustic video camera.

Unfortunately, the resolution of an acoustic daylight system cannot match that of the eye. That feat would require a parabolic reflector 600 meters wide. Moreover, one trades off resolution for imaging distance. High-resolution requires high-frequency sound, which does not travel very far before attenuating. The investigators are now aiming to make images of objects 10 centimeters in size at a range of 10 meters.

Still, there are several applications, most notably those concerning submarines. "Subs don't like using active sonar," Buckingham says, because that gives away their position. On the other hand, acoustic daylight is a covert system that can provide the necessary vision. Other possible uses might be monitoring harbor entrances, like security cameras in shopping malls, and tracking whales or schools of fish. *—Philip Yam*

Defining Dyslexia

Is it a distinct disorder or a problem of degree?

Dyslexia is primarily a reading problem, but for some children it is also a geographic one. Not all those getting remedial assistance in, for example, Vermont, which defines dyslexia as a 22-point discrepancy between I.Q. and reading achievement scores, would qualify for help in California, which defines it differently; only about a dozen of the 50 states apply the same criteria. The quilt of conflicting standards that covers the U.S. is a by-product of the difficulty that educators, policymakers and researchers have faced in defining dyslexia.

Some new research now questions whether it is a distinct disorder at all: investigators at Yale University have concluded that dyslexia differs in degree but not in kind from other forms of poor reading achievement. Such a change in the view of dyslexia would affect the direction of future research and the provision of services to children.

The absence of a unique diagnostic peculiarity has been an obstacle to identifying the disorder. Dyslexia has become synonymous with perceived reversals of letters or words on the page, but such illusions are actually uncommon. In the early 1970s the World Federation of Neurology defined dyslexia as "difficulty in learning to read despite conventional instruction, adequate intelligence and sociocultural opportunity."

That description, which did not specify precisely how to measure dyslexia, led the federal government to mandate the development of special education programs to help children identified as dyslexic. Implicit within these actions was the notion that dyslexia represented a special category of disability unlike others causing underachievement in normal children.

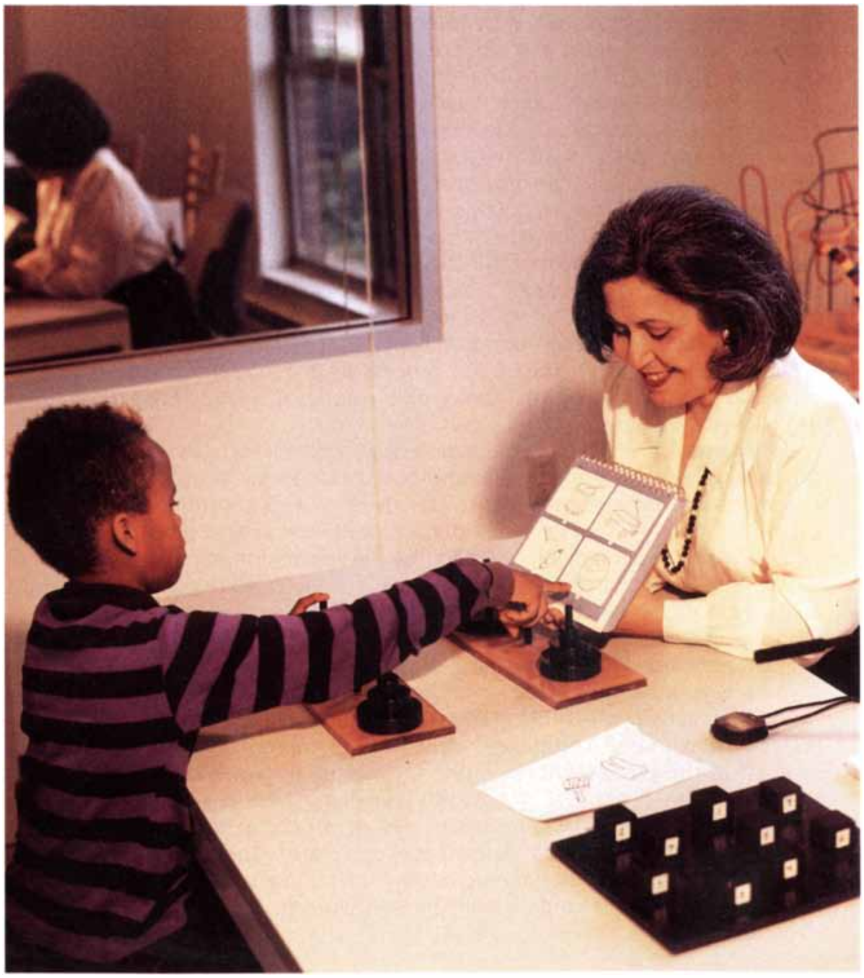
Sally E. Shaywitz, professor of pediatrics at Yale, thinks this categorical view of dyslexia is mistaken. This past January she and her colleagues published a long-term study indicating that dyslexia is only part of a smooth continuum of reading difficulties found throughout the population.

The Yale researchers studied the development of several hundred Connecticut children as they progressed from kindergarten through sixth grade. Each year the researchers administered reading achievement tests to the children. Dyslexic children were identified as those whose reading skills were significantly below the level predicted from their I.Q.'s. Shaywitz found that these dyslexics did not form a distinct subgroup of poor readers; instead they fell within the low end of the normal distribution of reading scores.

More important, as Shaywitz had predicted statistically, two thirds of the children labeled dyslexic in the first grade no longer fit the study's criteria by the third grade. An approximately equal number of normal but poor readers moved the other way. These reclassifications did not occur because the children's abilities had changed dramatically, Shaywitz emphasizes. Rather, because the dyslexic children were not qualitatively different from others with slightly better scores, it was possible for children to drift from one category to the other with only minor changes in test scores. "The differences are in severity, not in type," Shaywitz says.

If dyslexic children are not fundamentally different from others, why have so many educators tagged them as a distinct group? A landmark study of children on the Isle of Wight conducted more than two decades ago did conclude that dyslexics had unique disabilities, but Shaywitz says those findings have never been replicated. She also argues that the criteria for measuring dyslexia have been highly subjective and sometimes biased. In 1990 Shaywitz's group documented unconscious biases in the identification of dyslexics: "Boys who are hyperactive or disruptive are more likely to be identified than girls who may not read but are not bothering anyone," she explains.

Yet some experts also criticize the Yale group's approach to spotting dyslexics. "They took a group of children defined by substandard reading achievement quite early in their education and then a couple of years later. That isn't really dyslexia. That's just substandard reading achievement, period. It has many explanations," fumes Albert M. Galaburda, a dyslexia researcher at Beth Israel Hospital in Boston. He suspects that by looking at scores instead of individuals, the Yale group failed to recognize many true dyslexics and misla-



SALLY E. SHAYWITZ of Yale University argues that dyslexia is only part of the spectrum of reading skills and not a distinct disability. Photo: Jason Goltz.

beled other poor readers. The net effect of these inaccuracies, he argues, is to make the dyslexic group appear more normal. "It's a self-fulfilling prophecy," he says.

Paula Tallal, a language disorder specialist at Rutgers University, believes that the children were too young to be reliably classified early in the study. "How can children be diagnosed as having reading difficulties if they haven't yet had the opportunity to learn to read?" she asks. The children who seemed to move from the dyslexic to the normal categories, she thinks, were not correctly diagnosed in the first place.

Moreover, Tallal considers the argument over the arbitrariness of the dyslexia definition a straw man. Instead she maintains that the important task confronting researchers is to understand the mechanisms underlying reading problems so that dyslexic children can be helped more efficiently.

Most investigators now suspect that dyslexic children read poorly because of a highly specific linguistic problem sometimes called "phonological unawareness." They cannot easily learn to read because they have extraordinary trouble associating printed letters with the sounds of speech. (Similarly, many congenitally deaf people who have mastered the linguistic complexities and subtleties of sign language have trouble learning to read.)

Several lines of evidence further suggest that the root cause for much dyslexia is a problem with processing very rapidly changing sensory stimuli. Tallal, for example, has shown that dyslexic children have trouble making accurate distinctions between similar auditory signals: they often cannot hear the difference between speech sounds such as "bah" and "dah." Galaburda and his colleague Margaret S. Livingstone of Harvard Medical School have recently seen variations in the visual pathway of dyslexics that suggest an analogous problem with fast-changing visual stimuli. Galaburda has also found several other neuroanatomical abnormalities in the temporal lobe and other areas of the brain.

Shaywitz fully endorses the idea that dyslexia has biological underpinnings but insists that "we have to move away from this set of anecdotal, impressionistic standards." Until a rigorous and empirical definition of dyslexia is developed, she contends, all the research on mechanisms is blunted by unresolved issues about whom to study.

Whom to help is also an issue. Galaburda worries that if dyslexics cease to be distinguished from other poor readers, they will again fall prey to the old stigma that they are "just stupid." More worrisome is the idea that the Yale study could be misinterpreted as meaning that dyslexia goes away on its own. Such views could lead to cutbacks in remedial assistance to dyslexic children.

To the contrary, Shaywitz hopes her work will make more assistance available. "Current school policy is very much based on an all-or-none model of dyslexia," she says. Consequently, many children are not allowed access to reading assistance programs because their abilities, although poor, are not bad enough for them to be classified as dyslexic. Shaywitz believes that if dyslexics are no longer seen as a group apart, inappropriate and arbitrary cutoffs limiting the access to remedial aid will be repealed. *—John Rennie*

Rapid Recall

Name a movie that has an extended dream sequence, red shoes, a lion and a scarecrow. Researchers have spent years trying to get computers to emulate the human ability to sift through tens of thousands of pieces of information and find unsuspected connections. But it could take hours for a conventional machine programmed with film information to come up with *The Wizard of Oz,* according to computer scientist James A. Hendler of the University of Maryland.

That kind of response time makes conversing with artificial intelligences an exercise in patience. Until recently, the only way to make the programs faster was to limit their range to carefully defined problems requiring only a few hundred to a few thousand pieces of knowledge. For the past decade, artificial-intelligence mavens have blamed slow hardware for the field's relative lack of progress.

Now the researchers' bluff is being called. Lack of computer power "is not an excuse anymore," says David Waltz of Thinking Machines Corporation, a parallel-computer company in Cambridge, Mass. By developing new algorithms that run on massively parallel computers such as the company's Connection Machine, Hendler and other researchers can pick the right datum out of a vast collection of facts in one ten-thousandth the time taken by other machines. Hendler's system, called PARKA, answers queries containing multiple conditions (*The Wizard of Oz* question has four) nearly as fast as those calling for a single fact. What might take a conventional algorithm the better part of a day takes PARKA less than a second.

The program that Hendler and his student Matthew P. Evett have built relies on semantic networks. These software structures represent knowledge using nodes that stand for objects and coded links that stand for relationships between them, such as "part-of," "location" or "is-a." (For example, Clyde, a zebra, might be connected by an is-a link to "Zebra," which would be linked in turn to "Herbivore" and to "Animal," thence to "Life-Form" and ultimately to "Thing.") Because each of the 65,536 processors in a full-size Connection Machine handles a node and all its links, all the nodes in a knowledge base can be checked simultaneously.

Other researchers have adopted somewhat different techniques. Waltz and his colleagues are exploring memory-based reasoning, which relies on an enormous data base of examples to classify new cases according to the ones they most closely resemble. Memory-based methods proved faster and more reliable than the people employed to pick key words for indexing news stories at Dow-Jones, Waltz says. Both Hendler and Waltz point out that a large, fast memory changes the way intelligent programs should be designed—long, complex chains of reasoning working from a few key facts give way to simpler deductions based on what did or didn't work in similar cases.

Waltz says a handful of other projects are under way, including one to tackle the nearly insuperable task of translating medical diagnoses into the codes that insurance companies insist on. Hendler, meanwhile, is looking forward to the chance of converting CyC (pronounced "psych") to a form that PARKA can digest. CyC, the largest knowledge base now in existence, is under construction at MCC in Austin, Tex. It contains more than two million rules encoding everyday knowledge. Its creator, Douglas B. Lenat, estimates that 10 million rules should endow it with intelligence. Parallel algorithms might allow such an intelligence to work fast enough for humans to recognize. *—Paul Wallich*

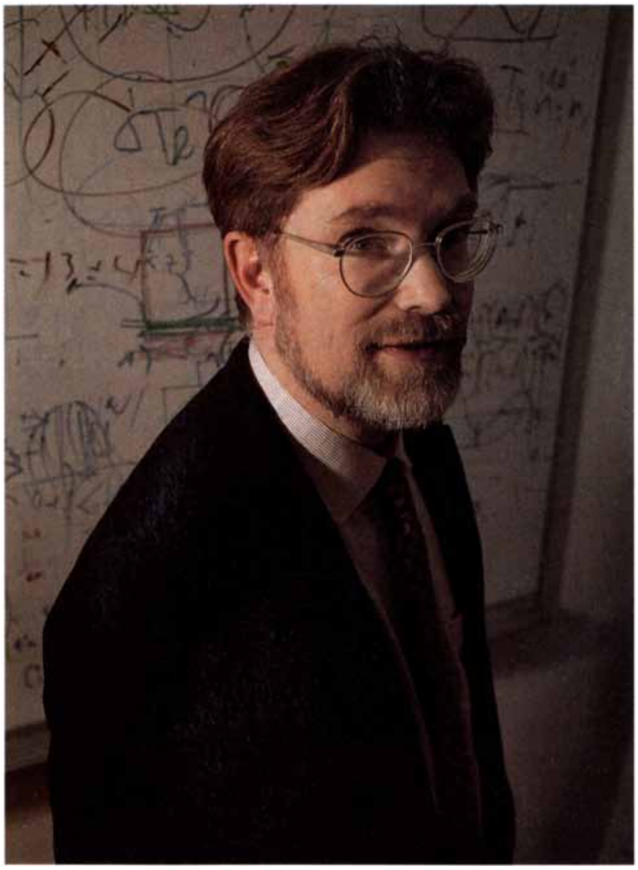
COBE's Cosmic Cartographer

The team from *People* magazine has already come and gone, and George F. Smoot is now trying to answer SCIENTIFIC AMERICAN's only slightly more scientific questions. It's not easy. A former colleague telephones to congratulate him. A graduate student strides into Smoot's office and shoves a fax into his hand, an invitation to a conference on religion, science and the environment. The student smirks at the mishmash of issues, but Smoot seems intrigued. "Cosmology is where science and religion meet, right?"

A photographer pokes her head in the door to ask when Smoot will be free. "I'm getting tired of having my picture taken," he mutters. The photographer jokes that his colleagues are worried about the *People* photographs. "*They're* worried," Smoot snorts. "They're not in 'em." Smoot is scheduled to go to a conference in Amsterdam the following week, and he is thinking of leaving a few days early. "Success is disruptive," he sighs.

Just two weeks earlier Smoot was a relatively obscure, 47-year-old astrophysicist at the University of California at Berkeley and the Lawrence Berkeley Laboratory. His specialty was the cosmic microwave background, the faint radiation that is thought to be the lingering afterglow of the dawn of creation. Smoot was perhaps best known in the astrophysics community as the team leader for the Differential Microwave Radiometers on board the *Cosmic Background Explorer* (*COBE*) satellite. The radiometers were designed to find inhomogeneities in the microwave background.

Astronomers had been nervously awaiting the *COBE* results. The failure to find inhomogeneities would make it difficult—and perhaps impossible—to explain how the cosmos came to be so clumpy. But during an April 23 meeting of the American Physical Society (APS), Smoot unveiled a garishly colored, ovoid-shaped map that he said showed the long-sought "ripples" in the primordial radiation corresponding to the vast vacuums and clusters of galaxies populating the modern universe. "If you're religious," Smoot said, "this is like looking at God."



SMOOT has been scrutinizing the cosmic microwave radiation since the early 1970s. Photo: Stephanie Rausser.

Although Smoot occasionally went to church as a child, he says he is not particularly religious himself; he just tries to use language that conveys the profundity of the subject. "When you start talking about 10^{-43} second [the instant after the big bang at which the universe becomes describable by physics], nobody knows what that is. Not even physicists, really."

Smoot is on a committee at Berkeley that seeks to "give a better public appreciation of science and show that scientists are real people—that's why I try not to be too stiff." Indeed, for someone in a field so other-worldly, his demeanor is distinctly down to earth. With his wire-rim glasses, neatly cropped beard, oxford shirt and khaki slacks he looks more like a lawyer—a hip public defender, say—than someone obsessed with such issues as dipole anisotropies resulting from galactic rotation.

The publicity is beginning to pall. Since the APS meeting, Smoot has been able to work only late at night, when he is already exhausted. Moreover, other team members have complained—and Smoot agrees—that he has received too much attention and credit for what was, after all, very much a group effort. "Even when you tell reporters it's a team effort, the best they end up doing is saying 'team leader.'" On the other hand, Smoot says emphatically, "I personally had a very strong involvement for a long time."

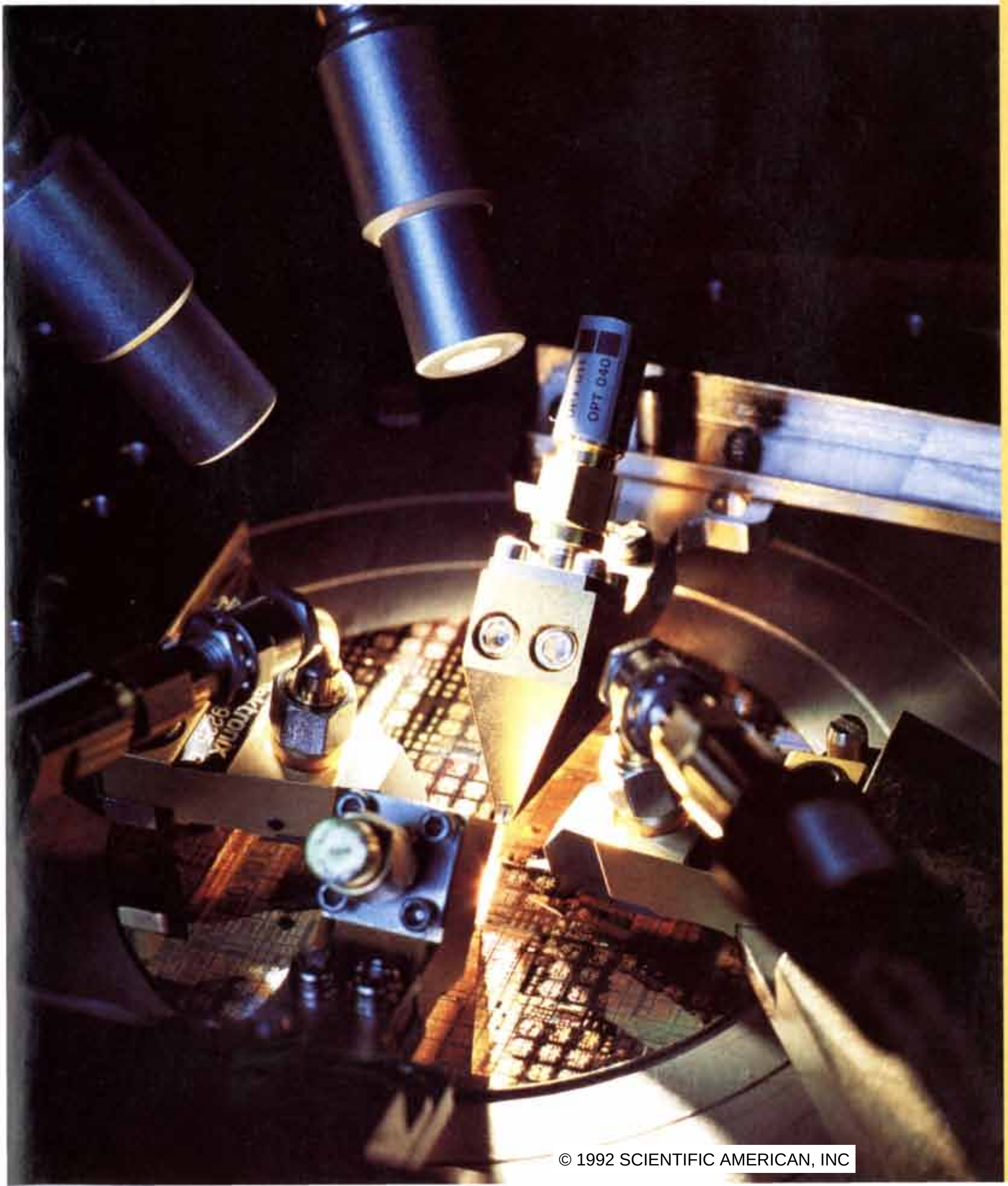
As the son of a geologist whose job with the U.S. Geological Survey took him from Florida to Alaska, Smoot grew up wanting to become a scientist. He eventually chose physics, obtaining his doctorate from the Massachusetts Institute of Technology. After moving to the University of California at Berkeley in 1971, he began working with the eminent particle physicist Luis W. Alvarez.

Under Alvarez's guidance, Smoot drifted into astrophysics, sending balloons aloft to study cosmic rays, mysterious high-energy particles found hurtling in the vacuum of space. Urged by Alvarez to consider other phenomena that could be investigated with high-altitude observations, Smoot became increasingly fascinated by the cosmic microwave background.

At that time, the definitive text on cosmology was one written by P. James E. Peebles of Princeton University. "He's influenced a whole generation of cos-

Toshiba's

21st Century Technologies



This special supplement outlines Toshiba Corp.'s ongoing commitment to 21st Century technologies, particularly emphasizing activities of the Research and Development Center in Kawasaki. The center, which employs approximately 2,200 individuals—including 1,550 researchers and engineers—marked its 30th anniversary last November.

Toshiba's
21st Century Technologies

By Roger Schreffler

Naohisa Shimomura, *Director of* The Research and Development Center

DR. NAOHISA SHIMOMURA, WHO WAS INSTRUMENTAL IN THE DEVELOPMENT OF TOSHIBA'S FIRST LARGE SCALE INTEGRATION DEVICE MORE THAN TWO DECADES AGO, IS DIRECTOR OF THE COMPANY'S RESEARCH AND DEVELOPMENT CENTER.

QUESTION: Dr. Shimomura, could you please outline the basic structure of Toshiba's R&D activities?

ANSWER: We presently have a three-layer research organization structured around the nature of the research all the way through applied research and product development. Four corporate laboratories, including the R&D Center, concentrate on basic research. The minimum perspective of these labs is five years. Under these corporate facilities are 10 development laboratories, which cover a broad range of essentially product engineering-related activities from information processing and semiconductors to medical engineering; development labs generally take three to five years to complete work. And finally, each of our 10 business groups or divisions has its own engineering department to focus on product engineering as it relates to commercialization. The time frame here is one to three years.

Smooth transfer and sharing of technologies is enhanced through regular meetings between the center, the engineering labs and the business groups, as well as through frequent exchange of personnel.

Q. What is the center's primary research focus?

A. Our current structure has three overriding themes. Our primary work continues to be basic research and development of materials, products, and systems within the context of the company's overall business strategy. That is not going to change. In addition, we also provide R&D support for the divisions, assisting all 10 Toshiba engineering laboratories. Finally, the center generates new ideas and establishes intellectual property rights, mainly the filing of patents.

AN "UNDER THE TABLE" PROGRAM ENCOURAGES TOSHIBA RESEARCHERS TO EXPLORE PROJECTS ON THEIR OWN.

Q. How is the corporate R&D strategy decided?

A. Our overall research strategy is formed by the corporate technology committee, comprised of executives from the five corporate labs, including the R&D Center, the heads of research and development of each business group, plus relevant divisions, including corporate planning, technology planning and coordination, productivity, intellectual property, and the design center. Some 20 to 25 individuals meet monthly to discuss the latest R&D issues, the allocation of the budget, and what new research themes to target. They then present proposals to the president.

Q. Which product areas now receive the greatest emphasis?

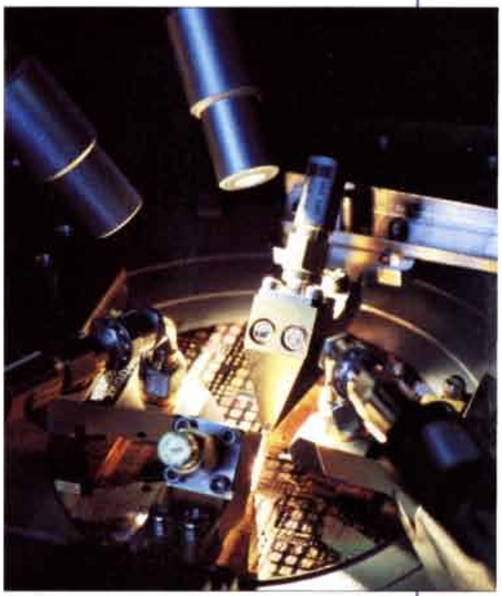
A. Computers, telecommunications, and various electronic devices, which is consistent with our corporate business strategy. In the consumer products field, the primary focus is on advanced audio and visual technologies such as high-definition television and digitalized imaging products.

To illustrate our changing focus, two years ago we set up a new communication systems laboratory. Then last year we established new information systems, display, and video and media electronics labs. Concurrently, we've expanded the activities of our older information systems and electron device laboratories. The electron device lab currently is developing high-speed switching devices and laser diodes, both essential to the multimedia concept.

In 1988, reflecting the growing importance of semiconductors and silicon materials, we separated our VLSI research laboratory from the main center, establishing its autonomy as an independent research function known as the ULSI Research Center.

Q. What would you characterize as the center's major achievements over the past five years?

A. I'd have to include the Josephson junction device for high-temperature superconductive materials, announced last October; in addition, our asynchronous transfer mode switch for multimedia systems and 10-gigabit gallium arsenide 8-bit multiplexer for very high speed optical communications, both developed in 1990; a NAND type EEPROM, which is 30% smaller than existing flash memory chips, in 1988; and 2-million pixel CCD image sensor used for high-definition TV, in 1987.



GALLIUM ARSENIDE WILL ALLOW TOSHIBA RESEARCHERS TO DEVELOP SUPERCONDUCTORS WITH MUCH FASTER OPERATING SPEEDS.

Q. What is the center's annual budget?

A. In the most recent fiscal year (ending March 1992), we spent approximately 26 billion yen, 9% more than in 1990. Generally, our budget rises about 10% annually. Of last year's total, headquarters funded 55% and the operating divisions 40%; the remaining 5% came from outside private and public sector contracts.

Q. How does Toshiba encourage creativity and initiative?

A. We've introduced several systems to encourage young scientists to initiate research activities of their own. One of these we call "under the table" research. All of our researchers are free to use 10% of their time and budget to pursue research of their own choosing. They need not report this. Second, we have set up a structure permitting scientists to seek direct approval from the director for project proposals, in effect bypassing the chain command. And lastly, each year we dispatch six or seven individuals overseas to further their research studies.

Materials Research

The Polysilane Race

Shuji Hayase, *Chief Research Scientist,*
Chemical Laboratory

EVERY MORNING like clockwork, Shuji Hayase leaves his small company apartment for the fourth floor of the Toshiba R&D Center. After a five-minute walk, he arrives several minutes ahead of a regular planning session scheduled for half past eight.

In recent months, these meeting between members of Hayase's chemical lab have become more intense as a six-year project on an advanced silicon polymer nears completion. This air of urgency has been fueled by the recent entry of several new and powerful competitors. "If we come up with a new silicon device ahead of anyone else, it will give us a tremendous advantage," declares the 37-year-old chemist, who joined Toshiba in 1978. "Failure to do so, conversely, could have serious consequences."

The key, according to Hayase, could be "polysilane," a silicon-based organic polymer which last summer Toshiba first synthesized with phenol molecules; included among these is the chemically active hydroxyl ("OH") group. Until then, polysilane was limited to laboratory experiments. "This material is extremely difficult to work with," maintains Hayase, who began work on the material in 1986 as an "under the table" project, a private undertaking funded and encouraged by the research center. His main focus at the time involved developing a new type of resist material for large scale integrated circuits. "In that endeavor," he explains, "I came across several scientific papers which detailed a number of unique photoreactions exhibited by the polymer."

The polymer, which was discovered by Dr. Robert West of the University of Wisconsin in 1980, is one of as many as 10 materials that Toshiba is currently evaluating for future use in electronic devices such as photoresists, displays, memory devices, optical switches, and composite materials. Suspecting it as one of the most promising, Hayase meets monthly with as many as 10 different product groups inside Toshiba, including those involved in displays, optoelectronics, LSIs, communications, and composite materials. "In this sense," he declares, "our research is both basic and applications-driven."

DR. SHUJI HAYASE: "IF WE COME UP WITH A NEW SILICON DEVICE AHEAD OF ANYONE ELSE, WE WILL HAVE A TREMENDOUS ADVANTAGE. FAILURET O DO SO COULD HAVE SERIOUS CONSEQUENCES."

The material's most important feature is its semiconductivity, which sets it apart from conventional carbon-based polymers such as polyethylene and polycarbonate. In addition, it is transparent, giving polysilane strong potential for optical devices. Soluble in water, the polymer boasts a high molecular weight—more than 500,000. Thus it can be used in long silicon chains of as many as 5,000 "repeating" silicon molecules. Moreover, it can be formed into both thin films and powders.

Although the project is now in its seventh year, efforts to polymerize polysilane began just two years ago. Polymerization is only possible in melted sodium, a highly reactive reagent, says Hayase, who received his doctorate from Osaka University while employed by Toshiba. Because of this, the reported chemical structures were limited to very stable substituents. These factors greatly restrict opportunities to utilize the material. And until last year it was virtually impossible to alter the polymer's basic structure.

ALL THAT REMAINS IS TO CHOOSE THE OPTIMAL MOLECULE TO ENHANCE PERFORMANCE AND INSURE DURABILITY. "WE'VE DONE THE HARD PART."

That changed with the "linkage" of hydroxl group molecules to polysilanes, which introduced an "active" side to the polymer's chain. Now all that remains is to choose the optimal molecule to enhance overall performance and insure durability in "cross-linking" of the chain. "We've done the hard part," affirms Hayase, who expects to complete that task early next year.

Besides difficulties associated with handling the material, polysilane's principal shortcoming was that it decomposes when exposed to ultraviolet light (less than 300 nanometers). Under normal light conditions, however, the polymer is stable.

Of the current array of materials under study, polyacetyline is probably the strongest competitor in terms of semiconductivity and nonlinear optical properties. However, the almost completely black polyacetyline is opaque, which gives the transparent polysilane a decided advantage for use in optical devices.

Furthermore, by using polysilane for LSIs, the Toshiba researcher believes that a finer pattern or drawing scale of 0.2 to 0.1 microns can be achieved. Presently the most advanced dynamic random access memory chips are based on 0.8-micron patterns. "A 0.1 or 0.2 micron pattern is equivalent to a 1-gigabit DRAM, or a billion circuits on a chip the size of a thumbnail," he asserts.

In memory and display devices, polysilane will have to compete with amine polymers such as polyvinyl carbazole. Here, too, the material holds a distinct advantage in that its "hole" mobility—the movement of positive charges back and forth throughout the silicon or "Si" chain—is much higher than conventional substances. This feature, says Hayase, could lead to devices 1,000 times faster than those currently used.

There is still no definite idea how polysilane might enhance composite materials. "Until now it hasn't been possible to make a composite because of difficulties in joining polysilane with other materials," says Hayase. "However with its high refractive index of 1.6 to 1.7, polysilane has one of the highest refractive indices among transparent organic molecules. Combined with glass, for instance, it can be used in lenses."

At present, Toshiba's main competitors are IBM and Nippon Telegraph and Telephone Corp., a situation which could change as the material has become the focus of an International Trade and Industry Ministry project on silicon-based polymers. Already eight groups, including Toshiba subsidiary Toshiba Silicone Co. and three universities, have joined the 10-year project.

Cryogenic Cooling:

A Supercool Refrigerator

Hideki Nakagome, *Chief Research Scientist,*
The Energy Science and Technology Laboratory

HIDEKI NAKAGOME has a dream: He wants to achieve a major scientific breakthrough through a total team effort. "Ten years from now, I want to look back with the knowledge that I've created something entirely new by pulling people together into a vibrant, self-contained unit," says the 40-year-old engineer. "While not discounting the importance of individual initiative, I believe that the dynamics of the group—bringing people together from different backgrounds with varied and diverse ideas—creates an atmosphere for truly innovative research."

The son of a high school science teacher, Nakagome currently is group leader of a liquid helium refrigerator project in Toshiba's energy science and technology laboratory. Working in concert with the R&D Center's metals and ceramic lab and Tokyo Institute of Technology, his job is to develop a new cryogenic cooling unit that would improve upon conventional helium refrigerators. Liquid helium, with a boiling point of minus 268.9°C, the lowest of any chemical element known to man, is used as a cooling agent to achieve superconductivity. Because it has such a low heat of vaporization, special refrigeration units are needed to slow the dissipation process.

THE NEW REFRIGERATOR LIKELY WILL COOL SUPERCONDUCTIVE MAGNETS OF A MAGLEV TRAIN, MAKING THE TRAINS TWICE AS ENERGY EFFICIENT.

Since the project began in the summer of 1986, Nakagome's 10-man team has been on a virtual nonstop roller coaster. After reaching their initial target of 10 Kelvin, or minus 263°C, within months, they struggled for nearly two years with virtually no progress. Unexpectedly in July 1988, in the heat of summer, the research team achieved an unprecedented minus 268°C, or 5K—which is only 5° above absolute zero. Just as suddenly, the following February, the refrigerator stopped functioning at that "super-cool" level. Then in November, again without warning, the temperature was lowered to 4.5K, only to revert to levels above 5K several months later. "To say we were confused is an understatement," exclaims the mild-tempered Nakagome, who has had to learn patience the hard way. "We tried repeatedly for the next seven months to get back to 5K, and nothing seemed to work. The frustrating thing is that we didn't know why."

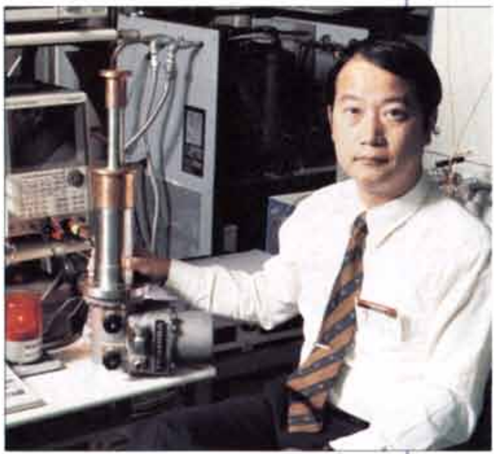
All the while, Nakagome, who joined Toshiba in 1979 after earning his doctorate at Tokyo University, was dogged by lingering questions about the project's findings from several prominent individuals in the Japanese scientific community. Those doubts have only recently have been put to rest.

Toward the end of 1990, through a series of modifications in the refrigerator's heat exchange unit, temperature was gradually lowered to 4K (minus 269°C), then 3K, and finally 2.5K. Although Nakagome would like to go lower, he believes that 2K is probably the limit. "At this point, a more important objective is to improve the refrigerator's performance at 4K." Based on current progress, he expects to complete that task within two years.

Toshiba feels that the key to the project's success to date was the 1988 development of a composite, erbium3-nickel (Er3Ni), a material capable of retaining five times more heat than conventional materials. Unlike lead or copper, which lose their heat preserving qualities at 10K or above, Er3Ni can function at liquid helium temperatures, or below 4.2K (minus 268.9°C). The new material subsequently was adopted for the second stage of the refrigerator's regenerator, a type of heat exchanger. A series of modifications aimed at raising the device's efficiency and minimizing its helium vaporization proved equally significant.

Among the major target applications are magnetic resonance imaging (MRI) equipment for medical diagnostics, magnetically levitated or "maglev" trains, radiowave electronics, and vacuum pumps for semiconductor manufacturing equipment. Longer term, Toshiba sets its sights on space and supercomputer applications.

The new refrigerator likely will be used to cool the superconductive magnets of a maglev train scheduled to begin trial service three years from now in Tsuru, 48 miles west of Tokyo. The planned 27-mile line, funded mainly by the Japanese government, is being built to accommodate trains running at over 310 miles per hour, more than twice the speed of Japan's famous "bullet" train. Engineers at Toshiba's Fuchu Works currently are developing magnets capable of powering trains at speeds of up to 350 miles per hour with 10,000-hour intervals between service. For more than 15 years, the company has supplied niobium and titanium (Nb-Ti) magnets for a 4.3-mile-long experimental maglev service in Miyazaki, Kyushu. Toshiba anticipates maglev trains employing the new cooling unit would be twice as energy efficient at 4K.



SINCE BEGINNING THE PROJECT IN 1986, DR. HIDEKI NAKAGOME'S 10-MAN TEAM HAS BEEN ON A VIRTUAL NONSTOP ROLLER COASTER.

in the case of MRI systems, for which Toshiba is Japan's leading manufacturer, liquid helium refrigerated at 10K must be replenished every three to six months. At 5K, the interval extends to between five and ten years. Similarly, for parabolic radio antennae, such as the one used by Japan's National Radio Observatory in Nobeyama, the weight of the cooling unit can be trimmed by as much as 75% (or approximately 100 kilograms) from the current level. Moreover, the time required to cool the 45-meter antenna's superconducting sensor can be cut from the current 12 hours to three. Finally, for supercomputers, Toshiba predicts that computing speed will increase by as much as 10 times. But industry experts claim that a cooling temperature of minus 269°C, or 4K, is necessary before Josephson computers can be realized.

Nearly a century has passed since a Dutch scientist, working alone, first liquified helium, leading several years later to the discovery of super conductivity. Now, a team from Japan shares a dream of making those discoveries more practical.

Voice Recognition

Dialog Between Man and Machine

Yoichi Takebayashi, *Chief Research Scientist,*
Information Systems Laboratory

A DEMONSTRATION commemorating the 30th anniversary of Toshiba's Research and Development Center last November came off with just one minor glitch: An experimental fast-food ordering system, designed to react to voice command, failed to respond to a foreign visitor's request for a hamburger. Ever alert, one of Toshiba's obliging hosts quickly stepped forward and reordered "hum-bah-gah" in perfect Japanese.

In all fairness, the prototype voice-recognition system was programmed with speakers of Japanese, not English, in mind. Built around an engineering workstation and a 19-inch color display, a simple graphical menu containing icons first meets a customer's eye. A cartoon-like visage appears on screen to take the order. The machine processes all of the menu items by voice command. In a crisp "synthesized" reply, it confirms the order, then extends a friendly thank you. If it doesn't understand, it politely asks the customer to order again.

SPEECH IS JUST PART OF A LARGER MULTI-MEDIA CONCEPT THAT INCLUDES GESTURES, FACIAL EXPRESSIONS, AND EYE MOVEMENTS.

Yoichi Takebayashi, who heads the voice-recognition project from Toshiba's Information Systems Lab, says the system represents "a spontaneous interaction between man and machine." It is "speaker independent," he explains, thus designed for unspecified users and indefinite, unrehearsed speech.

Generally, researchers have found that recognition accuracy in systems not restricted to a predetermined text is lower than if word order is fixed. But for systems to be truly user-friendly, they must deal with the same variations in vocabulary and sentence structure that occur in conversation. The Toshiba approach attempts to meet these conditions. Based on "key words" essential to the completion of the job at hand, the system ignores all other verbal input. The fast-food prototype recognizes just 49 words, including "hamburger," "cheeseburger," "coffee," "cola," "one," "two," and "give me." A more complex system that might be employed at a McDonald's or Pizza Hut would require a vocabulary of approximately 200 words.

Versions of such systems are expected to hit the real world within five years. Standing in the way are a need for more computing power and greater knowledge about spontaneous speed. Just 49 words require 500 megaflops of processing speed; a McDonald's system would need at least four gigaflops, or eight times more capacity.

Toshiba, which has conducted voice recognition R&D since 1977, developed a novel "noise immunity" method to deal with ambiguities of speech and language. Rather than rely on earlier approaches which attempt to minimize background noise, Toshiba has taken the opposite tack, gradually adding many different types of noise collected from the real world and simulated by computer. "Utterances of both intentional and unintentional words are included in the data base," explains the 40-year-old researcher, who earned his doctorate in 1979 from Sendai's Tohoku University. "As a result, we've developed voice-recognition software that is more `robust.' It can understand a wider mix of inflections and can even be 'taught' foreign languages."

In an attempt to measure and eventually raise performance, Toshiba researchers have plotted key word patterns on a "time-frequency" spectrum: Comprising 16 frequency channels and 12 sampling points (between the starting and ending points of words), computations for pattern matching between the input speech and key word reference are performed at intervals of eight milliseconds.

Besides restaurant applications, voice recognition is a prime candidate for airline reservation and ticketing systems, telephone dialing and information services, stock market trading, and automobile and home appliance controls. Simple forms of the technology already have been employed in remote banking systems, elevators, and telephones.

In the automotive field, a potential $1 billion market, Toshiba expects to concentrate on auxiliary functions such as radio tuning ("switch on/off," "volume up/down"), temperature control, and automatic doors and windows. Longer term, it hopes to include the basic car operation instructions: stopping, starting, and activating a growing number of electronic safety devices. Takebayashi points out that as many as half of accidents occur when the driver is not paying attention to driving. "Purely from a safety standpoint, there's a place for this technology."

Takebayashi, who spent several years of postdoctoral studies at the media lab at the Massachusetts Institute of Technology, admits that his research interests extend far beyond the automobile or simple interactive systems for the food service and airline industries. Like his friend Marvin Minsky, the founder of "artificial intelligence," the congenial scientist hopes to uncover the secrets of the mind. Presently using his leisure time to read such diverse subjects as psychology, child development, linguistics, and the body's immune system, Takebayashi declares: "If we ever hope to simulate human behavior by computer, we need to know as much about the human condition as possible."

DR. YOICHI TAKEBAYASHI: "WE'VE DEVELOPED VOICE-RECOGNITION SOFTWARE THAT IS `ROBUST.' IT CAN UNDERSTAND A MIX OF INFLECTIONS."

And speech, he insists, "is just part of a larger multi-media concept that includes gestures, facial expressions, and eye movements. The tone and inflection of the voice also must be considered, as should the almost infinite inventory of human emotions. Uncovering meaning," stresses Takebayashi, "involves much more than simply inputting text. Our objective is a more natural `human-computer' interaction. In the future there will be no boundaries between television, radio, computers, and telephones; they'll combine in a comprehensive network. Which is why," says Takebayashi, "I've chosen to work here."

Superconductors

An Open Door

Jiro Yoshida, *Chief Research Scientist,*
Advanced Research Laboratory

LAST AUGUST, a seven-person research team at Toshiba's Kawaski R&D Center developed what is believed to be the world's first Josephson junction using only oxide superconductors. Unlike conventional materials such as niobium and lead, these substances can perform at significantly higher temperatures—at minus 196°C, the boiling point of liquid nitrogen. "In the process," asserts Jiro Yoshida, the 40-year-old scientist heading one of four superconductor-related research groups at Toshiba, "the door has opened to a whole new field of superconductive electronics."

Though years away from practical application, oxide superconductors promise enormous potential. At the very least, they will likely find use in a future Josephson computer, which was part of a national supercomputer project completed two years ago involving most of Japan's big electronic firms. The technology also is expected to enhance high-sensitivity magnetic field sensors for application in medical diagnostic systems.

Longer term, Yoshida hopes that the technology will serve as a catalyst for a superconductive transistor, a device that could lead to dramatic increases in switching speeds for numerous futuristic information systems. "We're just at the beginning," he explains. "But our success with the Josephson junction is extremely encouraging. This is 21st Century technology."

Yoshida credits the development to advances in thin-film fabrication. By employing a process known as "multi-target" sputtering, the company produced an ultra-thin, three-layer junction. Undetectable to the human eye, the composite film—which is a mere 850 nanometers thick—consists of two layers of yttrium-barium-carbon-oxide sandwiched around a 100-nanometer thick middle layer of praseodymium-barium-copper-oxide.

DR. JIRO YOSHIDA ASSURES THAT A CLEAR MAGNETIC DEPENDENCE—A NECESSARY CONDITION FOR A JOSEPHSON EFFECT—HAS BEEN OBSERVED.

When launching the project in the fall of 1987, Toshiba tried to produce the junction both from high-temperature and conventional low-temperature materials such as lead. Three years later, with clear evidence of a magnetic field effect using the high-temperature material, the company narrowed its focus.

Prior to that discovery, researchers only had succeeded in demonstrating a Josephson effect—the flow of current between two pieces of superconducting material—with superconductors made of niobium and lead, both of which require refrigeration below minus 269°C (the boiling point of liquid helium). With the new material, the temperature has been raised 73°.

Yoshida assures that a clear magnetic dependence—a necessary condition for a Josephson effect—has been observed. More importantly, the current passed through a junction with a 100-nanometer thick middle layer, abnormally thick according to conventional theory. A published report reveals that 100 microamperes of superconducting current have passed through the junction; junction resistance stands at 0.2 ohm.

The Toshiba researcher claims that at least 10 times more superconducting current is needed to reach 1 millivolt output voltage levels (or, 1,000 microamperes at 1 ohm) and resistance must be increased at least five times before the company can apply the technology to magnetic field sensors such as SQUIDs, or superconducting-quantum interference devices. He hopes to achieve both levels by 1995.

THERE'S A VERY REAL POSSIBILITY THAT A COMPLETELY NEW PHYSICS IS HIDDEN IN OXIDE SUPERCONDUCTORS.

The experiences of Yoshida and his team recall those of Walter Brattain nearly half a century ago. When the future Nobel laureate first observed an electrical effect from a needle touching a crystal of germanium, the idea for the point-contact transistor was born. As with Brattain's discovery, no theory explains the effect, no proven causal relationship exists. With no explanation apparent, "There's a real possibility that a completely new physics is hidden in this high-temperature superconductor," asserts Yoshida, who holds a doctorate in physics from Tokyo University.

The Toshiba team is excited about achieving a fundamental breakthrough. "Many Japanese scientists want to do more basic research," says Yoshida. "When America was the world leader in science, all we had to do was apply the US results. Japan now has grown to a point on par with the United States and Europe, so if we don't create ideas of our own, we can't expect progress."

In 1988, prior to Toshiba's Josephson junction success, Japan's International Trade and Industry Ministry launched a 10-year national project on superconductivity. Now in its fifth year, project members will receive close to $24 million this year alone. In addition to Toshiba, other electronics firms teaming with MITI include Fujitsu, Hitachi, Mitsubishi Electrical, NEC, Oki Electric Industry, Sanyo Electric, and Sumitomo Electric Industries. The stakes are high. The company that can advance in the field of superconductive electronics ahead of competitors will be assured of a strong industry position early next century.

High-temperature superconductors clearly offer a number of advantages, cost being among the most significant. Liquid helium currently costs close to $8 per liter in Japan, compared with liquid nitrogen which is just a few cents. Moreover, liquid helium is harder to handle; those limited applications where it so far has found use generally require extensive and costly refrigeration.

Of long-term importance, many observers believe that existing semiconductor technology will reach its practical limits by the end of the decade. "We'll need a breakthrough to go further," warns Yoshida. Superconductive devices are among the more promising alternatives. "Should a superconductive material be developed that demonstrates its effects at room temperature, the world will change in ways we can only imagine."

mologists," Smoot says. Peebles argued that careful measurements of the apparently smooth microwave radiation should show anisotropies corresponding to such phenomena as galaxy clusters and the rotation of the Milky Way.

In 1977 a team headed by Smoot discovered for the first time that the cosmic background was not utterly isotropic but was slightly hotter in one direction of space than the other. The observations were carried out from a U2, an aircraft originally designed for spying. It was Alvarez, Smoot recalls, who suggested using a U2 rather than a balloon and who persuaded the National Aeronautics and Space Administration to make one available.

The so-called dipole anisotropy was surprisingly small, and, odder still, it did not seem to result from the rotation of the Milky Way. The workers concluded that the rotation was being offset by a slightly greater and opposite motion of the galaxy relative to the microwave sea. The discovery of this "peculiar motion" helped lead a decade later to the hypothesis that the Milky Way and other nearby galaxies are being tugged toward some huge, mysterious mass, dubbed the Great Attractor.

Meanwhile, back in 1974, Smoot, Alvarez and others at Berkeley had put together a proposal for a satellite-borne instrument that could make a much more detailed search for cosmic microwave anisotropies. NASA eventually combined the Berkeley proposal with others from the Goddard Space Flight Center and the Jet Propulsion Laboratory into *COBE*. Smoot was given responsibility for the Differential Microwave Radiometers, one of three types of instruments on the satellite. "It came down to who wanted to spend the time doing it and whose design we wanted to follow," Smoot says.

Smoot never envisioned himself as a manager. "I was pretty shy when I was young," he says. "Even at the time of the proposals. I just wanted to do the science." According to some, Smoot evolved into an occasionally abrasive, even autocratic, team leader. One theorist likened him to Carlo Rubbia, who is said to have won the 1984 Nobel Prize in part by relentlessly whipping a team of physicists into finding a particle confirming electroweak theory.

Smoot laughs at the comparison. "It's the office," he says. "As the principal investigator, looking out for the instrument is your job. It takes a while to learn that it is your responsibility, because you're used to cooperating with your fellow scientists, and running a project is a different scheme."

The *COBE* project changed course several times during the next decade. Originally scheduled for launch on an unmanned rocket, *COBE* had to be redesigned when NASA decided to loft it on one of its new fleet of shuttles. Then, in the early 1980s, many theorists came to believe that the universe had undergone a prodigious growth spurt in the instant after the big bang. A related theory held that the universe consisted largely of invisible, slowly moving particles.

These theories, commonly called inflation and cold dark matter, predicted much smaller fluctuations in the microwave background than previous theories had. Smoot fought for and won improvements in the sensitivity of the radiometers. Yet another crisis struck in 1986 when the *Challenger* space shuttle exploded, forcing the *COBE* team to redesign the spacecraft yet again to fit on a Delta rocket.

Smoot hopes that his team can keep gathering data from the COBE *satellite "until it croaks."*

Smoot went to Vandenberg Air Force Base in November 1989 to watch *COBE* be hurled skyward. "You're working so hard to get everything ready and make sure everything goes right, you don't let it cross your mind that it might blow up until it's too late." There was no backup plan; if the satellite had been destroyed, 15 years of work would have been wiped out. "It was a high-risk experiment," Smoot smiles.

As the results started accumulating, Smoot emphasized to the data-processing team that "no matter what we find it's really important. If we find fluctuations that is really important because then we know that dark matter is out there, and if we don't find it that's important because that means gravity is not what's causing structure, some new force is involved. Either one is a big discovery, so let's do it."

After the researchers started seeing islands of structure poking from the sea of noise last summer, Smoot pushed his colleagues to root for possible sources of error. He offered a free round-trip airplane ticket to anywhere to the team member who could find a flaw in the data. "You always have to worry, because you're talking about parts in a million here."

Finally, by the end of the year, he was convinced. How could he be sure? In response to the question, Smoot extracts an inch-thick notebook from a crowded bookcase. It is a compilation of "all the systematic errors I could ever think of over the years. It's not like we just thought about them recently; it was intrinsic to the design of the experiment from the beginning." As he leafs through the notebook, he murmurs, "This is great, this is historical."

As the workers went through the laborious process of writing up the findings—the four papers went through more than 100 revisions before being submitted to journals—they had to keep curious colleagues at bay. "I was trying to keep people confused so we had a little more time to check the results and so forth," he says. "People would say, 'I hear you found the [anisotropies] at such and such a level,' and I'd say, 'That's a little big, don't you think?' "

The announcement, when it came, meant sudden death for a number of cosmological theories, but Smoot has no regrets. "In some sense, it's hard on people to kill their theories," he says, "but on the other hand it's nice to really focus down to a small area. If there's too much freedom in the early universe, the theorists can make up all kinds of things."

Is the finding by the differential radiometers deserving of a Nobel Prize, as some astronomers have suggested? Smoot grimaces before answering the question. "You don't know how things are going to be viewed in the long run," he says. On the other hand, he continues, "How many things do we really know about the early universe? We know the night sky is dark, we know the universe is expanding, we know what the primordial material is, the hydrogen, helium and lithium, and we know about the cosmic microwave background. Those are the relics of the early universe. So this is one more. It's obviously, in terms of cosmology, very important."

Smoot says he is eager to return to the task of analyzing the data still flowing in from the differential radiometers. "We have two years of data, we're into our third year of collection and I have a formal request in for a fourth year. Probably we want to run it until it croaks, because, especially now if the Russians aren't able to launch their satellite, it'll be a long time before anybody builds a better map."

Somehow he also has to write a book about his discovery for William Morrow & Company, tentatively entitled *Wrinkles in Time,* with the help of a science writer. Smoot declines to reveal the size of his brand-new contract. "There is enough of a problem with me getting all this press," he says. "I don't want people thinking I'm rich, too." *—John Horgan*

SCIENTIFIC AMERICAN

Origins of Western Environmentalism

Strategies to preserve nature arose as newly colonialized tropical lands were exploited in the 17th and 18th centuries. Scientists played an important role in this burgeoning concern

by Richard H. Grove

Widespread misgivings about the effects of economic activity on the environment can seem a uniquely modern preoccupation—the result of industrialization, an expanding population and a science sophisticated enough to trace cause and effect. Theodore Roosevelt's pride, the U.S. Forest Service, and the myriad nature refuges established in England by naturalist Nathaniel Charles Rothschild are remembered more as attempts to preserve unspoiled nature than as responses to worries about impending environmental doom.

In truth, the roots of Western conservationism are at least 200 years old and grew in the tropics. Arising in a search for utopia, European-based environmentalism first took shape in the mid-18th century. At that time, colonial enterprise began to clash with Romantic idealism and with scientific findings.

RICHARD H. GROVE has traveled to the Caribbean, South Africa, India and Mauritius, among other places, to study the history of environmental concerns. He holds degrees in geography from the University of Oxford, in conservation biology from University College, London, and in history from the University of Cambridge. Grove, who has worked as a consultant for several organizations, including Friends of the Earth, runs the Global Environmental History Unit at Cambridge. He is a lecturer in environmental history at Churchill College and a research associate at the National Institute of Science, Technology and Development Studies in New Delhi.

The setting for this conflict was the threatened ecology of tropical islands and lands, from the Caribbean Sea to Asia. In London, Paris and other imperial capitals, these islands became allegories for the world at large. The power of this metaphor and the simultaneous emergence of a community of professional natural scientists spurred governments to protect the environment.

