SCIENTIFIC AMERICAN

MAY 1989
$2.95

Nuclear artistry: forging the heaviest elements.

Can the Amazon rain forest survive humankind?

Optical fibers open windows into the body.

***Curtains of auroral light** are shaped by electric fields recently discovered high in the ionosphere.*

THE BMW 7-SERIES. IT BRINGS NEW BLOOD TO A CLASS OF AUTOMOBILE THAT CAN CERTAINLY BENEFIT FROM IT.

An eight-year, billion-dollar quest to rethink every aspect of the luxury car has produced what one industry observer deems "the world's first sports limousine."

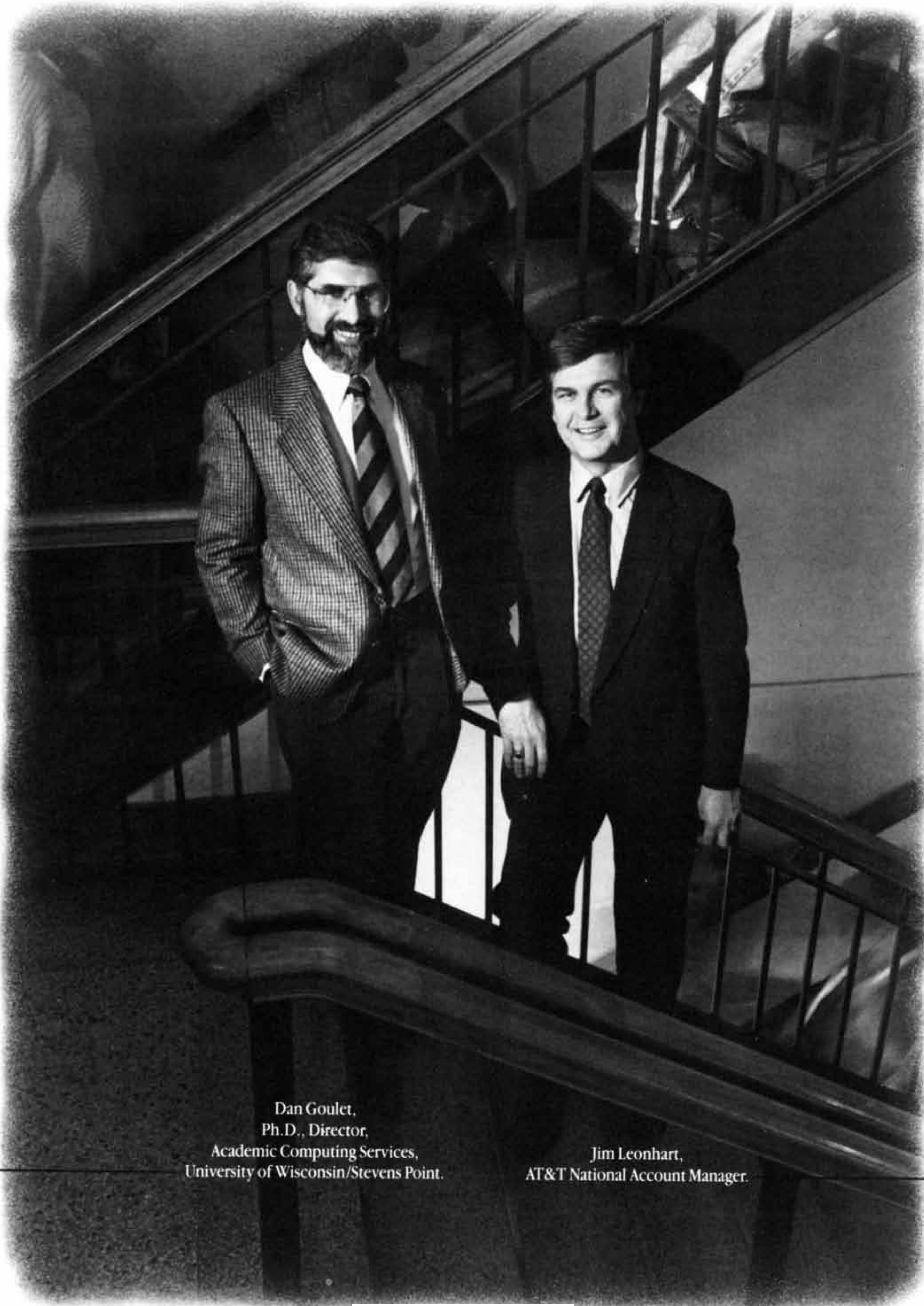
It is an astonishingly roomy, quiet sanctuary whose inventive amenities can even include a telephone as standard


OCALA.

SCIENTIFIC AMERICAN

May 1989 Volume 260 Number 5

DEPARTMENTS


Dan Goulet,
Ph.D., Director,
Academic Computing Services,
University of Wisconsin/Stevens Point.
Jim Leonhart,
AT&T National Account Manager.

THE COVER painting depicts a ray-streaked aurora sweeping across the clear night sky above a high northern-latitude region (see "The Dynamic Aurora," by Syun-Ichi Akasofu, page 90). The lights are emitted when oxygen and nitrogen in the lower ionosphere are struck by electrons, which stream in from the sun and are accelerated by a vast generator mechanism in the earth's magnetosphere. A peculiar electric field in the upper ionosphere accounts for the aurora's sheetlike shape.

THE ILLUSTRATIONS

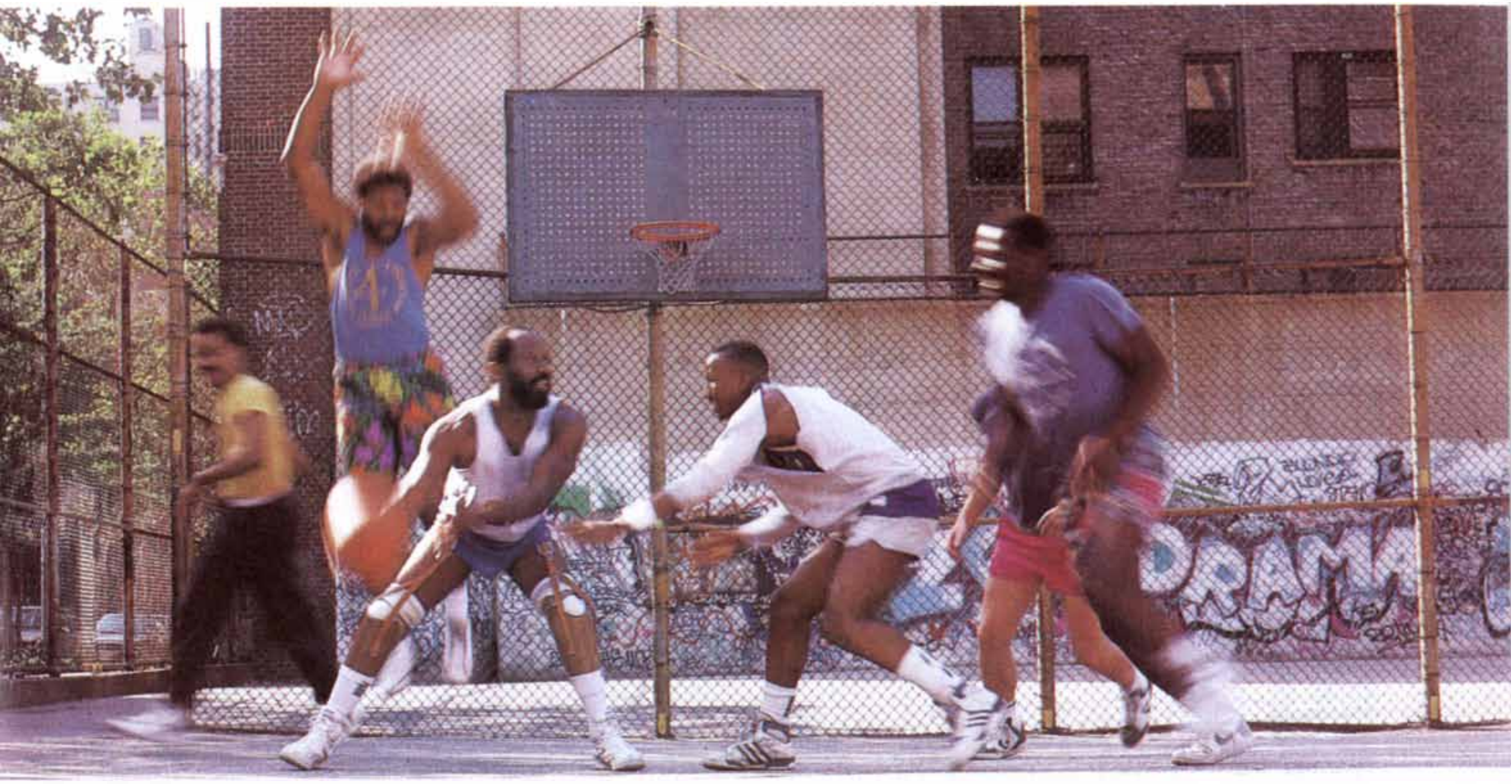
Cover painting by Ian Worpole

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017. Copyright © 1989 by Scientific American, Inc. All rights reserved. Printed in the U.S.A. No part of this issue may be reproduced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of the publisher. 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. Subscription rates: one year $24, two years $45, three years $60 (outside U.S. and possessions add $11 per year for postage). Subscription inquiries: U.S. only 800-333-1199; other 515-247-7631/32. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51593.

SCIENTIFIC AMERICAN

Established 1845

For Bill Demby, the difference means getting another shot.



When Bill Demby was in Vietnam, he used to dream of coming home and playing a little basketball with the guys.

A dream that all but died when he lost both his legs to a Viet Cong rocket.

But then, a group of researchers discovered that a remarkable Du Pont plastic could help make artificial limbs that were more resilient, more flexible, more like life itself.

Thanks to these efforts, Bill Demby is back. And some say, he hasn't lost a step.

At Du Pont, we make the things that make a difference.

Better things for better living. DU PONT

REG. U.S. PAT & TM. OFF

LETTERS

To the Editors:

"Fusion's Future," John Horgan's review of the prospects for fusion energy ["Science and the Citizen," SCIENTIFIC AMERICAN, February], is welcome and timely. As he points out, important scientific progress in magnetic-fusion research and new opportunities for international collaboration are generating renewed interest in fusion's prospects at a time when the need for alternative energy sources is becoming increasingly apparent.

We should like to call attention to a recent multiyear study, not cited by Horgan, that explores the possibilities for developing magnetic-fusion reactors well suited to society's requirements. The study was the work of a 10-person national committee reporting to the U.S. Department of Energy—the Senior Committee on Environmental, Safety and Economic Aspects of Magnetic Fusion Energy (ESECOM). Its findings were summarized in a 50-page article, "Exploring the Competitive Potential of Magnetic Fusion Energy: The Interaction of Economics with Safety and Environmental Characteristics," in the January 1988 issue of *Fusion Technology.*

The ESECOM study used state-of-the-art engineering-economic models and sophisticated methods for characterizing radiological hazards to compare a variety of fusion-reactor conceptual designs with one another and with advanced as well as contemporary fission reactors. The plasma-physics performance assumed in the fusion cases was based on optimistic but defensible extrapolations from the performance that has been achieved in experiments to date.

The ESECOM study found, in short, that fusion energy appears to have the potential to achieve significant advantages over fission energy with respect to safety, environmental impacts and links to nuclear weaponry, at energy costs roughly comparable to those of fission today. The ESECOM calculations indicated that "worst case" accidents in fusion reactors would produce population exposures to radiation from 10 to 1,000 times smaller than "worst case" accidents in fission reactors, and that radioactive-waste hazards (by the most meaningful indexes, which take into account the volume, radiotoxicity and longevity of the waste) could be from 100 to 1,000 times smaller for fusion than for fission. Economic costs associated with fusion technology's complexity might be offset, for the safest designs, by the savings resulting from easier siting and licensing of the plants and reduced requirements for plant components to be certified as being "nuclear grade."

Fusion reactors with such favorable environmental and safety characteristics, at competitive costs, will not materialize automatically. Years of difficult development work remain before any fusion reactor could be installed in the power grid, and achieving the full environmental and safety potential of this technology will entail working with advanced materials and innovative designs that may extend the development time and increase costs.

The difficulty is no reason not to proceed, but it does underline the importance of international cooperation to divide the work and share the costs. The history of East-West cooperation in magnetic fusion—which dates back to 1958 and in which the West has gained at least as much as it has given—shows that the concern with transferring advanced U.S. technology to the Soviet Union through cooperation in fusion is misplaced. The current embodiment of this cooperative effort, the International Thermonuclear Experimental Reactor (ITER) project mentioned by Horgan, deserves full U.S. participation and support.

Indeed, the international cooperation that has long characterized work in magnetic fusion should be replicated in building up R&D programs for other energy options, including increased end-use efficiency, renewable energy sources, more efficient and cleaner-burning fossil-fuel technologies and safer fission reactors. No single approach—not even the "cheap, efficient solar cell" postulated at the end of Horgan's story—will be able to do the entire job of meeting the energy needs of the world's huge and diverse civilization. Current uncertainties about future needs and the characteristics of the energy options that will be available to meet those needs are too great to allow a sensible decision now about which options can safely be forgone.

Instead of arguing on the basis of inadequate information about which of the long-term energy options is the "best" one, analysts ought to unite behind the proposition that we are investing far too little in the energy future all across the board. In inflation-corrected dollars, Federal energy R&D today is running at about a third of the level of a decade ago. The extra \$4 billion per year needed to regain the spending level of 1980 corresponds to an increase of four cents per gallon in the tax on gasoline or a tax of 75 cents per ton on carbon dioxide emissions (7.5 cents per million British thermal units of heat produced from coal, six cents for the same heat from oil and 4.5 cents for natural gas). It would be a wise investment, and hardly unaffordable.

JOHN P. HOLDREN

Energy and Resources Group
University of California, Berkeley

T. KENNETH FOWLER

Department of Nuclear Engineering
University of California, Berkeley

To the Editors:

As one of the 8 percent of males who have a red-green color deficiency, I was fascinated by "The Genes for Color Vision," by Jeremy Nathans [SCIENTIFIC AMERICAN, February]. I am a green-anomalous trichromat, although I had not realized that my "problem" had a name or understood its physiological basis.

Since the characteristics of these color-vision abnormalities are so well known, it would seem possible for human-factors-engineering standards to specify avoiding certain colors in the design of some signs, signals and displays. Examples are traffic signals (I can see only the slightest trace of green in most "green" lights) and computer displays (a personal-computer plotting program displays one type of data in a blue-green color that, to me, is almost indistinguishable from the "white" screen background).

The information in the article suggests that certain shades of green, red, yellow and orange probably appear brighter to red- and green-anomalous trichromats (and possibly to dichromats) than to people with normal vision. I wonder if this has influenced tastes in art. Van Gogh, Titian, El Greco and Kim Poor are among my favorite artists. Is it possible that the pigments available to the first three of these (Kim Poor is contemporary) included relatively large proportions of the colors I can see best? Or could it be that the artists chose such colors because they had color-vision abnormalities similar to mine or were influenced by teachers, patrons or critics with abnormal color vision?

CARL FIELDS

Farmer City, Ill.


© 1989 American Express Travel Related Services Company, Inc.

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

MAY, 1939: "When its temperature is lowered to 2.19 degrees Kelvin, liquid helium seems to stop boiling. This observation is merely one of many that show the temperature marks a transition point between two states of helium. Helium I is the warmer, helium II the colder. Helium II possesses an unusual heat-conducting capacity, far in excess of helium I; some time ago Peter Kapitza, the Russian scientist who spent many years in Lord Rutherford's laboratory at Cambridge, England, thought this might be due to an abnormally high fluidity—or low viscosity, as the physicists call it. This would mean that the liquid would circulate with extreme ease, carrying heat by means of convection currents. The viscosity of helium II now turns out to be about the same as that of gaseous hydrogen, low enough to set the theoretical physicists to try to explain these superfluid properties."

"The question of muscular flight is evidently not a fantastic one, and is being entertained in many responsible quarters. Dr. W. F. Gerhardt of Wayne University has built a 'girocycle,' which is in reality a helicopter of the simplest form. The airscrew is mounted on a vertical shaft driven through a two-to-one pinion gear by ordinary bicycle foot pedals. The girocycle is not yet a free-flying vehicle but is controlled laterally by a stationary tripod. Three records have been created with this unique device: the first vertical 'flight' by muscular power, the first 'flight' by muscular power ever achieved by a woman, and an ascent, or jump-off, of one eighth of an inch. At any rate, this is a beginning."

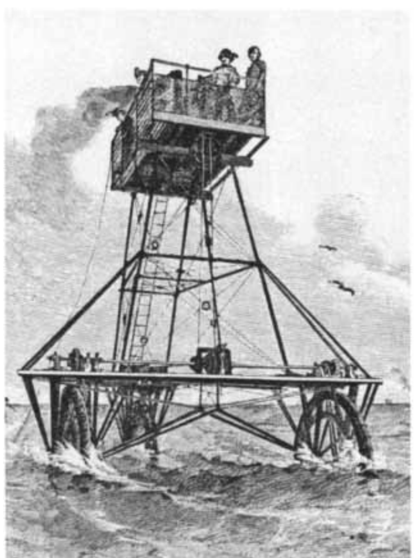
"From seeds estimated to be between 300 and 500 years old, lotus plants are being grown today at the Field Museum of Natural History in Chicago. So far as can be ascertained, this represents the longest duration of delayed germination on record. The lotus seeds had lain buried in a peat bed in southern Manchuria through several centuries."

"Air conditioning now being installed in submarines will make America's undersea fighting force a much more potent weapon in any future war. The use of cooling and dehumidifying equipment and the provision of oxygen from tanks to replace exhaled carbon dioxide will enable submarines to run submerged probably for periods of days, thus adding another factor to their effectiveness in attack or defense."

SCIENTIFIC AMERICAN

MAY, 1889: "The *Timberman* does not seem alarmed at the prospect of an early destruction of our timber supply. It asserts that Puget Sound has 1,800 miles of shore line, and all along this line, miles and miles farther than the eye can reach, is one vast and almost unbroken forest of enormous trees. The forests are so vast that, although the sawmills have been ripping 500,000,000 feet of lumber out of them every year for the past ten years, the spaces made by these inroads seem no more than garden patches."

"A French firm have succeeded, after long endeavors, in applying electricity to the art of tanning in such a way as is claimed to accelerate the process. What is actually done is very simple. The rawhides are placed in a large cylinder which revolves upon a horizontal axle. Provision is made for passing a current through the drum, the electrolyte being a decoction of tannin. The drum is kept slowly revolving until the process is complete. Light calf skins, and sheep and goat skins, which used to require from four to six months, are said to be completely tanned in twenty-four hours. The exact nature of the *role* played by the electric current is not at all clear. Prof. S. P. Thompson, who has examined the process, suggests that the effect may be in some way to open the pores of the hides and so permit a more rapid access of the tanning solution."



The ocean tricycle at Atlantic City, N.J.

"According to the *Virginia City* (Nev.) *Enterprise*, the fortune that awaits the inventor of a successful dry-placer machine, or any other method by which the gold in the loose dirt on the hills and mountains of Nevada can be separated, will make the present wealthy men of the world have, by comparison, dismal anticipations of the poorhouse."

"Mr. H. M. Stanley, the African explorer, gave an extremely interesting reference to the arrow poison employed by the natives of the Lower Congo district. Mr. Stanley says his party was much exercised as to what might be the poison on the heads of the arrows by which Lieut. Stairs and several others were wounded, and from the effects of which four persons died almost directly. The mystery was solved by finding at Arisibba several packets of dried red ants. The dried insects were ground into powder, cooked in palm oil, and smeared on the arrows. It is well known that formic acid exists in the free state in red ants; there is little ground for doubting that it was the 'deadly irritant.' "

"The illustration shows a vehicle designed to cause somewhat of a sensation in the world under the waters as well as above. The ocean tricycle, as it is called, consists of a high platform carried on an iron framework, the whole resting on three wheels. The engine, with its boiler, is placed on the platform, well above the reach of spray. As the engine works it turns a vertical shaft that descends within the framework. The height above the water at which the passengers are carried discloses many features of the bottom, otherwise invisible, such as variations in color, depth, beds of seaweed, etc., exactly the same as can be seen from the masthead of a sail boat in still, shallow water. We doubt not that the inhabitants of Atlantic City will have many enjoyable rides on this machine over the level bottom there."

The life busine drop-dead tech

In life insurance, the saying goes, the paperwork can kill you. The paperless office is every company's dream.

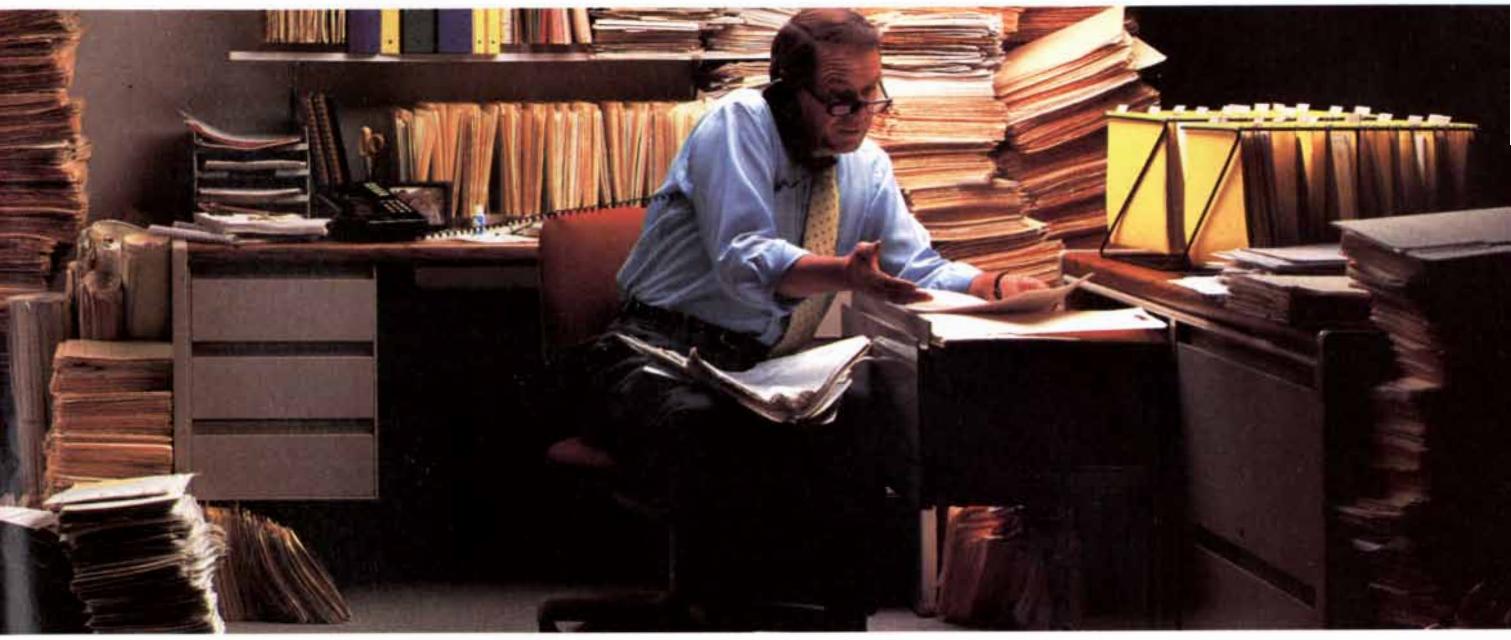
That's why lots of huge insurance companies are using NYNEX.

Recently, to simplify life for one of these colossi, NYNEX replaced acres of paper with an innovative new voice and data network, connecting their operations center with their headquarters. It makes their systems compatible.

And remarkably, it allows them to access their whole information network by plugging any work station into any phone jack.

This NYNEX technology isn't just convenient. It's saving them a lot of money.

ss needed some
nology.



Which is always the best policy.

The NYNEX family of companies would like to work with you, too. We offer you everything from computer networks and software, to mobile phones, to the services of New England Telephone and New York Telephone.

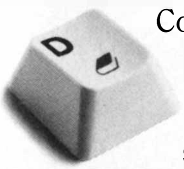
Call us at 1 800 535-1535.

When we say the answer is NYNEX, it's not just an empty claim.

Need to communicate? Need to compute? The answer is

NYNEX

Our frills require no skills.

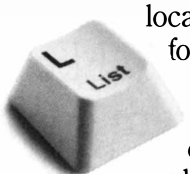


You don't have to be a genius to use the remarkable features of the Smith Corona XD 7500. You just have to know how to push the right buttons.

Detect typos and misspellings before anyone else does with the *Spell-Right*™ 75,000 word electronic dictionary. Tap *List* and it looks up the correct spelling.

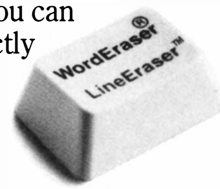
One touch of *Find* locates errors for you. Once you locate those errors, *Word-Eraser*®/*LineEraser*™ erases a whole word or a whole line automatically. There are even 10 lines of automatic memory correction to make revisions simpler than ever.

Smith Corona
Correcting Cassette



The XD 7500 also offers the exclusive Smith Corona fumble free *Correcting Cassette.* Just pop it in. There are no spools to unwind... no complicated threading... no tangles.

Using the word processing features of the XD 7500 is also a breeze. You can type directly into its 12,000 character memory using the easy-to-read, 40 character display. Then you can proofread your typing, make changes, and print once everything's perfect.



Of course, the XD 7500 is made right here in the U.S.A.

So come see how easy it is to use the XD 7500. We're sure you'll be frilled.

SMITH CORONA®
TOMORROW'S TECHNOLOGY AT YOUR TOUCH™

For more information on this product, write to Smith Corona Corporation, 65 Locust Avenue, New Canaan, CT 06840 or Smith Corona (Canada Ltd.), 440 Tapscott Road, Scarborough, Ontario, Canada M1B 1Y4.

SCIENCE AND THE CITIZEN

Aging Comes of Age

The new biology turns its acumen to an old topic

Advances in neurobiology, immunology and molecular biology are bringing the transformations associated with aging within the reach of analysis. "There's been a drastic change in the constituency of investigators," says Caleb E. Finch of the University of Southern California, one of the organizers of a recent meeting in Santa Fe devoted to the subject. The meeting "marks recognition by outside disciplines," according to Thomas E. Johnson of the University of Colorado at Boulder, Finch's co-organizer. "Five years ago there wasn't a molecular biology of aging: here, half to two thirds of the talks adopt a modern biotechnology approach."

Some venerable theories of aging have been swept away. "Nobody believes there's a single gene controlling aging—that's a change from 10 years ago," Finch remarks. Nor, according to Johnson, has experimental support been found for the plausible idea that a cell simply accumulates random errors in its genes and proteins until a fatal catastrophe happens. Indeed, some senescent cells (cells that have stopped dividing in culture) are still quite capable of replicating DNA. "We're not looking for single causes any more," Johnson says. "Aging is a mixture of multiple causes."

As the old theories fade workers are finding fresh leads. It is becoming clear that many of the changes that come with age are ordained by a genetic program. (Cancer cells, which multiply indefinitely, usually do not show changed gene expression over generations as normal cells do.) Some manipulations can affect the genetic program. For example, it is known that dietary restrictions can increase an animal's life span. In rats the increase is accompanied by enhanced expression of some genes in liver tissue, according to Arlan G. Richardson of Illinois State University. Yet some genetic changes also have a random component. Peter J. Hornsby of the Medical College of Georgia has shown that one bovine cell type loses its ability to express a certain gene with a fixed probability at each cell division. "The cells forget what they are supposed to be doing," Hornsby says.

Some genes affected by aging may have crucial roles. For instance, Paul S. Swerdlow of the University of Virginia has evidence that human cells approaching the end of their ability to divide in culture contain lowered levels of a bound form of the protein ubiquitin. One of ubiquitin's tasks is to mark defunct proteins for destruction. Swerdlow thinks senescent cells might have defective ubiquitin-binding systems that allow damaged proteins to accumulate.

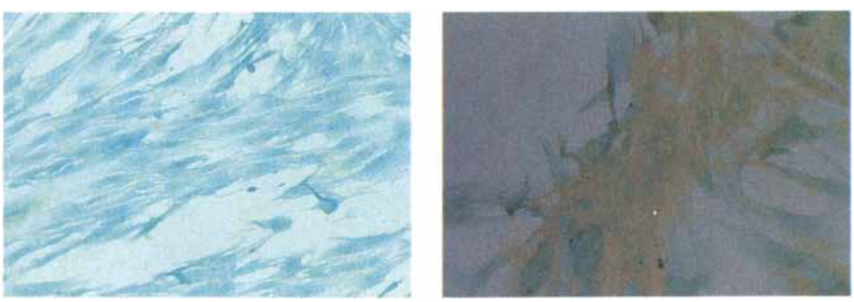
The mechanisms that prevent cell division in senescent cells offer another clue to the process of aging. Andrea L. Spiering, James R. Smith and their colleagues at the Baylor College of Medicine in Houston have found there is a substance (which they are now trying to isolate) that is present only in senescent cells and inhibits DNA synthesis. They have isolated a mutant human cell line that multiplies indefinitely but that still produces the inhibitor (which can affect other cells).

A line of inquiry that researchers find particularly exciting relates to fibronectin, a protein that anchors cells in position and helps young cells retain their shape. Mary Beth Porter, who also works at Baylor, has found a variant form of fibronectin that is present only in senescent cells. Her work dovetails with investigations at other laboratories showing that changes in gene transcription in old cells often affect fibronectin. It appears that the altered protein may be defective.

Despite such promising leads, there is still much ambiguity. One recurrent difficulty lies in deciding whether a particular genetic change is a cause or a consequence of cell senescence and indeed whether it is relevant to aging of the whole organism. Neurodegenerative diseases such as Alzheimer's, which seem to have a genetic component, are good subjects for study. Carol A. Miller of the University of Southern California has employed monoclonal antibodies to identify small subgroups of neurons that are destroyed in such diseases, which could give valuable clues to what is going wrong.

If important changes during aging are indeed genetically programmed, why should that be? Might there be some evolutionary benefit? Johnson has found a gene he calls *age-1* in the tiny worm *Caenorhabditis elegans* that shortens life span but compensates by increasing fecundity. A mutant worm with a defective *age-1* gene lives 70 percent longer than normal but pro-

Aging cells in culture exhibit an unusual protein



HUMAN CELLS that have attained their maximum number of divisions* (right) *have a variant form of the protein fibronectin, seen here as brown-stained patches. Patches show binding by a monoclonal antibody isolated by Mary Beth Porter of the Baylor College of Medicine. Young cells* (left) *do not show the variant protein.

Indulge yourself.

Introducing the Park Avenue Ultra.

The Buick Park Avenue is widely recognized as a standard for American luxury sedans.

Now, there is the new Buick Park Avenue Ultra — the ultimate full-size Buick sedan. The Ultra is both a spacious and yet very personal luxury car. It embodies a unique combination of premium styling, confident performance and sumptuous accommodations.

Naturally, the Ultra is equipped with a long list of sophisticated features, including: ■ All-leather seating with 20-way power front seats ■ Unique grille, tail lamps and 2-tone paint treatment ■ Exclusive full vinyl roof (available) ■ Special acoustic package for superior quietness ■ Anti-lock brake system ■ 165-horsepower 3800 engine with sequential-port fuel injection ■ And much more.

Visit your Buick dealer and experience the elegance of the new Park Avenue Ultra.

The Great American Road belongs to

BUICK®™

BUICK SMARTLEASE by GMAC

Brown, Greszczuk, Liebeck & Xerikos

These are the kind of Fellows who are making the world we'll live in tomorrow.

These are four top achievers in technology—newly named McDonnell Douglas Engineering and Research Fellows.

Our company searches out and develops men and women who are creative leaders in their work. We surround them with challenges to be met. We provide them with the facilities to fulfill their potential and with opportunities for advancement in the fields in which they do their best work.

Brown, Greszczuk, Liebeck & Xerikos, along with the previously named Fellows whose ranks they join, bring honor and progress to our company. They represent in the world of technology what the Albin Polasek sculpture shown here represents in art: Mankind carving its own destiny.

Dr. Ian Brown is acknowledged as a leading investigator of the kinetics of polymerization. Using the magnetic resonance technique to study the molecular behavior of polymers, polymeric matrices, and fibers, he has provided invaluable guidance for our design and manufacturing engineers as they work with various composite materials in the products we build.

Longin Greszczuk has gained national recognition as an authority on composite materials and structures. Models he has developed for predicting physical properties and characteristics, stresses and interactions in a variety of composite structures have helped produce success in several major space and national defense products we design and build.

Dr. Robert Liebeck has earned wide acclaim for his development of maximum lift low-speed airfoils—wing designs and propellers. His unique airfoil concepts have found their way into products as diverse as cooling fans and wind turbines; the man-powered Gossamer aircraft; and a winning Indianapolis 500 race car. And of course, into our own products as well: The Ultra High Bypass engine program, and our new MD-11 trijet.

Dr. James Xerikos clearly deserves his superior reputation in aerodynamic research. His work in numerical gasdynamics, supersonic and hypersonic flowfields, shock/boundary layer interactions, and advanced finite computation has led to vehicle designs ranging from hypersonic interceptors to underwater armaments. His current task: develop and verify the aero-optic analyses that will help put our new interceptor missile on target for the Strategic Defense Initiative.

"Man Carving Out His Destiny"

MCDONNELL DOUGLAS

A company of leaders.

duces five times fewer offspring. There may be similar genes in other species.

Changes in human immune function that occur with aging might also have beneficial aspects. Marc E. Weksler of Cornell University Medical College finds that only one quarter of T lymphocytes from old people retain their full ability to proliferate in response to an external stimulus. Such a decline could explain the increased susceptibility of the elderly to infectious disease. The change might nonetheless allow immune surveillance of the body's own tissues to be stepped up—which indeed seems to happen—and so lessen the likelihood of cancer.

For all the interest in genetic programs, it is also clear that naturally occurring chemical substances can destroy cells. Such destruction might be critical in some tissues. Attempts to develop drugs to block such damage are already under way. Retin-A (tretinoin), a drug that can reduce skin wrinkles, is a product that arose indirectly from aging research and is now a major success story for its manufacturer, Johnson & Johnson. This fact was not lost on venture capitalists at the Santa Fe meeting, who seemed particularly impressed with the promise of the fibronectin discoveries. If Wall Street was too late to back Hernando de Soto's 16th-century search for the fountain of youth, it seems to have caught up with his lab-coated successors. —*Tim Beardsley*

Cold Water on the Fire

A recent survey casts doubt on evidence for early use of fire

In a recent issue of *Nature* two investigators report finding what they call the earliest evidence for human use of fire: bits of charred animal bone from a South African cave that appear to be more than a million years old. The game was roasted, the authors argue, by hominid precursors of *Homo sapiens,* perhaps by members of the species *H. erectus.*

Now a review of the evidence for early use of fire suggests such claims should be regarded with caution. Writing in *Current Anthropology,* Steven R. James, a doctoral candidate at Arizona State University, considers evidence from sites where *H. erectus* is presumed to have made fires and concludes that all of it can be explained by natural causes. There is no solid evidence for controlled use of fire until about 200,000 years ago, James says. Since by that time early forms of *H. sapiens* were already present, it is possible *H. erectus* never used fire at all.

The main sites that James considers are in East Africa, South Africa, Europe and China. In each region, he claims, similar processes have led anthropologists to exaggerate the strength of the evidence for early fire. Some specimens once thought to be burned have actually been altered by other processes, including staining by manganese from the surrounding soil. Without specific chemical tests, which often were not applied in early excavations, such staining is difficult to distinguish from the effects of burning.

Even when prehistoric remains are clearly burned, it is difficult to prove that the fire was made and controlled by hominids. For example, burned clay found in association with stone tools and bone fragments in the Middle Awash Valley in East Africa has been taken as evidence for controlled use of fire by early hominids a little more than a million years ago. Some anthropologists have suggested that human ancestors there kept stumps burning as farmers still do in India—in order to have fire available while working far from their homes.

There are several problems with such interpretations, James says. Similar bits of burned clay are found elsewhere in the region unaccompanied by tools or bones. Furthermore, none of the animal remains are reported to have been burned, which suggests that if there were fires, they were not for roasting meat. In addition, the archaeological materials are not, for the most part, found in direct association with the burned clay. The simplest explanation, James argues, is that the clay was burned by natural fires or volcanic activity.

He is similarly skeptical about the evidence for fire from Zhoukoudian in China. Zhoukoudian is the famous cave site where "Peking man" (an example of *H. erectus*) was discovered in the 1920's. Among 50 meters of deposits in the cave are burned bones, ash, charcoal and stone tools, which have been taken as evidence that fire was a daily part of the life of *H. erectus* bands living there 400,000 to 500,000 years ago. Indeed, the findings from Zhoukoudian have been crucial in supporting the notion that *H. erectus* was the tamer of fire.

Yet recent reanalysis of specimens from the older layers at Zhoukoudian by Lewis R. Binford of the University of New Mexico calls the standard interpretation of the cave findings into question. Not only are there no well-defined hearths, but some of the layers first described as consisting of ash may actually include owl pellets, hyena scats and other accumulations that were burned by naturally caused fires. "Clearly the entire question of burned deposits at Zhoukoudian needs to be reexamined," James concludes.

The earliest evidence for controlled use of fire by hominids that the skeptical James is willing to accept (tentatively) comes from the beach at Nice. There, at a site called Terra Amata, excavations in the 1960's revealed what may be the foundations of huts, each with a hearth at its center. Recent dates obtained by new techniques indicate that the site is about 230,000 years old. If the date is correct, the hearths might be traces of ancestral Neanderthals, an early form of *H. sapiens*, later superseded by anatomically modern human beings.

"I'm certainly not ruling out the use of fire by *Homo erectus* or other precursors," James said recently, "but there have been some early claims that are not well substantiated. These need to be treated with caution. Our picture is changing: there's a lot of evidence now that early hominids were scavengers rather than hunters, and scavengers seem less likely to use fire. Anthropologists have invented stories to account for early use of fire—based on a model of hunters sharing meat around the campfire—but that's just not what you see in the archaeological record." *—John Benditt*

Case Proved

Changes in science curriculums may actually take place

The litany of surveys documenting the sorry state of U.S. science education seems to have convinced educators that the time has come for study to give way to action. The most recent survey, by the Educational Testing Service (ETS), found that 13-year-olds in the U.S. did worse in mathematics than those in Korea, five Canadian provinces, Spain, the United Kingdom and Ireland; in general science the U.S. was tied with Ireland and two Canadian provinces in the lowest tier of achievement. In response to the study U.S. Secretary of Education Lauro F. Cavazos asked, "How many times must this nation be reminded of its educational deficit?"

"People are tired of studies," says Bill G. Aldridge of the National Science Teachers Association. "We're calling for something to be done immediately." Aldridge is steering one of a number of disparate efforts to address scientific illiteracy by improving curriculums. His diagnosis is that too much science is now crammed into too little time. Aldridge points out that the few U.S. high school students who study physics at all must attempt to cover in one year the same amount of physics that students in many countries spread over five or more years. The science teachers' scheme would have each subject taught for several years, with the course content gradually becoming more theoretical; it would also strive for a coordination between subjects that is now lacking.

Aldridge says the project could start producing materials for high school teachers as soon as next year and that the curriculums would then begin to be introduced in five or more test sites, one of which may be the entire state of North Carolina. Aldridge expects half of the estimated $10 to $20 million total cost of the project to come from the National Science Foundation; the Department of Education and individual states might also chip in. Nobel laureates Sheldon Lee Glashow and Leon M. Lederman, as well as publishers and influential educators such as Bill Honig of the California State Department of Education, are said to have lent their support.

Less hurried efforts in curriculum development are under way at the National Research Council (NRC), where one committee is attempting to revamp mathematics education; as part of that initiative the National Council of Teachers of Mathematics published new standards for evaluating math curriculums in March. Another NRC committee is to issue reform recommendations on biology curriculums this fall; that panel is backed by the Howard Hughes Medical Institute in Bethesda, Md., which has a strong interest in improving biology education—and the financial muscle to make things happen.

Project 2061 (the name refers to the year when Halley's comet will next approach the sun) is yet another example. It is an enterprise of the American Association for the Advancement of Science (AAAS) that will attempt a radical redefinition of science curriculums. Progress has been slow. The first phase, which simply established what high school graduates should know, was intended to last six months but took five years. Moreover, an overview volume recently published by the AAAS, *Science for All Americans*, was the subject of much dissent among panels of scientists: each group felt its own subject had been underrepresented. Although unkind observers suggest that Halley's comet may be back before definitive curriculums and materials emerge, some interim materials from Project 2061 in fact will be available for Aldridge's faster-paced effort.

To be sure, curriculums are worth nothing unless school administrators and teachers adopt them. F. James Rutherford of the AAAS says Project 2061's long-term approach will allow time to build a constituency among teachers and principals. Will they be willing to abandon old ways? "We can hope there is enough recognition that there's a crisis," Rutherford observes. The ETS, though, is proposing just one more comparison of U.S. students with those overseas. Archie E. Lapointe of the ETS says a further study could establish just what educational standards are achievable and could also identify the characteristics of successful school systems. "The best way to mobilize is to appeal to our competitive instincts," according to Lapointe. *—T.M.B.*

PHYSICAL SCIENCES

Pinning Down Clouds

Scientists ponder the role of clouds in climatic change

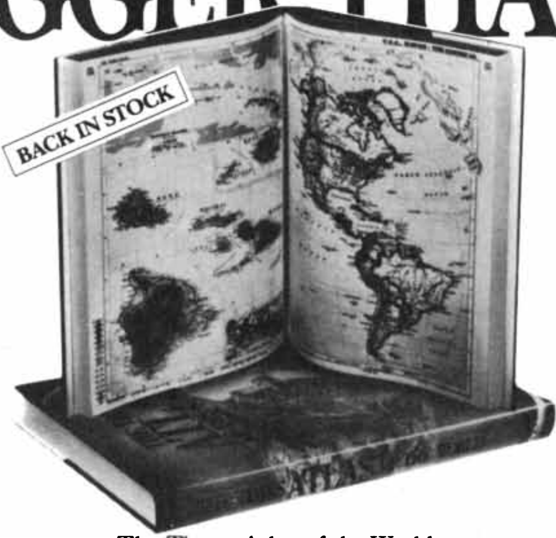
The more climatologists look at clouds, the more they realize they really don't know clouds at all. A recent comparison of 12 global climate models from six countries by Robert D. Cess of the State University of New York at Stony Brook underlined the problem. Cess eliminated equations simulating clouds' influence on the climate from the models and found that their forecasts of a warming induced by the greenhouse effect agreed "exceptionally well." When the cloud-feedback equations were restored, the models' forecasts diverged by a factor of three.

"There is no consensus on how clouds will affect the greenhouse warming," confirms David Randall of Colorado State University. All clouds, he says, act to some extent as "reflecting blankets," simultaneously cooling the earth by reflecting incoming sunlight back into space and warming the earth by preventing heat from escaping. Different clouds combine these properties in different proportions. Stratus clouds, which are dense and low-lying, have a net cooling effect

Choose one of these fine works and save over $290.

You simply agree to buy 4 books within the next 2 years.

BARGAINS DON'T GET ANY BIGGER THAN THIS.



Shakespeare: 6 Great Tragedies

for **$19.95** (List prices total $179.60)

Book-of-the-Month is proud to present 6 of Shakespeare's greatest tragedies: *Hamlet, King Lear, Macbeth, Othello, Romeo and Juliet* and *Julius Caesar.* Paul Scofield, Ralph Richardson, Albert Finney, Dame Edith Evans, Claire Bloom, John Mills and many more of England's most gifted Shakespearean actors perform memorably in complete readings of each play. 17 cassettes in all.

The Compact Edition of The Oxford English Dictionary

for **$27.95** (Pub. price $195)

"The most complete, most scholarly dictionary of the English language"—*The Christian Science Monitor.* Through photoreduction, the original 13-volume set has been reproduced in this two-volume *Compact Edition.* A Bausch & Lomb magnifying glass is included.

The Times Atlas of the World Seventh Comprehensive Edition

for **$29.95** (Pub. price $149.95)

"The best atlas now available in English."—*The New York Times.* The only atlas printed in 8 colors for clear reading of geographic and political features. This volume contains 292 pages of color maps and a 277-page index with over 200,000 entries which include longitude and latitude—a feature not offered in other atlases. Size: 18" x 12." Pages: 520.

The Encyclopedia of Philosophy

for **$24.95** (Pub. price $225)

Regarded as the most comprehensive encyclopedia of philosophy ever published, this superb set—compiled in collaboration with the world's foremost philosophers—encompasses aspects of ancient, medieval, modern, Eastern and Western thought. The four volumes represent an essential aid for students and a rewarding reference source.

The Story of Civilization by Will and Ariel Durant

for **$29.95** (Pub. prices total $335.45)

For almost half a century Will and Ariel Durant traced the continuity of world history—the religions and philosophies, the political and economic tides, the arts and sciences, the customs and conquests—to show the foundations of society today. A Book-of-the-Month Club exclusive for almost 50 years, the Durants' illustrated masterwork is history come alive.

Facts About Membership. As a member you will receive the *Book-of-the-Month Club News®* 15 times a year (about every 3½ weeks). Every issue reviews a Selection and more than 150 other books, which are carefully chosen by our editors. If you want the Selection, do nothing. It will be shipped to you automatically. If you want one or more other books—or no book at all—indicate your decision on the Reply Form and return it by the specified date. *Return Privilege:* If the *News* is delayed and you receive the Selection without having had 10 days to notify us, you may return it for credit. *Cancellations:* Membership may be discontinued, either by you or by the Club, at any time after you have bought 4 additional books. Join today. With savings and choices like these, Book-of-the-Month Club is where book lovers belong.

BOOK-OF-THE-MONTH CLUB®

Book-of-the-Month Club, Inc., P.O. Box 8803, Camp Hill, PA 17011-8803 A170-5-0-0

Check one box only.

- ☐ 923. Shakespeare (17 cassettes) **$19.95**
- ☐ 905. Compact OED **$27.95**
- ☐ 946. Times Atlas of the World **$29.95**
- ☐ 913. The Story of Civ. **$29.95**
- ☐ 917. Ency. of Philosophy **$24.95**

Please enroll me as a member of Book-of-the-Month Club and send me the item I have checked at left, billing me for the appropriate amount, plus shipping and handling charges. I agree to buy four more books during the next two years. A shipping and handling charge is added to each shipment.

Name________________ 9-64
(Please print plainly)

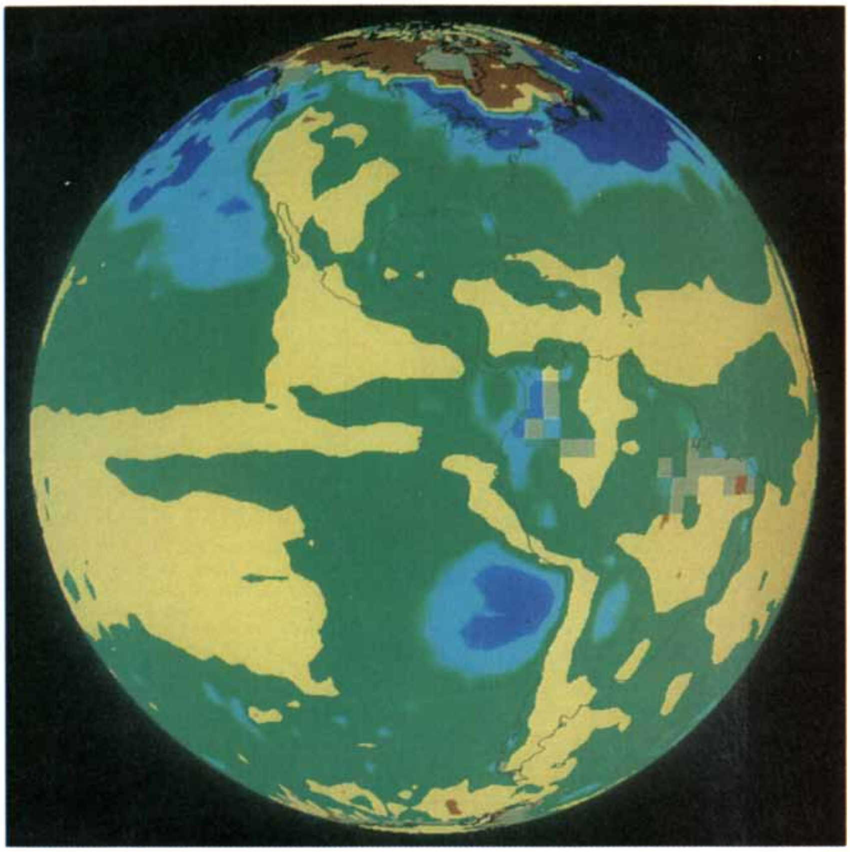
Address________________ Apt.______

City________________

State________________ Zip______

Prices shown are U.S. prices. Outside the U.S. prices are generally higher. All orders subject to approval.

The global cloud cover cools the earth now, but what will its effect be if the climate changes?



CLOUDS' EFFECTS on temperature were measured by the Earth Radiation Budget Experiment. Regions where clouds had a strong cooling effect are blue; green represents a smaller effect; yellow shows a negligible net effect and rust a slight heating effect. Image provided by V. Ramanathan of the University of Chicago.

because their albedo, or reflectivity, is relatively large. Wispy, high-flying cirrus clouds, because they are semitransparent to incoming sunlight but block infrared radiation emitted by the earth (as do greenhouse gases such as carbon dioxide), have more of a warming effect.

Beyond these essential facts, much remains unknown. What happens when cirrus clouds overlie stratus clouds, for example? "We don't have a handle on that yet," Randall says. "There is a tendency to avoid messy problems at this point."

Part of the difficulty in modeling clouds, according to Michael E. Schlesinger of Oregon State University, is that relevant phenomena range from microscopic particles called cloud nuclei, around which water vapor condenses, to the "waves" thousands of kilometers long observed rippling through the global cloud cover. "We will never have a computer powerful enough to model clouds at all scales," Schlesinger says.