The image of an untouched tropical island had long been associated with a Western vision of utopia. In the *Divine Comedy,* for example, Dante Alighieri set earthly paradise in a southern ocean. During the 15th and 16th centuries, voyages by Christopher Columbus and Ferdinand Magellan gave Europe its first glimpse of such islands.

As growing international trade extended Europe's commercial reach, it permitted "exploitation" of these sites for more philosophical needs. Exotic lands were seen as symbols for idealized landscapes: Edens, Arcadias or New Jerusalems. Eventually, as the large, uncharted terrains of India, Africa and America were explored, all wilderness became vulnerable to colonialization by an ever expanding myth.

During the 17th century, the full flowering of what could be called the Edenic island discourse led to the realization that European colonial rule could be environmentally destructive. Agriculture and the harvesting of timber, minerals and game by the government-run Dutch, British and French East India companies began to destroy idyllic terrain. The work of some contemporary artists communicated the extent of this degradation to Europeans. Drawings of Mauritius in 1677, for instance, forcefully depicted the stark reality of felled ebony forests. A coherent awareness of the ecological impact of capitalism and colonial rule began to emerge.

This insight was inextricably linked to the growing social leverage and often radical agenda of the scientific lobby of the time. During the late 17th and early 18th centuries, the urgent need to understand unfamiliar floras, faunas and geologies for commercial purposes attracted many scientists into employment with the trading companies.

These scientists, almost all of whom were medical surgeons or custodians of the early colonial botanical gardens, were an essential part of the administrative machinery of the East India companies. Hendrik B. Oldenland was a case in point: he served as curator of the botanical garden, doctor, town engineer and superintendent of roads for the Dutch Cape Colony in South Africa.

As companies extended territorial acquisitions, the associated research com-

MAP OF ST. HELENA from 1570 portrays the South Atlantic island as a paradise. In the early 1700s, as forests were cut for timber, St. Helena became one of the first sites for conservation.

munity grew proportionately. By 1838 more than 800 surgeons were employed by the British East India Company in India and in the East Indies. As time passed, increasingly complex administrative and technical demands were made on these highly educated and often independent-minded employees.

By the 19th century academies and scientific societies were established throughout the new territories. These institutions made it easy for scientists to communicate and debate their observations of the changes wrought by imperialism. Environmental theories and an ever growing flood of information about natural history and ethnology were diffused through meetings and publications. Thus, at the same time as it had promoted large-scale ecological change, the colonial enterprise had also created a coterie of men—and some women—predisposed to rigorous analytic thinking about the processes of ecological change and the need for land control.

One of the first places where science spurred conservation was Mauritius. Although initially visited by the Portuguese, this island in the Indian Ocean was claimed by the Dutch in 1598; it fell under French rule in 1721 and thereafter became directly associated with the utopic visions of Romanticism and French physiocracy, an economic philosophy based on the "laws of nature" and the methods of Isaac Newton.

The first Dutch settlers were not conservationists. By the time they left, Mauritius's vast hardwood forests were largely depleted in areas easily accessible from the coast. As a result, zealous anticapitalist French reformers attempted to forestall further deforestation. Mauritius became the site for some of the earliest experiments in conservation.

These initiatives were carried out by scientists who, characteristically, were followers of Jean-Jacques Rousseau and the rigorous empiricism associated with the French Enlightenment. They viewed responsible stewardship of the environment as an aesthetic and moral priority as well as a matter of economic necessity. On Mauritius, these men—including Philibert Commerson, Pierre Poivre and Jacques Henri Bernardin de St. Pierre—wished to construct a just society, uncorrupted by absolutist France.

The strategies of these scientists were founded on an awareness of the potentially global impact of modern economic activity. Commerson, a botanist trained by Linnaeus, had been the royal botanist accompanying Louis Antoine de Bougainville on his voyage around the world. Commerson's wife, Jeanne Baret, traveled with him, becoming the first woman to circumnavigate the globe. (She accomplished this feat by starting the trip disguised as a manservant.) As a result of this journey, she inspired her husband, the most experienced botanist of the 18th century, to take a post as state botanist on Mauritius in 1768.

Bernardin de St. Pierre, an engineer who joined Commerson on Mauritius, was shocked by the deforestation he witnessed on his arrival. Yet he also believed he had found an unequaled harmony between people and nature in Mauritius, a harmony celebrated in his novel *Paul et Virginie*, one of the first French Romantic novels. Bernardin de St. Pierre felt that the preservation of

Events and Ideas That Shaped Western Environmentalism

RENAISSANCE — BAROQUE

Year	Event
1271	Marco Polo sets off for Far East and Japan
1305	Dante Alighieri starts *Divine Comedy*
1492	Christopher Columbus reaches West Indies
1498	Columbus sights St. Vincent and Tobago
1502	Portuguese claim St. Helena
1516	Thomas More publishes *Utopia*
1519	Ferdinand Magellan sets out around the world
1560s	Deforestation starts in West Indies
1598	Dutch claim Mauritius
1600	British East India Company founded
1602	Dutch East India Company founded
1607	Jamestown colony established in America
1627	British settle Barbados; auroch extinct in Poland
1633	Dutch claim St. Helena
1640	Dutch take Tobago
1664	French East India Company founded; Dutch mandate forest protection in South Africa
1670s	Dodo extinct in Mauritius
1677	French take Tobago
1713	Redwood protection on St. Helena
1721	French claim Mauritius
1757	British colonial empire begins in India
1763	British get St. Vincent through Treaty of Paris

this harmony demanded rigorous protection of natural resources.

Commerson and Bernardin de St. Pierre in turn stirred the climatic and economic anxieties of Poivre, who had been appointed governor of Mauritius in 1767. An adherent of physiocracy, Poivre believed scientific knowledge should be applied to land management. Originally a Jesuit missionary, he had studied Indian and Chinese forestry and horticultural methods and had been greatly influenced by the botanical gardens and forest protection methods he had observed in the Dutch-run Cape Colony.

In the course of trying to acquire useful medical plants in India, the Dutch learned that their own classification methods were less sophisticated and efficient than the medicobotanical systems of an Indian caste—the Ezhava—from Malabar. Hendrik Adrian van Rheede Tot Drakenstein, who promoted forest conservation in the Cape Colony, organized a translation of the Ezhava texts into Latin. The resulting 12 volumes were published in Amsterdam as the *Hortus Malabaricus,* the garden of Malabar. These books formed the basis of all subsequent European classifications of South and Southeast Asian plants. Recognition of the superiority of the Ezhava system accounted for the first protection of trees and plants by the Dutch.

Although it may have been novel in a Western context, the protection of natural resources has been promoted since time immemorial. This recognition is especially significant today as researchers and others increasingly turn to native peoples for an understanding of the medicinal value of tropical plants or to small-scale efforts to stem desertification. Indigenous strategies have successfully combated soil erosion and deforestation in precolonial East Africa, the Cape Verde Islands, the Kingdom of Ghana and Mauryan India, as well as in the early colonial empires of China and Venice. As early as 450 B.C., for example, Artaxerxes I attempted to restrict cutting Lebanese cedar.

But the central, innovative aspect of French conservationism on Mauritius was the perceived relation between deforestation and local climatic change. A 1769 ordinance incorporated several stipulations prompted by this understanding: 25 percent of all landholdings were to be kept as forest, particularly on steep mountain slopes, to prevent soil erosion; all denuded areas were to be reforested; and all forests within 200 yards of water were to be protected. Eight years later a fully staffed forest service was set up. And in 1803 clearing of forest was forbidden higher than one third up a mountainside.

The early laws were not confined to forests. Pollution of water by effluent from indigo factories and sugar mills engendered more laws in 1791. In 1798 regulations were introduced to control vital but diminishing fish stocks.

The English were quick to imitate the example of the French policy in Mauritius. Again, these efforts were brought about by scientists, this time in the West Indies and in the Caribbean as well as, later, in India. In Tobago, an island in the eastern Caribbean, the work of Stephen Hales and Soame Jenyns was especially important. Hales, a plant physiologist who lived from 1677 to 1761, pioneered the study of transpiration, root pressure, the circulation of sap and the relation between green plants and the atmosphere. His friend Jenyns was the member of Parliament for Cambridge and one of the Lords Commissioners for Trade and Plantations, the group

SCIENTIFIC SOCIETIES urgently called for conservation in the 18th and 19th centuries. For instance, in the 1860s members of the Madras Literary and Scientific Society advocated establishment of Indian forest reserves. Scientists Edward Balfour (*standing*) and Hugh F. C. Cleghorn (*far right*) were instrumental in these efforts.

ENLIGHTENMENT — ROMANTICISM

Year	Event
1764	Forest reserves established on Tobago
1769	French pass conservation laws in Mauritius
1776	American Revolution
1780–85	Severe famines in India
1789–92	French Revolution
1791	Kings Hill Forest Act passed in St. Vincent
1802	British take back Tobago
1810	British take Mauritius from French
1830–33	Charles Lyell publishes *Principles of Geology*
1835	Charles Darwin lands in Galápagos
1835–39	Bad drought in South India
1852	British scientists report on deforestation in India
1856	Madras Forest Department founded
1858	Forest and Herbage Preservation Act passed in South Africa
1859	Darwin publishes *The Origin of Species*
1860s	Drought in India; bird protection laws enacted in Tasmania
1862	South African drought
1864	Indian Forest Service established
1868	Bird protection laws established in Britain
1877–78	Drought in India

that was responsible for settling Tobago.

Using techniques pioneered by Newton, Hales established a clear link between the atmosphere and plant processes. His experiments suggested a causal relation between trees and rainfall. Hales and his colleagues warned against the dangers of deforestation. Citing the examples of Jamaica and Barbados, where clearing for massive plantations had led to extensive soil erosion, Hales urged Jenyns and the Lords Commissioners to protect forests.

As a result, in 1764 forest reserves were established on Tobago. On land settlement maps these areas, which covered about 20 percent of the island, were marked as "reserved in wood for rains." Rain reserves were a revolutionary concept. They still exist today, although somewhat enlarged, as the oldest reserves of their kind in the world.

Similar measures were enacted on the West Indian island of St. Vincent in 1791. The Kings Hill Forest Act also protected the forests for climatic reasons. This piece of legislation was inspired by Alexander Anderson, the curator of the St. Vincent Botanic Garden—the first such garden to be founded in the Western Hemisphere. Anderson, like his French colleagues in Mauritius, was attracted by visions of utopian landscapes peopled by noble savages. By procuring protection of the St. Vincent forests, he hoped to prevent extinctions of species, protect the climate and preserve the island's idyllic quality. In practice, however, this vision was shattered. During the 1790s, the indigenous Carib people were uprooted and the culture stamped out.

The policies developed on Mauritius, Tobago and St. Vincent eventually provided the justification and practical models for the forest planting and protection systems that developed in India after 1847. Until then, it seems that concerns about environmental change had been delayed by the vastness of the subcontinent, which concealed the impact of soil erosion and deforestation.

The roots of environmentalism in India were strongly reinforced by the writings of Alexander von Humboldt, the famous German geographer and explorer. He promulgated a new ecological concept of the relation between people and the natural world: that of the fundamental interrelation of humankind and other forces in the cosmos. His ideas, which drew extensively from the holistic thinking of Hindu philosophers, presented a scientifically reasoned interpretation of the threat posed by unrestrained human activities.

Humboldt's views influenced some of the scientists working for the British East India Company. These men were receptive to a way of thinking that related deforestation, water supply, famine, climate and disease in a coherent fashion. Humboldt based his theories on detailed observations carried out over several years, supplemented by historical records of the level of Lake Valencia in Venezuela.

Several Scottish scientists, including Alexander Gibson, Edward Balfour and Hugh F. C. Cleghorn, became enthusiastic proselytizers of the conservationist message. They advocated establishing a forest system in India that was unequaled in scale. In an 1852 report, they warned that a failure to set up an extensive forest system would result in ecological and social disaster.

The study took a global approach, drawing on evidence and scientific papers from all over the world. Its authors argued that rapid deforestation might cause severe rainfall decline, reduced runoff and ultimately famine. They pointed to widespread deforestation and ensuing soil erosion on the southwestern coast of India, the Malabar Coast.

Their message struck at the heart of the British East India Company's concerns: revenue. The destruction in Malabar had caused commercially important harbors to silt up and become useless. This experience provided early evidence of what might happen in the absence of a state conservation program. (It should be mentioned that early warnings about deforestation in India came as much from indigenous rulers as from scientists. In 1830 the Rajah of Nilumbur alerted the governor of Bombay to the serious consequences of felling too many trees.)

The researchers' activities proved highly alarming to the British East India Company. Officials grasped the association between deforestation and famine fairly quickly, fearful as they always were of agrarian economic failure and social unrest. Unfortunately, it required an initial famine for scientists to gain credibility in the eyes of the government. Only then did the state take measures to protect the environment.

In India, periods of serious drought

DODO, which existed only on the Indian Ocean island of Mauritius, is shown with an unidentified animal. The flightless bird became extinct during the 1670s.

between 1835 and 1839, in the early 1860s and between 1877 and 1878 were all rapidly followed by state programs to strengthen forest protection. The forest conservation system set up in India, which was based in part on the Mauritius experience, later provided the model for most of the state conservation systems in Southeast Asia, Australia and Africa and, later, in North America.

Drought prompted environmental policy in other colonies as well. John Croumbie Brown, a pioneer of conservation in the Cape Colony, secured government agreement to conserve forests and prevent burning of grasslands only after the drought of 1862–1863 wreaked havoc on settler agriculture.

The South African drought of 1862, the worst ever recorded, had implications that extended far beyond conservation policies in Africa. It encouraged the development of an entire school of desiccationist theory that related the colonial experience to the world at large for the first time. Many scientists became convinced that most of the semiarid tropics were becoming arid as a result of colonial deforestation, an idea that has been confirmed by recent study.

Theories of widespread climatic change acquired further credibility in March 1865, when a paper by James Fox Wilson was presented at the Royal Geographical Society in London. The report, "On the progressing desiccation of the Orange River in Southern Africa," made a strong case. Wilson, a naturalist, believed that the Orange River was becoming deprived of moisture and that the Kalahari Desert was expanding. He attributed the desiccation to the "reckless burning of timber and the burning of pasture over many generations by natives."

Present at Wilson's lecture was the explorer David Livingstone. He vehemently disagreed, asserting that rainfall had declined because of natural geophysical phenomena. Another speaker, Sir Francis Galton, a cousin of Charles Darwin, believed the introduction of cheap axes into Africa by Europeans had promoted excessive deforestation and consequent drought. Yet another member of the audience, Colonel George Balfour of the Indian Army—brother to Edward—sounded a more caustic note. Rainfall decline in India, he asserted, was caused principally by the European community, including the plantation owners.

Balfour argued that countermeasures were necessary. He said he had been informed that the government of Trinidad had prohibited cutting trees near the capital in order to ensure a supply of rain. Balfour was quick to point out that in precolonial times it had been the practice of Indians to sink wells and "plant topes of trees" to encourage water retention. In 1866, in another Royal Geographical Society discussion, Balfour cited the example of Mauritius, where "the Government had passed laws to prevent the cutting down of trees, and the result has been to secure an abundant supply of rainfall." Thus, the debate about climatic change had become international in scope by the mid-1860s. Detailed research raising the possibility that the very composition of the atmosphere might be changing reinforced the concerns.

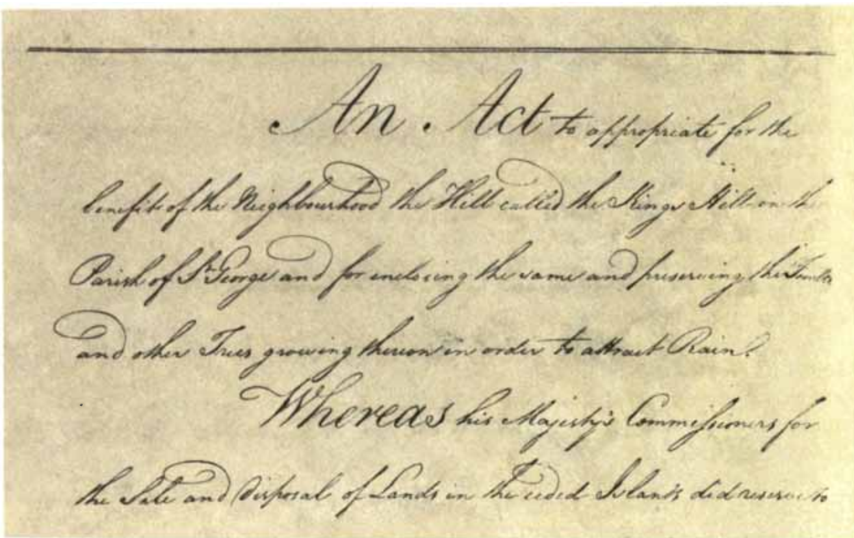
Such views, which presaged contemporary fears about global warming, found early advocacy in the writings of J. Spotswood Wilson. He presented a paper in 1858 to the British Association for the Advancement of Science on the "general and gradual desiccation of the earth and atmosphere." Upheaval of land, destruction of forests and waste by irrigation were not sufficient to explain the available facts on climatic change, Wilson stated in his paper. Instead, he believed, the cause lay in the changing proportions of oxygen and carbonic acid in the atmosphere. Wilson argued that their respective states were connected with the relative rates of their production and absorption by the animal and vegetable kingdom. This paper probably helped to influence the ideas of debaters at the Royal Geographical Society several years later.

Wilson concluded with a dismal set of remarks. Changes in the atmosphere were "in the usual course of geological changes, slowly approaching a state in which it will be impossible for man to continue as an inhabitant.... As inferior races preceded man and enjoyed existence before the earth has arrived at a state suitable to his constitution, it is more probable that others will succeed him when the conditions necessary for his existence have passed away."

Raising the specter of human extinction as a consequence of climatic change was a shocking psychological development in 1858. Yet it was consistent with fears that had been developing among the international scientific community for a long time. The concept of species rarity and the possibility of extinction had existed since the mid-17th century, when the scope of Western biological knowledge began to embrace the tropical world.

The demise of the auroch, a form of wild cattle, in 1627 in Poland and of the dodo in the 1670s in Mauritius had made a considerable impact. In 1680 the Polish government had set aside large areas of forest where hunting was prohibited. The contemporary survival of the wisent, or European bison, is attributed to this isolated effort. And in 1713 attempts were made to prevent the demise of redwood trees on the South Atlantic island of St. Helena.

The publication in the early 1830s by Charles Lyell of the *Principles of Geology* gave firm foundation to the confused awareness of extinction already shared by some East India Company scientists. The book questioned the permanence of species and laid the basis for modern understanding of geological change. Lyell questioned the ideas presented in Genesis, overturning notions about the speed of environmental processes. Paradoxically, this discussion emphasized


An Act to appropriate for the benefit of the Neighbourhood the Hill called the Kings Hill in the Parish of St George and for enclosing the same and preserving the Woods and other Trees growing thereon in order to attract Rain.
Whereas his Majesty's Commissioners for the Sale and Disposal of Lands in the ceded Islands did reserve to

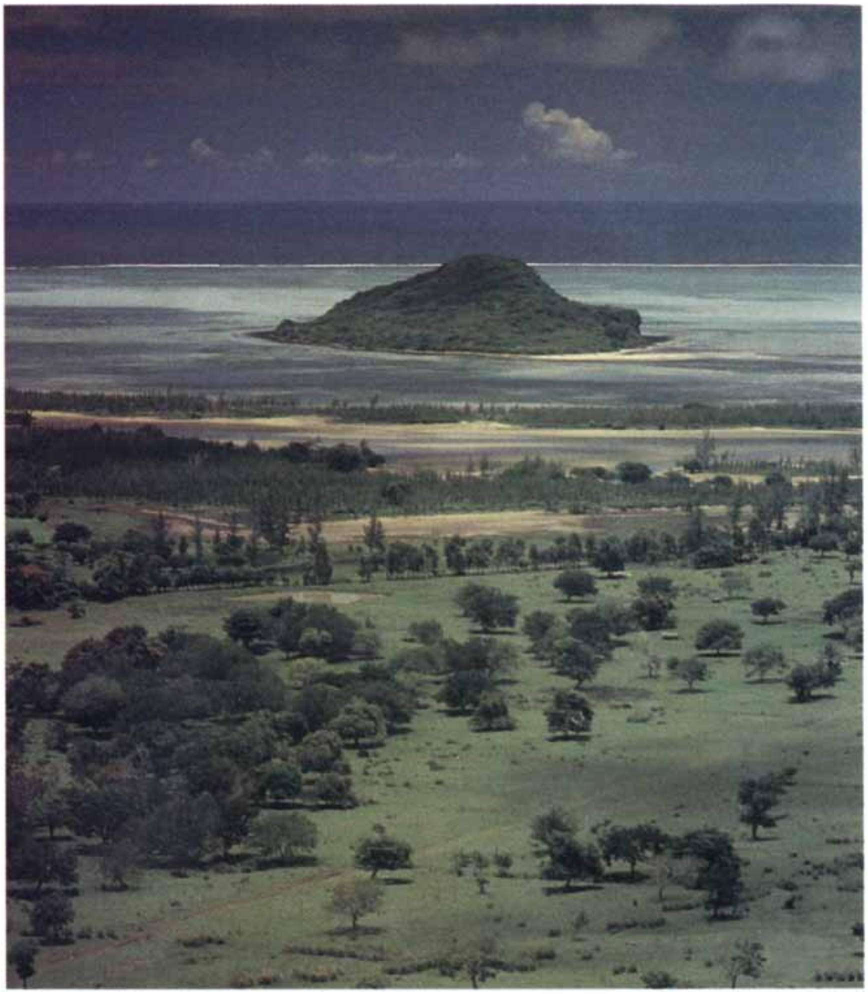
KINGS HILL FOREST ACT of 1791 protected trees on the island of St. Vincent in the West Indies. The British colonial legislation sought to forestall climatic change.

the apparent helplessness of humanity in the face of environmental change.

Other scientists were probing these ideas as well. In the 1840s Ernest Dieffenbach chronicled the fauna of Mauritius, New Zealand and the Chatham Islands, which lie off the coast of Chile. He too became acutely aware of the potential for further rapid extinction if European economic activity spread. Indeed, a paleontologist named Hugh Edwin Strickland, who understood the threat because of his work on the dodo and other vanished birds of the Mascarenes (Mauritius and Réunion islands), suggested that all of New Zealand be made a nature reserve.

The publication of *The Origin of Species* by Darwin in 1859 placed extinction in the dynamic context of natural selection. His theory served to sharpen the predicament of colonial scientists, many of whom were already aware of the part played by humans in hastening the demise of certain species.

A central part of the response to the existential havoc created by *Origin* served to fuel efforts to enact state conservation legislation. For instance, Cleghorn, who was the first inspector general of the Madras Forest Department, which was set up in 1856, stated that uncontrolled deforestation would both cause the loss of valuable species and prevent botanists from assembling evidence for evolution. (He was aware that such arguments might not carry great weight with government, and so he chose to emphasize the more obvious economic hazards of climatic change and resource depletion.)



UNTOUCHED LANDSCAPE in Mauritius today is for the most part the result of the creation of forest reserves by the French in the 1700s.

Origin made protection a more valid concept in the eyes of the government. Indeed, between 1860 and 1870, a flurry of protectionist legislation was enacted in Britain and its colonies. Once again the galvanizing force was an island colony: Tasmania. A comprehensive law designed to protect indigenous Tasmanian birds was introduced in 1860, supported principally by an amateur naturalist, J. Morton Allport.

Other territories rapidly followed suit. By 1865 the colonial legislatures of Natal in South Africa and Victoria in Australia introduced laws to protect several animals and birds. Somewhat belatedly, in 1868, the U.K. introduced its first measures to protect birds. Significantly, the architect of the British measure was Alfred Newton, a frequent correspondent with Allport and the first prominent scientist to recognize the validity of Darwin's theory. Such early measures to protect species, all closely connected to opinions of Lyell and Darwin, offered a symbolic as well as practical opportunity to try to reassert control over a process of environmental degradation that was now understood as global.

By the mid-19th century long-established anxieties about artificially induced climatic change and the loss of species had reached a climax. The spread of Western economic development, initially through colonial expansion, was increasingly seen by more perceptive scientists as eventually threatening the survival of humanity.

If a single lesson can be drawn from the early history of conservation, it is that states will act to prevent environmental degradation only when their economic interests are shown to be directly threatened. Philosophical ideas, science, indigenous knowledge and people and species are, unfortunately, not enough to precipitate such decisions. Time and again, from the 1850s onward, some scientists have discovered that the prospect of artificially induced climatic change, with the full weight of its implications, was one of the few effective instruments that could persuade governments of the extent of an environmental crisis.

Our contemporary understanding of the threat to the global environment is thus a reassertion of ideas that reached maturity over a century ago. It is to be regretted that it has taken so long for the warnings of early scientists to be taken seriously.

FURTHER READING

TRACES ON THE RHODIAN SHORE: NATURE AND CULTURE IN WESTERN THOUGHT FROM ANCIENT TIMES TO THE END OF THE EIGHTEENTH CENTURY. Clarence J. Glacken. University of California Press, 1967.

SCOTTISH MISSIONARIES, EVANGELICAL DISCOURSES AND THE ORIGINS OF CONSERVATION THINKING IN SOUTHERN AFRICA, 1820–1900. Richard Grove in *Journal of Southern African Studies,* Vol. 15, No. 2, pages 163–187; January 1989.

COLONIAL CONSERVATION, ECOLOGICAL HEGEMONY AND POPULAR RESISTANCE: TOWARDS A GLOBAL SYNTHESIS. Richard H. Grove in *Imperialism and the Natural World.* Edited by John M. MacKenzie. University of Manchester Press, 1990.

Building Molecular Crystals

By investigating how molecules assemble to form crystals, researchers hope to synthesize materials with novel electronic, magnetic or optical properties

by Paul J. Fagan and Michael D. Ward

The universe is full of things that take care of themselves. Atoms spontaneously assemble into molecules. Mountains steadily rise out of the sea. Stars continuously cluster into galaxies. Yet nowhere is self-organization more impressive than in crystals.

A snowflake, for example, begins its existence as countless water molecules in the clouds. The molecules initially come together in small groups because of a weak attractive force between oxygen and hydrogen atoms in each molecule. The same forces then organize the groups into a frozen molecular crystal, a perfectly ordered lattice of molecules. Finally, several molecular crystals join together to form a delicate, sparkling flake.

Scientists have realized for some time that the forces that assemble molecules into natural crystals can be utilized to produce a variety of important materials. They have determined the structure of more than 90,000 different molecular crystals. Some are well known, such as aspirin and mothballs. Others exhibit remarkable optical, electronic or magnetic properties—even superconductivity [see "Organic Superconductors," by Klaus Bechgaard and Denis Jérome; SCIENTIFIC AMERICAN, July 1982].

The material properties of a crystal depend greatly on the arrangement of the molecules in the crystal. Yet scientists know little about the factors that control the assembly of such crystals. In recent years, we and many other researchers have investigated how molecules organize themselves to form crystals. By studying such issues, we hope to understand better what types of molecules and what kinds of conditions will produce molecular crystals with unusual and useful properties.

Synthesizing a molecular crystal is similar to the task of designing a building. Before any construction can begin, engineers must specify the shapes and sizes of girders and decide the number and placement of rivets. Likewise, to produce new molecular crystals, chemists must choose molecules of the appropriate shape and size, and they must pick the molecular forces that will hold the crystal together. A chemist can usually find many molecules of various different shapes and sizes; the trick is to find ones that assemble in a predictable manner.

Molecules can attract one another in many ways, if they are designed properly. In general, these attractive forces are weaker than the bonds that form between atoms. Yet some attractive forces are relatively strong and act over a long range. For example, a positively charged atom embedded in one molecule will attract a negatively charged atom in another molecule, and the strength of this force decreases by a factor of only one fourth as the distance between the molecules doubles.

Other attractive forces are weak and influence only nearby molecules. For instance, if two molecules come close together, the charge may be redistributed so that one is negatively charged near the point of contact, whereas the other is positively charged. This interaction, known as an induced dipole, is generally referred to as van der Waals force. The strength of the dipole force decreases by a factor of 128 as the distance between the molecules doubles.

Chemists can exploit most of these attractive interactions to make molecular crystals. They can choose which interactions to use by selecting specific molecules. Once chemists understand how a molecule interacts and know its size and shape, they can begin to guess what kind of molecular crystals it might form. During the 1960s, Gerhard M. J. Schmidt pioneered this approach and coined the term "crystal engineering."

In 1973 Aleksandr I. Kitaigorodsky of the Academy of Sciences in Moscow pointed out one of the most important principles in the design of molecular crystals. He realized that molecules tend to pack together in ways that minimize space. (Molecules typically occupy 65 to 77 percent of the volume of a crystal; the remainder is simply empty space.) Hence, one would expect a "bump" on one molecule to fit into a "hollow" in some other species, or one might predict that a flat area on one molecule will fit flush against a flat area on another.

Molecules assemble in a way that minimizes space because dense packing balances the attractive and repulsive forces. More specifically, the arrangement that requires the least space is likely to be the one that minimizes the total energy of all the forces among all the molecules.

Although such principles are important, it is very difficult to predict exactly

PAUL J. FAGAN and MICHAEL D. WARD began collaborating six years ago to construct molecular crystals. Fagan is a scientist in the central research and development department of the E. I. du Pont de Nemours and Company, where he has worked since 1982. He earned his Ph.D. in inorganic chemistry from Northwestern University. Ward, who received his Ph.D. in chemistry from Princeton University, is associate professor in the department of chemical engineering and materials science at the University of Minnesota. He conducted research at Du Pont from 1984 to 1990.

DESIGNING MOLECULAR CRYSTALS requires building blocks that combine in predictable ways. Flat molecules known as HCTMCP (*top left*) pack closely next to cylindrical molecules called $Fe(MES)_2$ (*bottom left*). The flat molecule has a charge of –2 and is attracted to the cylindrical molecule, which has a charge of +2. The molecules stack to form columns, and the columns fit together to form the crystal (*right*). The arrangement maximizes the attractive forces among the molecules.

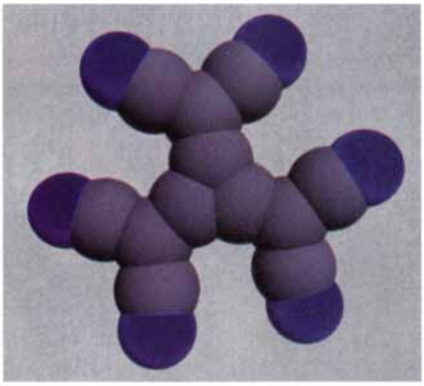
what type of crystal will form from any particular combination of molecules. Chemists would be able to make reliable predictions if they could deduce all the molecular interactions in any given lattice. Using this information, they could calculate the total energy of each possible arrangement and could predict the lowest energy configuration.

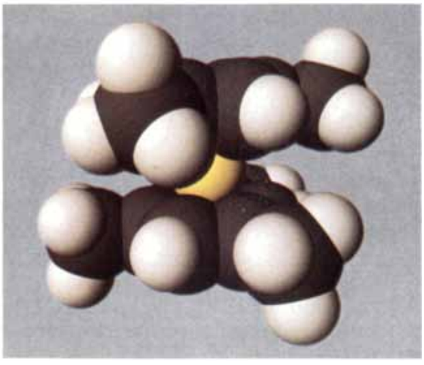
Unfortunately, the molecular interactions in a lattice become too complicated for any human or computer to handle. There are simply too many atoms involved, too many different kinds of interactions and too many possible arrangements. Furthermore, although the strength of the interactions could be approximated, such results might not be very helpful, because minute details can play a large role in determining structure.

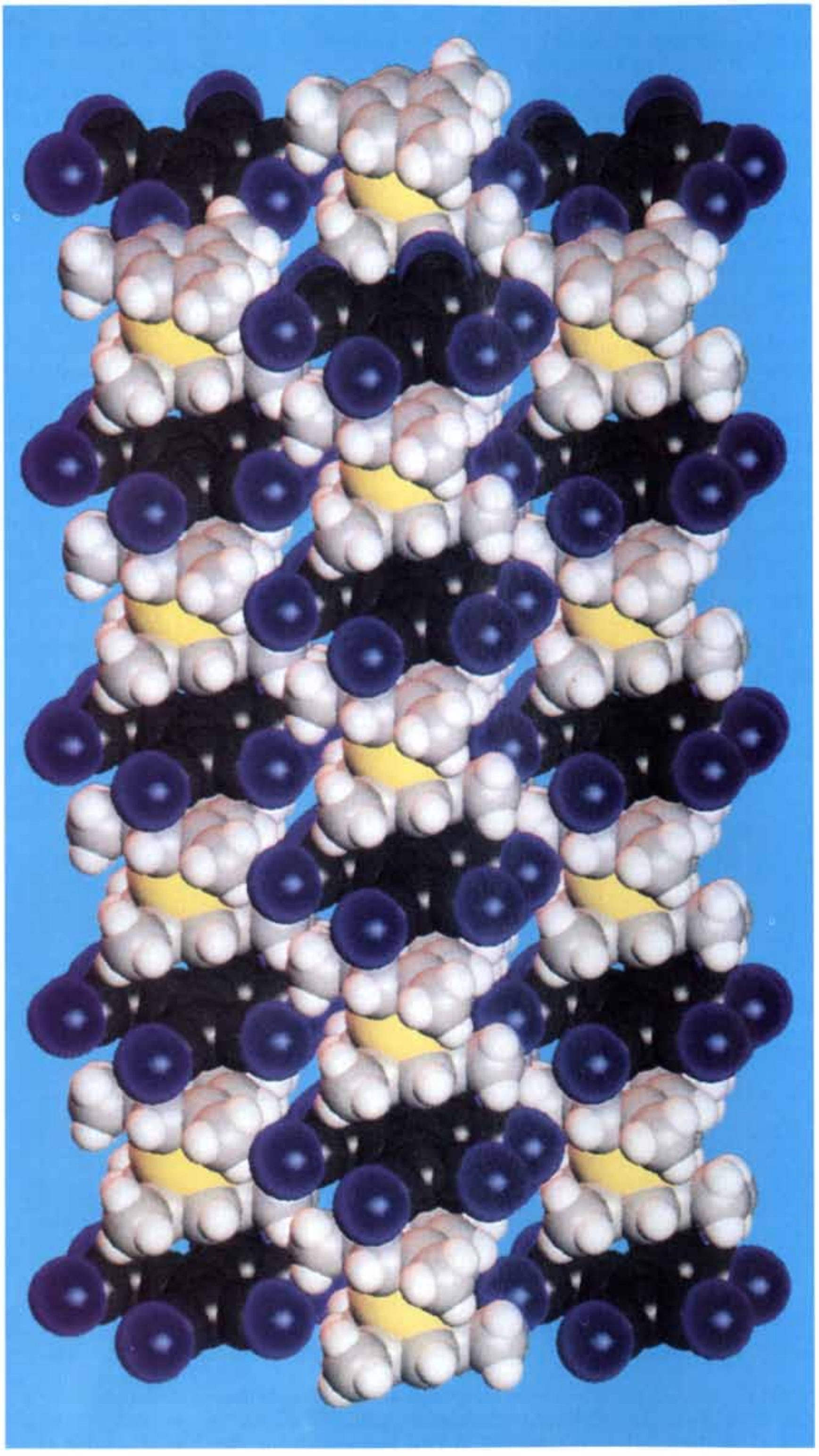
Even if chemists could somehow specify the preferred crystal structure for a given set of molecules, they might still not know how to produce such a crystal. For researchers have only begun to understand what conditions influence the structure of a molecular crystal as it grows.

As a start, investigators have identified several stages of crystal growth. Initially, molecules assemble into small aggregates that may differ slightly in structure. An aggregate may or may not combine with others depending on its shape and the ambient conditions. Those aggregates that do combine readily will become part of a larger aggregate and eventually become incorporated in a sizable crystal. In general, the growth of a molecular crystal depends on the structure of the aggregates and the ease with which they assemble.

During the past decade, chemists have deduced, in several cases, how the structures of molecules and their aggregates may be related to the structures of crystals. These insights have come from such researchers as Gautam R. Desiraju of the University of Hyderabad in India, Richard Robson of the University of Melbourne in Australia, Margaret C. Etter of the University of Minnesota and Angelo Gavezzotti of the University of Milan in Italy.







Desiraju has shown how molecules containing chlorine atoms can be guided into certain generic arrangements by exploiting the attractive interactions between those atoms. Robson has found that low-density crystalline structures tend to form when nitrogen atoms in metal compounds bond to metal ions. Etter has investigated how molecular structure is influenced by hydrogen bonding, an attractive interaction in which hydrogen atoms are shared between such atoms as oxygen.

Gavezzotti has examined aromatic rings, molecules characterized by special kinds of bonds between carbon atoms. For some time, researchers have realized that such aromatic rings as benzene and naphthalene pair up to form herringbone patterns. Gavezzotti has found, however, that if aromatic rings have high ratios of carbon atoms to hydrogen atoms, they generally organize into stacks of molecules in the crystalline lattice.

In 1987 we began to develop a different strategy to study how molecules assemble to form crystals. We searched for molecular ingredients that would combine in simple and predictable ways. Holding the crystal together required using the simple attractive force between negatively and positively charged atoms in a molecule. We therefore designed charged molecules with well-defined shapes that could pack easily into a crystalline lattice. In addition, we devised molecules that were identical in shape and size but had different electronic properties.

During the past five years, our strategy has yielded several novel materials, and we have gained many insights into how molecules combine to form aggregates and ultimately crystals.

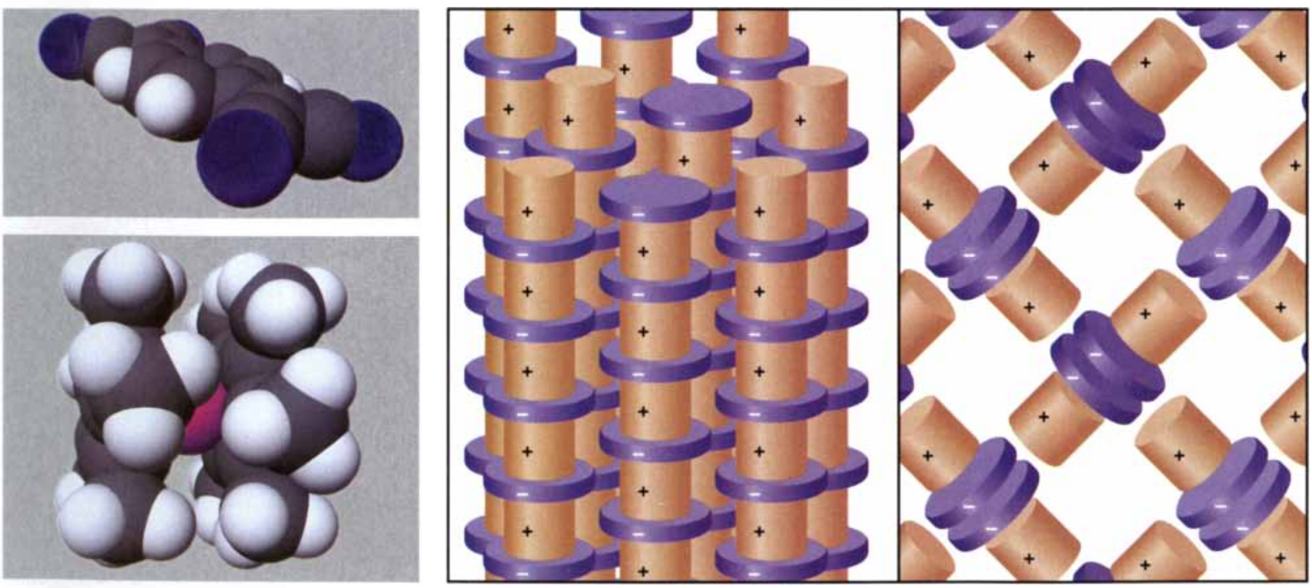
Most of our work involves two species of molecules: polycyanoanions and organometallic compounds. The polycyanoanions are a species of flat molecules containing carbon and nitrogen. The organometallic compounds are made from carbon, hydrogen and metal atoms such as iron. We selected these two families of molecules for many reasons.

We chose polycyanoanions mainly because they assemble in predictable ways. The polycyanoanions can hold a charge of –1 or –2 and can donate electrons to other molecules, resulting in a weak attraction between the molecules. Under some conditions, polycyanoanions can share electrons with each other, forming pairs or what chemists call a dimer. Under other conditions, polycyanoanions can share electrons among many molecules, and they then stack like pancakes to produce extended chains. By pairing or stacking together, the molecules share their outermost electron, thereby lowering the total energy of the arrangement. We were also intrigued by polycyanoanions because we suspected that they would form crystals with interesting electronic properties.

The negatively charged polycyanoanions combine readily with positively charged organometallic compounds. In particular, we decided to work with a subset of the organometallic molecules that chemists sometimes call "sandwich" compounds. In these molecules, each metal atom is nestled between rings of carbon atoms, and each can hold a charge of +1 or +2. The sandwich compounds include molecules with a wide variety of shapes and charge distributions. The simplest sandwich compounds are cylindrical molecules with a single charge in the center. The group also includes rod-shaped molecules with charges at both ends, tetrahedral molecules with charges at each vertex and even cubic or octahedral molecules.

Sandwich molecules have an important advantage. Whereas the carbon and hydrogen atoms determine the overall shape of sandwich compounds, the metal atoms are responsible for most of the electronic properties. Hence, two sandwich molecules might have identical structures but different electronic properties because one of the molecules was based on say, iron, and the other contained ruthenium. By using such pairs, we hoped to tune the electronic properties of molecular crystals.

When positively charged sandwich molecules are combined with negatively charged polycyanoanions, they attract one another, forming bound aggregates. The form of these aggregates—and of the crystals derived from them—derives from the shape of the sandwich molecules, the arrangement of charges within the molecules and the growth conditions. The crystal structures are also influenced by the electroneutrality prin-



SIMPLE MOLECULES tend to assemble into many different simple structures. The cylindrical molecule called $(Cp^*)(Ru^{1+})$ (HMB) (*bottom left*) can combine in two ways with the flat molecule $TCNQ^{1-}$ (*top left*). The molecules can stack together (*center*), or the two compounds can assemble to form a herringbone structure (*right*).

ciple, which requires that the number of negative charges equal the number of positive charges.

For the most part, we conducted our experiments with these molecules in a logical order. We started with molecules with the simplest shapes and charge distributions and worked our way up to the more complicated molecules.

The simplest molecular ingredient we used is a sandwich compound that is cylindrical and has a single positive charge in its center. The top of the molecule is a structure called pentamethylcyclopentadiene, or Cp*. It is a flat ring of five carbon atoms, each of which is bonded to a methyl group (that is, a carbon atom and three hydrogen atoms). The bottom of the molecule is hexamethylbenzene, or HMB. It is a flat ring of six carbon atoms instead of five, and as in Cp*, each of the carbon atoms is decorated with a methyl group. At the center of the molecule lies a ruthenium atom, which bonds to the carbon atoms in both Cp* and HMB. The ruthenium atom holds a charge of +1 and is surrounded by the carbon and hydrogen atoms. The entire molecule is abbreviated as (Cp*)(Ru^{1+})(HMB). (The superscript 1+ indicates the charge of the molecule.)

We wished to combine this cylindrical molecule with a polycyanoanion called tetracyanoquinodimethanide, or $TCNQ^{1-}$. The molecule consists of a flat network of 12 carbon atoms with four nitrogen atoms and four hydrogen atoms at the edges. TCNQ can carry a charge of -1.

We prepared $TCNQ^{1-}$ by placing a negatively charged electrode in a solution containing neutral TCNQ and (Cp*)(Ru^{1+})(HMB). This process yields two different types of molecular crystals: one is colored emerald green and the other, dark purple. Both kinds of crystal have the same molecular formula, but they differ in molecular organization.

To determine the structure of the two materials, we and Joseph C. Calabrese of Du Pont used a technique called x-ray crystallography. We bombarded the samples with x-rays and recorded the radiation that scattered off the atoms in the sample. With this information, we deduced the arrangement of atoms in the crystal.

The green crystals consist of columns that are parallel to one another. Each column contains both $TCNQ^{1-}$ and (Cp*)(Ru^{1+})(HMB). The flat surfaces of the $TCNQ^{1-}$ molecules fit flush against the carbon rings at the top and bottom of (Cp*)(Ru^{1+})(HMB). The two molecules alternate along the length of the column.

ELECTRODE in a glass flask charges TCNQ molecules in a solution, thereby promoting the growth of molecular crystals.

The purple crystals have a more complicated structure. The $TCNQ^{1-}$ molecules are joined together in pairs, one flat surface against the other. Each pair is sandwiched between two molecules of (Cp*)(Ru^{1+})(HMB). These cylindrical units then pack together in a herringbone pattern, so each unit is roughly perpendicular to its nearest neighbors [*see illustration on opposite page*].

We were not surprised that $TCNQ^{1-}$ and (Cp*)(Ru^{1+})(HMB) combined to form more than one type of molecular crystal. Because these molecules have a simple shape and charge distribution, they are free to assemble in at least two different ways.

We next examined what would happen if TCNQ was combined with a rodlike molecule with charges at each end. We searched for a molecule with this form and decided on a compound abbreviated as (Cp*)(Ru^{1+})(CYC)(Ru^{1+})(Cp*). This molecule is really nothing more than two (Cp*)(Ru^{1+})(HMB) molecules connected by two short bridges composed of carbon and hydrogen atoms.

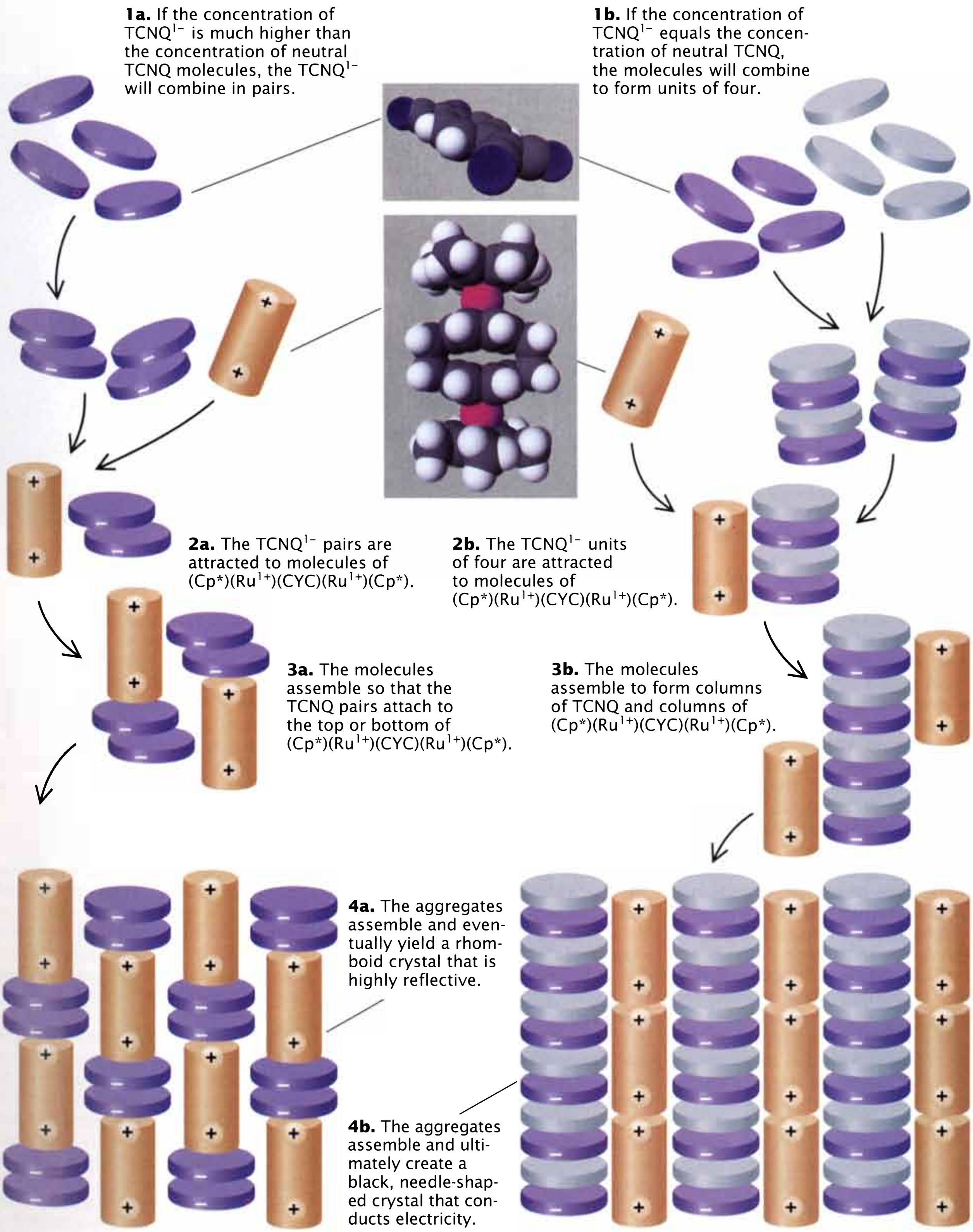
The top and bottom of the molecule consist of Cp*, the five-carbon ring with five methyl groups. Each Cp* is bonded to a ruthenium atom carrying a charge of +1. The ruthenium atoms are connected by cyclophane (CYC), which consists of two rings connected by two bridges, all made of carbon and hydrogen. Hence, the molecule (Cp*)(Ru^{1+})(CYC)(Ru^{1+})(Cp*) is roughly cylindrical, with the charged ruthenium atoms embedded near the ends.

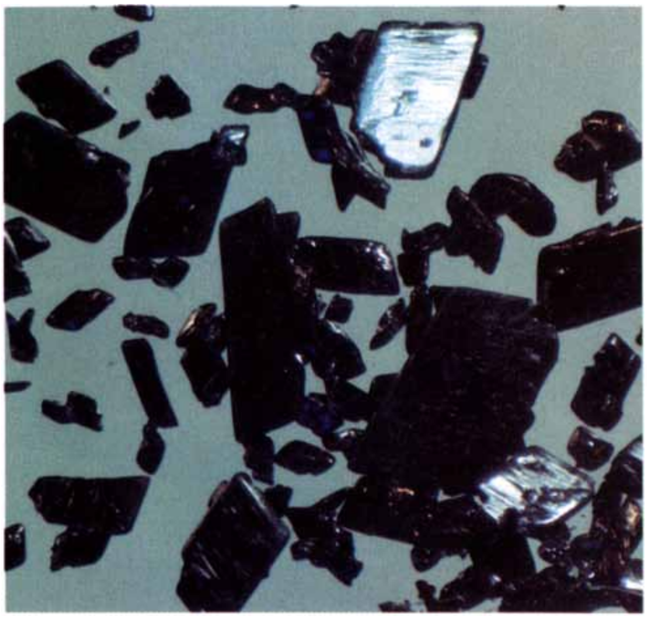
We mixed solutions of (Cp*)(Ru^{1+})(CYC)(Ru^{1+})(Cp*) and neutral TCNQ. We again formed $TCNQ^{1-}$ using a negatively charged electrode. When the electrode was set at a high negative voltage, we noticed that purple crystals formed on the electrode. Yet when the electrode was tuned to a low negative voltage, black, needle-shaped crystals grew.

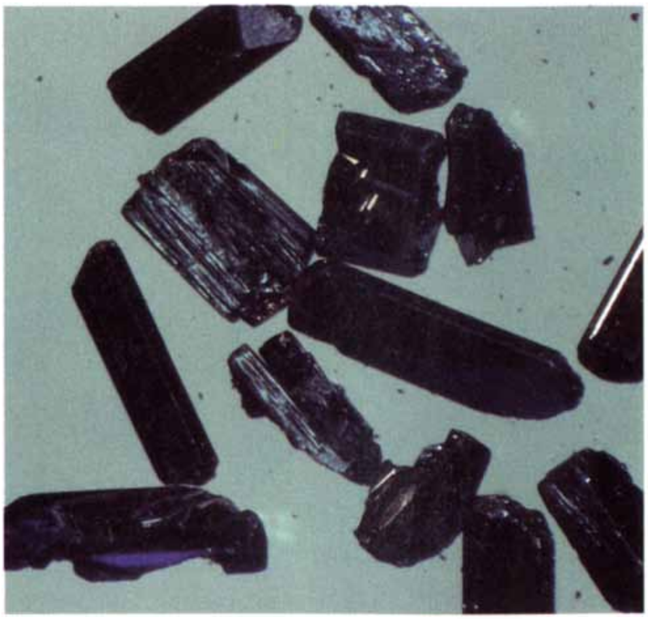
By analyzing the structure of purple and black crystals, we were able to deduce why these two different crystals form under different conditions. When the electrode is at a high negative voltage, many TCNQ molecules with a -1 charge are produced. These charged molecules then combine in pairs to form dimers. Each dimer, having a -2 charge, is then attracted to (Cp*)(Ru^{1+})(CYC)(Ru^{1+})(Cp*), which has

How Conditions Influence the Growth of Crystals

Molecular crystals with two different structures can be grown from a solution containing the cylindrical molecule (Cp*)(Ru^{1+})(CYC)(Ru^{1+})(Cp*) and the flat compound TCNQ. The growth process depends on how much electric current flows through the solution. A strong current produces many TCNQ molecules with a charge of −1. A weak current generates only a few TCNQ molecules with a −1 charge.







RHOMBOID CRYSTALS (*left*) were made from the molecules TCNQ and (Cp^*)(Ru^{1+})(CYC)(Ru^{1+})(Cp^*), as were the needle-shaped crystals (*right*). The crystals differ in their molecular structure (*see box on opposite page*).

a +2 charge. The flat side of each dimer fits against the flat surface of the ruthenium molecule to form a long cylindrical aggregate. These units then stack on top of one another to form columns in which each dimer is sandwiched between two ruthenium molecules. The columns align alongside one another to produce a purple, rhomboid crystal.

When the electrode is at a low negative voltage, molecules of TCNQ with a charge of –1 are formed rather slowly. Hence, the concentration of $TCNQ^{1-}$ is comparable to that of the neutral TCNQ molecules. Under these conditions, two charged TCNQ molecules combine with two neutral TCNQ molecules to form a stack of four. Each of these units, having a charge of –2, then aligns alongside one of the (Cp^*)(Ru^{1+})(CYC)(Ru^{1+})(Cp^*) molecules, which have a charge of +2. This alignment maximizes the attractive force between the unit of four TCNQ molecules and the ruthenium molecule because the stack of four TCNQ molecules and the (Cp^*)(Ru^{1+})(CYC)(Ru^{1+})(Cp^*) are equal in length. These aggregates then assemble in a manner that resembles the closing of a zipper. This process produces columns consisting of either TCNQ or ruthenium molecules. The columns assemble parallel to one another to form a black, needle-shaped crystal.