What detailed knowledge there is derives largely from meteorologists' studies of clouds—such as the towering cumulus—that generate "interesting" weather. But these clouds are not necessarily the most important ones for understanding climatic change. Marine stratus clouds, for example, which cover broad swaths of ocean at cool latitudes, "don't rain, or make lightning, or snow," notes Robert J. Charlson of the University of Washington. "They are very boring, but they are one of the dominant features governing the albedo of the earth."

Climatologists have created several large-scale studies of clouds. The International Satellite Cloud Climatology Project (ISCCP) pulls together data from geostationary weather satellites and polar orbiters. To improve the interpretation of the satellite data, the National Aeronautics and Space Administration has initiated a project called the First ISCCP Regional Experiment (FIRE), in which workers observe clouds from the earth's surface and from planes while the satellites scan them from above.

Yet another NASA project is the Earth Radiation Budget Experiment (ERBE), in which satellites measure the net effect of clouds on the earth's temperature. The project recently provided a valuable piece of information: after analyzing the reams of ERBE data gathered during one month in 1985, a team led by V. Ramanathan of the University of Chicago determined that the global cloud cover (which veils roughly half of the earth's surface) has a significant net cooling effect.

Randall points out that such findings, while crucial, say nothing about the behavior of clouds in a changed climate. Most global climate models suggest that a greenhouse warming would increase the formation of high-altitude cirrus clouds, which in turn would amplify the warming, "but there is no way of knowing if we're right," he notes. Although workers can test some aspects of their models against the paleoclimatic record—by examining ice cores, oceanic sediments or tree rings, for example—the record is mute on the subject of cloudiness.

Modelers also worry that they may be overlooking important cloud-related variables. Two years ago, for example, a group headed by Charlson created a furor by proposing that sulfur-containing excretions of plankton may exert a powerful influence on cloud formation and thereby on climate. Stephen E. Schwartz of the Brookhaven National Laboratory recently cast doubt on the theory. He argued that if Charlson were correct, then sulfur emissions from human activities in the Northern Hemisphere, which are roughly double the sulfur emissions from plankton, should have led to perceptible climatic changes. In fact, Schwartz reported, that is not the case.

Tony Slingo of the National Center for Atmospheric Research in Boulder, Colo., thinks the plankton issue is far from resolved. Indeed, he fears that modelers trying to forecast climatic change face so many uncertainties that "the real atmosphere will give us an answer before we can predict it, and then we'll just be trying to understand how it happened."

Is there any silver lining to the cloud problem? Perhaps, according to Cess. "Modeling groups used to do their own thing," he notes. "Now we're all working together." —*John Horgan*

Ambidextrous Universe

New particles blur distinction between fermions and bosons

The world may be divided into people who divide the world into two categories and those who do not, but any physics student knows the world is actually divided into fermions and bosons. Fermions are subatomic particles that include the electron, proton and neutron and are traditionally referred to as the building blocks of matter. Bosons include photons and are traditionally referred to as the particles that transmit the forces of nature.

Now a paper submitted to *Physical Review Letters* by Tom D. Imbo, Chandni Shah Imbo and E. C. G. Sudarshan of the University of Texas at Austin maintains that under certain circumstances the distinction breaks down altogether. The authors have introduced various new particles including a type they call "ambions," which act like both fermions and bosons.

Fermions and bosons differ in many ways. Fermions have half-integer spin, that is, 1/2, 3/2 and so on; bosons have integer spin, that is, 0, 1, etcetera. Fermions are described mathematically by an "antisymmetrical wave function." If the positions of two fermions are interchanged, the mathematical object that describes the state of the system is multiplied by −1. Bosons, on the other hand, have a "symmetrical wave function." If the positions of two bosons are interchanged, the wave function remains unchanged.

Such differences give fermions and bosons radically different properties. Two identical fermions cannot be at the same place at the same time, because in that case if the particles interchanged their positions, the wave function would have to remain unchanged but be its own negative—clearly impossible. Bosons, on the other hand, tend to stay together, a phenomenon that manifests itself in the famous "Bose-Einstein condensation."

Why should physicists abandon the classification? "Well," as Sudarshan says, "it's been around 50 years, even longer than the lifetime of a Volvo."

More seriously, until recently scientists have assumed that the particles are moving in ordinary three-dimensional space, which enables the distinction between fermions and bosons to be made. Yet on a two-dimensional plane, for instance, particles can act like something intermediate between bosons and fermions. They will display characteristics that are neither those of bosons nor of fermions but somewhere in between. Such particles have been well studied in the past five years: they obey "fractional" statistics and are often referred to as "anyons."

The spaces investigated by the Texas group have topological properties that make particles behave like both fermions and bosons at the same time. One example is provided by two identical particles moving on a sphere where the diametrically opposed points on the surface are identified. A two-boson wave function will become a two-fermion wave function after the particles traverse certain loops in this space. "It reminds us of Moses climbing the mountain," says Sudarshan. "When he returned he was a changed man." Such particles are the ambions.

Because anyons exist on two-dimensional surfaces, they are believed to be important in such widely studied phenomena as the quantum Hall effect and high-temperature superconductivity—which are surface phenomena. As for ambions, which exist on more bizarre two- and three-dimensional spaces, Sudarshan says, "We may have to climb the mountain before we know." —*Tony Rothman*

Gamma-Ray Bursts

Two groups take stabs at a persistent cosmic puzzle

More than 20 years ago American and Soviet satellites monitoring nuclear-testing treaties first detected powerful pulses of gamma rays emanating from outer space. Since then nearly 1,000 such gamma-ray bursts, lasting from less than a second to several minutes, have been detected in all regions of the sky. None has ever been linked to a particular object. Astronomers have even disagreed over whether the bursts originate in the Milky Way or far beyond it.

Two groups have now fingered highly magnetized objects in our galaxy as probable perpetrators of the bursts—or at least of a substantial number of them. One group, led by Donald Q. Lamb, Jr., of the University of Chicago and Edward E. Fenimore of the Los Alamos National Laboratory, drew its inspiration from high-resolution observations made by Japan's *Ginga* satellite of a burst last year. The burst included not only gamma rays but also X rays. The distinctive spectrum of the X rays indicated that they had been scattered by cyclotron radiation, which is produced when charged particles are accelerated by a magnetic field. At least 20 percent of all bursts exhibit a similar X-ray spectrum.

On comparing the data with computer models, Lamb and his colleagues concluded that the burst detected by *Ginga* must have originated in the presence of a magnetic field more than a trillion times more powerful than the earth's. Only a neutron star, created when an ordinary star collapses, Lamb says, can generate such a strong magnetic field. Others have proposed mechanisms that might have triggered the bursts—a thermonuclear fire storm at the surface of the neutron star perhaps or a "quake" beneath its surface—but Lamb emphasizes that these theories are speculative.

Other workers, led by M. Boer of the Center for Space Radiation Studies in Toulouse, have advanced a more precise explanation for a smaller subset of gamma-ray observations. Most bursts are singlets—they flash once in a specific region of the sky and never again—but three so-called gamma-ray repeaters have been identified. The repeating bursts are substantially softer, or lower in energy, than the spectrums of "classical" bursts.

In *Nature,* Boer and his colleagues point out that the spectrum of the repeating bursts resembles the spectrum of radiation emitted by matter falling onto a highly magnetized white dwarf star. The workers propose that the bursts begin when a comet passes near a white dwarf and breaks into pieces; each time a fragment of the comet tumbles onto the star, it spews out a shower of gamma rays.

Bohdan Paczyński of Princeton University thinks the two proposals are interesting but certainly not definitive, since they are based on analysis of so few events. "Gamma-ray bursts," he notes, "are quite a zoo of phenomena." He suggests that the *Gamma Ray Observatory,* due to be launched by the National Aeronautics and Space Administration next year, should help bring order to the zoo. —*J.H.*

TECHNOLOGY

Energy Re-energized

Congress plans to take a hard look at energy R&D

The gloomy prospect of future greenhouse warming, together with investment in new energy technologies by foreign nations, is brightening the outlook for renew-


1986
CAR and DRIVER
10 BEST
1987
CAR and DRIVER
10 BEST
1988
CAR and DRIVER
10 BEST
1989
CAR and DRIVER
10 BEST
Some of our best advertising isn't advertising.
If good press makes good advertising, Ford Taurus has had some of the best. Every year since Taurus was introduced, *Car and Driver* magazine has included it on their exclusive "Ten Best Cars" list. They have also said, "Cars as good as the Ford Taurus roll forth once per generation."
We appreciate such rave reviews, but we also seek another type of recognition that's even more important. The one that comes from the buying public. And year after year the public, like the press, keeps casting its votes for Taurus.
So, considering the glowing praise and consistent sales success, we've come to a simple conclusion. One of the best advertisements for Ford Taurus is Ford Taurus itself.
Buckle up—together we can save lives.
Transferable 6/60 Powertrain Warranty.
Covers you and future owners, with no transfer cost, on major powertrain components for 6 years/60,000 miles. Restrictions and deductible apply. Ask to see this limited warranty at your Ford Dealer.
Quality is Job 1.
Our goal is to build the highest quality cars and trucks in the world.
Ford Taurus
Have you driven a Ford...lately?
Ford

able-energy technologies and energy conservation. Federal spending in those fields peaked in the 1970's because of the oil crises and then after 1980 fell by 70 percent.

Emphasis on renewable-energy resources and improvements in energy efficiency might mitigate future global warming by reducing emissions of carbon dioxide and other gases. Although nuclear power offers similar advantages, its bad public image seems, for the present, to rule out any expansion in the U.S. Renewable resources such as sun, water and wind now account for 9 percent of the U.S. energy supply. Energy research is to be the principal focus this year of the House science and technology committee, and members of both houses have introduced "greenhouse" bills that would mandate budget increases for energy research. Representatives Marilyn Lloyd of Tennessee and Philip R. Sharp of Indiana have jointly sponsored a bill that would protect renewable-energy research by providing three-year funding authority.

Fear that the U.S. might lose its position in energy-supply markets and the growing reliance on oil imports are perhaps just as much motives for the new concern. Several witnesses at recent hearings noted that Germany and Japan have overtaken the U.S. in photovoltaic research. Japan spends twice as much as the U.S. and now has the largest share of the market. —*T.M.B.*

Fusion Breakthrough?

Remarkable claims generate excitement and skepticism

Separate investigators claimed in March to have demonstrated nuclear fusion at room temperature by an entirely new technique. Existing methods require either high temperatures and pressures to ignite the fusion fuel or subatomic particles known as "muons" to catalyze fusion at room temperature. The claims face profound skepticism: "I would bet my house that they're wrong," says William Happer, Jr., of Princeton University. If the results are confirmed, however, they could have far-reaching implications for the world's energy future.

The first claim came in a press conference held jointly by Martin Fleischmann of the University of Southampton, England, and B. Stanley Pons of the University of Utah. They reported detecting fusion in heavy water (deuterium oxide) that was electrolyzed with a cathode made of the metal palladium. Palladium can absorb large quantities of deuterium, an isotope of hydrogen. According to Fleischmann and Pons, deuterium atoms become entrapped in the palladium, where they fuse to form heavier nuclei.

The putative evidence for fusion is that large amounts of heat are generated: the energy density is about five megajoules per cubic centimeter, according to Fleischmann—10 times more than from any chemical reaction. The workers say the rate of energy output is greater than the input supplied as electricity. By-products such as neutrons and tritium (another hydrogen isotope) are also said to be produced, but in quantities a billion times smaller than would be expected from known fusion reactions.

The experiments were conducted over five years at the University of Utah at the investigators' own expense. The university has applied for patents covering the technique. A paper describing the results will be published this month in the *Journal of Electroanalytical Chemistry*. Fleischmann says, "Our indications are that the discovery will be relatively easy to make into a usable technology for generating heat and power."

A separate claim of "considerable evidence for a new form of cold nuclear fusion . . . (without muons)" is being made by Steven E. Jones of Brigham Young University. Jones, well known for his work on muon-catalyzed fusion, will describe his work at the May meeting of the American Physical Society in Baltimore. According to an investigator familiar with the work, Jones also relies on electrolysis but does not use palladium and finds a far smaller rate of fusion. —*T.M.B.*

Quantum Cryptography

Single-photon communications can outwit eavesdroppers

For as long as mathematicians have been creating ciphers to keep messages secret, other mathematicians have been breaking them. Whenever an eavesdropper intercepts a message, its encoding algorithm and key can eventually be deduced and the secret compromised. Matters have improved a bit in the 1980's. With the Data Encryption Standard, knowing the algorithm purportedly offers the code breaker no help, and with public-key cryptosystems, even knowing half of the key is not enough to break a cipher quickly.

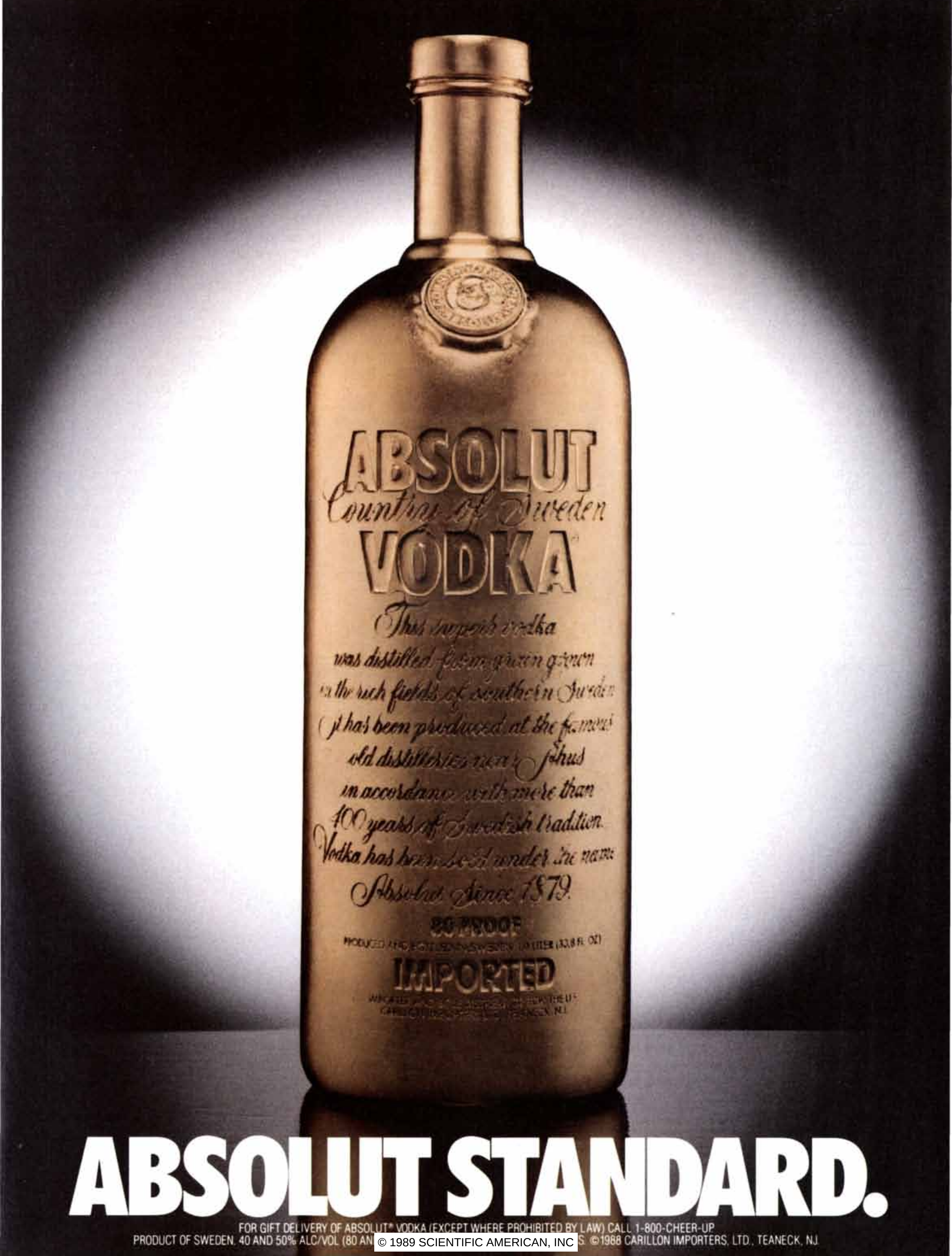
A new technique called quantum cryptography takes another approach: an eavesdropper cannot even listen in on a message without alerting the sender and receiver to the attack. There are a few hitches, of course—transmission is slow, the system can operate over distances of a few kilometers at best and only random bits can be communicated.

Developed by Charles H. Bennett of the IBM Thomas J. Watson Research Center and Gilles Brassard of the University of Montreal, quantum cryptography relies on the fact that measuring a quantum-mechanical system such as a photon irrevocably changes its state and destroys information about the details of the state before measurement. An eavesdropper must make a measurement to gain information, and so any eavesdropping scrambles the original message.

In the physical setup developed by Bennett and Brassard, the "quantum channel" consists of a device that sends single photons whose orientation is either rectilinear (vertical or horizontal) or diagonal (45 or 135 degrees from the horizontal) and a detector at the other end that reads the polarization of incoming photons. If the photon and the detector are both oriented in the same fashion (whether rectilinear or diagonal), the detector will be able to determine the polarization (vertical or horizontal, 45 or 135 degrees) correctly, but if the two are oriented differently, the results will be random: a diagonal photon always has a 50-50 chance of passing through a rectilinear detector, and vice versa.

To communicate the sender transmits a string of bits in a random series of orientations. The receiver chooses another random series of orientations for detection. After the transmission the receiver reports to the sender, through a conventional channel, the series of orientations selected for detection; the sender says which ones were correct. The resulting set of bits, known to both sender and receiver but to no one else, can then serve as a key for a conventional encryption scheme that is capable of sending actual messages.

An eavesdropper trying to measure photons and then retransmit them has an even chance of choosing the wrong orientation and thus passing on a photon that will be read by the intended receiver—using the correct orientation—but will produce a result that does not match the sender's. Any tampering would thus be readily detectable. Noise could mask some eavesdropping, but Bennett, Brassard


ABSOLUT
Country of Sweden
VODKA
IMPORTED
ABSOLUT STANDARD.
FOR GIFT DELIVERY OF ABSOLUT® VODKA (EXCEPT WHERE PROHIBITED BY LAW) CALL 1-800-CHEER-UP
PRODUCT OF SWEDEN. 40 AND 50% ALC/VOL (80 AN
S. ©1988 CARILLON IMPORTERS, LTD., TEANECK, NJ

and Jean-Marc Robert of McGill University in Montreal have worked out error-correcting techniques that make such eavesdropping ineffective.

How might all of this work in the real world? Photomultipliers can detect a photon every 20 to 50 nanoseconds, and so the basic transmission rate is about 20 million pulses per second, but as Bennett points out, each pulse should be attenuated to an average of about .1 photon to reduce the probability of generating a two-photon pulse that could be split and eavesdropped undetectably. Half of the pulses will be thrown away because they have been read in the wrong orientation, and photomultiplier tubes are only about 30 percent efficient in any case. The net rate is perhaps 300,000 bits per second—before factoring in random noise—and the range for sending a single photon is a few kilometers at best.

Bennett has set up a demonstration system, and Brassard is writing the software. They expect to transmit a few thousand bits per second early this summer, enough to show that the idea works. Although quantum cryptography may be impractical, Bennett says, the concept of security guaranteed by physical laws is both appealing and intriguing. *—Paul Wallich*

BIOLOGICAL SCIENCES

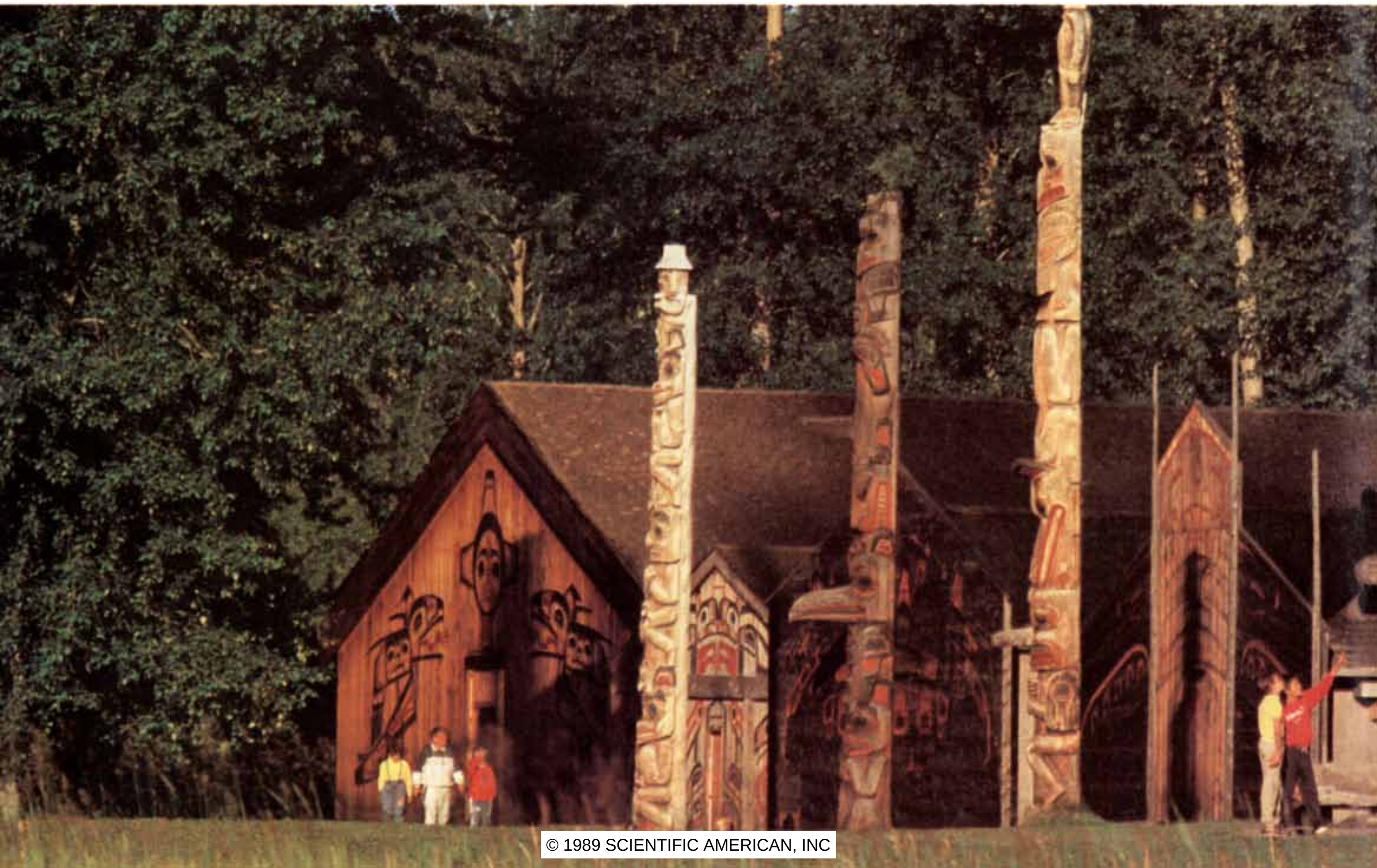
Playing Demigod

Biologists find limits to tinkering with reproduction

What ever happened to the idea of cloning a human being? Not long ago the prospect of generating duplicate individuals from single cells was topical enough to inspire a Doonesbury cartoon and seemed probable enough for the publishers of David M. Rorvik's 1978 book, *In His Image: The Cloning of a Man,* to preface the story with a disclaimer. Yet as biologists have become more familiar with the mechanisms of mammalian reproduction, they have also become increasingly aware of their own limitations.

Whereas procedures such as artificial insemination, in vitro fertilization and in utero genetic testing have become matter-of-fact, other technologies have drifted into the realm of science fiction. The experiments of Karl Illmensee, a former University of Geneva embryologist, offer a case in point. In the late 1970's Illmensee reported in the *Proceedings of the National Academy of Sciences* that he had produced so-called gynogenetic mice by removing the male chromosomes from a fertilized egg and artificially doubling the female chromosomes; in a 1981 article in *Cell* he claimed to have cloned mice by replacing the nuclei of fertilized eggs with nuclei from the cells of a single blastomere (an early embryo). In a 1982 article in the *Proceedings,* Illmensee described mice born from fertilized eggs whose nuclei had been replaced with those from parthenogenetic embryos, that is, embryos that had developed in the absence of the male gamete.

Other workers tried in vain to reproduce Illmensee's results. M. Azim Surani and his colleagues at the Agricultural and Food Research Council's Institute of Animal Physiology and Genetics Research in Cambridge wrote in *Science* that they could not get gynogenetic or parthenogenetic mouse embryos to survive beyond mid-gestation. Davor Solter and James McGrath of the Wistar Institute of Anatomy and Biology told *Science* that embryos created by nuclear transplantation from blastomeres would not develop to term. Meanwhile, accusations by members of Illmensee's laboratory prompted investigations of his research practices that culminated in his resignation in 1987.



Yet Illmensee's experiments were believable because the results he described were, and still are, theoretically possible. Whole plants can be regenerated from single cells, and the nuclei from tadpole cells, transplanted into fertilized frog eggs, can give rise to whole frogs. "In each and every cell is all the information necessary to make a full-grown duplicate of its host," says the mad hero of Jeremy Leven's novel *Creator,* who is trying to regenerate his departed wife. As Leven's hero discovers, however, the chromosomes of some species of mammal thwart such ambitions. Mammalian chromosomes undergo changes very early in development that render much of the information they contain inaccessible.

Why this inactivation happens has become the million-dollar question of 20th-century developmental biology. It does not occur, however, at the same point in development in every mammalian species. The chromosomes in the embryos of certain livestock, for example, retain their ability to direct development longer than do the nuclei in mouse embryos, a fact that Steen M. Willadsen of Alta Genetics, Inc., in Calgary has exploited to clone elite cattle and sheep. Willadsen creates the clones by separating individual cells from four-, eight- or even 16-cell embryos, culturing them for several days and placing them in surrogate mothers. He has also succeeded in fusing the separated cells to enucleated, fertilized eggs—a type of nuclear transfer.

Willadsen does not doubt that cloning will have an impact down on the farm. Research on parthenogenesis and nuclear transplantation might someday influence human breeding as well. One offshoot of such research is the discovery of imprinting, a process that distinguishes in an embryo the chromosomes contributed by the female from those contributed by the male. No one knows the mechanism of imprinting, but it seems to be involved in coordinating the developmental instructions in the two sets of chromosomes. It has also been implicated in malignancy, most recently in a *Nature* article exploring predispositions to bone cancer.

Many biologists think imprinting is the reason chromosomes from a male—not just any chromosomes—are required for normal development in a mammalian egg. It could also be the reason the genes in somatic cells lose their ability to direct development. "You could argue that, if you find the mechanism of imprinting, you could reverse it," says Surani in Cambridge. But he says the chances are remote—which may explain why, in New York City libraries, Rorvik's book is found on the shelves otherwise reserved for the occult. *—Karen Wright*

Quantum Biology

Hydrogen tunneling contributes to an enzyme reaction

It has been known for some years that electron tunneling—a quantum-mechanical effect that enables an electron to circumvent an energy barrier—has a crucial role in many biological reactions, such as photosynthesis. Now researchers at the University of California at Berkeley report that tunneling by hydrogen contributes to an enzyme-reaction mechanism under biologically relevant conditions.

The enzyme, a yeast alcohol dehydrogenase, speeds up the conversion of benzyl alcohol into benzaldehyde, a transformation that involves cleaving a hydrite (a hydrogen atom with an electron) from an alcohol molecule. The enzyme boosts the rate of this transformation enormously by lowering the energy barrier that must be surmounted in order for the reaction to take place.

A semiclassical model compares the binding of the hydrogen to the alcohol molecule with the behavior of a mass on a spring. The model predicts that an ordinary hydrogen nucleus—a proton—can hurdle the energy barrier more easily than the heavier isotopes deuterium and tritium. This means that the reaction rate for ordinary hydrogen should be faster than that for deuterium and tritium. How much faster can be calculated precisely.

If, however, quantum tunneling contributes to the reaction, the reaction rate for ordinary hydrogen should be greater than the rate predicted by the semiclassical model. According to quantum mechanics, a particle's position is uncertain; the probability of finding it at a given point is smeared out in space. Tunneling can occur if the region of uncertainty extends to the other side of an energy barrier. Because particles with smaller mass have a greater uncertainty in their position, they have a higher probability of tunneling. In biological molecules, electrons tunnel readily across distances of tens of angstroms, whereas a proton should tunnel less than one angstrom. The heavier hydrogen isotopes, deuterium and tritium, are even less likely to tunnel.

To see if tunneling is a factor in the reaction rate, Yuan Cha, Christopher J. Murray and Judith P. Klinman prepared two versions of the alcohol, with specific sites on the molecule occupied by an ordinary hydrogen and a tritium in one, and a deuterium and a tritium in the other. During the reaction a benzyl alcohol molecule loses a hydrogen atom to a molecule of nicotinamide adenine dinucleotide (NAD). The workers determined how much of each isotope became bound to NAD. As they report in a recent issue of *Science*, the rate at which ordinary hydrogen was transferred was greater than the rate predicted semiclassically, which would indicate that the reaction is assisted by tunneling.

What is more, the enzyme may facilitate tunneling not only by lowering the energy barrier but also by narrowing it. This could occur if the enzyme brings the active sites on the NAD and the alcohol very close together. "The next step is to see what happens if the molecules are kept farther apart," remarks Cha. She adds that it may be possible to design mutant enzymes that would alter the distance between the molecules.

The observation of hydrogen tunneling could have wide implications, because many biochemical reactions involve the transfer of hydrogen. Preliminary results from another Berkeley experiment suggest hydrogen tunneling occurs in a reaction mediated by plasma aminoxydase, an enzyme in blood. Indeed, Cha says, "hydrogen tunneling could turn out to be common in biology." —*June Kinoshita*

MEDICINE

Balloon Trial

Drugs and "watchful waiting" may suffice after heart attacks

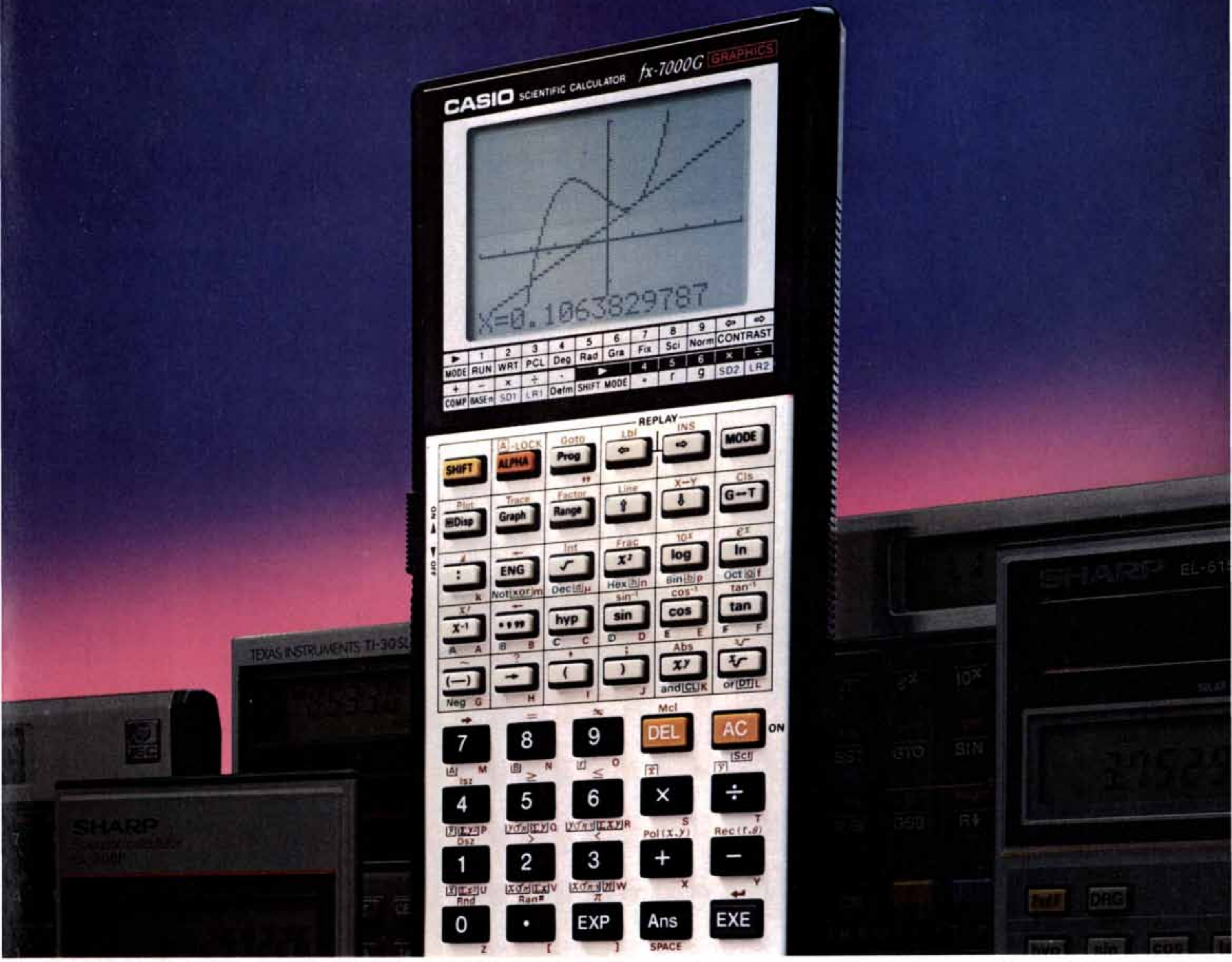
A clot lodges in a coronary artery already narrowed by fatty deposits, cutting off the flow of blood to ceaselessly pumping heart muscle. The victim has a heart attack: the muscle, starved of blood, is injured or killed. Newly available drugs such as tissue plasminogen activator (t-PA), given within a few hours, can dissolve the clot and restore blood flow to the muscle. The reopened artery remains narrow, however, leaving some heart-attack victims with continuing angina, or heart pain, and a high risk of a second attack. Should physicians routinely try to widen the artery so as to forestall such complications?

A large-scale study has tested the value of a common means of doing so, called percutaneous transluminal coronary angioplasty. In the procedure a catheter is inserted into an artery and snaked through the circulatory system until its tip reaches the narrowed coronary artery. A small balloon at the catheter tip is inflated, compressing the deposits and widening the artery. The procedure is valuable for treating existing angina; it seemed likely that, done preventively in the first few days after a heart attack, angioplasty would reduce the risk of future angina or a second attack. Many hospitals now perform it prophylactically.

The study's results, reported in the *New England Journal of Medicine,* suggest this costly practice of routine angioplasty is unnecessary. The study, sponsored by the National Heart, Lung, and Blood Institute and chaired by Eugene Braunwald of Harvard Medical School, included some 3,200 patients in 50 hospitals. All the patients were given t-PA within four hours of the onset of their heart attack.

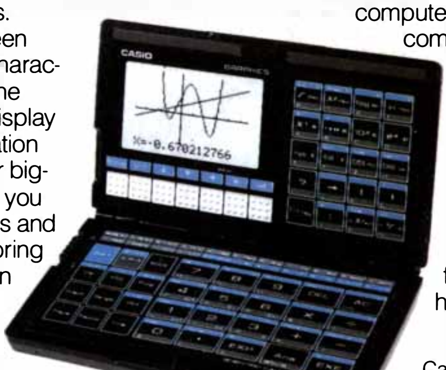
Subsequent treatment of half of the subjects followed an invasive strategy. Between 18 and 48 hours after the drug treatment, a dye was pumped

Graphic proof of high-tech leadership.



The Casio fx-7000G $89.95. It's the world's first programmable scientific calculator with a graphic display. It can give form to your figures, instantly drawing graphs that depict your formulas and calculations.

With a screen measuring 16 characters by 8 lines, the fx-7000G can display enough information for some of your biggest ideas. And you can store graphs and formulas, then bring them back for an instant replay.



With its 422 step memory and 193 functions—including linear regression, standard deviation and computer math—the fx-7000G lets you perform computer functions, without acquiring computer knowledge or incurring computer costs.

For additional capabilities, there's the fx-7500G ($109.95). It has a larger memory capacity, with 195 functions, 4,006 memory steps, instant graph enlargement or reduction feature, plus a convenient horizontal folding format.

If you need even more power, our fx-8000G ($119.95) techs you even higher, with 1,446 memory steps and 225 functions. And when hooked up with our optional FA-80 interface, it works with most parallel dot matrix printers or plotters including Casio's own models.

If you have any doubts as to who's the leader in high-tech calculators, just pick up an fx-7000G, fx-7500G or fx-8000G and draw your own conclusions.

CASIO.
Where miracles never cease

Casio, Inc. Consumer Products Division: 570 Mt. Pleasant Avenue, Dover, NJ 07801
 ...smere Road, Suite 240, Scarborough, Ontario M1H3B7


THE POWER OF
ORIGINAL THINKING.

through a catheter into the patients' coronary arteries to make them visible on X-ray images. If the artery in which the clot had lodged was still partly blocked and was otherwise suited to the procedure, angioplasty was done.

Treatment of the other group of patients was conservative. Cardiac catheterization and angioplasty were done only if recurrent angina or abnormal electrocardiograms indicated that the heart was still blood-starved. A total of 13.3 percent of the conservative-strategy patients ended up having angioplasty, in contrast to 56.7 percent of the invasive-strategy patients.

Yet the two groups fared equally well over the next six weeks. During that time 10.9 percent of invasive-strategy patients either died or suffered a second nonfatal heart attack; the figure was 9.7 percent for conservative-strategy patients. The groups were also comparable in other measures of cardiac function. The figures are quite favorable; even with clot-dissolving drugs, as many as 20 percent of patients in other studies have suffered reblockage of the artery within weeks.

That such favorable results can be had by drug treatment followed by "watchful waiting," with invasive measures held in reserve, should be welcome news to people concerned about the growing cost and complexity of medical care. It means, write the study's authors, that most heart-attack patients can be treated in community hospitals lacking facilities and personnel for cardiac catheterization, provided the patients can be transferred to a more sophisticated hospital if need be. An editorial accompanying the *New England Journal* report puts the potential savings at $200 million a year. *—Tim Appenzeller*

Risky Business

Has the danger of cancer from pesticides been underrated?

How healthy is the apple in a child's lunch box? One of Washington's most powerful environmental litigators, the Natural Resources Defense Council (NRDC), has claimed that residues of agricultural chemicals on fruits and vegetables eaten during childhood could be initiating between 5,500 and 6,200 cancers every year. The Environmental Protection Agency and many independent experts dispute the claim, but it has stirred widespread concern.

In a study called "Intolerable Risk: Pesticides in Our Children's Food," the NRDC presents a formula for assessing the risk of cancer that takes into account the many years that usually elapse between the initiation of a cancer and its diagnosis. The formula suggests that exposure to carcinogens in the diet before the age of six accounts for more than half the lifetime risk of diet-caused cancer.

The modified formula, by giving greater weight to children's preference for fruits and vegetables and to their low body weight, yields alarming estimates of the lifetime cancer risk from residues of agricultural chemicals. The calculations suggest that 90 percent of the total risk from artificial chemicals is from a metabolic by-product of Alar (daminozide), a growth regulator used on peanuts and on apples to stop them from dropping off before they have ripened.

Among other things, the NRDC's critics object to the starting point of the projections: the estimates of the chemicals' carcinogenicity, which are derived by feeding high doses to rats and mice. The effects on human beings of far lower doses are usually calculated by assuming that risk decreases linearly with decreasing dose and with increasing metabolic rate.

Bruce N. Ames of the University of California at Berkeley, a well-known skeptic about the dangers of pesticides, maintains that most evidence suggests this standard extrapolation greatly exaggerates the risk at low doses. Ames lambastes the NRDC study as "wildly inaccurate hysteria," saying he doubts that pesticide residues cause any cancer deaths at all; he argues further that many naturally occurring toxins have far more potential to cause cancer than pesticide residues do.

At a Congressional hearing in March, John A. Moore of the EPA also criticized the NRDC's figures for the carcinogenicity of Alar. The figures were rejected by the EPA's scientific advisory board when the agency first considered banning the chemical, in 1985. In addition, he challenged the NRDC's data on food consumption, saying that they are based on small surveys. The NRDC's Janet Hathaway replies that the potency data are still the best available and that the EPA's own consumption studies are inadequate because they date from the mid-1970's, when people consumed fewer fruits and vegetables.

Even though the EPA disputes the NRDC's figures, it agrees that Alar is hazardous. The agency, which was leaked the NRDC's conclusions in December, announced on February 1 that it intends to ban Alar, even though new cancer-potency studies the EPA has requested are still incomplete. The EPA currently estimates that Alar might cause 45 cancers per million people exposed for a lifetime; its declared policy is to accept no more than one death per million. Alar's manufacturer, Uniroyal, has the right to several levels of appeal, which could delay a ban for years.

The use of Alar on apples has decreased in recent years, although it is not clear by how much. It is clear, from one study on diet by the National Research Council, that people who eat a large quantity of fruits and vegetables seem to be less prone to cancer, not more, which suggests that the benefits outweigh any risks.

The putative dangers of Alar aside, the NRDC report includes a number of broader recommendations. The EPA's Moore testified that he agrees with some of them, such as the need to improve the agency's risk-assessment procedures. The EPA has already asked the National Research Council for advice on assessing risk in elderly people as well as in infants. "Sensitive populations is an issue we are going to have to address," concurs Ronald W. Hart, director of the Food and Drug Administration's National Center for Toxicological Research.

At the Congressional hearing Moore bemoaned "administrative impediments" that prevent the EPA from banning risky pesticides quickly. There are signs that Congress is willing to listen. —*T.M.B.*

PROFILE

Do Bees Think?

The discoverer of bat "sonar" thinks about animals' thoughts

In many respects he is the archetypal establishment scientist. Some 50 years ago, while still a mere student at Harvard College, he discovered that bats perceive the world by bombarding it with their cries and listening to the echoes. As a professor at Cornell University and then at Harvard, where he was eventually made chair of the biology department, he became one of the world's experts in the navigational skills of bats and birds. In 1965 he left Harvard to found a center for ethology at Rockefeller University in New York City, where he is now a professor emeritus. He has even written four articles for SCIENTIFIC AMERICAN (December 1948, August 1950, March 1954 and July 1958).

For the past 15 years, however, Donald R. Griffin has been rocking the establishment boat. In papers, speeches and two books he has challenged those behaviorists, cognitive psychologists and biologists who treat animals as mindless automatons responding reflexively to environmental and genetic prodding. Griffin contends that it is possible and even likely that our fellow species—from chimpanzees to bees—experience the world in a way not wholly unlike our own. He maintains, moreover, that if we scrutinize animals closely enough we may glimpse something of their mind-set.

At 73 Griffin himself looks somewhat like a bird of prey: long, lean face, prominent nose and deep-set eyes under a brow crowned by a silver crewcut. In an interview at his office at Rockefeller University, he gamely swoops in on the most elusive questions. Asked to define "consciousness," he talks about flexibility of response, about the ability to remember the past and to form a plan of action for the future, about having "intention." He proposes that, despite the contention of some philosophers, one can be conscious without being aware of one's self as a distinct entity. He tries to untangle the differences between consciousness, cognition and intelligence. Finally he pulls back, saying, "I try to avoid getting into semantic thickets."

He seems most comfortable simply describing the remarkable things animals do, things he believes the conventional stimulus-response paradigm poorly accounts for, things rife with intention: a chimpanzee searching for a stick that it can use to extract termites from their nest; a plover feigning injury to a wing to lure a potential predator away from its hatchlings; a honeybee waggling its abdomen to inform its fellows of the whereabouts of a field of flowers. "We aren't used to thinking of insects as conscious animals," he says, "but once you raise that possibility, I think a great deal of evidence supports it."

Griffin cheerfully concedes that his past successes allow him to indulge in his present iconoclasm. "If I had said these things 30 years earlier it may have been maladaptive, to use a biological term," he says, "but I just don't care anymore." On the other hand, half a century ago many scientists also looked askance at the notion that bats locate objects by sonar. Griffin recalls a meeting in 1940 at which he and an early collaborator, Robert Galambos, had described their findings: a prominent biologist grabbed Galambos and shook him while exclaiming, "You can't really mean that!"

Growing up in the wilds of Scarsdale, N.Y., and then Barnstable, Mass., Griffin was fascinated by all animals. He trapped deer mice and skunks and subscribed to the *Journal of Mammalogy* and *Fur-Fish-Game.* Bats were always a favorite research subject. As a teenager he crept through old houses and caves throughout Massachusetts in search of little brown bats, which he banded in order to learn about their migratory habits. The recovery of some of the banded animals more than 20 years later provided the first proof of bats' longevity.

The banding expeditions also led him to wonder how the bats managed to flit so deftly through the caves' dark, twisting passages. After entering Harvard in 1934, he learned that some biologists had speculated that bats might emit ultrasonic cries—at a pitch beyond the range of human hearing—in order to navigate. He also found a physics professor, George W. Pierce, who had an instrument capable of detecting ultrasound. "It was a happy accident," Griffin recalls.

In a series of experiments, first with Pierce and later with Galambos, a fellow student, Griffin discovered that bats do indeed emit staccato ultrasonic "chirps" as they fly; they also become disoriented when their ears are blocked or when ultrasonic noise drowns out the chirps' echoes. In subsequent studies, Griffin demonstrated that some bats can detect objects less than a millimeter in diameter and that they employ echolocation not only to avoid obstacles but also to hunt insects. As a brown bat closes in on a moth, Griffin says, the tempo of its chirps speeds up from the "idling of an outboard motor to the buzz of a model-airplane engine."

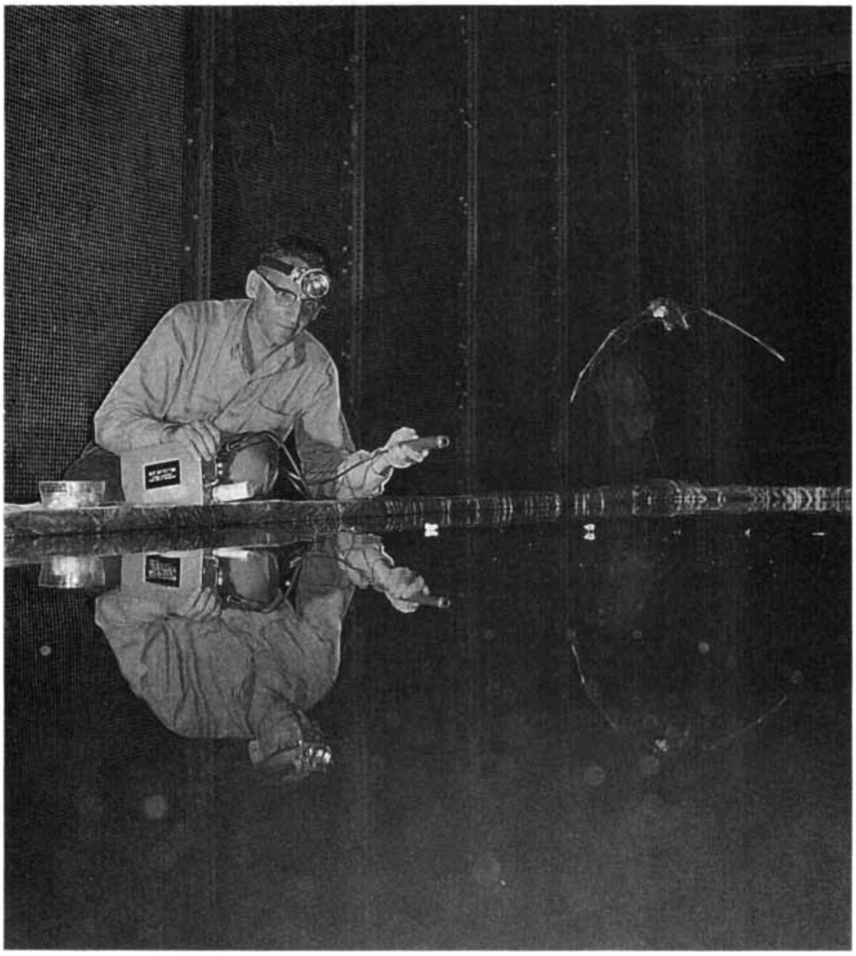
Griffin's other great obsession was the ability of birds to navigate and orient themselves with respect to the earth. In 1941 he learned to fly and even bought an airplane so that he could study the homing ability of gulls. In later years he tracked migrating birds with radar. As a result of his own work and experiments by others, Griffin has concluded that birds depend on landmarks and perhaps even the stars and sun as navigational aids, but he is skeptical of recent proposals that they orient themselves by sensing the earth's magnetic field. The issue remains, he says, "an uncracked nut."

Griffin did not test his teeth on the

nut of animal consciousness until 1974, after the philosopher Thomas Nagel visited him at Rockefeller. Nagel was writing a paper called "What Is It Like to Be a Bat?", which inquired whether human beings can know anything about the subjective experiences of other species. Although Nagel came to a negative conclusion, Griffin went on pondering the issue. He began to question why he and most other scientists had always shied away from any discussion of an animal's mental life as "unscientific." Some of the conventional objections were easy to dismiss, such as the position that ascribing mental states to animals smacks of anthropomorphism. "That's circular thinking," Griffin says. "You're arguing that animals can't think by saying that thinking is uniquely human."

Testability, that fundamental principle of science, posed a knottier problem. Griffin had to agree with Nagel that in an absolute sense it is impossible to either prove or disprove that animals have conscious thoughts. Yet to surrender completely to this fact seemed to Griffin to reflect a kind of solipsism—"species solipsism." If humans can infer each other's thoughts through language, he reasoned, might scientists not infer something about animals' thoughts, by studying the ways they communicate, for example? "We are stuck with a situation where you can only weigh likelihoods," he explains, "but that's not unusual in science. We speculate about the first few seconds of the universe without being able to be there and do nice tidy experiments."