The black crystal differs from the purple crystal not only in structure but also in electronic properties. In the purple crystal, charge is confined within the TCNQ and ruthenium molecules. In the black crystal, every two TCNQ molecules share a charge of –1, and the TCNQ stacks are not interrupted by the ruthenium molecules. As a result, electrons move through the columns of TCNQ easily, and the crystal is a good conductor of electricity [*see box on opposite page*].

This work demonstrated that a sandwich compound can define the range of possible molecular structures, whereas the experimental conditions determine which possibility actually emerges. In both cases, the structure of the molecular crystal is established by the aggregates that form during the early stages of growth.

Next we wanted to find out whether more complicated charge arrangements would combine in predictable ways. We used a tetrahedral sandwich molecule whose vertices hold positive charges. The molecule is based on ruthenium atoms, each of which holds a charge of +1. The ruthenium atoms are held in a tetrahedral configuration within a network of carbon and hydrogen atoms.

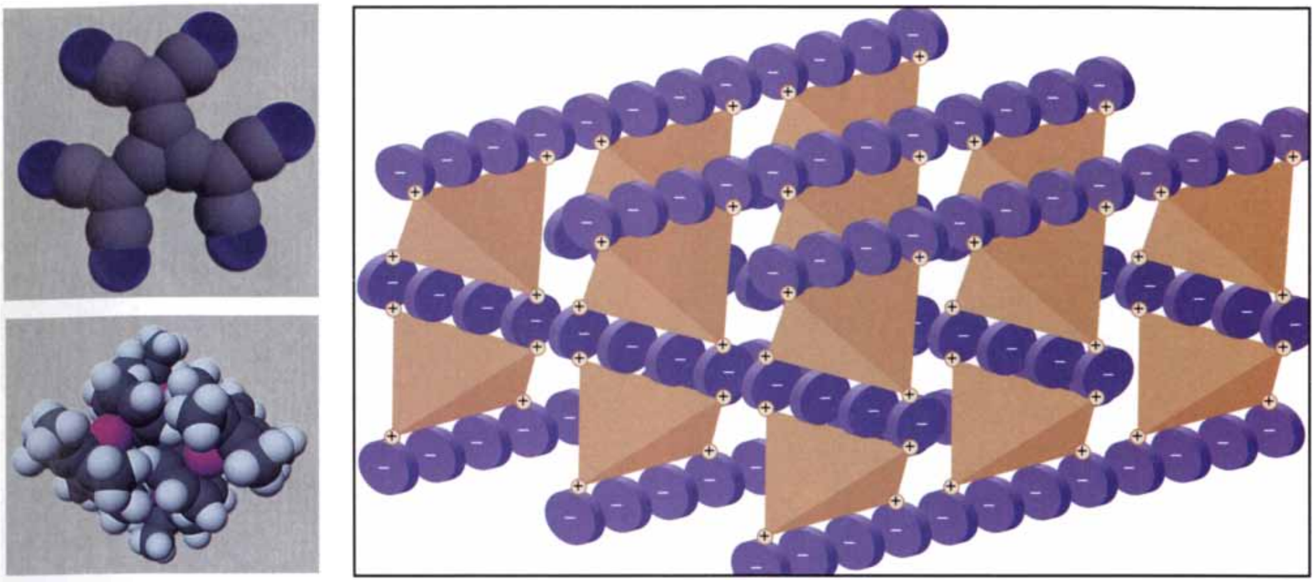
We combined this tetrahedral molecule with a polycyanoanion known as hexacyanotrimethylenecyclopropanide, or HCTMCP. This molecule consists of three Y-shaped branches, projecting from a triangular ring of three carbon atoms. These branches are made of three carbon atoms and two nitrogen atoms, which sit at the tips. The HCTMCP molecule is flat and can carry a charge of –1.

When we combined the negatively charged HCTMCP molecules with the positively charged tetrahedral molecules, they formed a crystal with a predictable structure. The crystal consists of units of four $HCTMCP^{1-}$ and one tetrahedral molecule. Two of the positive charges within the tetrahedral molecule encourage $HCTMCP^{1-}$ to stack in one direction, whereas the other two positive charges force $HCTMCP^{1-}$ to stack perpendicular to that direction. The result is an interweaving network of $HCTMCP^{1-}$ columns between tetrahedral molecules [*see illustration on next page*]. In these experiments, we believe the charge distribution of tetrahedral molecules is the governing factor in the growth of the crystal.

We and our colleagues continue to investigate how negatively charged molecules combine with sandwich compounds that have even more complicated charge distributions than that in the tetrahedral molecule. But we have also been interested in influencing the growth of crystals by changing the magnitude of the charge instead of the distribution of the charge.

To study this issue, we chose (Fe^{2+})$(MES)_2$, a cylindrical molecule with a single iron atom resting in the center that holds a charge of +2. The iron atom is bonded to two mesitylene rings (MES). Each MES ring consists of six carbon atoms connected to three hydrogen atoms and three methyl groups.

We combined (Fe^{2+})$(MES)_2$ with $HCTMCP^{1-}$. Each $HCTMCP^{1-}$ prefers to share one of its electrons with another $HCTMCP^{1-}$. Consequently, the HCTMCP



COMPLICATED MOLECULES tend to assemble into complex but predictable structures. The molecule called $[(Cp^*)(Ru^{1+})(C_6H_5)]_4C$ (*bottom left*) combines with the flat compound HCTMCP (*top left*), which is negatively charged. The molecule $[(Cp^*)(Ru^{1+})(C_6H_5)]_4C$ contains four positive charges, which are arranged at the vertices of a tetrahedron. These molecules encourage the HCTMCP molecules to stack in an array of interweaving columns (*right*).

molecules tend to stack together to form columns. Each HCTMCP column is then surrounded by columns of $(Fe^{2+})(MES)_2$, creating a crystal.

We obtained dramatically different results, however, when $(Fe^{2+})(MES)_2$ was combined with HCTMCP with a charge of –2. The $HCTMCP^{2-}$ molecule will not share its electrons with others and repels them. Rather $HCTMCP^{2-}$ prefers to stack on top of $(Fe^{2+})(MES)_2$. This preference leads to columns whose components are alternately $HCTMCP^{2-}$ and $(Fe^{2+})(MES)_2$. The columns assemble in such a way that each $HCTMCP^{2-}$ has eight neighbors: two $(Fe^{2+})(MES)_2$ in the same column and six $(Fe^{2+})(MES)_2$ in adjacent columns [*see illustration on page 49*].

In this crystal, $HCTMCP^{2-}$ prefers to share its electrons with $(Fe^{2+})(MES)_2$. As a result, a small amount of electric charge is transferred from the $HCTMCP^{2-}$ molecule to the $(Fe^{2+})(MES)_2$ molecule. This transfer reduces the total energy of the electrons. Ultimately, the process favors the formation of columns that contain both $HCTMCP^{2-}$ and $(Fe^{2+})(MES)_2$ molecules. This experiment showed that the assembly of molecular crystals depends not only on the structure of the template molecules but also on the electronic states of the constituents.

The crystal composed of $(Fe^{2+})(MES)_2$ and $HCTMCP^{2-}$ is remarkable for another reason. We can essentially change the electronic properties of the crystal without altering its structure. We can do so by creating a crystal of $(Ru^{2+})(MES)_2$ and $HCTMCP^{2-}$ instead of $(Fe^{2+})(MES)_2$ and $HCTMCP^{2-}$. Both kinds of crystals have the same structure. Yet whereas the iron-based crystal absorbs red light, the ruthenium crystal captures blue light. The ruthenium atom influences the electronic structure of the crystal in such a way that the crystal absorbs blue instead of red light.

We could also produce a crystal that absorbs both red and blue light when HCTMCP was mixed together with $(Ru^{2+})(MES)_2$ and $(Fe^{2+})(MES)_2$. And by altering the proportion of iron to ruthenium, we could control the amount of absorption of each color, thereby changing the color of the crystal.

This example demonstrates an important property of molecular crystals based on organometallic compounds. For the most part, the structure of an organometallic compound and the molecular crystal derived from it do not depend on whether the compound is based on iron, ruthenium or some other metal atom. The structure is determined mainly by the outer shell of carbon and hydrogen atoms. Yet the electronic properties of such a crystal will be influenced by the metal atoms, and in many cases, the effect of the metal atoms can be predicted. We believe that tuning the electronic properties of molecular crystals in this manner will lead to novel materials.

Crystal engineering offers exciting possibilities in several areas of technology. The techniques may make it feasible to fabricate molecular crystals with large pores. Such structures could be considered part of the class of zeolites [see "Synthetic Zeolites," by George T. Kerr; SCIENTIFIC AMERICAN, July 1989]. Porous molecular crystals could serve as size-selective filters for molecules, or they might be able to contain guest molecules, yielding materials with useful structural or electronic properties. Other approaches to crystal engineering may result in new materials with unique mechanical, optical or electronic characteristics. Molecular crystals may be used for magnetic storage, laser systems or electronic devices.

Our work demonstrates that it is possible to control the organization of matter in a reasonably predictable way using the attractive interactions between positively and negatively charged molecules. A tantalizing question is whether we can apply our knowledge to fabricate structures that are a few billionths of a meter in size. If so, the future may bring a new generation of electronics and minute mechanical devices.

FURTHER READING

MOLECULAR CRYSTALS AND MOLECULES. Aleksandr I. Kitaigorodsky. Academic Press, 1973.

LINEAR-CHAIN CONDUCTORS. Arthur J. Epstein and Joel S. Miller in *Scientific American*, Vol. 241, No. 4, pages 52–61; October 1979.

CRYSTAL ENGINEERING: THE DESIGN OF ORGANIC SOLIDS. Gautam R. Desiraju. Elsevier Science Publishing, 1989.

We Fit Hundreds Of Advancements Into A 4x4 Area.

The New Jeep® Grand Cherokee Limited

There's Only One Jeep®...

Advantage: Chrysler

Jeep Eagle

OWNER'S CHOICE PROTECTION PLAN 7/70 POWERTRAIN OR 3/36 BUMPER-TO-BUMPER

Chrysler's Owner's Choice Protection Plan includes: 7-year/70,000-mile powertrain protection or 3-year/36,000-mile bumper-to-bumper warranty. Your choice. And with no deductible. See limited warranties, restrictions, and details at dealer. Excludes normal maintenance, adjustments, and wear items. For further information about Jeep Grand Cherokee, call 1-800-JEEP-EAGLE. Jeep is a registered trademark of Chrysler Corporation.

Buckle up for safety.

OFFICIAL SPONSOR OF THE 1992 U.S. OLYMPIC TEAM 36 USC 380 USA

G Proteins

Tucked into the internal surface of the cell's outer membrane, these versatile molecules coordinate cellular responses to a multitude of signals that impinge from without

by Maurine E. Linder and Alfred G. Gilman

For a person to think, act or merely exist, the cells of the body must communicate with one another. They do so by dispatching such chemical messengers as circulating hormones and neurotransmitters. Perhaps surprisingly, few of these messengers actually enter target cells to induce changes in function. Instead, in a cellular version of the game "telephone," most of the agents convey information through intermediaries. They issue orders by binding at the outer surface of target cells to proteins that serve as specific receptors. Then, in a process called signal transduction, the receptors, which span the cell membrane, relay the information to a series of intracellular middlemen that ultimately pass the orders to the final executors.

Investigators have discovered dozens of extracellular messengers. Each could conceivably initiate a distinct series of molecular interactions. Yet it turns out that a great many messengers rely on just one class of molecules, called G proteins, to direct the flow of signals from the receptor to the rest of the cell. G proteins are so named because they bind to guanine nucleotides, which like all nucleotides consist of an organic base (in this case guanine), a sugar and one or more phosphates.

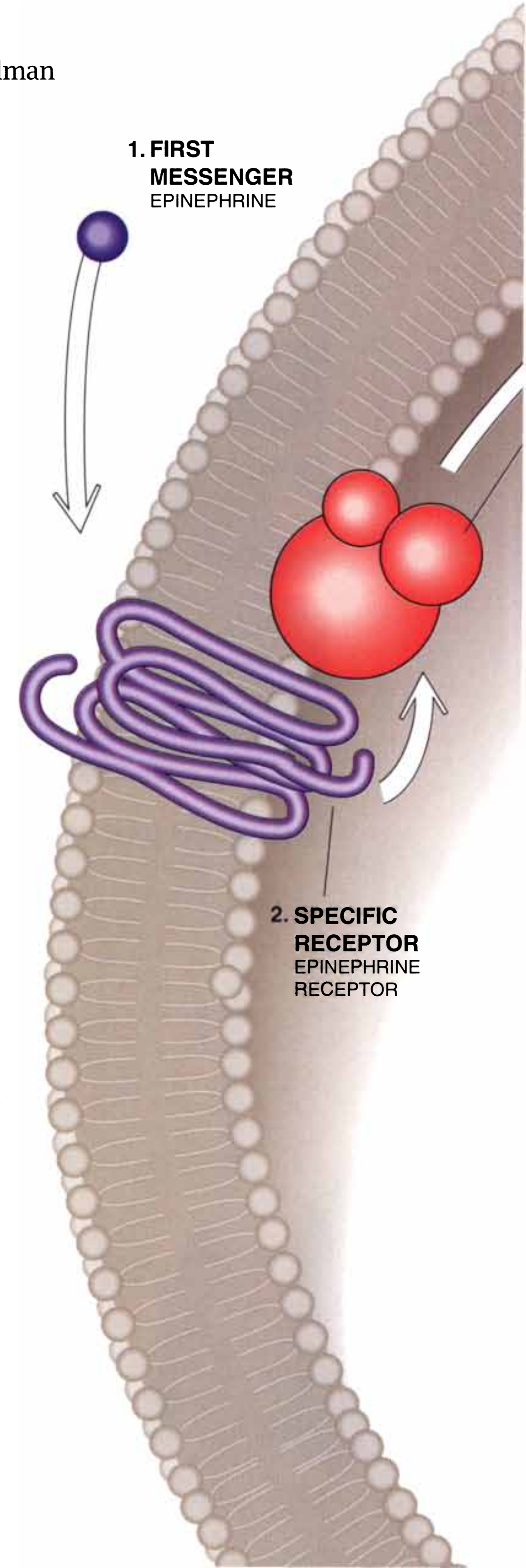
One of us (Gilman) and his colleagues identified G proteins as crucial intermediaries in signal transduction in the late 1970s at the University of Virginia. Soon after, Gilman moved to the University of Texas Southwestern Medical Center at Dallas, where his group has since deciphered a good deal of what is known about how these remarkable proteins work. We continue to be fascinated by their machinations, as well as by the central role they play in an ever increasing array of cellular activities—from mating in yeast, to chemically induced movement in slime molds, to vision, smell, hormone secretion, muscle contraction and cognition in humans.

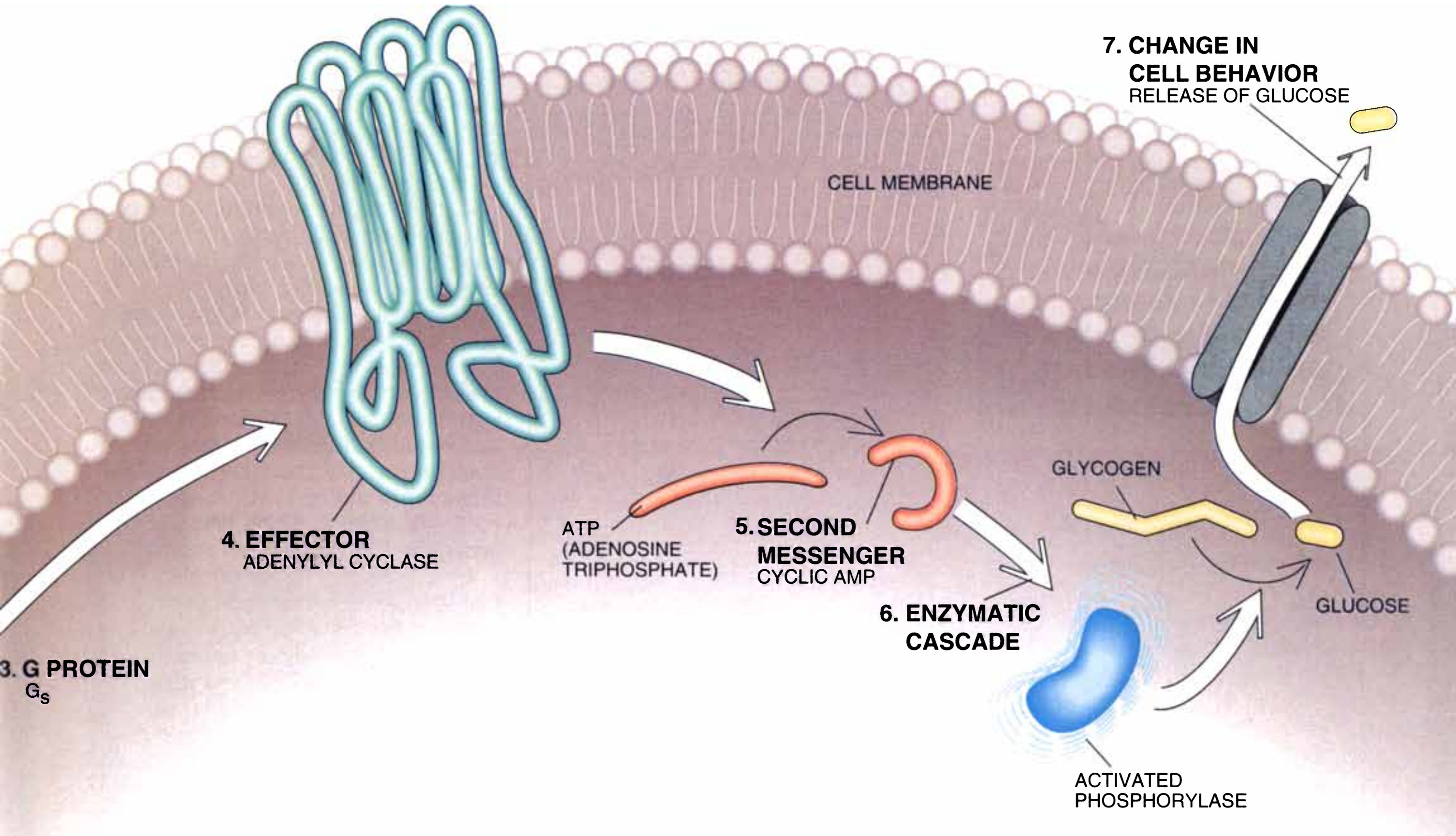
Over the years researchers have also learned that signal-transducing G proteins are part of a large superfamily of proteins that are regulated by guanine nucleotides. It has become apparent as well that aberrations in the functioning of signal-transducing G proteins and their relatives can contribute to diseases—among them, cholera, whooping cough and cancer. One day, treatments targeted to regulating specific G proteins may well become commonplace.

Researchers who study G proteins owe a major debt to Theodore W. Rall and the late Earl W. Sutherland, Jr. Those investigators carried out ground-breaking studies on cellular signaling in the late 1950s at Western Reserve University (now Case Western Reserve).

Scientists now know that many different receptors convey instructions of hormones or other extracellular "first messengers" by stimulating one or another G protein. Attached to the inner surface of the cell membrane (the plasma membrane), such proteins in turn act on membrane-bound intermediaries called effectors. Often the effector is an enzyme that converts an inactive precursor molecule into an active "second messenger," which can diffuse through the cytoplasm and thus carry the signal beyond the membrane. The second messenger triggers a cascade of molecular reactions leading to a change in

MAURINE E. LINDER and ALFRED G. GILMAN are in the department of pharmacology at the University of Texas Southwestern Medical Center at Dallas. Linder, an instructor, earned her Ph.D. in molecular and cell biology from the University of Texas at Dallas in 1987. Gilman was at the University of Virginia for a decade before becoming professor and chairman of the department of pharmacology at the Southwestern Medical Center in 1981. He was appointed to the Raymond and Ellen Willie Distinguished Chair in Molecular Neuropharmacology in 1987 and won the Albert Lasker Basic Medical Research Award in 1989.





G PROTEIN (*red complex*) in a liver cell relays signals (*arrows*) from the hormone epinephrine (*purple sphere at far left*) to a series of molecules that ultimately cause the cell to release glucose (*small yellow cylinder at far right*). The pathway traveled by such signals is typical of many that are regulated by G proteins; standard features are numbered and bold. In this case, epinephrine, a first messenger, opens the pathway by binding to its receptor (*lavender*), which stimulates the protein G_S to activate adenylyl cyclase (*green*). This enzyme, called an effector, converts ATP (*orange cylinder*) into the second messenger cyclic AMP (*orange curl*). Cyclic AMP then triggers a cascade of enzymatic reactions that stimulates the enzyme phosphorylase (*blue*) to convert glycogen (*yellow strip*) into glucose, which the cell finally extrudes.

the cell's behavior. For instance, the cell might secrete a hormone or release glucose.

Back in the 1950s, though, the workings of intracellular signaling pathways were almost completely obscure. Dogma held that the effects of hormones could be observed only in intact cells. Yet Rall and Sutherland, who were studying how the hormone epinephrine (adrenaline) stimulates liver cells to release glucose, thought otherwise. They insisted that to understand a system, one must study its component parts.

Their analyses of cell fragments revealed that epinephrine caused an enzyme in the plasma membrane to convert the nucleotide adenosine triphosphate (ATP) into a previously unknown substance: cyclic AMP (cyclic adenosine-3′,5′-monophosphate). The enzyme in question was another novel molecule they called adenyl cyclase (now adenylyl cyclase).

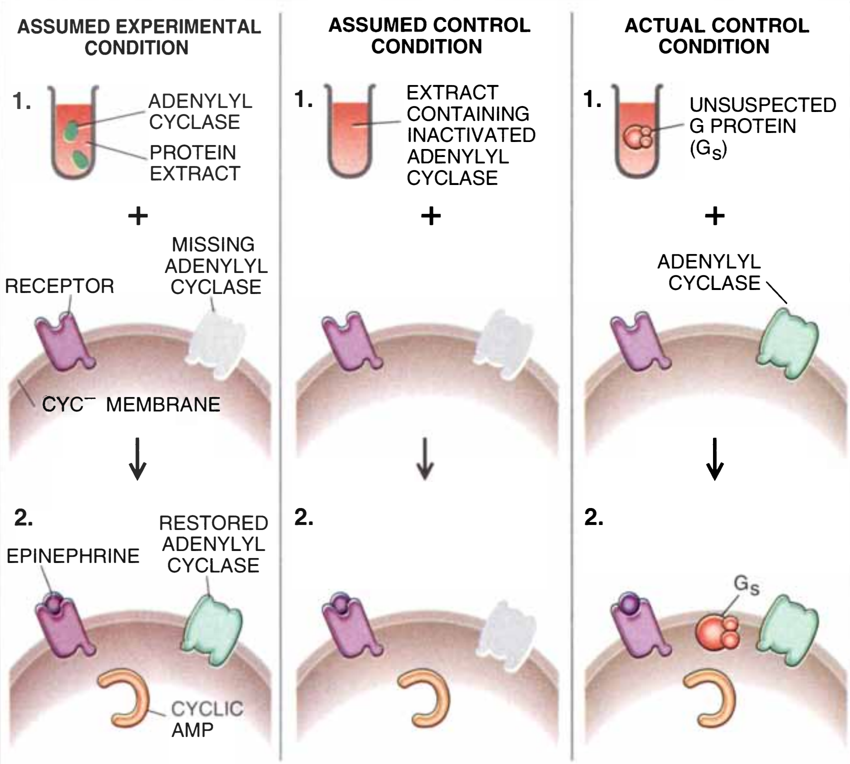
In today's terminology, the team revealed that a first messenger (epinephrine) works in part by stimulating an effector (adenylyl cyclase) to produce a second messenger (cyclic AMP). The steps between signaling by epinephrine outside the cell and activation of adenylyl cyclase in the membrane remained elusive, however.

Sutherland suspected that adenylyl cyclase and the receptor for epinephrine might be the same molecule, which would have meant the receptor served double duty as an effector, without need for a go-between. In some signaling pathways, that is indeed the case, but experiments done in the early 1970s established that the epinephrine receptor and adenylyl cyclase were in fact separate entities.

How, then, did the receptor communicate with adenylyl cyclase? Two important clues suggested guanine nucleotides were involved—and opened a line of inquiry that eventually led to discovery of the significance of G proteins in signal transduction. First, Martin Rodbell, Lutz Birnbaumer and their colleagues, then at the National Institutes of Health, demonstrated that the presence of the hormone glucagon (which also exerts effects through adenylyl cyclase) and its receptor was not sufficient to activate adenylyl cyclase. The nucleotide guanosine triphosphate (GTP) had to be present, too.

Then Danny Cassel and Zvi Selinger of the Hebrew University discovered that when epinephrine was added to fragments of plasma membranes, it not only stimulated adenylyl cyclase, it also caused GTP to be converted to guanosine diphosphate (GDP). The conversion occurred because an enzyme called a GTPase hydrolyzed GTP (split it by bringing it in contact with a water molecule), thereby freeing one of its three phosphate groups. The identity of the GTPase and the specific function of GTP would take some time to discern. But, at the least, both results indicated that guanine nucleotides participated in transmitting the epinephrine signal.

In the mid-1970s Gilman and Elliott M. Ross, who was a member of Gilman's laboratory, were among the investigators puzzling over the connections between epinephrine receptors,



Experiment Leading to Discovery of the Protein G_S

In the experimental condition, Elliott M. Ross in Gilman's laboratory added an extract of membrane proteins to so-called cyc$^-$ membranes (*left, 1*), which were thought to lack adenylyl cyclase, the enzyme that synthesizes cyclic AMP. Epinephrine then stimulated the membranes to produce cyclic AMP (*2*), which seemed to indicate that adenylyl cyclase (*green*) had been inserted into the deficient membranes. In the control experiment the adenylyl cyclase in the extract was destroyed (*center, 1*). Yet, even without it, epinephrine caused the cyc$^-$ membranes to make cyclic AMP (*2*). This puzzling finding led to the discovery that the cyc$^-$ membranes included adenylyl cyclase all along (*right, 1*) but lacked a third component necessary to activate it—a G protein (*red spheres*) that persisted in the extract even after adenylyl cyclase was eliminated. In the control experiment, restoration of the G protein to the membranes enabled the hidden adenylyl cyclase to produce cyclic AMP (*2*).

GTP and adenylyl cyclase. Like Rall and Sutherland, they felt the best way to understand a biological system was to take it apart and put it back together. They would have liked to isolate the various proteins associated with the plasma membrane and then systematically to add the proteins to, or remove them from, an artificial membrane. In that way, they could have teased out the minimum set of molecules required to activate adenylyl cyclase and could have identified the role of GTP in the sequence of events.

In those years, however, science lacked technology for readily purifying membrane components without damaging them. Gilman and Ross therefore set about testing various techniques for fitting proteins into membranes without first purifying the inserted material. As luck would have it, they found the solution to the signaling problem in the course of that effort.

Ross had devised an experiment to see whether he could restore adenylyl cyclase to a line of mutant cells that retained epinephrine receptors but apparently were unable to make adenylyl cyclase [*see box above*]. The late Gordon M. Tomkins, Henry R. Bourne and Philip Coffino of the University of California at San Francisco had discovered such cyclase-deficient, or cyc$^-$, cells in 1975.

Ross hoped to insert adenylyl cyclase into the cyc$^-$ membranes by mixing them with an extract of proteins derived from membranes that contained adenylyl cyclase. If he succeeded, the second messenger cyclic AMP would be produced when he exposed the treated cyc$^-$ membranes to epinephrine (as well as to ATP and GTP—elements that had been proved critical to production of the second messenger). Sure enough, when he ran the test, cyclic AMP appeared.

Naturally, Ross was quite pleased with his technical achievement. But the results of his control experiment proved even more satisfying.

In the control situation, Ross had unequivocally inactivated the extract's adenylyl cyclase by gently heating the extract or by adding chemicals to it. He assumed the protein mixture would then have no effect on the activity of cyc$^-$ membranes. In other words, after exposure to the extract, the membranes would continue to lack adenylyl cyclase and would thus display no response to epinephrine, ATP and GTP. Instead the membranes unexpectedly synthesized cyclic AMP, just as if they had taken up active adenylyl cyclase from the extract.

The explanation for the puzzling finding soon became obvious. Cyc$^-$ cells actually contain adenylyl cyclase. The enzyme had simply gone undiscovered because it was inactive. And it was inactive because another cellular component—one adenylyl cyclase required for stimulation—was missing from the cells. Fortunately, the component remained in the extract after adenylyl cyclase was eliminated, and it went on to "waken" the formerly silent enzyme in the cyc$^-$ cells.

Ross and Gilman quickly showed the activating substance to be a protein that was itself stimulated by GTP. Rather than referring to the substance by the cumbersome name "guanine nucleotide–binding protein," they simply called it a G protein. It did not take much imagination to deduce that epinephrine led to the production of cyclic AMP by passing information through its receptor to a G protein, which, when bound by GTP, could stimulate adenylyl cyclase to convert ATP into cyclic AMP.

By 1980 Paul C. Sternweis and John K. Northup, who were then in Gilman's laboratory, managed to purify the G protein, now known as G_S, that stimulates adenylyl cyclase. Later, when the epinephrine receptor and adenylyl cyclase had been isolated as well, Gilman and Ross prepared artificial cell membranes consisting of all three proteins embedded in phospholipids, the main structural components of plasma membranes. When epinephrine, ATP and GTP were added, cyclic AMP was produced.

Thus, in the mid-1980s, this experiment proved conclusively that all the essential elements of the pathway leading from epinephrine to production of the second messenger cyclic AMP had been identified. By then, too, other workers had learned that cyclic AMP activates protein kinases, enzymes that phosphorylate, or add phosphate groups to, other proteins. And they had filled in the rest of the epinephrine pathway in liver cells, demonstrating that cyclic AMP sets off an enzymatic cascade that ultimately stimulates phosphorylase, the enzyme that breaks glycogen (animal starch) into glucose. In other types of cells, cyclic AMP, acting through the same kinases, can lead to different effects, such as the synthesis and release of steroid hormones by the adrenal gland and gonads.

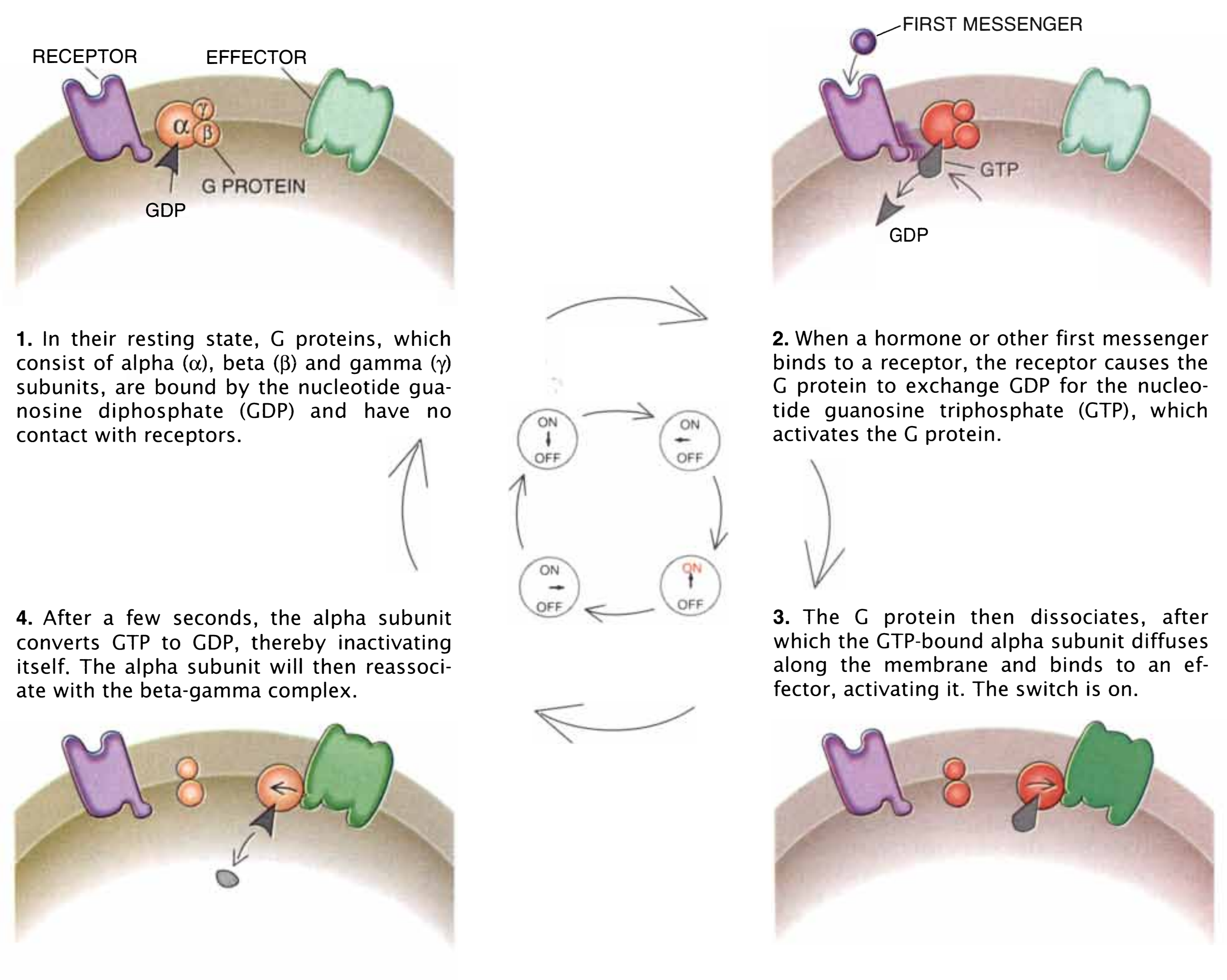
As often happens in science, at about the same time as Gilman and his coworkers were closing in on G_s as the stimulator of adenylyl cyclase, other investigators were about to make a similar discovery in the course of studying a seemingly different problem: how rod cells in the retina respond to light. Experiments by Mark W. Bitensky, then at Yale University, and independently by Lubert Stryer and his colleagues at Stanford University revealed a second G protein, now known as transducin, or G_t.

This protein acts as an intermediary between the receptor for light (rhodopsin) and an effector enzyme (phosphodiesterase) that regulates the levels of a second messenger (guanosine-3′,5′-monophosphate, or cyclic GMP). When rhodopsin is struck by photons, it activates G_t, which instructs phosphodiesterase to convert cyclic GMP to GMP. Because cyclic GMP keeps sodium channels open, conversion to GMP closes the channels, preventing sodium ions (Na^{2+}) from entering the cell. The negative charge inside the cell then increases, causing the cell to become hyperpolarized and so to generate an electrical signal that is relayed to the brain.

The striking parallels between hormonal stimulation of adenylyl cyclase and light activation of phosphodiesterase soon led other investigators to join the search for new G proteins. The field did not truly blossom, however, until 1984, when the genes for the first members of the family were cloned. (Melvin Simon and his associates at the California Institute of Technology have been especially prolific at isolating the genes for novel G proteins; they now number new members of the family instead of naming them.)

Since the early 1980s researchers have found that more than 100 receptors (or perhaps thousands, if odorant receptors are included) convey messages through G proteins, of which at least 20 distinct forms have been isolated. Scientists have identified several different G protein–dependent effectors as well. Along with adenylyl cyclase and

How G Proteins Switch Effectors On and Off



1. In their resting state, G proteins, which consist of alpha (α), beta (β) and gamma (γ) subunits, are bound by the nucleotide guanosine diphosphate (GDP) and have no contact with receptors.

2. When a hormone or other first messenger binds to a receptor, the receptor causes the G protein to exchange GDP for the nucleotide guanosine triphosphate (GTP), which activates the G protein.

3. The G protein then dissociates, after which the GTP-bound alpha subunit diffuses along the membrane and binds to an effector, activating it. The switch is on.

4. After a few seconds, the alpha subunit converts GTP to GDP, thereby inactivating itself. The alpha subunit will then reassociate with the beta-gamma complex.

cyclic GMP phosphodiesterase, the effectors include other enzymes and membrane channels that regulate the flow of inorganic ions into or out of the cell. Of the enzymes, one called phospholipase C is particularly interesting. It breaks down a phospholipid in the plasma membrane into two second messengers, one of which releases internal stores of yet another second messenger, calcium.

As the realization dawned that G proteins were important in virtually all cells, the need to understand how they regulate the flow of information between receptors and effectors became more urgent. Within a few years after discovery of G_S and transducin, the basic structure and mode of operation of G proteins were elucidated.

The G proteins that participate in transmembrane signaling are bound to the inner surface of the plasma membrane and consist of three protein chains, or subunits, called, from largest to smallest, alpha (α), beta (β) and gamma (γ). The alpha subunit differs in each of the G proteins isolated to date. The attached pair of beta and gamma subunits is not necessarily unique; different alpha chains may be linked to an identical beta-gamma pair or to differing pairs. So far workers have noted five distinct beta structures and possibly more than 10 gammas, suggesting that more than 1,000 combinations of alpha, beta and gamma could arise.

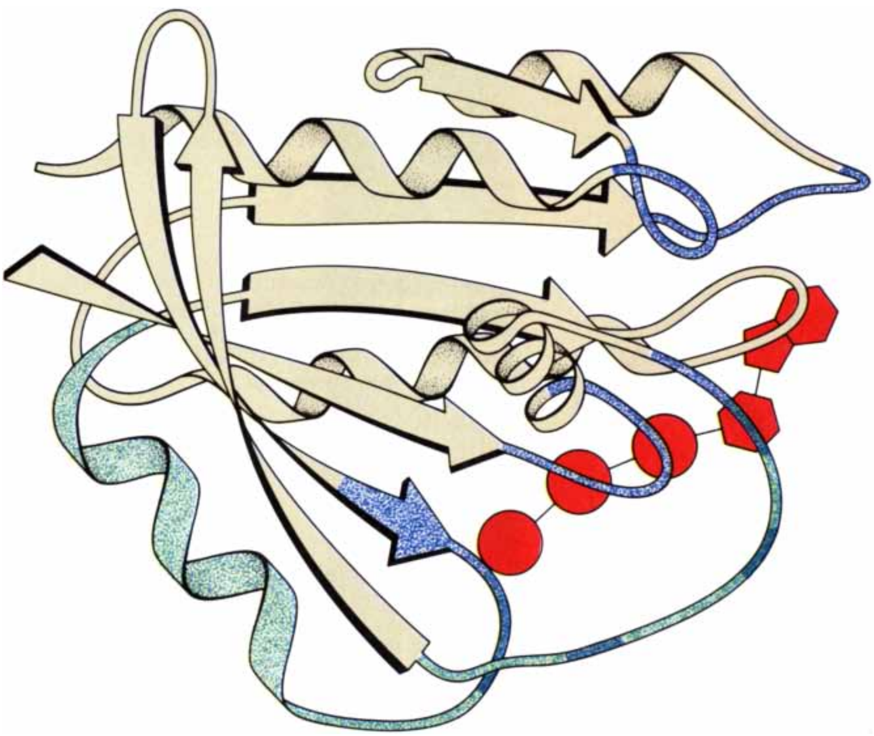
G proteins carry out their functions in a curious way. In the resting state, the alpha, beta and gamma chains form a complex, and GDP is bound to the alpha subunit. After a hormone or other first messenger docks with a receptor, the receptor, through a change in conformation, binds to the G protein. This interaction spurs the alpha subunit to release GDP. Then GTP, which is more abundant in cells, fills in the open binding site, altering the shape of the alpha subunit and activating it.

Once activated, the GTP-bound alpha subunit dissociates from the beta and gamma subunits and diffuses along the inner surface of the plasma membrane until it links up with an effector, such as adenylyl cyclase. After usually a few seconds, the alpha subunit hydrolyzes GTP to form GDP, thereby shutting itself off—a behavior that explains why Cassel and Selinger had earlier detected GTPase activity in epinephrine-stimulated plasma membranes. After the alpha subunit becomes inactive, it dissociates from the effector and then reassociates with free beta and gamma subunits.

Hence, G proteins serve as switches and timers, determining when and for how long signaling pathways are turned on and off. The switch flips on when the GTP-bound alpha chain binds to an effector, and the switch goes off when GTP is hydrolyzed to GDP. The rate of hydrolysis determines the time elapsing between the on and the off state.

G proteins also amplify signals. In the extremely efficient visual system, for example, one molecule of rhodopsin activates almost simultaneously more than 500 molecules of transducin. In the case of G_S, one alpha subunit bound to a molecule of adenylyl cyclase can stimulate synthesis of many molecules of cyclic AMP before the timer runs out and GTP is converted back to GDP.

Interestingly, self-hydrolysis of GTP also regulates the activity of many proteins that do not participate in transduction of external signals. Together these proteins and the signal transducers form a superfamily of GTPases. Some members of the superfamily participate in the synthesis of proteins on ribosomes. Other members, notably the products of *ras* genes, help to control the rate of cell division.

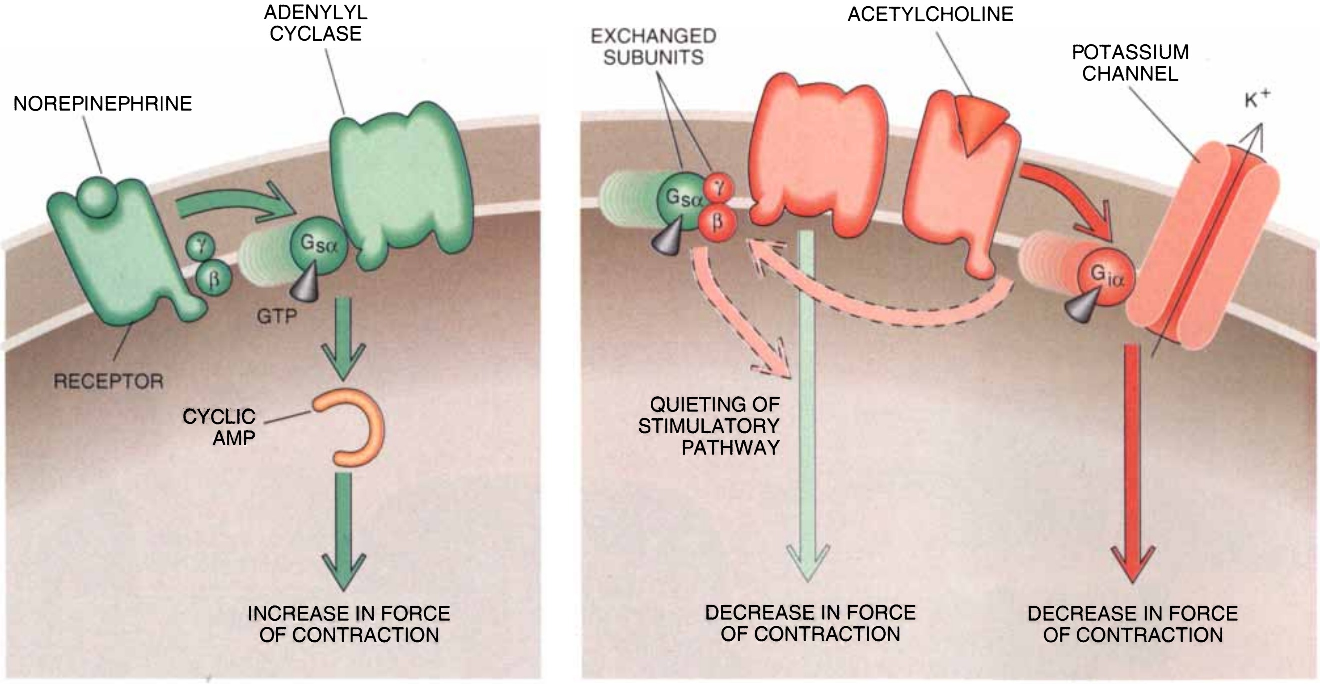


P21*ras* PROTEIN (*ribbonlike structure*) is bound to the nucleotide guanosine triphosphate (GTP) (*red*), which activates it. The GTP-binding region of ras is thought to closely resemble that of signal-transducing G proteins, whose three-dimensional structure has not yet been deciphered. Curls, arrows and loops indicate, respectively, that the backbone is helical, extended or of a less stereotyped nature. The blue areas convert GTP to GDP, causing the green regions to change shape and inactivate the molecule. (Blue-green areas interact with nucleotides and change shape.) Alfred Wittinghofer of the Max Planck Institute for Medical Research in Heidelberg and Sung-Hou Kim of the University of California at Berkeley determined the structure.

Clearly, the alpha subunit of signal-transducing G proteins plays a critical part in conveying messages to effectors, but whether beta-gamma pairs also regulate effectors is a subject of spirited debate. Some think not, and one partisan even has a license plate that reads "a not b"!

We are among those who have gathered evidence in support of the view that beta-gamma complexes, which remain bound to each other and act as a single unit, can be important in signaling. Work from our laboratory suggests that beta-gamma affects some signaling pathways more than others and is sometimes stimulatory and sometimes inhibitory. For example, Wei-Jen Tang has recently shown that beta-gamma can cooperate with G_S alpha to activate particular forms of adenylyl cyclase. Yet it can inhibit a type of adenylyl cyclase found in some neurons, probably by binding directly to that enzyme. In contrast, beta-gamma seems to have no direct effect on certain other kinds of adenylyl cyclase.

In the last case, the complex might sometimes exert an indirect influence. Such activity may explain the continuing mystery of how G_i, which interferes with the ability of adenylyl cyclase to



STIMULATORY PATHWAY in heart muscle cells (*green arrows in left panel*) might be inhibited in part by the process of subunit exchange (*left part of right panel*). Heart cells are stimulated to increase the force of contraction when $G_{s\alpha}$, the alpha subunit of the protein G_s, activates adenylyl cyclase (*left panel*). The cells decrease the force (*solid red arrows in right panel*) when acetylcholine spurs the alpha subunit of G_i (*large red sphere*) to open channels that enable potassium ions (K^+) to leave the cell. Subunit exchange would cause a further decrease (*broken red arrows*) if the beta and gamma subunits of G_i (*double red spheres*) combined with the alpha subunit of G_s, thus blocking alpha from acting on adenylyl cyclase.

generate cyclic AMP, achieves its effect. By analogy with G_s and transducin, Gilman and his colleagues anticipated that G_i, which was discovered in 1982, would inhibit adenylyl cyclase by releasing its alpha subunit. Yet when Northup and Toshiaki Katada in Gilman's group measured the ability of purified G_i alpha and beta-gamma proteins to inhibit adenylyl cyclase in plasma membranes, they found that the alpha protein was a much weaker inhibitor than was the beta-gamma complex.

Strangely, inhibition by beta-gamma occurred only when G_s was present. This finding led Gilman to suggest that a process called subunit exchange might be at work. According to this hypothesis, beta and gamma subunits from activated G_i molecules may combine with alpha subunits that have been released from G_s—something that could well occur if the beta-gamma complexes of G_s and G_i were identical. By engaging the stimulatory alpha subunit, the inhibitory beta-gamma complex would prevent it from interacting with adenylyl cyclase.

Heart muscle cells offer an example of how such subunit exchange could influence cellular behavior [*see illustration above*]. The neurotransmitter norepinephrine increases the vigor and rate of contraction through the G_s alpha subunit, adenylyl cyclase and cyclic AMP. When the neurotransmitter acetylcholine also acts on cells, it blunts the contraction, apparently by two mechanisms. Birnbaumer, Arthur M. Brown and their colleagues at Baylor College of Medicine have shown that acetylcholine stimulates the alpha subunit of G_i to open membrane channels that enable potassium ions (K^+) to flow out of the cells, ultimately impeding contraction. Acetylcholine also inhibits adenylyl cyclase and thus clamps the stimulatory pathway set in motion by G_s.

Inhibition of adenylyl cyclase in heart cells could well be the work of the G_i alpha subunit, but some evidence suggests the beta-gamma complex is responsible. It could achieve this effect by binding to the alpha subunit of G_s, thereby blocking its access to adenylyl cyclase. We are not wedded to the notion that subunit exchange occurs in cells, but the hypothesis continues to live because it is consistent with much experimental data and because no one has yet demonstrated that the alpha subunit of G_i can inhibit adenylyl cyclase directly.

A fuller understanding of the details of how signals are conveyed from receptors through G proteins to effectors awaits elucidation of the three-dimensional structure of such molecules by x-ray crystallography. A knowledge of shape can reveal which parts of a molecule might mesh well with other molecules.

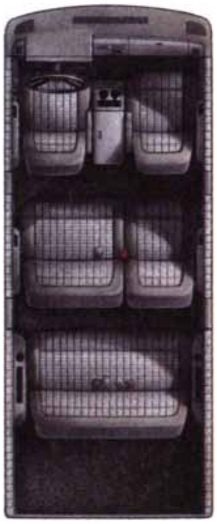
In the meantime, investigators are learning something about structure by watching for clues in the amino acid sequences of the proteins and by comparing the sequences with those of related proteins whose three-dimensional shapes are known. For instance, analyses of receptors that interact with G proteins indicate that most of them include seven separate regions that are rich in hydrophobic, or water-hating, amino acids. Those segments almost certainly span the hydrophobic plasma membrane (thus escaping the watery cytoplasm) and describe a pocket that is accessible to first messengers from the extracellular face of the membrane. In addition, the cytoplasmic segments that connect the hydrophobic regions include one or more loops that presumably bind to specific G proteins.

No one yet knows which sites on signal-transducing G proteins make contact with receptors nor how the subunits of a G protein contact one another. But investigators have a good idea of where the alpha subunit binds GTP and GDP and of its conformation in the active and inactive state [*see illustration on opposite page*]. This information is

INTRODUCING A TRUCK COMPANY'S NEW

As you'd expect, it extends beyond a simple station wagon. In fact, our all-new Suburban carries up to nine passengers facing forward, and is so powerful it can tow some of your biggest trailers and boats. And every commanding inch and ounce has been redesigned for 1992. To give you even more of the extraordinary comforts and capabilities that caused one automotive writer* to enthuse: "A better blend of utility, comfort, safety and style in a motor vehicle is difficult to imagine."

For example, Suburban now tows up to five tons,† a full 500 pounds more than before. New aerodynamics virtually eliminate wind noise. There's less bulk outside, yet more legroom, width, cargo space and front-seat shoulder room inside. Seats, instrument panel and interior trim are all-new; third-seat passengers now enjoy their own recessed foot-well and separate available climate control system. And in an age that increasingly sees six cylinders as sufficient, GMC Suburban gives you the assurance of standard 5.7 liter or available 7.4 liter V8 power. As well as the technical sophistication of an available electronically controlled automatic transmission on 3/4-ton models. And the advanced safety of 4-wheel anti-lock braking on all models.



JUST ONE OF SUBURBAN'S OPTIONAL SEATING CHOICES

* *Detroit News* 12/11/91
†Properly equipped including passengers, cargo and trailer

Idea Of The Perfect Way To Travel.

But the full appreciation of Suburban's newness comes from behind the wheel. Reporting on the benefits of the new longer wheelbase and independent front suspension, *Trailer Life* wrote: "A very smooth, controlled boulevard ride. We found steering characteristics impossible to fault. Road feel was excellent... requiring little more than a two-finger effort on the steering wheel." Driving accolades all the more impressive when you consider the editors were towing a 27-foot travel trailer at the time.

"FUN TO DRIVE SOLO ON MOUNTAIN ROADS." –*Trailer Life*

Inspect our all-new Suburban at your GMC Truck dealer's. It's the only vehicle that gives you the strengths of a truck, the versatility of a van and – for 1992 – the smooth, quiet ease and comfort of a luxury sedan. Combined with the strength of 90 years of GMC Truck experience.

For more Suburban information, call 1-800-879-4621.

GMC TRUCK®

THE STRENGTH OF EXPERIENCE

based on the crystal structure of a ras protein, which was deciphered independently by Alfred Wittinghofer of the Max Planck Institute for Medical Research in Heidelberg and Sung-Hou Kim of the University of California at Berkeley. The guanine nucleotide–binding part of ras closely resembles that of signal-transducing G proteins.

In contrast to receptors, G proteins do not seem to include strongly hydrophobic regions that would explain how the proteins attach to the plasma membrane. Nevertheless, Susanne M. Mumby and Patrick J. Casey in our laboratory and John A. Glomset and his associates at the University of Washington have independently discovered a possible dab of "glue": one end of the gamma protein is bound to a lipid molecule called an isoprenoid. We think the lipid might anchor G proteins to the cell membrane, for the simple reason that it, like the phospholipid membrane, is hydrophobic. Or the isoprenoid might help G proteins attach to other proteins in the membrane. The alpha subunits of some signal-transducing G proteins are probably also bound to the membrane with assistance from a second lipid known as myristic acid.

Little is known about the conformations of the effectors targeted by G proteins, but our laboratory has been making some headway with adenylyl cyclase. On the basis of the amino acid sequence of several such molecules, we suggest that adenylyl cyclase assumes a rather complex shape: 12 transmembrane segments collectively form a tunnel resembling the channels through which ions travel in and out of cells. Two hydrophilic (water-loving) domains that reside in the cytoplasm are necessary for the synthesis of cyclic AMP. At the moment, we can only speculate on why the structure is so intricate. Perhaps adenylyl cyclase is more than an enzyme and serves as a transporter of some kind as well.

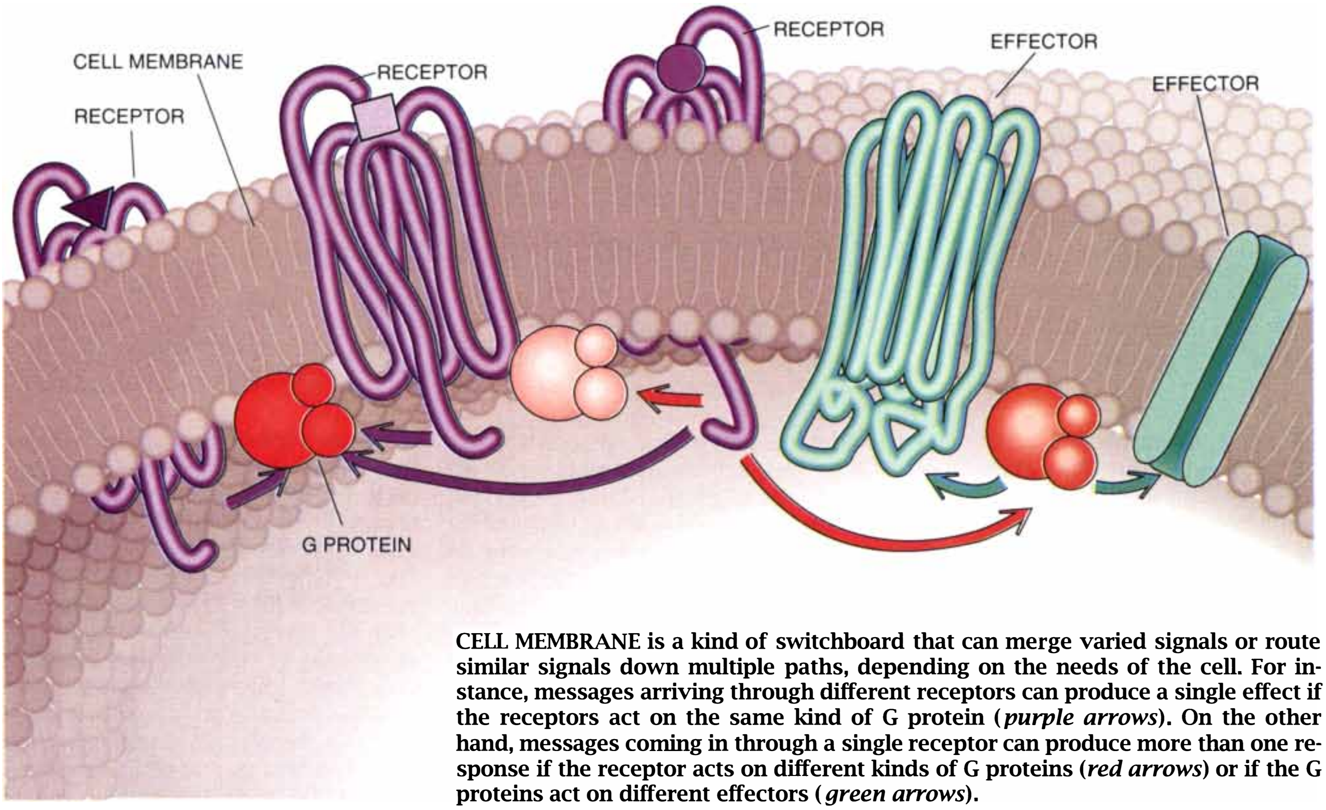
Up to this point, we have focused mainly on individual pathways of signal transduction. Yet, as the complexity of activity in heart muscle suggests, the signaling pathways mediated by G proteins moderate one another's effects.

When pathways converge, such as when different receptors act on the same type of G protein or when different G proteins act on one kind of effector, a cell may produce a single (albeit graded) response to different first messengers. On the other hand, when pathways diverge, such as when one kind of receptor acts on different types of G proteins or when one G protein acts on more than one effector, a cell may exhibit several concurrent responses to a single external message. The ability of receptors, G proteins and effectors to interact with more than one species of molecule inside the cell also means a cell can make different choices from time to time—now sending a signal down one pathway and now directing it along a somewhat different route.

Clearly, the cell membrane is a switchboard of considerable complexity, taking in a diversity of signals, assessing their relative strengths and relaying the summed signals to second messengers that will assure the cell reacts appropriately to a changing environment. The specific responses a cell makes depend on the precise combination of signals affecting it from without, as well as on both the mix of its receptors, G proteins and effectors and the repertoire of other specialized proteins it

A Sampling of Physiological Effects Mediated by G Proteins

STIMULUS	AFFECTED CELL TYPE	G PROTEIN	EFFECTOR	EFFECT
Epinephrine, glucagon	Liver cells	G_s	Adenylyl cyclase	Breakdown of glycogen
Epinephrine, glucagon	Fat cells	G_s	Adenylyl cyclase	Breakdown of fat
Luteinizing hormone	Ovarian follicles	G_s	Adenylyl cyclase	Increased synthesis of estrogen and progesterone
Antidiuretic hormone	Kidney cells	G_s	Adenylyl cyclase	Conservation of water by kidney
Acetylcholine	Heart muscle cells	G_i	Potassium channel	Slowed heart rate and decreased pumping force
Enkephalins, endorphins, opioids	Brain neurons	G_i/G_o	Calcium and potassium channels, adenylyl cyclase	Changed electrical activity of neurons
Angiotensin	Smooth muscle cells in blood vessels	G_q	Phospholipase C	Muscle contraction; elevation of blood pressure
Odorants	Neuroepithelial cells in nose	G_{olf}	Adenylyl cyclase	Detection of odorants
Light	Rod and cone cells in retina	G_t	Cyclic GMP phosphodiesterase	Detection of visual signals
Pheromone	Baker's yeast	GPA1	Unknown	Mating of cells



CELL MEMBRANE is a kind of switchboard that can merge varied signals or route similar signals down multiple paths, depending on the needs of the cell. For instance, messages arriving through different receptors can produce a single effect if the receptors act on the same kind of G protein (*purple arrows*). On the other hand, messages coming in through a single receptor can produce more than one response if the receptor acts on different kinds of G proteins (*red arrows*) or if the G proteins act on different effectors (*green arrows*).

produces. Of particular importance are the unique proteins in a cell. Hence, a liver cell, which contains phosphorylase and stores much glycogen, will respond to epinephrine through G_s by liberating glucose. Heart cells, which produce specialized channels and contractile proteins, will respond through G_s by contracting more forcefully and frequently.

There is obvious satisfaction in unraveling the mechanisms of transmembrane signaling, but these explorations have a practical side, too. Studies of signaling systems that rely on G proteins have already improved knowledge of several diseases, which is a first step to their rational control.

The bacterium responsible for cholera secretes a toxin that enters intestinal cells, where it prevents the alpha subunit of G_s from converting GTP to GDP. The toxin, then, prevents G_s from shutting itself off. Consequently, the cells accumulate excess cyclic AMP, which causes them to secrete large amounts of electrolytes and water into the lumen of the gut. The severe diarrhea causes potentially lethal dehydration.

The bacterium that causes whooping cough elaborates a related toxin that prevents receptors from activating G_i. Without an inhibitor, stimulatory pathways again remain functional for too long. This toxin affects many types of cells, apparently contributing to the immunodeficiency that can accompany the characteristic cough.

Finally, mutations of G proteins, including G_s and G_i, seem to participate in the development of some cancers. In the cells of pituitary tumors, for example, Bourne and his colleagues in San Francisco have found mutations in the gene that specifies the alpha subunit of G_s. The mutations cause the subunit to act on effectors for minutes rather than seconds, which can promote excessive replication of pituitary cells. Similarly, in a more widespread example, defects that impair the ability of ras proteins to convert GTP to GDP result in uncontrolled proliferation of cells.

Because G proteins control the highly specialized functions of most cells, investigators are also beginning to examine whether disturbances in G protein function can contribute to conditions as diverse as heart failure, diabetes and psychological depression. As knowledge of the structure and activity of G proteins expands, it should be possible to design drugs to interact selectively with specific types, thereby correcting defective functioning in disordered cells while leaving healthy cells untouched.

Ultimately, scientists will compile a complete wiring diagram of the plasma membrane in all the myriad cell types of the human organism. For each, they will know how dozens of types of receptors, G proteins and effectors are connected. And they will be able to predict how the cells will operate in response to any combination of signals. As one wag has said, for those who would hope to develop drug therapies, such discoveries would be like giving a thief a wiring diagram to the alarm system at a bank.

FURTHER READING

G PROTEINS: TRANSDUCERS OF RECEPTOR-GENERATED SIGNALS. Alfred G. Gilman in *Annual Review of Biochemistry,* Vol. 56, pages 615–649; 1987.

THE GTPASE SUPERFAMILY: A CONSERVED SWITCH FOR DIVERSE CELL FUNCTIONS. Henry R. Bourne, David A. Sanders and Frank McCormick in *Nature,* Vol. 348, No. 6297, pages 125–132; November 8, 1990.

THE GTPASE SUPERFAMILY: CONSERVED STRUCTURE AND MOLECULAR MECHANISM. Henry R. Bourne, David A. Sanders and Frank McCormick in *Nature,* Vol. 349, No. 6305, pages 117–127, January 10, 1991.

DIVERSITY OF G PROTEINS IN SIGNAL TRANSDUCTION. Melvin I. Simon, Michael P. Strathmann and Narasimhan Gautam in *Science,* Vol. 252, pages 802–808; May 10, 1991.

STRUCTURE AND FUNCTION OF SIGNAL-TRANSDUCING GTP-BINDING PROTEINS. Y. Kaziro, H. Itoh, T. Kozasa, M. Nakafuku and T. Satoh in *Annual Review of Biochemistry,* Vol. 60, pages 349–400; 1991.

Genetic Algorithms

Computer programs that "evolve" in ways that resemble natural selection can solve complex problems even their creators do not fully understand

by John H. Holland

Living organisms are consummate problem solvers. They exhibit a versatility that puts the best computer programs to shame. This observation is especially galling for computer scientists, who may spend months or years of intellectual effort on an algorithm, whereas organisms come by their abilities through the apparently undirected mechanism of evolution and natural selection.

Pragmatic researchers see evolution's remarkable power as something to be emulated rather than envied. Natural selection eliminates one of the greatest hurdles in software design: specifying in advance all the features of a problem and the actions a program should take to deal with them. By harnessing the mechanisms of evolution, researchers may be able to "breed" programs that solve problems even when no person can fully understand their structure. Indeed, these so-called genetic algorithms have already demonstrated the ability to make breakthroughs in the design of such complex systems as jet engines.

Genetic algorithms make it possible to explore a far greater range of potential solutions to a problem than do conventional programs. Furthermore, as researchers probe the natural selection of programs under controlled and well-understood conditions, the practical results they achieve may yield some insight into the details of how life and intelligence evolved in the natural world.

Most organisms evolve by means of two primary processes: natural selection and sexual reproduction. The first determines which members of a population survive to reproduce, and the second ensures mixing and recombination among the genes of their offspring. When sperm and ova fuse, matching chromosomes line up with one another and then cross over partway along their length, thus swapping genetic material. This mixing allows creatures to evolve much more rapidly than they would if each offspring simply contained a copy of the genes of a single parent, modified occasionally by mutation. (Although unicellular organisms do not engage in mating as humans like to think of it, they do exchange genetic material, and their evolution can be described in analogous terms.)

Selection is simple: if an organism fails some test of fitness, such as recognizing a predator and fleeing, it dies. Similarly, computer scientists have little trouble weeding out poorly performing algorithms. If a program is supposed to sort numbers in ascending order, for example, one need merely check whether each entry of the program's output is larger than the preceding one.

People have employed a combination of crossbreeding and selection for millennia to breed better crops, racehorses or ornamental roses. It is not as easy, however, to translate these procedures for use on computer programs. The chief problem is the construction of a "genetic code" that can represent the structure of different programs, just as DNA represents the structure of a person or a mouse. Mating or mutating the text of a FORTRAN program, for example, would in most cases not produce a better or worse FORTRAN program but rather no program at all.

The first attempts to mesh computer science and evolution, in the late 1950s and early 1960s, fared poorly because they followed the emphasis in biological texts of the time and relied on mutation rather than mating to generate new gene combinations. Then, in the early 1960s, Hans J. Bremermann of the University of California at Berkeley added a kind of mating: the characteristics of offspring were determined by summing up corresponding genes in the two parents. This mating procedure was limited, however, because it could apply only to characteristics that could be added together in a meaningful way.

During this time, I had been investigating mathematical analyses of adaptation and had become convinced that the recombination of groups of genes by means of mating was a critical part of evolution. By the mid-1960s I had developed a programming technique, the genetic algorithm, that is well suited to evolution by both mating and mutation. During the next decade, I worked to extend the scope of genetic algorithms by creating a genetic code that could represent the structure of any computer program.

The result was the classifier system, consisting of a set of rules, each of which performs particular actions every time its conditions are satisfied by some piece of information. The conditions and actions are represented by strings of bits corresponding to the presence or absence of specific characteristics in the rules' input and output. For each characteristic that was present, the string would contain a 1 in the appropriate position, and for each that was absent, it would contain a 0. For example, a classifier rule that recognized dogs might be encoded as a string containing 1's for the bits corresponding to "hairy," "slobbers," "barks," "loyal" and "chases sticks" and 0's for the bits corresponding to "metallic," "speaks Urdu" and "possesses credit cards." More realistically, the programmer should choose the simplest, most primitive characteristics so that they can be combined—as

JOHN H. HOLLAND has been investigating the theory and practice of algorithmic evolution for nearly 40 years. He is a professor of psychology and of electrical engineering and computer science at the University of Michigan. Holland received a B.S. in physics from the Massachusetts Institute of Technology in 1950 and served on the Logical Planning Group for IBM's first programmed electronic computer (the 701) from 1950 until 1952. He received an M.A. in mathematics and a Ph.D. in communication sciences from the University of Michigan. Holland has been a member of the Steering Committee of the Santa Fe Institute since its inception in 1987 and is an external professor there.

in the game of 20 Questions—to classify a wide range of objects and situations.

Although they excel at recognition, these rules can also be made to trigger actions by tying bits in their output to the appropriate behavior [*see illustration on page 69*]. Any program that can be written in a standard programming language such as FORTRAN or LISP can be rewritten as a classifier system.

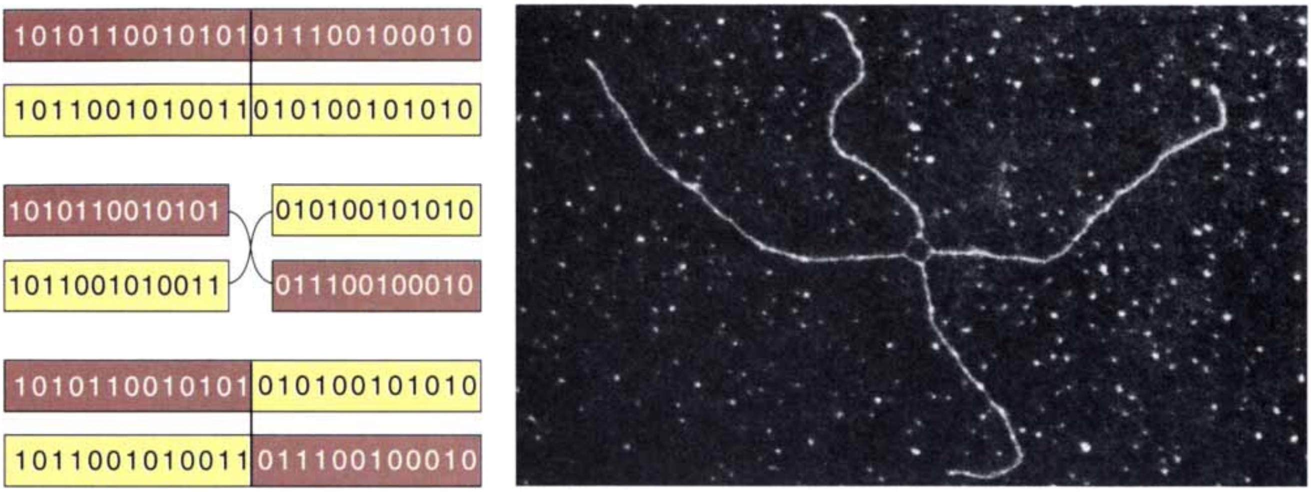
To evolve classifier rules that solve a particular problem, one simply starts with a population of random strings of 1's and 0's and rates each string according to the quality of its result. Depending on the problem, the measure of fitness could be business profitability, game payoff, error rate or any number of other criteria. High-quality strings mate; low-quality ones perish. As generations pass, strings associated with improved solutions will predominate. Furthermore, the mating process continually combines these strings in new ways, generating ever more sophisticated solutions. The kinds of problems that have yielded to the technique range from developing novel strategies in game theory to designing complex mechanical systems.

Recast in the language of genetic algorithms, the search for a good solution to a problem is a search for particular binary strings. The universe of all possible strings can be considered as an imaginary landscape; valleys mark the location of strings that encode poor solutions, and the landscape's highest point corresponds to the best possible string.

Regions in the solution space can also be defined by looking at strings that have 1's or 0's in specified places—a kind of binary equivalent of map coordinates. The set of all strings that start with a 1, for example, constitutes a region in the set of possibilities. So do all the strings that start with a 0 or that have a 1 in the fourth position, a 0 in the fifth and a 1 in the sixth and so on.

One conventional technique for exploring such a landscape is hill climbing: start at some random point, and if a slight modification improves the quality of your solution, continue in that direction; otherwise, go in the opposite direction. Complex problems, however, make landscapes with many high points. As the number of dimensions of the problem space increases, the countryside may contain tunnels, bridges and even more convoluted topological features. Finding the right hill or even determining which way is up becomes increasingly difficult. In addition, such

BEE ORCHID demonstrates the specificity with which natural genetic selection can match an organism to a particular niche. The flower, which resembles a female bumblebee, is fertilized by male bees that attempt to mate with it. Mechanisms similar to natural selection, the author says, can produce computer programs (so-called genetic algorithms) capable of solving such complex problems as the design of jet turbines or communications networks.



CROSSOVER is the fundamental mechanism of genetic rearrangement for both real organisms and genetic algorithms. Chromosomes line up and then swap the portions of their genetic code beyond the crossover point.

search spaces are usually enormous. If each move in a chess game, for example, has an average of 10 alternatives, and a typical game lasts for 30 moves on each side, then there are about 10^{60} strategies for playing chess (most of them bad).

Genetic algorithms cast a net over this landscape. The multitude of strings in an evolving population samples it in many regions simultaneously. Notably, the rate at which the genetic algorithm samples different regions corresponds directly to the regions' average "elevation"—that is, the probability of finding a good solution in that vicinity.

This remarkable ability of genetic algorithms to focus their attention on the most promising parts of a solution space is a direct outcome of their ability to combine strings containing partial solutions. First, each string in the population is evaluated to determine the performance of the strategy that it encodes. Second, the higher-ranking strings mate. Two strings line up, a point along the strings is selected at random and the portions to the left of that point are exchanged to produce two offspring: one containing the symbols of the first string up to the crossover point and those of the second beyond it, and the other containing the complementary cross [*see illustration above*]. Biological chromosomes cross over one another when two gametes meet to form a zygote, and so the process of crossover in genetic algorithms does in fact closely mimic its biological model. The offspring do not replace the parent strings; instead they replace low-fitness strings, which are discarded at each generation so that the total population remains the same size.

Third, mutations modify a small fraction of the strings: roughly one in every 10,000 symbols flips from 0 to 1, or vice versa. Mutation alone does not generally advance the search for a solution, but it does provide insurance against the development of a uniform population incapable of further evolution.

The genetic algorithm exploits the higher-payoff, or "target," regions of the solution space, because successive generations of reproduction and crossover produce increasing numbers of strings in those regions. The algorithm favors the fittest strings as parents, and so above-average strings (which fall in target regions) will have more offspring in the next generation.

Indeed, the number of strings in a given region increases at a rate proportional to the statistical estimate of that region's fitness. A statistician would need to evaluate dozens of samples from thousands or millions of regions to estimate the average fitness of each region. The genetic algorithm manages to achieve the same result with far fewer strings and virtually no computation.

The key to this rather surprising behavior is the fact that a single string belongs to all the regions in which any of its bits appear. For example, the string 11011001 is a member of regions 11****** (where the * indicates that a bit's value is unspecified), 1******1, **0**00* and so forth. The largest regions—those containing many unspecified bits—will typically be sampled by a large fraction of all the strings in a population. Thus, a genetic algorithm that manipulates a population of a few thousand strings actually samples a vastly larger number of regions. This implicit parallelism gives the genetic algorithm its central advantage over other problem-solving processes.

Crossover complicates the effects of implicit parallelism. The purpose of crossing strings in the genetic algorithm is to test new parts of target regions rather than testing the same string over and over again in successive generations. But the process can also "move" an offspring out of one region into another, causing the sampling rate of different regions to depart from a strict proportionality to average fitness. That departure will slow the rate of evolution.

The probability that the offspring of two strings will leave its parents' region depends on the distance between the 1's and 0's that define the region. The offspring of a string that samples 10****, for example, can be outside that region only if crossover begins at the second position in the string—one chance in five for a string containing six genes. (The same building block would run a risk of only one in 999 if contained in a 1,000-gene string.) The offspring of a six-gene string that samples region 1****1 runs the risk of leaving its parents' region no matter where crossover occurs.

Closely adjacent 1's or 0's that define a region are called compact building blocks. They are most likely to survive crossover intact and so be propagated into future generations at a rate proportional to the average fitness of strings that carry them. Although a reproduction mechanism that includes crossover does not manage to sample all regions at a rate proportional to their fitness, it does succeed in doing so for all regions defined by compact building blocks. The number of compactly defined building blocks in a population of strings still vastly exceeds the number of strings, and so the genetic algorithm still exhibits implicit parallelism.

Curiously, an operation in natural genetics called inversion occasionally rear-ranges genes so that those far apart in the parents may be placed close to one another in the offspring. This amounts to redefining a building block so that it is more compact and less subject to being broken up by crossover. If the building block specifies a region of high average fitness, then the more compact version automatically displaces the less compact one because it loses fewer offspring to copying error. As a result, an adaptive system using inversion can discover and favor compact versions of useful building blocks.

The genetic algorithm's implicit parallelism allows it to test and exploit large numbers of regions in the search space while manipulating relatively few strings. Implicit parallelism also helps genetic algorithms to cope with nonlinear problems—those in which the fitness of a string containing two particular building blocks may be much greater (or much smaller) than the sum of the fitnesses attributable to each building block alone.

Linear problems present a reduced search space because the presence of a 1 or a 0 at one position in a string has no effect on the fitness attributable to the presence of a 1 or 0 somewhere else. The space of 1,000-digit strings, for example, contains more than $3^{1,000}$ possibilities, but if the problem is linear, an algorithm need investigate only strings containing 1 or 0 at each position, a total of just 2,000 possibilities.

When the problem is nonlinear, the difficulty increases sharply. The average fitness of strings in the region *01***, for example, could be above the population average, even though the fitnesses associated with *0**** and **1*** are below the population average. Nonlinearity does not mean that no useful building blocks exist but merely that blocks consisting of single 1's or 0's will be inadequate. That characteristic, in turn, leads to an explosion of possibilities: the set of all strings 20 bits in length contains more than three billion building blocks. Fortunately, implicit parallelism can still be exploited. In a population of a few thousand strings, many compact building blocks will appear in 100 strings or more, enough to provide a good statistical sample. Building blocks that exploit nonlinearities to attain above-average performance will automatically be used more often in future generations.

In addition to coping with nonlinearity, the genetic algorithm helps to solve a conundrum that has long bedeviled conventional problem-solving methods: striking a balance between exploration and exploitation. Once one finds a good strategy for playing chess, for example, it is possible to concentrate on exploiting that strategy. But this choice carries a hidden cost because exploitation makes the discovery of truly novel strategies unlikely. Improvements come from trying new, risky things. Because many of the risks fail, exploration involves a degradation of performance. Deciding to what degree the present should be mortgaged for the future is a classic problem for all systems that adapt and learn.

The genetic algorithm's approach to this obstacle turns on crossover. Although crossover can interfere with the exploitation of a building block by breaking it up, this process of recombination tests building blocks in new combinations and new contexts. Crossover generates new samples of above-average regions, confirming or disproving the estimates produced by earlier samples. Furthermore, when crossover breaks up a building block, it produces a new block that enables the genetic algorithm to test regions it has not previously sampled.

The game known as the Prisoner's Dilemma illustrates the genetic algorithm's ability to balance exploration against exploitation. This long-studied game presents its two players with a choice between "cooperation" and "defection": if both players cooperate, they both receive a payoff; if one player defects, the defector receives a higher payoff and the cooperator receives nothing; if both defect, they both receive a minimal payoff [*see table on page 71*]. For example, if player A cooperates and player B defects, then player A receives no payoff and player B receives a payoff of five points.

Political scientists and sociologists have studied the Prisoner's Dilemma because it provides a simple, clear-cut ex-

How to Build a Classifier System

Building a computer algorithm that can evolve requires a way of representing the program so that any change in its genotype (the bits that compose the program) leads to a meaningful change in its phenotype (what the program does). A classifier consists simply of strings representing possible characteristics of the program's input and actions to take (*below*). Changing any symbol in a string changes its behavior.

CLASSIFIER ALPHABET	1 = YES	0 = NO	# = DON'T CARE

A classifier system to emulate a frog, for example, might contain strings that react to objects that the frog sees. Depending on an object's motion, size, location and other attributes, the frog would attack, pursue or ignore it. Several strings may match the same input string; the one with the fewest "don't care" symbols governs the system's actions.

MOVING	ON THE GROUND	LARGE	FAR	STRIPED	FLEE!	PURSUE!
INPUT					OUTPUT	

1	#	#	#	#	1	0

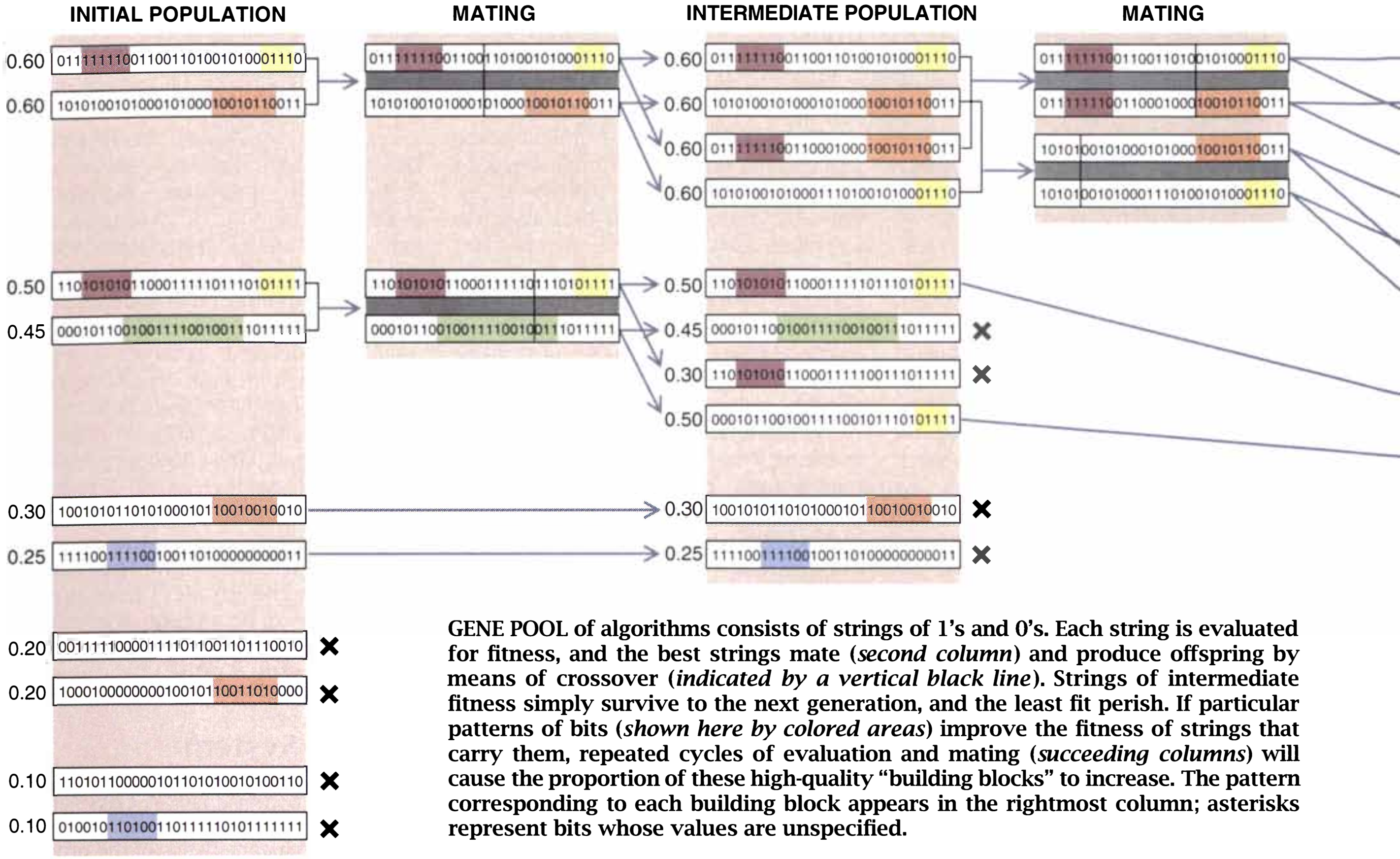
IF OBJECT IS MOVING, FLEE

1	0	0	0	#	0	1

IF OBJECT IS MOVING, IN THE AIR, SMALL AND NEAR, PURSUE

1	0	0	0	1	0	0

IF OBJECT IS MOVING, IN THE AIR, SMALL, NEAR AND STRIPED, DO NOTHING



GENE POOL of algorithms consists of strings of 1's and 0's. Each string is evaluated for fitness, and the best strings mate (*second column*) and produce offspring by means of crossover (*indicated by a vertical black line*). Strings of intermediate fitness simply survive to the next generation, and the least fit perish. If particular patterns of bits (*shown here by colored areas*) improve the fitness of strings that carry them, repeated cycles of evaluation and mating (*succeeding columns*) will cause the proportion of these high-quality "building blocks" to increase. The pattern corresponding to each building block appears in the rightmost column; asterisks represent bits whose values are unspecified.

ample of the difficulties of cooperation. Game theory predicts that each player should minimize the maximum damage the other player can inflict: that is, both players should defect. Yet when two people play the game together repeatedly, they typically learn to cooperate with each other to raise their joint payoff. One of the most effective known strategies for the Prisoner's Dilemma is "tit for tat," which begins by cooperating but thereafter mimics the last play of the other player. That is, it "punishes" a defection by defecting the next time, and it rewards cooperation by cooperating the next time.

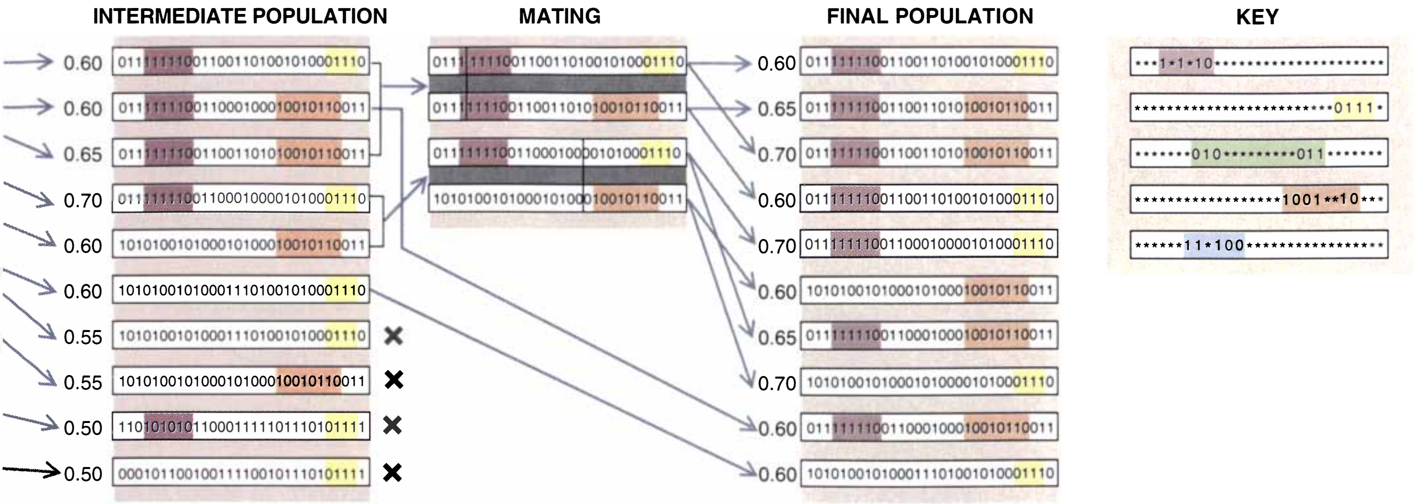
Robert Axelrod of the University of Michigan, working with Stephanie Forrest, now at the University of New Mexico, decided to find out if the genetic algorithm could discover the tit-for-tat strategy. Applying the genetic algorithm first requires translating possible strategies into strings. One simple way is to base the next response on the outcome of the last three plays. Each iteration has four possible outcomes, and so a sequence of three plays yields 64 possibilities. A 64-bit string contains one gene (or bit position) for each. The first gene, for instance, would be allocated to the case of three consecutive mutual cooperations and the last to three mutual defections. The value of each gene would be either 1 or 0 depending on whether the preferred response to its corresponding history was cooperation or defection. For example, the 64-bit string consisting of all 0's would designate the strategy that defects in all cases. Even for such a simple game, there are 2^{64} (approximately 16 quadrillion) different strategies.

Axelrod and Forrest supplied the genetic algorithm with a small random collection of strings representing strategies. The fitness of each string was simply the average of the payoffs its strategy received under repeated play. All these strings had low fitnesses because most strategies for playing the Prisoner's Dilemma are not very good. Quickly the genetic algorithm discovered and exploited tit for tat, but further evolution introduced an additional improvement. The new strategy, discovered while the genetic algorithm was already playing at a high level, exploited players that could be "bluffed"—lured into cooperating repeatedly in the face of defection. It reverted to tit for tat, however, when the history indicated the player could not be bluffed.

Biological evolution operates, of course, not to produce a single superindividual but rather to produce interacting species well adapted to one another. (Indeed, in the biological realm there is no such thing as a best individual.) Similarly, the genetic algorithm can be used, with modifications, to govern the evolution not merely of individual rules or strategies but of classifier-system "organisms" composed of many rules. Instead of selecting the fittest rules in isolation, competitive pressures can lead to the evolution of larger systems whose abilities are encoded in the strings that make them up.

Re-creating evolution at this higher level requires several modifications to the original genetic algorithm. Strings still represent condition-action rules, and each rule whose conditions are met generates an action as before. Rating each rule by the number of correct actions it generates, however, will favor the evolution of individual "superrules" instead of finding clusters of rules that interact usefully. To redirect the search toward interacting rules, the procedure is modified by forcing rules to compete for control of the system's actions. Each rule whose conditions are met competes with all other rules whose conditions are met, and the strongest rules determine what the system will do in that given situation. If the system's actions lead to a successful outcome, all the winning rules are strengthened; otherwise they are weakened.

Another way of looking at this method is to consider each rule string as a



hypothesis about the classifier's world. A rule enters the competition only when it "claims" to be relevant to the current situation. Its ability to compete depends on how much of a contribution it has made to solving similar problems. As the genetic algorithm proceeds, strong rules mate and form offspring rules that combine their parents' building blocks. These offspring, which replace the weakest rules, amount to plausible but untried hypotheses.

Competition among rules provides the system with a graceful way of handling perpetual novelty. When a system has strong rules that respond to a particular situation, that is the equivalent of saying that it has certain well-validated hypotheses. Offspring rules, which begin life weaker than do their parents, can win the competition and influence the system's behavior only when there are no strong rules whose conditions are satisfied—in other words, when the system does not know what to do. If their actions help, they survive; if not, they are soon replaced. Thus, the offspring do not interfere with the system's action in well-practiced situations but wait gracefully in the wings as hypotheses about what to do under novel circumstances.

Adding competition in this way strongly affects the evolution of a classifier system. Shortly after the system starts running, it evolves rules with simple conditions—treating a broad range of situations as if they were identical. The system exploits such rules as defaults that specify something to be done in the absence of more detailed information. Because the default rules make only coarse discriminations, however, they are often wrong and so do not grow in strength. As the system gains experience, reproduction and crossover lead to the development of more complex, specific rules that rapidly become strong enough to dictate behavior in particular cases.

Eventually the system develops a hierarchy: layers of exception rules at the lower levels handle most cases, but the default rules at the top level of the hierarchy come into play when none of the detailed rules has enough information to satisfy its conditions. Such default hierarchies bring relevant experience to bear on novel situations while preventing the system from becoming bogged down in overly detailed options.

The same characteristics that make evolving classifier systems adept at handling perpetual novelty also do a good job of handling situations where the payoff for a given action may come only long after the action is taken. The earliest moves of a chess game, for example, may set the stage for later victory or defeat.

To train a classifier system for such long-term goals, a programmer gives the system a payoff each time it completes a task. The credit for success (or the blame for failure) can propagate through the hierarchy to strengthen (or weaken) individual rules even if their actions had only a distant effect on the outcome. Over the course of many generations the system develops rules that act ever earlier to set the stage for later payoffs. It therefore becomes increasingly able to anticipate the consequences of its actions.

Genetic algorithms have now been tested in a wide variety of contexts. David E. Goldberg of the University of Illinois, for example, has developed algorithms that learn to control a gas pipeline system modeled on the one that carries natural gas from the Southwest to the Northeast. The pipeline complex consists of many branches, all carrying various amounts of gas; the only controls available are compressors that increase pressure in a particular branch of the pipeline and valves to regulate the flow of gas to and from storage tanks. Because of the tremendous lag between manipulating valves or compressors and the actual pressure changes in the lines, there is no analytic solution to the problem, and human controllers, like Goldberg's algorithm, must learn by apprenticeship.

Goldberg's system not only met gas demand at costs comparable to those achieved in practice, but it also developed a hierarchy of default rules capable of responding properly to holes punched in the pipeline (as happens all too often in reality at the blade of an errant bulldozer). Lawrence Davis of Tica Associates in Cambridge, Mass., has used similar techniques to design communications networks; his software's goal is to carry the maximum possible amount of data with the minimum number of transmission lines and switches interconnecting them.

A group of researchers at General Electric and Rensselaer Polytechnic In-

The Prisoner's Dilemma

PLAYER	(B) COOPERATE	(B) DEFECT
(A) COOPERATE	3/3	5/0
(A) DEFECT	0/5	0/0

IN PRISONER'S DILEMMA each player can either cooperate or defect and receives a payoff based on the other's choice. If both cooperate, for example, both receive three points. Mutual defection is the safest strategy, but repeated play often leads to cooperation instead.

SOFTWARE TO DESIGN JET TURBINE includes a genetic algorithm that combines the best features of designs produced by other programs. Engineers using the algorithm achieved better results than with more conventional software aids.

stitute recently put a genetic algorithm to good use in the design of a high-bypass jet engine turbine such as those that power commercial airliners. Such turbines, which consist of multiple stages of stationary and rotating blade rows enclosed in a roughly cylindrical duct, are at the center of engine-development projects that last five years or more and consume up to $2 billion.

The design of a turbine involves at least 100 variables, each of which can take on a different range of values. The resulting search space contains more than 10^{387} points. The "fitness" of the turbine depends on how well it satisfies a series of 50 or so constraints, such as the smooth shape of its inner and outer walls or the pressure, velocity and turbulence of the flow at various points inside the cylinder. Evaluating a single design requires running an engine simulation that takes about 30 seconds on a typical engineering workstation.

In one fairly typical case, an engineer working alone took about eight weeks to reach a satisfactory design. So-called expert systems, which use inference rules based on experience to predict the effects of a change of one or two variables, can help direct the designer in seeking out useful changes. An engineer using such an expert system took less than a day to design an engine with twice the improvements of the eight-week manual design.

Such expert systems, however, soon get stuck at points where further improvements can be made only by changing many variables simultaneously. These dead ends occur because it is practically impossible to sort out all the effects associated with different multiple changes, let alone to specify the regions of the design space within which previous experience remains valid.

To get away from such a point, the designer must find new building blocks for a solution. Here is where the genetic algorithm comes into play. Seeding the algorithm with designs produced by the expert system, an engineer took only two days to find a design with three times the improvements of the manual version (and half again as many as using the expert system alone).

This example points up both a strength and a limitation of simple genetic algorithms: they are at their best when exploring complex landscapes to locate regions of enhanced opportunity. But if a partial solution can be improved further by making small changes in a few variables, it is best to augment the genetic algorithm with other, more standard methods.

Although genetic algorithms mimic the effects of natural selection, until now they have operated on a much smaller scale than does biological evolution. My colleagues and I have run classifier systems containing as many as 8,000 rules, but this size is at the low end of viability for natural populations. Large animals that are not endangered may number in the millions, insect populations in the trillions and bacteria in the quintillions or more. These large numbers greatly enhance the advantages of implicit parallelism.

As massively parallel computers become more common, it will be feasible to evolve software populations whose size more closely approaches those of natural species. Indeed, the genetic algorithm lends itself nicely to such machines. Each processor can be devoted to a single string because the algorithm's operations focus on single strings or, at most, a pair of strings during crossover. As a result, the entire population can be processed in parallel.

We still have much to learn about classifier systems, but the work done so far suggests they will be capable of remarkably complex behavior. Rick L. Riolo of the University of Michigan has already observed genetic algorithms that display "latent learning" (a phenomenon in which an animal such as a rat explores a maze without reward and is subsequently able to find food placed in the maze much more quickly).

At the Santa Fe Institute, Forrest, W. Brian Arthur, John H. Miller, Richard G. Palmer and I have used classifier systems to simulate economic agents of limited rationality. These agents evolve to the point of acting on trends in a simple commodity market, producing speculative bubbles and crashes.

The simulated worlds these agents inhabit show many similarities to natural ecosystems: they exhibit counterparts to such phenomena as symbiosis, parasitism, biological "arms races," mimicry, niche formation and speciation. Still other work with genetic algorithms has shed light on the conditions under which evolution will favor sexual or asexual reproduction. Eventually artificial adaptation may repay its debt to nature by increasing researchers' understanding of natural ecosystems and other complex adaptive systems.

FURTHER READING

INDUCTION: PROCESSES OF INFERENCE, LEARNING, AND DISCOVERY. J. H. Holland, K. J. Holyoak, R. E. Nisbett and P. R. Thagard. MIT Press, 1986.

GENETIC ALGORITHMS AND SIMULATED ANNEALING. Edited by Lawrence Davis. Morgan Kaufmann, 1987.

GENETIC ALGORITHMS IN SEARCH, OPTIMIZATION, AND MACHINE LEARNING. D. E. Goldberg. Addison-Wesley, 1989.

GENETIC ALGORITHMS: PROCEEDINGS OF THE FOURTH INTERNATIONAL CONFERENCE. Edited by Richard Belew and Lashon Booker. Morgan Kaufmann, 1991.

ADAPTATION IN NATURAL AND ARTIFICIAL SYSTEMS. J. H. Holland. MIT Press, 1992.

COMPLEX ADAPTIVE SYSTEMS. J. H. Holland in *Dædalus,* Vol. 121, No. 1, pages 17–30; Winter 1992.

STANFORD GRAPHICS FOR WINDOWS

Finally...presentation graphics with numerical clout!

*"Stanford is a breakthrough product that combines technical charting, data analysis, and presentation tools. Basically, it's fabulous." **

– PC Magazine
March, 1992

"Stanford Graphics offers the best of both worlds. It combines top-notch presentation features with unprecedented technical sophistication. We were greatly impressed."

– Infoworld,
Sept., 1991

"Stanford Graphics slams the competition..."

– Infoworld
April, 1992

Does your presentation software give you all beauty and no brains? Where is the substance behind those pretty pictures? What if your presentation program had enough data-handling muscle to meet the needs of business *and* statistical *and* scientific users?

With on-screen curve fitting, formula solving, statistics, and data manipulation, Stanford Graphics advances both the art and science of technical presentations.

All this and Windows too. Stanford Graphics takes full advantage of the latest in Windows 3.0 and 3.1 technology, with sophisticated DDE, OLE, and TrueType font support.

Stanford Graphics for Windows: Presentation and Analysis.

Finally, form *and* substance.

STANFORD GRAPHICS™
PRESENTATION & ANALYSIS

3-D VISIONS
Call Toll Free for Details
1-800-729-4723
2780 Skypark Drive, Torrance, CA 90505
Tel: (310) 325-1339 FAX: (310) 325-1505

Breath Tests in Medicine

Since antiquity, it has been known that the breath contains clues to many diseases. Researchers are developing a new generation of breath tests as an alternative to invasive diagnostic procedures

by Michael Phillips

For many people, the idea of breath testing conjures up a vision of a roadside interrogation by a grim-faced police officer. A motorist suspected of being intoxicated must blow into a hand-held instrument that, after a few seconds, indicates the amount of alcohol in the blood. Despite this unflattering association with inebriation and arrest, breath testing is blossoming into an exciting area of medical technology. Physicians have begun using these tests to diagnose an increasingly wide variety of diseases without the hazards or discomforts of invasive procedures. Furthermore, breath tests are providing important new insights into the understanding of basic biochemical functions of the body.

The recent upsurge in interest stems from advances in analytic technology, which have made it possible to identify a rapidly growing number of compounds in the breath. Clinical studies with these improved assays have shown that the presence of abnormal chemicals in the breath can aid in the early diagnosis of many diseases.

The breath contains valuable information because only a slender barrier separates the air in the alveoli of the lung from the blood in the capillaries. This barrier is called the pulmonary alveolar membrane. Like water flowing down a hill, a volatile organic compound, such as alcohol or methyl mercaptan (the fragrant compound in garlic), will diffuse across the alveolar membrane from the compartment with the higher vapor pressure to the lower, from the air into the blood, or vice versa.

The detection of volatile organic compounds in the breath has a long history. Since the time of Hippocrates, physicians have known that the aroma of human breath can provide clues to diagnosis. The astute clinician is alert for the sweet, fruity odor of acetone in patients with uncontrolled diabetes, the musty, fishy reek of advanced liver disease, the urinelike smell that accompanies failing kidneys and the putrid stench of a lung abscess.

But without objective chemical analysis, breath testing as a medical tool could not have progressed beyond educated sniffing. This problem attracted the attention of scientists more than 200 years ago. Antoine Laurent Lavoisier, celebrated as the father of the "chemical revolution" and renowned for his discovery of the role of oxygen in combustion, was also a pioneer in breath testing.

In 1784 he and Pierre Simon Laplace, the French mathematical physicist, analyzed the breath of a guinea pig. They found that the animal consumed oxygen and expired carbon dioxide. This finding was the first direct evidence that food undergoes combustion in the body. In this discovery lies both the foundation of modern biochemistry and the expression "to be a guinea pig," as Lavoisier confirmed the finding by using himself as a subject.

Lavoisier's apparatus included an ingenious innovation that researchers today still use in various guises: the breath trap. This device accumulates and concentrates components of the breath. Lavoisier's trap was a chemical solution through which he bubbled a large volume of breath. The carbon dioxide in the breath reacted with the solution to form a visible precipitate.

Carbon dioxide is relatively easy to detect, because it makes up approximately 5 percent of the breath. Unfortunately, most other volatile compounds in the breath are present in much lower concentrations, in parts per million or less. Their detection had to wait until the mid-19th century, when colorimetric assays were introduced. In such an analysis, an organic compound interacts with reagents in a solution to produce a change in color.

One of the earliest to exploit colorimetric analysis was A. Nebelthau, a physician at the Marburg Polyclinic in Germany. He constructed a device to analyze the breath of patients suffering from diabetes mellitus. When uncontrolled, this condition causes the blood glucose to rise to excessively high levels; as a result, the body generates large quantities of acetone as a major metabolite. When Nebelthau bubbled a patient's breath through the alkaline iodine trap of his apparatus, he observed a rapid and intense change in color. The change demonstrated that unusual amounts of acetone were being excreted through the lungs.

In 1874 the British physician Francis E. Anstie applied colorimetric analysis to study the fate of alcohol in the human body. His objective was to resolve the controversy raging between physiologists, who believed the body broke down alcohol as it does food, and temperance activists, who thought the body excreted it unchanged as if it were a foreign substance. Anstie's breath trap contained a solution of chromic acid, which changed from red-brown to green in the presence of ethyl alcohol. In an

BREATH-COLLECTION DEVICE in the author's laboratory provides the patient with air purified by activated carbon (in containers at the left). Exhaled breath passes through a drying agent (in container at the right) and is trapped for later analysis. The nose clip ensures the patient breathes through the tubing.

MICHAEL PHILLIPS is associate director of medicine at St. Vincent's Medical Center of Richmond in Staten Island and clinical professor of medicine at New York Medical College in Valhalla. He received his degree from the University of Western Australia and has had appointments at the University of Connecticut and Georgetown University. He writes that his main interests are in the clinical pharmacology of alcohol abuse and in "the less popular exhalations of the human body—breath and sweat—for the neglected windows they open onto metabolic processes."

elegant series of experiments, he demonstrated that the amount of alcohol excreted via the breath and other routes fell far short of the amount consumed. Therefore, most of the alcohol must have been metabolized. (He is also remembered as the originator of Anstie's limit: more than two alcoholic drinks a day may be injurious to one's health, a concept strongly supported by modern findings.)

Current hand-held breath analyzers for alcohol are electronic: the donor blows into a plastic tube, and a spring-loaded piston withdraws approximately one cubic centimeter of breath, which is oxidized in a fuel cell. The resulting electric current is proportional to the concentration of alcohol in the breath. After a few seconds, the breath analyzer calculates and then displays digitally the blood alcohol concentration.

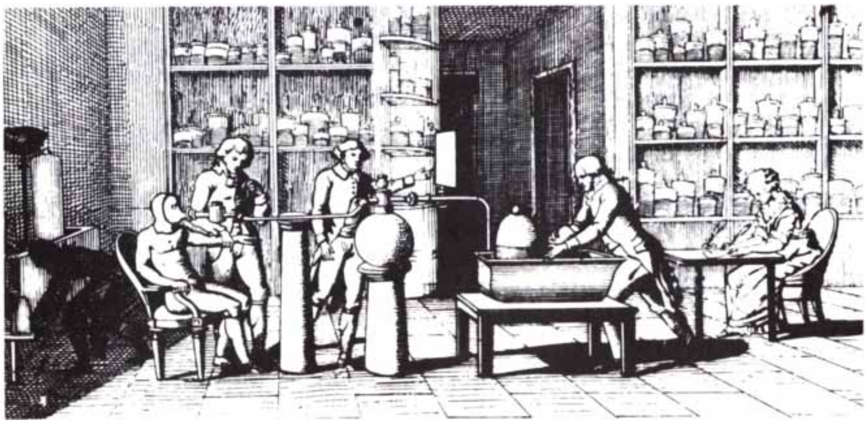
As testing for alcohol demonstrates, it is easy to detect a volatile organic compound in the breath if the substance has been consumed in large quantities beforehand. Consequently, many breath tests require a patient to consume a dose of a specific precursor to a volatile chemical. The disease reveals itself when an abnormal quantity of the breakdown products appears in the breath.

For example, physicians use this approach to diagnose an intestinal disorder known as malabsorption syndrome. This condition causes severe and chronic diarrhea because the diseased intestine cannot absorb food efficiently. One consequence of the syndrome is that most of the sugar in food arrives at the colon intact, where bacteria break it down. The bacterial action releases hydrogen, which is absorbed into the blood and excreted through the lungs. As a test, a physician gives the patient an oral dose of xylose, a five-carbon sugar normally absorbed completely by the small intestine. The appearance of large amounts of hydrogen in the breath within the next few hours confirms the diagnosis of malabsorption syndrome.

This test can also be refined to demonstrate less common kinds of malabsorption, in which the intestine fails to absorb a specific carbohydrate. For instance, a patient who has a deficiency of lactase in the small bowel will produce hydrogen after a dose of lactose but not after a dose of a different sugar, such as xylose or glucose. Other disorders of the small intestine, such as bacterial overgrowth, can also be diagnosed by breath testing. This condition may result from any cause of stagnation in the small intestine, including scarring resulting from a surgical procedure or damage to the nerves that control the propulsion rate of food through the gut.

Breath testing is also used to identify damage to the pancreas. Jay A. Perman of the Johns Hopkins University School of Medicine studied children with cystic fibrosis, an often fatal condition that severely harms the pancreas and lungs. He gave them a dose of rice starch, a complex carbohydrate. Children with impaired pancreatic function could not secrete sufficient amylase to digest the meal in the small intestine. Instead the bacteria in the colon completed the digestive task, producing hydrogen, which was detectable in the breath.

Although the presence of hydrogen can help diagnose several disorders of the intestine, other markers are also useful. For example, radioactive carbon 14 is used to identify diseases of the pancreas. The carbon 14 is combined with other compounds that are metabolized by pancreatic enzymes. The amount of radioactive carbon dioxide



LAVOISIER'S LABORATORY to study human breath may be the earliest recorded. In this woodcut, a subject breathes oxygen through a tube, so that the French chemist can study respiration. Madame Lavoisier, seated at the right, records the data.

($^{14}CO_2$) that subsequently appears in the breath indicates how well the pancreas is functioning. For instance, a diseased pancreas may not secrete sufficient lipase to break down dietary fats. A physician can identify the disorder if a reduced amount of $^{14}CO_2$ appears in the breath after the patient has consumed a dose of a radiolabeled triglyceride, a low-molecular-weight fat.

Although breath tests for disorders of the small intestine and pancreas are the primary ones now used by the medical community, recent studies have shown that the technique holds promise in recognizing other conditions. Breath tests employing radioactive carbon 14 labeling may be especially useful in diagnosing two of the most common stomach disorders seen in general medical practice: peptic ulcer disease and chronic gastritis. The conditions may result from *Helicobacter pylori* infections, which appear to be quite common. The infections are probably underdiagnosed because confirmation usually requires endoscopy, in which the physician passes an instrument into the stomach to collect a sample of the mucosal lining.