Griffin has challenged scientists studying animal behavior to overcome "species solipsism"



DONALD R. GRIFFIN records the ultrasonic chirps of a fish-catching bat at a research station of the New York Zoological Society in Trinidad in the early 1960's. The bats employ echolocation to detect fish disturbing the surface of the water. They gaff the fish with their long, curved talons.

Griffin has presented his case in two books so far: *The Question of Animal Awareness*, published by the Rockefeller University Press in 1976, and *Animal Thinking*, published by the Harvard University Press in 1984. The reception of his views has been mixed. "I've been accused of turning the clock back to Romanes [a 19th-century British biologist who inferred elaborate mental processes from animal behavior], but I like to think we are turning the clock ahead to the 21st century and escaping from behaviorism." Behaviorism rejects subjective experience as irrelevant, but its opposite number, cognitive psychology, also leaves Griffin cold. Although its practitioners reestablished such mental functions as memory, perception and learning as valid subjects for scientific inquiry, Griffin notes, they are more concerned with "information processing" than with subjective experience.

To those who suspect Griffin of romantic and even mystical notions about nature, he points out that he is an agnostic. "Natural selection is quite adequate for explaining nature," he observes. He is also neither a vegetarian nor an animal-rights enthusiast. "I think animal suffering should be kept to a minimum," he says, "but when the hard choice comes to human benefits versus animal benefits, I'm afraid we have to consider our own species as more important." He does not even own a pet. "My wife and I travel too much."

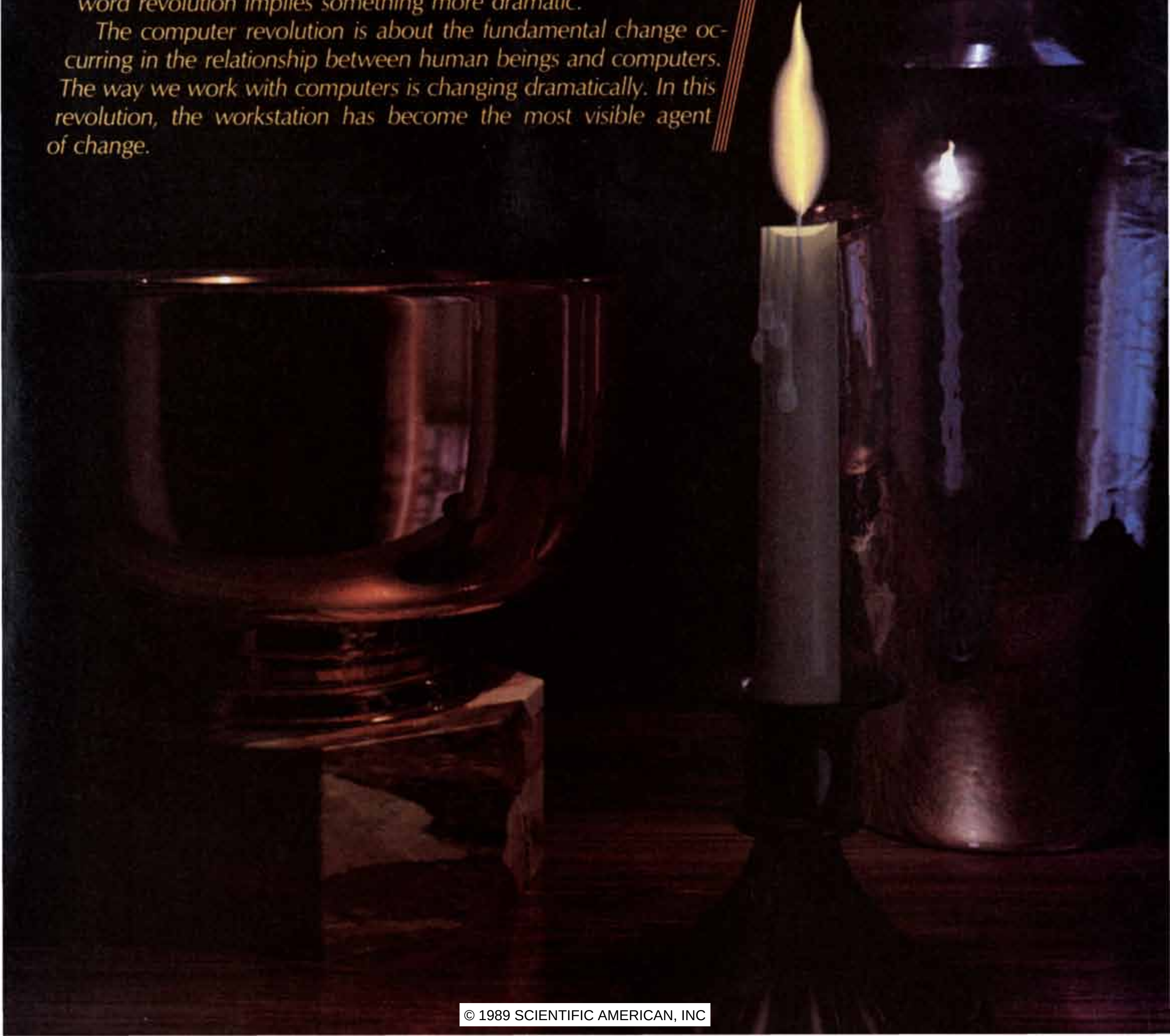
Although he officially retired from Rockefeller three years ago, Griffin remains busy. He is investigating the role of acoustics in the communication of honeybees with a group at Princeton University, near his home. He periodically tramps through the New Jersey Pine Barrens and the Adirondack Mountains in search of clues to the mind-set of another favorite animal, the beaver.

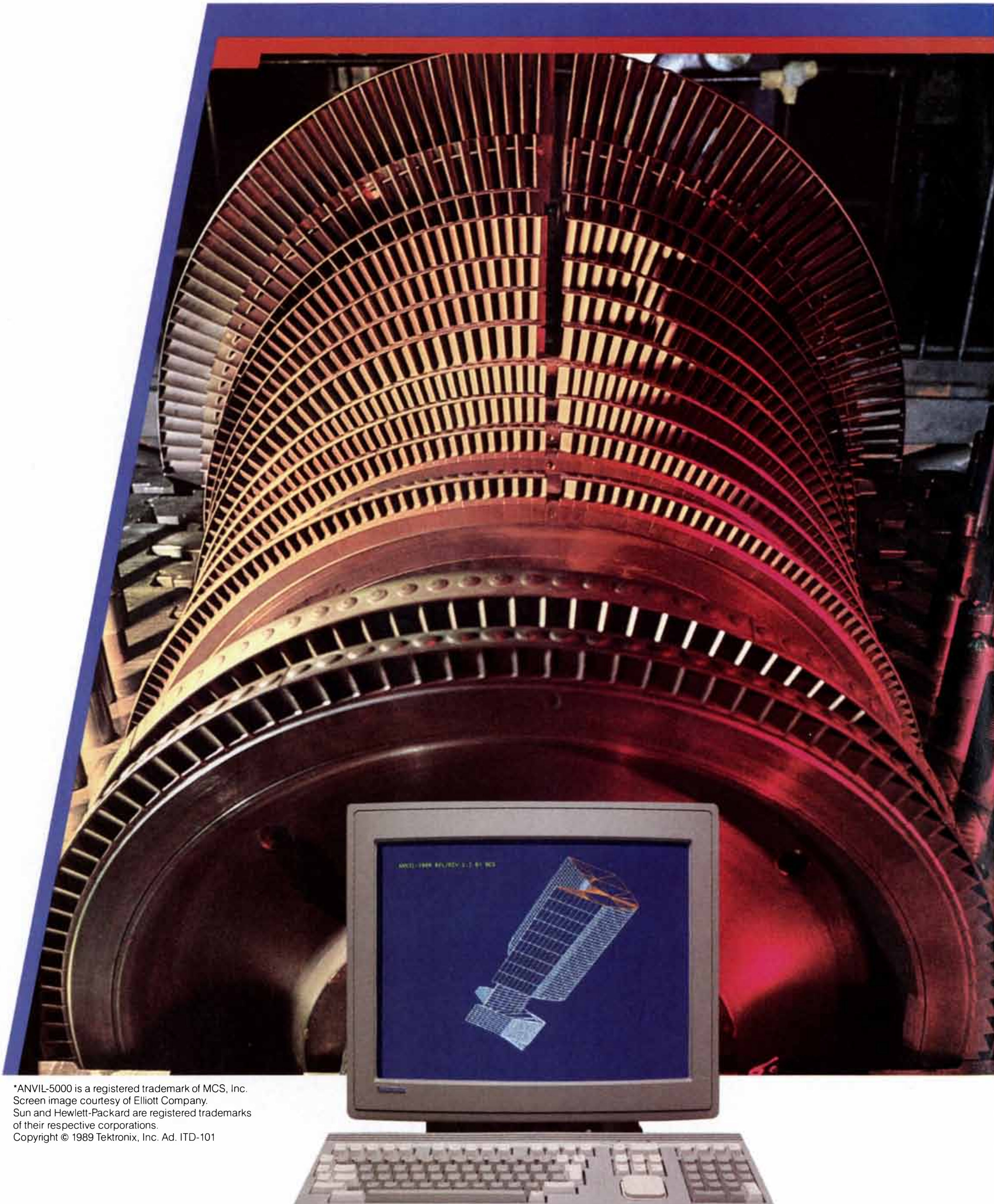
He is also searching for fresh examples of animal ingenuity to cite in yet another book he is writing on animal consciousness. One example he has turned up in the scientific literature is the honeyguide, an African bird that enlists the help of humans to obtain the honeycomb it craves. With distinctive cries and flight patterns the bird attracts the attention of a potential helper and leads him to the hive, which may be more than a kilometer away. If all goes well (for the honeyguide), the helper will expose the hive and leave enough of the honeycomb for the honeyguide to gorge itself. "This seems sort of intentional to me," Griffin remarks. —*John Horgan*

THE WORKSTATION REVOLUTION

Every year new stories about the emergence of faster, more functional, and, in defiance of generations of economists, cheaper computers appear. And every year come new claims that we are in the midst of a revolution. We are in the midst of a revolution. It is not simply because computers are getting smaller and cheaper. The word revolution implies something more dramatic.

The computer revolution is about the fundamental change occurring in the relationship between human beings and computers. The way we work with computers is changing dramatically. In this revolution, the workstation has become the most visible agent of change.





*ANVIL-5000 is a registered trademark of MCS, Inc.
Screen image courtesy of Elliott Company.
Sun and Hewlett-Packard are registered trademarks of their respective corporations.

Agents of Change:

Workstations are both cause and effect in a changing computer environment

by Richard M. Fichera

The very size of the first generation computers helped create the massive centralized computing complexes and equally centralized DP departments. Added to that was the fact that computers were very early woven into the main enterprise systems of corporations—payroll, accounting, etc.—where their cost could be easily justified. By their nature, these were centralized tasks. The model of computers as a central resource was perpetuated.

The problem with this early model of computing is that it did not reflect the way people like to work. Computers, like many other machines, are good servants but poor masters. When the emergence of computers forced organizations to modify the way people dealt with their work and with each other, distortions and inefficiencies crept into the working environment.

The Motivation—Individual Control of the Working Environment

According to Bill Paduska, founder of Apollo and now Chairman of the Board and CEO of Stellar Computer, Inc., one of the main reasons workstations are being adopted at such an overwhelming rate is that they address a "fundamental desire on the part of workers to exercise individual control over their own work environment and resources." The workstation provides this control.

The Enabling Factors

Workstations came into existence upon the confluence of a number of interrelated technologies. Key among these were the advances in semiconductors (particularly microprocessors and memory chips) and networking.

As the demand for chips grew, it enabled vendors to bring on-line large efficient manufacturing plants that could produce large quantities of chips at a lower cost per chip. This drove prices down, which in turn stimulated demand. The result has been that computers have shown a continuous decline in price relative to performance and they are likely to do so in the future. Only as the system cost contribution for semiconductors approaches and falls below that of electromechanical components, enclosures, displays and peripheral devices will the rate of decrease flatten out.

When microprocessors were first released as a commercial product they were regarded as little more than a toy. Yet the microprocessor price/performance evolution has been similar to that of memory. The performance of commercial microprocessors has continued to evolve from tiny devices suited for controlling sewing machines to processors that are capable of rivaling almost anything that today is called a minicomputer.

The cover illustration of this section, entitled *Candle,* was created by John Grower at Wavefront Technologies. The image demonstrates the high-end capabilities of Wavefront Technologies' software, producing accurate reflections and shadows. The database was created using Wavefront Technologies' Model and PreView programs on a Silicon Graphics Iris Workstation. The final image was rendered on an Edge Computer Server using Wavefront's Image program.

RISC—The Next Evolution in Microprocessors

Today there are two types of microprocessors vying for supremacy—the CISC (Complex Instruction Set Computers) chips and the RISC (Reduced Instruction Set Computers) chips.

On the CISC chip side, the Intel and the Motorola families of chips are the current market leaders for workstation processors. They share many of the architectural features of the larger computers of the 1970's, including the use of complex instructions composed of many steps of lower level microcode stored on the chip.

In recent years, work originating at IBM, Stanford University and the University of California at Berkeley has given rise to a new breed of chip, the RISC chip. These systems use a simpler instruction set that is directly executed by the hardware. Typical programs require more instructions, but the instructions are executed faster. RISC chips are physically smaller and cheaper to produce for equivalent performance levels, and they appear to offer a two to five times increase in performance over CISC processors implemented in equivalent technology.

The earliest major vendor to make a major investment in RISC technology was Hewlett-Packard, which began work on its RISC-based Precision Architecture in 1981. The public announcement of the product was made in 1986. According to Michael Mahon, system architect and manager of the Precision Architecture project, HP made a decision that RISC architecture offered sufficient benefits to make it the "convergence architecture" that would unite all current product lines and serve as a platform for future growth.

RISC is an appropriate architecture for a wide range of systems, Mahon says, adding that "people tend to think of it as a workstation solution only because the majority of products that have been implemented with RISC processors are workstations."

The two early commercial pioneers of RISC technology were MIPS Inc. and Sun Microsystems. While there are major technical differences between the products, both vendors succeeded in delivering products that had higher performance than alternative processors. With its announcement of a new series of RISC workstations based on the R2000 chipset from MIPS, Inc., DEC is now able to offer advanced workstations which are competitive with those from Sun, Apollo and other newer workstation vendors.

The continued healthy competition between new RISC architectures and existing CISC designs will continue to fuel the rapid evolution of workstation platforms into the 1990's, with no slowdown in the rate of performance growth.

Networking

In addition to improvements in high-performance processors, the rise of workstations was made possible through advances in networking that enabled users to share information and provided them access to remote corporate resources. The inclusion of integrated networking was a major differentiating feature of the first Apollo workstations, and it has become an assumed function on all subsequent workstation products. There are two basic types of networks in use in the workstation and PC world—Ethernet and Token Ring. Both have now become standards, with IBM initially promoting the Token Ring and most of the rest of the industry favoring Ethernet. Ethernet has been successful enough that IBM is also supporting it as a standard.

New Computing Architectures

But although semiconductor and networking advances were key ingredients in the development of workstations, they alone were not sufficient to cause their initial creation. Denser memories and faster semiconductors were also available to the designers of large central machines, but for some time they were used only to create better mainframes. The final enabling factor that gave rise to the workstation was a new system architecture based on a new vision of how the elements of the computer and the user should interact.

The new vision was one that called for the user to become actively involved with the system rather than remaining passively outside it. And the kind of architecture that allowed that to happen was one that combined a graphic user interface with an interactive operating system. The concept of rapid interactivity is crucial to bringing the user "inside" the creative loop.

According to co-founder Russel G. Barbour of Apollo Computer, now Vice President of Advanced Technology, "There is a 'magic' threshold for interactive system response time of around 150–200 milliseconds. Beyond this, a user's attention tends to wander, and productivity falls off."

PCs—From Toys to Giants

Just as the workstation was making its debut in the early 1980's, an equally important and not unrelated event was also occurring—namely the emergence of the IBM personal computer.

The PC's unexpected success was due in part to IBM's opening up of the system to allow third-party development of hardware and software. The opening of the architecture allowed massive investments by innovative third parties. When the initial wave of microcomputer software packages—some of them ported from earlier CP/M platforms—hit the streets, market demand began to climb.

Initially the PC user community consisted of a very much nonmainstream mixture of hobbyists and early adopters. But in short order, they gained commercial acceptance on the strength of word-processing and spreadsheet applications. Still, at that point, they were clearly not suited for technical problems, lacking adequate processing power, memory capacity and graphics capability.

But subsequent introductions of new Intel processors, more capable graphics and an adequate interim solution to memory limitations began to change that. Today operating systems such as Unix and OS/2 offer "big computer" software environments to PC software developers and users. The culmination of this evolution is a "PC" that today incorporates an Intel 80386 with a performance level of 3.5–5 MIPS, several hundred megabytes of disk, and much of the software functionality previously found only in technical workstations. Major technical applications, such as CAD/CAM and structural analysis, now run on these "PCs." In addition to PCs, the Apple Macintosh has emerged as a serious platform for a variety of CAD and other graphics oriented applications.

Workstations—Moving Down from the Top

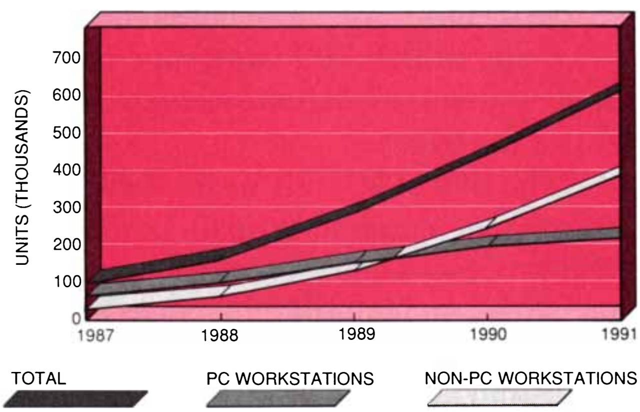
While PCs were acquiring workstationlike characteristics, workstations were expanding in both directions from their initial $25,000–$40,000 price range. On the low end, the drop in component prices has made it possible for workstation vendors to offer systems with all of the traditional features that differentiated workstations from PCs—virtual memory, window oriented graphics interfaces and transparent networking—for prices that now overlap those of high-end PCs. Depending on the configuration, low-end technical workstations now carry prices under $5,000.

Sun Microsystems was founded shortly after Apollo, and in launching its own line of workstations it displayed some major technical and philosophical differences. Most significantly it endorsed open systems and standards (Unix, Ethernet).

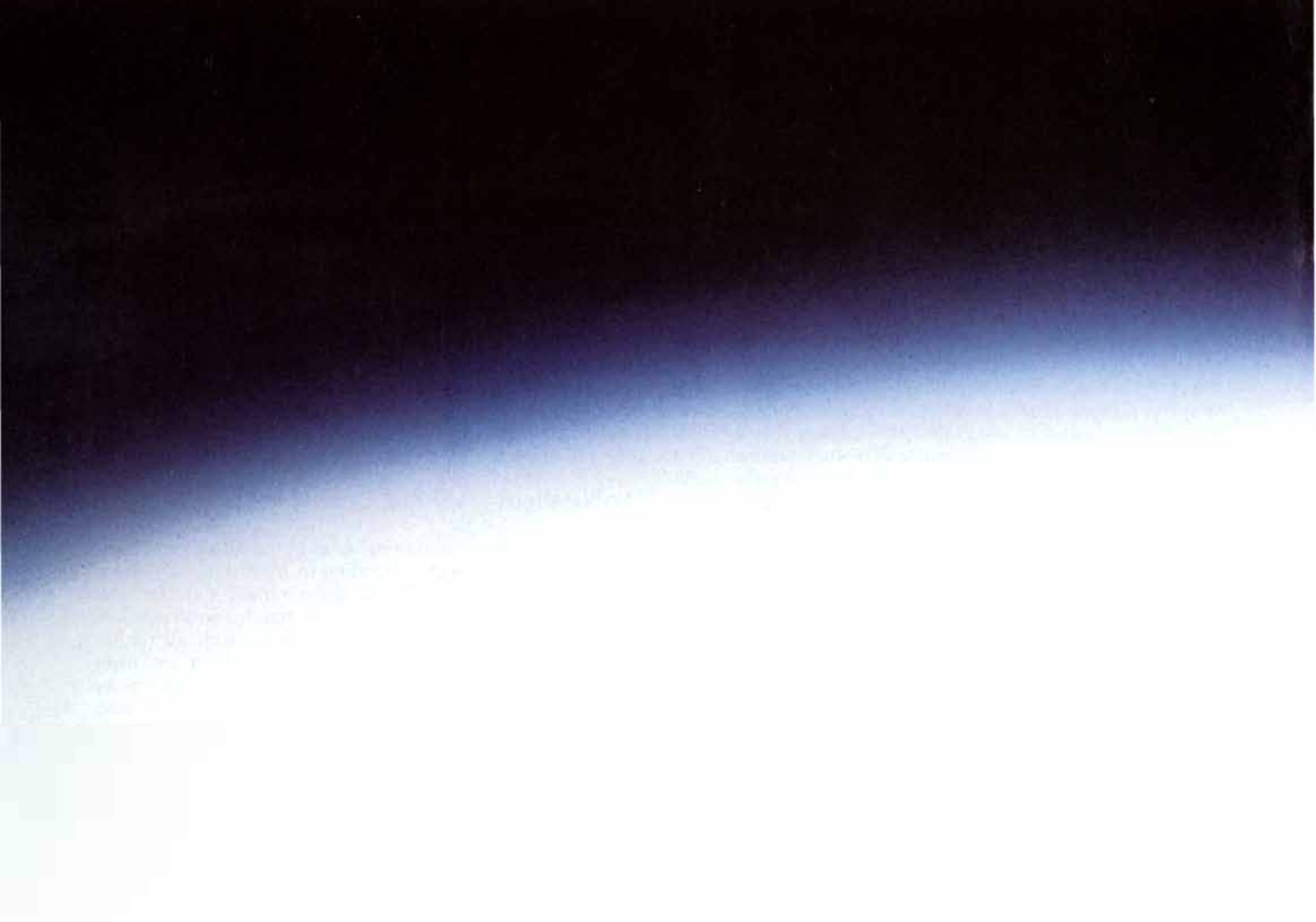
The Convergence

Although people still tend to think about personal computers and technical workstations as two distinct classes of machines, the changes and advances each have made in recent years has almost obliterated their traditional differences. Today processor performance, software capability and graphics features are almost identical. Most of the

TECHNICAL WORKSTATION FORECAST
ESTIMATED SHIPMENTS



Increasing workstation demand: **These figures showing traditional technical workstations and high-end PC forecasts indicate that shipments will approach 700,000 in 1991, from a base of 94,000 at the end of 1987.**



Before he pushes the edge, IBM takes you beyond it.

From miles away and thousands of feet below, you know exactly what he's feeling.

Using an IBM 3090 with supercomputer power, you've pushed your design through relentless iterations of finite element modeling and fluid dynamics analysis. You've defined the edge.

The IBM 3090 with Vector Facility, for its part, has redefined supercomputer performance, combining up to six vector processors with one of the industry's fastest scalar processors and most efficient memory management systems, parallel processing capabilities, and full IBM System/370 compatibility.

The result is balanced supercomputer performance that can improve your job turnaround by eliminating data bottlenecks, and save the added expense of a separate

front-end processor. Plus, the IBM Vector Facility's modular design lets you start with just the computer power you need today, and make economical in-field upgrades as your needs grow. And that gives the IBM 3090VF a big edge over more expensive and less flexible vector processors.

IBM's integrated system approach combines the power of the IBM 3090VF with the advantages of a single operating environment, low entry cost, low software and maintenance costs, a wide variety of application programs, and IBM support to create a winner in total cost-of-computing.

To find out more, or to arrange to have your IBM Marketing Representative contact you, simply call 1-800-IBM-2468, ext. 20. IBM®

remaining differentiation has been in terms of software environments and processor type.

Sun's 386i is an Intel-based workstation, running under the same SunOS (Unix), user interface, and network environment as the rest of Sun's line, but it also runs MS-DOS software as well. It is intended to be the universal workstation, integrating the world of PC software with the world of networked Unix workstations.

At the same time that the hardware platforms are converging, so too are many of the application requirements. New commercial applications, such as financial analysis and brokerage systems, now have system performance requirements similar to many technical computing applications.

The result of this overlapping of formerly separate systems and requirements is that corporate planners now are closer to being able to implement an integrated strategy for supporting all of their business and scientific applications with a limited number of platforms in a common network environment.

The ability to implement some form of consistent network environment, in some cases including networked workstations, has been available in various forms from established vendors for several years. What has been missing is the ability for users to create an integrated environment that did not tie them down to one vendor's equipment, software and possibly lagging price performance. That problem is beginning to be solved with the rise of open systems and standards.

The Rise of Standards

Standard software and hardware environments have been a major factor in the growth of the workstation and PC industry. Indeed, the incredible success of Sun workstations and personal computers in general can be credited to their active championing of standards, specifically the DOS and Unix operating systems.

Unix and PC-DOS—The Two Standards

The most rapidly growing operating system is PC-DOS and its public cousin, MS-DOS. It has created, in terms of the number of users, amount of software and total dollar revenue, the largest software industry in the world. It was the first operating system to be so well standardized that plug and play hardware clones could be built. And it was directly responsible for the success of the PC.

The workstation world, in contrast, has benefited from the emergence of Unix. Originally written at the AT&T Bell Laboratories, Unix was a tool written by very advanced developers with the intention that it would be used by other advanced developers. It was never intended to be a serious commercial operating system, and the commercially available versions today, while still identifiable as Unix, have been massively rewritten to accommodate required capabilities.

After initial official resistance, AT&T agreed to take an active role in the commercialization of Unix. It began gaining momentum when it was adopted by start-up workstation vendors, including Sun Microsystems. These companies saw that the existence of compatible machines would overcome resistance on the part of potential users to take a risk on an unproved vendor. In addition, they realized that by adopting an existing operating system, they could vastly reduce their product development cycle. Unix is now the second most rapidly growing operating system environment in the industry, outstripping all existing proprietary minicomputer and mainframe operating systems.

Open Systems

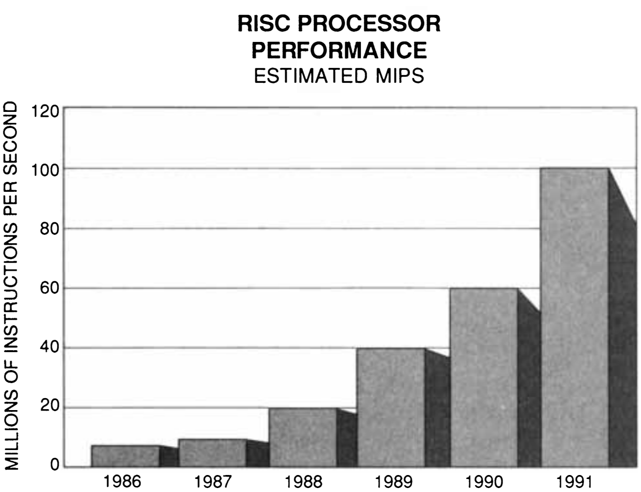
One difference between the DOS and Unix operating systems is that DOS has been an open and highly standardized system from the beginning. Only recently has the Unix world started to move strongly in the direction of providing an open system—that is, defining the interfaces at all levels of the system—from physical to high-level software.

The goal of open systems architecture is to allow third parties to add and interconnect both software and hardware with a high degree of confidence of success.

For users open systems have many advantages. They provide a guarantee that there will be a steady stream of competitive products. Through standardized software and operating procedures they lower the barriers to switching vendors. And most significantly they preserve current investments in peripherals and software. This last point is critical because the value of software often exceeds the cost of hardware for an increasing number of users.

Open systems in effect "levels the playing field" when it comes to evaluating new solutions and it factors out the need for users to consider during an evaluation the cost of converting from one vendor's hardware and software products to another. The net effect of this will be to ensure a stream of technologically aggressive products coming from vendors who can no longer rely on negative factors—expense and difficulty of migration—to keep their users loyal.

For established vendors attempting to protect their installed bases, open systems are a major new problem to grapple with. Some of the more established vendors have been accepting this change and have begun to actively seek out the best in current market technology and offer it to customers in a role of integrator. As basic workstation platforms become more and more of a commodity, this is quite likely to be a role many of the current major companies will have to adopt.



RISC processor performance: **The pace of performance improvement in RISC processors is rapid, with expectations that by 1991 RISC processors will eclipse all but the fastest of today's mainframe processors.**

The Future View

For all the hardware improvements we have already seen, the end of the rapidly improving price/performance curve is nowhere in sight. One can expect that significant advances will still be made on the strength of RISC processors, next generations of 4M-bit and 16M-bit RAM, and new storage options.

Indeed, Linda Fabel, IBM Director of Technical Workstations, noted that IBM has stated that within a few years desktop workstations will be available with the power of the current IBM 3090-200E with vector facility—a large mainframe by today's standards.

Better Tools

Although at first glance it may seem that there is no need for such incredibly powerful workstations, most industry participants believe that users will have no trouble finding ways to absorb these extra CPU cycles. These new products will be used both in performing today's work faster and enabling new solutions. This means that the growing population of workstations will not simply be used to offload the work of existing mainframes but, as Edward H. Robbins, IBM Director of Scientific and Technical Computer Systems observes, they will primarily be involved with new applications.

More specifically we can expect that the applications of tomorrow will include the following:

Expert systems—After a classic boom and bust cycle, expert system applications are beginning to gain some acceptance. Possible commercial applications will be interpreting high-volume data such as economic and financial data and for solving traditionally defined "well bounded specialist" problems where, to date, computers have not been used so successfully.

Massive computation—Major analytical models for finance, trading and brokerage, mechanical design and manufacturing process optimization will become more accessible. With resources on a desktop that will rival today's minisupercomputers, many sophisticated analytical tools that are too expensive today will become readily available.

Distributed DBMS—The problems of fully distributed database management systems (DBMS) have not yet been solved. What is certain is that they will. When that happens, it will remove one of the last organizational barriers to eliminating central MIS mainframe resources, except when they make sense as servers on the network.

Text database—Another growing application area that will eventually be linked with networks, distributed processing and expert systems, is text retrieval. Now a fairly expensive process, and one that can require massive computation, it will yield to advances in software technology and raw computing power. Text retrieval, once it is freed from constraints imposed by current technology, will be driven by its fit with the way users want to work: much of the information we deal with on a daily basis is in textual form rather than in fields and records.

A basic model for new applications, says Bill Joy, founder and current Vice President of R&D for Sun Microsystems is that "every order of magnitude of performance improvement makes a new class of applications possible to perform in an interactive environment." Joy also commented that the fundamental nature of innovation is that it is unpredictable, making it difficult to forecast which applications will be the next big winners.

Integration of Computers and Communications

As hardware technology improves, we will also continue to see a merging of computers and communications. It is a merger that has been driven by the requirements of very large corporations and it will yield systems that have tightly integrated computing and communications features, making it easier for organizations to achieve the right balance between the physical location of resources and physical location of workers.

Digital Equipment's John Adams, Group Manager, Local Area Networks, comments that within five years it will be commonplace to see large organizations managing networks with 10,000–100,000 nodes. Digital, currently managing the world's largest network with approximately 32,000 nodes, can expect to have well over 100,000 by then.

Integrated Systems & Global Enterprise

The increasing capabilities of advanced networking and the increasing ability to interconnect systems will make tomorrow's global enterprises that much more efficient. Increasingly, companies will be able to concentrate on what they are doing instead of how they go about getting it done.

This is an area where existing vendors, such as AT&T, IBM, DEC and Unisys, will have a major impact. With their advanced technology and experience in integrating large global networks they will be able to provide large-scale communications and computing resources on a global scale.

This will be important because, as DEC's Adams notes, finding ways to efficiently interconnect multiple large corporate networks will be a major problem in the future. Gerald V. Butler, Vice President, Engineering & Scientific Products, Prime Computer, agrees, adding that having the ability to solve these large system level problems will be an important value-added skill for established vendors. The ability to integrate existing applications with leading edge hardware and networking, he notes, is lagging by several years.

The Fluid Workplace

And what of the role of the workstation in this evolving world of networked computers? Apollo's Barbour describes it as "the users' window into an integrated computing environment."

It is a description that conjures up an image of a world where the users' work model is a logical task model, uncluttered by irrelevancies of physical resources location. Local and global resources will be flexibly available, and users will be able to move about in organizations and take their working environment with them.

Although this may sound like a technocrat's vision of the promised land, people such as Barbour, Poduska and the others who catalyzed one of the most significant revolutions in the relationship of people to computer technology believe that it is possible, practical and not too far away.

Richard M. Fichera is an independent consultant specializing in graphics, workstation and technical computing systems applications.

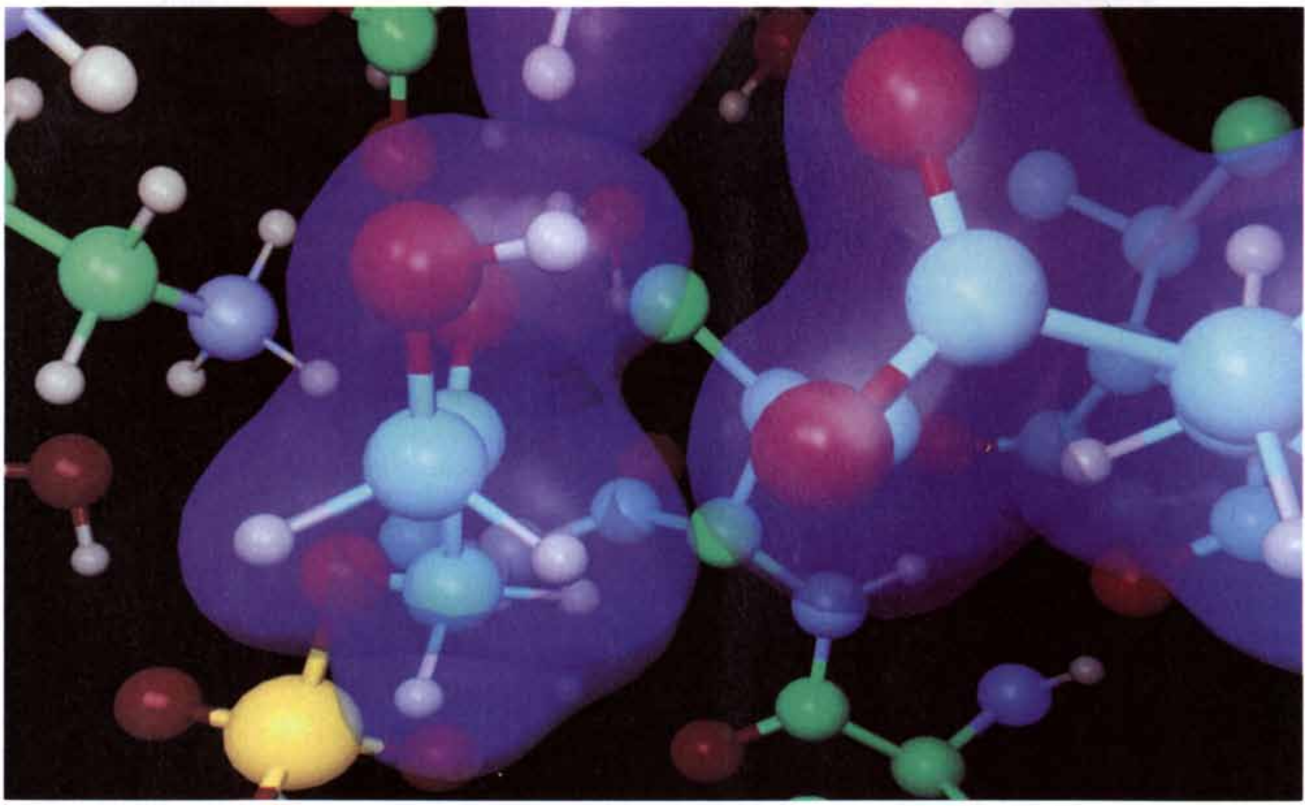
Scientific Visualization

by Ronald D. Levine

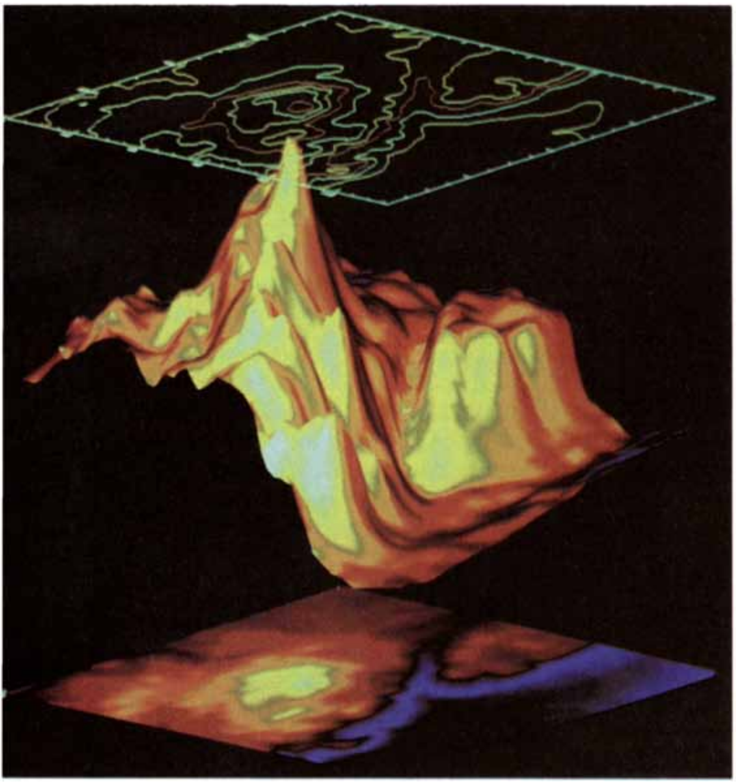
Visualization is the formation of mental images of what cannot be seen. It is seeing, in a way, that which is invisible, either because it is inaccessible to eye or camera, because its physical nature supports no real image formation or because its existence is abstract or hypothetical. Scientific visualization involves real images—drawings, photographs, holograms or video displays of graphs, charts, diagrams and other visual models—to serve as intermediate aids; it makes use of our real eyes to help our mind's eye in its task of grasping the structures, relationships and abstractions of science. The importance of real visual aids in scientific understanding may well be related to the fact that a very large proportion of our brains' neurons seem to be devoted to processing visual input.

Presently the workstation revolution, together with the associated proliferation of supercomputers and near-supercomputers into academic and industrial environments, is giving rise to a rapid and dramatic revolution in the art and practice of scientific visualization. The supercomputer centers create a demand for powerful, efficient visual presentation methods simply because of the immense volume of the data produced by supercomputer simulations. On today's machines the solution of a single, 3-D continuous-field simulation, as frequently carried out in such computational scientific domains as fluid dynamics, structural mechanics, meteorology, seismology, cosmology, molecular modeling and others, may well contain some hundreds of millions of numerical values. There is no hope of digesting such large data sets—of finding their essential features and exposing their hidden details—without the application of high-resolution, high-speed computer-based visualization tools.

High-resolution graphical rendering is computationally expensive in the sense that a large number of arithmetic operations are required to produce a picture from model data. However, these graphical computations are stereotyped and repetitive, so they may be performed by special dedicated processing circuitry and with a high degree of parallelism. The development of application-specific VLSI

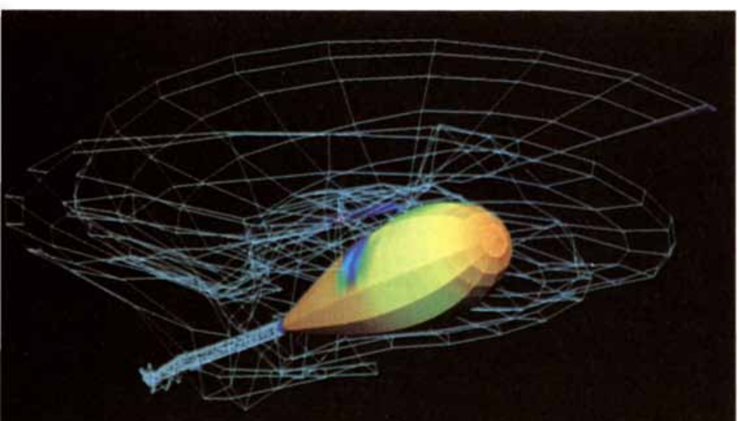


Active site of the enzyme triose phosphate isomerase, together with its natural substrate dihydroxyacetone phosphate. Atoms are represented as colored balls (red = oxygen, blue = nitrogen, green = carbon, yellow = phosphorus and white = hydrogen) with associated chemical bonds depicted by cylinders. The transparent blue surface represents a level surface of the electron density, calculated quantum mechanically. This image is a frame from a computer graphic movie, generated by Stephan Fangmeier at the National Center for Supercomputing Applications, from a simulation carried out by Paul Bash and Martin Karplus of Harvard University, using the original enzyme structure determined by Robert Davenport and Gregory Petsko of the Massachusetts Institute of Technology.



Three different but correlated bivariate distributions are represented using three kinds of visualization functions. In the middle, terrain elevation data around Pike's Peak is rendered as a shaded surface. Above, a correlated distribution of temperatures is depicted by a contour curve plot. Below, a correlated distribution of snow pack depth is represented by a color shaded contour plot. This image was produced using the software package PV Wave, a toolkit which gives scientists the capability of interactive construction and viewing of visual representations of functions and data sets.

Display of the results of a supercomputer simulation of the dynamical interaction of the wake of the rotor blades with the body of a helicopter. The vortex sheet created by the blade motion is depicted in wire frame style. Colors on the helicopter body denote pressure. The high-pressure spot *(blue)* on the starboard side is caused by the blade tip vortex, and is seen running rapidly aft in phase with the blade rotation in the movie of which this represents one frame. The fluid dynamical simulation was performed using the USAERO program from Analytical Methods, Inc. This image was produced using Ardent Computer's CFD Viewer.



circuits for advanced graphical processing (chips which can be produced relatively cheaply and ganged in parallel or pipelined fashion in order to multiply performance) has been a major factor enabling widespread affordable access to previously unheard of levels of graphics functionality in the new workstations and image computers. These hardware platforms have the potential for supporting highly innovative modes of scientific visualization.

Visualization innovations will not come automatically with the hardware, however, but will require further inventiveness. While the rendering of images from surface models may be regarded as stereotyped computation, there remains considerable freedom for choosing the methods of associating geometric objects (for example, curves, surfaces, volumes) and their visually relevant properties (color, opacity, illumination), with the abstractions of any particular computational science (pressure, velocity in fluid dynamics). And, of the possible methods of mapping physics to graphics, call them *visualization functions,* we have, in general, no way of knowing *a priori* which will be best in any particular case.

Moreover, we have to contend with the terrible 2-D bottleneck: until computer holographics becomes a widely available technology, all graphical display media are essentially 2-D surfaces, including the retina of the eye. There is no way to project a general 3-D model scene to a single 2-D view without losing information, in fact, most of it. For simple mechanical models a few views, as in drafting, may suffice to communicate all the relevant 3-D information, but the situation is much more difficult in the case of continuously varying spatial distributions, or in the case of very large biomolecular models, more serious still when these complex systems also vary in time. But it is these complex cases which give rise to the largest solution data sets and so are most in need of efficient visualization modes.

Even in the static case, whatever visualization functions and rendering algorithms are used, it is clear that many 2-D views are needed to get the 3-D picture. The researcher is obliged to view them in some sequence in time and then must somehow mentally integrate these viewing experiences to form his picture of the static 3-D situation. This process is aided considerably when the viewer has interactive control over the view selection. This means he is able to vary viewing parameters (for example, view point, view direction, etc.) smoothly and at will and see the corresponding views appear without noticeable delay. In the ideal case, using a joystick or other natural and comfortable control device, he would be able to roam about in his massive 3-D data set rotating objects to see their shapes from all sides, choosing section planes, panning to search for features of interest and zooming in to inspect them. Another dimension of assistance to the process is added when the researcher is also given interactive control of the selection of visualization functions.

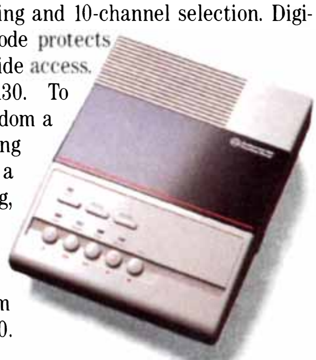
Interactive graphics is, of course, a principal raison d'être of the workstations and it is also one of the primary characteristics distinguishing this visualization revolution from the earlier modes. The illustrations on these pages, while attractive and informative, do not begin to convey the power of interactivity. And even though these illustrations all have markedly different visualization functions, they will probably look trite in comparison with the newer styles of scientific visualization which will appear during the next few years as the workstations and minisupercomputers become available to workers in more and more diverse scientific and engineering areas.

THE LIFESTYLE RESOURCE™

At The Lifestyle Resource we give you all the facts and details necessary to make an informed purchase. Your satisfaction is our primary concern. If your purchase doesn't meet your expectations, return it in original condition within 30-days for prompt refund.

WORTH CUTTING THE CORD FOR

A leading consumer magazine article likens a person's first conversation walking around talking on a cordless telephone to the exhilaration of that first bike ride minus training wheels. In the magazine's 1988 Buying Guide, the Southwestern Bell FF-1700 Cordless Phone is still rated tops for range in controlled tests among 21 brands and models. By handling incoming and outgoing calls to a range of 1000′ (the article rated a maximum of 1500 ft.), with outstanding speech quality and convenience features, the FF-1700 ended up on top of the consumer magazine ratings reports. Base unit serves as freestanding speakerphone with dialpad, so you get two phones in one. Plus intercom, paging and 10-channel selection. Digital security code protects line from outside access. **$179.95** #2130. To take your freedom a step further, Southwestern Bell's FA-450 Telephone Answering Machine gives you the latest technology and newest features at a most attractive price. Single cassette operation, call screening, household memo function, voice-activated record, one-touch playback. Two-way record for messages or conversations. Beeperless remote lets you retrieve messages from any pushbutton phone, also allows remote announcement changes. These Freedom Phones connect you in without tying you down. **$99.95** #2140.



FROM CHINA TO YOUR HEALTH

Ancient mandarins dating back 800 years believed that these Chinese Exercise Balls induced well-being of the body and serenity of spirit. These treasured gifts were given to President Reagan and his wife while visiting the Peoples Republic of China. The Chinese say that rotating the balls in the palm of each hand stimulates the fingers and acupuncture points, and improves circulation of vital energy throughout the body. Sports enthusiasts, musicians, computer users and health-conscious people everywhere consider them great muscle conditioners. Arthritis sufferers feel a decided benefit from this gentle but challenging exercise. Very effective for relaxation and meditation, Chinese Exercise Balls emit a distantly mysterious chime as you turn them. Beautifully handcrafted, 45mm. hollow polished chrome balls are perfectly weighted and fit comfortably into the average man's or woman's hand. In exquisite silk brocade box. **$29.95** #1701.

SHARP AS A RAZOR

The Chef'sChoice® Knife Sharpener safely and easily hones the dullest of knives to razor sharpness in minutes. According to *Food and Wine*, the Chef'sChoice Knife Sharpener is "unquestionably the best." *Popular Science* calls it "idiot proof." Its ingenious patented design is the brainchild of former DuPont engineer, Daniel Friel. His unique sharpening method uses diamond abrasives to sharpen a knife in three stages to form a progressively sharper edge. Since magnets hold the knife at the precise angle required for each operation, the guesswork of traditional methods is eliminated. The resulting triple bevel edge stays sharper longer than conventional edges. Only the minimum amount of steel is removed during each stage, making it safe for all high quality steel or alloy knives. World class chef, Pierre Franey, in his *N.Y. Times* column, hailed the Chef's-Choice as "unparalleled, in my experience, for putting an exceedingly sharp and long-lasting edge on knives of all kinds." Manufacturer's two year limited warranty, UL listed, measures 9″ x 3½″ x 4″. **$79.95** #2450.

A WATCH FOR ALL REASONS

This handsome digital-analog chronometer will swim with you (to a depth of 150 feet), fly with you, calculate your speed, wake you in the morning, and tell you the day and date. A push of a button will convert the digital display to 24-hour timekeeping. All functions—digital and analog timekeeping, calendar, alarm, chronometer and tachymeter—are powered by a highly accurate electronic quartz movement. The analog hands and hour markers are luminous. The stopwatch times to 1/100th of a second, and the rotating bezel can be used for such practical matters as telling you when the parking meter will need another quarter. Adjustable fit. **$39.95.** Teflon coated durable black matte anodized stainless steel #1051; Stainless steel and gold-plated #1061.

THE LIFELINE GYM

The Lifeline Gym is the most space efficient, time efficient home fitness device on the market today. You can simulate just about any of the exercises done on the expensive machines found in health spas and gymnasiums — the reason is latex. With this unique stretchable latex cable, resistance increases with movement — your muscles are challenged through their full range of motion. These variable resistance exercises provide an excellent workout for a beginner as well as a challenge for a pro. The Lifeline Gym is so successful it is used by members of the Chicago Bears, the U.S. Ice Hockey team, and the UCLA track team to name a few. Adjustable for individual strength levels and for different body shaping and toning. The gym includes a lifting bar, resistance cable, stirrups, exercise belt, door attachment and a comprehensive fitness guide that outlines over 25 exercises. Everything packs into a compact carry case, so you can enjoy the benefits of a complete workout anywhere. **$49.95** #1841 Ladies or #1851 Men's.

SCIENTIFIC AMERICAN May 1989 Volume 260 Number 5

Technology, Employment and U.S. Competitiveness

Technological change is essential to U.S. economic progress. To facilitate it, workers need to be more adaptable, research needs to focus more on applications and managers need to be more open-minded

by Richard M. Cyert and David C. Mowery

In his 1950 book, *The Human Use of Human Beings,* the American mathematician Norbert Wiener wrote: "The factory of the future... will be controlled by something like a modern high-speed computing machine.... Industry will be flooded with the new tools to the extent that they appear to yield immediate profits.... It is perfectly clear that this will produce an unemployment situation, in comparison with which the present recession and even the depression of the thirties will seem a pleasant joke." Recent history has shown that Wiener was on the right track as far as the computer-controlled production equipment is concerned, but he was much less prescient about the equipment's economic impact.