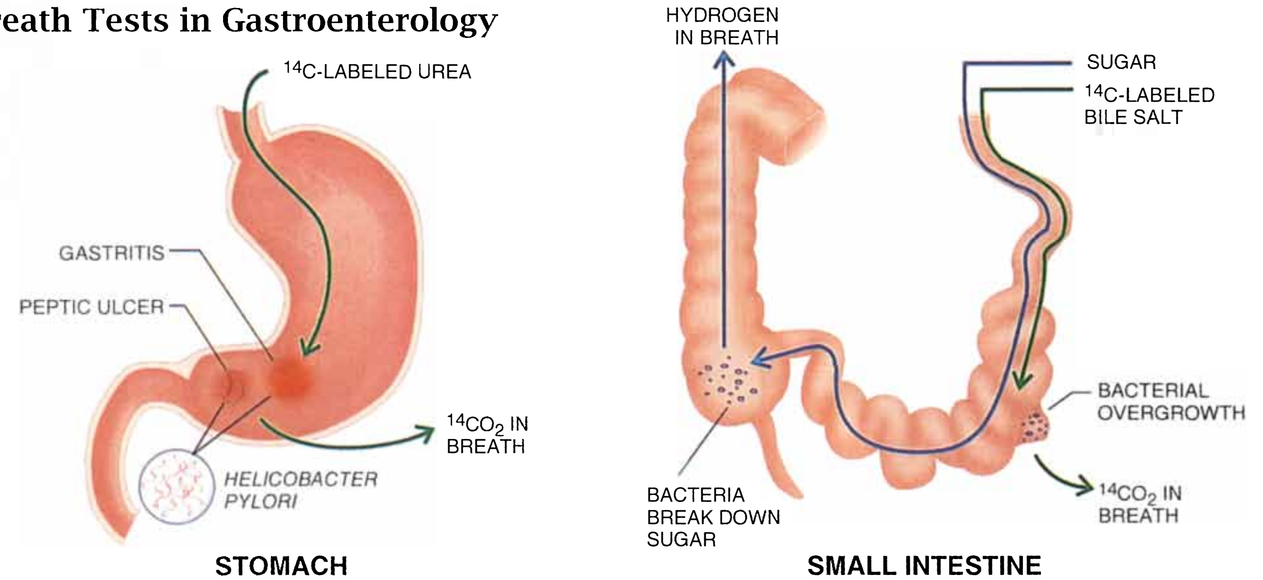
Barry J. Marshall of the University of Virginia Health Sciences Center has shown that breath tests are a highly reliable noninvasive alternative to endoscopy in the detection of *Helicobacter* infections. *Helicobacter* possesses an enzyme called urease, which is not present in humans. If an infected patient consumes a dose of urea labeled with carbon 14, the urease in *Helicobacter* will break down the urea in the stomach. One of the metabolites will be $^{14}CO_2$, which is absorbed by the blood and excreted in the breath.

A similar test is also proving useful in uncovering damage to the liver. Conditions such as cirrhosis and hepatitis often go unnoticed in their early stages. By the time bilirubin, a major metabolite of hemoglobin, builds up in the blood to cause the characteristic jaundice, a patient may have lost more than 50 percent of the cells in the liver.

Breath testing can measure the impairment of specific metabolic pathways before jaundice sets in. Specifically, it can test the *N*-demethylation pathway in the hepatic microsomes, which is affected early by any kind of liver damage. The demethylation process in a healthy liver releases substantial amounts of carbon dioxide. To detect liver injury, a physician can give a patient a dose of a tracer compound, such as aminopyrine, phenacetin or galactose, labeled with radioactive carbon 14. Any impairment of *N*-demethylation capacity will reveal itself through a marked reduction in the rate at which $^{14}CO_2$ appears in the breath.

There are other markers of liver disease. For example, the liver normally

Breath Tests in Gastroenterology



Peptic ulcers and chronic gastritis may result from infections of *Helicobacter pylori*. The bacteria will break down a dose of ^{14}C-labeled urea into $^{14}CO_2$, which is absorbed by the blood and will appear in the breath.

In malabsorption syndrome, the small intestine is unable to absorb a sugar, such as xylose, lactose, sucrose, glucose, fructose or lactose. Hydrogen will appear in the breath. A dose of a bile salt labeled with carbon 14 will reveal bacterial overgrowth.

breaks down dimethyl sulfide, a volatile metabolite of the amino acid methionine. But if the liver's functions are impaired, a greater quantity of dimethyl sulfide "spills over" into the breath.

Diagnosis of disease is only one realm in which breath testing can be useful. Drug monitoring is a neglected but potentially fruitful field, because most prescribed and illicit drugs have low molecular weights. Their vapor pressures at body temperature may be sufficiently high so that they or their metabolites are propelled into the alveolar air in measurable quantities. Indeed, breath tests for alcohol are so widely employed that it is curious that testing for other drugs is still comparatively unexplored. Metabolites of marijuana can be readily measured in the breath for several days after use.

Breath testing can also illuminate differences in metabolism. It had long been observed that Asians react to the consumption of alcohol in a different manner from most whites. Many Asians experience flushing after a drink, as well as more severe reactions, such as palpitations, nausea and malaise. Breath tests have confirmed that concentrations of acetaldehyde, a volatile metabolite of alcohol, are considerably higher in Asians than in whites after consumption of an alcoholic beverage. Further research has shown that more than 80 percent of Asians generally have relatively low levels of aldehyde dehydrogenase in the liver microsomes. A dose of alcohol results in a greater buildup of acetaldehyde in the blood, causing flushing and other unpleasant symptoms.

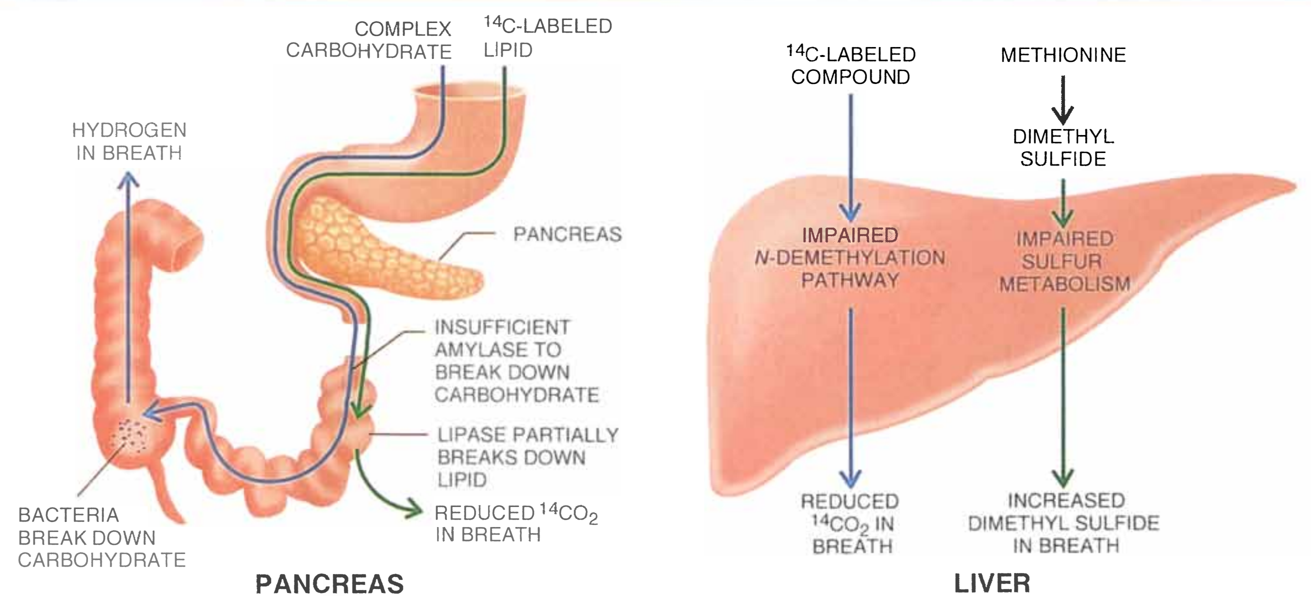
Breath testing is already proving useful in monitoring the exposure of industrial workers to potentially hazardous solvents and petrochemicals. Several of these toxic compounds can be readily measured in human breath. Sensitive breath tests developed by Edo D. Pellizzari, now at the Research Triangle Institute in North Carolina, and his colleagues at the Environmental Protection Agency have shown that toxic chemicals can pollute air far from the factory floor. It is now clear that indoor air in the workplace, office and home may be polluted with small amounts of several compounds, some of them carcinogens. In one study Pellizzari and his colleagues found that workers in hardware stores had ingested volatile compounds that may have originated from carpet solvents and plastics.

All the breath tests described so far detect either a volatile organic compound that had been previously consumed (such as alcohol) or the metabolites of a precursor. Researchers have attempted to advance one step further by detecting the volatile compounds present in the breath under normal circumstances. In 1971 Linus Pauling described an elegantly simple method for microanalysis of normal breath. He first passed breath through a cold trap consisting of a stainless-steel tube chilled by dry ice. He then assayed the condensate by gas chromatography and mass spectroscopy. He observed approximately 250 different compounds. This surprisingly large number indicated that the composition of human breath was more complex than had been previously suspected. Other laboratories have confirmed his findings, and researchers have now isolated nearly 400 volatile organic compounds in normal human breath.

Why have scientists not routinely assayed these compounds, if the technology has been available for more than 20 years? Let me attempt to answer this question by inviting the reader to accompany me on an imaginary visit to my laboratory at Bayley Seton Hospital in Staten Island. As you enter, I point out the instruments we use to analyze volatile compounds in breath: thermal desorbers, gas chromatographs and a mass spectrometer, their functions all orchestrated by microcomputers.

But, you wonder, how does one actually collect breath samples in the first place? I now display our collecting apparatus, which is something of an anticlimax. It is a handcrafted and somewhat bizarre-appearing device on a wheeled cart. Several meters of plastic tubing wind in seemingly disparate directions; a plywood container conceals a pump and even more tubing. The unit contrasts oddly with the polished elegance of the analytic instruments.

This clash of technological styles is



A lipid labeled with carbon 14 will be broken down by lipase, so any deficiency will result in a below normal amount of $^{14}CO_2$ in the breath. Insufficient production of amylase, which breaks down carbohydrates, will be revealed by the presence of hydrogen.

Impaired functioning of the *N*-demethylation pathway can be detected by giving aminopyrine, phenacetin or galactose labeled with carbon 14. Sulfur metabolism can be gauged by methionine, which is broken down by the body into dimethyl sulfide.

often encountered in laboratories where breath compounds are studied. The analytic instruments are purchased "off the shelf," while the collection technology consists of the unstandardized gadgetry created by individual tinkerers. Despite two centuries of experience gained since Lavoisier's time, there is still no general agreement on the best way to trap a sample of concentrated breath.

Much of the debate is caused by several difficulties in designing breath-collection equipment. First and foremost, a patient should have no trouble exhaling into the device. Blowing through a liquid trap can be hard work, which probably explains why Lavoisier's subjects stripped to the waist before using the instrument. Even Pauling's device, which included a tube five feet long and 0.2 inch in diameter, could not have been especially easy to blow through. In addition, mouthpiece parts must be replaceable to avoid any risk of transmitting infections from patient to patient.

Contamination of the sample is also a major consideration. Even the cleanest room is polluted with volatile chemicals from such sources as ventilation units and solvents in carpeting adhesive. The trace compounds are measurable as background "noise." In addition, the moisture in breath will condense into droplets on the tubing, which may contaminate the chemically clean trap.

For meaningful results, the apparatus must be able to sample breath that has originated deep in the lungs, because human breath is not a homogeneous gas. The first 150 milliliters of every expiration consists of "dead space" air from the upper airways, where no gas exchange has occurred (a healthy individual exhales half a liter or more with each breath). The trap should also capture only the volatile compounds of interest and allow the remainder, including nitrogen and oxygen (which together constitute more than 90 percent of the breath), to pass unimpeded.

Generally speaking, there are three types of breath traps: chemical, cryogenic and adsorptive. None is perfect. Chemical trapping usually uses traditional "wet chemistry": breath is bubbled through a reagent solution that captures a specific compound, such as ethanol or acetone. The method is simple and direct, and the trapped derivative, often colored, can be measured easily. The disadvantages are poor sensitivity and the physical effort required of the donor.

In cryogenic trapping, the volatile compounds are captured by freezing. The breath travels through a tube immersed in such cooling fluids as liquid nitrogen, which is at –196 degrees Celsius. Unfortunately, a cold trap also freezes the water and carbon dioxide in the breath and may rapidly become plugged by ice.

Adsorptive trapping has become the most convenient and most widely used method today. It captures volatile compounds by binding them to agents such as activated carbon and adsorptive resins. Recent advances in microprocessor-controlled thermal desorbers, which automatically "bake off" and concentrate trapped compounds, have made the technique even more convenient.

Several investigators have developed highly sensitive assays using adsorptive traps. They include Pellizzari, M. Sydney Gordon of Battelle, Jack O'Neill of the IIT Research Institute in Chicago and E. S. Kovaleva and R. Liedeman of the All-Union Mental Health Research Center in Moscow.

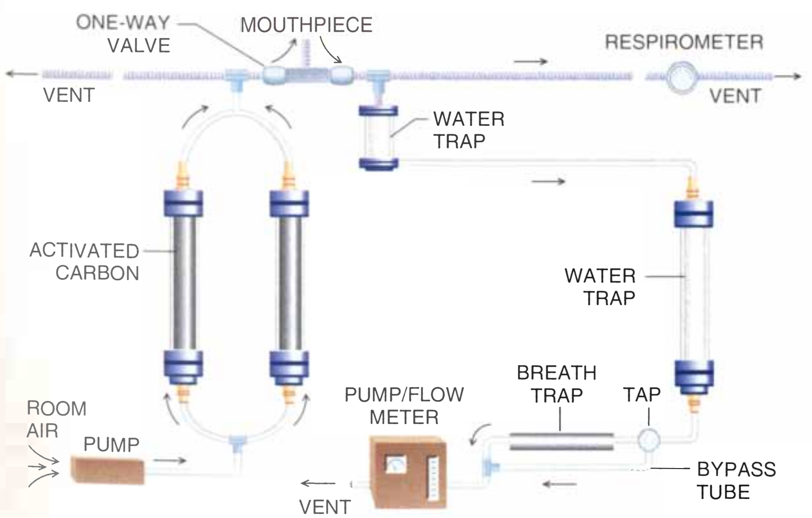
The apparatus that my colleague Joel Greenberg and I have made illustrates how adsorptive trapping can overcome most of the technical difficulties [*see illustration on this page*]. In our device the donor inhales purified air and breathes out into an exhaust tube. One-way valves ensure that there is no reflux of breath into the reservoir of purified air. There is virtually no resistance to breathing because the tubing is wide (nearly three centimeters in diameter), and the one-way valves require very little pressure to open and close. Even severely ill patients can donate a sample without much discomfort.

No sample is collected from the first part of the breath, which contains the dead-space air. Only alveolar breath from deep in the lungs is extracted from the exhaust tube. Five minutes of collection yields 10 liters of alveolar breath. The breath is dessicated in water traps and drawn through a stainless-steel tube, where activated carbon and an adsorptive resin capture the volatile organic compounds. Later, a thermal desorber automatically strips the volatile compounds from the trap and concentrates them more than 200,000 times, for separation by gas chromatography and assay by mass spectrometry.

The main drawback of adsorptive trapping is that it may have a greater affinity for some compounds than for others. The gas chromatograph is also a source of variation. It separates samples in a long, narrow tube (called the separating column), whose chemical composition influences the number of detectable compounds. Thus, our method yields only about 50 to 60 compounds. This comparatively small number has proved more than sufficient for research purposes.

Collecting Human Breath

This device, in the author's laboratory, is mounted on a cart, which can be wheeled to a patient's bedside [*see illustration on page 75*]. Room air is pumped through the activated carbon, supplying the volunteer with chemically clean air. Exhaled alveolar breath, drawn through the water trap to remove moisture, passes through the stainless-steel breath trap, which captures the volatile organic compounds. The one-way valves prevent backflow. The respirometer measures the amount of each exhalation.



At our current state of knowledge, microanalysis of volatile compounds in the breath has raised more questions than it has answered. The first and most pressing question is, from where do the compounds come? The origin of most is still unknown. Some may be environmental pollutants. For instance, we found carbon disulfide in the breath of all our human volunteers. The compound might be a breakdown product of sulfur-containing amino acids. On the other hand, we have also found carbon disulfide in air samples collected in and around New York City. The origin of the compound is not a trivial concern, because studies have shown that environmental exposure to carbon disulfide can accelerate atherosclerosis and coronary artery disease.

Another problem is the sheer overwhelming mass of data. It is difficult to identify a disease marker among nearly 400 different compounds. Even when abnormal quantities of a compound are observed in patients suffering from a particular illness, physicians must still use rigorous statistical analysis before they can be confident that the findings are not due to chance.

Studies by Gordon and O'Neill have illustrated both the promise and the potential pitfalls of identifying new disease markers in the breath. They assayed volatile chemicals in the breath of patients with lung cancer and found higher than normal concentrations of acetone, methylethylketone, *n*-propanol, tolualdehyde and oxepanone. These findings are exciting, because a breath test for the early detection of lung cancer could become an important new tool in medical practice. But physicians will need more "hard" clinical information about such tests, including its sensitivity and specificity and its positive predictive value. Until then, breath testing for lung cancer will remain only a provocative possibility.

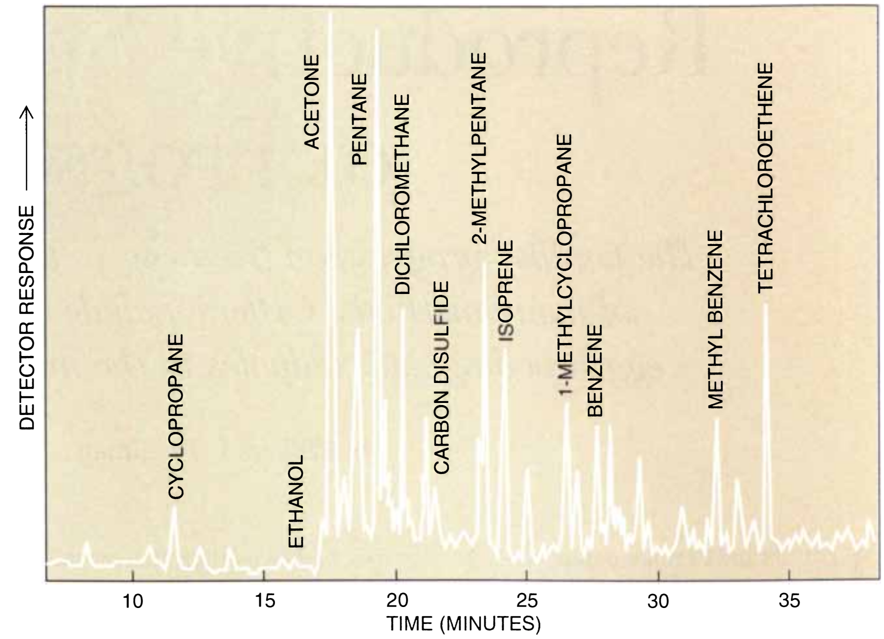
Lack of information now limits the clinical value of most breath tests. Only some are accepted as diagnostic tools. Yet all are fascinating signposts to possible future developments. For example, the measurement of dimethylamine and volatile fatty acids in the breath might enable physicians to recognize diseases of the kidneys and liver in their earliest stages, before any symptoms have developed, without invasive procedures or the administration of a precursor.

Breath tests also offer intriguing clues to the biochemical basis of many diseases whose causes are still unknown. In particular, two alkanes in the breath, pentane and ethane, are known to be elevated in several conditions. Edward J. Zarling of Loyola University Medical Center has found increased levels of pentane in the breath of patients suffering from acute myocardial infarction, arthritis or multiple sclerosis. Kovaleva and Liedeman discovered that acutely psychotic schizophrenic patients had high concentrations of pentane, which rapidly returned to normal as their clinical condition improved during treatment. Elevated levels of ethane have also been observed in the breath of humans and animals deficient in vitamin E and such trace metals as selenium and copper.

The excretion of pentane and ethane in the breath appears to result from an accumulation of free radicals of oxygen in a damaged cell. These free radicals attack the fatty acids in cell membranes, a process that releases the alkanes. The toxic effects of the free radicals have assumed an increasing importance in recent years, because they seem to mediate the final stages of cellular damage arising from a wide range of causes.

How can breath testing make the transition from the realm of research into the mainstream of clinical practice? The two most pressing needs are for simplification and clinical evaluation. Simplification is essential, because the methods currently employed to collect, concentrate and assay breath samples are too complicated and expensive for general use.

This situation may change during the next decade. New assay instruments on the horizon, such as spectroscopes that use infrared lasers and mass spectrometers that provide instantaneous readouts, may be able to detect low levels of volatile organic compounds in an unconcentrated sample of breath. Improved collection procedures combined with rapid assays should enable researchers to conduct detailed clinical studies of the diagnostic value of breath tests. Physicians and patients of the 21st century may eventually come to think of a breath test in much the same way we now think of a chest x-ray: as an inexpensive and convenient screening test that can help the physician diagnose several diseases in their earliest and most treatable stages.



VOLATILE ORGANIC COMPOUNDS in human breath are identified by mass spectroscopy; they show up as large peaks. Only some of the 50 to 60 chemicals detected are labeled. Other assays have discerned nearly 400 compounds.

FURTHER READING

BREATH TESTS: PRINCIPLES, PROBLEMS, AND PROMISE. C. W. Lo, E. A. Carter and W. A. Walker in *Advances in Pediatrics,* Vol. 29. Edited by Lewis A. Barness. Year Book Medical Publishers, 1982.

THE DIAGNOSTIC POTENTIAL OF BREATH ANALYSIS. A. Manolis in *Clinical Chemistry,* Vol. 29, No. 1, pages 5–15; January 1983.

CLINICAL APPLICATION OF BREATH HYDROGEN MEASUREMENTS. Jay A. Perman in *Canadian Journal of Physiology and Pharmacology,* Vol. 69, No. 1, pages 111–115; January 1991.

ION-TRAP DETECTION OF VOLATILE ORGANIC COMPOUNDS IN ALVEOLAR BREATH. Michael Phillips and Joel Greenberg in *Clinical Chemistry,* Vol. 38, No. 1, pages 60–65; January 1992.

Reproductive Strategies of Frogs

The familiar progression from egg to tadpole is only one of many methods. Others include egg to froglet, egg brooding, and tadpoles in the mother's stomach

by William E. Duellman

In 1758 the German naturalist A. J. Rösel von Rosenhoft published what was then the definitive study of frog development. His magnum opus, with its detailed descriptions and meticulously executed drawings, provided the world with a close-up view of a frog's life cycle—from egg to tadpole to four-legged adult.

Yet Rosenhoft's work, which was based entirely on the common European frog, *Rana esculenta,* is now known to describe only one of many life-history strategies exhibited by frogs as a group. Among herpetologists (those who study frogs and their relatives), including myself, it is increasingly apparent that many species do not conform to the life cycle of the common European frog. Although some frogs adhere to its stereotypical mode of development (depositing thousands of small eggs that become free-swimming tadpoles), many more have evolved reproductive strategies that border on the improbable. Indeed, in reflecting on my 40 years as a field biologist, two thoughts come to mind. The first is frogs are never-ending sources of fascination; the second is they adhere to few conventions.

Frogs and toads (hereafter referred to collectively as frogs) are members of the amphibian order Anura, a term that means tailless in Greek and appropriately describes the adult form. Over the course of the 200 million years since anurans first appeared in the fossil record, they have differentiated into some 24 families and more than 3,800 species. The precise number of species is unknown because new frogs are discovered at a remarkable rate: several dozen are formally described every year. Another indicator of the group's success is the degree to which its members have dispersed across the planet. Today frogs inhabit all continents except Antarctica and have adapted to a breadth of environments, including deserts, forests and grasslands; some species even live at altitudes greater than 5,000 meters (19,685 feet) in the Himalayan and Andean mountains.

Not surprisingly, though, the group is most abundant in tropical regions. One of the richest localities I know, for instance, is a small outpost in Ecuador's Amazon basin called Santa Cecilia, where my colleagues and I once collected 56 different frogs in a single evening. A total of 81 species actually live there (precisely the number of frog species in the entire U.S.). Although the diversity of frogs at Santa Cecilia is unusually high, 40 or more species are commonly found in tracts of rain forest no bigger than two square kilometers.

But the frogs of Santa Cecilia are noteworthy for another reason: they epitomize a key trend in frog evolution, notably the movement away from water, toward land. Of the 81 species we found there, 35 are terrestrial during one or more developmental stages. Yet their proclivity for life on land belies a fundamental requirement of the group: amphibian eggs—like those of fishes—must be kept continuously moist because the semipermeable egg membrane offers little protection against desiccation. Moreover, virtually all males (like fishes) lack an intromittent organ, or penis, and thus must inseminate the eggs externally, after they pass from the female's cloaca (the combined reproductive and excretory opening).

Such a feat is accomplished when the male grasps the female (generally

MATING FROGS of the genus *Centrolene* were photographed in Colombia. Their popular name, which derives from their transparency, is glass frogs; at the right the transparency makes the female's eggs visible. The eggs (*below*) develop on leaves above streams.

WILLIAM E. DUELLMAN took the opportunity to study frogs when he was a student at the University of Michigan. Through this work he acquired a lifelong interest in frogs and their reproductive biology. After obtaining his Ph.D. at Michigan, he joined the faculty of the University of Kansas, where he is professor in the department of systematics and ecology, curator of the division of herpetology in the Museum of Natural History and director of the Center for Neotropical Biological Diversity. His field studies have taken him throughout North and South America and to Africa and Australia. His studies have resulted in some 250 publications, including *Hylid Frogs of Middle America* and, with Linda Trueb, *Biology of Amphibians.*

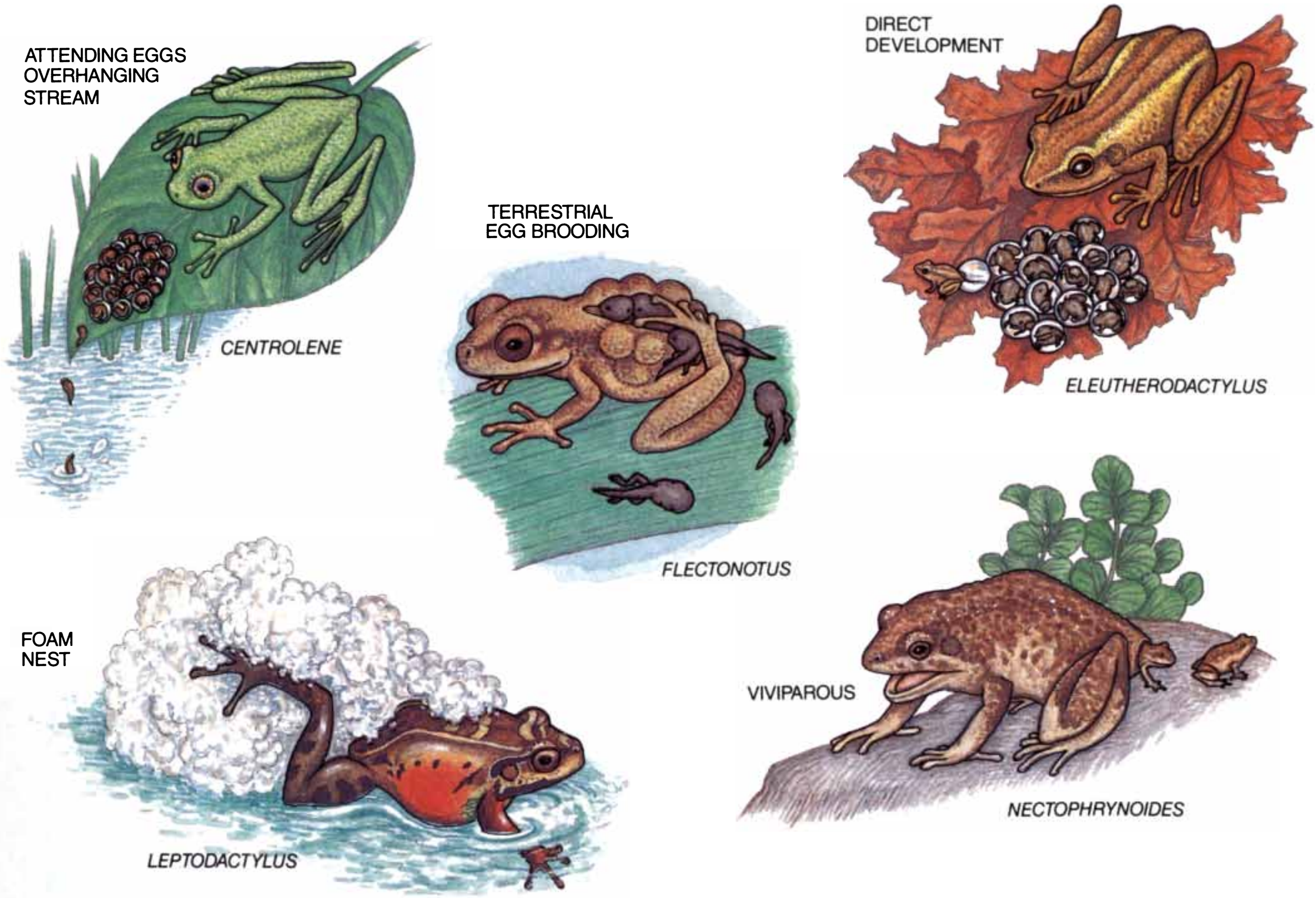
around her armpits) in a mating embrace known as amplexus. In most temperate species, mating takes place in the water, and so the egg clutch (and the aquatic tadpoles that hatch subsequently from it) is never at risk of desiccation. The situation is considerably different on land, however, where rainfall is not guaranteed and humidity levels may fluctuate.

In view of such physiological constraints, how is it frogs have colonized land so successfully? And why, given the harsh conditions that prevail on land, has it been evolutionarily advantageous for them to do so? The answer to the first question has several parts. To begin with, most terrestrial frogs inhabit tropical rain forests, where high humidity levels (close to 90 percent) reduce the risk of dehydration and where ambient air temperatures (25 to 30 degrees Celsius on average) foster rapid growth.

More significant perhaps is that terrestrial frogs adhere to surprisingly few stereotypes; during the 200 million years since their debut, they have evolved numerous—and in many cases startling—adaptations for life on land. Because many of these adaptations have arisen multiple times, they are construed to have a powerful selective advantage. Finally, it should be said that life on land does not necessarily imply life without water; many species lay their eggs in the small, temporary pools that form after a heavy rain or in places such as the spray zone of waterfalls where the humidity is exceptionally high. Thus, they are not terrestrial in the strict sense of the word.

The second question is harder because it is not so easily proved. Nevertheless, most herpetologists believe land offers a reprieve from aquatic predators such as the fishes and insects (and their aquatic larvae) that abound in tropical ponds and streams. The problem is most acute during the egg stage, when mortality may reach as high as 100 percent. Rich in protein and polysaccharides, the frog's membranous egg capsule provides prize fodder for an abundance of predators. According to my colleague Linda Trueb of the University of Kansas, frogs have achieved the best of both worlds: by adapting to land, they have reduced the risk of predation, yet by retaining the ability to live in water, they are able to move as conditions change. The jumping hind legs, of course, enable adults to change habitats quickly, and the muscular tail of some tadpoles allows them to wriggle across the ground.

In fact, many species complete only the egg stage on land, returning to water as tadpoles. In short, adaptive strategies among terrestrial frogs cover a broad gamut of behaviors. One of the simplest and most common calls for laying eggs near water. Indeed, egg clutches are often seen glued to the



REPRODUCTIVE TACTICS of frog species vary widely; some border on the improbable. In this depiction the genus name appears at the lower right of each panel, and the reproductive strategy is summarized at the upper left. One of the most un-

rocks and vegetation that overhang a tropical pond or mountain stream; from this vantage, the tadpoles drop neatly into the water when they hatch. All leaf frogs and glass frogs (so called because the skin on their ventrum, or underside, is transparent) lay their eggs in this fashion. Other species deposit their eggs in shallow depressions adjacent to ponds; when the nest floods, the tadpoles wash directly into the larger body of water.

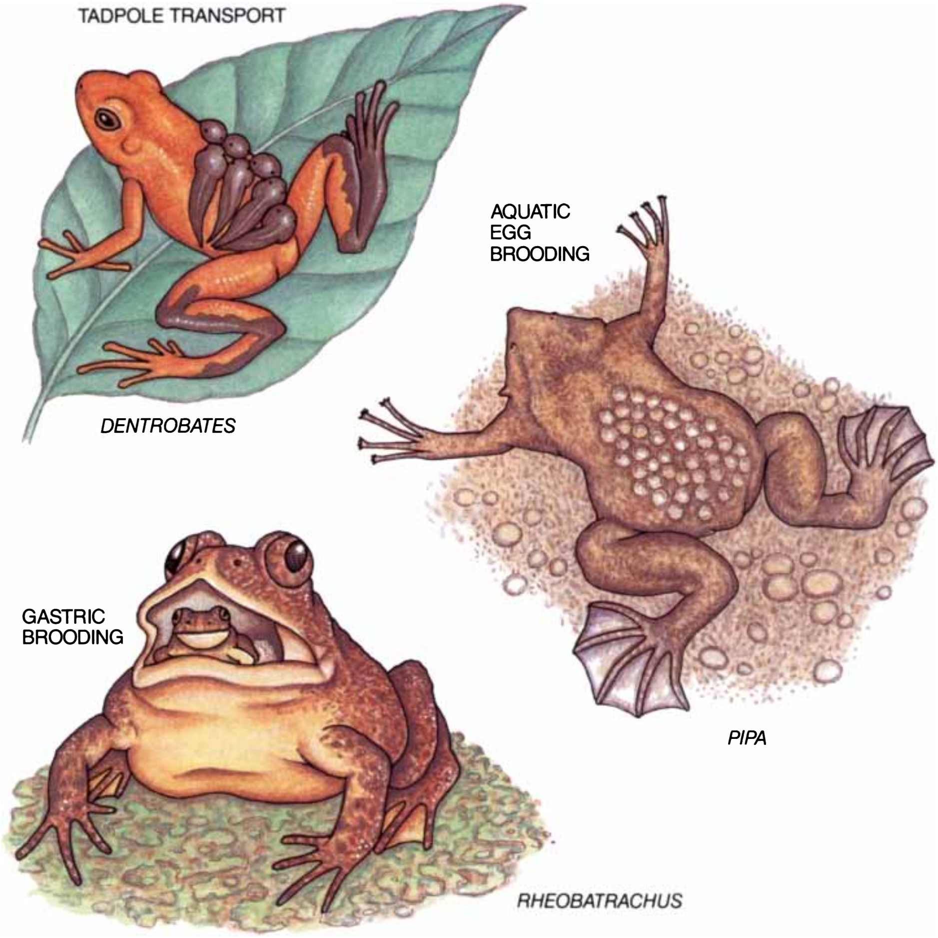
A more sophisticated alternative, which has been documented in at least four families, is to envelop the eggs in protective foam. As the eggs are laid, one or the other parent vigorously kicks its hind legs, mixing together air, sperm, eggs, cloacal secretions and sometimes water. When the frothy mass dries, its exterior turns tacky (sometimes hard) and so protects the eggs from physical harm. At the same time, the egg-filled interior remains liquid for as long as 10 days, which enables the eggs to survive short bouts of drought.

Strategies for laying eggs on land may lead to increased survivorship within the clutch, but they clearly have no effect on mortality during the aquatic tadpole stage. Indeed, predation pressure operating at the tadpole level probably explains another trend in frog biology: the shift away from streams and ponds to small pools or puddles that contain water for only a few days or weeks at a time. Such temporary pools often form in tree holes, at the base of arboreal plants called bromeliads or in small cavities in the ground. Because these water pockets are too ephemeral to support a predator population, they provide a comparatively safe environment in which all stages of the frog's life cycle can unfold. The great risk, of course, is that the water will evaporate before the eggs hatch or the tadpoles complete their development, in which case the entire clutch perishes.

A more derived evolutionary strategy involves prolongation of the egg stage. Rather than producing eggs that hatch into free-living tadpoles, some frogs lay eggs that hatch into four-legged froglets. Known as direct development, the process significantly reduces mortality, although fewer eggs are produced because each embryo needs a large quantity of yolk to see it through the long maturation period. Direct development is associated with nine families of frogs, including the immense tropical American genus *Eleutherodactylus* and related genera; it also occurs in many Australian, Asian and African species. In fact, the elimination of a free-living, feeding tadpole stage is characteristic of nearly 800 frog species—20 percent of the species known.

Although most direct-development frogs live in rain or cloud forests, one of the more interesting but poorly understood genera of this kind is *Breviceps,* the African rain frog. As the name suggests, the animal comes above ground only during heavy downpours. Although much about the life history of this elusive animal remains undocumented, the adults are known to form pairs during the breeding season. Adults emerge from their underground burrows and absorb rainwater through the skin, thus replenishing fluid in their bladder and other body tissues.

Because of his comparatively diminu-



usual strategies is found in the Australian genus *Rheobatrachus:* after the male has fertilized the eggs, the female swallows them and broods them in her stomach.

tive size, the male cannot achieve amplexus with the larger, rotund female but instead glues himself to her back. With the male riding on her back, the female burrows into the ground and proceeds to lay eggs that are then fertilized by her amplectic partner. Periodically the female wets the eggs with water from her bladder, keeping them moist until they hatch as froglets. Most species of *Breviceps* are true desert dwellers, living in the arid regions of sub-Saharan Africa, and one (*B. macrops*) inhabits coastal sand dunes, on the Atlantic coast of South Africa, where there is no fresh water.

Some species undergo a form of highly specialized direct development known as viviparity. Literally meaning live birth, viviparity occurs when the eggs are retained inside the female, hatching as advanced tadpoles or froglets after a period of incubation more reminiscent of mammals than of amphibians. Although energetically taxing for the female, viviparity protects the young against terrestrial predators and profoundly influences the chances of survival.

Viviparity has been documented in only five species of frogs (of those, four belong to the same genus), but the number is bound to increase as biologists learn more about the life history of tropical anurans. In three—*Eleutherodactylus jasperi* from Puerto Rico and *Nectophrynoides tornieri* and *N. viviparus* from Africa—the tadpoles rely solely on egg yolk for their development and so are considered ovoviviparous (meaning live birth from eggs). True viviparity, in which the female supplements yolk with secretions from her oviduct, giving birth to newly metamorphosed froglets, is known to occur in only two African species: *N. liberiensis* and *N. occidentalis.*

Parental care, the range of behaviors exhibited by individuals protecting their offspring, is well documented among birds and mammals, but it is also unexpectedly common among frogs. As with other animals, parental care among frogs typically involves one or both parents. Some species display only rudimentary behavior, as in the case of the African hairy frog (so named because males grow long, hairlike extensions of the epidermis when breeding), which simply sits on its egg clutches in water, but by doing so protects them from aquatic predators such as insect larvae and fish. Or parental care may be more energetically costly, as in the case of the small African toad, *Nectophryne afra,* the male of which stays with its eggs throughout their development, kicking his feet in the water near the eggs and thus increasing oxygen flow through the membranes. Among more specialized frogs, care ranges from retention of the young before they hatch (as in viviparous species) to the transportation, feeding and incubation of the young after they hatch.

Males act as guardians of the offspring in some species; in others the females do. In only a few are both sexes known to cooperate as providers for their young. Peter Weygoldt of Albert-Ludwigs University in Germany studied a captive laboratory population and reported seeing a most unusual form of cooperation among males and females of *Dendrobates pumilio,* a small dart-poison frog from Costa Rica. In this species the two sexes divide parenting duties: males guard the eggs for the first 10 to 12 days until they hatch; thereafter females assume care for the young. The females begin by transporting each newborn tadpole to a bromeliad, at the base of which a small pool of water has collected. Although protected from desiccation and predation, the tadpole has no food supply and therefore relies entirely on its mother for nutrition during the six- to eight-week period to metamorphosis. The mother accommodates her offsprings' needs by returning periodically to each bromeliad to deposit an unfertilized egg, which provides the tadpole with essential proteins and carbohydrates [see "Dart-Poison Frogs," by Charles W. Myers and John W. Daly; SCIENTIFIC AMERICAN, February 1983].

Progressive provisioning of this type has subsequently been reported in other dart-poison frogs, both in the laboratory and in the field. It has also been observed in the large Jamaican tree frog, *Osteopilus brunneus,* which like *Dendrobates* deposits its eggs in bromeliads. Leaf frogs (so named because many species encase their eggs in leaves) have evolved a distinctly different strategy. At the time they deposit the egg clutch, female leaf frogs also lay many eggless, water-filled capsules. Although the capsules have no nutritive value, they are a vital source of water, which reaches the eggs by osmosis.

In some species the task of parental

FOAM NEST of the gray tree frog, genus *Chiromantis,* is visible below the two mating pairs in the photograph at the left. The foam serves to protect the eggs. At a later stage (*above*), tadpoles are seen dropping from the foam nest into the water below. The photographs were made in South Africa.

care falls to the males, who aggressively defend their eggs against a variety of predators. The aptly named gladiator frogs, for instance, frequently engage in violent territorial disputes with other males, wielding needlelike spines at the base of the thumb to slash the skin and eardrums of any intruder. In this species the males invest considerable time and energy in their offspring, initially by forming nests in which the fertilized eggs will be laid and subsequently by staying with the eggs until they hatch.

MOUTH BREEDING by the adult male is part of the reproductive strategy of Darwin's frog, *Rhinoderma darwinii*. The male takes the newly hatched tadpoles into his mouth; they move into his vocal sacs and stay there for several weeks until their development is complete. Developed tadpoles have just emerged from the mouth of this male, photographed in Chile.

Males create the nests by scooping mud from banks adjacent to slow-moving streams or ponds, thus forming shallow basins that fill with water. The resulting predator-free site has the advantage of being close to a permanent body of water but is nonetheless risky because the eggs, which are deposited as a thin film on the water's surface, are vulnerable to attack. Any rupture in the egg film, brought about, say, by the leap of another frog or the aerial dive of a dragonfly, will cause the eggs to sink to the bottom, where they die from lack of oxygen. Consequently, males vigorously defend their eggs and rush to head off any intruder.

Other species forcefully defend their calling sites (which may include egg deposition sites). In Tandayapa, Ecuador, for example, I have seen *Centrolene* males spring into action when approached by conspecifics (males of the same species). Thus confronted, they grapple, attempting to dislodge one another from a favored leaf overhanging a stream. Males of the Australian tusked frog, *Adelotus brevis,* also engage in battles. Equipped with long, pointed tusks that protrude from the lower jaw, they jump and snap at intruders.

Not all males are so confrontational. In some species, males carry the eggs with them until they hatch and so have no need to defend a nest site. Such behavior was first observed more than 200 years ago, when males of the midwife toad (appropriately named *Alytes obstetricans*) were found with strings of eggs entwined around their hind legs. Since then, egg transport has been documented in at least eight families of frogs and has been noted in both males and females. It ranges from simple piggybacking (when the eggs or tadpoles adhere to the parent's back) to egg brooding (when the eggs develop in a special pouch somewhat analogous to the marsupial mammal pouch).

Although egg brooding is most common among females, in the minute Australian hip-pocket frog, *Assa darlingtoni*, it is the males who brood the young. Shortly after hatching, the tadpoles wriggle up their father's legs and into special lateral pouches, where they continue their development. In other species, males play a facilitating role during mating, guiding the eggs to the female. In the South American frog *Pipa*, for example, egg laying is accompanied by a truly graceful display of underwater acrobatics: as the swimming pair flips and turns, the male spreads the emerging eggs across his partner's back, where they adhere to her skin. In some species the skin then swells, enveloping the eggs in a protective matrix. After spending about two months thus embedded in the female's back, the tadpoles—or in some species tiny froglets—burst through the skin to become free-living frogs.

One of my favorite examples of egg brooding occurs among South American tree frogs in the family Hylidae. Indeed, because the frogs brood their eggs in a manner analogous to that of mammals, they are often referred to as the marsupial frogs. The analogy stems from the presence of a unique epidermal pocket on the back of the female. As the eggs are extruded and fertilized, they are immediately pushed into the pouch by the male who propels them across the female's back with his feet. Once inside the pouch, the eggs develop special gill-like structures that exchange oxygen and carbon dioxide osmotically with the highly vascularized tissues lining the pouch, enabling the embryos to breathe.

In most species the eggs hatch as froglets, which then drop off the mother's back or crawl from her pouch. But in some species of *Gastrotheca* living at high elevations in the Andes, development terminates at the tadpole stage; the female sits in shallow ponds and releases her tadpoles into the water. The stage at hatching is determined by the amount of yolk present in each egg, which in turn reflects the number of eggs produced per clutch. Species in which the eggs hatch as small tadpoles produce 100 or more ova, each about two millimeters in diameter. Direct-development species may produce only six ova, but each is about 10 millimeters in diameter [see "Marsupial Frogs," by Eugenia M. del Pino; SCIENTIFIC AMERICAN, May 1989].

Whereas the number of eggs tends to reflect clutch size, development time mirrors climate conditions. Frogs throughout the tropics, for example, develop relatively rapidly, sometimes spending only two or three weeks in the tadpole stage, whereas those that live in cool, temperate climates develop much more slowly. One such species is the spotted frog, *Rana pretiosa*, which lives in the cold streams that cascade down the Rocky Mountains. Because the cold water in which the frogs live slows their metabolism, more than one year is needed to produce fully yolked eggs, and females lay eggs only once every two or three years.

Tadpoles, too, metamorphose slowly. Populations of bullfrogs, *Rana catesbeiana*, living in the northern U.S. and Canada, typically spend two years in the tadpole stage; another species, *Ascaphus truei*, needs three years to reach

TRANSPORT OF TADPOLES takes place on the back of a female dart-poison frog, *Dendrobates reticulatus*, in the Amazon rain forest of Peru. This female is carrying two tadpoles side by side. Once she gets them to water, often in a temporary pool formed at the base of the type of arboreal plant called a bromeliad, she will feed them with unfertilized eggs for about six weeks.

MARSUPIAL FROGS get their name from the pouch on the back of the female. As eggs are fertilized, the male pushes them into the pouch, where they develop to the tadpole or froglet stage. Three newly born young are shown with their mother, and another is shown at birth from this female's pouch. She is a marsupial frog of the species *Gastrotheca ovifera* and was photographed in Venezuela.

adulthood. Although bullfrogs are enormously fecund, producing from 10,000 to 20,000 eggs per year, mortality during the long development period is also extremely high, reaching as much as 99 percent.

In arid habitats, development is limited not by temperature but by moisture. Desert-dwelling frogs such as *Breviceps* may breed only on one or two nights a year, when it rains heavily. But once fertilization occurs, growth proceeds rapidly. Spadefoot toads in the southwestern U.S., for example, dispense with the tadpole stage in less than two weeks, a pace deemed necessary in a habitat where water may last for only a few weeks. Frogs in comparable habitats elsewhere in the world show similar patterns.

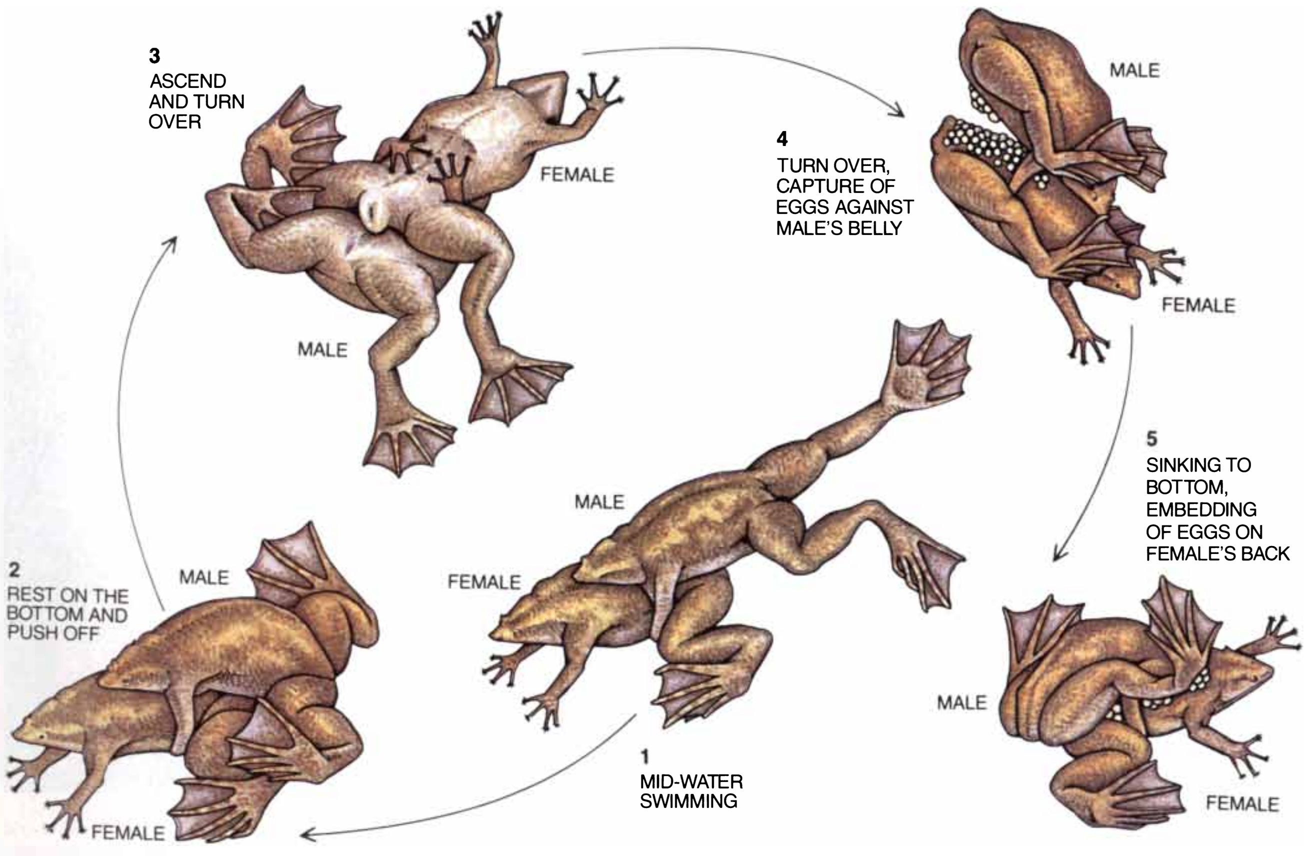
Fascinating as such diverse life-history strategies are, few would be scientifically documented were it not for the willingness of herpetologists to spend countless hours observing frogs under natural field conditions. Indeed, I am often asked why I choose to spend so much time wandering through the world's jungles at night looking for snakes, frogs and other cold-blooded vertebrates. Yet the answer is obvious, I think, to any field biologist: the natural world offers an endless cornucopia of wonderment and surprise. My colleagues and I never know, for example, when something unique will suddenly appear in the beam of our headlamps, even though we might be walking along a trail that we have traversed a dozen nights in succession. Perhaps we will find a species we have never seen before; perhaps we will come across a species no one has ever seen before.

One of the most thrilling moments of my life, in fact, took place during an expedition to southern Chile in 1975. One day I came across Darwin's frog, *Rhinoderma darwinii,* just as it was giving "birth." Although the frog is small and rather drab, it has evolved a mode of development unlike any ever described. Adult males take the newly hatched tadpoles into their mouth, and from there the young migrate to the vocal sacs. They remain for several weeks until growth is complete and then emerge from the father's mouth.

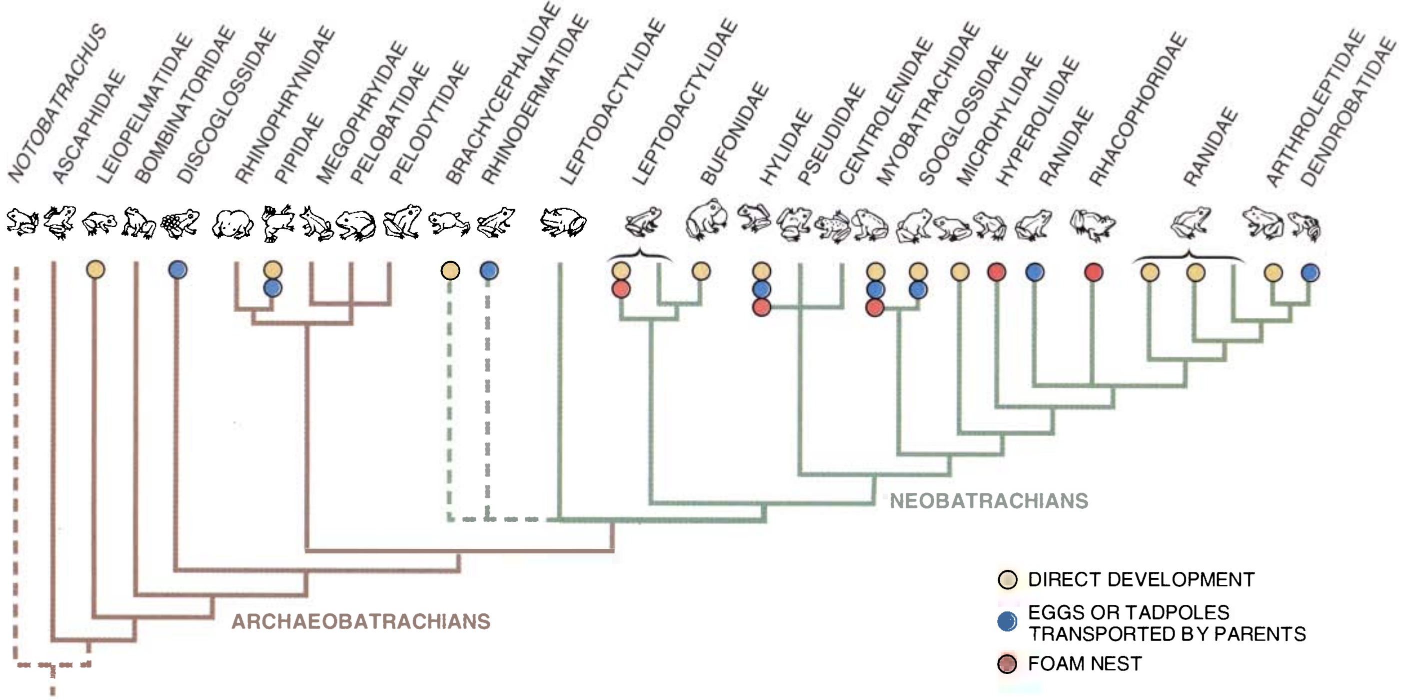
I was aware that *Rhinoderma* engaged in such behavior but nonetheless was excited beyond words to see it finally for myself. In fact, soon after witnessing a succession of froglets parade from their father's mouth, I dashed a letter off to Michael J. Tyler, a colleague and fellow frog enthusiast at the University of Adelaide in Australia. Imagine my chagrin a few weeks later when I received a reply that put my small announcement to shame. Tyler and his colleagues C. J. Corben and G. J. Ingram had recently observed a previously unknown yet more bizarre form of parental care in an Australian frog, *Rheobatrachus silus.* In that species, it is the females who swallow the eggs after fertilization and then brood them in the stomach. A photocopy of the paper detailing the discovery was inscribed simply, "Touché! Mike." My small observation had clearly been overshadowed.

Gastric brooding, as the phenomenon is called, ranks as one of the strangest examples of animal reproduction known. Having finally observed these frogs for myself in 1979 while visiting southern Queensland, I can truthfully say that *Rheobatrachus* has evolved a reproductive mode unlike any other frogs. Who would believe that tadpoles could spend six weeks in their mother's stomach without being digested? How is such a feat possible?

To begin with, nurturing females



MATING PROCEDURE of the South American frog *Pipa carvalhoi* gives rise to graceful underwater acrobatics. With the male riding on the female, the pair flips and turns, and the male spreads the emerging eggs across the female's back. There they adhere and become embedded in a swelling of the skin, hatching two months later as tadpoles or tiny froglets.



MAJOR FROG FAMILIES are classified according to their reproductive strategy (*colored circles*) and family relationships (*colored lines*). The broken lines in the diagram of family relationships represent associations that are questionable.

cease to feed during the breeding period. Tyler and his colleagues have shown that production of hydrochloric acid and pepsin are halted in the stomach by the hormonelike substance prostaglandin E_2, which is secreted first by the egg capsules and then by the tadpoles. With the shutting down of normal functioning in this way, the stomach is transformed from a digestive organ into a protective, gestational sac.

The eggs, which range in number from 21 to 26, are relatively large—about five millimeters in diameter—and therefore rich in yolk. Consequently, the tadpoles do not need an external source of nutrition but feed exclusively on yolk throughout their six-week developmental period. During birth, the female's esophagus dilates in a manner that recalls the birth of mammalian young through the vaginal canal, and the young froglets are propelled from her mouth. Within a few days after expulsion of the young, the stomach begins to function again as a digestive organ, and the frog resumes feeding.

In 1984 a second species of gastric brooder, *Rheobatrachus vitellinus,* was found farther north in Queensland. Sadly, intensive searching during the past few years has failed to turn up individuals of either species, leading my colleagues and me to conclude that both species are now extinct. The loss is especially unfortunate because preliminary studies by Tyler and his associates indicate that prostaglandin extracts from these species might have been effective in the treatment of human stomach ulcers.

After the discovery of gastric brooding and tadpole feeding, I would not be surprised by anything. But why is it that frogs demonstrate such reproductive plasticity when other vertebrates do not? All birds, for example, lay shelled eggs, and nearly all mammals bear living young that are then suckled by the female. Yet frogs conform to few if any reproductive stereotypes. Moreover, it is evident that their reproductive tactics are not the result of simple mutation but the outcome of complex physiological, morphological and behavioral interactions.

The fact that some members of the three amphibian groups (caecilians, salamanders and frogs), as well as fishes, deposit their eggs in water suggests aquatic egg laying is a primitive (generalized) trait. Other strategies, which have evolved independently in many lineages of frogs, therefore can be viewed as evolutionarily derived. One conclusion is certain: frogs, which seem distinctly unsuited for a terrestrial existence, have adapted in extraordinary ways to life on land and in doing so have managed to colonize much of the planet.

Is it not therefore ironic and terribly unfortunate that frogs, having radiated so successfully, are now experiencing extinction at unprecedented rates, victims of human inability to preserve the natural world? Within the past five years alone, an untold number of species have been lost, many to habitat destruction. Although the decline and extinction of anurans is not well understood, a number of biologists (myself included) think the cumulative effects of air and water pollution, acid rain and increased radiation brought about by the loss of stratospheric ozone are taking their toll. Regrettably, the loss of so many species not only affects the overall stability of ecosystems but brings to an end evolutionary lineages that have survived for millions of years. The magnitude of such loss is immeasurable. Our only hope is that people and their governments will recognize the consequences of biological extinction and take steps to prevent it before too much of our global gene pool is gone.

FURTHER READING

THE GASTRIC BROODING FROG. Edited by Michael J. Tyler. Croom Helm, 1983.

BIOLOGY OF AMPHIBIANS. William E. Duellman and Linda Trueb. McGraw-Hill Book Company, 1986.

ALTERNATIVE LIFE-HISTORY STYLES IN ANURAN AMPHIBIANS: EVOLUTIONARY AND ECOLOGICAL IMPLICATIONS. William E. Duellman in *Alternative Life History Styles of Animals.* Edited by Michael N. Bruton. Kluwer Academic Publishers, 1989.

GUILDS OF ANURAN LARVAE: RELATIONSHIPS AMONG DEVELOPMENTAL MODES, MORPHOLOGIES, AND HABITATS. Ronald Altig and Gail F. Johnston in *Herpetological Monographs,* No. 3, pages 81–109; June 22, 1989.

The Last Stone Ax Makers

A chance encounter with a people living in New Guinea's highlands has given archaeologists a priceless opportunity to model prehistoric technology

by Nicholas Toth, Desmond Clark and Giancarlo Ligabue

Only 10,000 years ago all human societies made and used stone implements. In modern times the relentless advance of more complex technologies has left but few remnants of that primeval world, and even these will surely disappear before this century is out. Anthropologists must therefore hurry to study contemporary Stone Age craftsmen for clues they may provide about our early ancestors.

The study of modern humans from an archaeological perspective is called ethnoarchaeology. Unfortunately, it comes on the scene very late in the day. The scholars of the 19th century often paid more heed to such artisans as Britain's gunflint makers than to the contemporary stone knappers of the Americas, southern Africa, Australia, New Guinea and the Pacific Islands.

We sought to help fill this gap in 1990 by visiting a group of horticulturalists of Langda village, in the central mountains of Irian Jaya, western New Guinea. These people had made and traded stone axes in almost complete isolation from the outside world until 1984, when Gunter Konrad, a German urologist, and one of us (Ligabue) met them while on an expedition to the area. Because the later establishment of a Christian mission and the encroachment of a cash economy had begun to transform the traditional culture, the Ligabue Research and Study Center funded the expedition explicitly for the purpose of studying the stone technology. This report is based on that study.

NICHOLAS TOTH, DESMOND CLARK and GIANCARLO LIGABUE ventured into the highlands of New Guinea to study the last known makers of stone tools. Toth is associate professor of anthropology at the University of Indiana, where he also co-directs the Center for Research into the Anthropological Foundations of Technology (CRAFT). Clark is emeritus professor of anthropology at the University of California, Berkeley. Ligabue, a businessman and anthropologist, is the director of the Ligabue Research and Study Center in Venice, Italy.

The Langda ax makers live some 2,000 meters (about a mile) up in the southern slopes of the central cordillera of New Guinea, an altitude that makes for a generally cold and damp climate. Clouds often shroud the village, but when they clear, a superb vista of snow-capped peaks appears. We saw the peaks at their most awe-inspiring while flying by helicopter at a level halfway up the V-shaped walls of a river valley that runs from Wamena, 200 kilometers (120 miles) to the northwest.

The villagers call themselves the Kim-Yal; their language, which belongs to one of the many Papuan families, they call Uni. The average Kim-Yal man stands about four feet, six inches tall, far shorter than populations who live at lower altitudes, such as the Dani of the central highlands and the Asmat of the southwest coast. It is not yet clear whether the highlanders' short stature stems from an adaptation to local conditions caused by natural selection or merely from nutritional factors.

The stone ax makers cultivate sweet potato and taro root and raise pigs and chickens, a diet they supplement with wild plants and small game. They grow their crops in fields, usually on steeply terraced slopes of the river valleys. To clear a new field, they must often fell trees and chop out the residual roots. Historically they have done such work with stone axes, although these implements are slowly yielding to imported metal axes.

The Kim-Yal form open village clusters in which about 200 people occupy from 10 to 15 huts. Each hut typically contains a man, his wives and their children. Adult males may also choose to reside part of the time in a communal hut that serves as a kind of men's club. The residence pattern is patrilocal: men generally stay in the village in which they were born, whereas women marry men from different villages. There are 10 village clusters in the Langda area, most of them situated on a plateau or at the top of a mountain ridge.

The craft of ax making confers high status to the male specialists who practice it and to their sons, who traditionally learn the work in lengthy apprenticeships. Today every man at Langda owns at least two stone axes and each woman at least one. Most children older than five years also tote an ax to the fields, where they work alongside their elders.

An archaeologist would normally classify these implements as adzes, because in side view their ground edges assume an asymmetric, plano-convex shape rather than the symmetric shape typical of axes. Moreover, they are hafted with the working edge at right angles to their handles, whereas ax edges generally lie in the same plane as the handle. But we call them axes because they are used to chop wood and fell trees and because the literature has generally classified such implements according to their use rather than their design. In some other groups, adzes are commonly used not to chop down trees but to shape wood.

An ax begins as a large piece of unflawed stone of a kind that will fracture predictably and carry an edge. Best of all are the large boulders of fine-grained, blue-gray andesite lava that are carried from their volcanic source by the Hei River, which flows some 800 meters below the Langda village. The ax makers descend into the river valley along steep trails that wind through forest and field to the stream bed, a walk that normally takes about 60 minutes going down and about 90 minutes coming back up. They seek not only boulders but also smaller lava stones to use as hammers.

The worker typically holds in both hands a large stone hammer, some 25 centimeters (10 inches) in diameter, and then swings it against the edge of a lava boulder core. If he strikes it correctly, he detaches large flakes from the core. We were quite surprised to find that workers usually swing the hammerstone through their legs, much like the motion of an American football center. This technique, to our knowledge, has never before been described ethnographically; we believe it could have been used in antiquity for detaching such large flakes.

Suitable flakes may also be produced by throwing one boulder hammer against another boulder core or by kindling a fire alongside a boulder to initiate fractures. On rare occasions a properly proportioned, water-worn cobble will also serve as a blank. A typical blank measures about 24 centimeters long, 14 centimeters wide and seven centimeters deep (about 10 by six by three inches).

Half an hour of quarrying will produce, on average, from five to 10 blanks for ax manufacture. Hundreds or even thousands of smaller fragments, used-up boulder cores and hammerstones also come out of the quarrying operation. The workers discard such detritus by the riverside, where the next flood will sweep it away.

Paleolithic archaeologists would recognize the three major manufacturing steps that the Langda stone workers themselves regard as integral. Each stage has its characteristic ax morphology and waste products, which tend to be discarded in different places on the landscape. In the first stage the craftsman uses a hammerstone from 10 to 12 centimeters (four to five inches) long to remove large flakes from both faces of the lava blank, leaving bold scars on the flake surface. Normally he does this work at the riverside quarry. The resultant biface is quite crude: a jagged outline and an irregular, sinuous edge run around its entire circumference. The "rough-outs" are then reduced by flaking so as to minimize carrying weight—especially if a hidden defect should cause the stone to break.

The second stage of ax manufacture may take place back at the village, although it is often carried out in a field hut situated on a terrace about 100 meters above the riverbank. There the ax maker can work in the shade of the hut, build a fire if it is cold and perhaps roast sweet potatoes or game as he continues flaking. He normally does so while sitting on his haunches, holding the hammerstone in his right hand and the rough-out in the left. (All of the stone workers we observed were right-handed.)

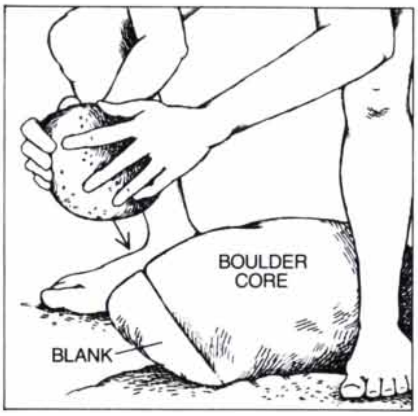
The ax maker keeps several stone hammers of different sizes, shapes and degrees of hardness, and he switches from one to another as flaking proceeds. As our earlier reconstructions of

STONE TOOL OF THE KIM-YAL resembles an adze but works as an ax, cutting rather than shaping wood. The T bar measures 30 centimeters (12 inches) long.

Making a Stone Ax

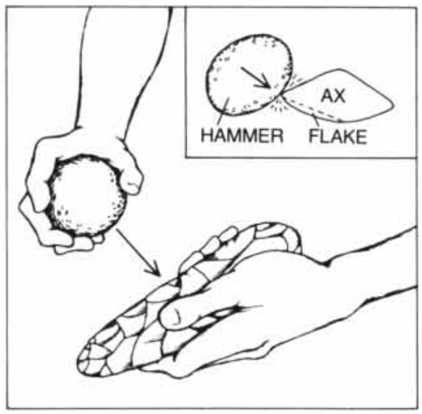
Working the riverside quarry

Flaking the stone blanks



Quarrying

Rough flaking



Fine flaking

stone-knapping techniques had predicted, the craftsman tends to use ever smaller hammerstones as his ax approaches its final shape. But we did not foresee that such hard hammerstones would be used for fine work. Most archaeologists would have labeled this retouch flaking as "soft hammer"—that which is done with a piece of wood, bone or other soft material.

Each flake sets up further flaking by preparing what is called the striking platform [*see inset in illustration above, right*]. The object is to steepen the edge angles by lightly flaking, crushing and abrading them with the hammerstone. If all goes well, flat, thin flakes can then be removed from the ax surface. In this way, the overall shape of the ax becomes progressively thinner, narrower and more symmetric.

A man may work alone or sit with other ax makers, creating an occasion for talking, singing and showing off one's handiwork. Members of such groups invariably sit in a line and face in the same direction so that sharp waste flakes will not fly into anyone's face. These flakes, which become smaller and proportionately thinner as flaking proceeds, are usually found outside the perimeter of the field hut. Sometimes, however, they appear around a hut in the village, evidence that some rough-outs are carried there before being worked to any great extent.

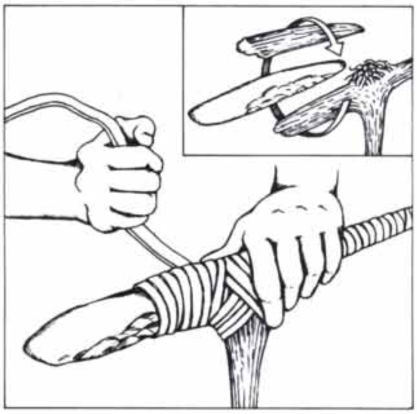
Women from the village, most often relatives of the stone workers, typically transport rough-outs from the valley to the village. The women protect themselves from cuts and scratches by wrapping the axes in vegetation. They put the implements in carrying nets that they sling over their backs, supporting the load on a strap that loops around the forehead.

The final reduction of the ax normally takes place back at Langda, often next to the hut of an ax maker. Here the worker prepares his striking platform with extreme delicacy so as to detach tiny trimming flakes; he also abrades the surface. The final product of flaking is a long, thin ax with nearly parallel

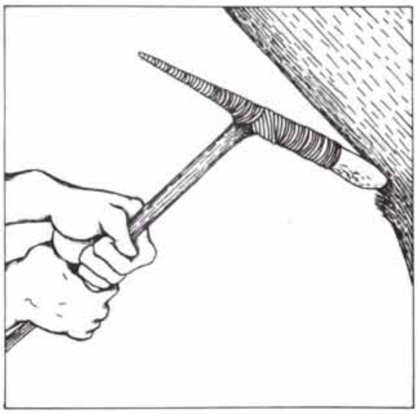
Finishing the flaked ax

Binding the ax to the haft

Grinding



Hafting



Chopping

sides, a triangular cross section and a bit end having a slightly convex outline and an edge beveled at about 50 degrees. A typical stage-three ax measures about 20 by four by three centimeters, but substantially larger and smaller ones are occasionally made as well. The time spent flaking—from quarry blank to stage three—varies from about 30 to 50 minutes. Between 10 and 20 percent of the blanks tend to be lost to breakage.

The waste flakes produced in Langda village are much smaller than those produced in the earlier stages. The complete sequence of ax manufacture, from the quarry blank to the finished flaked ax, in a typical case produced a total of 225 flakes and fragments larger than two centimeters in length (and many more smaller bits, grading to microscopic size). Of these, approximately 100 could be expected to end up at Langda village, dropped there during the later stages of flaking. They are generally strewn in a circular or oval-shaped pattern, often near an ax maker's hut, with the vast majority of fragments concentrated in an area of one square meter. Such a concentration would probably be dispersed over time by humans or pigs walking over the ground. Ultimately, the flaking debris would be trampled underground, at times when the ground is soft and muddy.

After the final flaking, the craftsman grinds the working end of the ax against a wetted slab of fine-grained sandstone. These grinding stones are highly prized and come from a distant river valley some six hours' walk away. They are stationed close to a water source in the village. The worker squats down with the grindstone in front of him, adds water to the worn surface on the stone and then grinds the ax against the stone. The fine sandstone and lava particles mix with the water to form a paste that serves as the grinding medium. The axhead is inspected at regular intervals and more water added, with special attention being paid to the working edge. An hour or so of grinding turns out an axhead with a polished bit and an extremely regular edge.

One might wonder why the Kim-Yal would grind flaked edges, which are already as sharp as they can be. The reason is simple: ground tools hold their edges longer and lend themselves more readily to resharpening. Grinding thus makes sense mainly for woodworking tools and other implements that are used intensively over long periods. It is no accident that the so-called age of polished stone began with the development of agriculture, when people began clearing forests in order to plant their crops.

Aesthetic considerations also motivate the Kim-Yal grinders, otherwise they might not bother to polish more than a few millimeters from the leading edge. In fact, they polish nearly to the point where the axhead disappears into the binding of the haft. Prehistoric craftsmen in New Guinea and other parts of the world often went further, polishing the entire surface of the tool. Additional evidence of aesthetic intent appears in the Kim-Yal's use of red ochre and other locally occurring pigments to fill in the unground depression spots on their axes. Such use of pigments for aesthetic or symbolic purposes has been widely documented in modern nonindustrial societies, as well as in prehistoric cultures over the past 35,000 years.

Hafting allows a tool to be used at length and to greater effect. The workers haft the finished axhead onto a T-shaped wooden handle, made from a flattened tree trunk or branch from which a thinner branch projects at an angle of approximately 45 degrees [*see illustration on preceding page*]. The stone rests against several strips of wood and frayed bark, which help to absorb shock. The worker then lashes the composite tightly with split vine, creating a design that distributes the force of a chopping blow throughout a flexible binding. Without this flexibility, the stone would tend to snap across its short axis.

A Langda villager equipped with such an ax can fell a tree as thick as a telephone pole in five to 10 minutes. The villagers also use these axes to cut out roots from a field, to build fences to protect crops from pigs and to make wooden digging sticks, containers and other wooden tools. A well-ground edge can often last for several hours of use without requiring resharpening. When it begins to dull, it can easily be resharpened by rubbing it against a grindstone using water as a lubricant; the axhead generally remains mounted in its haft.

The villagers will use an ax until it breaks or until regrinding has reduced the tool to the point where it can no longer function. Broken-off tips may be lost in the field, but interestingly, most worn-out or broken specimens end up in the village. The ax makers say they "feel sorry" for their handiwork and take pains to bring it home for final discard. Although such personal attachment to tools figures in the *Iliad* and the *Odyssey*, few anthropologists probably have suspected that it might explain distribution patterns in prehistoric sites. Henceforth, we and our colleagues must take such factors into account.

The stone axes carry cultural significance that goes beyond their practical function. Feasting villagers often exchange axes to cement their social bonds, and men give them in payment of the bride price—compensation to a father for the privilege of marrying his daughter. Our informants tell us that in the days before the cash economy intruded, a stone ax could be traded for a pig, the most prized food. Today three stone axes can be traded for one metal ax.

The Langda people trade their axes to other villages that lack suitable stone for such implements. Some six days' journey northward, the people of Bima barter their goods for Langda axes. Southward, the main trading post, at Lukan village on the edge of the highlands, attracts Kopkaka tribespeople from the lowlands, who pay for the axes with bowstaves, vine, fibers for mats, bird-of-paradise feathers or animal pelts. In the past the Langda axes were traded as far away as the southern coast, mainly in exchange for shells. Such specialization of labor as the ax industry requires must, in any case, have depended on long-distance trade. Traces of such far-reaching trade networks become obvious in the prehistoric record only after about 35,000 years ago.

Interestingly, the Langda villagers do not commonly use other stone tools in their daily activities. The exception is a small, flat knife ground from slate or some softer stone. Villagers use these knives—which come in oval, kidney or subtriangular form—to cut off the leaves and stem from a taro root, scrape the outer surface clean and then split the root for cooking. They also use the knives to split the long leaves from which women's skirts are made. For other purposes, however, organic materials work better. We saw men butchering pigs, for example, with razor-sharp pieces of split bamboo. One resharpens such a tool by simply tearing off a strip of bamboo with the thumbnail or teeth. The serviceability of bamboo has been invoked to explain the simplicity of stone tools in East Asian sites that were inhabited by protohumans for many hundreds of thousands of years.

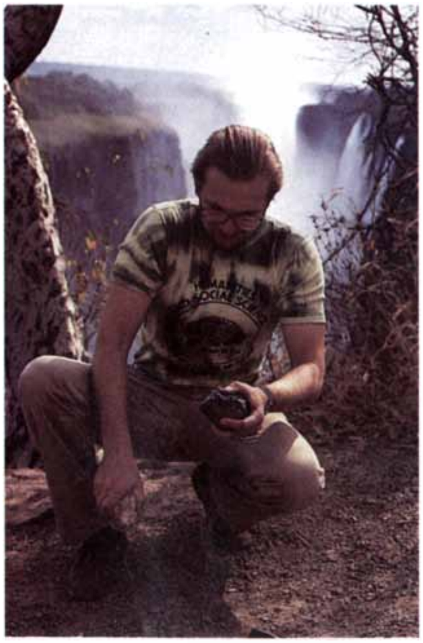
The Langda ax makers provide many other interesting insights to the archaeologist. They show how horticultural-

TWO AXMEN cut through this tree, some 20 centimeters (eight inches) thick, in about three minutes. They do such work to clear fields and get materials for building huts.

ists use stone tools in a natural and economically sensible context. The byproducts of their labor, at each of the three stages of manufacture, may mirror patterns preserved in the archaeological record. Finally, the ax makers are of intense archaeological interest because various stages of their ax manufacture appear to recapitulate the evolution of stone technology.

The first stage—quarrying flake blanks from boulder cores—strikingly recalls the manufacture of the Acheulean hand axes and cleavers of the Early Stone Age. Acheulean technology was invented by *Homo erectus* some 1.7 million years ago, and it continued unchanged into the age of archaic *H. sapiens* some 200,000 years ago [see "The First Technology," by Nicholas Toth; SCIENTIFIC AMERICAN, April 1987]. The second stage is recapitulated in the Langda villagers' roughouts, which resemble the so-called Sangoan and Lupemban core axes of the Middle Stone Age of Equatorial Africa. Unhafted Sangoan tools are believed to have been used to work wood in tropical Africa between 200,000 and 100,000 years ago. The third stage of ground and hafted axes corresponds almost perfectly to the axes associated with early farming communities in both the Old and New World.

The earliest evidence of ground stone axes comes, however, from sites in New Guinea and Australia that date to more than 20,000 years ago. Still, such tools do not appear frequently in the archaeological record until after the rise of farming communities some 10,000 years ago. The Langda axes are clearly one of the last remnants of this Stone Age tradition.



EXPERIMENTAL ARCHAEOLOGISTS: the author (Toth) and Kanzi, a bonobo, flake stone in a project to assess the toolmaking skills of long-extinct hominids.

We do not wish to create the impression, however, that any ethnoarchaeological study that might have been undertaken in modern times could have revealed all the secrets of Early Stone Age technological culture. The Kim-Yal, unlike some of those earlier inventors, are anatomically modern humans possessed of language systems and cultures as complex as anyone else's today. Their modes of operation are almost certainly much more complex than those of protohuman toolmakers.

We suspect, for example, that the Kim-Yal's tendency to "pity" the ax that has outlived its usefulness would not have been seen in many archaic hominids that could probably have mastered the ax-making craft. Anthropologists caution that one should not confuse similarity of performance with identity of mind. One could easily teach modern hunter-gatherers how to fly and land a Boeing 747 airliner. But such might not be the case for an *H. erectus* a million years ago, even one that had been adopted at birth into modern society.

One can try to reconstruct the earliest stone technology in other ways. Experimental archaeologists make and use stone tools with materials and methods that were available to prehistoric people. Primatologists study tool use among chimpanzees in the wild. Recently the two lines of inquiry converged in a fascinating experiment at the Language Research Center in Atlanta, operated by Georgia State University and the Yerkes Regional Primate Research Center.

There Susan Savage-Rumbaugh, Duane Rumbaugh and Rose A. Sevcik collaborated with one of us (Toth) and Toth's co-worker, Kathy Schick, in teaching stonecraft to a bonobo, or pygmy chimpanzee. The bonobo, named Kanzi, has demonstrated the ability to grasp the most basic skills required to detach flakes from stones and to use the flakes as cutting tools [*see illustration above*] and has even taught himself new skills—such as breaking up big stones by throwing them on a tiled floor. Ongoing research also seeks to discover whether Kanzi can impart toolmaking knowledge to other bonobos. Such studies may help workers estimate the limits bounding toolmaking among modern humans, nonhuman primates and protohuman populations.

How much longer will stone ax manufacture and its complex exchange network continue? It is likely that the network will disappear in a few short years as stores providing metal tools are introduced in the highlands and mountains and a cash economy replaces the traditional system of barter. Given that none of the younger members of the tribe is currently apprenticed in this craft, it is likely that most of this skilled ax-making technology will be lost within one or two generations.

We have felt privileged to visit and learn from these friendly peoples, to examine their stone ax exchange system, to observe the relationships between technological processes and their products, and to describe this technology from an archaeological perspective. We hope these observations will provide information that archaeologists will be able to draw on when attempting to decipher the Stone Age past.

FURTHER READING

EXPLORATIONS IN ETHNOARCHAEOLOGY. Edited by Richard A. Gould. The School of American Research, University of New Mexico Press, 1978.

A PREHISTORY OF AUSTRALIA, NEW GUINEA AND SAHUL. J. Peter White and James F. O'Connell. Academic Press, 1983.

QUARRYING IN A TRIBAL SOCIETY. John Burton in *World Archaeology*, Vol. 16, No. 2, pages 234–247; October 1984.

INDONESIA: LA GRANDE DERIVA ETNICA. Edited by Gilda M. Ronzoni. Erizzo Editrice, 1986.

THE FIRST TECHNOLOGY. Nicholas Toth in *Scientific American*, Vol. 256, No. 4, pages 112–121; April 1987.

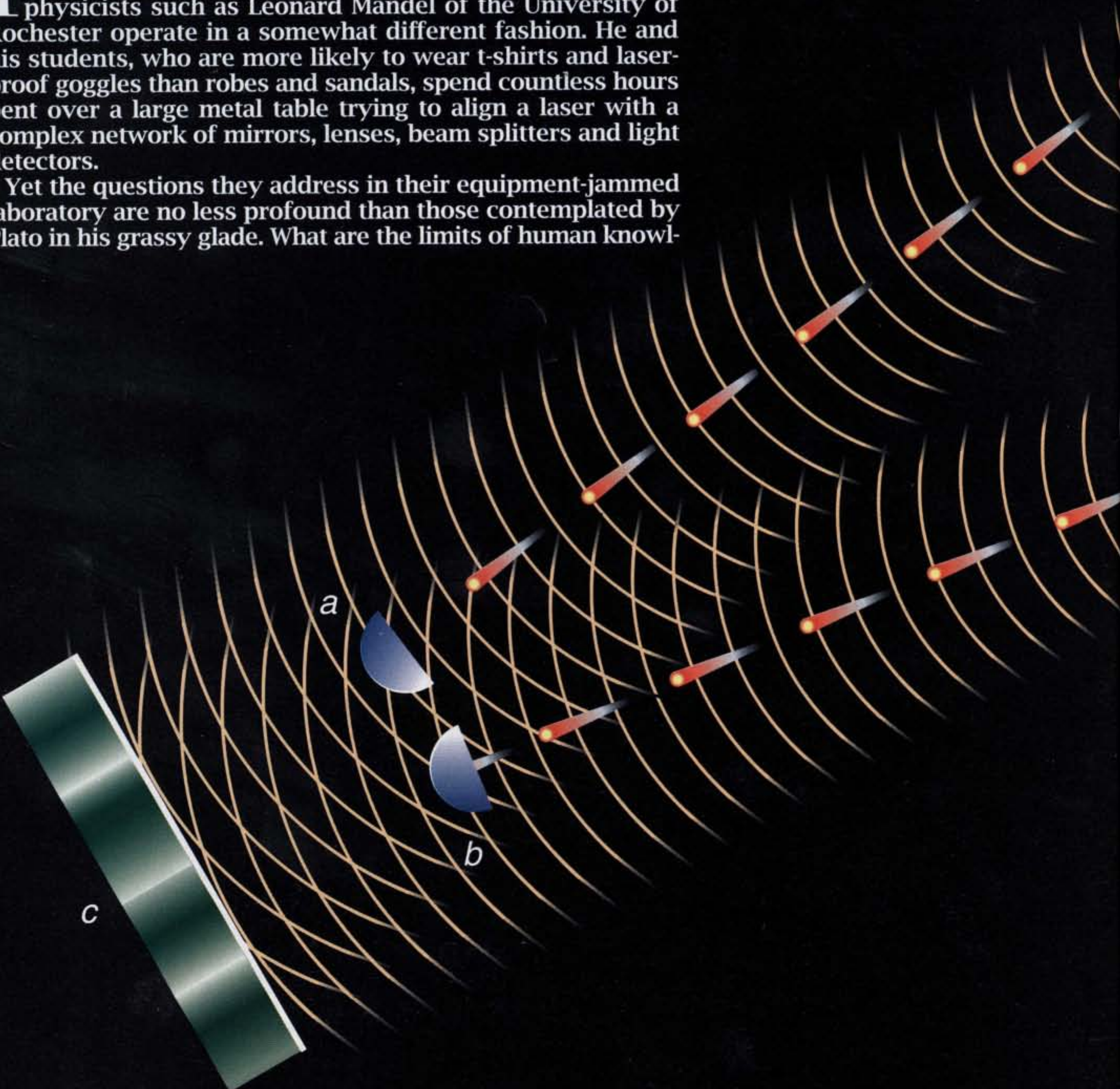
QUANTUM PHILOSOPHY

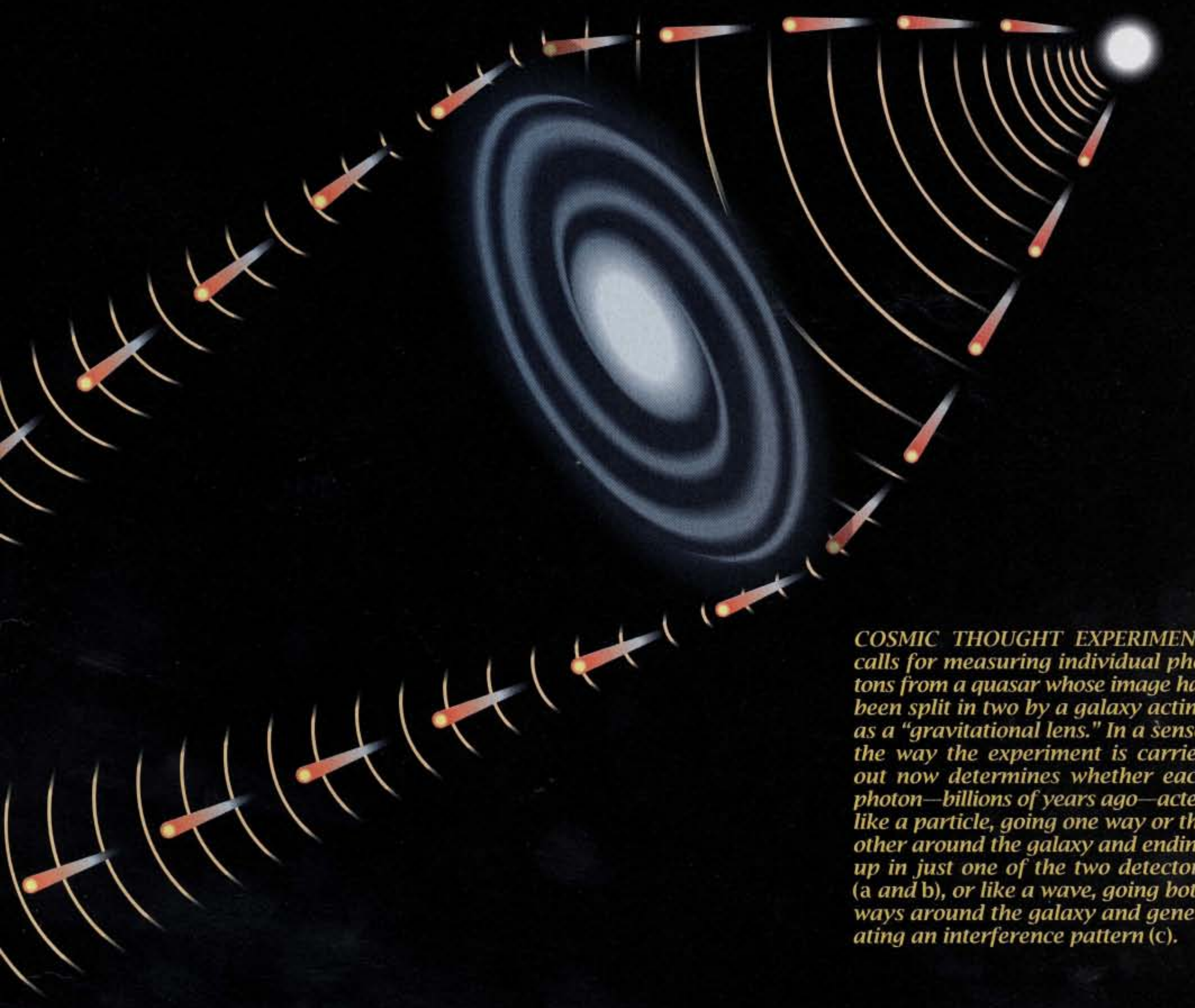
by John Horgan, *senior writer*

New experiments—real and imagined—are probing ever more deeply into the surreal quantum realm.

In ancient Greece, Plato tried to think and talk his way to the truth in extended dialogues with his disciples. Today physicists such as Leonard Mandel of the University of Rochester operate in a somewhat different fashion. He and his students, who are more likely to wear t-shirts and laser-proof goggles than robes and sandals, spend countless hours bent over a large metal table trying to align a laser with a complex network of mirrors, lenses, beam splitters and light detectors.

Yet the questions they address in their equipment-jammed laboratory are no less profound than those contemplated by Plato in his grassy glade. What are the limits of human knowl-





COSMIC THOUGHT EXPERIMENT calls for measuring individual photons from a quasar whose image has been split in two by a galaxy acting as a "gravitational lens." In a sense, the way the experiment is carried out now determines whether each photon—billions of years ago—acted like a particle, going one way or the other around the galaxy and ending up in just one of the two detectors (a *and* b), *or like a wave, going both ways around the galaxy and generating an interference pattern* (c).

edge? Is the physical world shaped in some sense by our perception of it? Is there an element of randomness in the universe, or are all events predetermined?

Mandel, being inclined toward understatement, offers a more modest description of his mission. "We are trying to understand the implications of quantum mechanics," he says. "The subject is very old, but we are still learning."

Indeed, it has been nearly a century since Max Planck proposed that electromagnetic radiation comes in tidy bundles of energy called quanta. Building on this seemingly tenuous supposition, scientists erected what is by far the most successful theory in the history of science. In addition to yielding theories for all the fundamental forces of nature except gravity, quantum mechanics has accounted for such disparate phenomena as the shining of stars and the order of the periodic table. From it have sprung technologies ranging from nuclear reactors to lasers.

Still, quantum theory has deeply disturbing implications. For one, it shattered traditional notions of causality. The elegant equation devised by Erwin Schrödinger in 1926 to describe the unfolding of quantum events offered not certainties, as Newtonian mechanics did, but only an undulating wave of possibilities. Werner Heisenberg's uncertainty principle then showed that our knowledge of nature is fundamentally limited—as soon as we grasp one part, another part slips through our fingers.

The founders of quantum physics wrestled with these issues. Albert Einstein, who in 1905 showed how Planck's electromagnetic quanta, now called photons, could explain the photoelectric effect (in which light striking metal induces an electric current), insisted later that a more detailed, wholly deterministic theory must underlie the vagaries of quantum mechanics. Arguing that "God does not play dice," he designed imaginary "thought" experiments to demonstrate the theory's "unreasonableness." Defenders of the theory such as Niels Bohr, armed with thought experiments of their own, asserted that Einstein's objections reflected an obsolete view of reality. It is not the job of scientists, Bohr chided his friend, "to prescribe to God how He should run the world."

Until recently, the prevailing attitude of most physicists has been utilitarian: if the theory can foretell the performance of a doped gallium arsenide semiconductor, why worry about its epistemological implications? In the past decade or so, however, a growing cadre of researchers has been probing the ghostly underpinnings of their craft. New technologies, some based on the very quantum phenomena that they test, have enabled investigators to carry out experiments Einstein and Bohr could only imagine. These achievements, in turn, have inspired theorists to dream up even more challenging—and sometimes bizarre—tests.

The goal of the quantum truth-seekers is not to build faster computers or communications devices—although that could be an outcome of the research. And few expect to "disprove" a theory that has been confirmed in countless experiments. Instead their goal is to lay bare the curious reality of the quantum realm. "For me, the main purpose of doing experiments is to show people how strange quantum physics is," says Anton Zeilinger of the University of Innsbruck, who is both a theorist and experimentalist. "Most physicists are very naive; most still believe in real waves or particles."

So far the experiments are confirming Einstein's worst fears. Photons, neutrons and even whole atoms act sometimes like waves, sometimes like particles, but they actually have no definite form until they are measured. Measurements, once made, can also be erased, altering the outcome of an experiment that has already occurred. A measurement of one quantum entity can instantaneously influence another far away. This odd behavior can occur not only in the microscopic realm but even in objects large enough to be seen with the naked eye.

These findings have spurred a revival of interest in "interpretations" of quantum mechanics, which attempt to place it in a sensible framework. But the current interpretations seem anything but sensible. Some conjure up multitudes of universes. Others require belief in a logic that allows two contradictory statements to be true. "Einstein said that if quantum mechanics is right, then the world is crazy," says Daniel Greenberger, a theorist at the City College of New York. "Well, Einstein was right. The world is crazy."

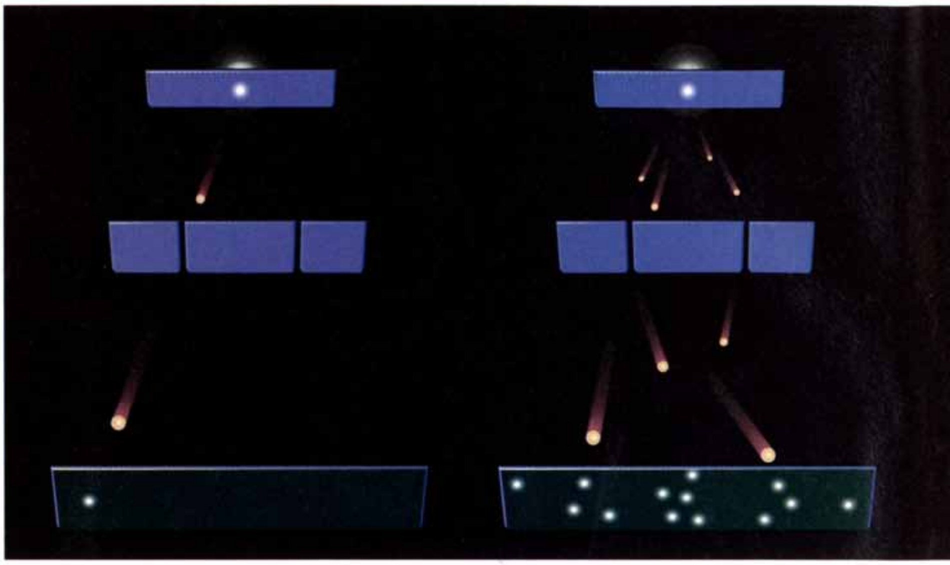
The root cause of this pathology is the schizophrenic personality of quantum phenomena, which act like waves one moment and particles the next. The mystery of wave-particle duality is an old one, at least in the case of light. No less an authority than Newton proposed that light consisted of "corpuscles," but a classic experiment by Thomas Young in the early 1800s convinced most scientists that light was essentially wavelike.

Young aimed a beam of light through a plate containing two narrow slits, illuminating a screen on the other side. If the light consisted of particles, just two bright lines should have appeared on the screen. Instead a series of lines formed. The lines could be explained only by assuming that the light was propagating as waves, which were split into pairs of wavelets by the two-slit apparatus. The pattern on the screen was formed by the overlapping, or interference, of the wavelet pairs. The screen was bright where crests coincided and dark where crests met troughs, canceling each other out.

But more recent two-slit experiments suggest that Newton was also right. Modern photodetectors (which exploit the photoelectric effect explained by Einstein) can show individual photons plinking against the screen behind the slits in a particular spot at a particular time—just like particles. But as the photons continue striking the screen, the interference pattern gradually emerges, a sure sign that each individual photon went through both slits, like a wave.

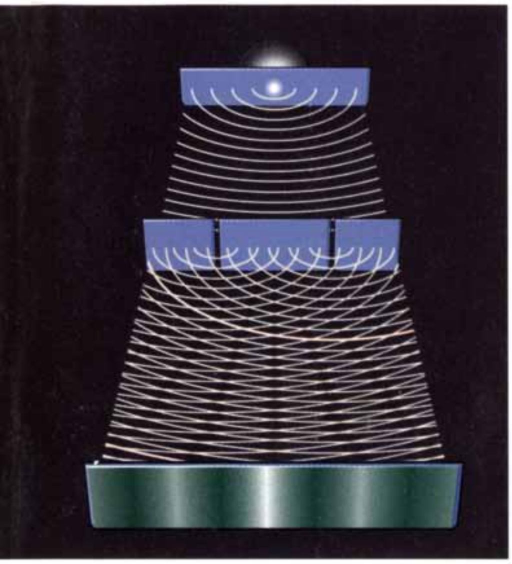
Moreover, if the researcher either leaves just one slit at a time open or moves the detectors close enough to the two slits to determine which path a photon took, the photons go through one slit or the other, and the interference pattern disappears. Photons, it seems, act like waves as long as they are permitted to act like waves, spread out through space with no definite position. But the moment someone asks where the photons are—by determining which slit they went through or making them hit a screen—they abruptly become particles.

Actually, wave-particle duality is even more baffling than this explanation suggests, as John A. Wheeler of Princeton University demonstrated with a thought



Revealing the Split Personality of Light

Two-slit experiments reveal that photons, the quantum entities giving rise to light and other forms of electromagnetic radiation, act both like particles and like waves. A single photon will strike the screen in a particular place, like a par-



ticle (*left*). But as more photons strike the screen, they begin to create an interference pattern (*center*). Such a pattern could occur only if each photon had actually gone through both slits, like a wave (*right*).

experiment he devised in 1980. "Bohr used to say that if you aren't confused by quantum physics, then you haven't really understood it," remarks Wheeler, who studied under Bohr in the 1930s and went on to become one of the most adventurous explorers of the quantum world.

In the two-slit experiments, the physicist's choice of apparatus forces the photon to choose between going through both slits, like a wave, or just one slit, like a particle. But what would happen, Wheeler asked, if the researcher could somehow wait until after the light had passed the two slits before deciding how to observe it?

Five years after Wheeler outlined what he called the delayed-choice experiment, it was carried out independently by groups at the University of Maryland and the University of Munich. They aimed a laser beam not at a plate with two slits but at a beam splitter, a mirror coated with just enough silver to reflect half of the photons impinging on it and let the other half pass through. After diverging at the beam splitter, the two beams were guided back together by mirrors and fed into a detector.

This initial setup provided no way for the investigators to tell whether any individual photon had gone right or left at the beam splitter. Consequently, each photon went both ways, splitting into two wavelets that ended up interfering with each other at the detector.

Then the workers installed a customized crystal called a Pockels cell in the middle of one route. When an electric current was applied to the Pockels cell, it diffracted photons to an auxiliary detector. Otherwise, photons passed through the cell unhindered. A random-signal generator made it possible to turn the cell on or off after the photon had already passed the beam splitter but before it reached the detector, as Wheeler had specified.

When the Pockels-cell detector was switched on, the photon would behave like a particle and travel one route or the other, triggering either the auxiliary detector or the primary detector, but not both at once. If the Pockels-cell detector was off, an interference pattern would appear in the detector at the end of both paths, indicating that the photon had traveled both routes.

To underscore the weirdness of this effect, Wheeler points out that astronomers could perform a delayed-choice experiment on light from quasars, extremely bright, mysterious objects found near the edges of the universe. In place of a beam splitter and mirrors, the experiment requires a gravitational lens, a galaxy or other massive object that splits the light from a quasar and refocuses it in the direction of a distant observer, creating two or more images of the quasar.

Psychic Photons

The astronomer's choice of how to observe photons from the quasar here in the present apparently determines whether each photon took both paths or just one path around the gravitational lens—billions of years ago. As they approached the galactic beam splitter, the photons must have had something like a premonition telling them how to behave in order to satisfy a choice to be made by unborn beings on a still nonexistent planet.

The fallacy giving rise to such speculations, Wheeler explains, is the assumption that a photon had some physical form before the astronomer observed it. Either it was a wave or a particle; either it went both ways around the quasar or only one way. Actually, Wheeler says, quantum phenomena are neither waves nor particles but are intrinsically undefined until the moment they are measured. In a sense, the British philosopher Bishop Berkeley was right when he asserted two centuries ago that "to be is to be perceived."

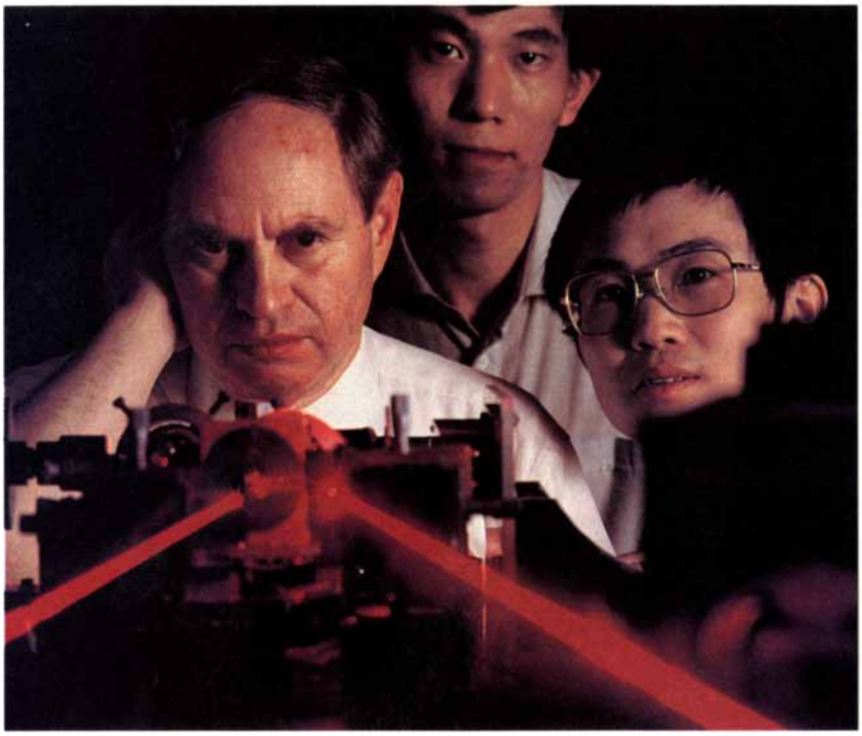
Reflecting on quantum mechanics some 60 years ago, the British physicist Sir Arthur Eddington complained that the theory made as much sense as Lewis Carroll's poem "Jabberwocky," in which "slithy toves did gyre and gimble in the wabe." Unfortunately, the jargon of quantum mechanics is rather less lively. An unobserved quantum entity is said to exist in a "coherent superposition" of all the possible "states" permitted by its "wave function." But as soon as an observer makes a measurement capable of distinguishing between these states, the wave function "collapses," and the entity is forced into a single state.

Yet even this deliberately abstract language contains some misleading implications. One is that measurement requires direct physical intervention. Physicists often explain the uncertainty principle in this way: in measuring the position of a quantum entity, one inevitably knocks it off its course, losing information about its direction and about its phase, the relative position of its crests and troughs.

Most experiments do in fact involve intrusive measurements. For example, blocking one path or the other or moving detectors close to the slits obviously disturbs the photons' passage in the two-slit experiments, as does placing a detector along one route of the delayed-choice experiment. But an experiment done last year by Mandel's team at the University of Rochester shows that a photon can be forced to switch from wavelike to particlelike behavior by something much more subtle than direct intervention.

The experiment relies on a parametric down-converter, an unusual lens that splits a photon of a given energy into two photons whose energy is half as great. Although the device was developed in the 1960s, the Rochester group pioneered its use in tests of quantum mechanics. In the experiment, a laser fires light at a beam splitter. Reflected photons are directed to one down-converter, and transmitted photons go to another down-converter. Each down-converter splits any photon impinging on it into two lower-frequency photons, one called the signal and the other called the idler. The two down-converters are arranged so that the two idler beams merge into a single beam. Mirrors steer the overlapping idlers to one detector and the two signal beams to a separate detector.

This design does not permit an observer to tell which way any single photon went after encountering the beam splitter. Each photon therefore goes both right and left at the beam splitter, like a wave, and passes through both down-converters, producing two signal wavelets and two idler wavelets. The sig-



LEONARD MANDEL (*at left*) and co-workers at the University of Rochester gather around a parametric down-converter, an unusual crystal that converts any photon striking it into two photons with half as much energy. Mandel's group pioneered the use of the device in tests of quantum mechanics.

nal wavelets generate an interference pattern at their detector. The pattern is revealed by gradually lengthening the distance that signals from one down-converter must go to reach the detector. The rate of detection then rises and falls as the crests and troughs of the interfering wavelets shift in relation to each other, going in and out of phase.

Now comes the odd part. The signal photons and the idler photons, once emitted by the down-converters, never again cross paths; they proceed to their respective detectors independently of each other. Nevertheless, simply by blocking the path of one set of idler photons, the researchers destroy the interference pattern of the signal photons. What has changed?

The answer is that the observer's potential knowledge has changed. He can now determine which route the signal photons took to their detector by comparing their arrival times with those of the remaining, unblocked idlers. The original photon can no longer go both ways at the beam splitter, like a wave, but must either bounce off or pass through, like a particle.

The comparison of arrival times need not actually be performed to destroy the interference pattern. The mere "threat" of obtaining information about which way the photon traveled, Mandel explains, forces it to travel only one route. "The quantum state reflects not only what we know about the system but what is in principle knowable," Mandel says.

Can the threat of obtaining incriminating information, once made, be retracted? In other words, are measurements reversible? Many theorists, including Bohr, thought not, and the phrase "collapse of the wave function" reflects that belief. But since 1983 Marlan O. Scully, a theorist at the University of New Mexico, has argued that it should be possible to gain information about the state of a quantum phenomenon, thereby destroying its wavelike properties, and then restore those properties by "erasing" the information.

Several groups working with optical interferometry, including Mandel's, claim to have demonstrated what Scully has dubbed a "quantum eraser." The group that has come closest, according to Scully, is one led by Raymond Y. Chiao of the University of California at Berkeley.

Earlier this year Chiao's group passed a beam of light through a down-conversion crystal, generating two identical photons. After being directed by mirrors along separate paths, the two photons crossed paths again at a half-silvered mirror and then entered two detectors. Because it was impossible to know which photon ended up in which detector, each photon seemed to go both ways. As in Mandel's experiment, the interference pattern was revealed by lengthening one arm of the interferometer; a device called a coincidence counter showed the simultaneous firings of the two photon detectors rising and falling as the two wavelets entering each detector went in and out of phase.

Then the workers added a device to the interferometer that shifted the polarization of one set of photons by 90 degrees. If one thinks of a ray of light as an arrow, polarization is the orientation of the plane of the arrowhead. One of the peculiarities of polarization is that it is a strictly binary property; photons are always polarized either vertically or horizontally. The altered polarization served as a tag; by putting polarization detectors in front of the simple light detectors at the end of the routes, one could determine which route each photon had taken. The two paths were no longer indistinguishable, and so the interference pattern disappeared.

Finally, Chiao's group inserted two devices that admitted only light polarized in one direction just in front of the detectors. The paths were indistinguishable again, and the interference pattern reappeared. Unlike Humpty-Dumpty, a collapsed wave function can be put back together again.

Spooky Action

Following up another proposal by Scully, Chiao has even suggested a way to delay the choice of whether or not to restore the interference pattern until *after* the photons have struck the detectors. The simple polarizing filters in front of the detectors are replaced with polarizing beam splitters, which direct photons with opposite polarization to different detectors. A computer then stores the data on the arrival times of all the photons in one file and the polarization of all the photons in another file. Viewed all at once without regard to polarization, the arrival times show no interference pattern. But if one separates differently polarized photons and plots them independently, two distinct interference patterns emerge.

Such possibilities provoke consternation in some quarters. Edwin T. Jaynes of Washington University, a prominent theorist whose work helped to inspire Scully to conceive the quantum eraser, has nonetheless dubbed it "medieval necromancy." Scully was so pleased by Jaynes's remark that he included it in a recent article on the quantum eraser.

Necromancy cannot hold a candle to nonlocality. Einstein, Boris Podolsky and Nathan Rosen first drew attention to this bizarre quantum property (which is now often called the EPR effect in their honor) in 1935 with a thought experiment designed to prove that quantum mechanics was hopelessly flawed. What would happen, Einstein and his colleagues asked, if a particle consisting of two protons decayed, sending the protons in opposite directions? According to quantum mechanics, as long as both protons remain unobserved their properties remain indefinite, in a superposition of all possible states; that means each one travels in all possible directions.

But because of their common origin, the properties of the protons are tightly correlated, or "entangled." For example, through simple conservation of momentum, one knows that if one proton heads north, the other must have headed south. Consequently, measuring the momentum of one proton instantaneously determines the momentum of the other proton—even if it has traveled to the opposite end of the universe. Einstein said that this "spooky action at a distance" was incompatible with any "realistic" model of reality; all the properties of each proton must be fixed from the moment they first fly apart.