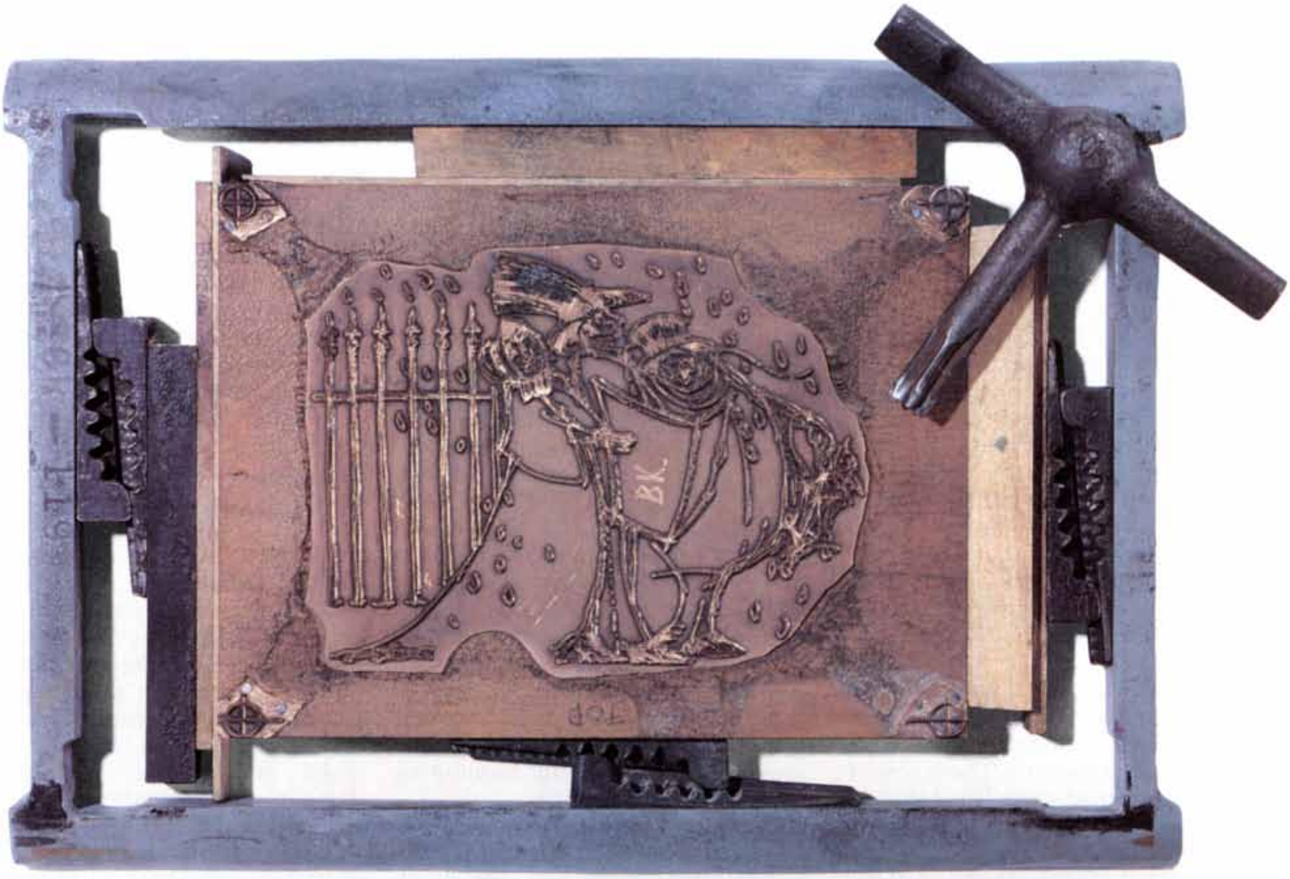
That is not to say that all has gone well with the U.S. economy since Wiener's time. Although the unemployment rate has declined from a peak of more than 10 percent in 1982, inequality in the distribution of household incomes has increased since the late 1960's, and real hourly earnings for U.S. workers have been stagnant since 1973. More puzzling perhaps is the extremely slow growth of labor productivity (measured as output per hour of work) between 1973 and 1979. Although labor productivity in the U.S. has recovered somewhat since, its rate of growth still lags behind those in a number of other industrialized nations. Furthermore, the American performance in international trade has been dismal.

Could Wiener's basic premise—that automation in the factory leads naturally to less employment for the worker and more profits for the factory owner—be partly right after all? Are robots and computers creating a two-tiered society in which employment security, job advancement and increased earnings are available only to a small minority of workers?

Recently a panel of experts drawn from business, labor, government and academia [*see illustration on page 56*] was assembled under the auspices of the National Academy of Sciences, the National Academy of Engineering and the Institute of Medicine to analyze technology's effects on employment and the U.S. economy between now and the year 2000. Based on its review of an extensive body of evidence, the panel concluded unanimously that the adoption of new technology is not a cause of the high unemployment and low growth rate of earnings that have afflicted the U.S. economy for much of the past 15 years. On the contrary it will have to be part of the cure for those very problems. Indeed, changes in the structure of the domestic and world economies have made the rapid and effective adoption of new technologies more important for the U.S. than ever before.

It is easy to see why someone might think that the introduction of new technology in the workplace leads to unemployment. After all, technological innovations are generally incorporated into production processes in order to decrease the amount of resources—particularly labor—required for a unit of output. Such thinking does not go deep enough, however. New production technology is likely to have other economic repercussions that offset any potential reductions in aggregate labor demand resulting

RICHARD M. CYERT and DAVID C. MOWERY were respectively chairman and study director of a panel organized by the National Academy of Sciences (NAS), the National Academy of Engineering (NAE) and the Institute of Medicine (IOM) to assess the effects of new technology on employment and the U.S. economy. Cyert, who holds a bachelor's degree from the University of Minnesota and a doctorate from Columbia University, is president of Carnegie-Mellon University. Before being appointed president he served for 10 years as dean of the university's Graduate School of Industrial Administration. Mowery is associate professor of business administration at the School of Business of the University of California, Berkeley. He studied at Stanford University, receiving undergraduate and graduate degrees in economics. The opinions expressed in the article are those of the authors alone and do not necessarily represent those of the entire panel, its sponsors or the NAS, NAE and IOM.



TECHNOLOGICAL CHANGE has transformed the art of printing, as it has many other occupations. To print an illustration 30 years ago, copper plates (*top*) had to be etched (one for each color ink) and mounted on frames. Today a printer can do much the same job on a computer (*bottom*) and in substantially less time. Clearly the skills necessary for the job have changed.

from its adoption. For example, by decreasing manufacturing costs and thereby lowering the price of a product, a new technology can expand consumer demand for the product and so lead to increased production and ultimately a greater—not a lesser—demand for labor.

The effect of a new production technology on employment therefore depends as much on the consumer's response to reduced prices (or improved quality) as it does on the changes in the amount of labor required for a unit of output. Even if consumer demand for the product does not increase significantly, a net rise in employment can still result: consumers are likely to apply the savings gained from price reductions on the product to the purchase of other goods, thereby causing employment in other consumer-goods industries to expand. Moreover, the widespread utilization of a new production process may require new machine tools, materials and supplies. Those new requirements may stimulate increases in the employment levels in the supplier industries as well.

The employment effects of a new technology are further complicated by the fact that managers and workers often need to restructure their relations to accommodate the technology. Such restructuring can directly affect worker satisfaction and safety, which in turn affect productivity. Indeed, organizational changes made possible by the introduction of a new technology may influence the quality of a company's products and its profitability as much as engineering advances in the technology itself.

The General Motors Corporation, for example, was able to reap the full benefits of mass-production technologies in the 1920's only after it decentralized its product divisions and established a system for pricing internal sales among the divisions. In contrast the Ford Motor Company (which was the first automobile manufacturer to apply mass-production techniques) resisted such organizational innovations and as a result steadily lost market share to its rival.

The future demand for workers also depends on what particular skills a new technology requires, which in turn depends on how the technology fits into a company's organizational structure. Production workers, for example, may have to learn much more about a plant's overall operations and management if they are to assume greater responsibility for monitoring the production process.

In addition, the skill requirements of new technologies change as the technologies themselves are developed, modified and incorporated into production processes. Compare, for example, the level of skill necessary to operate a large mainframe computer of the early 1960's with that needed to perform similar tasks on a personal computer in 1989. Continued innovation in the complementary software and hardware technologies not only has made today's personal computers more powerful than the large mainframes of two decades ago but also has made it possible for them to be used by elementary school pupils as well as university graduates.

Predicting what cumulative impact all these various effects will have on the unemployment rate and the training of workers is beyond the capabilities of current forecasting techniques. Just to predict how an innovation in the manufacture of a particular product will affect employment, for example, requires detailed information about a number of variables: the rate of diffusion of the technology and its impact on the amount of labor required to manufacture a unit of output (a datum that almost certainly changes as the technology itself diffuses throughout the economy); the response of demand for the product to a reduction in price (the so-called price elasticity of demand); the change in aggregate labor demand resulting from any increase in the demand for the product; and the response of individuals to changes in the cost of the bundle of goods they consume. Unfortunately, the data needed for estimating these numerous and interacting variables simply do not exist. Foretelling how the development of an entirely new good or service will affect employment is even more difficult, because the new good or service may actually create industries that did not exist before.

Predicting the future demand for certain worker skills is hampered not only by the lack of data but also by deficiencies in the methods applied to gauge skills. A clerk from the 1950's and a clerk from the 1980's are alike only in name. Clerical personnel today perform tasks that differ substantially in content and that require entirely different abilities. Yet this change is completely obscured in most occupational data, which assume that the skills necessary for clerical occupations are the same today as they were in the 1950's.

Because the skill requirements of a job can vary over time in occupations as much as they vary across different occupations at any given time, a projection of the future demand for job-related skills that is based on the likely growth of occupations is bound to be inaccurate. Moreover, occupational trends themselves are difficult to forecast, because they depend on the structural changes that take place in an economy. As mentioned above, a detailed prediction of those changes

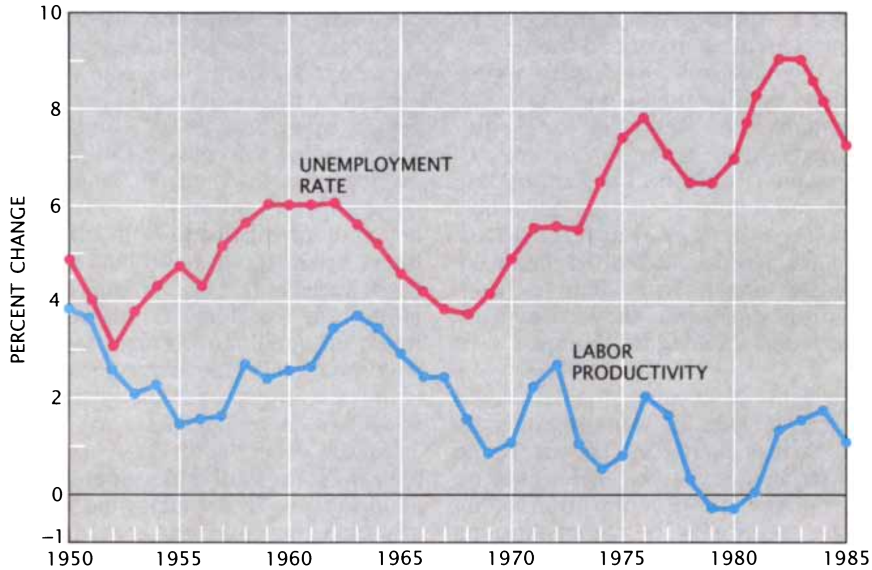
PANEL ON TECHNOLOGY AND EMPLOYMENT

Richard M. Cyert (Chairman)
Carnegie-Mellon University
Morton Bahr
Communications Workers of America
David Cass
University of Pennsylvania
Alonzo A. Crim
Superintendent, Atlanta Public Schools
Douglas A. Fraser
Wayne State University
Richard B. Freeman
Harvard University
Samuel H. Fuller
Digital Equipment Corporation
Judith M. Gueron
Manpower Demonstration Research Corporation
Anne O. Krueger
The World Bank
Lawrence Lewin
Lewin and Associates, Inc.
James N. Morgan
University of Michigan
Thomas J. Murrin
Westinghouse Electric Corporation
Eleanor Holmes Norton
Georgetown University Law Center
D. Raj Reddy
Carnegie-Mellon University
Nathan Rosenberg
Stanford University
William W. Scranton III
Lieutenant Governor, Commonwealth of Pennsylvania
G. Russell Sutherland
Deere & Company
Marta Tienda
University of Wisconsin, Madison
Louise Tilly
New School for Social Research
Amy D. Wohl
Wohl Associates
David C. Mowery (Director)
University of California, Berkeley

MEMBERS of the Panel on Technology and Employment were drawn from business, labor, government and academia. The panel, which assessed technology's future effects on employment and the U.S. economy, was organized by the National Academy of Sciences, the National Academy of Engineering and the Institute of Medicine.

is nearly impossible for an economy as large and complex as that of the U.S.

Given that the introduction of a new technology in the workplace is often accompanied by restructuring of the production organization, retraining of workers and numerous technical modifications of the technology itself, considerable time is required before the technological possibilities of a laboratory invention are fully realized in the U.S. economy. Projections of the economic effects of new technologies that do not allow for the lengthy period required to perfect and to adopt new technologies therefore may overstate the magnitude and suddenness of any changes in the unemployment rate or in the demand for certain job-related skills. The gradual pace of technological diffusion allows relatively extended periods during which people and organizations can adapt. For that reason the policy challenges created by the effects of new technologies on employment may actually be more tractable than those resulting from other, more volatile sources of economic change.



GROWTH RATES OF UNEMPLOYMENT AND LABOR PRODUCTIVITY (the total value of goods and services produced divided by the number of labor hours that went into their production) have been diverging in the U.S. since 1970. Since new technology in the workplace increases productivity, it might be thought that technological change leads to a lower aggregate demand for labor. Yet if that were truly the case, unemployment should increase in step with productivity, contrary to what the data show.

Because predicting the economic impact of technological change is fraught with so many uncertainties, an economist has no alternative but to rely heavily on historical evidence, and the effects of new technology on employment in retrospect have been unambiguous [*see illustration on this page*]. The record shows that when the U.S. unemployment rate was high (as it was in the 1970's and the 1980's) productivity growth rates were low. Because labor productivity depends on the introduction of new technologies, the data provide strong evidence that technological change is not associated with increases in the nation's unemployment rate.

Technological change also does not appear to be responsible for the lower rates of growth in average hourly earnings in the U.S. [*see illustration on page 59*]. Hourly compensation (including employee benefits as well as wages and salaries) has grown in step with labor productivity. If the introduction of new technology had resulted in lower wages, then hourly compensation should have fallen as productivity rose.

Much has also been made of the increasing inequality in the distribution of household incomes in the U.S. during the 1970's and 1980's. According to McKinley L. Blackburn of the University of South Carolina and David E. Bloom of Columbia University, for example, the share of the nation's total family income accounted for by the incomes of middle-class households (which range between 100 and 160 percent of the median household income) declined from 27.5 percent in 1967 to 21.3 percent in 1985. Yet available data indicate that this increased inequality stems mainly from factors other than technological change, such as the increasing number of households headed by single women and changes in federal budgetary and tax policies. Blackburn and Bloom also find little evidence of a significant change in the distribution of worker earnings, which would reflect more directly any increasing inequality in the work force. If the adoption of new technology were indeed widening the gap between high-paying and low-paying jobs, their analysis would have revealed it.

In the light of such evidence it is reasonable to conclude that technology's effect on employment is probably less important than the influence of other economic factors. The rate of economic growth, aggregate demand and the impact of extraneous shocks to the U.S. economy (such as the 1973 and 1979 oil-price hikes) appear to be more significant determinants of the nation's overall employment level than technological change.

Does the past also offer some insight into how workers should be educated and trained to meet future demands for labor? The historical evidence clearly shows that technological change has increased the value of such basic skills as reading, writing, numerical reasoning and problem solving. The computer-based technologies that are likely to be associated with jobs in the future will place even greater demands on the mental rather than the physical capabilities of factory and office workers. Furthermore, expanded use of work teams in many production systems also means that a premium will be placed on the ability to communicate effectively.

Most important for future U.S. workers, however, will be the ability to adapt to structural changes in the economy that reshape the nature of their jobs. They will have to learn how to cope better with more frequent moves—from employer to employer, from town to town and even from industry to industry. Surveys made during the early 1980's, for example, suggested that as many as 30 percent of the workers who lost their jobs in the automobile and steel industries were hampered in their search for new jobs by deficiencies in basic skills.

The fact that employment dislocations in specific industrial sectors will unavoidably take place as new machines and new organizational relations are introduced into the workplace poses a dilemma for the U.S., because the economy as a whole stands to benefit from such technological and organizational change. Yet the dilemma can be solved to a large extent by directing a portion of the economic benefits of technological change—in the form of programs for monetary compensation, counseling and retraining—to those individuals who are subject to its adverse effects. Indeed, because such programs make it easier for workers to adapt to changes in the workplace, they can actually promote technological change.

How exactly does the U.S. economy gain by promoting technological change? The answer to that question becomes clear as soon as one views the economy from a global perspective. In the discussion thus far the U.S. economy has been regarded as an independent entity, but of course it is only one of many national economies that interact with one another by means of trade. Indeed, the vagaries of international trade have come to exert a strong influence on the U.S. economy; foreign trade now accounts for nearly twice as large a share of the U.S. gross national product as it did during the 1960's.

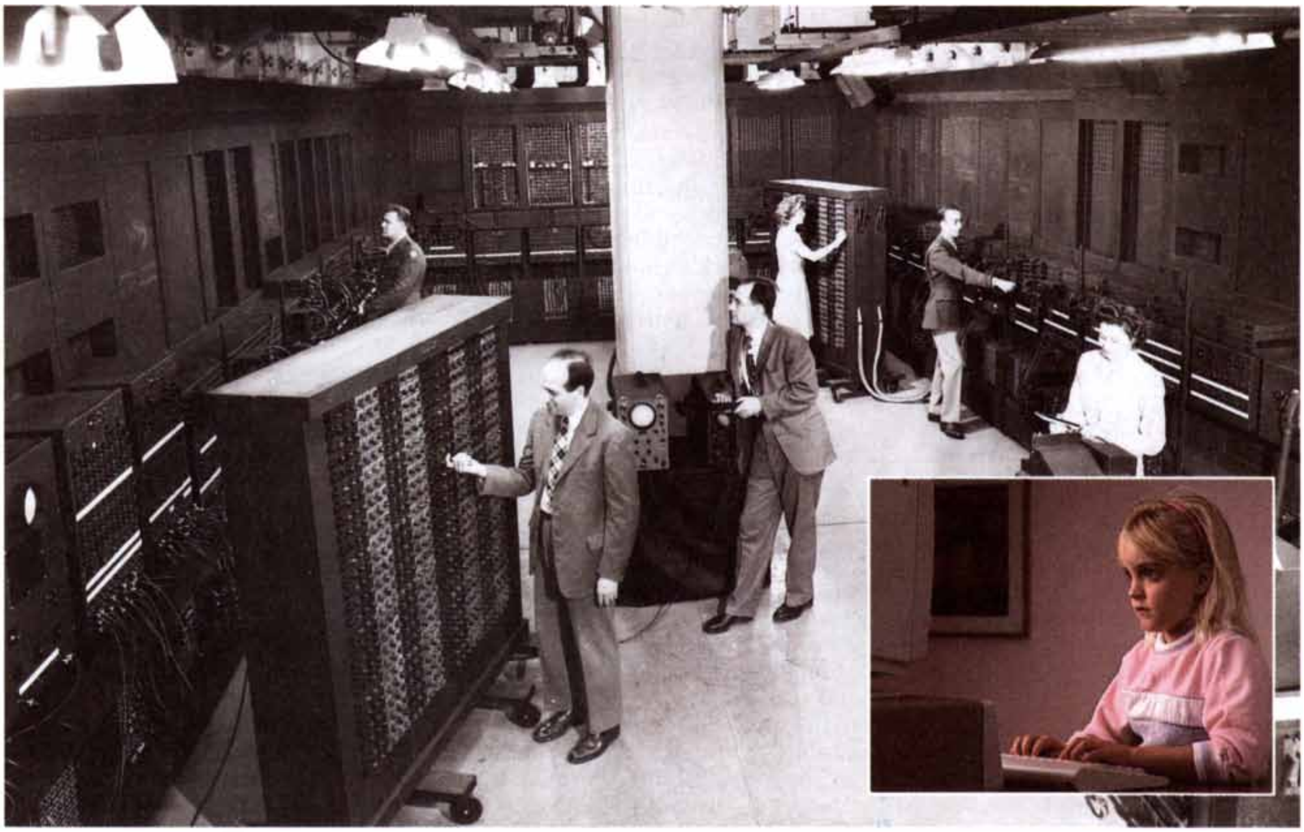
In such an environment, if foreign companies develop and adopt new technologies faster than U.S. companies do, the production costs of foreign companies will tend to fall more rapidly than those of U.S. companies. The reductions in costs enable foreign producers to lower prices, squeezing U.S. companies out of the global market and forcing workers in affected industries to accept fewer jobs or lower earnings. In fact, the extensive penetration of U.S. domestic markets by imported goods during the past decade in part reflects the greater speed with which foreign producers tend to adopt new technologies.

On the other hand if U.S. companies are successful in developing and adopting new technologies more rapidly than their overseas competitors, the economic picture is likely to be quite different. Higher productivity growth in U.S. industries supports reductions in production costs, which are reflected in lower prices. And lower prices in turn stimulate the expansion of new, higher-wage jobs within the American economy.

Yet such technological leadership is difficult to maintain, because the scientific, engineering and practical knowledge that underpins commercial applications of new technologies can easily spread from one nation to another. Hence, a competitive advantage based on a specific technology or skill erodes over time as the underlying technical know-how diffuses throughout the world. Moreover, the speed with which technologies move across international boundaries has increased in recent years as a result of improved worldwide communications and transportation, growing direct foreign investment in the U.S. and the greater scientific and engineering capabilities of many foreign research laboratories.

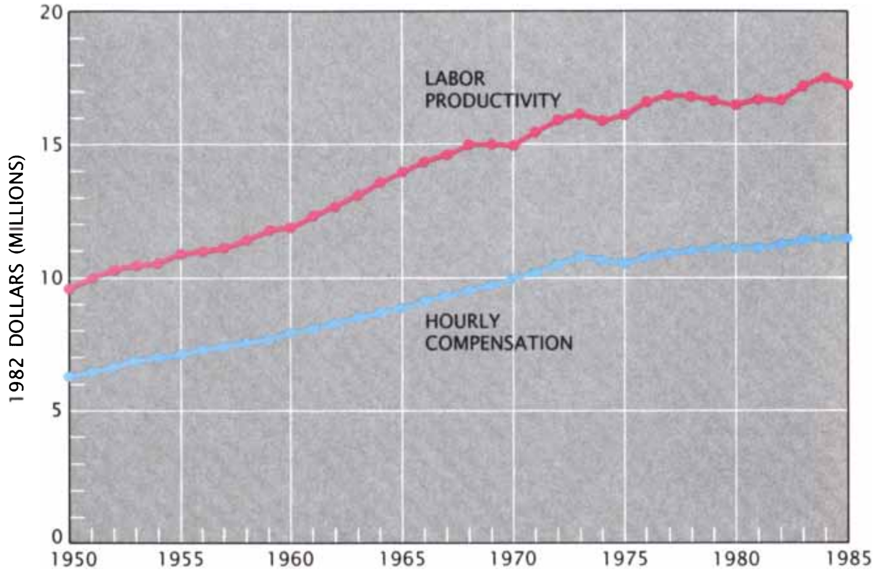
In the future any technology-based advantage held by a U.S. company over foreign rivals is therefore likely to be even more fleeting. Indeed, in the global economy of the 1990's, U.S. employment losses resulting from slow or inefficient adoption of new technologies are likely to outweigh by a considerable margin any employment losses resulting from the rapid adoption of those same technologies.

That conclusion does not bode well for the U.S., since the country already lags behind other industrial nations in the adoption of such critical manufacturing technologies as robots and



TRAINING necessary to operate a computer has been drastically reduced as a result of advances in both hardware and software. Only highly trained operators could run the first computers, which were developed some 40 years ago. Today's personal computers (which surpass those first machines in both memory and computing power) can be operated by a child.

computer-controlled machine tools. Ramchandran Jaikumar of the Harvard Business School estimates that between 1982 and 1986 Japanese companies outspent U.S. companies by more than two to one on factory automation. The reasons for the U.S. industry lag in adopting advanced manufacturing technologies are not precisely known. They probably include differences between the U.S. and other industrial countries in the cost of capital and rates of investment, in the skills of the production work force and in the approach managers take to evaluating and investing in new technologies [see "U.S. Economic Growth," by Ralph Landau; SCIENTIFIC AMERICAN, June, 1988].



LABOR PRODUCTIVITY AND WORKER COMPENSATION (the total value of all wages, salaries and employee benefits divided by the total number of work hours) have grown in step with each other in the U.S., suggesting that technological change is not responsible for the lower growth rates in worker compensation seen in recent years.

Much of the formidable task that confronts American managers and policymakers therefore amounts to doing everything possible to ensure that future technologies are rapidly and effectively adopted. Achieving this goal almost certainly will require larger investments by public and private institutions in the education of the U.S. work force. Yet the uncertainty as to what job-related skills new technologies will demand makes it a risky strategy to invest resources in educational programs designed to prepare young people for a specific occupational vision of the future.

A more prudent strategy is to improve the adaptive capabilities of members of the work force by investing in better basic-skills education. Even if technological progress were to stop tomorrow, the employment prospects for labor-force entrants who lack strong fundamental skills would be dismal. Those same skills will be even more important in finding and retaining quality jobs in the future. This state of affairs could portend a decidedly bleak future for minority entrants and for the economy itself. A large fraction of those entering the work force between now and the year 2000 will consist of blacks and Hispanics, many of whom have not been well served by public institutions of primary and secondary education.

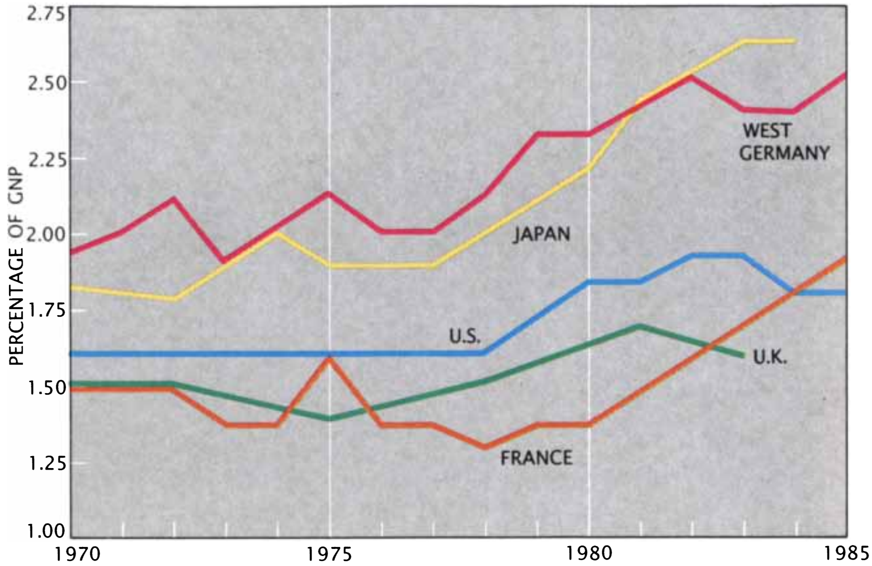
Policies designed to improve the quality and adaptability of the U.S. work force will have to devote attention not only to new entrants but also to those currently employed. Indeed, experienced workers should probably receive greater attention from public and private assistance programs, since most of the workers who will be in the labor force in the year 2000 are currently employed. Although retraining may help some of them, the potential benefits should not be overstated. For some workers, particularly those who have lost relatively high-paying jobs in durable-goods manufacturing, retraining is unlikely to restore their previous level of earnings.

At the very least, basic skills should be taught to those displaced workers who lack them, because such remedial education may improve their reemployment prospects considerably. After all, many if not most displaced workers are interested mainly in finding another job. For that reason retraining programs that do not incorporate job-placement assistance are unlikely to attract substantial participation. Programs for displaced workers must also offer other services: even an experienced worker who has adequate basic skills may need counseling in the event he or she is laid off.

Although considerable sums have been spent on federal programs for workers who have little if any work experience, there exist few programs that serve the needs of experienced workers who have lost their jobs. The Trade Adjustment Assistance (TAA) program, created in 1962, provides income support and retraining for workers able to show they were laid off by companies that were undercut by foreign companies exporting to the U.S., but the only federal program serving all displaced workers—regardless of the cause of their displacement—is Title III of the 1982 Job Training Partnership Act (JTPA).

The main purpose of the Title III program, which had an average annual budget of roughly $200 million during 1982–1988, is to provide a displaced worker with a new job as quickly as possible. But Title III services have focused mainly on workers who are relatively easy to place in new jobs. As a result the program does not serve most of the displaced worker population. According to a 1987 estimate by the U.S. General Accounting Office, no more than 7 percent of the annual flow of displaced workers received JTPA assistance. In addition, JTPA services do not provide training in basic or job-related skills, and they rarely provide income support to workers.

In 1988 annual funding for Title III was increased to as much as $1 billion, and some reforms of the JTPA were mandated. Nevertheless, further reforms are necessary to expand coverage to include all displaced workers, provide better training in basic and job-related skills and increase income support for workers who receive the training. In addition the web of federal and state employment services, unemployment insurance, retraining and worker-assistance programs needs to be rationalized and made internally consistent. It accomplishes little, for instance, to prohibit people enrolled in training courses from receiving unemployment compensation, which is the most important source of tempo-



FUNDING FOR CIVILIAN RESEARCH AND DEVELOPMENT in the U.S. has in recent years represented a smaller share of national economic output than it does in West Germany or Japan. The U.S., however, invests more money in military R&D than the other two countries. Although the technological "spillover" from military R&D did assist the early development of the civilian semiconductor, jet engine and computer industries, the authors argue that it now has a less important role in stimulating industry.

rary income support for many displaced workers. Yet a number of states impose such restrictions.

The eligibility for worker assistance should also not be made contingent on proof that technological change caused the displacement. The intended assistance can be sidetracked by the administrative difficulties and delays in service that such a requirement entails. Eligibility restrictions have hampered TAA's usefulness to workers, for example. Even when TAA was operating most efficiently, the average length of time needed to certify eligibility was nine months. By the time some displaced automobile workers began receiving their TAA income-support payments in the 1970's, they had already been called back to work.

Worker-assistance services are most effective when provided to workers before they are displaced. In order to do that, however, workers must receive advance notice of permanent layoffs. Recent federal action requiring 60 days' advance notice of permanent, large-scale layoffs or plant closures can serve not only to reduce public expenditures on unemployment compensation but also to smooth labor-market adjustments.

Changes in Government worker-assistance programs alone are not enough to speed up the adoption of new technology in the U.S. The country must also fundamentally alter its approach toward science and technology. U.S. policies in this area have tended to concentrate more on the generation of new knowledge than on the adoption of product or process innovations based on that knowledge. Because technological and scientific knowledge flows across national boundaries more rapidly today than ever before, the economic returns on investments in basic scientific research cannot always be gathered by the nation making the investment.

In the past the U.S. was able to rely on investments in applied research of military technologies to yield important civilian applications. Yet the contribution of that kind of research to commercial applications in such technologies as microelectronics appears to have declined. The direction of influence may even have been reversed: military technologies now appear to depend on advances in civilian applications. Currently the U.S. devotes a smaller share of national output to nondefense research than does West Germany or Japan. Yet an increase in the funding of civilian basic research may now contribute at least as much to the military technology base as an equal increase in the defense research budget.

In funding nonmilitary R&D, the Government must also avoid its past tendency to support radical and expensive advances in technologies. Several energy research programs in the 1970's typified this flawed approach. If publicly funded R&D programs are intended to support faster commercialization and adoption of new technologies, they should take a more incremental and diversified approach. They should also draw heavily on private-sector input.

One model for such programs is the aeronautics research program of the National Aeronautics and Space Administration. The program contributed significantly to the development of an innovative and internationally competitive commercial aircraft industry. Further examples of programs that have adopted a diversified focus and serve a broad industrial clientele include the National Science Foundation's Engineering Research Centers and its other initiatives in university-industry research cooperation. Recent Congressional action to expand the responsibilities of the National Bureau of Standards (which included changing its name to the National Institute of Standards and Technology) is also aimed primarily at facilitating the adoption of new technology throughout the country.

As the goals of diffusion and commercialization become more prominent in federal science and technology policies, however, the U.S. must guard against assuming an excessively nationalistic or even protectionist policy line. The fact is that scientific and engineering research done in the U.S. today depends greatly on the research done in other countries, and hampering international scientific communication and collaboration may ultimately work to this country's disadvantage. If foreign participation in a publicly supported research program is restricted as a result of a desire to retain all of the program's potential economic benefits for the U.S., American companies and researchers may find their access to foreign research restricted in retaliation. For that reason the U.S. may end up paying a heavy price for prohibiting the participation of foreign companies in the Semiconductor Manufacturing Technology (SEMATECH) consortium and the National Center for Manufacturing Sciences, Inc., in Ann Arbor, Mich., both of which are research programs supported in part by public funds.

Rather than restricting the access of foreign scientists and engineers to the results of publicly funded basic research, the U.S. should strive to ease the access of American scientists and engineers to the results of basic re-

search done in laboratories abroad. In addition U.S. companies should more closely monitor foreign scientific and engineering advances and hasten the speed with which they apply commercial innovations based on them. Protectionism in scientific and technological research does no more to promote U.S. economic welfare than protectionism in trade policy does.

Because rapid and widespread adoption of new technologies is critical to enhanced earnings and employment for the work force, both labor and management stand to gain from encouraging technological change. Yet that fact has not led to better labor-management cooperation in some U.S. industries.

New management policies could improve the situation, however. Employment security is one of the greatest concerns of workers in factories and offices that are adopting new technologies. Policies to address this concern include reassignment of employees to other jobs within a company, early retirement programs, retraining and job-placement assistance for those workers who are fired. Without such policies labor-management cooperation can break down easily.

On the other hand unions must be more willing to consider revisions in job classifications and pay schemes in order to give the company the necessary organizational flexibility to implement new technologies. Recent examples of union concessions include the basing of wage increases in the automobile and aerospace industries on profit-sharing rather than cost-of-living provisions. Some U.S. companies and unions have also explored a quid-pro-quo arrangement: in exchange for revising job classifications and compensation systems, management assures the union that it will seek to minimize layoffs and maximize employment security. Managers and workers in a wider range of industries could benefit from similar reciprocal arrangements.

Changes in a company's organization that are necessary for it to exploit advanced manufacturing processes also have important implications for its employees. Such reorganization often shifts more responsibility for product quality and process control to workers, reducing the importance (and possibly the number) of middle managers. As a result middle managers may resist the organizational changes needed to adopt these technologies successfully.

This shifting of responsibility also means that worker training should not take place only when new equipment is installed in an office or factory. It should be an ongoing activity. If production workers are to assume many of the responsibilities of middle managers or technical employees, their skills must be periodically renewed through training programs similar to the programs developed by some U.S. companies for their middle-management and technical staff.

Another private-sector obstacle to the adoption of new technology in the workplace is the fault of management alone. Many U.S. corporations evaluate investments in new technology by applying analytical techniques that overemphasize direct labor costs (production workers' wages and salaries) and ignore potential savings in overhead. Yet a reduction in overhead costs is one of the most important benefits derived from the adoption of computer-based technologies and the organizational changes that should accompany their adoption. Indeed, in many industries, Japanese companies have a greater comparative advantage over their U.S. competitors in overhead costs than in direct labor costs.

In addition, standard accounting techniques often fail to incorporate the cost and profitability consequences of improvements in product quality, reductions in inventory and other intangible gains from the adoption of advanced manufacturing technologies. Because managers do not quantify such cost savings, they may also be unable to quantify properly the benefits of the reorganization that often must accompany the introduction of new production processes.

In sum, the education of managers in the U.S. appears to focus on outmoded or biased techniques for financial analysis; it overlooks the potential competitive advantages inherent in the creative application of modern technology. That situation should be corrected. The development of new analytical techniques that incorporate the technological, organizational and financial dimensions of innovation is long overdue.

Technological change and its concomitant structural change pervade the U.S. economy, as they do any dynamic economy. Far from reducing economic welfare, such changes are essential to the further growth of earnings and employment opportunities. The expanded role of international trade in the U.S. economy and the quickened pace of technology transfer have made the rapid development and adoption of new technologies a critical determinant of the country's competitiveness in the global market.

Dealing with the budget and trade deficits of the 1980's will require that both the Government and the citizenry of the U.S. live within their means, increase savings and investment, and reduce spending on consumption. These adjustments will be far less difficult if they are supported by growth in productivity that exceeds the rate of the past decade. Here again technological change plays a crucial role, but it requires improved public and private policies to tackle the changing employment patterns it inevitably causes. Better policies must be developed to assist displaced workers, to enhance the effectiveness of the national educational system and to encourage cooperation between labor and management. Such policies, which address the adverse economic effects of new technology, can serve to facilitate its introduction and diffusion throughout the economy.

The U.S. has little choice but to manage the development and adoption of new technologies more efficiently and humanely. Remaining at the leading edge of technology is essential to expanding economic opportunity for all members of society. Indeed, technological change appears to offer an unusual opportunity to enhance economic welfare while also improving the nation's economic efficiency.

FURTHER READING

TECHNOLOGICAL CHANGE: MEASUREMENT, DETERMINANTS, AND DIFFUSION. E. Mansfield in *The Employment Impact of Technological Change.* U.S. National Commission on Technology, Automation, and Economic Progress, U.S. Government Printing Office, 1966.

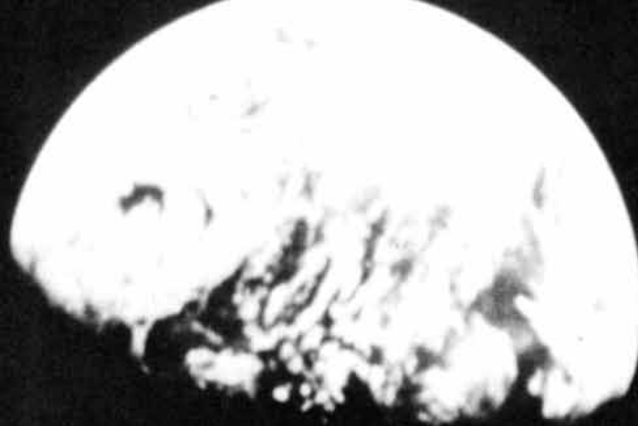
TECHNOLOGY AND EMPLOYMENT: INNOVATION AND GROWTH IN THE U.S. ECONOMY. Edited by Richard M. Cyert and David C. Mowery. Panel on Technology and Employment, Committee on Science, Engineering and Public Policy, National Academy Press, 1987.

EMPLOYMENT FUTURES: REORGANIZATION, DISLOCATION, AND PUBLIC POLICY. Paul Osterman. Oxford University Press, 1988.

THE FUTURE OF TECHNOLOGY AND WORK: RESEARCH AND POLICY ISSUES. Edited by Bruce E. Henderson and David C. Mowery. Committee on Science, Engineering and Public Policy, National Academy Press, 1988.

THE IMPACT OF TECHNOLOGICAL CHANGE ON EMPLOYMENT AND ECONOMIC GROWTH. Edited by Richard M. Cyert and David C. Mowery. Ballinger Publishing Company, 1988.

EARTH



SPECIAL ADVANCE ANNOUNCEMENT

Due to the expected high interest in this September 1989 single-topic issue and a limited print run, the publisher is offering a special opportunity to our readers to order extra or bulk copies in advance.

"Managing Planet Earth" will be a path-breaking issue that belongs on everyone's reading list and should be included in next year's curriculum. Share this vital information with your colleagues, students and anyone else for whom these pressing issues are a concern. Clip the Pass-Along Order Form below and bring it to their attention!

ORDER FORM

Yes, enter my order for _____copies of the September 1989 single-topic issue, "Managing Planet Earth", at $_____ per copy. Total order $_____.

Bulk and prepaid orders receive substantial discounts! Quantities limited. Order by June 30, 1989.

*1 to 9 issues $2.95 per issue
*10 to 24 issues $2.75 per issue
*25 to 99 issues $2.50 per issue
*100 or more issues $2.20 per issue

NOTE Deduct 10% from total with paid order.

PRICES INCLUDE SHIPPING (U.S. ONLY)*

☐ $_____ Payment enclosed ☐ Please bill me
☐ Charge my ☐ Purchase order attached
☐ VISA ☐ MasterCard ☐ Access ☐ Eurocard

Card # _____ Exp. Date _____

Signature _____

Name _____ School _____

Address _____

City _____ State _____ Zip _____

For more information, call Ms. Davis at 1-212-754-0539.
Cut and mail this coupon to: 415 Madison Avenue, Dept. MPE, NY NY 10017
*Outside the U.S. add $1.00 per copy.

P9AA94

PASS-ALONG ORDER FORM

Yes, enter my order for _____copies of the September 1989 single-topic issue, "Managing Planet Earth", at $_____ per copy. Total order $_____.

Bulk and prepaid orders receive substantial discounts! Quantities limited. Order by June 30, 1989.

*1 to 9 issues $2.95 per issue
*10 to 24 issues $2.75 per issue
*25 to 99 issues $2.50 per issue
*100 or more issues $2.20 per issue

NOTE Deduct 10% from total with paid order.

PRICES INCLUDE SHIPPING (U.S. ONLY)*

☐ $_____ Payment enclosed ☐ Please bill me
☐ Charge my ☐ Purchase order attached
☐ VISA ☐ MasterCard ☐ Access ☐ Eurocard

Card # _____ Exp. Date _____

Signature _____

Name _____ School _____

Address _____

City _____ State _____ Zip _____

For more information, call Ms. Davis at 1-212-754-0539.
Cut and mail this coupon to: 415 Madison Avenue, Dept. MPE, NY NY 10017
*Outside the U.S. add $1.00 per copy.

P9AA102

Creating Superheavy Elements

Nuclei far more massive than those found in nature are stabilized by subtle quantum-mechanical effects. Experimenters have had to revise ideas of how best to synthesize them

by Peter Armbruster and Gottfried Münzenberg

During the past 20 years physicists around the world have been engaged in the task of producing superheavy elements. At the Institute for Heavy-Ion Research (GSI) in Darmstadt we have met with some success, synthesizing the nuclei of elements 107, 108 and 109. These nuclei lie beyond the 106-proton threshold that marked the limits of previous techniques for creating and identifying heavy elements.

Experimental mass measurements, followed up by theory, show that the new elements owe their stability to the microscopic arrangement of their protons and neutrons rather than the macroscopic properties that stabilize lighter nuclei. On the other hand we have met with problems that have thus far prevented us from reaching the goals set in the late 1960's, when elements up to 114 seemed within reach. Working to overcome obstacles to further progress, however, has deepened our understanding of nuclear structure and of the dynamics of fusion reactions between nuclei.

Nucleosynthesis has come a long way from the earliest years when elements not found in nature were created in nuclear reactors. Physicists have employed ever-heavier projectiles to bombard target atoms; the latest development is "cold fusion," in which masses and bombardment energies are carefully chosen to minimize the excitation of newly formed nuclei.

In the course of our work early ideas about how to synthesize superheavy elements have almost all turned out to be wrong: the nuclei of the elements that can be synthesized are deformed and not, as postulated in 1966, spherical. In the fusion process we use stable, naturally plentiful spherical nuclei and medium-weight projectiles, not, as expected earlier, the heaviest radioactive man-made nuclei and correspondingly light projectiles. Fusion must take place at the lowest possible bombardment energy—as gently as possible, and not, as believed earlier, with excess impact energy to assist the process by brute force.

The idea of synthesizing transuranic elements (those with atomic numbers greater than 92) first appeared in the 1930's. In 1934 Enrico Fermi bombarded thallium with slow neutrons to produce lead by beta decay (the decay of a neutron into a proton and an electron). Neutron capture followed by beta decay produces elements one atomic number higher than the original element.

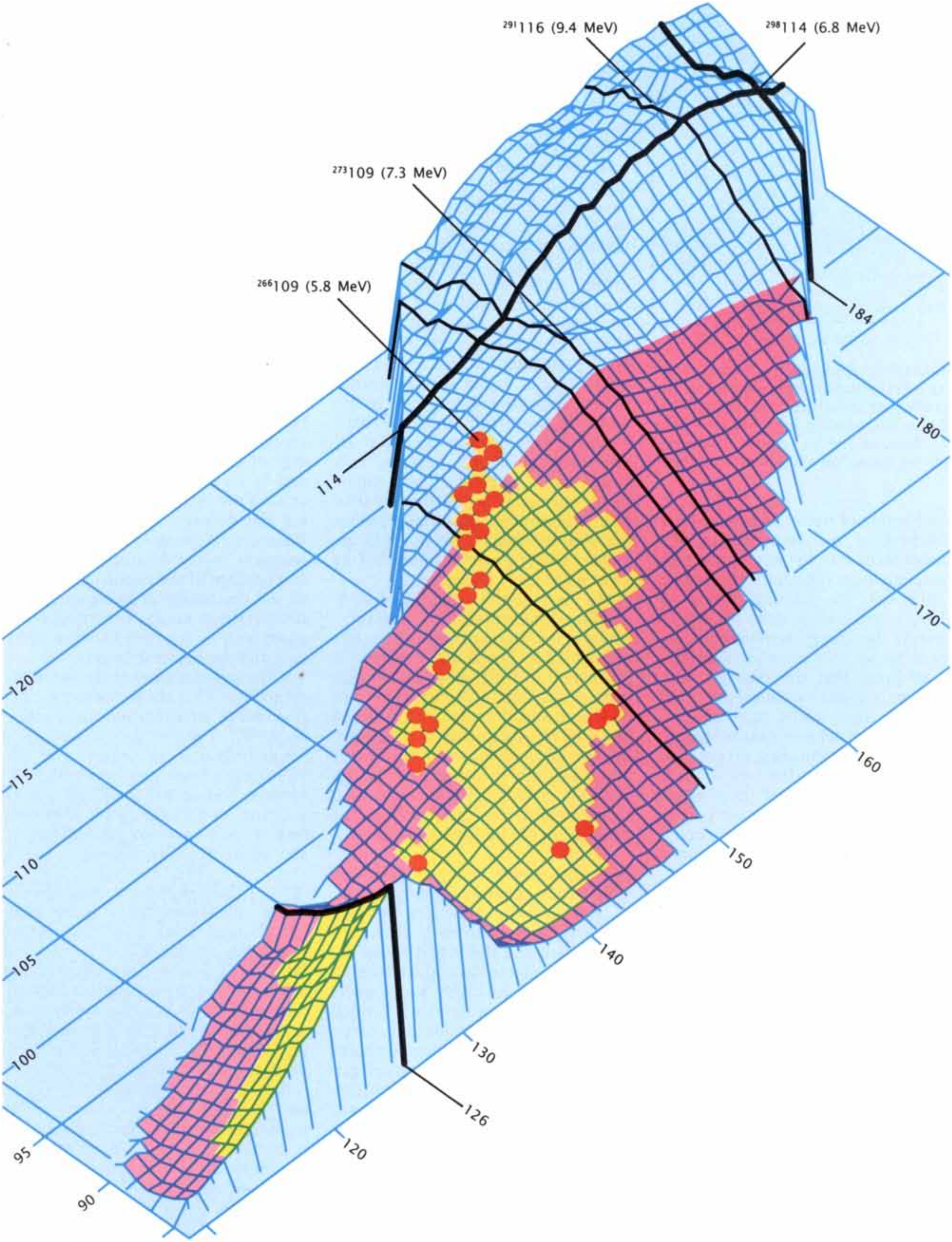
Between 1940 and the mid-1950's the elements 93, 94, 99 and 100 were produced by neutron irradiation. Fittingly fermium 100 was the last of the chain of elements that could be produced by Fermi's process of neutron capture and beta decay: none of its isotopes displays beta decay. During the same period, elements 95 to 98 and 101 were produced by alpha irradiation, a process in which heavy nuclei absorb two protons and two neutrons, thus ratcheting the atomic number up two steps at a time. Like all heavy elements the transuranics contain more neutrons than protons. Plutonium (element 94), for example, typically contains 145 neutrons for a total mass of 239; the longest-lived isotope of fermium has 157 neutrons for a total mass of 257.

To create elements beyond 100 the obvious approach was to fuse the nuclei of the heaviest elements with the nuclei of light elements containing more protons and neutrons than helium. Elements up to 99 were suitable because they could be produced in weighable macroscopic quantities. Accelerators were developed at Berkeley in the U.S. and Dubna in the U.S.S.R. that could give heavier ions enough energy to overcome the electrostatic forces that hinder fusion. Between 1958 and 1974 these heavy-ion accelerators made possible the synthesis of elements 102 to 106. Claims to the first discovery of these elements, and thus the privilege of naming them, remain disputed.

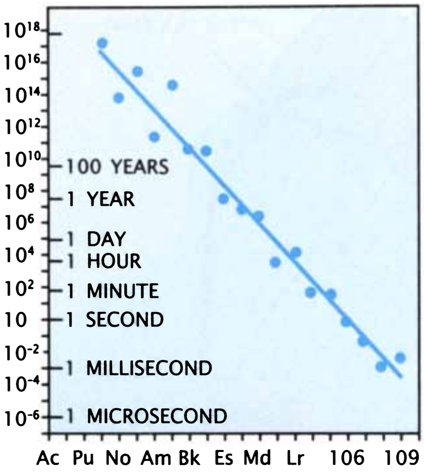
The procedures that served so well at Berkeley and Dubna have failed to produce elements beyond 106. To understand why superheavy elements should be so difficult to synthesize—and why some of them may be particularly stable—it is necessary to understand how nuclei in general hold together or break apart and how the balance of the different forces that govern their stability changes with increasing mass. Effects that are negligible in smaller nuclei turn out to make the difference between total instability and relatively long life in larger ones.