Until the early 1960s, most physicists considered the issue entirely academic, since no one could imagine how to resolve it experimentally. Then, in 1964, John S. Bell of CERN, the European laboratory for particle physics, showed that quantum mechanics predicted stronger statistical correlations between entangled particles than the so-called local realistic theory that Einstein preferred. Bell's papers triggered a flurry of laboratory work, culminating in a classic (but not classical) experiment performed a decade ago by Alain Aspect of the University of Paris.

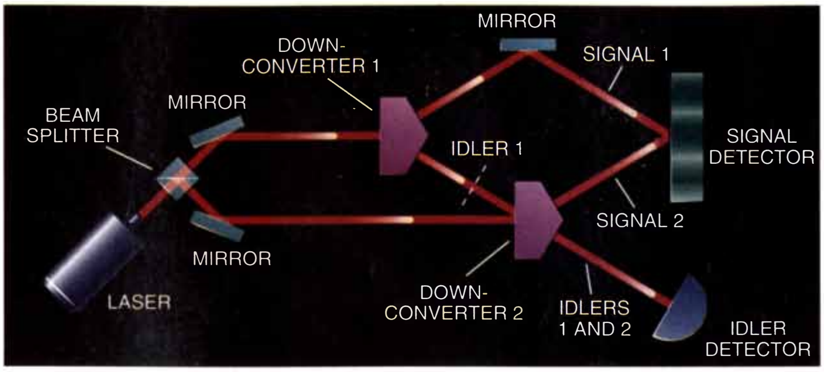
Instead of the momentum of protons, Aspect analyzed the polarization of pairs of photons emitted by a single source toward separate detectors. Measured independently, the polarization of each set of photons fluctuated in a seemingly random way. But when the two sets of measurements were compared, they displayed an agreement stronger than could be accounted for by any local realistic theory—just as Bell had predicted. Einstein's spooky action at a distance was real.

Until recently, no experiment had successfully shown that the EPR effect held true for momentum, as Einstein, Podolsky and Rosen had originally proposed. Two years ago John G. Rarity and Paul R. Tapster of the Royal Signals and Radar Establishment in England finally achieved that feat.

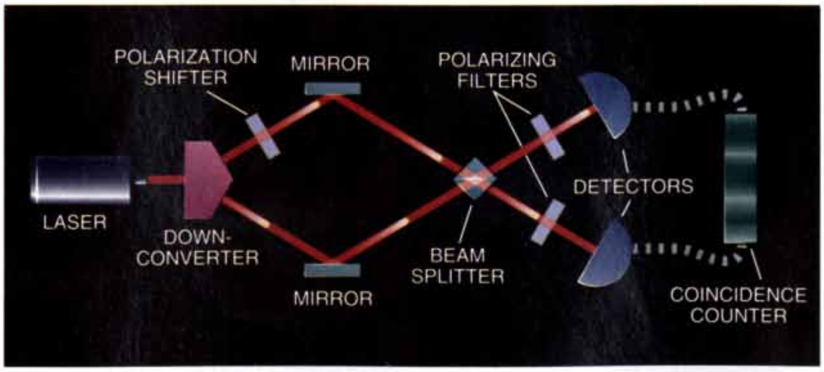
The experiment began with a laser firing into a down-converter, which produced pairs of correlated photons. Each of these photons then passed through a separate two-slit apparatus and thence to a photon detector. Through conservation of momentum, one could determine the route of each photon if one knew the route of its partner. But the arrangement of mirrors and beam splitters made it impossible to determine the actual route of either photon.

Next, the workers slightly lengthened one of the four routes, as Chiao did in his quantum eraser experiment. Although the rate at which photons struck each detector did not change, the rate of simultaneous firings recorded by a coincidence counter oscillated, forming a telltale interference pattern like the one

How to Destroy—and Revive—a Light Wave



Information rather than direct intervention destroys wavelike behavior in an experiment done at the University of Rochester. A laser fires photons past a half-silvered mirror, or beam splitter, to two down-converters, labeled 1 and 2. These convert each incident photon into two lower-energy photons, called signals and idlers. Because the signal detector cannot tell how the signals arrived, each signal takes both routes, like a wave, generating an interference pattern at the signal detector. But the pattern can be destroyed merely by blocking idlers from down-converter 1 (*dotted line*). The reason is that each signal's path can now be retraced; simultaneous detection of a signal and idler indicates that both came from a photon reflected by the beam splitter into down-converter 2.



Erasing information about the path of a photon restores wavelike behavior in an experiment done at the University of California at Berkeley. Pairs of identically polarized photons produced by a down-converter bounce off mirrors, converge again at a beam splitter and pass into two detectors. A coincidence counter observes an interference pattern in the rate of simultaneous detections by the two detectors, indicating that each photon has gone both ways at the beam splitter, like a wave. Adding a polarization shifter to one path destroys the pattern by making it possible to distinguish the photons. But placing two polarizing filters in front of the detectors makes the photons identical again, erasing the polarization distinction and restoring the interference pattern.

observed by Chiao. Such a pattern could occur only if each photon, the one on the left and the one on the right, was passing through both slits to its respective detector, its momentum fundamentally undefined and yet still correlated with the momentum of its distant partner.

Still more ambitious EPR experiments have been proposed but not yet carried out. Greenberger, Zeilinger and Michael Horne of Stonehill College have shown that three or more particles sprung from a single source will exhibit much stronger nonlocal correlations than those between just two particles. Bernard Yurke and David Stoler of AT&T Bell Laboratories have even suggested a way in which three particles emitted from separate locations can exhibit the EPR effect.

Unfortunately, the EPR effect does not provide a loophole in the theory of relativity, which prohibits communications faster than light, since each isolated observer of a correlated particle sees only an apparently random fluctuation of properties. But the effect does allow one safely to transmit a random number that can then serve as the numerical "key" for an encryption system. In fact, such a device has been built by Charles H. Bennett of the IBM Thomas J. Watson Research Center.

A die-hard realist might dismiss the experiments described above, since they all involve that quintessence of ineffability, light. But electrons, neutrons, protons and even whole atoms—the stuff our own bodies are made of—also display pathological behavior. Researchers observed wavelike behavior in electrons through indirect means as early as the 1920s, and they began carrying out two-slit experiments with electrons several decades ago.

Superposed Philosophers

A new round of electron experiments may be carried out soon if Yakir Aharonov of Tel-Aviv University has his way. Noting that superposition is generally inferred from observations of large numbers of particles, Aharonov contends that a single electron bound to a hydrogen atom could be detected smeared out in a relatively large cavity—say, 10 centimeters across—by very delicately scattering photons off it.

Aharonov has not yet published his idea—"I am a very fast thinker but a very slow writer," he says—and some physicists he has discussed it with are skeptical. On the other hand, many were skeptical in 1958, when Aharonov and David Bohm of the University of London suggested a way in which a magnetic field could influence an electron that, strictly speaking, lay completely beyond the field's range. The so-called Aharonov-Bohm effect has now been confirmed in laboratories.

Since the mid-1970s various workers have done interference experiments with neutrons, which are almost 2,000 times heavier than electrons. Some 15 years ago, for example, Samuel A. Werner of the University of Missouri at Columbia and others found that the interference pattern formed by neutrons diffracted along two paths by a sculpted silicon crystal could be altered simply by changing the interferometer's orientation relative to the earth's gravitational field. It was the first demonstration that the Schrödinger equation holds true under the sway of gravity.

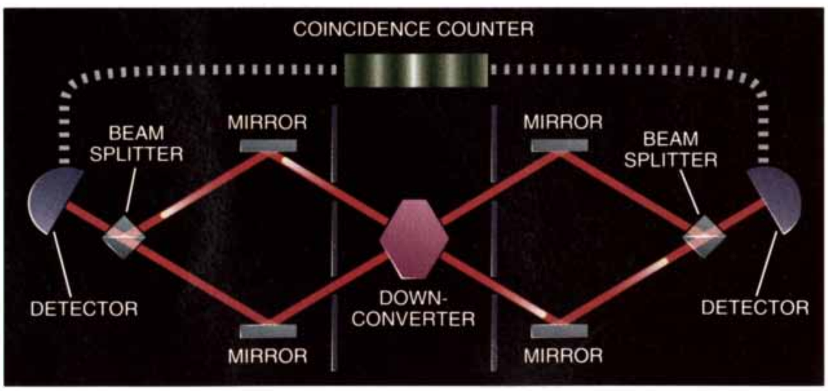
Investigators have begun doing interferometry with whole atoms only in the past few years. Such experiments are extraordinarily difficult. Atoms cannot pass through lenses or crystals, as photons, electrons and even neutrons can. Moreover, since the wavelength of an object is inversely proportional to its mass and velocity, the particle must move slowly for its wavelength to be detectable. Yet workers such as David E. Pritchard of the Massachusetts Institute of Technology have created the equivalent of beam splitters, mirrors and lenses for atoms out of metal plates with precisely machined grooves and even standing waves of light, formed when a wave of light reflects back on itself in such a way that its crests and troughs match precisely.

Pritchard says physicists may one day be able to pass biologically significant molecules such as proteins or nucleic acids through an interferometer. In principle, one could even observe wavelike behavior in a whole organism, such as an amoeba. There are some obstacles, though: the amoeba would have to travel very slowly, so slowly, in fact, that it would take some three years to get through the interferometer, according to Pritchard. The experiment would also have to be conducted in an environment completely free of gravitational or other influences—that is, in outer space.

Getting a slightly larger and more intelligent organism, for instance, a philosopher, to take two paths through a two-slit apparatus would be even trickier. "It would take longer than the age of the universe," Pritchard says.

While physicists may never nudge a philosopher into a superposition of states, they are hard at work trying to induce wavelike behavior in objects literally large enough to see. The research has rekindled interest in a famous thought experiment posed by Schrödinger in 1935. In a version altered by John Bell, the EPR theorist, to be more palatable to animal lovers, a cat is placed in a box containing a lump of radioactive matter, which has a 50 per-

How Distant Particles Keep in Touch

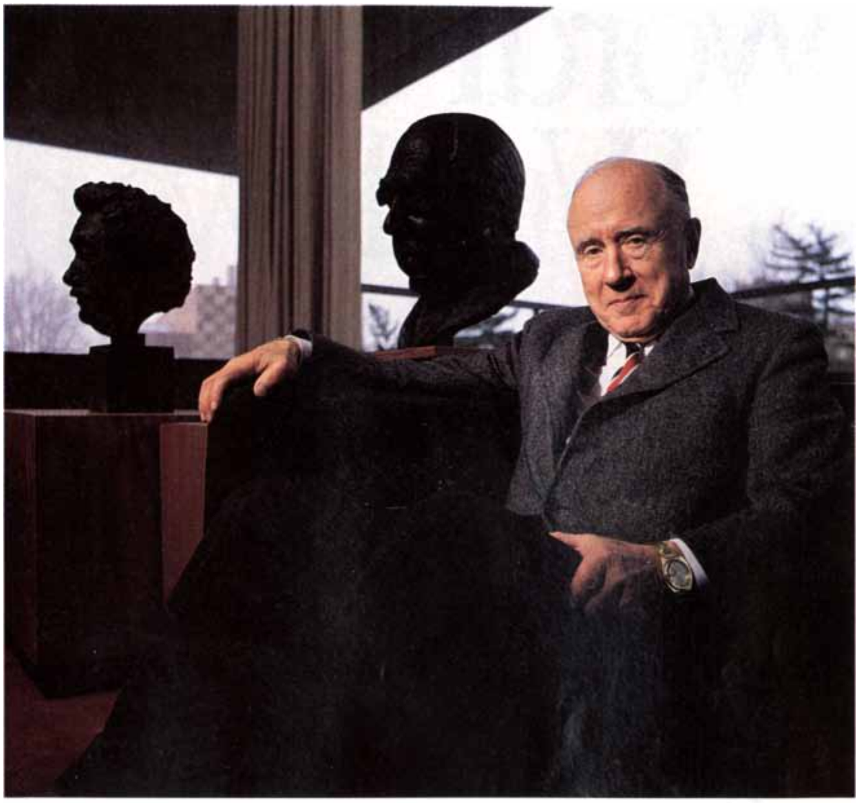


Spooky correlations between separate photons were demonstrated in an experiment at the Royal Signals and Radar Establishment in England. In this simplified depiction, a down-converter sends pairs of photons in opposite directions. Each photon passes through a separate two-slit apparatus and is directed by mirrors to a detector. Because the detectors cannot distinguish which slit a photon passes through, each photon goes both ways, generating an interference pattern in the coincidence counter. Yet each photon's direction, or momentum, is also correlated with its partner's. A measurement showing a photon going through the upper left slit would instantaneously force its distant partner to go through the lower slit on the right.

cent chance of emitting a particle in a one-hour period. When the particle decays, it triggers a Geiger counter, which in turn causes a flask of milk to pour into a bowl, feeding the cat. (In Schrödinger's version, a hammer smashes a flask of poison gas, killing the cat.)

Common sense dictates that a cat cannot have a stomach both empty and full. But quantum mechanics dictates that after one hour, if no one has looked in the box, the radioactive lump and so the cat exist in a superposition of indistinguishable states; the former is both decayed and undecayed, and the latter is both hungry and full.

Various resolutions to the paradox have been suggested. Wojciech H. Zurek, a theorist at Los Alamos National Laboratory, contends that as a quantum phenomenon propagates, its interaction with the environment inevitably causes its superposed states to become distinguishable and thus to collapse into a single state. Mandel of the University of Rochester thinks this view is supported by his experiment, in which the mere potential for knowledge of a photon's path destroyed its interference pattern. After all, one can easily learn whether the cat has been fed—say, by making the box transparent—without actually disturbing it.



JOHN A. WHEELER, seen here with the likenesses of two earlier explorers of the quantum realm, Einstein and Bohr, thinks the deepest lesson of quantum mechanics may be that reality is defined by the questions we put to it.

But since the early 1980s Anthony J. Leggett, a theorist at the University of Illinois, has argued that one should be able to observe a superconducting quantum interference device, more commonly called a SQUID, in a superposition of states. A SQUID, which is typically the size of a pinhead and therefore huge in comparison with atoms or other quantum objects, consists of a loop of superconducting material, through which electrons flow without resistance, broken by a thin slice of insulating material called a Josephson junction. In a classical world the electrons would be completely blocked by the insulator, but the quantum indefiniteness of the electrons' positions allows hordes of them to "tunnel" blithely through the gap.

Inspired by Leggett's calculations, Claudia D. Tesche of the IBM Watson center proposed an experiment that could show the superposition quite directly. Given certain conditions, Tesche notes, the current in a SQUID has an equal chance of flowing in either direction. According to quantum mechanics, then, it should flow both ways, creating an interference pattern analogous to the one formed in a two-slit experiment.

Tesche's design calls for placing two extremely sensitive switches around the SQUID, each of which is tripped when the current is going in a different direction. Of course, once a switch is tripped, the wave function collapses, and the interference pattern is destroyed. Tesche hopes to infer the pattern from those microseconds during which the switches are not activated—making measurements, in effect, by not making them.

Orthodoxy under Attack

Other theorists note that Tesche's experiment is extremely difficult, since even minute disturbances from the environment can cause the SQUID's wave function to collapse. In fact, Tesche recently turned to other, more conventional pursuits, at least temporarily setting aside the experiment. "It wasn't working very well," she concedes.

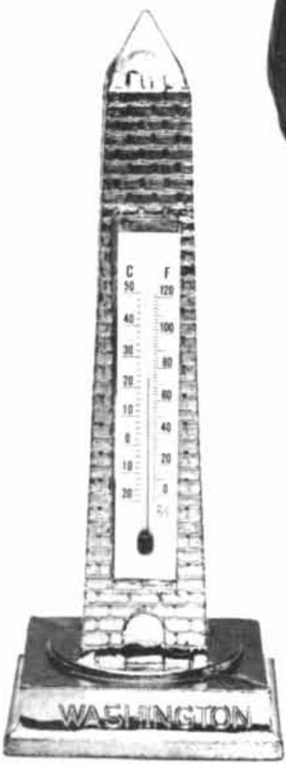
Yet less ambitious experiments by John Clarke of the University of California at Berkeley, Richard A. Webb of IBM and others have produced strong circumstantial evidence that a SQUID can in fact exist in a superposition of two states. The experiments involve a property known as flux, which is the area of the superconducting ring multiplied by the strength of the magnetic field perpendicular to the ring. In an ordinary superconducting ring the flux would be constant, but measurements with magnetometers show the flux of the SQUID spontaneously jumping from one value to another. Such jumps can occur only if the flux is in a superposition of states—hungry and full at the same time, as it were.

All the recent experiments, completed and proposed, have hardly led to a consensus on what exactly quantum mechanics means. If only by default, the "orthodox" view of quantum mechanics is still the one set forth in the 1920s by Bohr. Called the Copenhagen interpretation, its basic assertion is that what we observe is all we can know; any speculation about what a photon, an atom or even a SQUID "really is" or what it is doing when we're not looking is just that—speculation.

To be sure, the Copenhagen interpretation has come under attack from theorists in recent years, most notably from John Bell, author of the brilliant proof of the divergence between "realistic" and quantum predictions for EPR experiments. In a television interview just before his sudden death from a stroke two years ago, the Irish physicist expressed his dissatisfaction with the Copenhagen interpretation, noting that it "says we must accept meaninglessness." Does that make you afraid? the interviewer asked. "No, just disgusted," Bell replied, smiling.

WordPerfect users pre
Windows no matter

Portland: 92% of WordPerfect users found Microsoft Word for Windows easier to learn.



Washington, D.C.: 72% of WordPerfect users found Microsoft Word for Windows easier to learn.

Dallas: 88% of WordPerfect users preferred Microsoft Word for Windows over WordPerfect for Windows.

Los Angeles: 72% of WordPerfect users, if given a choice, would buy Microsoft Word for Windows over WordPerfect for Windows.

The 1992 Coast-to-

The 1992 Word Challenge is over.

And as you can see, the results speak for themselves. WordPerfect® for DOS users all over the country prefer Word for Windows™ for everyday word processing tasks. Over the past two months, the National Software Testing Labs visited 10 different cities to ask WordPerfect users to compare WordPerfect for Windows and Microsoft® Word for Windows side-by-side. WordPerfect users were amazed at how Word put them one step away from accomplishing everyday word processing tasks with, in many instances, one simple click of the mouse. They found it that easy.

The tour was not only fun, but it definitely confirmed two very strong hunches we've had for some time. One, WordPerfect

*Offer good for current licensees of WordPerfect, MultiMate,® WordStar,® MS® Word for MS-DOS® and DisplayWrite.® Please allow 2-4 weeks for delivery upon receipt of order by Microsoft. Offer expires 9/30/92. Limit one per customer. Reseller prices may vary. Call for system

fer Microsoft Word for what state they're in.

Chicago: 68% of WordPerfect users, if given a choice, would buy Microsoft Word for Windows over WordPerfect for Windows.

St. Louis: 71% of WordPerfect users preferred Microsoft Word for Windows over WordPerfect for Windows.

Salt Lake City: 76% of WordPerfect users preferred Microsoft Word for Windows over WordPerfect for Windows.

Minneapolis: 88% of WordPerfect users felt that Microsoft Word for Windows was the easiest to use.

Atlanta: 74% of WordPerfect users felt that Microsoft Word for Windows was the easiest to use.

New York: 80% of WordPerfect users think that Microsoft Word for Windows is easier to use.

Coast Word Challenge.

users prefer Word for Windows for everyday word processing tasks. And two, airports sell some very tacky souvenirs.

But you can see for yourself. Simply call us for a free "Word Challenge Kit." Included is a videocassette highlighting the actual test, and interviews with WordPerfect users who took the challenge. Also in the kit are the files you need to test Word on your own. Call (800) 323-3577, Department HD4. We think once you get a chance to judge for yourself, you'll no doubt want to take advantage of our special $129 upgrade offer.*

No matter what state you're in.

Microsoft
Making it easier

requirements. Offer good only in the 50 United States. To receive only the NSTL test results: In the United States, call (800) 323-3577, Dept. HD4. For information only: In Canada, call (800) 563-9048; outside the U.S. and Canada, call (206) 936-8661.

Bell's exhortations helped to revive interest in a realistic theory originally proposed in the 1950s by Bohm. In Bohm's view, a quantum entity such as an electron does in fact exist in a particular place at a particular time, but its behavior is governed by an unusual field, or pilot wave, whose properties are defined by the Schrödinger wave function. The hypothesis does allow one quantum quirk, nonlocality, but it eliminates another, the indefiniteness of position of a particle. Its predictions are identical to those of standard quantum mechanics.

Bell also boosted the standing of a theory developed six years ago by Gian-Carlo Ghirardi and Tullio Weber of the University of Trieste and Alberto Rimini of the University of Pavia and refined more recently by Philip Pearle of Hamilton College. By adding a nonlinear term to the Schrödinger equation, the theory causes superposed states of a system to converge into a single state as the system approaches macroscopic dimensions, thereby eliminating the Schrödinger's cat paradox, among other embarrassments.

Unlike Bohm's pilot-wave concept, the theory of Ghirardi's group offers predictions that diverge from those of orthodox quantum physics, albeit subtly. "If you shine a neutron through two slits, you get an interference pattern," Pearle says. "But if our theory is correct, the interference should disappear if you make the measurement far enough away." The theory also requires slight violations of the law of conservation of energy. Zeilinger of the University of Innsbruck was sufficiently interested in the theory to test the neutron prediction, which was not borne out. "This approach is one of those dead-end roads that has to be walked by someone," he sighs.

Yet another view currently enjoying some attention, although not as a result of Bell's efforts, is the many-worlds interpretation, which was invented in the 1950s by Hugh Everett III of Princeton. The theory sought to answer the question of why, when we observe a quantum phenomenon, we see only one outcome of the many allowed by its wave function. Everett proposed that whenever a measurement forces a particle to make a choice, for instance, between going left or right in a two-slit apparatus, the entire universe splits into two separate universes; the particle goes left in one universe and right in the other.

Although the theory was long dismissed as more science fiction than science, it has been revived in a modified form by Murray Gell-Mann of the California Institute of Technology and James B. Hartle of the University of California at Santa Barbara. They call their version the many-histories interpretation and emphasize that the histories are "potentialities" rather than physical actualities. Gell-Mann has reportedly predicted that this view will dominate the field by the end of the century.

An intriguing alternative, called the many-minds view, has been advanced by David Z. Albert, a physicist-turned-philosopher at Columbia University, and Barry Loewer, a philosopher from Rutgers University. Each observer, they explain, or "sentient physical system," is associated with an infinite set of minds, which experience different possible outcomes of any quantum measurement. The array of choices embedded in the Schrödinger equation corresponds to the myriad experiences undergone by these minds rather than to an infinitude of universes. The concept may sound far-fetched, but it is no more radical, Albert argues, than the many-histories theory or even the Copenhagen interpretation itself.

The It from Bit

Other philosophers call for a sea change in our very modes of thought. After Einstein introduced his theory of relativity, notes Jeffrey Bub, a philosopher at the University of Maryland, "we threw out the old Euclidean notion of space and time, and now we have a more generalized notion." Quantum theory may demand a similar revamping of our concepts of rationality and logic, Bub says. Boolean logic, which is based on either-or propositions, suffices for a world in which an atom goes either through one slit or the other, but not both slits. "Quantum mechanical logic is non-Boolean," he comments. "Once you understand that, it may make sense." Bub concedes, however, that none of the so-called quantum logic systems devised so far has proved very convincing.

A different kind of paradigm shift is envisioned by Wheeler. The most profound lesson of quantum mechanics, he remarks, is that physical phenomena are somehow defined by the questions we ask of them. "This is in some sense a participatory universe," he says. The basis of reality may not be the quantum, which despite its elusiveness is still a physical phenomenon, but the bit, the answer to a yes-or-no question, which is the fundamental currency of computing and communications. Wheeler calls his idea "the it from bit."

Following Wheeler's lead, various theorists are trying to recast quantum physics in terms of information theory, which was developed 44 years ago to maximize the amount of information transmitted over communications channels. Already these investigators have found that Heisenberg's uncertainty principle, wave-particle duality and nonlocality can be formulated more powerfully in the context of information theory, according to William K. Wootters of Williams College, a former Wheeler student who is pursuing the it-from-bit concept.

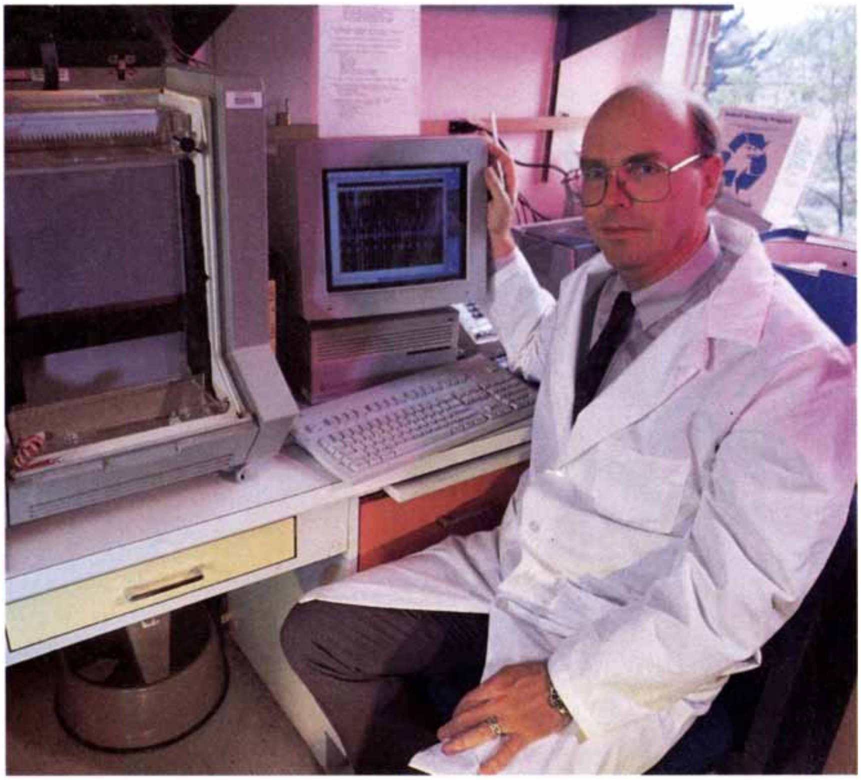
Meanwhile theorists at the surreal frontier of quantum theory are conjuring up thought experiments that could unveil the riddle in the enigma once and for all. David Deutsch of the University of Oxford thinks it should be possible, at least in principle, to build a "quantum computer," one that achieves superposition of states. Deutsch has shown that if different superposed states of the computer can work on separate parts of a problem at the same time, the computer may achieve a kind of quantum parallelism, solving certain problems more quickly than classical computers.

Taking this idea further, Albert—with just one of his minds—has conceived of a quantum computer capable of making certain measurements of itself and its environment. Such a "quantum automaton" would be capable of knowing more about itself than any outside observer could ever know—and even more than is ordinarily permitted by the uncertainty principle. The automaton could also serve as a kind of eyewitness of the quantum world, resolving questions about whether wave functions truly collapse, for example. Albert says he has no idea how actually to engineer such a machine, but his calculations show the Schrödinger equation allows such a possibility.

If that doesn't work, there is always Aharonov's time machine. The machine, which is based not only on quantum theory but also on general relativity, is a massive sphere that can rapidly expand or contract. Einstein's theory predicts that time will speed up for an occupant of the sphere as it expands and gravity becomes proportionately weaker, and time will slow down as the sphere contracts. If the machine and its occupant can be induced into a superposition of states corresponding to different sizes and so different rates of time, Aharonov says, they may "tunnel" into the future. The occupant can then disembark, ask physicists of the future to explain the mysteries of quantum mechanics and then bring the answers—assuming there are any—back to the present. Until then, like Plato's benighted cave dwellers, we can only stare at the shadows of quanta flickering on the walls of our cave and wonder what they mean.

Piecemeal Patents

The U.S. reconsiders patents on DNA fragments



J. CRAIG VENTER with DNA sequencing machine, one of seven in his laboratory that are tagging human genes. Photo: Randy Santos/Randolph Photography.

When J. Craig Venter, chief of receptor biochemistry and molecular biology at the National Institute of Neurological Disorders and Stroke, developed a rapid, inexpensive technique for labeling genes, he expected to advance genetics, not generate a major controversy. But for the past nine months, scientists, lawyers and representatives of the biotechnology industry have been furiously debating whether the early fruits of Venter's technique—more than 2,700 gene fragments—can or should be patented by the National Institutes of Health.

The dispute is not likely to be resolved soon. Although the U.S. Patent and Trademark Office may issue a preliminary decision as early as this summer on whether the sequences are patentable, appeals could grind on for years. Meanwhile Louis W. Sullivan, secretary of Health and Human Services, is formally faced with determining whether the NIH will continue to press its patent claims. That question is also being aired in interagency discussions coordinated by the president's science adviser, D. Allan Bromley.

The controversy began in October 1991, after the NIH applied for patents on 347 pieces of DNA isolated and sequenced by Venter. A follow-up application in February of this year covered another 2,375 fragments. Each fragment, which Venter calls an expressed sequence tag (EST), contains about 400 base pairs of DNA corresponding to part of a gene that is expressed in human brain cells. Venter identified the sequences in commercially available collections of human genes.

By searching through computer data bases to see whether an EST corresponds to a known gene, Venter can guess what kind of gene some of his ESTs might correspond to. But Venter knows nothing about most of his ESTs except their sequence. Even so, he estimates that his technique could be used to map and partially sequence most human genes for less than $10 million, far less than the several billion dollars the international human genome program is planning to spend to map and sequence the entire genome. Few doubt him: Venter's laboratory, equipped with seven automatic sequencing machines and three robots, is now tagging 150 genes a day. According to one estimate, Venter's laboratory alone could triple the work load of the Patent and Trademark Office's biotechnology section.

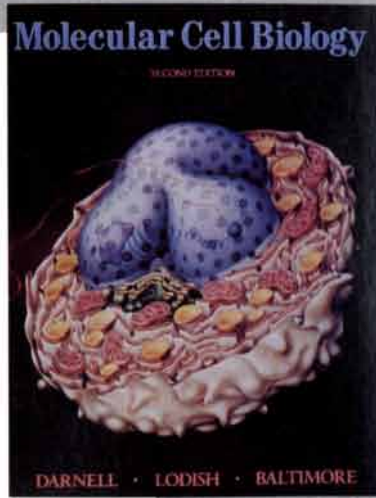
The initial decision to apply for patents on Venter's sequences was made by Reid G. Adler, the NIH's director of technology transfer. Adler realized that the sequences would not be patentable if Venter published them without first laying claim in the form of a patent application. If nobody could patent the ESTs, Adler reasoned, drug companies would have little incentive to develop products based on them. The NIH opted to rescue the situation by filing for patents so it could license the corresponding genes and possibly their proteins to U.S. companies. Both Adler and NIH director Bernadine Healy have repeatedly asserted that the filing was a defensive move. "We're not in this to make money," Adler insists.

The principal U.S. trade organizations, the Industrial Biotechnology Association (IBA) and the Association of Biotechnology Companies, initially supported the NIH. But a committee of scientific advisers to the Human Genome Project opposed the NIH action, as did the American Society for Human Genetics. "There is no question that the patenting of ESTs...will precipitate a race to isolate ESTs in many countries," the society said in a position paper. "It is virtually certain that under such conditions the information would not be shared between competing groups." Says David Botstein, professor of genetics at Stanford University: "A catalogue of sequences that says 'Kilroy was here' is of modest value."

James D. Watson, the co-discoverer of the double helical structure of DNA and until recently head of the National Center for Human Genome Research at the NIH, criticized Venter's approach and publicly disagreed with Healy over the decision to file for patents. Watson angrily resigned the post in April.

Recently biotechnology companies have also begun tilting toward opposing the NIH claims, according to Lisa J.

KEEPING YOU CURRENT IN TODAY'S CHANGING WORLD



The Eagerly Awaited Revision of a Groundbreaking Book

MOLECULAR CELL BIOLOGY

Second Edition

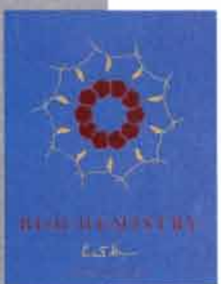
James Darnell, Harvey Lodish, and David Baltimore
A Scientific American Book

The most authoritative, comprehensive guide to the molecular biology of the cell, MOLECULAR CELL BIOLOGY is the only book to fully present modern techniques and findings and tie these developments to basic principles.

Thoroughly updated to cover the latest advances and fully revised for greater accessibility and convenience, MOLECULAR CELL BIOLOGY is an ideal reference. The Second Edition features a brand new full color art program that adds a wealth of visual detail. Don't miss it!

Order your copy today and save 20%!

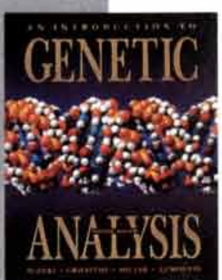
1105 pages, 1095 illustrations, hardbound.



BIOCHEMISTRY, Third Edition
Lubert Stryer

"Destined to remain one of the 'bibles' of our subject for many years to come."—*Trends in Biochemical Sciences*

The world's leading biochemistry book is an indispensable desk reference for anyone in the biochemical sciences.
1988, 1098 pages, 1476 illustrations, hardbound



AN INTRODUCTION TO GENETIC ANALYSIS
Fourth Edition
David T. Suzuki, Anthony J. F. Griffiths, Jeffrey H. Miller, and Richard C. Lewontin

This profusely illustrated book is an accessible introduction to the principles, methods, and findings of genetic analysis. Including all recent advances, it presents balanced coverage of classical and modern genetics.
1989, 768 pages, 831 illustrations, hardbound

Your satisfaction matters!
Return any book in good condition within 15 days of receipt and we'll send you a prompt refund.

To order, mail this coupon or a copy to:
W. H. Freeman and Company
4419 West 1980 South
Salt Lake City, UT 84104

☐ I enclose my check or money order made payable to W.H. Freeman and Company.

☐ Please use my: ☐Visa☐MasterCard
Exp. Date__________
Account #__________
Signature__________
(Credit card orders must be signed)

Name__________
Address__________
City/State/Zip__________

Title	Author	ISBN	Price	Total
Molecular Cell Biology, 2/e	Darnell et al.	1981-1	$45.99*	
Biochemistry,3/e	Stryer	1843-X	$56.95	
Intro to Genetic Analysis, 4/e	Suzuki et al.	1956-8	$46.95	

Subtotal ______
(Add $1.95 for shipping & handling of first book, $1.25 for each additional book.) ______
Sales Tax (NY, CA, UT) ______
TOTAL ______

* Darnell, et al. regular price $56.95

For Canadian residents add $2.25 for postage and handling, $1.50 each additional book. Please allow four weeks for delivery. All orders should be prepaid in US dollars.

4917

W. H. Freeman and Company
The book publishing arm of Scientific American

Raines, vice president of government affairs at the IBA. "A typical biotechnology patent on a gene is the result of 300 person-years of work," Raines says. "These claims are based on an average of 20 person-minutes of work. Many companies question whether the NIH should have the right to license what is really no more than a research tool."

Some companies fear that if the NIH receives its patents and licenses the genes to companies, it could invoke a clause that any products developed with them are "reasonably priced." (The AIDS drug dideoxyinosine, or ddI, which the NIH licenses to Bristol-Myers Squibb, is subject to such a restriction, Raines notes.) Small companies are worried about having to search through thousands of patented EST sequences to see whether a gene has been tagged by the NIH. And licensing could be a nightmare if different ESTs from the same gene have been claimed by different researchers or countries.

Raines emphasizes that IBA members think the crucial question is whether a gene of unknown function can be patented. A patentable invention has to be useful and novel, as well as not obvious. The requirements are roughly similar in other countries. Venter's technique is based on procedures well known to molecular biologists, raising questions about its nonobviousness. And it is far from clear whether ESTs have any uses other than as tools for more research. Adler argues that the ESTs can be used as markers to find genes. But some patent lawyers are not sure.

On the other side of the coin, IBA members worry that if the U.S. does not patent but other countries do, U.S. companies could be locked out. Norton D. Zinder, a biologist at the Rockefeller University and former head of the board of science advisers to the genome project, derisively calls the NIH filing "preemptive patenting" and points out that a "gene race" is already under way. Workers in Britain, France and Japan, he says, are churning out ESTs as fast as or faster than Venter is. Indeed, the British Medical Research Council was to file for patents on some 4,000 ESTs in May and was keeping the sequences secret until it did so.

Zinder also points out that Jonathan Marc Rothberg of Yale University has discovered that some of Venter's ESTs contain parts of the commercial molecular systems used to copy them. That provides further evidence that "everything is not being done scientifically," Zinder says. Venter counters that such errors are an insignificant result of publishing the data as quickly as possible. And he contends there would be a race to tag genes whether ESTs are patentable or not.

British officials say they, too, are only filing defensively: they are seeking an international understanding on when DNA sequences can be patented. "It's the prevailing view that it is inappropriate to receive patents on sequences of unknown utility," says David Owen of the British Medical Research Council. Owen says Britain, like most European countries, supports a "patent bonfire" for ESTs. The IBA also sees a possible solution in an international agreement that would make sequences of unknown function unpatentable. According to Raines, the NIH might be able to fulfill its legal obligation to commercialize its discoveries without taking out patents.

The NIH was scheduled to participate in late May in a public meeting on its patent claims. Other countries, including Britain, were considering using the occasion to urge the NIH to withdraw. And in May, Healy gave a hint that the NIH may be ready to listen, provided other countries agree to play along.

All eyes are now on Japan, which has been keeping its options open. But the public meeting might, Healy said, lead to a decision to put a hold on a third large patent filing planned for June. For his part, Venter would be "delighted" if an agreement made gene fragments unpatentable. *—Tim Beardsley*

U.S. Semiconductor Toolmakers Regain Ground

After an excruciating slide in market share, the U.S. companies that make equipment for manufacturing semiconductor memory and integrated circuit chips are inching back. According to the San Jose–based consulting firm VLSI Research, the U.S. toolmakers increased their worldwide market share by about 3 percent last year—the first increase in a decade.

The U.S. gains came in spite of a tough year for chip machine makers worldwide. Total purchases of semiconductor manufacturing equipment dipped from $10.2 billion in 1990 to $9.9 billion. Only U.S. toolmakers posted a gain; Japanese manufacturers as well as other producers suffered losses.

Until last year, the story of U.S. makers of such equipment as chemical-vapor deposition and optical lithography tools was one of decline. In 1981 U.S. producers controlled more than 73 percent of the international market for semiconductor manufacturing equipment. That share had slipped below 43 percent by 1990. After the 1991 gains, however, U.S. and Japanese chip toolmakers may be standing head to head. Half the top 10 firms in the industry are based in Japan; the other half are in the U.S.

Those who read the tea leaves of the semiconductor trade found last year's results particularly heartening because sales of U.S. producers rose both at home and abroad. U.S. firms won more than two thirds of their gains last year in Japan. Sales were up in their home market as well, reversing the steady erosion that has been taking place. "Americans now seem more willing to buy American equipment," declares Jerry Hutcheson, chief executive officer of VLSI Research.

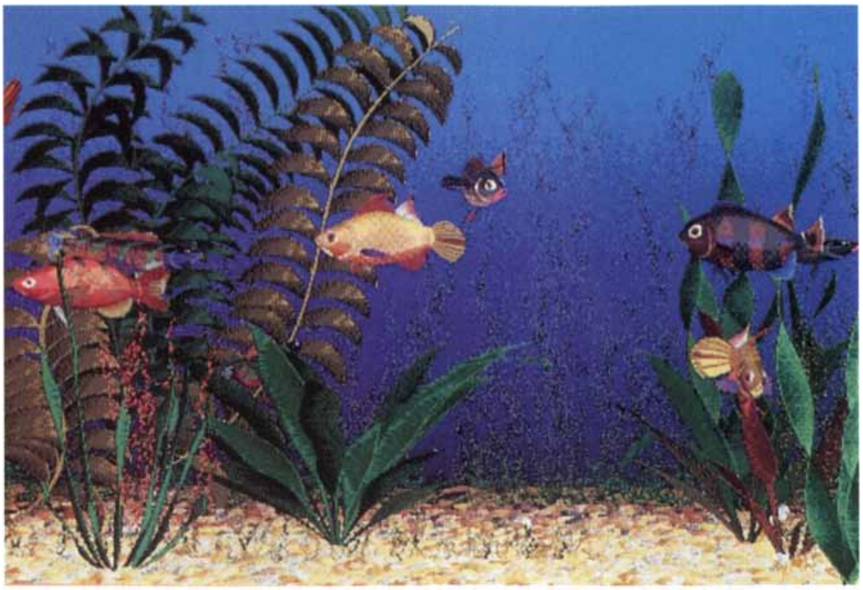
Hutcheson attributes much of the equipment makers' comeback to the support extended to the industry by Sematech, the semiconductor manufacturing research consortium sponsored jointly by industry and government. "I don't see anything else out there that would have caused the increase," he says.

That assessment can only help Sematech, whose continued funding is up for debate in Washington. When Sematech was established in 1987, the U.S. government promised to carry half of its annual $200-million budget; industry provided the remaining half of the tab. With the five-year honeymoon over, Sematech, like any other program, must now cajole the government for support.

In its fiscal 1993 budget proposals, the administration leaned toward shaving its contributions to Sematech, offering only $80 million. Not surprisingly, Sematech is asking for an increase above its previous allowance. Congress aims to resolve this and all other funding squabbles before the fiscal year begins on October 1. According to congressional aides, the legislators seem inclined to try to keep Sematech's budget intact. *—Elizabeth Corcoran*

One Fish, Two Fish

How to raise a school of tempting software toys



COMPUTER-GENERATED FISH, spawned from algorithms built by the Russian company AnimaTek, are at the heart of a novel computer toy. Image courtesy of Maxis.

These days Vladimir I. Pokhilko spends his time thinking about fish. Not delicious morsels of shad and salmon—not even endangered creatures such as short-nosed sturgeon. Instead Pokhilko is interested in virtual ichthyology. "We have very, very beautiful fish," boasts the Russian, as a few waft gracefully across his computer screen. And all are figments of algorithmic manipulations.

Pokhilko is the U.S. representative for the Moscow-based start-up AnimaTek. For the past three years, he and a team of 15 scraggly-haired scientists and computer programmers have painstakingly toiled to fulfill a dream: write software that electronically conceives images of glamorous fish. "They're driven by a sense of aesthetics," observes Will Wright, chief designer at the software toy company Maxis in Orinda, Calif. "They keep thinking about how people will compete to breed beautiful fish."

That said, Maxis now hopes to help the Russians use their El-Fish to reel in profits. As they see it, El-Fish will be a computer toy that people can use to create novel-looking fish by electronically "breeding" images. Users will endow the fish with characteristics that would enable them to survive under specific environmental conditions. Unlike computer games aimed at squashing aliens or conquering obstacles, El-Fish is a plaything for pacificists built from surprisingly sophisticated algorithms.

Seven years ago neither Pokhilko nor any of his compatriots imagined themselves as entrepreneurs. "El-Fish is a child of *perestroika,*" Pokhilko declares. Computers had permeated the premiere research institutions, and scientists from many disciplines were tinkering with ways in which these tools might advance their studies. As a psychologist, Pokhilko became intrigued with using computers to model—and tease—human intelligence and so began devising computer games.

In 1989, shortly after he had joined Moscow State University, Pokhilko met hacker Alexey Pajitnov, the author of Tetris, the first Russian computer game to slip across the border and become a commercial hit in the West. They began tinkering with a number of computer projects, including one that would automatically generate artificial life-forms. Although they were not familiar with U.S. work on genetic algorithms, developing software that could selectively "breed," or recombine, images to generate new versions seemed a promising approach [see "Genetic Algorithms," by John H. Holland, page 66]. Although the Russians first thought about software that would generate flowers, they settled on fish instead because the creatures were mobile, pretty and popular pets. "So it was serious art and industry," Pokhilko says.

In the spring of 1989 the Russians found an enthusiastic backer in a small U.S. firm, Bullet-Proof Software in Redmond, Wash. With funding from Bullet-Proof, Pokhilko and Pajitnov bought computer hardware and software, hired a team of scientists in Moscow (most of whom were theoretical and experimental physicists) and plunged into the messy business of algorithmic modeling. "It was a good time for private projects," Pokhilko recalls. "Some people took the job just to survive."

Their first geometric models proved so flexible that the software churned out many unusual images but no fish, Pokhilko sighs. By adding 800 parameters—essentially a genetic code to describe generic fish bodies—the workers finally ensured that their system would spew out only piscine figures.

Coaxing the fish to act like fish required more characteristics. The programmers imbued the fish with sensitivity to water temperature and depth. (They also designed lake environments and families of seaweed to provide a familiar setting.) To enable the fish to swim, the scientists tuned the software to sketch out 220 frames portraying fish in various angles and positions. Just as cartoons spring to life when an artist quickly flips through many versions, El-Fish meander across the screen when the computer runs through these frames in different orders.

Lately El-Fish are surfacing at conferences ranging from scientific symposiums on artificial life to trade shows on computer games. In demonstrations, Pokhilko pulls up a menu of several dozen fish and picks two parents. In less than 30 seconds the first progeny pop onto the screen, soon followed by a school of siblings. Pokhilko can then put the fish into a virtual lake stocked with seaweed and other fish and monitor how the newest permutation survives. Alternatively, he can save the fish image in a "museum" (which is environmentally benign).

The software has gone through its own evolution since Pokhilko and Pajitov brought it to the U.S. about a year ago. Even now the images consume much computer memory: 10 fish swimming across a screen need about eight megabytes of room, says Steve G. Beckert, a project manager at Maxis who is working with the Russians. And then there are those extras, such as easy-to-use interfaces, documentation and packaging—elements that are at the forefront of any commercial vendor's agenda but are new to the entrepreneurs.

At this point, Maxis and AnimaTek hope to unveil a commercial El-Fish package in time for Christmas. Precisely how the final product will fare is as uncertain as the outcome of breeding fish. Nevertheless, as Pokhilko points out, "sometimes a little bit of mutation is absolutely necessary to generate a beautiful fish." —*Elizabeth Corcoran*

Sensible Biology

Scientists have been attempting since the 1950s to meld biology and electronics into detection devices known as biosensors. But it is only now that technologies for mass-producing biologically active compounds, sophisticated membranes and integrated circuits are converging to make commercial products a reality.

Leading the way are a handful of companies, including Molecular Devices, i-Stat, Idetek and Biocircuits. The devices they are bringing to market will enable physicians to perform pretreatment tests of chemotherapy on a patient's own cancer cells or monitor drug levels in the blood. Manufacturers expect biosensors to speed pharmaceutical research and perhaps help replace animal testing. They will also be useful in industrial production to look at yield and by-products and to screen for contaminants.

Combining technologies as disparate as semiconductors and molecular biology has not been an easy undertaking, observes Harden M. McConnell, founder of Molecular Devices in Menlo Park, Calif., and chairman of Stanford University's department of chemistry. "If we'd known anything about silicon, we'd never have started all this," he chuckles. The firm recently introduced its Cytosensor Microphysiometer, which monitors the metabolic rate of living cells. By detecting changes in the acidity of cellular waste products, researchers can determine the effects of drugs or toxins in real time.

The principle of measuring pH in living cells is simple, McConnell says; building the biosensor was anything but. The current device can monitor eight samples at a time. Silicon chips sit like burners on a stove; on each is placed a small vessel containing a dollop of cells. A nutrient solution is pumped through the vessels, but periodically flow is halted to allow acid from the cells' waste to accumulate in the solution. As ions build up, they pass through a permeable membrane that forms the bottom of the vessel, carrying their charge to the chip below.

The Cytosensor is proving particularly useful in the areas of neuropharmacology and immunology, the company claims, because it can pick up changes in a cell's metabolism in about 30 seconds—fast enough to ascertain the influence of neurotransmitters and other quick-acting substances.

Processing speed is the key to the blood-screening biosensor sold by i-Stat in Cambridge, Mass., explains William Moffit, the firm's president. Doctors currently rely on clinical laboratories to process blood samples "stat," or as fast as possible, he notes. The company's device, meant to be used at bedside, relies on patented algorithms and a microprocessor to predict accurately what the electrodes will ultimately register. In 90 seconds the device displays levels of six critical blood components, such as glucose, sodium and potassium. "Without algorithms, you'd have a device just like ours, but it would take 10 minutes," Moffitt asserts.

The clinical diagnostic market is also the target of Biocircuits in Burlingame, Calif. It plans to introduce by mid-1993 a biosensor to detect thyroid dysfunction and, later, biosensors for other metabolic disorders. Disposable cartridges that slip into a tabletop reader will contain sections of a fluorescent membrane. When target substances bind to the membrane, the fluorescence is quenched, and the change can be measured by a simple photodiode.

The "definitely low-tech user" is the customer sought by Idetek for its biosensors, says Mark Platshon, president of the Sunnyvale, Calif., company. Truck drivers who go from farm to farm collecting milk will use the company's dipsticks to test for antibiotic residues in milk. Idetek's biosensor adds biology to a physical phenomenon first described in post-Renaissance Italy. It was noted then that a grooved mirror gives off a distinctive diffraction pattern because white light is broken into its component colors. "The ultimate mirror is a silicon wafer," Platshon says.

The manufacturer coats silicon wafers with antibodies that bind to antibiotics, then cuts stripes of alternating grooves with photolithographic techniques. The chips are dipped in milk and then exposed to laser light. If no antibiotics have bound to the chip, the laser simply reflects off the grooves; in the presence of contamination, a bright diffraction pattern emerges.

Regardless of the markets that early biosensor companies seek, all face the same basic obstacle: transforming research into high-quality products that can be made and sold in large quantities. Companies are betting that their biosensors will soon be detecting dollars. *—Deborah Erickson*

Metastatic Machinations

Blood-clotting factors may help cancer cells spread in the body

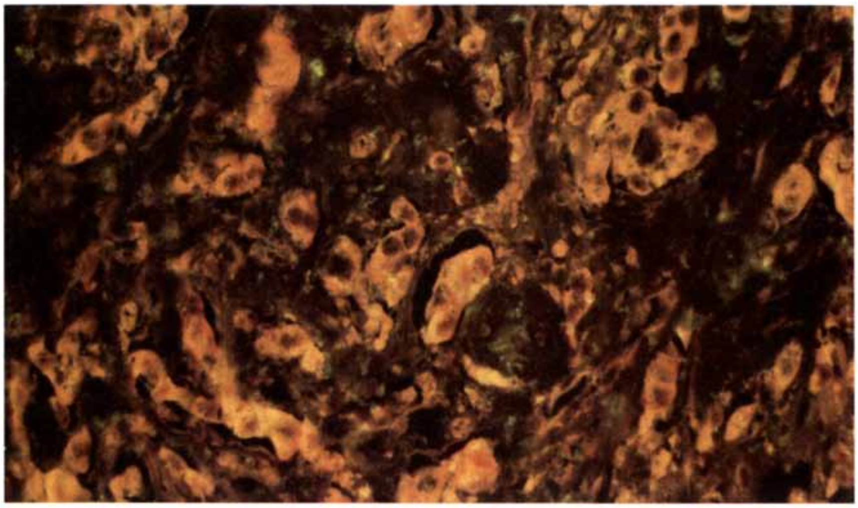
Something intriguing happened as scientists at Abbott Laboratories were trying to improve one of the firm's drugs for treating heart attacks. Researchers at the north Chicago facility were adding fragments of the clot-dissolving enzyme urokinase to cultured cell lines that they suspected might have receptors for the molecule. To their surprise, urokinase dramatically enhanced the ability of prostate, breast and colon cancer cells to invade normal tissue.

The company's research is now suggesting ways to prevent metastasis, the spread of cancer to distant sites in the body. "Cancers in general have been known to promote clot formation, presumably to protect themselves from the body's immune system, so we were intrigued to see this connection with a clot-dissolving compound," says Arthur A. Sasahara, director of Abbott's Thrombolytics Venture. To gain insight into these initial observations, the company is providing researchers around the country with the enzyme.

Gradually, the scientists are piecing together a picture of how urokinase assists cancer. Ordinarily, urokinase initiates a cascade of activity against blood clots that form at sites of injury. The process includes the conversion of an inactive precursor molecule called plasminogen to active plasmin—a compound that lyses fibrin and can also degrade other proteins. Studies suggest that in certain cancers, plasmin is turned against collagen and laminin, both essential components of the basement membrane that surrounds the walls of blood vessels. "We think urokinase wiggles between the cells in the wall, and then cancer cells move to the breach and into the bloodstream," Sasahara explains.

Douglas D. Boyd, a tumor biologist at the M.D. Anderson Cancer Center in Houston, contends that cell-surface receptors are critical to urokinase's effects. When he inserted the gene for the enzyme into colon cancer cell lines that had few receptors for the molecule, the cells produced more of the compound but did not become any more invasive. "There is strong circumstantial evidence that receptor-bound urokinase is important," Boyd declares.

Abbott researchers are inclined to agree, because certain highly malignant cell lines have large numbers of receptors for urokinase. Sasahara cites several recent studies that suggest a receptor



BREAST CANCER CELLS glow fluorescent yellow, indicating the presence of the clot-dissolving enzyme urokinase. Photo: Leo R. Zacharski.

count taken at the time of biopsy might help assess a cancer's ability to metastasize and so guide clinicians to better courses of therapy. By the same token, drugs that blocked those membrane receptors might retard metastasis.

But Richard M. Schultz, a biochemist at Loyola University Medical Center, discounts the importance of receptors. "My evidence is that urokinase secreted by the cells—and not taken up by them—can promote metastasis," he contends. Putting the human urokinase gene into mouse cancer cells dramatically promoted their metastatic ability, even though the human form does not bind to mouse receptors. Schultz believes cancer cells secrete urokinase to degrade extracellular matrices, not for their own internal use. "One can think of urokinase inhibitors, but I'm not sure if membrane receptor inhibitors would work," he says.

Some research suggests that urokinase's clot-dissolving capacity is only incidental to its true mode of action as a growth factor. The part of the molecule that attaches to the receptor closely resembles domains in proteins known to spur cell growth, observes David A. Hart of the University of Calgary. He suspects that urokinase production may be switched on by cancer-causing genes, called oncogenes, that use the enzyme as a tool for multiplication. Other oncogenes have been shown to stimulate different protein-cleaving enzymes.