The essential feature of all nuclei is the interplay between the strong nuclear force, which binds both protons and neutrons together, and the electrostatic force, which pushes protons apart. As nuclei get heavier they become relatively more neutron-rich, which offsets some of the repulsive forces between protons. Nonetheless,

PETER ARMBRUSTER and GOTTFRIED MÜNZENBERG work on the synthesis of superheavy elements at the Institute for Heavy-Ion Research (GSI) in Darmstadt. Armbruster has been chief scientist at GSI since 1971. He leads a research group in nuclear chemistry and atomic physics; his personal interests are reactions between heavy nuclei and the use of heavy ions to study questions in atomic and solid-state physics. He received his doctorate from the Technical University of Munich. Münzenberg has worked in Armbruster's group at GSI since 1976. His research interests include the ground-state properties of heavy nuclei, cold fusion of heavy ions and the production and separation of exotic nuclei. He received his doctorate from the University of Giessen; while there he designed the SHIP velocity filter for GSI and oversaw its construction. Now that a heavy-ion synchrotron is being added to GSI's facilities, he is leading the construction of a separator for relativistic beams of exotic nuclei.



CLOSED NUCLEAR SHELLS (*black lines*) distinguish nuclei with numbers of protons and neutrons that contribute to extra stability. The shell corrections plotted here increase nuclear binding energy and make possible the existence of isotopes that would not be stable if they consisted simply of drops of "nuclear fluid." The fission barrier for liquid drops (*pink*) vanishes at about element 107. The shell effects peak in the vicinity of lead (*lower left*) and then climb again into the region of superheavy elements. Circles mark the proton and neutron numbers of isotopes synthesized by the authors' group.



HALF-LIVES of radioactive elements decrease rapidly with increasing mass. The heaviest elements produced so far decay almost immediately after they have been synthesized and so require new methods for detecting and identifying them.

the binding of nucleons peaks at iron (26 protons and 30 neutrons), less than a quarter of the way up the table of elements, and declines thereafter.

Although the splitting of any nucleus heavier than iron can release energy, the energy required to fragment those less massive than lead is so great that the reaction takes place only under extraordinary conditions. Because nuclei more massive than lead can achieve a more strongly bound state by emitting just a small part of the nucleus, they are unstable. The natural isotopes of thorium and uranium decay for the most part by emitting an alpha particle. Only in uranium and heavier elements do unexcited nuclei fission spontaneously.

In general atomic nuclei are increasingly unstable as their atomic number (the number of protons in the nucleus) goes up: half-lives decrease from thousands of years to millionths of a second. Theories of nuclear structure lead us to believe, however, that elements only slightly heavier than those so far created will be more rather than less stable.

Nuclei with certain combinations of neutrons and protons have particularly high binding energies; helium 4, oxygen 16, calcium 40, calcium 48 and lead 208 are very stable in comparison with their neighbors. These high binding energies are due to a shell structure that is the nuclear equivalent of the shells into which electrons organize themselves around the nucleus. Nucleon configurations that form fully occupied shells are particularly stable. The shell structure increases the binding energy of a nucleus by as much as 11 million electron volts (MeV), in the case of lead, over that of a hypothetical structureless droplet that has the same number of neutrons and protons. For most nuclei, which have binding energies of up to two billion electron volts, this increase is relatively insignificant. For heavier elements with marginal stability, however, the "shell stabilization" can mean the difference between instant disintegration and a relatively long life.

Nuclei with "closed" neutron and proton shells are particularly stable; the next such closed shell after lead occurs at 114 protons and 184 neutrons. The success of the shell theory in predicting binding energies for lighter nuclei fueled the speculation that nuclei with masses close to 298 might be so strongly stabilized that—as in the case of uranium and thorium—they could form a zone of relatively stable elements. These superheavy shell-stabilized elements, however, in contrast to the elements of the uranium-thorium region, would be unstable as homogeneous droplets of nuclear material.

The first of the superheavy shell-stabilized elements, 107, whose properties match those of the ekarhenium postulated by Fermi, was identified in Darmstadt in 1981, 47 years after his prediction. Since then we have also produced and identified elements 108 and 109. Measurements of their binding energy show that we have already stumbled into the region of the superheavy elements. We are now investigating the limits that apply to the production of still heavier elements.

Synthesizing heavy elements by fusion requires the experimenter to walk a fine line between methods of bombardment that yield no reaction and those that cause the fused nucleus to fission instead of settling into a relatively stable state. Reduced heating of the newly formed nucleus is the most important reason for switching from light-ion bombardment of relatively heavy target nuclei to bombardment of less massive targets with relatively heavier ions (a change pioneered by Yuri Oganessian and his colleagues at Dubna). For example, when lead 208 or bismuth 209 fuses with chromium 54 or iron 58, the excitation energy of the new nucleus is about 20 MeV. On the other hand, in the fusion of heavy actinide targets (californium 249, berkelium 249 or curium 248) with carbon 12, nitrogen 15 or oxygen 18, an excitation energy of about 45 MeV results.

The nucleus formed from light ions and actinide targets relaxes by means of the emission of four neutrons. In contrast the one formed from lead or bismuth and heavier ions relaxes by emitting only a single neutron. Because the probability that a nucleus cools down by emission of a neutron is only a few percent of its likelihood of undergoing fission, each neutron-emission step that a nucleus undergoes significantly reduces the ultimate yield of superheavy nuclei. The one-neutron relaxation mechanism is much more likely to allow a newly formed nucleus to remain intact.

Unfortunately cold fusion has a disadvantage as well: the repulsive electrostatic forces between the two nuclei make it more difficult for the two nuclei to fuse. When two nuclei approach each other part of their kinetic energy is converted into excitation energy of the transient collision system and is no longer available for overcoming the fusion barrier, thus reducing the chances of fusion. For cold fusion using heavy ions more kinetic energy is converted during the ascent and passage of the fusion barrier, and so the probability of overcoming the fusion barrier is much reduced, compared with a reaction between light ions and the heaviest targets.

If the initial energy is increased to compensate for these losses, excitation energy increases and the number of surviving nuclei goes down. As a result only beyond element 106 are there clear advantages for cold fusion. We have shown that the largest cross sections for forming heavy elements lie in a narrow energy range of about 5 MeV above the fusion barrier.

While the theory of how to produce superheavy nuclei may be an interesting exercise, the practice is a much more complex affair. Theories must mesh with accelerator and target designs as well as with the development of detectors that can verify the existence of a superheavy nucleus once it has been synthesized. When the idea of superheavy elements caught the imagination of physicists and chemists at the end of the 1960's, no one in Germany was experienced in nucleosynthesis. For newcomers to the field many doors stood open. There was much to learn from the old hands in Berkeley and Dubna, but it was clear that further progress would not be made by copying them. What was needed was an accelerator for heavy ions, fast separation methods to isolate the new elements and corresponding techniques for identifying them. And the question of exactly which nuclear reactions would lead to success was still unanswered.

In 1969 the German central government and the state of Hessen agreed on joint funding of a new institute for heavy-ion research (Gesellschaft für Schwerionenforschung), to be located in Darmstadt. The Universal Linear Accelerator (UNILAC), around which the work of the GSI is built, has been operational since 1975.

The UNILAC can accelerate all ions up to and including uranium to energies above the Coulomb barrier. The machine was designed from the beginning to produce the highest possible ion currents. Particular effort was made to enable the ion energy to be varied in small steps and to allow it to be reproducibly set at a given value. The accelerator's design was initially worked out by Christoph Schmelzer and his colleagues in Heidelberg. The design also builds on previous work elsewhere: the ion sources are a modified version of sources used in Dubna to produce highly charged ions, and the Alvarez structure developed at Berkeley forms part of the high-frequency linear accelerator.

As the UNILAC was being built many groups addressed the question of how best to use the machine: Which reactions and which techniques should be used? In its infancy the UNILAC was used to test all manner of ideas, but eventually one strategy—cold fusion coupled with recoil transport of fusion products—won out.

Since the discovery of plutonium in 1941, about 400 tons of this element has been synthesized, a quantity corresponding to 10^{30} atoms. On the other hand only a few atoms of element 109 have been produced and identified. Why have the heaviest elements been produced in such vanishingly small quantities? The answer is that tons of neutrons have bombarded blocks of uranium 238 several centimeters thick or more to make plutonium; the UNILAC, on the other hand, has accelerated only 100 micrograms of iron 58 onto lead 208 targets a few hundred nanometers thick. In addition the cross section for neutron capture, the reaction that produces plutonium 239, is roughly 10 trillion times that of the cross section for the nuclear fusion that yields element 109.

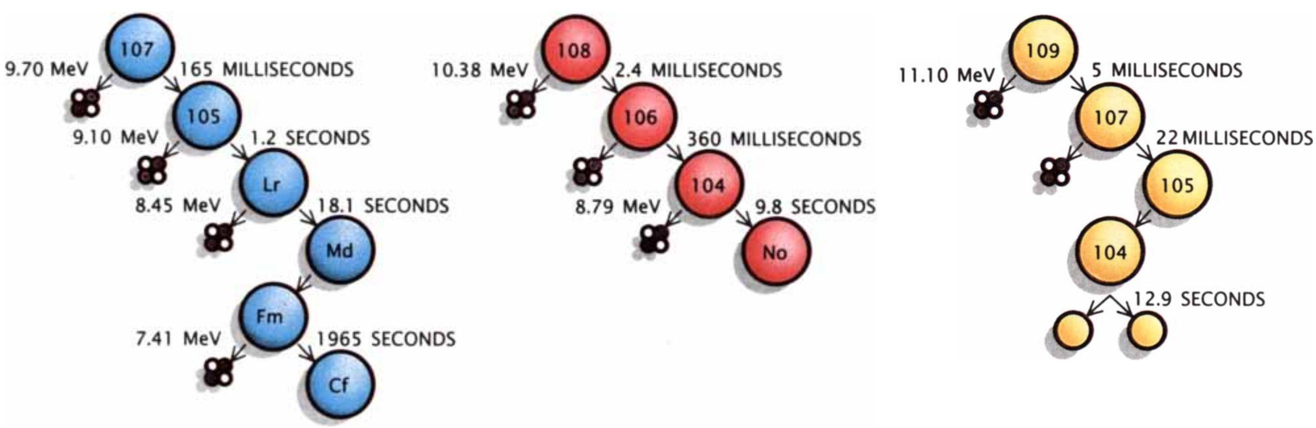
The increasing difficulty of producing ever-heavier elements is only part of the story. Once synthesized, elements such as 109 decay so rapidly that synthesis cannot keep up with decay. The heavier elements are so short-lived that by the end of the irradiation all atoms created have already decayed. These atoms must therefore be detected and identified during the production process itself.

For elements up to 106, production and detection techniques relied essentially on mechanical means to transport the newly produced elements out of the reaction area to a detector. The transport time between creation and detection of the reaction products was determined by the flow velocities of the reaction products in gases, by their diffusion times out of solid surfaces or by the speed of rotating target wheels. These techniques, however, were not good enough for detecting elements heavier than 106. They imposed unacceptable trade offs between speed and accuracy of detection, so that the faster methods were not able to identify newly produced isotopes with certainty.

The technique we adopted in Darmstadt was to let the recoil velocity imparted to the reaction products by the incoming heavy ions transport newly formed nuclei to our detector. When a heavy ion strikes a target atom and fuses with it, the resulting nucleus continues along the ion's original path at a velocity up to several percent of the speed of light. As a result, nuclei with half-lives as short as 100 nanoseconds can be detected.

Although the recoil transport technique makes possible the detection and identification of very short-lived nuclei, it also complicates the detection apparatus. Not only do the individual nuclei formed by the fusion reaction travel out of the reaction area at high speed, but so do trillions of heavy ions as well as thousands of atoms knocked out of the target foil. To separate the superheavy nuclei from the rest of the beam, we installed a velocity filter (SHIP—separator for heavy-ion reaction products), developed in collaboration with the Second Physics Institute of the University of Giessen. Because of the kinematics of the collision and fusion reaction, all the fusion products recoil at a velocity that can be calculated beforehand. Therefore, they can be separated in a relatively straightforward fashion.

The filter consists of two stages, each employing both electric and magnetic fields. The two fields tend to deflect charged objects in opposite directions; only for a nucleus traveling at the correct velocity will the effects cancel so that it continues along the midplane of the device. Our tandem filter reduces the number of accelerated ions entering the detection area by a factor of 100 billion and the number of recoil nuclei from the target by a factor of 1,000.



SUPERHEAVY ELEMENTS 107, 108 and 109 were identified on the basis of decay patterns observed at the Institute for Heavy-Ion Research (GSI) in Darmstadt. Detectors registered the mass of the incoming fusion product, followed by a series of alpha-particle emissions and, in the case of element 104, spontaneous fission. The probability of such correlated signals arising by chance is less than one in 1,000 trillion; a single decay series can thus prove the existence of a superheavy element.

Although the SHIP spectrometer eliminates almost all of the unwanted particles from the beam, it allows more than 40 percent of the fusion products to pass. The detectors that follow that spectrometer record the decay chains of the particles coming through the spectrometer, so that we can identify the fusion products with complete certainty.

The first element in the final detector system is a time-of-flight arrangement that measures a particle's velocity for the third time (the first two measurements are the basis for the velocity filters). After passing through the time-of-flight detector, the particle implants itself in spatially sensitive silicon surface-barrier detectors that register its position and energy. The combination of flight time and energy allows a rough determination of the mass, and so fusion products can be distinguished from scattered ions and recoiling target nuclei.

A firm identification of a nucleus nonetheless requires that its decay be correlated with the decays of its radioactive daughter products. Decay processes stemming from a single nucleus will have the same spatial coordinates, and the resulting daughter nuclei will undergo decay processes whose type, energy and half-life are known from previous measurements.

Given such correlated decays it is possible to identify each fusion-product nucleus unambiguously. Although a spurious particle located at the same place as our putative fusion product might decay and cause a spatially correlated signal, it is extremely unlikely that its decay energy, half-life and type would match that predicted for the fusion product. We have observed such decay chains up to the fourth generation; the likelihood that such a series of correlated events would occur by chance is somewhere between one in a quintillion and one in a sextillion. If correlated events from a desired isotope occur at a rate of one per day, then one would have to wait for 100 times the age of the earth to see one spurious signal comprising four generations of decays. As a result even a single event can demonstrate unequivocally the existence of a given superheavy isotope.

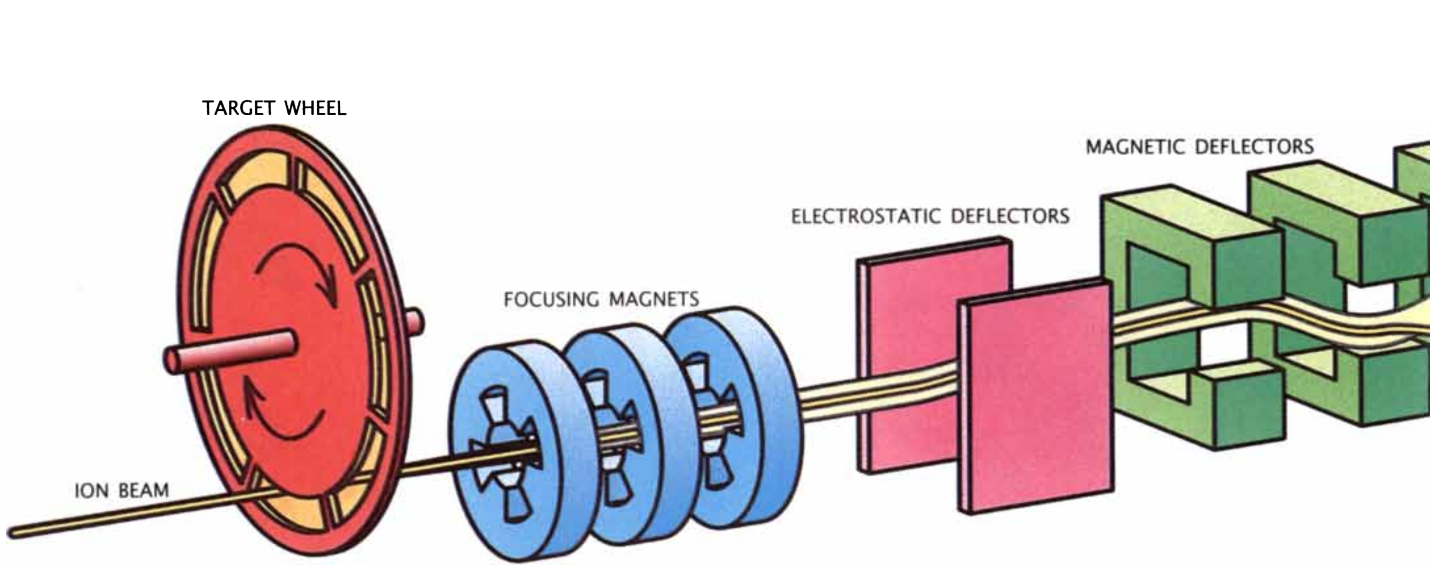
Between 1981 and 1986 we and our collaborators Peter Hessberger, Sigurd Hofmann, Matti Leino, Willibord Reisdorf and Karl-Heinz Schmidt employed the UNILAC and its detector system to synthesize and identify elements 107 to 109. In the course of our experiments we synthesized 14 isotopes of elements 104 to 109 (of which five were known previously) as well as two further isotopes of elements 107 and 108 with mass numbers 261 and 264, respectively.

In 1981 we produced an isotope of element 107 with mass number 262 by bombarding bismuth 209 with chromium 54. For the odd-odd isotope of element 107—in which both the proton and the neutron numbers are odd—we know five alpha-decay energies, which provide insight into the energy levels of the nucleus; we can tell that the isotope also possesses an isomer (a long-lived excited state).

Element 109 was identified on the basis of a single decay chain observed at 4:10 P.M. on August 29, 1982, in the reaction between iron 58 and bismuth 209. The nucleus of $^{266}109$ survived for five milliseconds before emitting an 11.1-MeV alpha particle; the resulting nucleus of element 107 decayed into 105 after 22 milliseconds; 105 decayed into 104, followed 12.9 seconds later by spontaneous fission of the nucleus of element 104. From this one chain the decay energy, half-life and cross section could be determined, although with limited accuracy. Two more decay chains were observed in early 1988—six years after the identification of element 109. They confirm the interpretation of the event observed in 1982.

In 1984 we identified three decay chains of the isotope $^{265}108$ in the reaction between iron 58 and lead 208. The two identified isotopes of the elements 107 and 109 are odd-odd isotopes with strongly reduced fission probability, but the isotope of element 108 has an even number of protons and an odd number of neutrons. Even though the fission probability of even-odd isotopes is significantly larger than for odd-odd ones, $^{265}108$ displays alpha decay.

It is of particular interest that no isotope of elements 107 to 109 displays spontaneous fission and that the even-even isotopes $^{256}104$, $^{260}106$ and $^{264}108$ all show nearly the same stability with respect to spontaneous fission. The approximately constant level of stability shows how shell-stabilization effects counteract the over-



EFFICIENT FILTER is crucial to the detection and identification of superheavy elements. Signals from the ion beam that bombards the target wheel would otherwise overwhelm the detector system. The SHIP (separator for heavy-ion reaction products) filter at GSI consists of two sets of electrostatic and magnetic deflectors that allow only ions having the proper

all decrease in stability that occurs with increasing mass.

On the far side of elements 104 and 105 lies a small island of nuclei that decay by alpha emission into known isotopes of lighter elements. These alpha decays let us determine the binding energy of the superheavy elements. In each step of the decay, if the binding energy of the daughter nucleus is known, it is possible to calculate the binding energy of the initial nucleus from its alpha-decay energy. For a chain of alpha decays it is possible to arrive at the binding energy of the initial member of the series if the binding energy of the end product is known. The decay chain from $^{264}108$ to $^{260}106$ to $^{256}104$ to $^{252}102$ can be reconstructed because one event has been observed for the decay of each of elements 108 and 104 and several for element 106. The binding energies for these nuclei are 120, 106 and 94 MeV, respectively.

The shell correction to the binding energy increases steadily for all isotopes in the range from uranium 232 to $^{264}108$ that are connected by alpha-decay processes; the corresponding values rise from 1 to 2 MeV up to 6 to 7 MeV. In fact, the elements from uranium to element 108 all have equally high fission barriers of about 6 MeV. In contrast to uranium, which would still be stable as a nuclear droplet, elements 106 and 108 owe their stability entirely to the quantum-mechanical arrangement of their many-body fermion systems. Recent theoretical work predicts fission barriers that agree with our measurements.

The height of the fission barrier and its width for the most part determine the half-life of an element with respect to fission. Shell corrections increase the lifetimes of elements 106 and 108 by 15 orders of magnitude. On a logarithmic scale the observed half-lives lie midway between the *eigenzeit* of a nucleus (the roughly 10^{21} seconds it would take for an unbound collection of nucleons to fly apart) and the age of the universe (10^{18} seconds). It is only in comparison with a human lifetime (2×10^9 seconds) that the new elements are unstable. To become stable on this scale the half-lives would have to be greater by 12 orders of magnitude. Nuclear physics, however, is not based on human time scales.

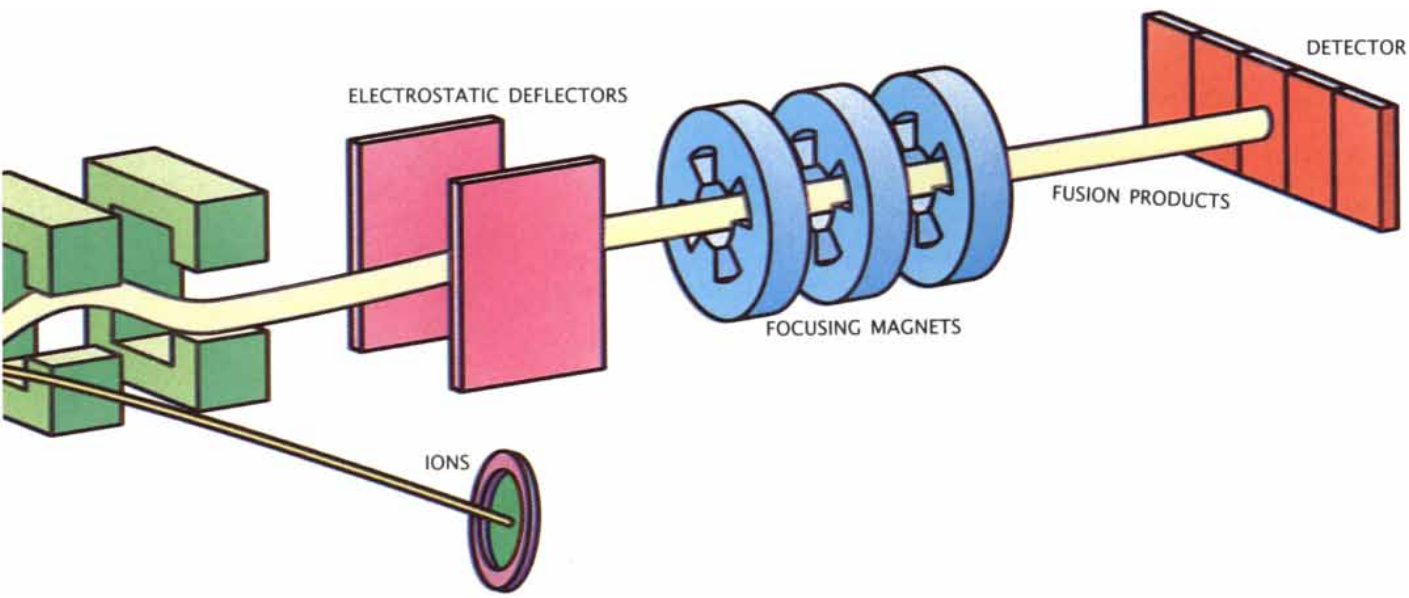
The island of alpha activity that we have found is a direct consequence of the stabilization of these isotopes by shell corrections. Thus, the stabilization predicted in the late 1960's for the spherical superheavy nuclei around element 114 actually begins much earlier than expected and grows steadily. In the narrow instability minimum on the far side of lead, between elements 83 and 90, the shell corrections decrease. But between elements 92 and 114 the shell-correction energies increase slowly and smoothly.

Even the dikes of the island of superheavy nuclei are stabilized by quantal ordering of the fermion system, whereas the mainland of nuclei is stabilized by the bulk properties of liquid drops. The nuclei of elements 107 to 109 lie on the dike between the island and the mainland, and it is merely a question of viewpoint whether the new isotopes belong to the island or to the mainland. In any case they too—like the superheavy elements—can only be detected because of the shell stabilization of their ground states.

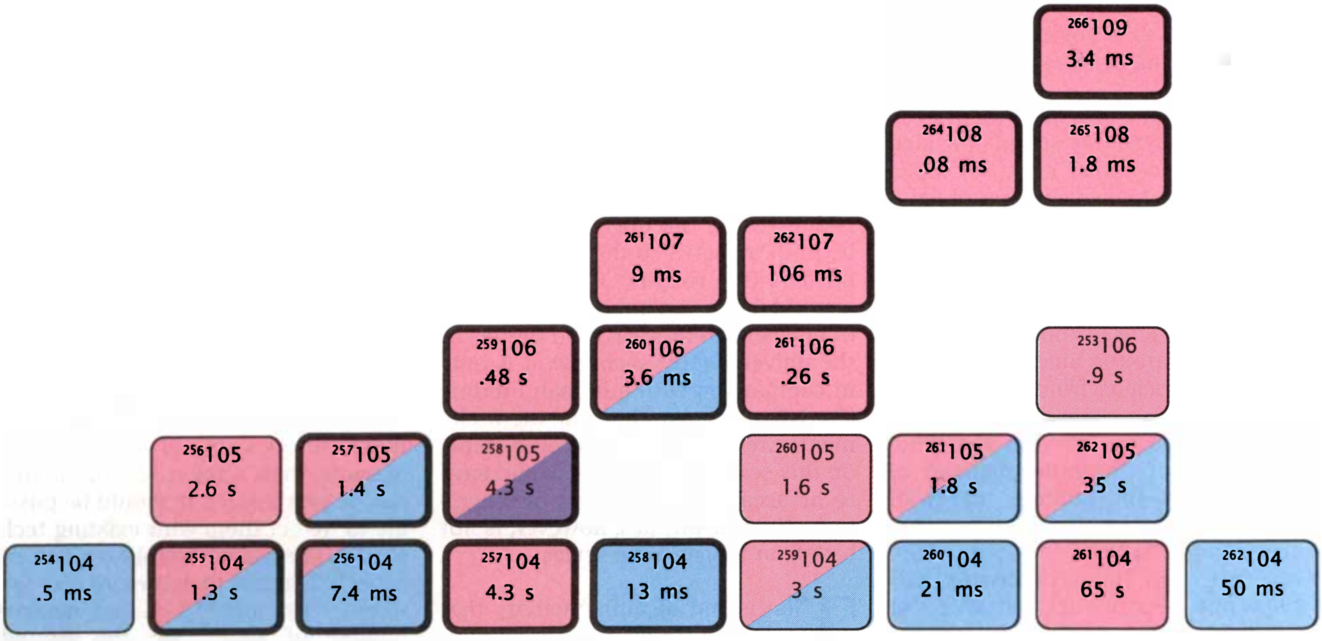
If we look at the latest predictions of the theory for the shell corrections to the binding energy, we see a region of almost 400 superheavy nuclei between elements 106 and 126 whose fission barriers exceed 4 MeV. All those isotopes should have half-lives of more than a microsecond; if they can be synthesized it should be possible to detect them with existing techniques. Particularly stable regions are expected in the vicinities of the isotopes $^{273}109$ and $^{291}115$. At a neutron number of about 166, the ground-state deformation changes. Isotopes lighter than this are deformed, whereas the heavier ones are spherical.

In the past 20 years all attempts to produce isotopes that lie in the expected center of the stabilization at $^{298}114$ have failed. Neither fusion nor any other heavy-ion reaction has led to the detection of these superheavy isotopes. Nonetheless, the basic idea of possible shell-stabilized nucleon systems beyond the stable nuclear droplets has been confirmed by the experiments described here. There remains every reason to believe that the theory can be extrapolated to even heavier elements.

An interesting question which now



velocity to pass. Ions that do not match the characteristics of the expected fusion products are diverted out of the beam path. SHIP reduces the number of heavy ions in the beam by a factor of 100 billion and the number of recoiling target nuclei by a factor of 1,000. A superheavy nucleus, on the other hand, has better than a two in five chance of reaching the detector.



SUPERHEAVY ISOTOPES are distinguished by their resistance to decay by fission (*blue*). Instead of fissioning they decay by emitting alpha particles (*red*) or by electron capture (a process in which one proton turns into a neutron, reducing the atomic number of the nucleus by one). In contrast to the general trend of reduced stability with increasing atomic number, the half-life of element 109 is greater than that of 108. Isotopes outlined in bold are those synthesized by the authors' group.

arises is: What is it that finally limits the creation of these fragile objects? Our intensive studies of fusion reactions have presented us with some important clues. Shell-stabilized nuclei with spherical ground states can be destroyed by excitation at energies of as little as 15 MeV—this was demonstrated experimentally by Karl-Heinz Schmidt as early as 1979—whereas deformed nuclei can survive excitations of up to 40 MeV. Even in the reaction between calcium 48 and curium 248—the most suitable accessible reaction—one finds an excitation energy of about 30 MeV. This means that we can only produce superheavy elements with deformed nuclei. These attempts, however, have to date only succeeded for elements with atomic numbers smaller than 110.

The fusion of two nuclei to form a superheavy nucleus is, as mentioned earlier, hampered from the very beginning by the need to overcome the fusion barrier. For a given product nucleus the probability of overcoming the fusion barrier is smallest when the lightest possible ions bombard the heaviest possible targets. In spite of its advantage in ease of fusion, this most asymmetric combination has the disadvantage of producing maximum heating of the product nucleus, thereby leading to the largest losses through fission on deexcitation. The less asymmetric the combination, the smaller the losses at the cooling-down stage. The best compromise between small losses at the final stage and probability of large formation at the initial stage is offered by the more symmetric combinations with target nuclei in the vicinity of lead.

The use of lead and bismuth as targets makes twofold use of the shell correction of these nuclei: the strong binding of these nuclei with their double-closed shells leads to a reduction in the energy delivered to the product nucleus of more than 10 MeV and to a corresponding reduction in the fission losses. Moreover, the chance of overcoming the fusion barrier is greater when spherical, strongly bound and relatively hard nuclei are used as reaction partners. Here again, the strong shell effects of lead come into their own—this time, however, in the dynamics of the process.

We are now beginning to understand why even heavier elements will be very difficult to produce. It is only the interplay of shell corrections in fusion partners that have closed shells, shell effects in the dynamics of the fusion reaction and the higher stability of deformed superheavy nuclei under excitation that enables us to create a few isotopes of the lightest superheavy elements. We have to extend the original question about the existence of shell-stabilized nuclei to the effect of shell corrections in all stages of the synthesis process. Introduction of already existing ordering and avoidance of unnecessary disordering in the fusion reaction are the essentials for the creation of those complex and fragile objects.

How will further superheavy elements be produced? For elements 110 and 111 it should be possible to use the methods we have developed for reactions between nickel 62 and lead 208 or bismuth 209. Once the elements are produced, detecting them is less a question of essential new expertise and more a matter of enriched-isotope requirements coupled with the patience to master and apply our technique over periods of months.

FURTHER READING

IDENTIFICATION OF ELEMENT 107 BY α-CORRELATION CHAINS. G. Münzenberg et al. in *Zeitschrift für Physik A*, Vol. 300, No. 1, pages 107–108; June, 1981.

EVIDENCE FOR ELEMENT 109 FROM ONE CORRELATED DECAY SEQUENCE FOLLOWING THE FUSION OF ^{58}FE WITH ^{209}BI. G. Münzenberg et al. in *Zeitschrift für Physik A*, Vol. 315, No. 1, pages 145–158; January, 1984.

ON THE PRODUCTION OF HEAVY ELEMENTS BY COLD FUSION: THE ELEMENTS 106 TO 109. Peter Armbruster in *Annual Review of Nuclear and Particle Science*, Vol. 35, pages 135–194; 1985.

OBSERVATION OF THE ISOTOPES ^{264}108 AND ^{265}108. G. Münzenberg et al. in *Zeitschrift für Physik A*, Vol. 328, No. 1, pages 49–59; September, 1987.

RECENT ADVANCES IN THE DISCOVERY OF TRANSURANIUM ELEMENTS. Gottfried Münzenberg in *Reports on Progress in Physics*, Vol. 51, No. 1, pages 57–104; January, 1988.

AN EAGLE IN

PRESENTING PREMIER ES LIMITED.

The only thing "limited" about this new Eagle is the number we're producing.

Everything else about this pure white beauty is taken to the limit. A quick scan of the details will reveal four-wheel disc brakes and 15-inch performance tires and wheels. Air conditioning, power windows, and power door locks are also standard equipment. Then there's its shape—one of the most slippery sedans ever to emerge from a wind tunnel.

Yes, the ES Limited will stop traffic. But it also pulls away from it. Quickly and easily. With power from features that maximize engine output, like hemispherical combustion chambers.

The ES Limited has independent suspension at all four wheels. And an advanced torsion-bar rear suspension. A major reason for its world-class ride and handling. Inside: roomy, comfortable, and generously fitted with genuine leather. With functionally positioned instruments and controls that would do credit to a modern fighter jet.

Like every Eagle, the Premier ES Limited is protected by Chrysler's exclusive 7-year/70,000-mile Protection Plan.*

7/70

Take a test flight in the Premier ES Limited at your Eagle Dealer. It's the only way to fly.

For further information, call 1-800-JEEP-EAGLE. Call this same number to find out about the Eagle Information/Game PC Disk.

Jeep
Eagle

Expect the Best.

FULL FLIGHT.

*Protects engine and powertrain for 7 years or 70,000 miles and against outerbody rust-through for 7 years or 100,000 miles. See limited warranty at dealer. Restrictions apply.

Buckle up for safety.

Topobiology

The word refers to "place-dependent" interactions, which appear most strikingly in the developing embryo. Surprisingly, their study has provided a major clue to the origins of the immune system

by Gerald M. Edelman

One of the most gratifying things about science is that it is impossible to predict just where a particular trail of research will lead. Indeed, an extended investigation may wind up shedding light on a long-standing puzzle in a field that initially appeared to be quite distant. Just such an unexpected connection has recently appeared between the two areas of biology on which I have spent most of my scientific career working. The first of these areas was the structure of antibodies, the immune system molecules that recognize invaders and trigger their destruction. That work reached its culmination in the late 1960's with the elucidation of the complex structure of antibodies.

Then, in the early 1970's, it became apparent that antibodies belong with certain other immune system molecules in a single evolutionary entity called the immunoglobulin superfamily. Although that was a striking finding, I was perhaps not quite as intrigued as I might have been a few years before, because by that time my interest had turned to another problem: how cells interact in a developing embryo to yield an organism. The new work led to the discovery of cell-adhesion molecules, or CAM's, which are proteins that mediate interactions between cells in the embryo. More recently insights derived from the analysis of CAM's and related molecules have begun to lay the groundwork for a molecular embryology, which links the form and function of embryonic tissues to evolution and genetics. Most of this article will be devoted to describing those recent investigations.

Yet before the article ends I shall return to the subject of my earlier work. The reason for the recursion is that, ironically, the discovery and analysis of CAM's have resolved the long-outstanding problem of the evolutionary origin of the immunoglobulin superfamily. Recently it has been shown that the genes for CAM's and those for immunoglobulins have so many similar DNA and protein sequences that they must be related in evolution. Moreover, CAM's are widespread in the animal kingdom, whereas the adaptive immune system is limited to vertebrates. Therefore, it seems likely that the immune system molecules arose from the genes of the cell-adhesion system in a remarkable piece of evolutionary opportunism.

GERALD M. EDELMAN is Vincent Astor Professor at Rockefeller University and director of the Neurosciences Institute on the Rockefeller campus. After getting a B.S. and an M.D. Edelman earned a Ph.D. at Rockefeller (then known as the Rockefeller Institute for Medical Research) for work on the structure of gamma globulin, a crucial immune system protein. After receiving his doctorate in 1960 he stayed at Rockefeller to continue the research. By 1969 he had described the amino acid sequence and internal structure of gamma globulin, work for which he received the Nobel prize in physiology or medicine in 1972. Thereafter he turned his attention to the problem of how form emerges in the animal embryo. This article, his third for SCIENTIFIC AMERICAN, describes how those two apparently unconnected areas of research ultimately intersected.

Both antibodies and cell-adhesion molecules have their main effects at the surfaces of cells. Surface interactions between cells can lead to changes in gene expression and in cell shape, movement and function. Just which reactions will take place when cells interact depends in part on the history of the cells: on what interactions they have had with other cells in the past. What is more, because these interactions depend on what cells surround any given cell, cells react differently at different places. Such "place-dependent" interactions do occur in the immune system, but they are particularly important in embryonic development. Indeed, the fates of cells—what they will become in the mature organism—are crucially dependent on the place or neighborhood of the cells in the embryo.

Although embryologists have long been aware in general of the possibility that place is a critical element in determining cell fate, it seemed necessary to refocus the subject by giving a name to the study of those place-dependent interactions at the cell surface that lead to cell regulation. About two years ago I coined the term topobiology (from the Greek word *topos*, meaning "place") to describe this dynamic, interactive process.

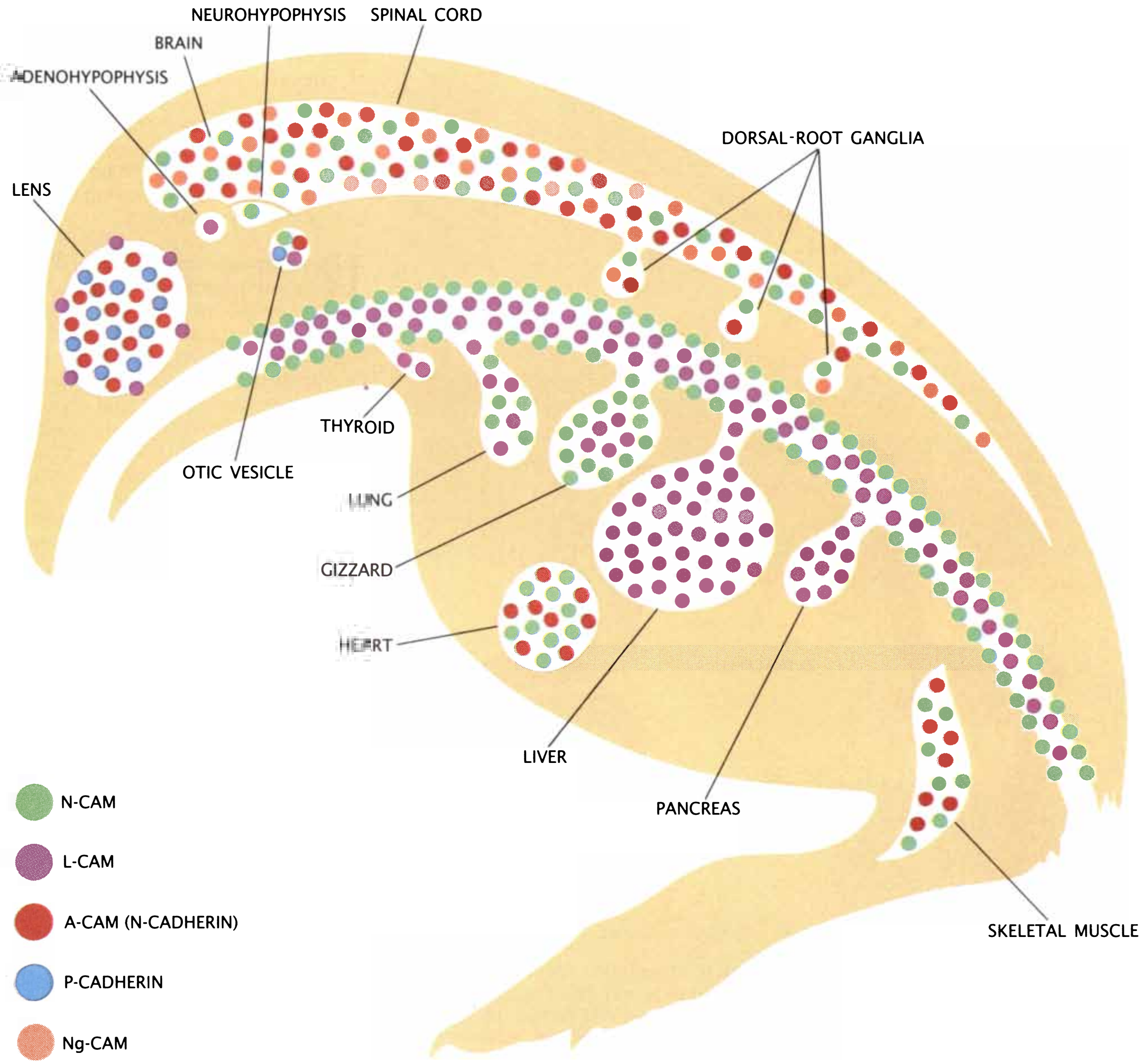
As the discovery of CAM's made clear, one of the key factors determining the place of an embryonic cell (and ultimately the form and pattern of the tissues) is the presence of cell-adhesion molecules. In the past few years several families of molecules have been shown to mediate such adhesions and the related intercellular transactions [see "Cell-Adhesion Molecules: A Molecular Basis for Animal Form," by Gerald M. Edelman; SCIENTIFIC AMERICAN, April, 1984]. Aside from the CAM's themselves, the most important of these are substrate-adhesion molecules, or SAM's, and cell-junctional molecules, or CJM's. Whereas CAM's are always on the cell surface and mediate cell-to-cell interaction, SAM's are thrown out by cells into the environment, forming part of the complex extracellular matrix to which cells sometimes attach. CJM's serve in the formation of the complex structures (including the so-called tight junctions, gap junctions and adherens junctions) that join cells together in tissues.

All of the known CAM's, SAM's and CJM's are complex proteins, and the structures of some of them are becoming reasonably well understood. Attention is now focusing on how each of these adhesion molecules functions in the embryo and whether the functions of different molecules are interrelated. It has been established that almost all known CAM's bind to one another by a mechanism that is homophilic: a CAM on one cell joins a CAM of the same type on an apposing cell. Yet there are differences among CAM's; they show different binding specificities and varying dependence on ions, for example the calcium ion.

One of the recurrent motifs of embryonic development is the transition from epithelium to mesenchyme and back. A mesenchyme is a group of loose or migrating cells that are not firmly organized in a geometric sense. An epithelium, on the other hand, is a sheet of cells that is highly organized geometrically and often joined at its base to the extracellular matrix. The joining of the epithelium to the matrix is accomplished by SAM's, and the epithelial cells are often linked by junctions composed of various CJM's. Because the transition from one collective form to another is an underlying theme of development, it would be valuable to know which of these molecules is responsible for the initial conversion of mesenchyme to epithelial sheet.

Recent experiments in my laboratory suggest that the initial role in linking epithelial cells is played by CAM's and that the formation of junctions such as gap junctions and adherens junctions depends on CAM linkage. That result was obtained by inserting CAM DNA, by means of the process called transfection, into cultured cells that ordinarily lack it. The CAM's that were chosen were of two types—neural CAM (N-CAM) and liver CAM (L-CAM). These were the first CAM's to be discovered, and they were named for the tissues in which they were first found. They are now known to be much more widely distributed in the embryo.

Before transfection of L-CAM the cells were separate and resembled a mesenchyme. After transfection of the



DISTRIBUTION OF FIVE KNOWN CAM'S (cell-adhesion molecules) is shown for a chick embryo just before hatching. The distribution of CAM's (each with a different binding specificity) suggests they are involved in the generation of form in the embryo. Their spatial distribution changes over time: earlier, the areas overlap and cover most of the depicted area.

L-CAM gene the cells were linked in a sheet that had some of the characteristics of an epithelium. After the sheet formed, gap junctions and adherens junctions appeared between its cells. It seems likely that the action of CAM genes caused existing CJM's to assemble into junctions, and so it seems that this process is CAM-dependent. Moreover, adding fragments of antibodies against L-CAM, which block the typical homophilic binding of that molecule, led to dissolution of the sheets and a marked decrease of junctions between cells.

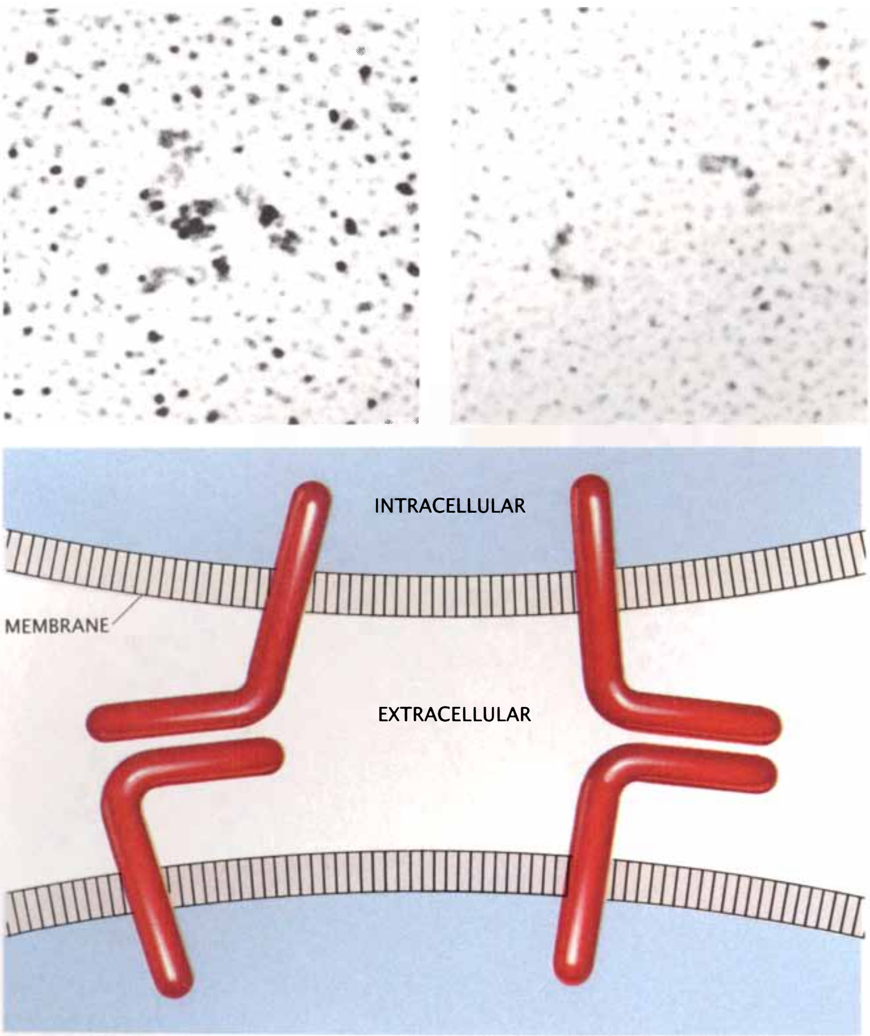
Such results give a hint of the significance of CAM's in the development of the embryo. As I noted above, the transition from epithelium to mesenchyme and back is one of the fundamental motifs of embryonic development. But it is not the only one. Among the others are certain precise tissue movements and the formation of new boundaries between tissues. These processes are part of the overall pattern of morphogenesis: the appearance of distinct, specialized organs and tissues in the animal. The significant role of CAM's in these spatial processes is shown by the fact that CAM's with different specificities are distributed differently, each one in a characteristic pattern that emerges in time and space in the developing embryo.

An excellent example of all of these topobiological processes is provided by the formation of feathers, a crucial part of the development of the chick embryo. Feathers are formed by a series of coordinated processes that include cells moving, dividing, dying in patterned ways, adhering to each other and differentiating. In its early stage much of this coordinated development depends on interactions between two types of embryonic tissue, known as mesoderm and ectoderm. From mesoderm most bones and muscles arise; from ectoderm come the skin and nervous system.

Now, in some parts of the developing chick embryo, a layer of a specialized ectodermal tissue—the epidermis—lies above mesodermal tissues. The first step in feather formation is the induction in the epidermis of structures called feather germs that later give rise to the feather. The appearance of the feather germs requires passage of an intricate pattern of chemical signals between mesoderm and ectoderm—a process called embryonic induction. The efficacy of that signaling in turn depends on the action of CAM's, which can be studied by transferring embryonic skin tissues into laboratory culture.

Recent experiments carried out in my laboratory shed light on the relations between CAM linkage and the signaling that leads to feather-germ patterning. Simply put, the thrust of the experiments was to disturb the CAM linkage and observe the effect on the intricate normal geometry of the feather germs. In the normally developing embryo the induced feather germs begin at the embryonic midline and form a hexagonal array as they spread laterally. That pattern is seen both in the placodes, the precursors of the feather germs that will form the structures of the mature feather, and in the dermal condensations, the underlying mesodermal cells that give the inducing signal needed for the formation of the placodes in the overlying epidermis.

The rationale of the experiment was that the cells of the epidermal placodes were linked by L-CAM but had no N-CAM. Conversely the dermal condensations were linked by N-CAM and had no L-CAM. We proposed that a signaling loop, set up by the CAM's, was needed for proper pattern formation. In such a loop, signals would move not merely from mesoderm to epidermis but also from epidermis to mesoderm. Therefore, we placed antibodies against L-CAM in the culture. These antibodies could only affect epidermal linkages, not those of the dermal condensations. Nonetheless, the pattern of the dermal condensations



CAM'S BIND CELLS to each other by a homophilic mechanism: a CAM on one cell binds to the same type of CAM on an apposing cell (*lower panel*). The images in the upper panel were made by shadowing CAM's with platinum and viewing them in the electron microscope. One (*left*) shows three neural cell-adhesion molecules (N-CAM's) linked by a "hub" made up of the areas that would have protruded through the cell's outer membrane. The other (*right*) shows a single liver cell-adhesion molecule (L-CAM). CAM's appear to be hinged linear molecules. The lower panel depicts two hypothetical binding mechanisms based on the hinged structure of N-CAM.