One study that may clear up the confusion is being conducted by Leo R. Zacharski, a hematologist at the Veterans Administration Medical Center in White River Junction, Vt. He and his colleagues at Dartmouth College are categorizing various types of cancer by their ability to express urokinase and other blood-clotting factors. "Lo and behold, there is heterogeneity," he says. Certain cancers, such as prostate, breast and colon cancer, do produce urokinase and show no evidence of clot-promoting factors on the tumor cell. In contrast, other kinds, such as kidney cancer and malignant melanoma, lack urokinase and often show the presence of factors that encourage blood clots to form immediately adjacent to the tumor.

Early results indicate that giving urokinase to patients with clot-promoting cancers, such as small cell lung cancer, makes the tumors more susceptible to chemotherapy and the patient's own immune system, leading to more frequent regression, Zacharski declares. "In urokinase-expressing tumors, which seem to be the mirror image of small cell lung cancer, the idea would be to try to block urokinase," he explains.

Abbott's drugs to block urokinase are not yet ready for human trials, Zacharski sighs, but experimental models show that doing so reduces tumor growth and retards invasion. Studies of any urokinase inhibitors will have to consider the effects of temporarily depriving healthy cells of a vital enzyme, however. Boyd points out that immune system cells rely on the enzyme to get to sites of infection. "It remains to be seen whether patients can tolerate a little macrophage shutdown," he says.

Yet the prospect of a treatment window, perhaps at the time of surgery when cancer cells are sometimes inadvertently dispersed, brightens the outlook of many clinicians. "If you can somehow deprive cancer cells of something they need, they will be just as dead as if you had poisoned them with chemotherapy," Zacharski says. "This may be a more natural way of killing cancer." —*Deborah Erickson*

Fickle Fluids

Applications still elude fluids that solidify in an electric field

When development of electrorheological fluids began in earnest in the 1970s, analysts predicted they would soon become a billion-dollar technology. Because they can be transformed rapidly from a lubricant to a solid and back again by varying an electric field, developers envisioned a variety of applications in which electrical signals could be converted to mechanical work, such as automotive shock absorbers, engine mounts and suspension systems—even girders for buildings and bridges that would adjust to changing loads.

But dozens of companies in the U.S., the U.K., Germany and Japan have yet to find a way to convert electrorheological (ER) fluids into a single product that can compete widely and aggressively in the market. "At present, applications of ER fluids must be confined to small, specialized niche markets, areas in which the essence of the problem is high-speed response to electronic signals," says James E. Stangroom, one of the founders of ER Fluid Developments in England.

The problem, according to Stangroom and other industry experts, is that engineers have been forced to resort to complex and unconventional design principles in dealing with these fickle materials. Although ER fluids can have simple recipes—such as cornstarch in corn oil—they can display very complicated behavior.

All ER fluids are suspensions of dielectric particles in a dielectric fluid; today they typically consist of ceramic, polymer or graphite particles in silicone or mineral oil. When the fluid is placed between two plates of metal, it acts very much like a simple oil, and the plates can slide against each other. When an electric field is applied across the plates, the dielectric particles in the fluid line up, forming chains. As the field is increased, the chains grow in strength and number, increasing the resistance to the flow of the fluid. At some voltage, the fluid solidifies into a jellylike substance. The effect can be turned on and off more than 1,000 times a second.

Unfortunately, existing ER fluids do not function well outside a narrow range of conditions. The best fluids can resist sliding at a shear stress of 3,000 pascals in a field of 4,000 volts per millimeter. Most analysts believe the industry needs an ER fluid with 10 times that strength. Furthermore, most ER fluids operate over a temperature range of only 30 degrees Celsius. Yet the automotive industry requires the fluids to work from -40 to 150 degrees C. "In recent years, we have made fluids that are more reliable, more stable and stronger and that have better temperature properties," comments Theodore G. Duclos of Lord Corporation, which investigates ER fluids for use in automobiles. "There is still a little ways to go."

Stangroom believes engineers should concentrate on what can now be done with ER fluids. "Engineering would indeed be easy if we had ER fluids that were three times stronger," he says. "Of course, the same would be true if we had steel that was three times stronger." His firm has patented a clutch system based on ER fluids; it may be used for drawing thin wires because it can maintain a constant tension. The company has also produced a pump whose displacement can be varied continuously and quickly; the pump may be suited for hydrostatic drives for various vehicles.

Yet some researchers think ER fluid technology demands more than good engineering and better materials. Edward A. Collins, a consultant in the field of ER fluids who is co-heading a Department of Energy study of the technology, argues that universities should be encouraged to do more fundamental research. "Right now, we don't have the basic theories to explain many of the properties of the fluids," he says.

Frank E. Filisko of the University of Michigan, who contributed to the development of many new ER fluids, agrees. He tells the story of an engineer from the 1950s who wanted to try out the newly invented transistor. In a rush of excitement, he took apart his radio and replaced one of the vacuum tubes with a transistor. To his disappointment, the radio short-circuited. "Likewise, people are trying to apply ER fluids in conventional ways," Filisko remarks. "Instead they should be searching for new radical designs to take advantage of the fluid's dramatic properties." *—Russell Ruthen*

Machine Vision

New ways for the image makers to invent reality

You will likely be seeing much more of things your eyes won't believe. Police officers becoming the people they touch, arms that turn into swords and back again, cars changing into cheetahs and dancing men made of water—these are just some of the astonishing visual tricks already made possible by a computer graphics technique called "morphing." The term, which is derived from the word "metamorphosis," describes a method of smoothly dissolving between two images on film or videotape, so that one appears to become the other.

Entertainment insiders say that morphing has permanently altered the public's expectations of visual media. "When we first started doing this, it was just one effect for a movie," says Doug Smythe, who is senior technical director of computer graphics at Industrial Light and Magic (ILM) in San Rafael, Calif. "Now it's huge."

Morphing made its commercial debut in the 1987 film *Willow.* When director Ron Howard wanted a sorceress to change into different animal forms, he turned to ILM, the special effects studio owned by filmmaker George Lucas. "It was a call for guerrilla programming—just go in there and do it, damn what the neighbors or government thinks," Smythe recounts. He has since done morphing for several commercials and the film *Terminator 2.*

The basic technique had been demonstrated in crude fashion a few years before by Tom Brigham of the New York Institute of Technology. Smythe added grids to the screen to simplify decomposition of the images. The mosaic pieces could then be pulled and pushed, to match a left eye with a left eye, a nose with a nose—a procedure known as mesh warping. "It's a resampling process. The idea is, we have spline curves in horizontalish and verticalish directions," Smythe explains. The "ish" is because curves are allowed to bend and snake around images in either of the two images, as long as they do not overlap or touch. Samples of pixels are taken from along those curves in the so-called source image (the film as shot) and transferred to a destination image. Changing the rate of sampling produces more or less stretching. A transformation that lasts two seconds can take two weeks to accomplish.

Less exacting techniques have been developed, asserts Thaddeus Beier, a software engineer at Silicon Graphics Computer Systems in Mountain View, Calif. Details of the method he developed with Shawn Neely of Pacific Data Images (PDI) in Sunnyvale, Calif., will be revealed in a paper at SIGGRAPH, the annual computer graphics conference to be held in Chicago this July. "Rather than trying to specify explicitly how an image should change, we let the artist specify just by drawing features and then let the computer figure out what moves around," Beier explains. In his technique, lines added by the animator are surrounded by a field of influence that automatically warps the image a predetermined amount. For instance, drawing lines around the silhouette of a person's head pulls the hairline along without any extra work between the hair and the forehead.

"I think it's 1,000 times more efficient—you're concentrating on the artistic rather than the exasperating technical aspects," raves Jamie Dixon, digital effects supervisor for PDI in Hollywood. "It's a very different approach that lends itself to more dynamic morphing sequences," he says. Dixon employed the technique in Michael Jackson's acclaimed music video *Black or White,* seamlessly transforming singers into individuals of different races and Jackson himself into a panther.

MORPHING TECHNIQUES such as those used in **Terminator 2** *meld images for dazzling effects. Photo: ILM.*

The largest market for morphs achieved by any technique is likely to be the invisible manipulation of otherwise real images. Morphs could serve as Band-Aids, to join the best sections of an actor's cuts, or simply to remove an unwanted section of a picture. "We have the technology now to realize in film or video just about any vision," observes Smythe, who says he will stick with mesh warping. "We're not limited by what human beings can really do with their bodies—or gravity or any other physical laws." But then, was Hollywood ever? *—Deborah Erickson*

Doubling Up on Payoffs for Schooling

Once a year Twinsburg, Ohio, hosts a convention that has the rest of the town seeing double. The Annual Twins Day Festival is a magnet for almost 3,000 sets of twins (not to mention triplets, quadruplets and "mirror twins"). Almost everything shows up in pairs: there are double coffee mugs, twin birthday cards and two sit-down dinners for the look-alike celebrants. Last year there were even two economists.

For the economists, however, the trip to Twinsburg was strictly business. In addition to being a Mecca for twins, the yearly jamboree offers scientists of many stripes a seemingly ideal opportunity to separate the effects of nature from those of nurturing. "You can hold a lot of things constant," points out Alan B. Krueger, an economist from Princeton University. In particular, Krueger and his colleague (but not twin) Orley Ashenfelter aimed to cast light on a long-standing question: What is the value of education?

Liberal arts graduates may speak of broadened horizons and the life of the mind, but economists reduce the issues to a less subjective denominator: expected income. All other things being equal, how much does spending another year in school boost a person's future paycheck?

Economists—and common sense—have long maintained that more education is the path to greater fortune. Most estimates suggest that each additional year of school boosts income by about 8 to 10 percent. Beyond that generalization, however, lies much debate. For example, some studies suggest that women's income increases regularly for each year of schooling beyond high school, whereas the incremental gains for men are lower but jump disproportionately when they earn a diploma, reports Joseph G. Altonji, an economist at Northwestern University. And spending 10 years on a doctorate in early Saxon literature does not typically promise more income than devoting two years to a master's of business program.

What is more, other influences on success are not negligible. A child's home environment, inherent talents, interests, role models and social group all affect both completed years of schooling and future income. Trying to eliminate these factors statistically requires uncomfortable data contortions. Except, of course, in the case of twins—and particularly in the case of identical twins.

The first study of twins began as an instance of "pure scholarship," says Paul Taubman, an economist at the University of Pennsylvania. In the course of teaching a senior economics class in the late 1970s, Taubman recounted his research on the educational and economic achievements of World War II veterans. The students proposed repeating the study based on a sample of twins; Taubman said he doubted that the data existed. To his surprise, the students waded through military records and found a 2,000-person sample of twins, comprised of white men born between 1917 and 1927.

Along with his colleagues Jere R. Behrman, Zdenek Hrubec and Terence Wales, Taubman combed through the data and found some unexpected results. When evaluated individually, the men seemed to benefit from education at roughly the predicted rate—reaping about an 8 percent return for every year of completed schooling. But when the economists factored out "environmental" effects by looking at how fraternal twins fared, they found the return for an additional year of schooling had slipped to about 5 percent. Even more unsettling: when they looked at identical twins, who have presumably the same genetic makeup (and thus, perhaps, the same inherent abilities), the real returns for education were only about 3 percent. Such findings seem to indicate that environmental and inherent abilities play a larger role than education.

All other things being equal, how much will another year in school boost a person's paycheck?

Their data, nonetheless, were still plagued by inconsistencies. For instance, many people tend to forget—or misrepresent—how much time they spent in school. Because of the nature of the statistical analysis, even small inaccuracies in such data can dramatically warp the results.

Last year Ashenfelter and Krueger decided to reinvestigate the education of twins—this time by undertaking a survey themselves. The Twinsburg festival was the obvious place for such a task. "Where else can you find twin nuns and twin accountants?" Krueger asks. Over the course of the three-day gathering, the economists (and several assistants) interviewed some 250 pairs of twins. To reduce errors in establishing the records, they separately asked each member of a pair how many years of schooling he (or she) and his (or her) sibling had achieved.

The results indicate that economists may have badly underestimated the affect schooling may have on income. According to Ashenfelter and Krueger, each additional year of schooling completed by one twin but not the other boosted the wages of the more educated one by about 17 percent. A twin who completed college would earn nearly twice as much as an identical sibling with a high school diploma.

That said, the economists concede the data may have their own bugs. Those who attend twins festivals tend to be better educated and more affluent than the average population. Furthermore, more than half the twins surveyed had spent the same number of years in school.

Taubman proposes another reason for the different results: higher education may be more valuable today than ever before. Even though the supply of college graduates has gone up, demand for more sophisticated skills may have gone up even faster, he suggests. Alternatively, there may be even fewer—and worse-paying—jobs available for those without college degrees. Some studies have shown that education pays off more in industries that have high rates of technological progress, Altonji points out.

Like any other group of scientists, economists need to do far more research to build reliable collections of data. Only then will their work offer insights into complex social issues such as the value of education. Ashenfelter and Krueger have learned that lesson well: this August will find them back in Twinsburg, looking to double their data.

—*Elizabeth Corcoran and Paul Wallich*

Survival of the Fittest Bits

In Darwinian terms, life is a struggle in which only the fittest survive to reproduce. What has proved successful to life is also useful to problem solving on a computer. In particular, programs using so-called genetic algorithms find solutions by applying the strategies of natural selection. The algorithm first advances possible answers. Then, like biological organisms, the solutions "cross over," or exchange "genes," with every generation. Sometimes there are mutations, or changes, in the genes. Only the "fittest" solutions survive so that they can reproduce to create even better answers [see "Genetic Algorithms," by John Holland; page 66].

One way to understand how genetic algorithms work is to implement a simple one and use it in some experiments. I chose the C language for my algorithm, but any other computer language would do as well. It took me two days to complete the version described here. I spent the most time by far writing the utilities to make the program easy to run, to change parameter settings and to display the population and statistics about its performance.

I recommend developing a genetic algorithm in two steps. First, write utilities (some are described below) to track the population's average fitness and the most fit members of each generation. Your version should generate random populations of individuals and have subroutines to assign fitness. The program should also display the individuals and their fitnesses, sort by fitness and calculate the average fitness of the population. Second, add subroutines to implement the genetic algorithm itself. That is, your program should select parents from the current population and modify the offspring of those parents through crossover and mutation.

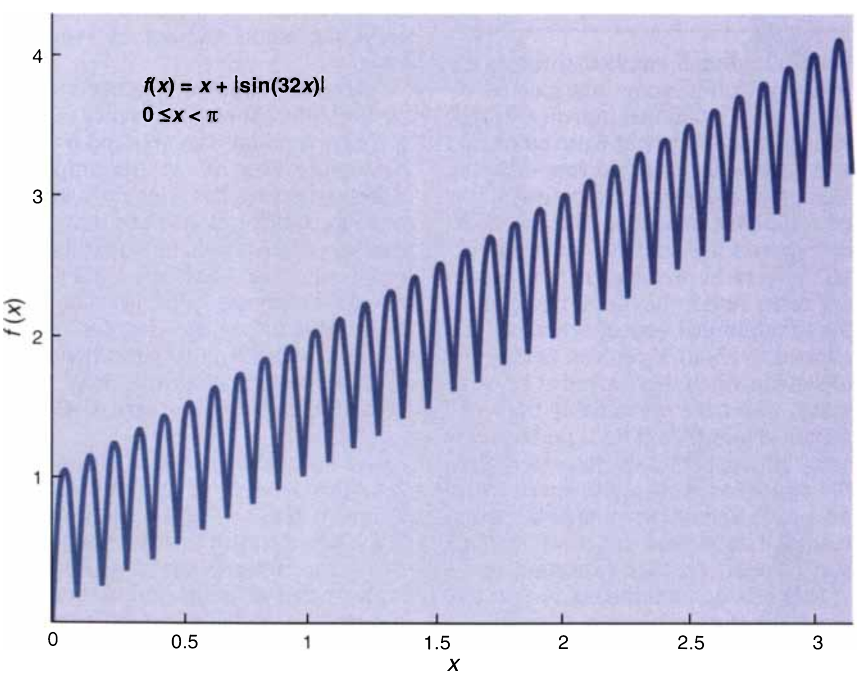
In writing a genetic algorithm for the first time, I found it best to use a simple problem for which I know the answer. This approach made it easier to debug and to understand the algorithm. Adjusting it for more complex problems is not too difficult.

I chose the problem of finding the optimum value of the function $f(x) = x + |\sin(32x)|$. I limited the values of x (the domain) to between 0 and π (I am working in radians, where $\pi = 180$ degrees.) Thus, this function has 32 regular oscillations added to the straight line $f(x) = x$ [*see illustration below*]. Because $f(x)$ is always positive, it can be used directly as a measure of fitness. The goal is to find an individual string (a value of x) that yields the largest $f(x)$.

The first step is deciding how to represent structures from the domain in a form your genetic algorithm can process. I represented the values of x as binary strings, as described in Holland's article. For my initial experiments, I chose the length, L, of my strings to be 16 bits ($L = 16$). I thus have $2^L = 2^{16} = 65{,}536$ possible values of x. (Increasing the length of the string improves the precision of the solution.) I assigned these 2^L string values to values of x: I let the string 0000000000000000 represent $x = 0$, 0000000000000001 to be $x = 1(\pi/2^L)$, 0000000000000010 to be $x = 2(\pi/2^L)$, and so on, all the way to 1111111111111111, which represents the maximum value of the domain, $x = (2^L - 1)(\pi/2^L) = \pi - \pi/2^L$. (The value of π itself is excluded because there is no string left to represent it.) To test your intuition, what value of x does the string 1000000000000000 represent?

In short, the binary strings represent 2^L values of x that are equally spaced between 0 and π. Hence, as the algorithm searches the space of binary strings, it is in effect searching through the possible values of x. The C fragment on page 116 (the mapping algorithm) shows one way to implement this mapping from a binary string, which is stored as a char array in C. For example, it will map 1000000000000000 onto (approximately) $\pi/2$. The fitness of an individual binary string S is just $f(x)$, where x = MapStringToX(S).

Given this mapping from binary strings to the domain of $f(x)$, it is instructive to map schemata, or target regions, as specified by Holland. For



FITTEST SOLUTION of f(x) = x + |sin (32x)| *is the value of* x *that yields the highest* f(x). *The answer is* x = 3.09346401 *(in radians), represented by 1111110000010100.*

RICK L. RIOLO is a researcher at the University of Michigan at Ann Arbor, where he received his Ph.D. in computer science. He thanks Robert Axelrod, Arthur Burks, Michael Cohen, John Holland, Melanie Mitchell and Jim Sterken for comments and suggestions.

example, the schema 0*************** consists of all the strings that start with 0. The * is the "don't care" placeholder: it does not matter whether it is 0 or 1. The schema 0*************** maps into the first half of the domain—that is, between 0 and $\pi/2$. The schema 1*************** corresponds to the other half, between $\pi/2$ and π. For the function $f(x)$, the average value of $f(x)$ over 1*************** is greater than that over 0***************. Thus, in a random population of strings, you should find that the average fitness of individuals in the former region is greater than that of members in the latter.

My implementation includes some subroutines that track individuals in any schemata specified by the user (for example, all strings that start with 1 or all that end in 001 and so forth). At each generation, the program will count the number of individuals in the population that are members of each region. It will also calculate the average fitness of those individuals, thus explicitly estimating the fitness of the regions.

To further hone your intuition, try mapping some other schemata into the domain x and examine the random individuals that fall into those regions. Can you specify schemata that partition each oscillation in $f(x)$ into eight contiguous regions? What is the average fitness of individuals in each region?

Once you are confident the first version of your program is working, you should add the subroutines that implement the heart of the genetic algorithm. The algorithm I used appears on the next page.

The EvaluatePopulation subroutine simply loops over all the individuals in the population, assigning each an $f(x)$ value as a measure of fitness. The DoTournamentSelection subroutine conducts a series of tournaments between randomly selected individuals and usually, but not always, copies the fittest individual of the pair (i,j) to NewPop. A sample tournament selection algorithm appears on the next page. In the algorithm, URand01 returns a (uniformly distributed) random value between 0 and 1. The net effect is that more fit individuals tend to have more offspring in the new population than do less fit ones. You may find it useful, nonetheless, to copy the fittest individual into the new populations every generation.

Note that if URand01 is greater than 0.75, the selective pressure will be greater (it will be harder for less fit strings to be copied into NewPop), whereas a smaller value will yield less selective pressure. If 0.5 is used, there will be no selective pressure at all.

The ModifyPopulation subroutine should randomly perform the crossover operation on some of the pairs in NewPop; a typical crossover rate is 0.75 (that is, cross 75 percent of the pairs). I used the "single point" crossover operation described in Holland's article. My version also mutates a small, randomly selected number of bits in the individual strings, changing the value from 1 to 0, or vice versa. A typical mutation rate is 0.005 per bit—in other words, each bit has a 0.005 chance of being mutated.

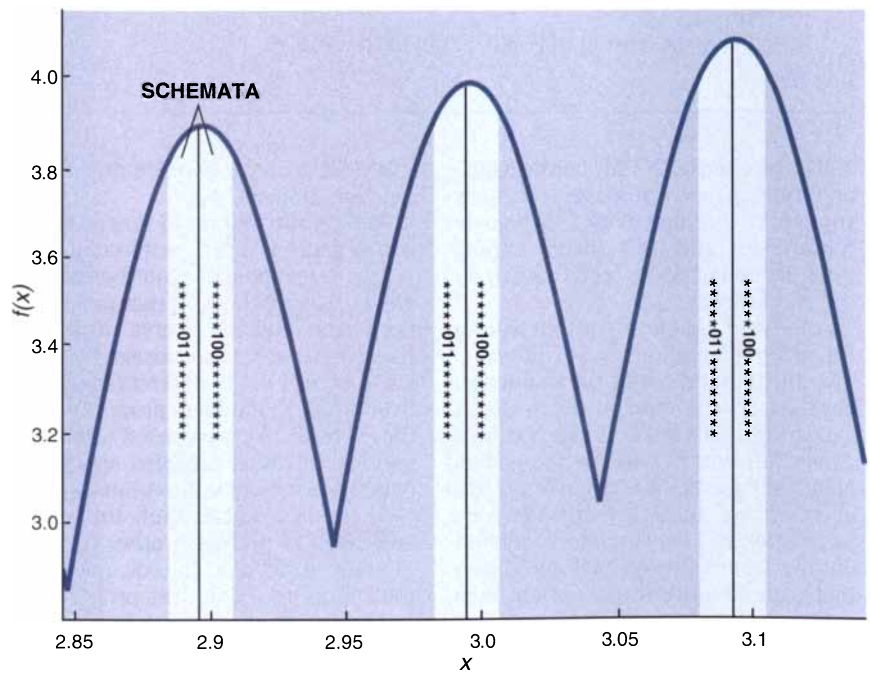
Once the implementation is complete, you can begin to run experiments. Try starting the system with a random population of 200 individuals and run it for 50 generations, collecting statistics on how the average fitness changes and noting the highest individual fitness found at each generation. You might compare the results you get with different population sizes and with crossover and mutation rates. I found it best to compare average results from several runs started with different seeds of the random number generator.

One especially useful experiment for understanding the dynamics of how populations evolve is to track the individuals in the eight regions that partition the 32 oscillations of $f(x)$, to which I referred earlier. The regions are defined by the schemata *****000********, *****001******** and so on, to *****111********. Notice that *****011******** and *****100******** occupy the center part of each peak, which includes the points of highest fitness.

In an initially random population, there should be approximately equal numbers of individuals in each of the eight schemata. As the genetic algorithm runs, you should see, on average, more individuals in the two center schemata than in the other six. This result happens even though many members of the center schemata map into low values of x and so have low fitnesses. The reason is that the average fitness in the center schemata is higher than that in the others.

Using these eight schemata, I tried to track how well the algorithm can bias its search for individuals of higher fitness. I set up my program to record at each generation the number of individuals in each schema, the average fitness of the individuals in those regions and the average fitness for the population as a whole. The theorems mentioned in Holland's article predict that the number of individuals in regions with an estimated average fitness above the population average should increase and the number of those in regions having below-average fitness should decrease.

For a population of 400, a crossover rate of 0.75 and a mutation rate of 0.001, the program confirmed the predictions about 67 percent of the time. I found that the greater the mutation and crossover rates and the smaller the population size, the less frequently the predictions were confirmed. These re-



TWO SCHEMATA occupy the central regions of each peak in f(x)*. They include the points of highest fitness. Only the last three peaks are illustrated.*

Programming Fragments

MAPPING ALGORITHM—allows binary strings to represent values of *x*
(Note: the compiler defines M_PI as π)

```
double scalefactor = (double) M_PI / pow ((double) 2.0,(double) L);
double MapStringToX ( char *S )
{ /* put bits from S in rightmost end of r and move them left */
  int i, r;
  for ( i = r = 0; i < L; ++i, ++S ) {
    r << = 1;      /* shift bits left, place 0 in right bit */
    if ( *S == '1' ) /* if S has a 1, then... */
      ++r;         /* change rightmost bit to 1 */
  }
  return ( (double) r * scalefactor );
}
```

BASIC GENETIC ALGORITHM—performs crossover, mutation and selection operations repeatedly

```
/* create structure to store the populations */
struct PopStr {
  char   Genome[L + 1]; double  Fitness;
} Pop[PopSize], NewPop[PopSize];
/* implement algorithm */
GenerateRandomPopulation ( Pop );
EvaluatePopulation ( Pop );
while ( Not_Done ) {
  DoTournamentSelection ( Pop, NewPop );
  ModifyPopulation ( NewPop );
  Switch ( Pop, NewPop );
  EvaluatePopulation ( Pop );
}
```

TOURNAMENT SELECTION ALGORITHM—selects for and copies the more fit individuals into the next generation

```
while ( NewPopNotFull ) {
  /* pick two individuals from Pop at random */
  i = random () % PopSize; j = random () % PopSize;
  /* return a random value between 0 and 1 */
  if ( URand01 () < 0.75 )
    copy the most fit of Pop[i], Pop[j] to NewPop
  else
    copy the least fit of Pop[i], Pop[j] to NewPop
}
```

sults make sense, I think, because higher mutation and crossover rates disrupt good "building blocks" (schemata) more often, and, for smaller populations, sampling errors tend to wash out the predictions.

After your genetic algorithm has run for many generations, you will probably notice that most of the population consists of very similar, if not identical, strings. This result is called convergence and occurs because the genetic algorithm pushes the population into ever narrower target regions. Often the population will converge to an individual that is not the optimal one. Essentially, the algorithm has gotten stuck on a local optimum. Convergence can be particularly debilitating, because it means that crossover will not contribute much to the search for better individuals. Crossing two identical strings will yield the same two strings, so nothing new happens.

Finding ways to avoid inappropriate convergence is a very active area of research. One possible mechanism involves biasing the selection process to keep the population diverse. In essence, these mechanisms encourage individuals to occupy many different regions in the domain in numbers proportional to the average value associated with those regions, much as different species inhabit niches in natural evolution.

Of course, you can apply the genetic algorithm to problems other than optimizing functions. It took me about four hours to modify the program described here to create a version in which the individuals represent strategies for playing a form of the Prisoner's Dilemma. In this game, two prisoners face a choice: cooperate with the other (by remaining silent about the crime) and receive a three-year sentence, or try to get off free by "defecting," or squealing. Unfortunately, you do not know what the other player is going to do: if you cooperate and he defects, then you get a 10-year sentence and he goes free. If you both defect, then you both get seven years.

My modified genetic algorithm explored the strategies involved in playing the game many times with the same opponent. Each individual's binary string represents two numbers (p, q), where p is the probability the individual will cooperate if the opponent cooperated on the previous play and q is the probability of cooperation if the opponent had defected. For instance, the strategy (1,0) is tit for tat, and (0.1,0.1) is approximately the "always defect" approach.

The genetic algorithm generates some very interesting dynamics in the evolution of populations of such strategies, especially if each individual's binary string also contains some bits that specify an arbitrary "tag." Another set of bits could then indicate the propensity of the individual to play only with those that have a similarly tagged string. In effect, the population can evolve individuals that recognize external "markings" (the tags) associated with tit-for-tat players. Thus, they can prosper by looking for cooperative individuals and by avoiding defectors.

From relatively simple gene structures, complicated evolutionary dynamics may emerge. For instance, mimicry and other forms of deception may evolve: a defector could have a tag associated with cooperators. With some imagination, you can construct other simple gene structures and use the genetic algorithm to coax complex evolution out of your computer.

For a copy of the optimization program (written in C language), please write to: The Amateur Scientist, Scientific American, 415 Madison Avenue, New York, NY 10017-1111.

FURTHER READING

A COMPARATIVE ANALYSIS OF SELECTION SCHEMES USED IN GENETIC ALGORITHMS. David E. Goldberg and Kalyanmoy Deb in *Foundations of Genetic Algorithms.* Edited by Gregory J. E. Rawlins. Morgan Kaufmann, 1991.

HANDBOOK OF GENETIC ALGORITHMS. Edited by Lawrence Davis. Van Nostrand Reinhold, 1991.

TIT FOR TAT IN HETEROGENEOUS POPULATIONS. Martin A. Nowak and Karl Sigmund in *Nature*, Vol. 355, No. 6357, pages 250–253; January 16, 1992.

No other system of keeping up can compare with ours.

YOUR SYSTEM: a time-consuming, futile struggle to keep up with the information explosion.

The classic texts are convenient references—but the information they contain is obsolete before publication.

Like many physicians, you probably rely on the texts you first used in medical school. But even using the most recent editions, you find material that no longer reflects current clinical thinking—and they lack the latest information on such topics as herpes, oncogenes, AIDS, and photon imaging.

Reading stacks of journals alerts you to recent developments—but can't give you quick answers on patient management.

Struggling through the hundreds of journal pages published each month—even on only the really significant advances in the field—is arduous and memory-taxing. And it's a task that costs physicians valuable time—their most precious resource.

Review courses cover clinical advances—but, months later, do you recall the details of a new procedure or unfamiliar drug?

Seminars can also be costly and make you lose valuable time away from your practice—expenses that may amount to several thousand dollars. And, the speaker's skill often determines how much you learn.

OUR SYSTEM: a rewarding, efficient way to keep yourself up-to-date—and save hundreds of hours of your time for patient care.

A comprehensive, 2,300-page text in three loose-leaf volumes, incorporating the latest advances in medical practice as of the month you subscribe.

This superbly designed, heavily illustrated resource, called "the best written of all [the internal medicine] books" by *JAMA* (251:807, 1984), provides a practical, comprehensive description of patient care in 15 subspecialties. And, because the text is updated each month, the clinical recommendations reflect all the current findings. A practice-oriented index and bibliography of recent articles further enhance the efficiency of the text.

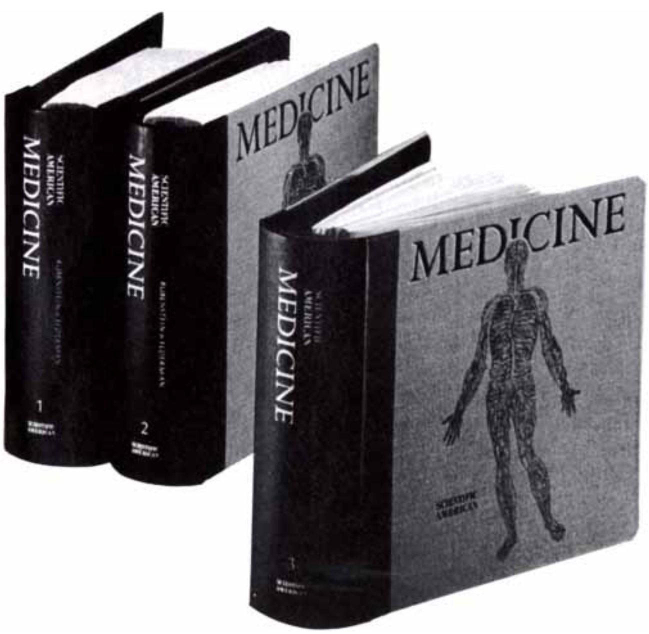
Each month, six to nine replacement chapters to update your text *plus* new references, a four- to five-page news bulletin, and a completely new index.

You'd have to read hundreds of journal pages each month—and memorize the contents—to get the same information SCIENTIFIC AMERICAN *Medicine* contains. With our updated text, you read only the information you really need. Our authors, largely from Harvard and Stanford, sort through the literature and monitor developments, incorporating the significant advances into our chapters.

At a nominal cost, a 32-credit CME program, to save you valuable patient-care time and the expense of attending review courses.

Earn 32 Category 1 or prescribed credits per year with our convenient self-study patient management problems; each simulates a real-life clinical situation. Choose either the complimentary printed version or, at a modest extra charge, the disk version for your IBM® PC or PS/2, Macintosh™, or Apple® (or compatible).

Your subscription includes: the three-volume, 2,300-page loose-leaf text and, each month, six to nine replacement chapters, a newsletter, new references, and a completely revised index.



☐ Yes, I'd like to try the SCIENTIFIC AMERICAN *Medicine* system. Please enter my subscription at a first-year price of US$285* plus $7 shipping for a total of US$292.

☐ Also enroll me in the CME program. I'd prefer:
- ☐ the printed version at US$19.95 (U.S. and Canada only; foreign US$97.95 + US$8 shipping)
- ☐ the disk version at US$102* additional. Computer type or compatible:
  - ☐ IBM® ☐ 5¼" ☐ 3½"
  - ☐ Macintosh™ (512K) ☐ Apple® IIc/IIe/IIGS/II+

☐ Enroll me in the disk CME only at US$159.* (Note model above.)

☐ Check enclosed.* ☐ Bill me ☐ VISA ☐ MasterCard

Exp. Date______ Acct. No.______

Signature______ Specialty______

Name______

Address______

City______ State______ Zip Code______

*Add sales tax for CA, DC, IA, IL, MA, MI, NY, UT or Canada (Canadian GST #R127387652). Canada: add US$10 shipping, payment in U.S. funds or equivalent. Foreign: add US$22 shipping, U.S. funds only. Apple®, IBM®, and Macintosh™ are registered trademarks of their respective companies.

Or, call toll free 1 800 545-0554

SCIENTIFIC AMERICAN MEDICINE

415 Madison Avenue, New York, NY 10017 212 754-0550

V9125

 by Stephen Jay Gould

Impeaching a Self-Appointed Judge

DARWIN ON TRIAL, by Phillip E. Johnson. Washington, D.C., Regnery Gateway, 1991 ($19.95).

I teach a course at Harvard with philosopher Robert Nozick and lawyer Alan Dershowitz. We take major issues engaged by each of our professions—from abortion to racism to right-to-die—and we try to explore and integrate our various approaches. We raise many questions and reach no solutions.

Clearly, I believe in this interdisciplinary exercise, and I accept the enlightenment that intelligent outsiders can bring to the puzzles of a discipline. The differences in approach are so fascinating—and each valid in its own realm. Philosophers will dissect the logic of an argument, an exercise devoid of empirical content, well past the point of glaze over scientific eyes (and here I blame scientists for their parochiality, for all the world's empirics cannot save an argument falsely formulated). Lawyers face a still different problem that makes their enterprise even more divergent from science—and for two major reasons.

First, the law must reach a decision even when insufficient evidence exists for confident judgment. (Scientists often err in the opposite direction of overcaution, even when the evidence is compelling, if not watertight.) Thus, in capital cases, the law must free a probably guilty man whose malfeasance cannot be proved beyond a doubt (a moral principle that seems admirable to me but would not work well in science). We operate with probabilities; the law must often traffic in absolutes.

Second, there is no "natural law" waiting to be discovered "out there" (*pace* Clarence Thomas in his recent testimony). Legal systems are human inventions, based on a history of human thought and practice. Consequently, the law gives decisive weight to the history of its own development—hence the rule of precedent in deciding cases. Scientists work in an opposite way: we search continually for new signals from nature to invalidate a history of past argument. (As a sometime historian of science, I wish that scientists, like lawyers, would pay more attention to, and have more reverence for, their pasts—but I understand why this is not likely to happen.)

DARWIN'S LAST BOOK, on worms, received this commentary in the October 1881* Punch. *(The illustration does not appear in the book under review.)

Phillip E. Johnson is a law professor at Berkeley and "a philosophical theist and a Christian" who strongly believes in "a Creator who plays an active role in worldly affairs." His book has received great "play" in print and television, largely (I suppose) because such unconventional products rarely emanate from the symbolic home of California "flowerpower." The press loves an oddity. This publicity is certainly no measure of the book's merit, as I shall argue. Now, I most emphatically do not claim that a lawyer shouldn't poke his nose into our domain; nor do I hold that an attorney couldn't write a good book about evolution. A law professor might well compose a classic about the rhetoric and style of evolutionary discourse; subtlety of argument, after all, is a lawyer's business. But, to be useful in this way, a lawyer would have to understand and use our norms and rules, or at least tell us where we err in our procedures; he cannot simply trot out some applicable criteria from his own world and falsely condemn us from a mixture of ignorance and inappropriateness. Johnson, unfortunately, has taken the low road in writing a very bad book entitled *Darwin on Trial*.

In a "classic" of antievolutionary literature from the generation just past, lawyer Norman Macbeth (1971) wrote a much better book from the same standpoint, entitled *Darwin Retried* (titles are not subject to copyright). Macbeth ultimately failed (though he raised some disturbing points along the way) because he used an inappropriate legal criterion: the defendant (an opponent of evolution) is accused by the scientific establishment and must be acquitted if the faintest shadow of doubt can be raised against Darwinism. (As science is not a discipline that claims to establish certainty, all its conclusions would fall by this inappropriate procedure.)

Johnson's current incarnation of this false strategy, *Darwin on Trial,* hardly deserves to be called a book at all. It is, at best, a long magazine article promoted to hard covers—a clumsy, repetitious abstract argument with no weighing of evidence, no careful reading of literature on all sides, no full citation of sources (the book does not even contain a bibliography) and occasional use of scientific literature only to score rhetorical points. I see no evidence that Johnson has ever visited a scientist's laboratory, has any concept of quotidian work in the field or has read widely beyond writing for nonspecialists and the most "newsworthy" of professional claims.

The book, in short, is full of errors, badly argued, based on false criteria, and abysmally written. Didn't anyone ever teach Johnson not to end chapters with "announcement sentences" or to begin subsequent sections with summaries? Chapter 6, for example, ends with a real zinger: "We will look at that claim in the next chapter." The very next chapter then begins with the maximally lively: "Before we try to get any answers out of the molecular evidence, we had better review where we stand." Mrs. McInerney, my tough but beloved third-grade teacher, would have rapped his knuckles sore for such a construction, used by Johnson at almost every chapter transition.

Johnson is not a "scientific creationist" of Duane Gish's ilk—the "young earth" Biblical literalists who have caused so much political trouble of late, but whom we beat in the Supreme Court in 1987. He accepts the earth's great age and allows that God may have chosen to work via natural selection and other evolutionary principles (though He may also operate by miraculous intervention if and when He chooses). Johnson encapsulates his major insistence by writing: "In the broadest sense, a 'creationist' is simply a person who believes that the world (and especially mankind) was *designed,* and exists for a *purpose.*" Darwinism, Johnson claims, inherently and explicitly denies such a belief and therefore constitutes a naturalistic philosophy intrinsically opposed to religion.

But this is the oldest canard and non sequitur in the debater's book. To say it for all my colleagues and for the umpteenth millionth time (from college bull sessions to learned treatises): science simply cannot (by its legitimate methods) adjudicate the issue of God's possible superintendence of nature. We neither affirm nor deny it; we simply can't comment on it as scientists. If some of our crowd have made untoward statements claiming that Darwinism disproves God, then I will find Mrs. McInerney and have their knuckles rapped for it (as long as she can equally treat those members of our crowd who have argued that Darwinism must be God's method of action). Science can work only with naturalistic explanations; it can neither affirm nor deny other types of actors (like God) in other spheres (the moral realm, for example).

Forget philosophy for a moment; the simple empirics of the past hundred years should suffice. Darwin himself was agnostic (having lost his religious beliefs upon the tragic death of his favorite daughter), but the great American botanist Asa Gray, who favored natural selection and wrote a book entitled *Darwiniana,* was a devout Christian. Move forward 50 years: Charles D. Walcott, discoverer of the Burgess Shale fossils, was a convinced Darwinian and an equally firm Christian, who believed that God had ordained natural selection to construct a history of life according to His plans and purposes. Move on another 50 years to the two greatest evolutionists of our generation: G. G. Simpson was a humanistic agnostic, Theodosius Dobzhansky a believing Russian Orthodox. Either half my colleagues are enormously stupid, or else the science of Darwinism is fully compatible with conventional religious beliefs—and equally compatible with atheism, thus proving that the two great realms of nature's factuality and the source of human morality do not strongly overlap.

But Johnson's major premise—the inherent Godlessness of Darwinism—could be wrong, and he might still have a good argument for the major thrust of his text: the attempt to show that Darwinism is a dogma, unsupported by substantial and meaningful evidence, and propped up by false logic. But here he fails utterly, almost comically (Macbeth's 1971 book is much better).

Johnson's line of argument collapses in two major ways, the second more serious than the first. I feel a bit more forgiveness in this first category—familiarity with the facts of biology—because the field is immense and alien to Johnson's training. Still, the density of simple error is so high that I must question wider competence when attempts at extension yield such poor results. To cite just a few examples from the compendium of Johnson's factual and terminological errors: On page 16, he claims that all immediate variation for natural selection comes from mutation: "Darwinian evolution postulates two elements. The first is what Darwin called 'variation,' and what scientists today call *mutation*." He then realizes that he has neglected sexual recombination, the vastly predominant source of immediate variation in sexual species, but he makes his error worse by including recombination as a category of mutation. On page 30, he reports that "sexual selection is a relatively minor component in Darwinist theory today." But sexual selection is perhaps the hottest Darwinian topic of the past decade, subject of at least a dozen books (which Johnson has neither noted nor read—a sure sign of his unfamiliarity with current thinking in evolutionary theory). On page 41, he states that polyploidy (as a result of doubling of chromosomes) can occur only in "hermaphrodite species capable of self-fertilization"—and therefore can play little role in major change (for self-doubling does not yield markedly new qualities). But the evolutionarily potent form of polyploidy is not the autoploidy that he equates with the entire phenomenon, but alloploidy, or doubling of both male and female components after fertilization with pollen of a different species.

On page 60, he calls the German paleontologist Otto Schindewolf a saltationist, whereas Schindewolf's subtle theory contained a central element of insensible change in a process that he called proterogenesis (gradual seepage of juvenile traits into adult stages). Schindewolf spent most of his career studying small and continuous changes in ammonite suture patterns. On page 103, Johnson raises the old chestnut against a natural origin of earthly life by arguing: "the possibility that such a complex entity could assemble itself by chance is fantastically unlikely." Sure, and no scientist has used that argument for 20 years, now that we understand so much more about the self-organizing properties of molecules and other physical systems. The list goes on.

Second, and more important for documenting Johnson's inadequacy in his own realm of expertise, he performs abysmally in the lawyer's domain of the art of argument. To begin, he simply does not grasp (or chooses not to understand) the purpose and logic of evolutionary argument. I have already illustrated his central conflation of Darwinism with hostility to religion. I was particularly offended by his false and unkind accusation that scientists are being dishonest when they claim equal respect for science and religion: "Scientific naturalists do not see a contradiction, because they never meant that the realms of science and religion are of equal dignity and importance. Science for them is the realm of objective knowledge; religion is a matter of subjective belief. The two should not conflict because a rational person always prefers objective knowledge to subjective belief." Speak for yourself, Attorney Johnson. I regard the two as of equal dignity and limited contact. "The two should not conflict," because science treats factual reality, while religion struggles with human morality. I do not view moral argument as a whit less important than factual investigation.

Johnson then upholds the narrow and blinkered caricature of science as experiment and immediate observation only. Doesn't he realize that all historical science, not just evolution, would disappear by his silly restriction? Darwin, he writes, "described *The Origin of Species* as 'one long argument,' and the point of the argument was that the common ancestry thesis was so logically appealing that rigorous empirical testing was not required. He proposed no daring experimental tests, and thereby started his science on the wrong road." But Darwin spent 20 years collecting facts for evolution. The *Origin* is one long compendium of observations and empirical confirmations. To be sure, Darwin's method is not generally experimental, for singular and complex past events are not so explained by any historical science. Darwin thought long and hard about proper methodology of confirmation for historical science and used Whewell's "consilience of induction," or bringing of widely disparate information under a uniquely consistent explanation. Darwin wrote of his method in 1868: "This hypothesis may be tested...by trying whether it explains several large and independent classes of facts; such as the geological succession of organic beings, their distribution in past and present times, and their mutual affinities and homologies."

Not only does Johnson misconstrue the basic principles of our science (as I have shown), but he also fails to present cogent arguments in his own brief as well. His development of a case is fatally marred by three pervasive techniques of careless or unfair discourse.

First, omissions that unjustly castigate a person or a claim. On page 5, Johnson recounts the tale of H. F. Osborn and his error in identifying a pig tooth as a human ancestor: "Osborn prominently featured 'Nebraska Man'... in his antifundamentalist newspaper articles and radio broadcasts, until the tooth was discovered to be from a peccary." True, but who made the correction? Although Johnson does not tell us, the answer is H. F. Osborn, who properly tested his claim by mounting further collecting expeditions, discovering his error and correcting it—in other words, science working at its best.

On page 74, in his lick-and-promise tour through the history of vertebrates, we learn that no intermediary has ever been discovered between rhipidistian fishes and early amphibians. Yet Johnson never mentions the first amphibians, *Ichthyostega* and *Acanthostega* (featured in all paleontological texts) with their conserved features of a fishy past: small tail fins, lateral line systems, and six to eight digits on each limb. On page 76, he admits my own claim for intermediacy in the defining anatomical transition between reptiles and mammals: passage of the reptilian jaw-joint bones into the mammalian middle ear. Trying to turn clear defeat to advantage, he writes: "We may concede Gould's narrow point." Narrow indeed; what more does he want? Then we find out: "On the other hand, there are many important features by which mammals differ from reptiles besides the jaw and ear bones, including the all-important reproductive systems." Now how am I supposed to uncover fossil evidence of hair, lactation and live birth? A profession finds the very best evidence it could, in exactly the predicted form and time, and a lawyer still tries to impeach us by rhetorical trickery. No wonder lawyer jokes are so popular in our culture.

Second, consider Johnson's false use of synecdoche. The art of having an item or part stand for the whole is a noble trope in poetry and *the* classical, unfair trick of debate. Professions are big, and everyone makes a stupid statement now and again. As an honorable opponent, you cannot use a single dumb argument to characterize an entire field. Yet Johnson does so again and again—and this, I suppose, represents the legal tactic of "poke any hole and win acquittal." Thus, Johnson quotes a few ill-informed statements, representing opposite extremes around a golden mean held by nearly every evolutionist—that natural selection is either meaningless as a tautology or necessarily and encompassingly true as an a priori universal principle. Now both claims have been advanced, but they are held by tiny minorities and unsustained by any strong or enduring argument. The principle of natural selection does not collapse because a few individuals fall into fallacies from opposite sides of claiming too little or too much. Similarly, the consensus that science and religion are separate and equally valuable is not brought down by the fact that Julian Huxley unites them on one side, while Will Provine holds that science implies atheism on the other. Minorities are not necessarily wrong (or science would never advance), but only the cogency of their data and arguments, not the mere fact of their existence, brings down old theories.

As his third trick, Johnson continues to castigate evolutionists for old and acknowledged errors. T. H. Huxley, paraphrasing Dryden's famous line about Alexander the Great's drunken boasting, stated that life is too short to occupy oneself with the slaying of the slain more than once. In law, the illogicality of an important precedent might bring down a current structure like a house of cards. But in science, a bad old argument is just a superfluous fossil. Nothing is gained by exposing a 30-year-old error—save the obvious point that science improves by correcting its past mistakes. Yet Johnson continually tilts at such rotted windmills. He attacks Simpson's data from the 1950s on mammalian polyphyly (while we have all accepted the data of mammalian monophyly for at least 15 years). He quotes Ernst Mayr from 1963, denying neutrality of genes in principle. But much has changed in 30 years, and Mayr is as active as ever at age 87. Why not ask him what he thinks now?

Johnson's grandiose claims, backed by such poor support in fact and argument, recall a variety of phrases from a mutually favorite source: "He that troubleth his own house shall inherit the wind" (Proverbs 11:29, and source for the famous play that dramatized the Scopes trial); "They have sown the wind, and they shall reap the whirlwind" (Hosea 8:7). But *Darwin on Trial* just isn't good enough to merit such worrisome retorts. The book is scarcely more than an acrid little puff—and I therefore close with a famous line from Darwin's soulmate, born on the same day of February 12, 1809. Abraham Lincoln wrote: "'And this, too, shall pass away.' How much it expresses! How chastening in the hour of pride! How consoling in the depths of affliction!"

STEPHEN JAY GOULD teaches biology, geology and history of science at Harvard University.

Women and the Discourse of Science

To make the point of male bias in language, a computer magazine recently substituted *she* for *he* throughout one issue as its standard third-person pronoun. Impassioned responses poured in, ranging from those appreciative of the reverse sexism to those canceling their subscriptions. ("Dear Editor," one irate reader responded, "I... suggest [you] adopt the subtitle, *A Magazine for High Tech Women and Eunuchs.*")

Indeed, passion runs high among the normally dispassionate when it comes to women and the language of science. True, generic *he* (*he* used for all humans, rather than specifically for males) has disappeared from most scientific publications, after a fierce fight. But in a domain as traditionally masculine as science, that's only a beginning. Further work awaits the linguistic reformer, for in this hard-fought battle, pronouns are but prelude. The discourse of science is soaked in testosterone.

Consider, for instance, the historical naming of science as male and nature as female. Together the couple forms a deeply embedded, central metaphor of scientific discovery, one that is a favorite of Francis Bacon. In it, the scientist, always a bold, rational he, pursues nature, inevitably a passive, mysterious she.

"If any man there be who, not content to rest in and use the knowledge which has already been discovered, aspires to penetrate further... to seek... certain and demonstrable knowledge," Bacon says in *Novum Organum,* "I invite all such to join themselves, as true sons of knowledge, with me, that passing by the outer courts of nature... we may find a way at length into her inner chambers."

Bacon regularly summons his sons of knowledge to an aggressive male wooing of female nature's secrets. When nature proves approachable, stern science treats her well, deciphering her mysteries and imposing order and reason; when she resists, he puts her on the rack. "I am come in very truth leading to you nature with all her children to bind her to your service and make her your slave," Bacon says in *The Masculine Birth of Time.* In *Advancement of Learning,* he comments, "You have but to follow and as it were hound nature in her wanderings, and you will be able, when you like, to lead and drive her afterwards to the same place again."

The historical identification of science with male domination is an enduring one. In the metaphor, masculine science and feminine nature often produce offspring. Richard P. Feynman used the trope in his 1965 Nobel lecture. "So what happened to the old theory that I fell in love with as a youth?" he asked. "It's become an old lady, who has very little that's attractive left in her.... But, we can say the best we can for any old woman, that she has been a very good mother and has given birth to some very good children."

The language of modern science has a decidedly masculine bent—and not just in its pronouns and metaphors. Even its praise is skewed. Edwin Hubble congratulated the brilliant astrophysicist Cecilia Payne-Gaposchkin by calling her "the best man at Harvard."

Indeed, the discourse of modern science is replete with arguments and asides meant to demonstrate that only males have the intellectual, physical and psychological qualities necessary to do good science. Women just don't have the right stuff. Nineteenth-century neuroanatomists and craniologists, for example, diligently measured and weighed female brains to prove women lacked a talent for the hard task of scientific reasoning. Sir David Brewster, Newton's biographer, announced that "the mould in which Providence has cast the female mind, does not present to us those rough phases of masculine strength which can sound depths, and grasp syllogisms, and cross-examine nature." The womb, too, came in for its share of blame. "[Woman] is less under the influence of the brain than the uterine system," wrote Dr. J. G. Millingen in 1848.

Women were sternly warned that any effort to hone their inferior brains, particularly in science, would lead to damage both to themselves and to their unborn children. "Over-activity of the brain during the critical period of the middle and late teens will interfere with the full development of mammary power and of the functions essential for the full transmission of life generally," warned G. Stanley Hall, president of Clark University, in 1906. As ever, female nature was called to support male science in this argument. "It cannot be emphasized enough," Max Planck said, "that nature herself prescribed to the woman her function as mother and housewife, and that laws of nature cannot be ignored under any circumstances without grave damage, which... would especially manifest itself in the following generation."

Women who managed to do science despite these injunctions were historically portrayed in language that minimized or trivialized their accomplishments. Caroline Herschel, for instance, sister of William Herschel and an important astronomer in her own right, is described in one account as someone who took care of the "tedious minutiae that required a trained mind but would have consumed too much of Sir William's time." Another commentator felt called on to explain in *Westminster Review* (1902) that scientific work done by women "is either done in conjunction with men, or is obviously under their guidance and supervision, and much is made about it out of gallantry."

Nowadays women's treatment in the public discourse of science is looking up. Rampant sexism appears to have expired, although occasionally there is an eerie echo of Planck's warning to those unborn generations. In September 1990, for instance, a respected chemist at the University of Alberta published a peer-reviewed article in the *Canadian Journal of Physics* (*CJP*) that blamed mothers who work for many of the ills of modern society, including drug use, cheating and corrupt politics. But while Planck's comments caused no stir whatsoever, the *CJP* paper led to a very public uproar. Nine issues later the editor in chief apologized, saying that the "article does not comprise science and has no place in a scientific journal."

That's progress for you, however slow it may sometimes seem in coming. We'll just have to accelerate the pace a bit, for, as we all know, time waits for no woman.

ANNE EISENBERG, a professor at Polytechnic University in Brooklyn, is the author of four books on scientific and technical writing.


Ford
Buckle up—together we can save lives.


PRODUCT OF ENGLAND
BOMBAY
SAPPHIRE
Distilled
LONDON
DRY GIN
Distilled from 100% Grain Neutral Spirit from a 1761 recipe.
1 Litre
Alc.47%/vol.(94° Proof)
SOLE DISTRIBUTORS FOR THE U.S.A
CARILLON IMPORTERS LTD
TEANECK, NEW JERSEY
BOMBAY
IMPORTED
ENGLISH GIN
THE BOMBAY SAPPHIRE MARTINI. AS ENVISIONED BY MICHAEL GRAVES.
POUR SOMETHING PRICELESS.
Bombay® Sapphire™ Gin. 47% alc/vol (94 Proof). 100% grain neutral spirits. ©1992 Carillon Importers, Ltd., Teaneck, N.J. ©1992 Michael Graves.