Finally. A sophisticated weapon in the war against plaque.

These days, it seems like every product from mouthwash to toothpaste wants to help you fight plaque.

And for very good reason. Plaque buildup is a leading cause of early gum disease which can have a number of very serious complications.

But among the so-called "plaque attackers," the INTERPLAK® Home Plaque Removal Instrument stands out as a true technological breakthrough.

Bristles rotating 4,200 times a minute literally scour plaque off your teeth.

The INTERPLAK instrument cleans teeth nearly plaque-free.

If plaque is not removed daily, its bacterial film can lead to gingivitis, an early stage of gum disease, and tooth decay. But clinical studies have shown that manual brushing removes only some of the plaque build-up. Those same studies, on the other hand, show that the INTERPLAK instrument cleans teeth nearly plaque-free, and reduces gingivitis to improve the health of your gums.

It cleans circles around ordinary brushing.

With manual or even electric toothbrushes, you move the bristles up and down or back and forth. But with the INTERPLAK instrument's patented design, the brush remains still while the bristles rotate. Ten tufts of bristles rotate 4,200 times a minute, reversing direction 46 times a second. They literally scour off plaque and stimulate your gums. And at the precise moment they reverse direction, the tufts fully extend to clean deep between teeth and under gums. Yet because the bristles are four times softer than the softest toothbrush, it's no more abrasive than manual brushing with toothpaste.

Dental professionals approve.

The INTERPLAK instrument has received rave reviews from dentists and periodontists across the country.

"I am recommending it to all my patients."–Dr. L.K. Yorn, Cedar Grove, NJ

"At last my patients enjoy using a product which we recommend."–Dr. J.W. Blackman, III, Winston-Salem, NC

"Since my patients have been using the INTERPLAK instrument, I have seen a dramatic improvement in the health of their teeth and gums."–Dr. S.G. Newhart, Orthodontist, Beverly Hills, CA

"It's a technological breakthrough in home dental care."–Dr. Alan Kushner, Chicago, IL

Ask your own dentist about the benefits.

Serious plaque removal for the whole family.

Each INTERPLAK Home Plaque Removal Instrument is cordless, recharging between uses in its own stand. It comes with two interchangeable brush heads. You can also purchase additional brush heads so every family member can benefit from cleaner teeth and healthier gums.

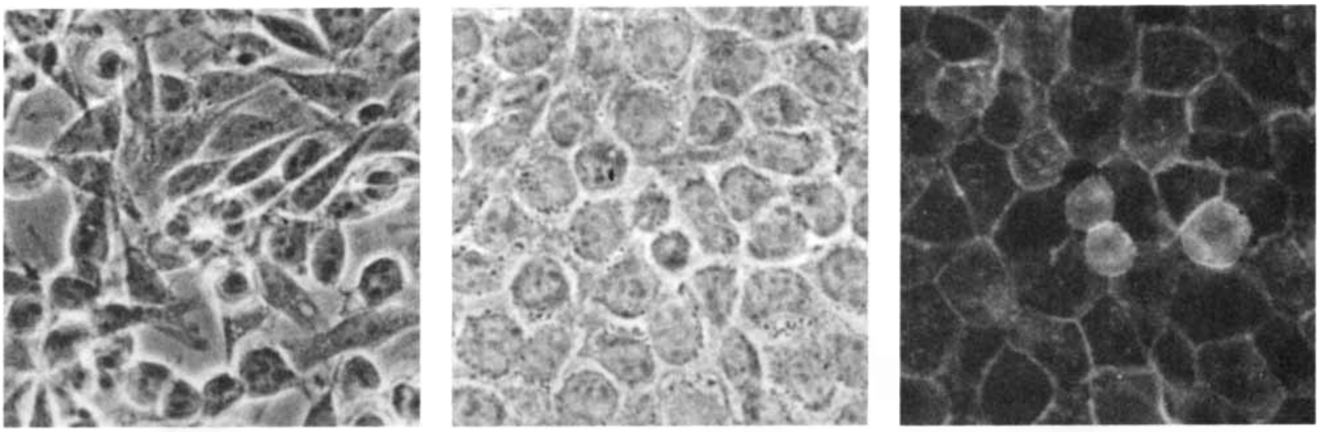
For more information or the name of a retailer near you, call toll free **1-800-544-3536,** and ask for **Operator 14.** In Canada call 1-800-554-3636.

INTERPLAK®
HOME PLAQUE REMOVAL INSTRUMENT

Interplak is Acceptable as an effective cleaning device for use as part of a program for good oral hygiene to supplement the regular professional care required for good oral health.
Council on Dental Materials, Instruments and Equipment, American Dental Association

ADA ACCEPTABLE American Dental Association

INTERPLAK® is the registered trademark of Bausch & Lomb Oral Care Division, Inc.
©Bausch & Lomb Oral Care Division, Inc., 1989



CAM'S CHANGE THE FORM of groups of cells, as shown in micrographs made in the author's laboratory. Cultured cells lacking the L-CAM gene assume a loose form similar to one known in the embryo as a mesenchyme (*left*). When the L-CAM gene is inserted into the cells and activated to make L-CAM, the cells form a more regular arrangement resembling the one known as an epithelium (*middle*). Fluorescent techniques show that L-CAM is on the cell's surface (*lighted regions in right panel*).

was also changed after the introduction of the antibodies into the culture [*see upper illustration on page 82*].

These results indicated that changing the CAM linkages in one layer of cells can alter the signaling loop on which patterning depends. Either the response of epidermal cells to signals from below may be altered, or their own downward messages may be changed, but in either case the result is the same: normal pattern formation is interrupted. What is more, the effect is long-lasting. When the antibodies were washed out and the perturbed cells cultured for 10 days, the normal pattern of feather development was severely disturbed. As these experiments indicate, correct formation of patterns depends on the correlation between CAM linkages and cellular responses to inducing signals.

That correlation retains its significance as the feather develops. The mature feather is a complex structure with a central spine, or rachis, from which smaller spines called barbs diverge. The barbs are held together by even smaller filaments that project from them called barbules. All these structures are made up of a fibrous protein called keratin, which is deposited by the cells of the feather as they die at the end of the process of maturation. The barbs and barbules are formed in so-called feather filaments from folds called barb ridges [*see illustration on opposite page*]. After swelling out from the inner surface of the filament, the folds ultimately pinch off to yield cylindrical structures that interconnect; these cylinders are the barbs, ultimately linked by barbules.

The cells of the barb ridge arise from precursors that originally express both L-CAM and N-CAM. As the ridges develop, however, an intriguing specialization occurs that distinguishes the ridges from the valleys between them. As the ridges mature they come to express only L-CAM. At a certain point the cells in the valleys lose L-CAM, express only N-CAM and proliferate, which leads to the formation of regions bordering each ridge called marginal plates. Soon afterward all the cells expressing L-CAM begin synthesizing keratin and all those expressing N-CAM die. As a result, the borders between N-CAM and L-CAM expression are converted to the edges of new structures: the barbs and barbules.

Thus, the expression of specific CAM's, coupled with cell differentiation and with cell death, can lead to morphogenesis. It would be misleading, however, to imply that CAM's are the only morphoregulatory molecules. An example based on recent experiments in my laboratory shows how SAM's can also influence patterning events in the embryo. The example concerns a group of cells that give rise to the dorsal-root ganglia: bundles of nerves inserted into each vertebra in the mature organism. The vertebrae themselves develop from mesodermal tissue that is grouped into segments called somites, each somite corresponding to a single vertebra. Somites give rise to a number of other tissues as well; the mesodermal condensations of the feather, for example, come from a mesenchyme arising in somites. The dorsal-root ganglia originate in ectodermal cells that migrate as a mesenchyme from a structure called the neural crest.

In order for the correct ganglionic pattern to be made, the neural-crest cells must enter only the mesenchymal portion of the somite (the sclerotome) in its front part. How is this entry guided? Earlier studies by Jean-Paul Thiery of the National Center for Scientific Research (CNRS) in Paris had shown that neural-crest cells migrate on pathways made of fibronectin (among other molecules); fibronectin is a SAM. In addition, my colleagues and I had found that another SAM, cytotactin, is distributed in distinct patterns during the formation of the embryo. Cytotactin is capable of binding to fibronectin as well as to a third SAM known as cytotactin-binding proteoglycan, or CTBP. It seemed worthwhile to find out whether the periodic pattern produced by the neural-crest cells as they migrate into the sclerotome was correlated with the pattern of the three SAM's in the somites.

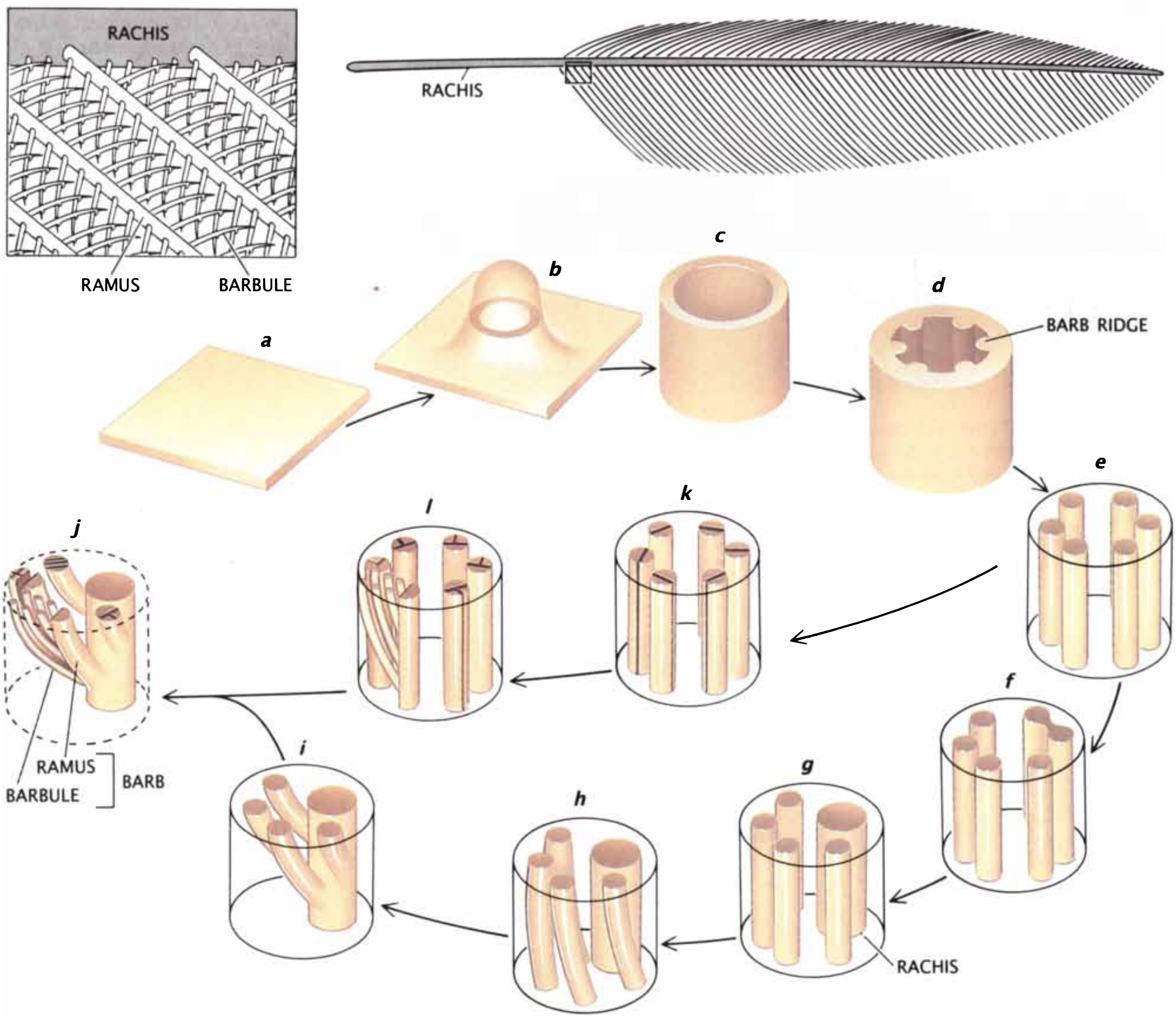
A tantalizing clue was provided by examination of the distribution of the SAM's during the invasion of the somites by the neural-crest cells. At early stages of embryogenesis cytotactin, fibronectin and CTBP were distributed evenly over the entire extent of the somite, but as development proceeded, a periodic pattern appeared: although fibronectin was still spread more or less evenly throughout the somite, cytotactin was concentrated in the front part, and later CTBP became concentrated in the rear portion. This alternating distribution of SAM's, which emerged just as the migrating cells entered the sclerotome, resulted from molecules synthesized by the cells of the somite and not by the arriving cells of the neural crest.

Separate experiments were conducted in an attempt to tease out the effects of the SAM's on cell shape and movement. Work in tissue culture showed that cytotactin and CTBP cause neural cells to assume the somewhat rounder shape characteristic of cells that are not migrating. What is more, such rounded cells would not migrate into regions of tissue containing the SAM's. In contrast, on a mesh of fibronectin alone, crest cells were flatter and migrated readily; they showed intermediate behavior on mixtures of fibronectin and either of the other two SAM's. Although the details remain to be sorted out, the general conclusion of this work is that various admixtures of different SAM's, linked to one another, form networks that can have different effects on cell behavior and movement in the embryo.

It is striking that such SAM-mediated events are topobiologically coordinated with the expression of CAM's: migrating neural-crest cells lose N-CAM at their surfaces as they move, but as they interact in the somite to form dorsal-root ganglia, N-CAM again appears on their surfaces. (Indeed, N-CAM is present on most neural cells in their fully mature, differentiated state.) Clearly CAM's and SAM's are regulated in a coordinated manner during development; both kinds of molecules can affect tissue patterning. I have already mentioned that cells probably must be bound by CAM's before CJM's can form junctions between them; it therefore seems that in general intricate dependencies come into play among these three classes of morphoregulatory molecules as an organism develops.

Yet these interactions, complex as they are, are not sufficient to account for how the embryo develops. The differentiation of embryonic tissues into mature tissues requires that they produce tissue-specific proteins: liver proteins in liver cells, muscle proteins in muscles and so on. The genes that code for such proteins are separate from the genes for cell-adhesion molecules. Yet the actions of these two classes of genes are not independent, and I have proposed the so-called morphoregulator hypothesis to account for the cycle that connects them.

In the morphoregulator hypothesis cells are controlled mechanochemically by cycles of CAM expression and by SAM networks. A good exam-



FEATHER FORMS from a precursor called a placode. Mesenchymal cells of a type called mesodermal migrate beneath a simple epidermal sheet (*a*) and form a collective that sends signals upward to induce formation of the placode. The sheet bulges to form a bud (*b*) and then a filament consisting of a cylinder of epidermis (*c*). The cylinder develops ridges of epithelium called barb ridges (*d*), which subdivide to form the branched structure characteristic of the mature feather (*e–l*).

ple of such control is provided by the changes in shape and movement of neural-crest cells caused by their binding to particular SAM's. The expression of CAM and SAM genes alters their shape and response to signaling patterns by controlling the formation of cell collectives that exchange signals at a given place. By these means expression of CAM and SAM genes affects the expression of other genes, including historegulatory genes that code for tissue-specific proteins. In topobiology the interaction of cell surfaces controls the mechanochemical driving forces, creating collectives of cells whose signaling is thereby altered, along with their state of differentiation. Such modulation of cell state by CAM's and SAM's must play a major role in the evolution of animal form and in the development of tissue pattern.

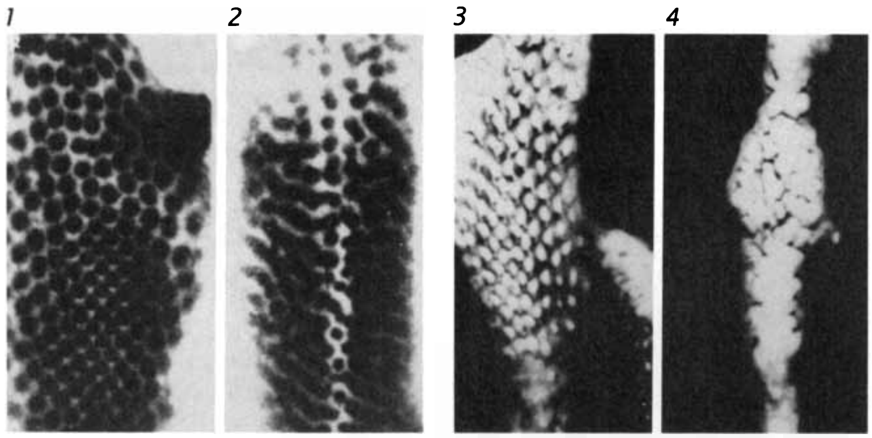
The morphoregulator hypothesis focused the attention of my colleagues and me on the genes that code for the morphoregulatory molecules, CAM's in particular. Bruce A. Cunningham and I worked out the structure of the genes for N-CAM and L-CAM by determining the DNA sequences that code for these CAM's. Our work showed that N-CAM and L-CAM are each specified by a single gene and that the genes differ substantially in their detailed structure, which suggests that the two molecules are not closely related in evolutionary terms.

The differences between the L-CAM and N-CAM genes lie partly in how they are processed to yield the ultimate message specifying a protein. Almost all mammalian genes contain coding regions called exons and noncoding regions called introns. After RNA has been transcribed from the gene, the noncoding portions are removed and the coding portions are spliced together to yield the finished messenger RNA (mRNA) molecule. In some instances alternative splicing of the mRNA can yield different proteins. That is the case for N-CAM: its 19 (or more) exons can be spliced in a variety of ways, some of which can give rise to CAM's that differ somewhat in the region that attaches the molecule to the cell membrane. Changes in that region may alter the strength of the binding of N-CAM's to each other and to the cytoskeleton, the cell's internal skeleton. The L-CAM gene, on the other hand, although it comprises a number of exons, gives no evidence of alternative splicing.

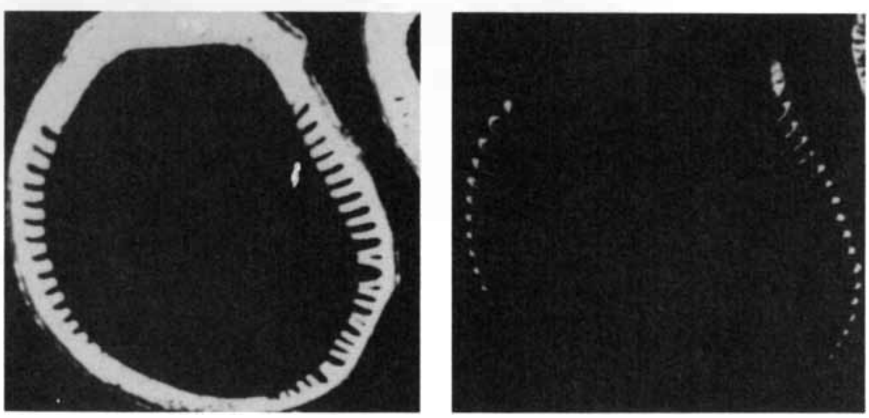
It should be emphasized that the alternative forms of splicing seen in the case of the N-CAM gene do not alter the specificity of the homophilic binding of the CAM: N-CAM molecules bind to each other no matter how their genes have been spliced. Only the collective strength of the attachment is affected by the alternative splicing of the genetic message. This is so because the number and arrangement of CAM molecules, and therefore the total binding efficacy, would be changed by altering the way these molecules attach to the cell membrane. This arrangement fits nicely with the notion that there are perhaps only a few dozen or so CAM's of different binding specificities, whose dynamic regulation at the cell surface by a variety of means can result in a very wide range of nuances of binding.

Some additional members of this small molecular family have already emerged from the laboratory. Not long after Cunningham and I reported our results, several studies indicated that N-CAM resembles another protein found in tissues of the nervous system. It is now suspected that this molecule, myelin-associated glycoprotein, or MAG, is also a CAM. Subsequent DNA work by Masotoshi Takeichi and his colleagues at the University of Kyoto showed that several other CAM's, called cadherins, exist, similar to L-CAM in DNA sequence and structure but with a different distribution in the embryonic tissues.

All these findings on the genetic structure of CAM's were very interesting, but perhaps the most striking



PERTURBATION EXPERIMENT shows CAM links are required in feather formation. As the placode develops, the mesodermal cells beneath it take on a symmetric arrangement of "blobs" (*1*) and later undergo further changes (*3*). If the tissue is cultured in the presence of antibodies against L-CAM, the "blob" geometry is perturbed to yield "stripes" (*2*); later the overall form is disrupted (*4*). The antibodies affect only L-CAM links (between epidermal cells), not N-CAM links (between mesodermal cells). Yet the mesodermal pattern is disrupted, implying a signaling loop between these tissues.



ALTERNATING PATTERN of L-CAM and N-CAM is seen as the barb ridges develop. Each panel shows a cross section through a developing feather follicle stained with fluorescent antibodies against a specific CAM. L-CAM (*left*) links ridge cells, whereas N-CAM (*right*) links the cells in the valleys between the ridges. Shortly after this stage cells with L-CAM make a fibrous protein called keratin while cells with N-CAM die.

MERCURY

INTRODUCING A COMPLETELY DIFFERENT MERCURY COUGAR.

We did more than give Cougar a sleek, new exterior. Much more. We gave it a longer wheelbase for a more comfortable ride. And an all-new suspension design with a lower, wider stance for greater stability and control. Plus new, speed-sensitive variable-assist power steering for improved road "feel" and control at highway speeds. Inside, it was obviously designed by people who know that cars are for driving. Experience the comfort and control of a Mercury Cougar at your Lincoln-Mercury dealer today. For more Cougar information, call 1-800-822-9292.

MERCURY
LINCOLN

LINCOLN-MERCURY DIVISION Ford

Buckle up—together we can save lives.

Quality is Job 1.

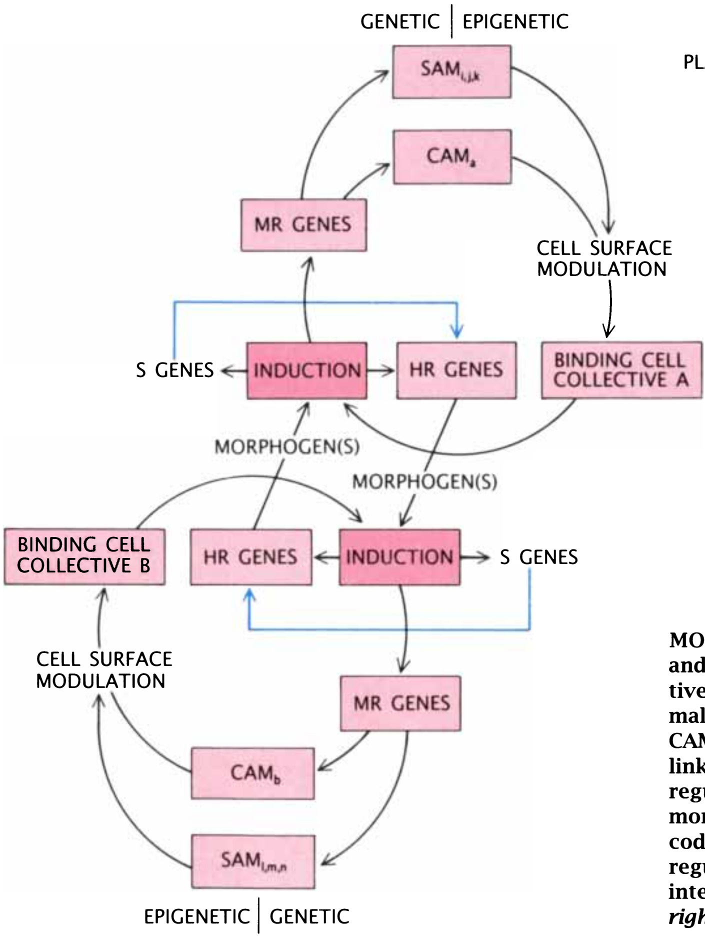
was the discovery that N-CAM shares structural homologies with the antibody molecule. When my colleagues and I first completed the analysis of the amino acid sequence of the entire antibody molecule in 1969, a number of very beautiful structural, genetic and evolutionary relations became apparent [see "The Structure and Function of Antibodies," by Gerald M. Edelman; SCIENTIFIC AMERICAN, August, 1970]. These relations were embedded in the "domain hypothesis," which held that immunoglobulins are made up of two types of structural and functional subunits, each about 100 amino acids long. These are the variable, or V, domains, which differ from molecule to molecule within a functional class, and the constant, or C, regions, which do not.

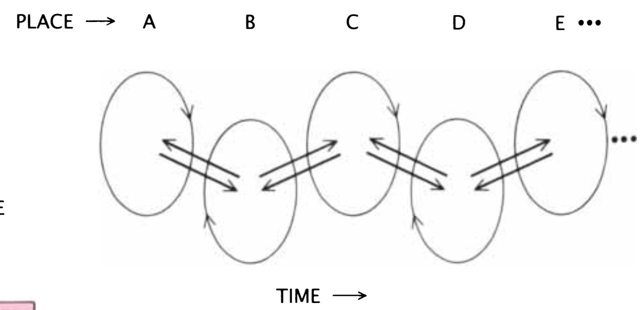
The characteristic T shape of the antibody molecule arises from a specific assemblage of V and C domains. Each antibody includes a pair of what are known as light chains, consisting of one variable and one constant region, and a pair of heavy chains, consisting of one variable and as many as three constant domains. Within this assembly the V regions are responsible for binding to the antigen, whereas the C regions carry out various "effector" functions, such as promoting the ingestion of a bound foreign protein antigen by immune system cells called macrophages. Intriguingly, all the domains, both C and V, of the antibody molecules share certain homologies of amino acid sequence, and it was proposed that they arose in evolution as the result of repeated duplication of a precursor gene that specified a region no longer than one domain.

Since its formulation in 1969 the domain hypothesis has been amply confirmed. Indeed, the evolutionary family so defined has been broadened to include other molecules; among them are growth-factor receptors and the histocompatibility antigens that mediate certain crucial interactions among cells of the immune system. This evolutionary clan has been called the immunoglobulin superfamily. As its members emerged one question remained: How did this diverse group arise in evolution?

The finding that N-CAM is homologous to these molecules suggested a possible answer to that question. I have proposed that the entire adaptive immune system, which is characterized by the presence of the immunoglobulin superfamily, arose from a more ancient cell-adhesion system. The reasoning underlying this hypothesis rests on the fact that the adaptive immune system is found only in vertebrates, implying that it arose late in evolution, whereas the system of CAM's appears to be much more widespread.

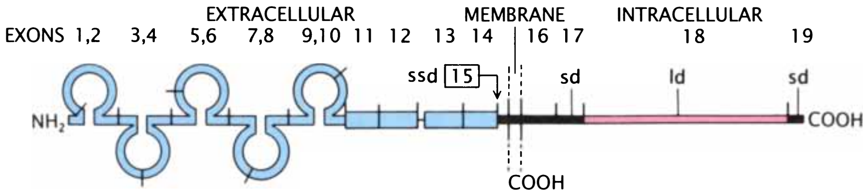
Some recent observations strongly support this notion. Thomas C. Kaufman and Mark Seeger of Indiana University have found a DNA sequence in the *antennapedia* gene complex of the fruit fly *Drosophila melanogaster* that is homologous to approximately two and one-half domains of the N-CAM gene sequence; this sequence specifies a protein called amalgam. Later Corey S. Goodman and his colleagues at the University of California at Berkeley showed that the sequence of a suspected CAM called fasciclin II, found on nerve cells of *Drosophila*, is also homologous to N-CAM. Fruit flies have nothing that resembles a system of adaptive immunity. Because insects have N-CAM-like molecules and only vertebrates have an antibody-based immune system, it appears likely that





MORPHOREGULATOR HYPOTHESIS seeks to explain how CAM's and substrate-adhesion molecules, or SAM's, create cell collectives and interact with them to determine the shape of an animal's body. In CAM cycles (*left*) collectives of cells linked by CAM's make molecules called morphogens that induce cells linked by different CAM's to alter the activity of their morphoregulatory (*MR*) genes: genes encoding CAM's and SAM's. The morphogens also affect historegulatory (*HR*) genes: genes encoding tissue-specific proteins. In some cells the *HR* genes are regulated in turn by selector (*S*) genes. As the cell collectives interact they alter their environment; new CAM cycles (*above right*) come into play, signaling to each other (*double arrows*).


Lufthansa
Smooth sailing to 84 countries around the world.
It may surprise you that Lufthansa, the airline of Germany, also flies to Australia, India, Japan, Tanzania, Yemen, anywhere business takes you.
But what shouldn't surprise you is that, no matter how unusual your destination, you can always expect our usual commitment to getting you there without the slightest turbulence.
People expect the world of us:
Lufthansa
German Airlines
Lufthansa is a participant in the mileage programs of United, Delta, USAir and Continental/Eastern. See your Travel Agent for details.

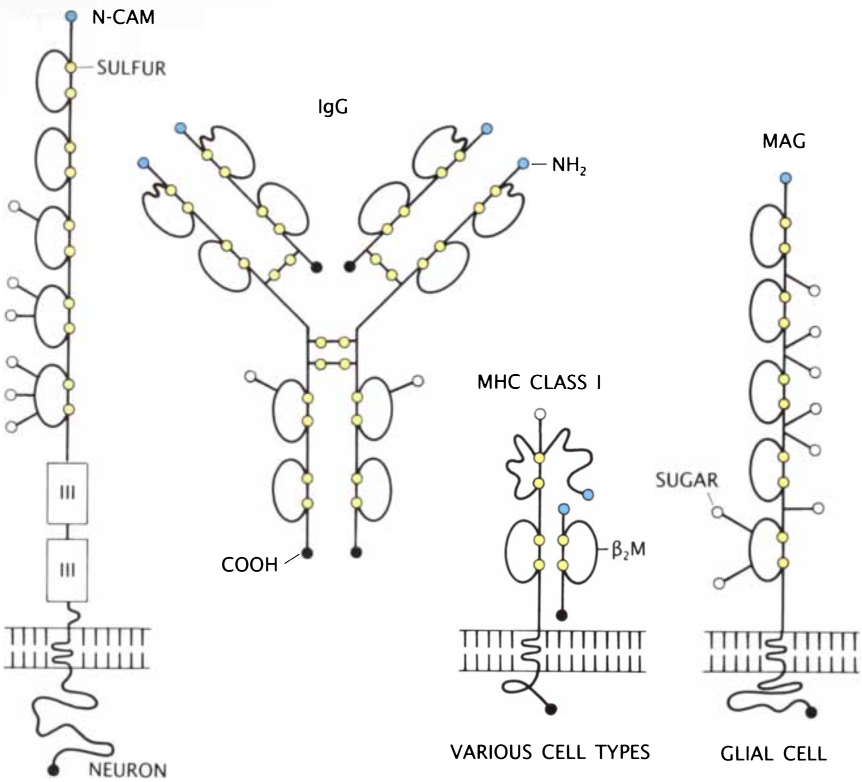


STRUCTURE OF N-CAM has been deduced from protein chemistry and determination of the DNA sequence in the N-CAM gene. N-CAM comes in several variants that differ according to how many exons (coding regions) of the N-CAM gene they incorporate. Nineteen exons are shown here; more have recently been discovered. All known variants include five loops that are linked by disulfide bonds (*left*). The illustration shows three variants: *ssd* stops at exon 15 and has no transmembrane region; *sd* includes exons 17 and 19 but lacks 15; and *ld* includes exon 18 but lacks 15.

CAM genes, present in some early precursor of both insects and vertebrates, provided the basis for the emergence of the molecules of the adaptive immune system.

Several additional facts help to provide a scheme for the evolution of the cell-adhesion molecules and their descendants that is a satisfying extension of the domain hypothesis. Unlike the CAM's, each immunoglobulin is specified by multiple genes for V and C regions that have arisen by duplication of an original precursor during evolution, and these families of duplicated genes show homologies from species to species of vertebrate. Yet it would seem that unless special mechanisms are operating, independent mutations accumulating in different genes of such a duplicated family would destroy the homology among species. What might the mechanism be?

In 1969 my colleague Joseph A. Gally and I proposed that if such genes underwent a process called "democratic gene conversion" then the families they belonged to could evolve in parallel between species. The details of the process are somewhat beyond the scope of this article, but its essence is that the genes of the family recombine with one another in a particular way that makes the coevolution possible. What is more, the families may actually serve as "mutation nets" that spread favorable mutations among their members. Since a mutation favorable for one species' adaptive immune system may well be favorable for another's, these nets, combined with selective pressure, might also help preserve the homologies in the families of immunoglobulin genes. This idea, important as it is for antibodies, is a general one and applies to the evolution of all multigene families.



IMMUNOGLOBULIN SUPERFAMILY, now known to include CAM's, is made up of molecules that are structurally and functionally diverse. Each of these molecules includes several domains. IgG, the antibody molecule, circulates in body fluids and binds to foreign substances. MHC class I proteins, components of almost all animal cells, associate with processed antigens; they are required for effective presentation of antigen to the white blood cells called *T* lymphocytes. The small chain of the MHC class I proteins is a single domain, β_2-microglobulin. MAG is a suspected cell-adhesion molecule found in membranes of glial cells, nonexcitable nervous system cells. Glial cells (and MAG) play a role in the myelination of certain neurons.

The functional collaboration of CAM's and antibodies has now been shown in the immune system in the revelation that various lymphocytes, which are among the most important effector cells of the immune system, require adhesion to carry out their functions. Recently Timothy A. Springer of the Harvard Medical School discovered a molecule called I-CAM on a variety of cells. I-CAM is homologous to N-CAM, and Springer has shown that it binds to a molecule called LFA-1, which is found on lymphocytes. LFA-1 resembles cell-surface molecules called integrins, which serve as receptors for SAM's.

If all of this information is combined into a general scheme, a magnificent example of evolutionary opportunism emerges. Its starting point is a stretch of DNA about half the size of an N-CAM domain. Analysis of the structure of the N-CAM gene reveals that this half-size precursor was duplicated to give rise to N-CAM-like genes in the ancestors of insects and vertebrates. Then, through the genetic shuffling of exons, DNA regions resembling SAM's such as fibronectin were introduced. The single gene for N-CAM-like adhesion molecules was then itself duplicated, giving rise to all the other related CAM's, including those in the central nervous system such as MAG.

All of these molecules were employed topobiologically in morphogenesis. But in some early vertebrate (or an immediate precursor), a gene in this family was turned to a different purpose altogether. By duplication from a CAM-like piece of DNA, there arose the V and C regions of immunoglobulin genes, receptors on certain lymphocytes and histocompatibility antigens. Thereafter, genes for the V

THE LIFESTYLE RESOURCE™

At The Lifestyle Resource we give you all the facts and details necessary to make an informed purchase. Your satisfaction is our primary concern. If your purchase doesn't meet your expectations, return it in original condition within 30-days for prompt refund.

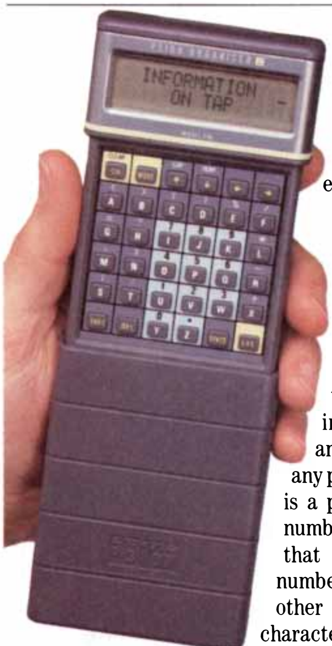
YOUR SMILE IS BEAUTIFUL

But you're only fooling yourself if you think it's proof of healthy teeth and gums. The unseen enemy is a natural mouth substance deposited on the teeth — plaque. Unchecked, plaque develops into rock-hard tartar behind which decay breaks down the tooth enamel and invades the soft, sensitive substructure. You have to remove the plaque before it hardens to prevent the process. Until recently there was no easy solution to this problem, but now there's Interplak, a precision - engineered, university-tested, three-speed plaque removal instrument. Its 10 independently driven bristle tufts rotate around their own axes: each tuft spins in the direction opposite to the spin of its neighbor at 4,200 rpm. The bristles are soft and rounded, adhering to dentistry recommendations, so that virtually every particle of plaque — *98% of it* — is removed gently and without irritation. Interplak even reaches between teeth and around orthodontic braces. Children think it's fun, so they're eager to brush. Interplak is cordless and stores in a free-standing or wall-mounted recharger/holder. Individual color-coded brush heads are available. The Interplak Home Plaque Removal Instrument with charger and two brush heads. **$99.95** #1110. Two additional brush heads **$29.95** #1111.

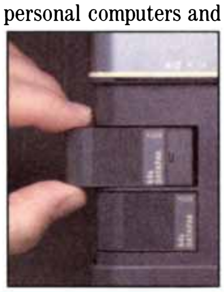
SIT UP STRAIGHT!

Sitting in a conventional chair, bent over a desk or table for long periods of time can stress your spine and back muscles. Posture or "back" chairs have long been accepted as a way to relieve your back from the pain of cramping or kinking up due to this stress. Our posture chair is designed to allow your body to properly align itself by shifting weight from your lower back to your knees, making long work periods far more comfortable. *The Lifestyle Resource* is one of the first to introduce the Criss-Cross Posture Chair. This new and innovative design with a criss-crossed base provides added stability, and an easy exit from your seat. The chair is crafted of laminated hardwood with an attractive oak finish. The seat and kneepad are thickly padded and fully upholstered in a classic neutral grey fabric. The swiveling, carpet-friendly, double wheel casters make maneuvering a breeze. Suited for use at any workstation, desk or table. Requires minimal assembly; a screwdriver is all you need. **$79.95** #2260.

IF YOU FORGET THE WHO, WHAT, WHERE OR WHEN...IT WON'T



Imagine being able to have every address, telephone number, appointment and every fact and figure you need right at your fingertips. The Psion Organiser II replaces your appointment schedule, telephone directory, memo pad, calculator and alarm clock. With 32K of built in software, Psion is ready to use right out of the box and is easily mastered. Most importantly though, you don't have to know anything about computers. All of Psion's functions are accessed from the "menu" which is automatically displayed when you switch on. Unlike other electronic "diaries" and organizers, with its interchangeable Datapaks, Psion *is the only "diary" we know of with unlimited storage capacity!* There is nothing else like it with anywhere near the powers of the Psion. It enables you to carry your desk in a pocket, to put all your important information where it can be easily retrieved. With Psion's amazing cross-referencing, you can find information by using any part of what you stored. For example, if all you can remember is a person's first name, John-and you want to find his phone number, simply enter "John" and Psion will display every entry that contains "John" along with his home and office phone numbers, last name, company, address, his wife's name, or any other information you stored for him. You can enter up to 254 characters in each line and when displayed, the whole line scrolls across the screen. Psion is so powerful you can even find your plumber without knowing his name! The Psion Organiser II is simply the smallest, most sophisticated, hand-held tool of information management yet devised. Psion can be equipped to communicate with personal computers and printers and has two-way data transfer capability. Plug-in Datapaks are available to give you up to a total of 320K of "on-board" memory - *five times more than any other "diary" we know of.* Over 350,000 Organiser II's have been sold worldwide in the past two years and a substantial amount of software is available. The Psion puts the power of information where it will always do the most good - right in the palm of your hand. Operates on a single 9V battery (included) *and does not lose stored information* when you change batteries. Weighs only 9 oz. Measures 5½" x 3" x 1". Model XP 32K Psion Organiser II **$249.95** #2041; 64K Plug-in Datapaks **$89.95** each #2051.



THE LIFESTYLE RESOURCE™ — FOR FASTEST SERVICE — CREDIT CARD ORDERS CALL TOLL-FREE 24 HRS A DAY **800-872-5200**

DEPT. SFAE19; 921 EASTWIND DR. SUITE 114; WESTERVILLE, OH 43081

SEND TO (PLEASE PRINT)

ADDRESS

CITY STATE ZIP

☐ CHECK OR MONEY ORDER ☐ MASTERCARD ☐ VISA ☐ AMEX SIGNATURE

ACCT. # EXP. DATE

ITEM NO.	QTY.	DESCRIPTION	ITEM PRICE	TOTAL PRICE
			SUB TOTAL	
			SHIPPING (See table at left)	
			TOTAL	

Shipping Charge covers UPS, handling and insurance for guaranteed delivery. UPS Second Day available for an additional $7.50 per order.

UPS 2ND DAY AIR

Up to $20 .... $ 3.95	$50.01 to $60 .... $ 7.95
$20.01 to $30 .... $ 4.95	$60.01 to $70 .... $ 8.95
$30.01 to $40 .... $ 5.95	$70.01 to $100 .... $10.95
$40.01 to $50 .... $ 6.95	Over $100 .... $12.95

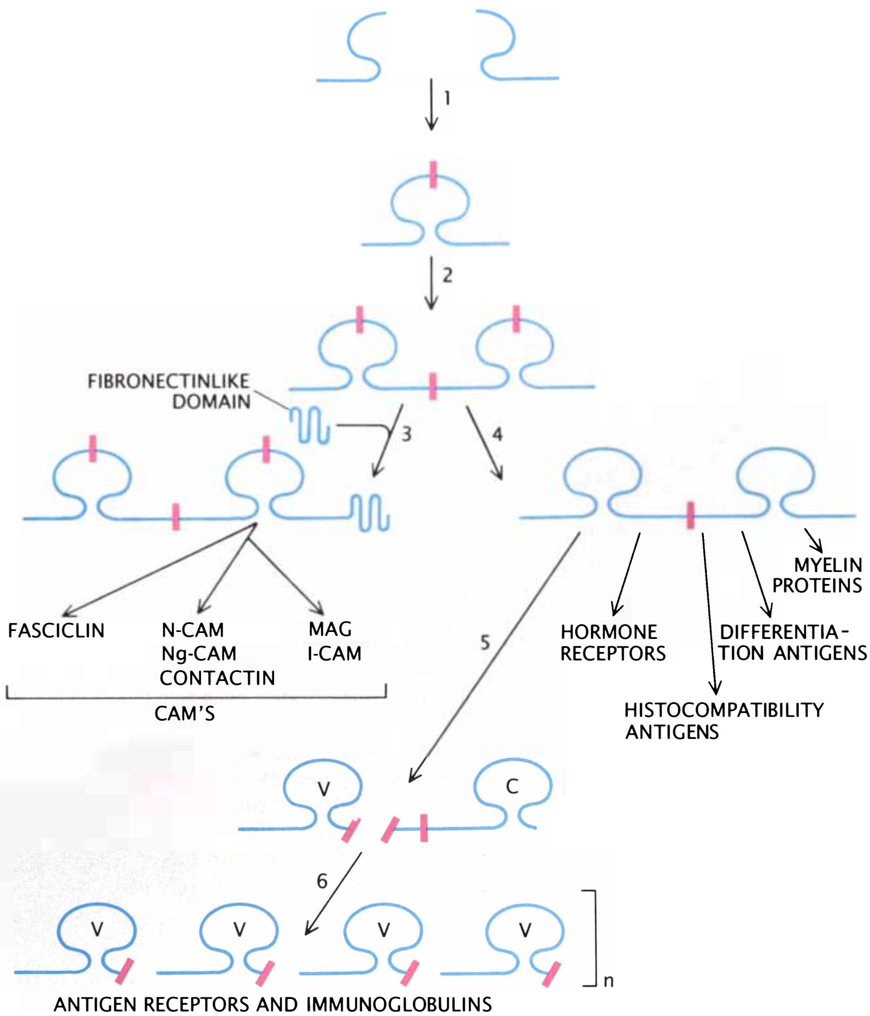
Canadian residents please call (614) 794-2662 for ordering information.

ORDER WITH CONFIDENCE

- *Most orders ship within 48 hours of receipt.*
- *Credit Card orders billed only upon shipment.*
- *No risk 30-day return privilege.*

regions duplicated frequently to yield families that were kept similar by democratic gene conversion. Similar conversion events took place in the histocompatibility family. CAM's such as I-CAM arose to serve as attachment points for LFA-1, the independently evolved receptor on the surface of the lymphocyte.

Thus, the key functions of the adaptive immune system arose from an early cell-adhesion system that operated topobiologically to regulate the cell movements and tissue patterns that give rise to animal form. Although lymphocytes are not dedicated to establishing form, their functions do rely on intricately regulated interactions between cells. That regulation in turn depends on specific binding mechanisms that may hail back to the cell-adhesion origins of the adaptive immune system. A specific type of protein folding is seen at all of the binding sites of immunoglobulins—and probably all those of CAM's as well. This type of folding is called the beta-pleated sheet; two such sheets form a special sandwich in the immunoglobulin domain. The beta-pleated sheet may be part of an ancient "language" of binding and cell regulation that the adaptive immune system has inherited from its CAM-like forebears.



EVOLUTIONARY SCHEME shows how the adaptive immune system arose from early CAM's. Exons encoding two "half-domains" combined to yield the gene for a single domain: a functional unit of a protein (*1*). Gene duplication produced the gene for a multidomain protein, an early N-CAM (*2*). "Exon shuffling" joined a fibronectinlike domain to those for the N-CAM family (*3*). Gene duplication and divergence gave rise to a variety of CAM's. Other genes lost the introns (stretches of noncoding DNA) within the domains but not the ones between domains (*4*). Subsequent divergence in that group gave rise to other members of the immunoglobulin superfamily. Insertion of a genetic element (possibly from a virus) created the potential for protein diversification in different lymphocytes (*5*). Along with this change went an increase in the number of variable (*V*) regions, providing the basis for generation of a vast population of antigen receptors and antibodies during the life of an individual (*6*).

The interaction between antibody and antigen, which can take place in the test tube in a fluid medium without respect to position, is frequently thought of as emblematic of the entire immune system. It must not be forgotten, however, that the immune system has solid tissues (such as the lymph nodes) as well as fluid ones (such as the blood). Even among free, circulating lymphocytes there are populations that "home" to particular bodily tissues. Although immune responses of single cells can occur in the test tube, it is likely that position is as critical in the overall immune response as it is in morphogenesis. Indeed, analysis of the morphoregulatory molecules and the immunoglobulins shows that these two systems are deeply linked both in genetics and in evolution.

It may not be amiss to add a final word about paths of research. The line of thought that links immune system molecules and CAM's would not have been possible within any one area of scientific specialization. It required following a vague surmise through a long trail of research (on CAM's) to the solution of a problem (the origins of immunoglobulins) posed at the end of a previous, apparently unrelated trail of research. This line of thought would have been much impoverished if it had been conceived within the limits of a specialty such as immunology or embryology. Only by seeing biology in its broadest evolutionary, genetic and developmental perspectives while pursuing specialized research can one connect what may first appear to be unrelated matters into a whole that is both organic and intellectually satisfying.

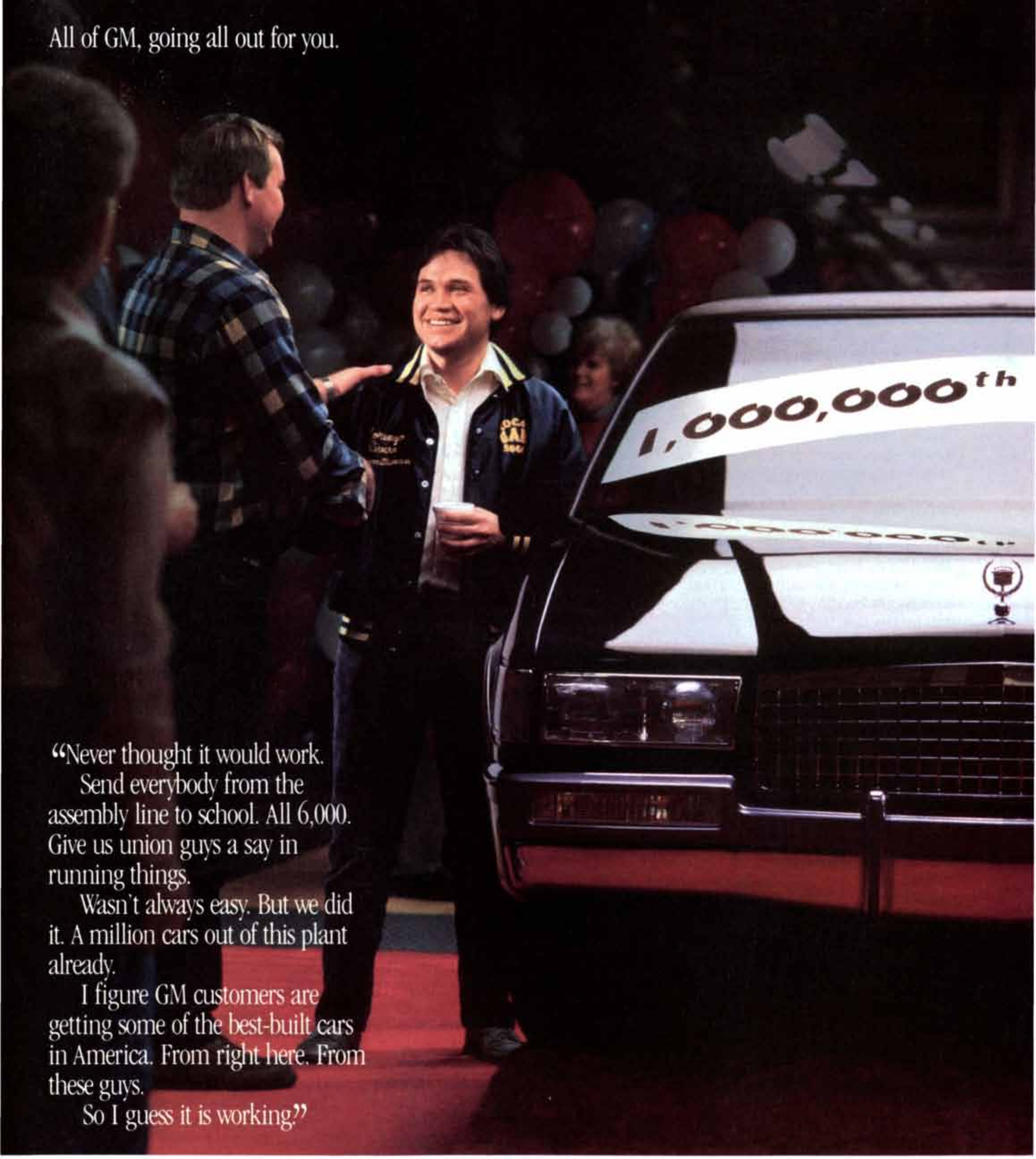
FURTHER READING

ARRANGEMENT AND EVOLUTION OF EUKARYOTIC GENES. G. M. Edelman and J. A. Gally in *The Neurosciences*, edited by Francis O. Schmitt. The Rockefeller University Press, 1970.

CAMS AND IGS: CELL ADHESION AND THE EVOLUTIONARY ORIGINS OF IMMUNITY. Gerald M. Edelman in *Immunological Reviews*, No. 100, pages 11–45; December, 1987.

ASYMMETRIC EXPRESSION IN SOMITES OF CYTOTACTIN AND ITS PROTEOGLYCAN LIGAND IS CORRELATED WITH NEURAL CREST CELL DISTRIBUTION. Seong-Seng Tan, Kathryn L. Crossin, Stanley Hoffman and Gerald M. Edelman in *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 84, No. 22, pages 7977–7981; November, 1987.

TOPOBIOLOGY: AN INTRODUCTION TO MOLECULAR EMBRYOLOGY. Gerald M. Edelman. Basic Books, Inc., 1988.



Fact: The UAW-GM training program is now the largest private educational program in the world.

GM
MARK OF EXCELLENCE

1989 Cadillac Fleetwood at the GM Assembly Center in Orion Township, Michigan

Chevrolet, Pontiac, Oldsmobile, Buick, Cadillac, GMC Truck

The Dynamic Aurora

Interactions of the earth's magnetic field and the solar wind give rise to a vast generator that powers the luminous displays. A similar process may prevail in other astrophysical phenomena

by Syun-Ichi Akasofu

Some early observers of the aurora borealis imagined that these spectacular displays in the Arctic sky resulted from sunlight refracted, rainbowlike, in the atmosphere. The shimmering undulations, they speculated, were caused by the movement of air. Auroral physicists today know that auroras are lights emitted when atoms and molecules in the ionosphere are struck by electrons blowing in from the sun. The apparent motion of the auroral curtain is caused not by atmospheric turbulence but by changes in the electromagnetic conditions that propel the electrons, just as motion on a television screen is an illusion created by changes in the magnetic field that directs electrons from a cathode tube onto the screen.

In the case of the aurora, what serves as the cathode tube? Where is its power supply? Why does that power seem to fluctuate from time to time, causing the aurora to ebb and flow across the polar sky? More than 20 years ago it was determined that auroral emissions occur because the ionosphere is bombarded by electron beams generated by a complex interaction between the solar wind and the earth's magnetic envelope [see "The Aurora," by Syun-Ichi Akasofu; SCIENTIFIC AMERICAN, December, 1965]. The geomagnetic nature of the aurora can be seen clearly from outer space. Centered on each of the geomagnetic poles is a great luminous oval that is a permanent feature of the planet. The one around the North Pole is the aurora borealis. Its counterpart in the south is the aurora australis.

My colleagues and I have now found a numerical relation between the solar wind and the generation of the power that drives the aurora, as well as other disturbances of the geomagnetic field. We are also gaining a better idea of how the sun's activity affects the solar wind. This progress raises the interesting possibility of developing a computer method to predict the intensity of auroral activity. Because auroras can interfere significantly with radio and satellite communications, as well as with power lines and some defense systems, the ability to predict auroral intensity has become increasingly important as human activity expands into polar regions and into space.

SYUN-ICHI AKASOFU, internationally recognized for his pioneering work in auroral physics, is director of the Geophysical Institute at the University of Alaska, Fairbanks. He was born in Japan and got a B.S. and an M.S. at Tohoku University. He received a Ph.D. in geophysics from the University of Alaska in 1961 and has been professor of physics there since 1964.

Generator in the Magnetosphere

The understanding of how auroras are generated came about through a revolution in scientists' knowledge of the earth's magnetic environment. The earth's magnetic field had long been assumed to be essentially a dipole field, like that of a bar magnet, in which magnetic field lines loop from the south pole to the north pole symmetrically about the geomagnetic axis. The earth does not exist in a vacuum, however. It is constantly buffeted by the solar wind, a dilute plasma of hydrogen ions (protons) and electrons streaming from the sun's corona.

The solar wind confines the earth's field inside a comet-shaped volume called the magnetosphere. On the sunward side the solar wind compresses the magnetosphere to a distance of about 10 earth radii. On the downwind side the solar wind confines the magnetosphere within a windsock-like structure known as the magnetotail, which stretches for more than 1,000 earth radii. The magnetosphere is filled with tenuous plasmas of different densities and temperatures, which originate from the solar wind and the ionosphere [see "The Earth's Magnetotail," by Edward W. Hones, Jr.; SCIENTIFIC AMERICAN, March, 1986].

In the early 1960's solar physicists began to recognize that the solar wind can stretch and carry the coronal magnetic field out into the far reaches of the solar system. This extension of the sun's magnetic field is called the interplanetary magnetic field. James W. Dungey of the Imperial College of Science and Technology in London proposed that this magnetic field could join with geomagnetic field lines originating in the polar region of the earth. This phenomenon, called magnetic reconnection, proceeds most efficiently when the solar-wind magnetic field is oriented southward—that is, antiparallel to the earth's field.

It had been widely assumed that reconnection would be a stable process, but a decade ago Christopher T. Russell of the University of California at Los Angeles found that it is not. Instead, "bundles" or "ropes" of field lines form and eventually cleave from the magnetosphere and are swept into the magnetotail. The instabilities occur even if the interplanetary magnetic field has a steady southward orientation. In reality the field constantly changes its strength and direction. The resulting complex interactions have been elucidated in a computer simulation by Lou-Chuang Lee of the University of Alaska at Fairbanks.

As solar-wind particles flow along the magnetosphere's boundary (the magnetopause) they cross the reconnected magnetic field lines. The ions and electrons, being of opposite electric charge, are deflected in opposite directions (according to the familiar right-hand rule), thereby generating an electric current [*see illustration at right on page 93*]. This process is the same as magnetohydrodynamic power generation. Indeed, the entire magnetopause constitutes a giant generator that converts the kinetic energy of solar-wind particles into electric energy, producing more than a million megawatts of power. The mechanism is called the solar-wind-magnetosphere generator, or auroral generator.

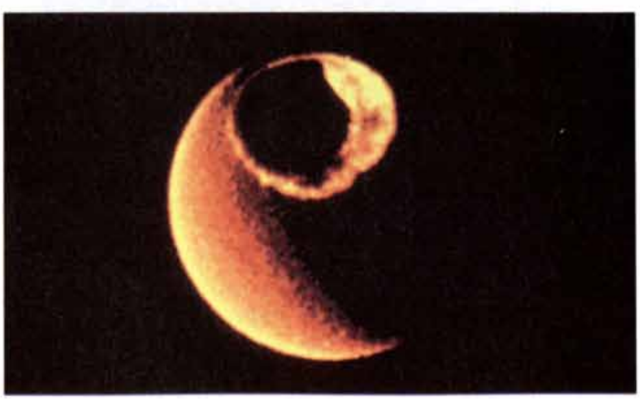
Field-Aligned Currents

The generator mechanism drives positive ions to the dawn side of the equatorial plane of the magnetopause, creating a kind of positive terminal; the electrons are driven to the dusk side, or negative terminal. In rarefied plasmas permeated by magnetic field lines, such as those that fill the magnetosphere, electrons travel along corkscrew trajectories that wind around the field lines. The process by which these so-called field-aligned currents are generated has been worked out in detail by Akira Hasegawa of the AT&T Bell Laboratories and others. Thomas A. Potemra of Johns Hopkins University, Takeshi Iijima of the University of Tokyo and others have confirmed the existence of the currents using satellite-borne magnetometers.

The field-aligned currents connect the magnetopause to the ionosphere. The dawn side of the magnetopause is thereby projected to the dawn half of the auroral oval, and the dusk side of the magnetopause is projected to the dusk half of the oval. The dawn side of the oval, then, becomes electrically positive and the dusk side becomes electrically negative. The resulting potential drop, or voltage difference, across the oval is about 100 kilovolts.

The field-aligned currents are carried by electrons rather than protons because electrons are more mobile. In the scenario described above, electron currents flow downward to the dusk half of the oval and flow upward from the dawn half. Since auroras result when electrons crash down into the ionosphere, one might well ask: Why are there emissions on the dawn side of the oval? It turns out that the oval in the ionosphere is highly conductive and so current flows between the inner and outer edges of the oval. The current then flows back up along the magnetic field lines, resulting in a secondary current that points in the opposite direction from the primary one. Secondary electron currents result in auroral emissions on the dawn side.

The auroral lights occur when in-



WESTWARD-TRAVELING SURGE along an active aurora appears in the evening sky above Fairbanks, Alaska. The whitish light is emitted by oxygen atoms. Ionized nitrogen molecules give off the pink light near the lower fringe of the curtain. At the left is a false-color image of the auroral oval above the North Pole. The image was made by the *Dynamics Explorer* satellite from a distance of three earth radii. It records emissions from oxygen atoms at a wavelength of 130 nanometers. The bright crescent on the left is the day side of the earth. The image was provided by Louis Frank of the University of Iowa.

coming electron beams collide energetically with the ionosphere, exciting or ionizing atoms and smashing apart molecules to create other excited atoms. The excited and ionized atoms emit radiation over a wide spectral range (from extreme ultraviolet to infrared) as excited atoms jump down to lower energy states and as ions combine with free electrons.

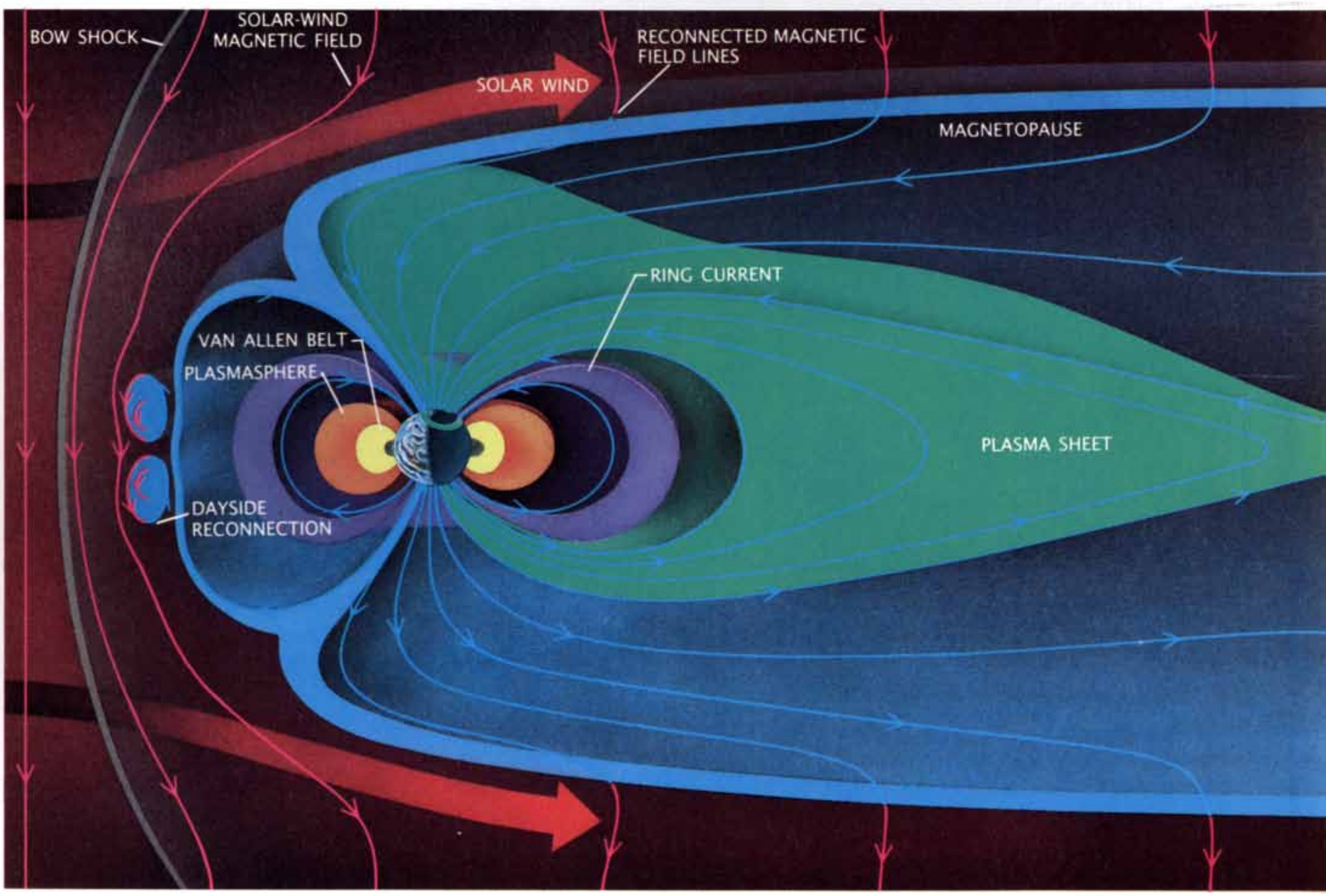
The commonest auroral emission is a whitish-green light with a wavelength of 557.7 nanometers, which is emitted by oxygen atoms. A beautiful pink emission comes from excited molecules of nitrogen. Various ionospheric atoms and molecules produce auroral emissions in the extreme ultraviolet, ultraviolet and infrared wavelengths, but these cannot be observed on the ground because they are absorbed by the intervening atmosphere. Images made by the Swedish *Viking* satellite of extreme ultraviolet emissions show that the aurora is surprisingly active on the sunlit side, often more active than on the dark side.

Why the Curtain Shape?

From the ground the aurora appears to be a curtain of light, streaked with rays. The curtain begins at an altitude of several hundred kilometers and ends about 100 kilometers above the ground, where the atmosphere becomes so dense that it stops most of the incoming electrons. The sheet is less than a kilometer thick, whereas it extends laterally for thousands of kilometers. What could account for this?

It is in fact surprising that field-aligned electron currents are able to reach the lower ionosphere at all. Like other charged particles in the Van Allen radiation belts, the helical trajectory of an electron tends to increase in pitch as the electron nears the earth (where the field is stronger). Indeed, its motion becomes completely circular at heights well above the ionosphere. At that point the electron is reflected back upward. Yet auroras testify to the fact that electrons are able to penetrate deep into the ionosphere.

The process appears to begin when the electrons in the field-aligned currents form thin, sheetlike beams. In a way that is not yet understood, when enough power is pumped into the magnetosphere, and when the sheet beams reach a sufficiently high intensity, a peculiar electric field called an auroral potential structure develops around the beams at an altitude of between 10,000 and 20,000 kilometers. The region inside the structure appears to separate into positively and negatively charged layers, producing a strong electric field between them [*see top illustration on page 94*].



SOLAR WIND, a diffuse plasma of protons and electrons streaming from the sun, confines the earth's magnetic field in a comet-shaped cavity called the magnetosphere. The wind compresses the magnetosphere on the day side to a distance of about 10 earth radii. On the night side the wind sweeps the earth's magnetic field into an elongated volume known as the magnetotail, which extends for at least 1,000 earth radii. The boundary of the magnetotail is called the magnetopause. The solar wind has a magnetic field (*red*). When it is directed southward, as is shown here, it can "reconnect" efficiently with the earth's field (*blue*). Solar-wind particles flow into the magnetosphere along the reconnected field lines. Magnetic field

Hannes Alfvén at the University of California at San Diego first suggested the presence of such a "double layer" above the aurora.

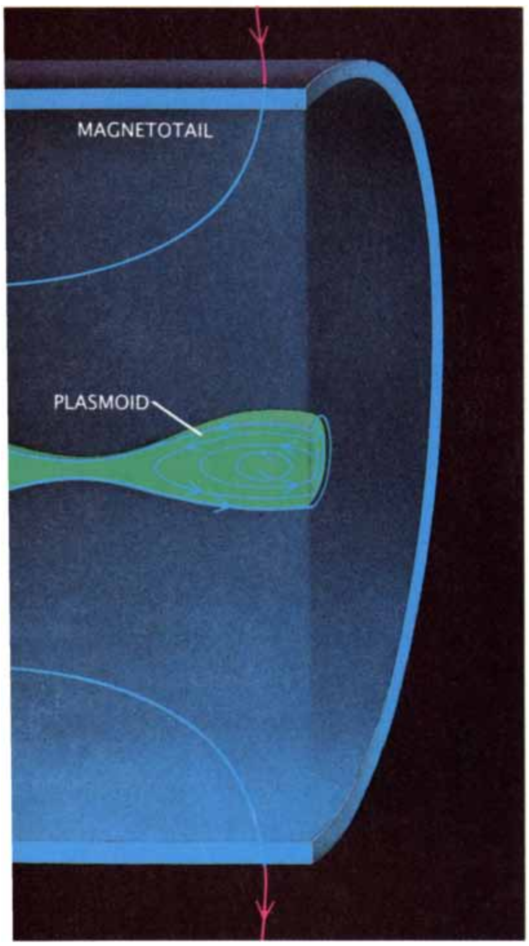
A structure resembling a double layer is known to form at the surface of the electrode in a neon light, but the exact nature of the one in the earth's upper atmosphere is at present controversial. Electrons seem to be accelerated downward by the electric field that accompanies the double layer. The electrons acquire a few thousand electron volts of energy by the time they reach the bottom of the auroral potential structure—enough to penetrate to an altitude where the atmosphere is dense enough to make auroral emissions visible to the unaided eye. The flat shape of the potential structure is related to the aurora's curtainlike form.

The rays shooting through the auroral curtain are actually a series of vortexes that form when antiparallel electric fields (associated with the auroral potential structure) along the face of the curtain cause electrons to flow in opposite directions, just as whirlpools form at the boundary between oppositely flowing currents of water. High-speed television cameras pointing upward at the bottom of the curtain have captured images of such vortexes, and remarkably similar structures have been generated in computer models of vortex formation.

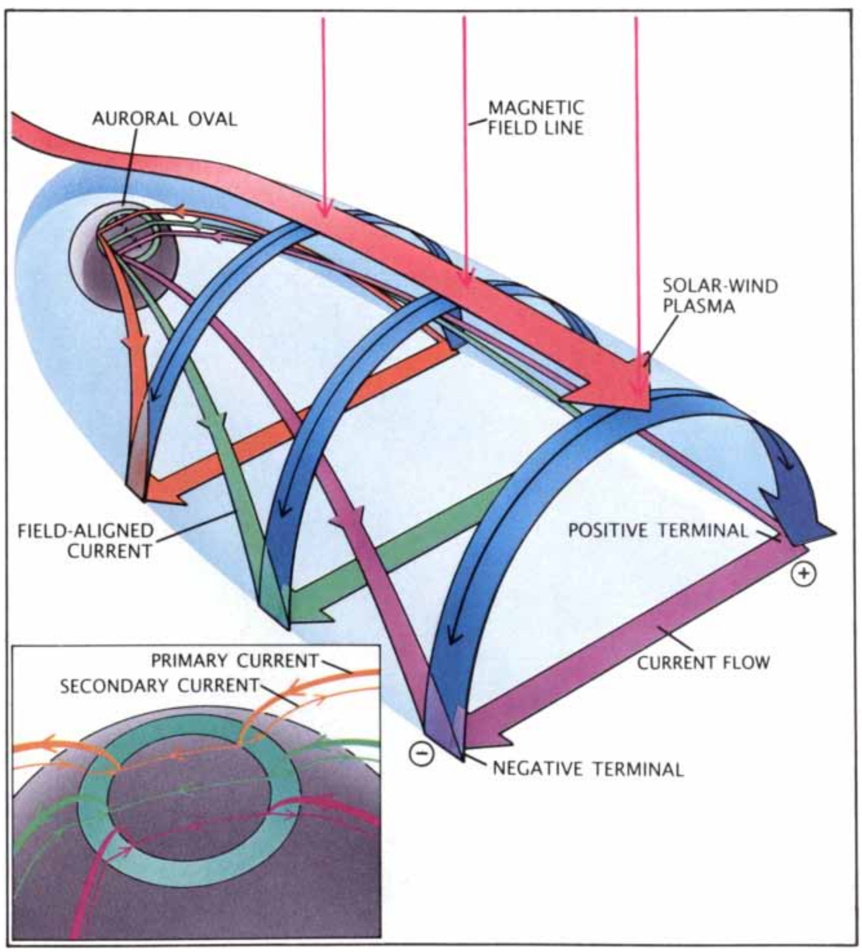
The energy of auroral electrons in relation to the spatial distribution of the potential structure has been studied extensively with instruments on board rockets and satellites by Louis A. Frank of the University of Iowa, James L. Burch of the Southwest Research Institute, Patricia H. Reiff of Rice University and Bengt Hultgvist and his collaborators on the Swedish Viking satellite program. Investigators at the Lockheed Missiles and Space Company have confirmed that the auroral potential structure also accelerates positive ions upward; indeed, such ions become at times a significant part of magnetospheric plasma.

Interactions of particles and electromagnetic waves in a plasma also occur in the auroral potential structure, generating intense radio waves. Donald A. Gurnett of Iowa noted that these emissions are so intense that an extraterrestrial intelligence system would detect them well before it could visibly detect the earth. These emissions do not interfere with radio broadcasts on the earth, because the ionosphere reflects them upward, just as it reflects earth-based emissions groundward.

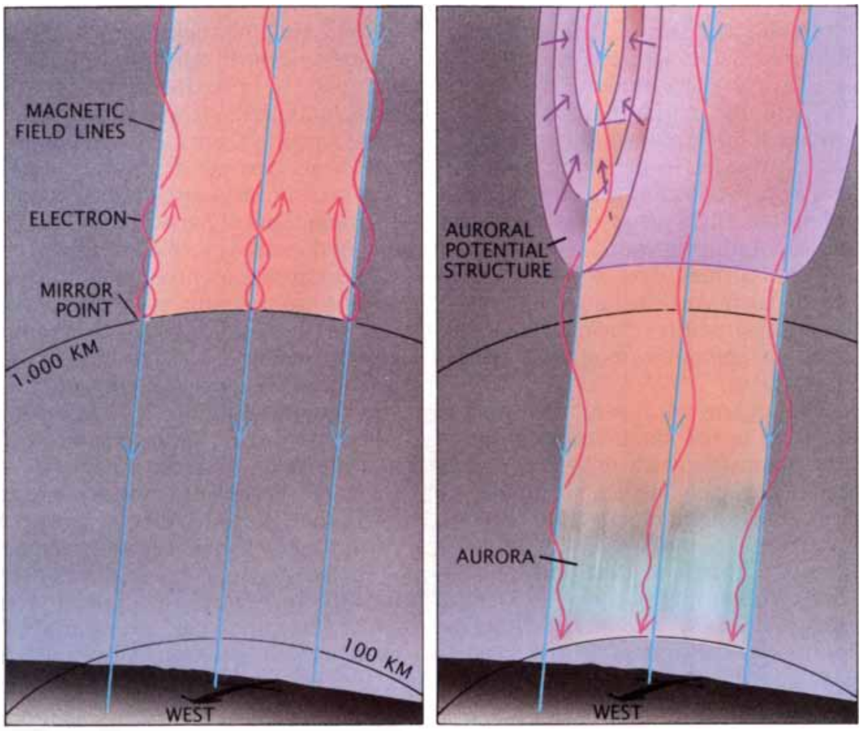
The existence of the auroral potential structure should be of great interest to astrophysicists and solar physicists, as well as to auroral physicists,



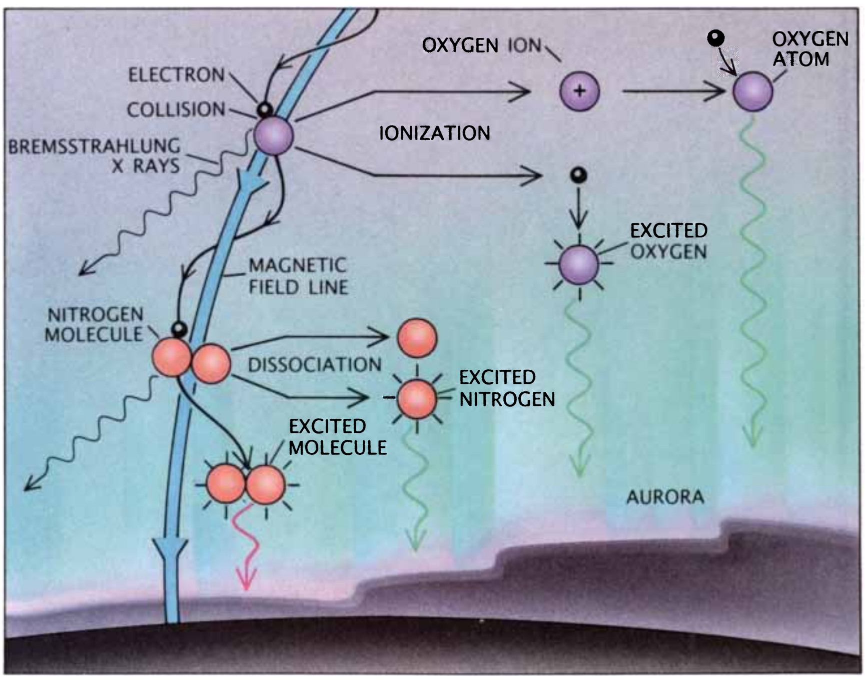
lines in the north lobe of the magnetotail point toward the earth; those in the south point away. Reconnection of field lines in the magnetotail can pinch off clumps of plasma known as plasmoids, which are ejected from the magnetotail.



AS SOLAR-WIND PLASMA flows across magnetic field lines in the magnetopause, protons are deflected to the dawn side of the magnetotail and electrons to the dusk side. Current flows between the two regions, most of it directly across the magnetotail but some of it along magnetic field lines to and from a vast oval in the ionosphere. This field-aligned current is carried by electrons, which excite the auroral emissions. Secondary currents develop along paths that are parallel to the primary currents (*inset*).



FIELD-ALIGNED ELECTRONS trace a helical path around magnetic field lines. As they descend toward the ionosphere their pitch angle increases until they are deflected back upward (*left*). In certain situations (*right*) a peculiar electric field called an auroral potential structure develops around the electron currents. Electrons accelerated downward by the potential structure are able to penetrate deeper into the ionosphere. The potential structure is very thin from north to south but extends from east to west for thousands of kilometers, giving the aurora its curtain shape.



ATOMS AND MOLECULES in the lower ionosphere emit radiation when they are struck by electrons that have been accelerated by the auroral potential structure. Electrons slowed by collisions emit bremsstrahlung X rays. Collisions break up molecules into excited atoms that emit radiation as they fall to a lower energy state. Electrons knocked off by the collisions strike and excite atoms, which then emit radiation. The electrons also ionize atoms, which emit radiation on recombining with electrons.

because it is still widely believed that a significant electric field cannot be maintained along magnetic field lines in a rarefied plasma, and therefore charged particles cannot be accelerated in this manner. The observation of such a field in connection with auroras suggests that similar fields may exist in astrophysical conditions such as those in the Crab nebula, which has a plasma density and energy like that of the magnetosphere.

Electrojets and Substorms

The largest injection of energy into the ionosphere by the auroral generator is through a pair of electric currents, called westward and eastward electrojets. These currents are driven along the auroral oval in the lower ionosphere and produce intense heating that contributes to the generation of large-scale winds in the polar upper atmosphere.

How do the electrojets arise? Bear in mind that electric currents flow between the outer and inner edges of the oval. The currents are induced by an electric field that is parallel to the earth's surface and therefore perpendicular to the earth's magnetic field at the poles [*see top illustration on opposite page*]. In the presence of these mutually perpendicular fields, charged particles undergo something called E-cross-B drift, according to which both positive and negative particles drift in the same direction, from the night side to the day side.

In the upper ionosphere the drifting particles travel at the same speed and so have no net current, but they do impart momentum to neutral particles, contributing further to wind generation. In the lower ionosphere, however, the protons collide too frequently with neutral particles and become filtered out of the E-cross-B drift. As a result only electrons can flow along the auroral oval. Hence there is a net electric current pointing eastward in the evening sector and westward in the morning sector.

In the International Geophysical Year (1957–1958) observations by camera of the entire sky revealed a hitherto unknown systematic aspect of auroral activity over the polar region: the auroral substorm. The first indication of a substorm is a sudden brightening of the auroral curtain from late evening to midnight. This brightening spreads rapidly along the curtain in both directions, so that in a matter of several minutes the entire section of the curtain in the dark hemisphere becomes bright.

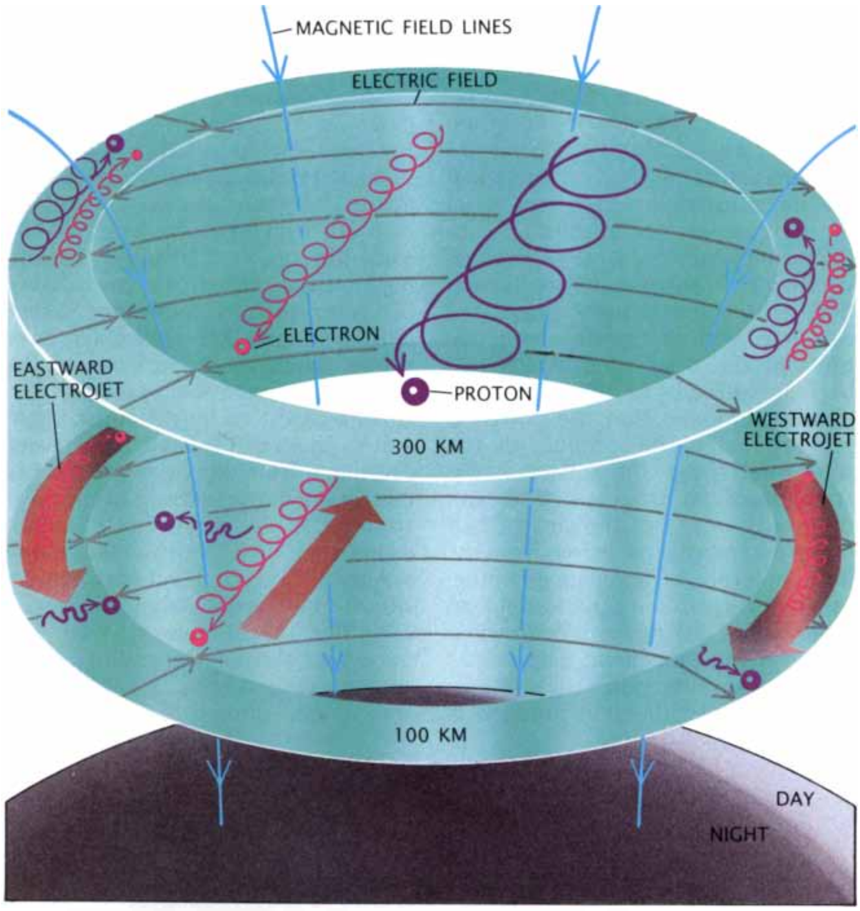
The bright curtain begins to move poleward in the midnight sector at a speed of a few hundred meters per second. At the same time a large-scale bulging structure is generated there. Various wavy motions appear close to the western end of the bulge and propagate westward at a speed of one kilometer per second. Halfway around the pole, in the morning sector, the auroral curtains disintegrate into many "patches." The poleward motion in the midnight sector typically lasts from about 30 minutes to an hour. After this poleward advance reaches its highest latitude the auroral activity begins to subside. The substorm typically lasts between one and three hours.

The auroral substorm is a manifestation of what is called a magnetospheric substorm, about four or five of which occur every day. Other manifestations include greatly enhanced electrojets, which in turn cause intense geomagnetic disturbances known as polar magnetic substorms. One such storm on March 18, 1978, was recorded in detail in an international effort that involved setting up more than 70 magnetometers in the Arctic region, arranged along six "spokes" radiating from the magnetic north pole. Powerful computer codes enabled Yosuke Kamide of Kyoto Sangyo University and Yasha I. Feldstein of the Academy of Sciences of the U.S.S.R. to reconstruct the electric current pattern from the magnetic records. They were also able to estimate the heat-production rate associated with these ionospheric currents.

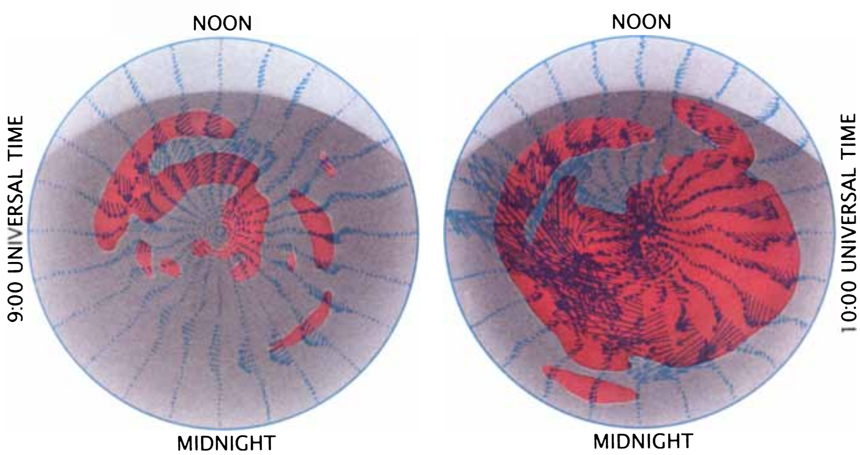
Substorm Dynamics

Many investigators, including Robert McPherron of the University of California at Los Angeles and Daniel Baker of the National Aeronautics and Space Administration Goddard Space Flight Center, speculate that substorms are driven by processes within the magnetotail. The generator process drives two loop currents that flow in opposite directions around the northern and southern lobes of the magnetotail. The circulating currents induce antiparallel magnetic fields along the length of the tail. It is generally thought such antiparallel fields would spontaneously and explosively reconnect, releasing the energy to drive the substorms.

It has become increasingly clear, however, that the growth and decay of magnetospheric substorms is partly controlled by the rise and fall of the power produced by the generator



FIELD-ALIGNED CURRENTS induce electric fields across the auroral oval that are perpendicular to the geomagnetic field. In such a situation electrons and protons are driven in the same direction, in what is known as the E-cross-B drift. The drift over the polar cap is in the direction opposite to that along the oval because the electric fields point in opposite directions. (The particles follow spiral paths because the Lorentz force causes them to move in circles around the magnetic field lines.) In the upper ionosphere both electrons and protons proceed at the same speed and there is no net current. In the lower ionosphere the protons undergo many collisions and move a small amount in the direction of the electric field rather than of the E-cross-B drift. Hence, there is a net current of electrons, which produces the electrojets.



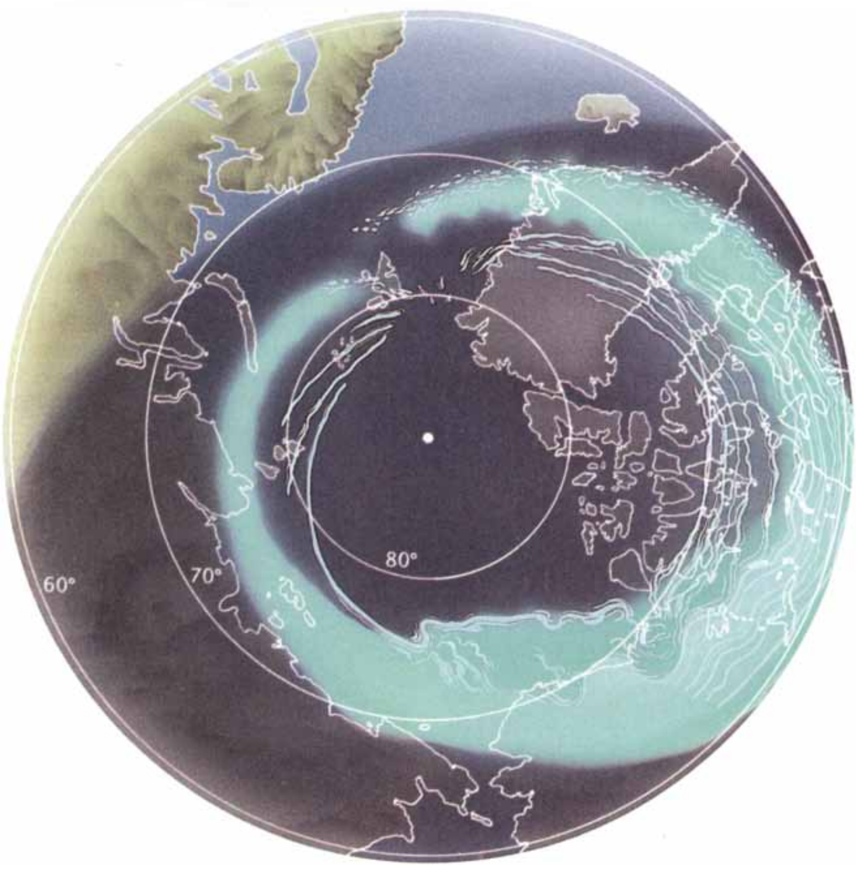
IONOSPHERIC CURRENTS recorded on March 18, 1978, during a substorm (*right*), are much larger than those recorded just one hour earlier (*left*). The activity is shown from above the magnetic north pole, with the day side at the top and the night side at the bottom. The arrows point in the direction of the currents, and their length indicates the strength of the current. During the substorm a westward electrojet was observed in the dark sector and an eastward electrojet in the afternoon sector. The amount of heat generated by the electrical activity is indicated by shades of red.

process. In the early 1970's my graduate student Paul Perreault and I set out to try to relate the generator power to various characteristics of the solar wind. We assumed that the actual measured total energy dissipated in the inner magnetosphere was equal to the rate of power injected into the magnetosphere by the solar wind. We then tried to see whether fluctuations in the dissipated energy were correlated with changes in certain features of the solar wind, which had been measured by satellites.

We determined that the power is proportional to the product of the solar-wind speed, the square of the strength of its magnetic field and the fourth power of the sine of half of the polar angle (measured from the north pole) at which this field impinges on the earth's magnetic field. In other words, the power is zero when the solar-wind magnetic field points north, because the polar angle is zero degrees and so the sine function is zero. Conversely, the power is at a peak when the field points south (everything else being equal), because the polar angle is 180 degrees and the sine function reaches its maximum value.

Mikhail Pudovkin and his colleagues at the University of Leningrad, along with other theorists, have derived the same formula theoretically by assuming that the magnetosphere behaves like a magnetohydrodynamic generator. Reiff and her colleagues have shown that measured variations in the potential drop across the polar cap (the voltage produced by the auroral generator) are closely related to the power calculated from our equation. What is more, each major increase of the power above 10,000 megawatts was associated not only with an increase of the potential drop but also with an intensification of auroral substorms.

We find that the auroral oval contracts and enlarges depending on the amount of power supplied by the auroral generator, which is in turn a function of the north-south component of the interplanetary magnetic field. Such observations indicate that substorms occur most often when the solar-wind magnetic field vector turns southward—persuasive evidence that substorms are closely controlled by the solar wind rather than by spontaneous events within the magnetosphere.



DETAIL OF AURORAL OVAL is depicted in this view from above the geographic north pole. A diffuse aurora sweeps out a smooth belt from the evening sector to the midnight sector. It has a fairly uniform luminosity in the evening sector and is often wide enough to fill the entire field of view of an observer on the ground. Discrete, curtainlike auroras are distributed along the poleward side of the diffuse aurora. At the height of a substorm bright forms advance toward the pole at about midnight, as is shown here. Toward morning the diffuse portion appears to disintegrate into numerous curtainlike structures and patches on the outer perimeter of the oval.

Some Remaining Questions

It is not yet well understood what triggers the substorms. In particular an increase in generator power by itself cannot account for the sudden brightening of the auroral curtain. Joseph R. Kan of the University of Alaska at Fairbanks theorizes that the brightening occurs because field-aligned currents become amplified by effects that arise in the ionosphere when the E-cross-B drift is enhanced by a surge in generator power.

What happens when, instead of pointing southward, the magnetic field is large and points northward for an extended period? As the power of the auroral generator decreases, the aurora becomes dim and the electrojets become weak. An unexpected auroral phenomenon takes place under these conditions. Several auroral curtains stretch across the polar cap in a direction parallel to the noon-midnight meridian, and patchy subvisual auroras drift across the polar cap in the same direction. These auroras cannot be understood simply in terms of the decreasing power of the generator and are beyond the scope of this article.

The model of auroral power generation that I have described so far concerns magnetic reconnection on the sunward side of the magnetosphere. Reconnection is expected to occur in the magnetotail as well, in a process that is also driven and partly controlled by the generator. There is now evidence, gathered by James A. Slavin and Bruce T. Tsurutani of the Jet Propulsion Laboratory, that the speed of the plasma flow away from the earth at a distance downstream of some 200 earth radii correlates with the intensity of the auroral electrojets. The downstream flow is thought to be propelled by the energy released by magnetic reconnection in the magnetotail.

As the aurora brightens, plasma sheets in the magnetotail first become very thin and a short time later give rise to a variety of activities in the magnetotail. These are expected to be in some way related to auroral dynamics during substorms. The International Solar-Terrestrial Physics project, a major multisatellite effort now in the final planning stage, is designed to find answers to these questions.

Beyond Auroras

The power of the auroral generator is ultimately controlled by activity in the sun, in particular by such transient events as solar flares and massive ejections of coronal gases, which generate shock waves that propagate in the solar wind. Just behind the wavefront the solar wind reaches a speed of from 500 to 1,000 kilometers per second, and the magnetic field is compressed and therefore increased in magnitude. As the shock wave collides with the magnetosphere, the generator power can surge to 10 million megawatts or more, provided the magnetic field is directed southward.

This situation can give rise to a geomagnetic storm. During such a storm the auroral oval expands abnormally. As the oval spreads southward it may disappear from the Alaskan sky and appear at the latitude of the border of the U.S. and Canada, or even lower. The 630-nanometer red emission from oxygen atoms is greatly enhanced in such auroras, perhaps because the increased power thermally excites the oxygen atoms to higher energy states. At the same time, intensified currents in the Van Allen belts create large magnetic fields even in the lower latitudes and on the ground.

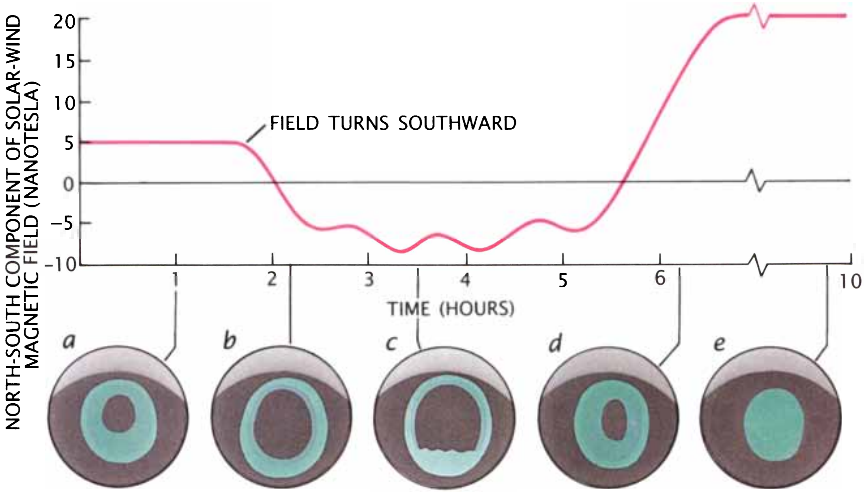
Recently my colleagues in solar physics and I have been trying to understand the effect of solar shock waves on the magnetosphere. A solar flare at the center of the solar disk would generate a shock wave that propagates toward the earth along the sun-earth line. In this situation the wavefront would collide nearly head on with the nose of the magnetosphere; the interplanetary magnetic field would undergo a large compression and would therefore increase in magnitude. This in turn would greatly increase the power of the auroral generator. If a flare occurs near the limb of the solar disk, its shock wave would propagate in a direction perpendicular to the sun-earth line and so would glance off the nose of the magnetosphere, compressing it by an insignificant amount. In this situation even an intense flare may not cause a major auroral display.

Another solar phenomenon that affects auroral activity are "holes" in the corona—regions devoid of sunspots. Such holes generate high-speed torrents of solar wind. They are particularly well developed during the declining epoch of a sunspot cycle. Often two large holes appear simultaneously, one extending from the sun's north pole and the other extending from the south pole. Each hole expels solar wind in a broad stream.

Since the sun rotates with a period of about 27 days with respect to the earth, there is a rotating-sprinkler effect: one stream arrives at the earth and is followed two weeks later by the other. The earth is immersed in each stream for about a week. During this time the generator power is high and variable. The coronal holes tend to last for many months, and so in the declining epoch of a sunspot cycle there will be two one-week intervals of auroral activity every 27 days over many months. The solar wind appears to be ejected at a greater speed from higher latitudes of the coronal hole; this may be one reason auroral activity intensifies in the spring and fall equinoctial months, when the earth is at the highest heliographic latitudes.

The key to understanding certain solar events may in fact lie in the auroras themselves. Like auroras, solar flares are caused by atmospheric emissions of excited atoms and appear as bright, curtainlike formations, and they are no doubt manifestations of similar processes. It has long been suggested that the energy for flares is supplied by magnetic reconnection in the sun's magnetic field. The theory requires the existence of so-called force-free fields, which are identical to the field associated with field-aligned currents in the earth's magnetosphere. It is essential to find a generator mechanism that can supply the electric power for force-free fields and solar flares. Perhaps gaseous motions in the sun's visible surface, which resembles the earth's ionosphere, generate the necessary power.

It is my hope that the study of auroras will contribute to the understanding of a diversity of astrophysical phenomena. After all, rarefied plasmas permeated by magnetic fields are intrinsic to most astronomical objects; the interactions of magnetized plasma flows and the atmosphere of such magnetized celestial bodies as stars, planets and comets may be quite common. Of all such interactions, only the aurora provides a ready laboratory in which scientists can test and confirm theories through direct observation.



WHEN SOLAR WIND'S magnetic field vector points northward, the auroral oval is small and mostly filled by a subvisual glow (*a*). As the field turns southward the oval brightens and expands rapidly; the subvisual glow disappears except along a narrow belt on the inside of the oval (*b*). About an hour later a substorm begins, and bright curtains advance toward the pole. The substorm reaches its peak one or two hours after its onset (*c*). After the field vector tips northward again, the aurora becomes dim, and curtains parallel to the noon-midnight meridian appear across the polar cap (*d*). When the field vector has had a large northward component for many hours, the oval rim may vanish, leaving a glow over the entire polar region (*e*).

FURTHER READING

SOLAR-TERRESTRIAL PHYSICS. Syun-Ichi Akasofu and Sydney Chapman. Oxford University Press, 1972.

AURORA. A. Vallance Jones. D. Reidel Publishing Co., 1974.

MAJESTIC LIGHTS. Robert H. Eather. American Geophysical Union, 1980.

THE NORTHERN LIGHTS: FROM MYTHOLOGY TO SPACE RESEARCH. Asgeir Brekke and Alf Egeland. Springer-Verlag, 1983.

THE SOLAR WIND AND THE EARTH. Syun-Ichi Akasofu and Yosuke Kamide. D. Reidel Publishing Co., 1987.

Technology Transfer

Research consortiums look for ways to work with backers

Fertile collaboration, high-tech houses, truffle technology

Just north of downtown Austin lies the Microelectronics and Computer Technology Corporation (MCC), which opened shop in 1983 as one of the earliest industrial research consortiums in the U.S. Five years later the Semiconductor Manufacturing Technology (SEMATECH) consortium, a joint industry-Government research project, took up residence some 25 miles south of MCC. They are not alone: scattered across the U.S. are more than 75 research consortiums, in which companies pool resources to investigate everything from cement to semiconductors.

Research consortiums have become the corporate fad of the 1980's. Their organizational details vary. Some projects are funded by local governments eager to attract industrial investment to the area; others aim to team companies and universities. For the universities these projects may mean more money for research, better equipment and added appeal for top-ranking graduate students. For the companies such collaboration can keep them abreast of basic research in more areas than they could manage in-house; it also gives them access to a pool of well-trained investigators.

Nevertheless, the consortiums that have attracted the most attention and funding are ones such as MCC and SEMATECH, which are meant to keep U.S. industries competitive by developing generic technologies that member companies can hone for their specific purposes. Government and industry have pinned high hopes on these consortiums, hoping they will invigorate flagging industries or boost U.S. efforts in fledgling technologies such as high-definition television or superconductors.

As consortiums mature, however, it is becoming obvious they cannot simply be research laboratories that member firms dip into and pull out new technology. Nor can consortiums set the research directions of member companies. Instead, consortiums must find ways to marry their corporate sponsors, so that consortium and sponsor can then jointly conceive and raise ideas into prototypes.

"The science is the easy part," says Grant A. Dove, MCC's chief executive. The difficulties cut two ways, Dove observes: motivating a corporate sponsor to plunge ahead with a new technology and instilling in researchers a better appreciation of the needs and concerns of the sponsor. "The creation of new technology—across the board—happens in the U.S. at a pace not found anywhere else," says Admiral Bobby R. Inman, MCC's first chief executive. It is in the race to commercialize or even apply new technology that the U.S. loses ground, he adds.

Most consortiums, intended by their founders to be fountains of new technology, spent little time wondering about how to transfer technology as they established their research agendas. Typically "technology transfer" meant that as project developers reached a milestone they summarized the work and turned it over to a recipient, says Dan P. Dimancescu, cofounder of the Technology & Strategy Group consulting firm in Cambridge, Mass. "It was a very linear process," he says.

"We once thought that one could have technology transferred through a receptor or a designated structure," such as one person or a series of documents, says Eugene Lowenthal, director of MCC's advanced computer architecture program.

Sponsoring firms, meanwhile, saw consortiums as places where a one-dollar investment would buy many dollars of research. Now they are realizing that gaining the full benefits of their participation requires a much greater commitment than simply earmarking dollars. "Those companies that just invest in a consortium and watch will find that nothing much happens," says Matthew Kuhn, who spent several years as the liaison between his employer, Bell Northern Research, and the Microelectronics Center of North Carolina. (He was recently appointed president of the center.)

What both consortiums and sponsors have learned, Dimancescu says, is that technology transfer means developing a project jointly. Technology transfer is evolving into a continuous process in which workers who are primarily developing the technology and those who intend to use it toss ideas back and forth, Dimancescu says.

MCC learned that lesson the hard way in its design-automation project. Workers knew that the sponsoring companies were eager to have new computer-aided-design (CAD) tools. MCC began developing algorithms in a programming language called LISP. What the members had hoped for, however, were fully completed tools, not research algorithms that they would have to adapt. "We handed them new algorithms, and it was like feeding grass to a tiger," says John T. Pinkston, MCC's chief scientist. The consortium's CAD program was reorganized last year.

The value of collaborative work was driven home during the development of an expert system for creating integrated-circuit designs. The project, called Proteus, caught the eye of a senior consultant to the NCR Corp. He recommended the company assign an engineer to work on the project. For a year the NCR engineer spent one week every month with MCC's Proteus team. She took the software the team developed back to her research group in Fort Collins, Colo., where she and other NCR engineers were building an application. Then she flew back to MCC with questions and suggestions about how to make Proteus fit NCR's needs better. The time was well spent; NCR's application of Proteus, called the Design Advisor, became the first MCC technology to reach the market.

Consortiums such as MCC also discovered that they need to be as sensitive as any in-house worker to the pressures on corporations. For instance, regardless of the merit of a technology, piling additional costs on a project will kill the interest of the most enthusiastic sponsor. When one company found it had to spend at least two dollars at home for every one it spent at MCC, it quit the program. "It isn't as simple as saying, 'we spent a million dollars, so we're going to use the technology,'" says James D. Babcock, a manager in MCC's software technology group. Lowenthal recollect-

ed one software project that would have benefited only a company with a parallel-processing computer; member companies balked at buying the expensive hardware. MCC workers now try to develop systems that are not likely to be so expensive to use. "We're moving some software we developed on expensive workstations to personal computers," Babcock says.

Companies in turn are devoting more people in-house to the task of working with outside projects. Digital Electronics, which helps to support some 200 external research projects, such as those at MCC, SEMATECH and the Microelectronics Center of North Carolina, now requires that every project have an individual sponsor within the company. Including the funding for external projects in the budgets for internal research projects also encourages Digital managers to pay close attention to the work at the consortiums, says John W. McCredie, Digital's director of external research. "So it's not much different from internal research," he adds.

These lessons have not been lost on SEMATECH, the consortium founded to develop manufacturing technology for semiconductor makers. Unlike most other consortiums, SEMATECH is narrowly focused on developing manufacturing techniques for producing five-inch (and eventually eight-inch) silicon wafers, which are the bedrock of most microprocessors.

"If we do our job well, we'll have 90 percent of the technology transferred by the time we do the 'final' transfer of technology," says Larry T. Novak, SEMATECH's manager of technology transfer. One means to the end: roughly half of SEMATECH's staff consists of investigators assigned by their employers to serve two-year stints at SEMATECH. (In contrast, only 15 percent of MCC's staff is loaned by member companies.) Robert N. Noyce, president of SEMATECH, also points out that all of its members are refining the consortium's work at home; SEMATECH is creating a manufacturing "tool kit," and the members in turn improve and use those technologies to manufacture proprietary products.

Yet consortiums represent only a partial step toward enhancing the competitiveness of U.S. companies. The most significant changes must take place within the companies themselves. Not surprisingly, the companies that receive the highest marks for taking advantage of technology from experienced consortiums are those that have reputations for fostering innovation at home, such as the 3M Corp. 3M emphasizes that consortiums play only one part in moving technology into use. It would be hard to imagine a consortium independently developing technology that was central to the company, says Thomas E. Wollner, a senior vice president at 3M.

Inman, moreover, argues that manufacturers have to develop iterative product cycles: design a product, make and sell it, collect feedback from the consumers, modify the product and sell the adapted version. "A consortium is a terrific way to put focus on high-risk projects," he says; companies in turn must learn to evolve technology. —*Elizabeth Corcoran*

Housing Developments

Increasing competition pushes more technology into houses

The U.S. house-building industry, one of the most fragmented and decentralized manufacturing businesses, has been called "the industry that capitalism forgot." But it might be described even more aptly as the industry that largely forgot about research. Both descriptions, however, may eventually become passé. Concerns about foreign competition in almost every aspect of the housing business—from materials to complete houses—are spurring the industry to take a hard look at the house of the future.

From the vantage of the Government, if not of the builders themselves, the threat of foreign competition begins with the materials and products that go into homes, including plywood and gypsum. After a building-materials surplus of $700 million in 1981, the U.S. ran a $2.34 billion deficit in 1987. When trade imbalances in building appliances, such as plumbing fixtures and ovens, are added, the deficit widens to $6 billion.

Lack of investment in research and development makes the industry's long-term prognosis also look bleak. For most U.S. industries the annual R&D expenditures are about 2.6 percent of revenues; "no matter how you look at the housing numbers, they're sure a long way from that," says Henry Kelly, a senior associate with the Office of Technology Assessment. Most of the roughly $80 million the U.S. government annually spends on commercial and residential construction research goes into efforts to improve energy efficiency. "There's exciting work going on in window technology and light fixtures," Kelly says, which is aimed at developing highly energy-efficient systems. But, he adds, "energy research only gets at part of the housing problem."

Foreign manufacturers are making inroads into everything from sophisticated appliances to innovative houses. Forget the tacky, modular houses of the past; "we prefer to call our product 'factory-crafted housing,'" says Michael Mononen, a program manager for building systems in the Swedish Trade Council.

Twenty years ago fewer than half of the new houses in Sweden were assembled from factory-made modules, Mononen says. But the 1973 energy crisis convinced the Swedish government to draft a single national building code requiring high energy efficiency. (In contrast the U.S. has four model codes, which thousands of communities have amended with local requirements.) As a result about 95 percent of all the new single-family houses built in Sweden are now factory crafted. "Stick built" houses have not been able to match the energy efficiency and quality of the factory-crafted modules, Mononen says.

Importers are betting that energy efficiency and sturdy Swedish construction will appeal to U.S. home buyers as well. Components of the home, such as walls and floors, are shipped from Sweden and can be snapped together by trained workers in three days; total construction time, including installation of appliances, is about eight weeks, says James E. Metivier, who is putting up about 175 Swedish-manufactured duplexes in Laconia, N.H. The houses are not cheap; Metivier says selling prices are in the $150,000 range.

Although factory-built houses are common in Japan, Michael L. Joroff, director of the laboratory for architecture and planning at the Massachusetts Institute of Technology, worries more about competition from Japanese manufacturers of house components, such as home automation and electronics systems. "Americans build houses as shells. Someone else does the furnishing," he says. "In Japan, a house is a product, and manufacturers continue to sell to their customers." Even though Japanese companies may not be building houses for the U.S., he says, they are aiming to service them.

One U.S. foray into home automation is the "smart house" project, initiated by the National Association of Home Builders about four years ago. Under the smart-house rubric, more than 100 North American manufacturers of housing components have been trying to design a system for distributing energy and communications throughout a house via a single cable. Any device—a lamp, a telephone or a cable television system—could be plugged into any wall outlet. A homeowner could program the system to turn components—lamps, thermostats, alarms and so on—on and off at prescribed times. The system might also incorporate a computer that would learn the habits of the homeowner and adjust devices accordingly.

"Intelligent buildings are a grand notion," says Steven Winter, founder of the Steven Winter consulting group in Washington, D.C. "It's unclear whether the (U.S.) smart house will be the winning system," says David Engel, an analyst at the Department of Housing and Urban Development. European and Japanese manufacturers are devoting much energy to a variety of home automation systems, he adds; Joroff predicts that the Japanese will be first to introduce smart technology for homes. "Whoever wins that war," Engel says, "will win the market for future American building products." —*E.C.*

A Truffling Matter

Lab coats replace pigs in the hunt for truffles

Picture this: a crisp winter day in southern France, oak branches creaking in the wind, a pig tugging impatiently on her leash as she catches a whiff of the truffles nestled near the tree roots. For centuries, epicures have had to rely on such traditional means for securing truffles. Now they need wait only for the next production run. A French company is marketing an artificial truffle oil that replicates the aroma of the fungus. Meanwhile, two Californian entrepreneurs are cultivating miniature truffles indoors.

Although by ancestry it belongs with fungi as mundane as the chestnut blight and the penicillin mold, *Tuber melanosporum*—the black truffle—has often been called the "diamond" of haute cuisine. Words fail to describe its delicate flavor, according to David P. Kellaway, a master chef at the Culinary Institute of America. ("Musty" might satisfy a less cultivated palate; to female pigs, truffles are attractively reminiscent of the smell of potential mates.) The fungus' price is similarly distinctive: a four-ounce jar of black truffles at a New York City market typically fetches $60.

For true gastronomes, declining truffle harvests are the real worry. Truffles, which take roughly seven years to reach their mature diameter of an inch or two, are found primarily in a few areas of France, Italy and Spain. Since 1925, when the annual truffle take was about 450 tons, truffle harvests have plummeted, to only 20 tons in 1986. The decline may reflect the loss of the stands of oak in which the fungus tends to grow.

Enter Thierry Talou, a chemical research engineer at the agroresources chemistry department of the National Polytechnic Institute in Toulouse, who is finishing his doctoral thesis on truffles this year. At the suggestion of Pébeyre S.A., a leading French distributor of truffles, Talou spent the past few years chemically dissecting the truffle aroma. He employed gas chromatography to isolate the volatile compounds that contribute to the smell. Unlike some fruits and vegetables, in which one compound produces the characteristic scent, mature truffles owe their aroma to a blend of at least 14 compounds, says Talou. Among them he found that dimethyl sulfide plays a leading albeit not overbearing role. Talou replicated the truffle aroma by mixing the compounds with vegetable oil.

How closely does the oil mimic the original? The chemical match is good, Talou says. The truffle oil has also been tested by French chefs; Talou says they have assured him that the oil is "a perfect imitation of the famous black truffle aroma." Perhaps more critical, however, were the field tests. "Samples of our flavoring buried in soil were detected by trained dogs with the same success as for truffles," reports Talou. (Two years ago Talou tried searching for truffles with a portable gas detector instead of animals. The suction-equipped sensor included a vacuum pump and a flame-ionization gas detector. Although it easily registered the presence of exhumed truffles, the gas detector failed to find buried truffles in an area where a dog sniffed out three. Consequently, Talou sets great store by animal tests.)

Still, there is more than one way to develop a truffle, as Moshe Shifrine, a retired immunologist in Woodland, Calif., is proving. During the past 30 years Shifrine has spent his spare time tinkering with truffle spores. Several years ago he hit on a growth medium that was conducive to growing truffles in a petri dish. With the help of Randy Dorian, formerly a biotechnology investigator, Shifrine has scaled up to harvests of roughly 60 pounds per week. Because Shifrine harvests his truffles after one year's growth, however, his fungi are tiny compared with their wild brethren. As a result he sells truffle oil, paste, powder and juice, but no fresh truffles.

Shifrine's "La Truffe" products began to reach U.S. consumers a year ago; a subsidiary of Pébeyre started selling Talou's "Arome Artificiel de Truffe" to professional chefs in Europe last July. Although both investigators say that consumers have been pleased with the taste of their products, there are caveats. Talou notes that according to "sensory technologists," aroma contributes no more than 80 percent of a food's flavor. Of the Californian truffles Talou says that in France "La Truffe" products would have to meet rigorous standards—including minimum spore count—before they could earn the appellation *truffes.* —*E.C.*

BLACK TRUFFLES grow underground, typically near the roots of oak trees. They are harvested in the winter months with the help of trained dogs and pigs. (Pictures courtesy of Pébeyre S.A.)

The Past and Future Amazon

The climatic history of the Amazon rain forest indicates that the ecosystem is well adapted to certain natural disturbances. Does it have the resilience to tolerate human exploitation?

by Paul A. Colinvaux

The rain forest covering the vast Amazon River basin of South America looks from the air like a uniform green carpet cut here and there by water. Actually the forest is anything but uniform. The "carpet" is the forest's canopy, formed by the broad leaves of many different kinds of giant trees, and this canopy is but the topmost layer of an ecosystem supporting more species than any other region on earth. The rain forest is home to perhaps 80,000 plant species (including 600 kinds of palm alone) and possibly 30 million animal species, most of them insects.

This remarkable diversity was once presumed to be a product of an everlastingly stable climate of abundant rain and warmth. Spared the short-term catastrophe of winter every year and the long-term disaster of glaciation, theory said, the tropical domain known simply as "the Amazon" would lose few species and accumulate many. Yet evidence from the field suggests that the Amazon is subject to climatic change on all time scales, including cooling when ice-age climates loose glaciers over more northern territories. Moreover, far from being disastrous to life in the Amazon, the moderate climatic disturbances in the region may actually help account for the splendid diversity of the Amazon rain forest today.

The influence of past climates on species diversity in the Amazon ecosystem is of more than academic interest; it is an invaluable predictor of the forest's ability to endure change. As human beings lay waste to massive tracts of vegetation, an incalculable and unprecedented number of species are rapidly becoming extinct. Nations with jurisdiction over the rain forest—notably Brazil, Venezuela, Colombia, Ecuador and Peru—need to identify management strategies that can save as many species as possible and yet be compatible with the need for economic development. Those who devise such strategies must take into account the kinds of stresses that affect species diversity, and so it is important to learn what insults the ecosystem has tolerated throughout history.

PAUL A. COLINVAUX, author of several books on ecology, has been professor of zoology and anthropology at Ohio State University since 1972. A native of Great Britain, he obtained bachelor's and master's degrees at the University of Cambridge before moving to the U.S. and earning a Ph.D. in zoology from Duke University in 1962; he went to Ohio State in 1964. When he is not engaged in his own work, Colinvaux enjoys accompanying his wife, a coral reef biologist at Ohio State, on her diving expeditions.

The old belief that climatic stability accounted for species diversity in the Amazon emerged, strangely enough, from observations of the deep sea. Howard L. Sanders of the Marine Biological Laboratory in Woods Hole discovered high diversity among the mud-dwelling animals of the deep-ocean floor in spite of the cold darkness and low biological productivity there. He argued that such hostilities to life are offset by the perpetual sameness of the place. Without significant fluctuations in physical conditions, the extinction of species that are adapted to prevailing conditions should be rare. In the course of time new species would continue to evolve, and so the rate of speciation would be greater than the rate of extinction, resulting in the accumulation of great diversity.

Sanders suggested that the Amazon forest and other tropical forests could be thought of as analogous to (albeit more productive than) the deep sea: places with a stable supply of annual moisture and warmth in which extinction should be rare. The absence of winter and glaciation and evidence that rain-forest trees have persisted for some 30 million years or longer somewhere in the 3,000-kilometer-wide, 1,000-kilometer-long Amazon basin supported this view.

Then, in 1969, several observations cast doubt on the validity of the stability theory for the Amazon, implying that the climate there has fluctuated significantly in the past. Jürgen Haffer, then of the Mobil Research and Development Corporation, noted that different corners of the basin held their own, separate bird faunas in spite of the fact that essentially unbroken green forest spread from the western edge of the Amazon to the Atlantic Ocean in the east. Haffer and other biologists also noted species disjunctions among butterflies and a few other groups. This pattern presented a glorious puzzle to workers studying the distributions of plants and animals: Why would the populations have become isolated if the habitat in which they lived was continuous?

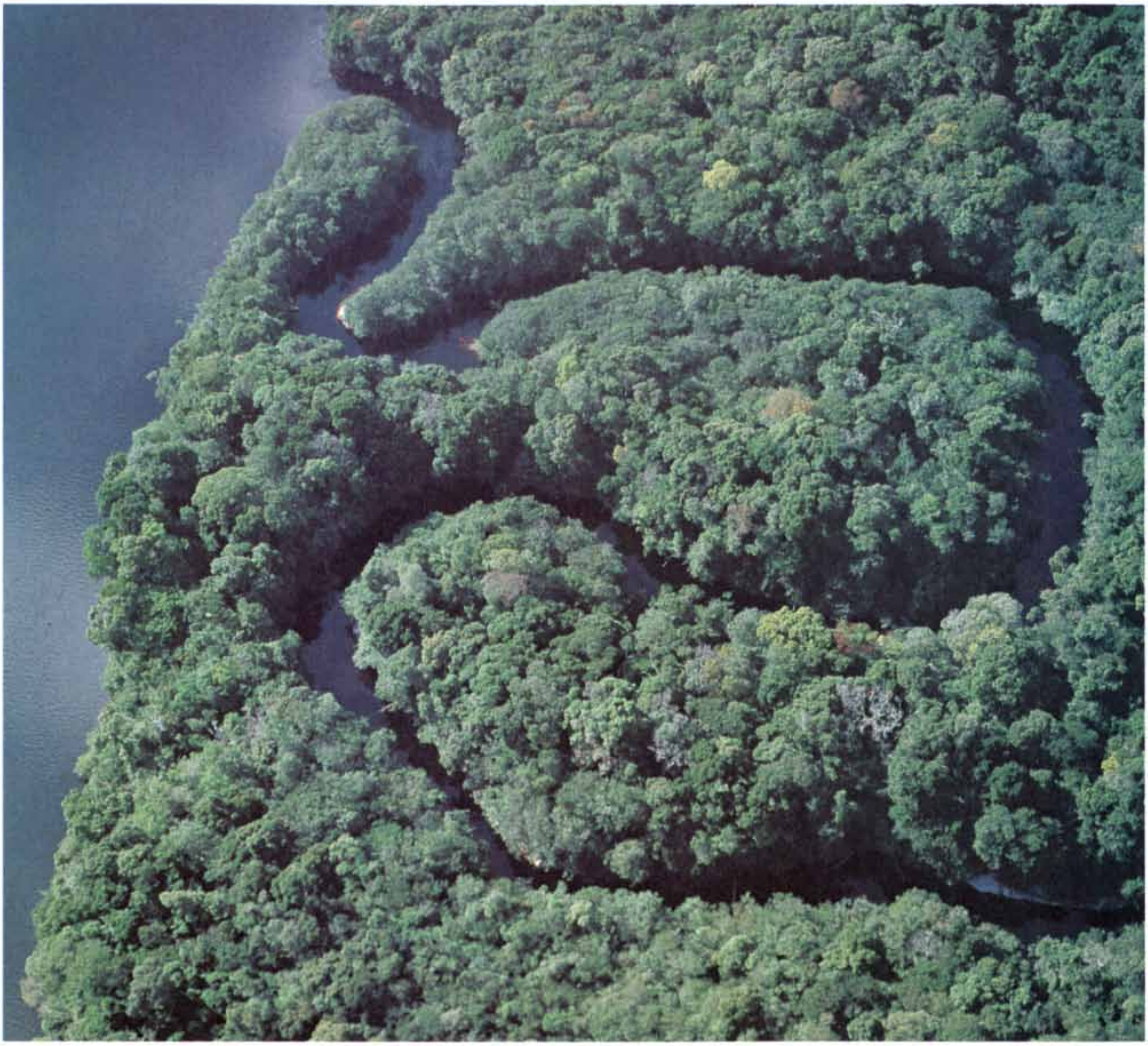
Haffer proposed an explanation so bold and logical that many of us wished we had thought of it first. He suggested that the modern isolation has its roots in times past, in the latest ice age. Observing that the regions of species isolation generally centered on outcrops of high ground and that lowlands are drier than uplands, he proposed that in glacial times the Amazon lowlands become a sea of near-desert arid plain; meanwhile the more elevated regions become islands of moisture and hence serve as refuges for the fauna and flora of the rain forest. Populations that were once continuous diverge along this ecological gradient and become permanently separated as they adapt to their upland habitats.

This hypothesis has great appeal because it appears to explain not only species disjunction in the Amazon basin but also the unusual species diversity there. The ice-age refuges would, as the earlier stability hypothesis predicted, protect existing species from extinction. But the periodic geographic isolation of related populations (there have been an estimated 13 ice ages to date) would facilitate ever-increasing divergence.

For the refuge hypothesis to stand there has to be proof that the lowlands of the Amazon, which now support wet forest, do indeed dry out in glacial times. No one has yet found compelling evidence that the Amazon was actually arid, although circumstantial evidence has been provided by computer models and also by sediment samples drawn from several sites in the tropics.

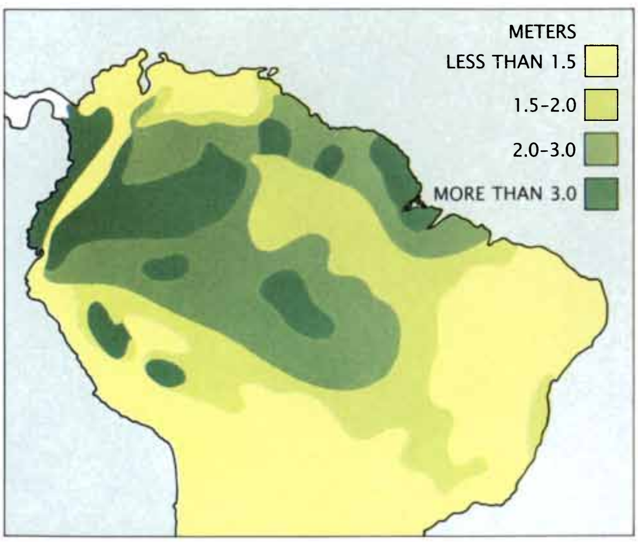
For example, John E. Kutzbach and Peter J. Guetter of the University of Wisconsin at Madison have experimented with computer models of the global ice-age climate, predicting reductions of up to 20 percent in monsoon rains in the tropics. Such a reduction would certainly spread aridity into strongly seasonal regions—those with true dry periods each year—including at least some parts of the Amazon. (Many so-called seasonal regions in the tropics do not have true annual dry periods, merely ones that are less rainy than other seasons.)

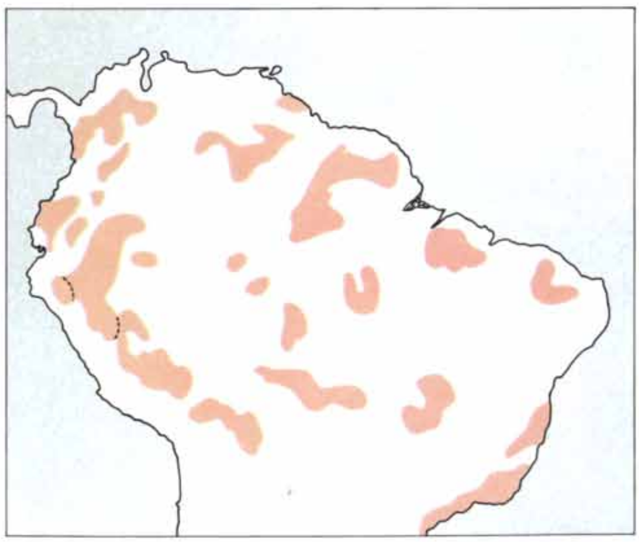
Daniel A. Livingstone of Duke University provided some of the earliest field data addressing the issue of aridity. He found by examining fossilized tree pollen, which is a good indicator of the local flora, that several tropical rain forests in modern Africa grow where dry woodland or savanna grew between 12,000 and 20,000 years ago; that is, late in the last ice age (from about 10,000 to 70,000 years ago). Hints of similar ice-age aridity in what are now wet tropical regions also came from outside the periphery of the Amazon basin, suggesting that what was true for ice-age Africa might also have



WATER-CUT PATCH of the Amazon rain forest is superficially uniform but, like the rest of the forest, actually includes many tree species growing side by side. Indeed, the wet, warm forest supports more plant and animal species per unit area (as well as overall) than any other place on earth. This diversity was once thought to stem from an eternally stable climate, but it now appears that the Amazon has always been perturbed by such disturbances as local electrical storms and floods. In contrast to, say, the glaciation of a continent, which causes mass extinctions, these more moderate hazards prevent extinctions. By felling some individuals of dominant species, they give weaker species a chance to establish themselves.

AMAZON RIVER BASIN in South America is bordered on the east by the Atlantic Ocean and on all other sides by lines of tall hills: the Venezuelan and Guiana Highlands to the north, the formidable Andes (rising to 6,000 meters) to the west and the Mato Grosso to the south. The rain forest (*dark green*) covering the basin persists up to an elevation of about 1,200 meters.





MAPS of average annual rainfall in the Amazon (*left*) and the ranges of distinct butterfly species (*right*) show that regions of species isolation generally overlap with the wetter parts of the basin; they also center over hills (not shown). A similar pattern, first noted in birds, gave rise to the hypothesis that the separated ranges were established long ago, perhaps in the last ice age (roughly 10,000 to 70,000 years ago). The theory postulates that the lowlands become arid in ice ages and the wetter uplands become isolated refuges for rain-forest species; the species then evolve in isolation and remain isolated even when the forest regenerates. New findings suggest that the postulated ice-age refuges were in fact too cool to support rain forests; instead, the forest persisted somewhere in the presumably warmer lowlands, which did not become arid. The new conclusion rejects the refuge theory but supports the notion that the Amazon's climate fluctuates appreciably.

been the case for the Amazon basin.

I myself gathered data suggestive of Amazonian aridity from the Galápagos Islands, which are about 2,000 kilometers directly west of the westernmost edge of the Amazon basin on the Equator. A well-dated sediment core, or pipeful of undisturbed sediment, from under the floor of the Galápagos' one freshwater lake shows that the lake was without water in the last ice age. It did not fill until the onset of the Holocene (the current, postglacial period).

On the continent itself others have found similar evidence at sites north of the Amazon. In northern Venezuela what is now deep Lake Valencia was without water at the end of the last ice age, and what is now a swamp on the coast of Guyana held dry vegetation. Different work suggests that a strongly seasonal forest to the south of the Amazon basin, near São Paulo, Brazil, was arid as well.

More evidence that the glacial-age Amazon lowlands might have been arid has been obtained from deep-sea cores from the Atlantic Ocean, right near the mouth of the Amazon River. John E. Damuth and Rhodes W. Fairbridge, then of Columbia University, have found that sediment delivered from the east-flowing river during the last ice age included a small fraction of feldspars, minerals from crystalline rocks, that could conceivably have been the erosion products of an arid landscape.

In spite of the abundance of the circumstantial evidence for ice-age aridity, I have become convinced that the evidence is not strong. Georg Irion of the Research Institute and Natural History Museum in Frankfurt has shown that the feldspars in the Atlantic deposits probably came from rocks in the Amazon basin, but not from rocks on the surface. The sea level fell in the ice age, and the Amazon's many tributaries presumably cut their channels deeper at the same time, carving out feldspars in the process.

The sediment-core data from the Galápagos Islands and around the Amazon on the mainland are also weak because the regions from which the sediments were taken, although closer than Africa, are still quite a distance from the Amazon rain forest. Moreover, the regions frequently come under the influence of different climatic forces. Between the study sites and the Amazon basin are the tall Andes, the Venezuelan and Guiana Highlands and another immense line of hills known as the Mato Grosso (the great scrub), which constitute respectively the western, northern and southern boundaries of the basin. Each of those boundaries is a site of climatic tension, so that the climate on one side of the heights is often quite different from the climate on the other side.

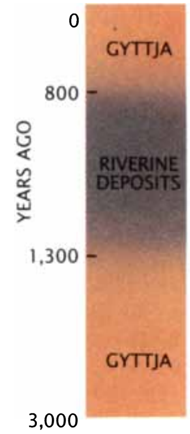
It may be that in the strongly seasonal regions of the modern Amazon reductions in monsoon rains do result in semiaridity during glacial times—perhaps just below the Guiana Highlands in the northeastern corner, immediately to the west of the Mato Grosso in the southwest and in some parts of the central basin in Brazil. For most other, wetter regions of the Amazon, where annual rainfall is from two to five meters, it is quite unlikely that even a 20 percent reduction in the rains would make any practical difference to life.

In 1985 the fact that there were no radiocarbon-dated, glacial-age fossil data of any kind from Haffer's postulated refuges or the lowlands of the Amazon prompted me and my colleagues at Ohio State University to seek data from the Amazon proper. We found the only fossils that have yet been dated to an ice age. They suggest that the Amazon was not dry. The region did undergo a significant change, however: it became cooler by several degrees Celsius.

We found our specimens by serendipity, while searching the westernmost edge of the rain forest, in eastern Ecuador, for old lakes whose sediments we hoped would hold ancient records. At one point we found ourselves on a dirt road in Mera, whose 1,100-meter elevation in the eastern foothills of the Andes places it close to the upper limits of today's rain forest and within one of the major putative refuges. There we noticed old tree stumps and logs that were embedded in an exposed patch of sediment. Because we had a bit of time on our hands as we waited for a flight to a distant lake, we took a collection of sediment and wood samples. Radiocarbon dating of the wood indicated that some samples were 35,000 years old and the others 26,000 years old. Although earlier than the time of maximum glaciation in the most recent ice age (18,000 years ago), these are dates of glacial times.

Another analysis of the wood samples showed they included softwoods, meaning conifers. Now, the only conifers in modern Ecuador are *Podocarpus* trees of the Andean forests, which today grow at elevations that are at least 700 meters above the Mera site. Indeed, pollen analyses of the sediment samples showed that the ancient forest at Mera had a strong Andean

BROAD-NEEDLE CONIFER, *Podocarpus oleifolius,* is one of several *Podocarpus* species that grow at some 1,800 meters and higher in the Andes of Ecuador but not in the Ecuadoran part of the Amazon rain forest, presumably because the trees require the relative cool of high elevations. The author recently found *Podocarpus* fossils dated to the ice age in the Ecuadoran rain forest. This discovery is the primary one suggesting that the refuge theory of species disjunction is incorrect. It indicates that the fossil site, at Mera, which is within one of the major postulated refuges, was cool in the ice age and hence was not a refuge. If Mera was too cool to support a rain forest, then some or all of the other postulated refuges were probably too cool also.



LAKE AÑANGUCOCHA (*left*) is one of four Ecuadoran lakes whose sediment records offer evidence that the western part of the Amazon rain forest underwent a long, unusually stormy period about 1,000 years ago. The top layer of sediment (*right*) is gyttja, or typical lake mud. Below this mud is a layer of typical river sediment, which was deposited over another layer of lake mud between about 800 and 1,300 years ago. The pattern shows that Añangucocha, once an isolated lake, was reentered and filled for some 500 years by its parent river about 1,300 years ago, probably because the river was swollen by repeated, powerful rainstorms. The river abandoned the lake about 800 years ago, presumably when the storms subsided.

cast to it and that it has no direct analogue in modern tropical America.

We conclude that the *Podocarpus* trees, which require both moisture and the relative coolness now found only in the mountains, grew at least 700 meters lower in glacial times. Taken together with the pollen data, this conclusion suggested to us that Mera was moist during the last ice age but was too cool to support a modern kind of rain forest. A standard formula suggests that the temperature in the Ecuadoran foothills was at least 4.5 degrees Celsius cooler than it is today. (The formula assumes that in moist air the temperature cools six degrees Celsius for every 1,000 meters of ascent up a mountain.)

There was, then, no refuge for warmth-loving species in Mera. If the temperature depression was widespread in the Amazon, as seems reasonable, then other high-lying parts of today's rain forest might also have been too cold to support rain-forest refuges. Given the well-established fact that typical rain-forest trees persisted somewhere in the Amazon basin, the data suggest that the rain forest was restricted to somewhere in the lowest region, where refuge theorists had predicted aridity. If the trees did in fact survive only in the lowlands, it follows that the lowlands were not arid, or at least were not uniformly so.

Obviously more than one data set from the ice age and data from the coldest part of the period are needed before a picture of the ice-age Amazon can be reconstructed with confidence. Still, the existing evidence suggests that the refuge scenario is no more supportable than the notion of imperturbability that it replaced. It also suggests that the rain forest shrinks in response to ice-age cooling and expands again during warmer interglacial periods, such as the one we are in now. In essence we have the inverse of the refuge picture: some part of the lowlands remains as a relatively warm, wet preserve for rain-forest trees and the uplands—formerly thought to be refuges—become inhospitable.

What then explains the species diversity and the species disjunctions identified by biogeographers? Changes in the climate between glacial and interglacial periods may still be important, albeit not in the way the refuge theorists predict. The migration I propose of rain-forest species to higher ground in warm interglacial periods such as the present one would provide opportunities for the formation of novel mixtures of species. Populations from the warm lowlands of glacial times could then become adapted to their new upland conditions and diverge from their ancestors. Perhaps something like that happened to the birds and butterflies now living in the uplands.

Another part of the answer certainly has to do with the varied climate and geography of the Amazon. The game of speciation has always been played across a huge playing board: the Amazon basin is almost the size of the continental U.S. All parts of the board are superficially alike and indeed have some species in common, but the territory does not hold one monolithic ecosystem. Rather, there are large regional and local differences in rainfall, seasonality, soil and susceptibility to flooding, all of which can affect the mixture and evolution of species in any given area.

A further explanation for the diversity is known as the intermediate-disturbance hypothesis, put forward by Joseph H. Connell of the University of California at Santa Barbara and independently by Stephen P. Hubbel, now of Princeton University. This hypothesis suggests that the highest species richness will be found not where the climate is stable but rather where environmental disturbance is frequent but not excessive.

The intermediate-disturbance hypothesis concedes that massive catastrophes lead to large-scale extinction. This is what would happen if asteroids struck the earth, as they may have done at the time of dinosaur extinction. But lesser and local catastrophes, such as storms and flooding, do not generally extinguish entire species. Instead, by killing some fraction of a dominant species, the smaller hazards prevent winner-take-all competitions from going to their conclusion (extinction for one species) and give initially weaker organisms the opportunity to establish themselves.

The effects of intermediate disturbances can be seen quite readily in rain-forest trees. Great gashes and succession communities are everywhere in the Amazon and other tropical rain forests [see "The Tropical Rain Forest," by Paul W. Richards; SCIENTIFIC AMERICAN, December, 1973]. Wherever gaps are cut in the forest by the felling of large, sun-blocking trees (such as rubber trees and balsa), various species of short-lived, sun-loving plants, together with their associated fauna, populate the gaps. This process initiates a series of successions that, if they go on undisturbed, culminate centuries later in the groupings of giant trees characteristic of mature rain forests. The big rain-forest trees actually fall fairly easily because their roots are shallow: often half of their root mass is in the top 20 centimeters of soil. All it takes is a strong push by wind or the washing away of topsoil at the base of the trees by floods or moving streams.

A growing body of data suggests that at least throughout the Holocene (and probably throughout history) the Amazon has been perturbed in one region or another by storms, erosion and other forces. Indeed, the Amazon's topography apparently changes from century to century and even from decade to decade.

Some of what is known about the Holocene's history comes from the sediment cores my colleagues and I obtained from the Ecuadoran lakes alluded to above. Our oldest data are from lakes in volcanic-explosion craters, the only crater lakes yet known in the Amazon. One of these lakes, known as Kumpak[a] in the Shuar language of the local populace, has muddy, oxygen-depleted water typical of an Amazon river-fed lake. The other lake, Ayauch[i], is transparent and blue, with oxygen deep in the water. It is almost like an alpine lake—a wildly exciting and improbable thing to find in the jungle.

Mark B. Bush, a postdoctoral fellow in my laboratory, has completed pollen analyses of the 7,000-year record from Ayauch[i], finding evidence for rain forest throughout most of that time, although there was a prolonged local drought 4,000 years ago, long after the ice age ended. (He also found evidence that people had grown maize in the vicinity 3,000 years ago—the earliest record of maize cultivation in the Amazon.)

Kam-biu Liu, now of Louisiana State University, found a different kind of history in the Kumpak[a] sediments. These are banded throughout with layers of different textures, suggesting a history of heavy rainstorms that washed sediment in pulses from the banks into the lake. In other words storms have been remodeling the land around Kumpak[a] for at least the past 5,000 years.

Again in the Ecuadoran rain forest, but to the north, cores we obtained from four lowland lakes abandoned by their parent rivers centuries ago have also yielded clear evidence of at least one extended period of unusually powerful storms and forest perturbation in the west. The sediment samples from each lake were topped by a layer of gyttja, typical lake mud, first laid down some 800 years ago, according to radiocarbon dating. Beneath this layer in all four lakes was river sediment that had been deposited continuously since about 1,300 years ago. The lakes, therefore, were left by their parent rivers roughly 800 years ago and have not been reentered since.

We postulate that between 800 and 1,300 years ago excessive rains in the mountains to the west caused massive flooding and forced rivers to spread back to old, once-abandoned channels in the lowlands, which became largely inundated. Our pollen analysis of the old river mud indicates that early succession forests were present on temporary sandbanks that formed when the floods, which presumably killed mature trees, subsided temporarily.

Marcia L. Absy of the Amazon Fisheries Institute in Manaus, Brazil, has also found evidence of flooding at the appropriate time in lakes of the central Amazon. Absy cored five *varzea* lakes (which are filled by rivers during the wettest seasons and then are abandoned to evaporate slowly in the driest seasons) from different watersheds near Manaus, including lakes spawned by south-flowing rivers from the dry Guiana Highlands, by north-flowing rivers from the Mato Grosso and by east-flowing rivers draining the Andes of Ecuador and Peru. Her data show that only lakes in the path of the western drainage hold the flood record—not those in the paths of the north- and south-flowing streams. In other words the stormy period was a local phenomenon of the west and had long-range effects solely on regions in the path of the swollen Amazon tributaries.

The western Amazon, then, has been disturbed in the Holocene by at least one unusually drastic climate change, whose effects lasted almost a millennium. Doubtless other climatic events of equal long-term effect await discovery in the turbulent basin.

On time scales of a few centuries the perturbing effects of erosion by moving streams are particularly evident. Some rivers are fast-moving and wear away sediment rapidly; others are sluggish but nonetheless erode sediment as they move. Many rivers flow now through one channel, now through another. As they

NAPO RIVER in Ecuador scours its banks and topples trees. It is one of many rivers that constantly rework the topography of the rain forest. The author argues that the Amazon ecosystem is adapted to such assaults, which leave survivors, but is not adapted to the catastrophe of clear-cutting now being inflicted by human beings.

change direction they topple trees in their path; they also leave behind sediment ripe for colonization. All are subject to flooding in rainy periods, at which time they wash away trees on their banks and redecorate the landscape. The turbid western tributaries of the Amazon are particularly active. Indeed, various estimates suggest that 80 percent of suspended solids in the lower Amazon comes from the west.

On the basis of published measures of the sediment discharged at the mouth of the Amazon River each year, my group tried to calculate how fast the huge rivers draining the Andes are eroding the western region. That is an exercise in unsafe extrapolation, but it does suggest that many centimeters of land surface are lost every century. Perhaps half the rooting depth of a typical, mature Amazon tree can be washed away in less than the tree's normal lifetime.

Jukka S. Salo and his team at the University of Turku in Finland, working in the Peruvian Amazon, have obtained better evidence of the disturbance caused by erosion in the west. With the help of satellites they have constructed maps of different forest types, including both early succession communities (which typically grow on sediment left behind as rivers progressively move their channels) and mature communities (which become established only centuries after an old riverbed has been abandoned). From these maps they determined that in the past couple of centuries as much as a quarter of the Peruvian forest has been swept away and rebuilt as new communities.

Rivers also perturb the land on smaller time scales. With aerial photographs taken 13 years apart, Salo and his group showed that a single small river had reworked 3.7 percent of its floodplain in these few years; on the average, it had eroded 12 meters of land per year.

Most people are not surprised to learn that the Amazon forest is disturbed by storms, flooding and erosion. Few, however, expect a rain forest to be perturbed by natural fires, and yet rain-forest trees apparently do burn in the Amazon. Recently Robert L. Sanford of the University of California at Berkeley and other investigators found layers of charcoal in soil pits in southern Venezuela, at the northern tip of the Amazon basin. Radiocarbon dating showed that some of the samples were 6,000 years old and hence were deposited before human beings are thought to have entered the region. At least some of the charcoal layers, then, must be products of natural wildfires.

What would cause such fires? Rainforest trees can be thought of as water-cooled energy traps. They spend their days absorbing intense solar radiation on the huge surface areas of their thin, broad leaves, an undertaking that is possible only if the trees can dissipate heat by evaporating immense quantities of water. Sanford thinks it possible that in a mere month without rain the trees might expend all the water within reach of their shallow roots, after which their leaves would wilt, overheat and be lost. The sun could then scorch the litter under the leafless trees, enabling spontaneous combustion or lightning to set the dried forest alight. The random properties of weather can be expected to produce such a dry spell in the rain forest once or twice in a span of several thousand years.

It becomes clear that, indeed, the Amazon basin has always been a place disturbed. In ice ages, if our single data set is representative of the region, the land apparently cools and the forest is displaced. In warmer times the Amazon forest is probably subject to the kinds of disturbance identified from lake-sediment records and other records of the Holocene. Different places at different times are beset by storms, flooding and erosion, and they burn in the rare periods when no rain falls for days on end. Few patches are likely to go undisturbed for more than a century or two. The result today is a mosaic of gaps, successions and mature forest—and a wonderful assortment of plant and animal species.

What does all of this say about the future of the Amazon ecosystem, now severely imperiled by humankind? The fact that species have accumulated in a place of constant change suggests that the fauna and flora can tolerate some human activity, if the activity is on the order of natural intermediate disturbances that always leave survivors. It must be emphasized, however, that nothing in the history of the Amazon seems to have approached the clear-cutting now being inflicted on the system by human beings. Such activities are more akin to the catastrophic forms of natural disturbance that have led to the extinction of vast numbers of species in the past.

Moreover, the larger animals that dwell in the rain forest cannot survive the devastating effect of modern firearms. Indeed, the plant-eating primates and sloths that graze in the canopy and their flying predators, such as the harpy eagle (an engine of destruction so powerful that it can tear monkeys out of the canopy), are desperately vulnerable to shotguns. One person with a 16-gauge shotgun can remove all the harpy eagles and less motile primates within 10 kilometers of a camp in a year, and thousands of people have done just that. Protected preserves are the only hope for such animals, and the governments of the Amazon are setting aside land for this purpose. The proper size for these refuges is now an active area of research.

The trees are another story. The contemporary reality is that much of the Amazon basin will be turned into pasture as people clear the land for cattle grazing. The only hope for trees and perhaps for other plants and insects probably lies in the development of wise uses: ways to generate cash flow from the remaining forest that inflict no more disturbance than the rain-forest species are accustomed to. Perhaps parts of the forest could be set aside for vacation retreats or retirement communities or for industries that manufacture products that do not require enormous amounts of power and would not pollute the ecosystem. History does suggest that parts of the Amazon can be exploited productively without causing mass extinction, but wise use must be the overriding theme.

FURTHER READING

EQUATORIAL ATLANTIC DEEP-SEA ARKOSIC SANDS AND ICE-AGE ARIDITY IN TROPICAL SOUTH AMERICA. John E. Damuth and Rhodes W. Fairbridge in *Geological Society of America Bulletin*, Vol. 81, No. 1, pages 189–206; January, 1970.

SEDIMENTATION AND SEDIMENTS OF AMAZONIAN RIVERS AND EVOLUTION OF THE AMAZONIAN LANDSCAPE SINCE PLIOCENE TIMES. G. Irion in *The Amazon: Limnology and Landscape Ecology of a Mighty Tropical River and Its Basin*, edited by Harald Sioli. Dr. W. Junk Publishers, 1984.

BIOGEOGRAPHY AND QUATERNARY HISTORY IN TROPICAL AMERICA. Edited by T. C. Whitmore and G. T. Prance. Oxford University Press, 1987.

AMAZON DIVERSITY IN LIGHT OF THE PALEOECOLOGICAL RECORD. Paul Colinvaux in *Quaternary Science Reviews*, Vol. 6, No. 2, pages 93–114; 1987.

PLEISTOCENE FOREST REFUGES IN THE AMAZON: EVALUATION OF THE BIOSTRATIGRAPHICAL, LITHOSTRATIGRAPHICAL AND GEOMORPHOLOGICAL DATA. Jukka Salo in *Annales Zoologici Fennici*, Vol. 24, pages 203–211; 1987.

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 two 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, an eight-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.

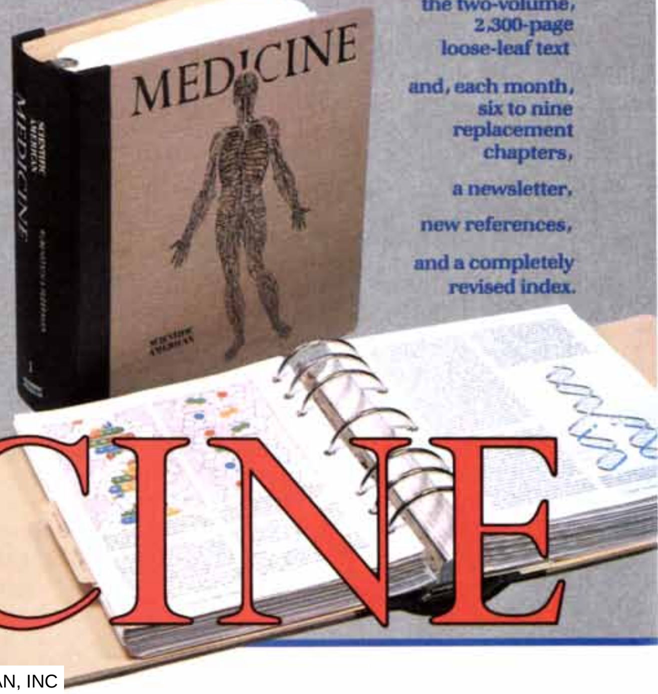
At no additional cost, a 32-credit CME program, to save you valuable patient-care time and the expense of attending review courses.
Earn 32 Category 1 or prescribed credits per year with our convenient self-study patient management problems; each simulates a real-life clinical situation. Choose either 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).

☐ Yes, I'd like to try the SCIENTIFIC AMERICAN *Medicine* system. Please enter my subscription at a first-year price of US$265* plus $7 shipping for a total of US$272.
☐ Also enroll me in the CME program. I'd prefer:
  ☐ the printed version at no additional charge.
  ☐ the disk version at US$97* additional. Computer type or compatible:
    ☐ IBM® PC (256K) or compatible (5 ¼") ☐ IBM® PS/2 (3½") ☐ Macintosh™ (512K)
    ☐ Apple® IIc/IIe/IIGS ☐ Apple® II + with 80-col. card by: ☐ Apple® ☐ Videx®
☐ Enroll me in the disk CME only at US$152.* (Note model above.)
☐ Check enclosed* ☐ Bill me
☐ VISA ☐ MasterCard Exp. Date ______ Account No. ______
Name ______ Specialty ______
Address ______
City ______ State ______ Zip Code ______
7104

*Add sales tax if for CA, DC, IL, MI or NY. Allow 8 weeks for delivery. Add US$10 for shipping to Canada. IBM is a registered trademark of International Business Machines Corporation. Apple is a registered trademark of Apple Computer, Inc. Videx is a registered trademark of Videx, Inc. Macintosh™ is a trademark of McIntosh Laboratory, Inc. and is used by Apple Computer, Inc. with its express permission.

Or, call toll-free: 1-800-345-8112

Your subscription includes:
the two-volume, 2,300-page loose-leaf text
and, each month, six to nine replacement chapters,
a newsletter,
new references,
and a completely revised index.



SCIENTIFIC AMERICAN MEDICINE

Marsupial Frogs

Some tropical frogs incubate eggs on the mother's back, often in a special pouch. Certain features of the adaptation recall pregnancy in mammals and the eggs and embryos of birds; others are unique

by Eugenia M. del Pino

Mammals give birth to live young and suckle them; birds lay eggs and brood them. Frogs, being amphibians, deposit their eggs in the water and have aquatic larvae, tadpoles. As these generalizations suggest, a common reproductive mode transcends much of the diversity of body forms within any one class of animals, uniting bats with elephants, hummingbirds with ostriches.

In the rain forests and mountains of Central America and tropical South America live 60-odd species of tree frogs that belie this truism. As adults these frogs look like other tree frogs, and yet their mode of reproduction is not at all froglike. Their eggs develop not in the water but on the mother's back, often inside a special pouch. In many species the young emerge as fully formed froglets, in others as advanced tadpoles.

This unusual mode of reproduction has called for some remarkable specializations. Protected on the mother's back, the embryo can develop at a leisurely pace. Prolonged development in turn means the embryo must be supplied with nutrients and able to exchange gases and fluids with its environment. All of this has required changes in the physiology of the mother, the structure and molecular characteristics of the egg, and the pattern of embryonic development. Many of these specializations invite comparisons between egg-brooding frogs and animals that are evolutionarily very distant.

EUGENIA M. DEL PINO is professor of biology at the Pontifical Catholic University of Ecuador in Quito. A native of Ecuador, she did her undergraduate studies at the university where she now teaches and then came to the U.S., earning an M.S. at Vassar College in 1969 and a Ph.D. from Emory University in 1972. In that year she returned to Ecuador to join the faculty at Quito. From 1984 to 1985 she held an Alexander von Humboldt Foundation fellowship, which enabled her to work with Michael F. Trendelenburg at the German Cancer Research Center in Heidelberg. In 1987 del Pino was made a member of the Latin American Academy of Sciences; she is Ecuador's sole member to date.

The pouch, of course, recalls the pouch of marsupials. The internal structure of the frog pouch and the hormonal controls that govern incubation there parallel the anatomical and hormonal apparatus of pregnancy in mammals. The large, yolk-filled egg looks more bird- or reptilelike than froglike, and so in some respects does the embryonic development that unfolds in it. In still other aspects of their reproduction, however, egg-brooding frogs are unique—a reminder that the parallels with other classes of animals are the result not of evolution along one line of descent but of evolutionary convergence, bridging hundreds of millions of years of separate history.

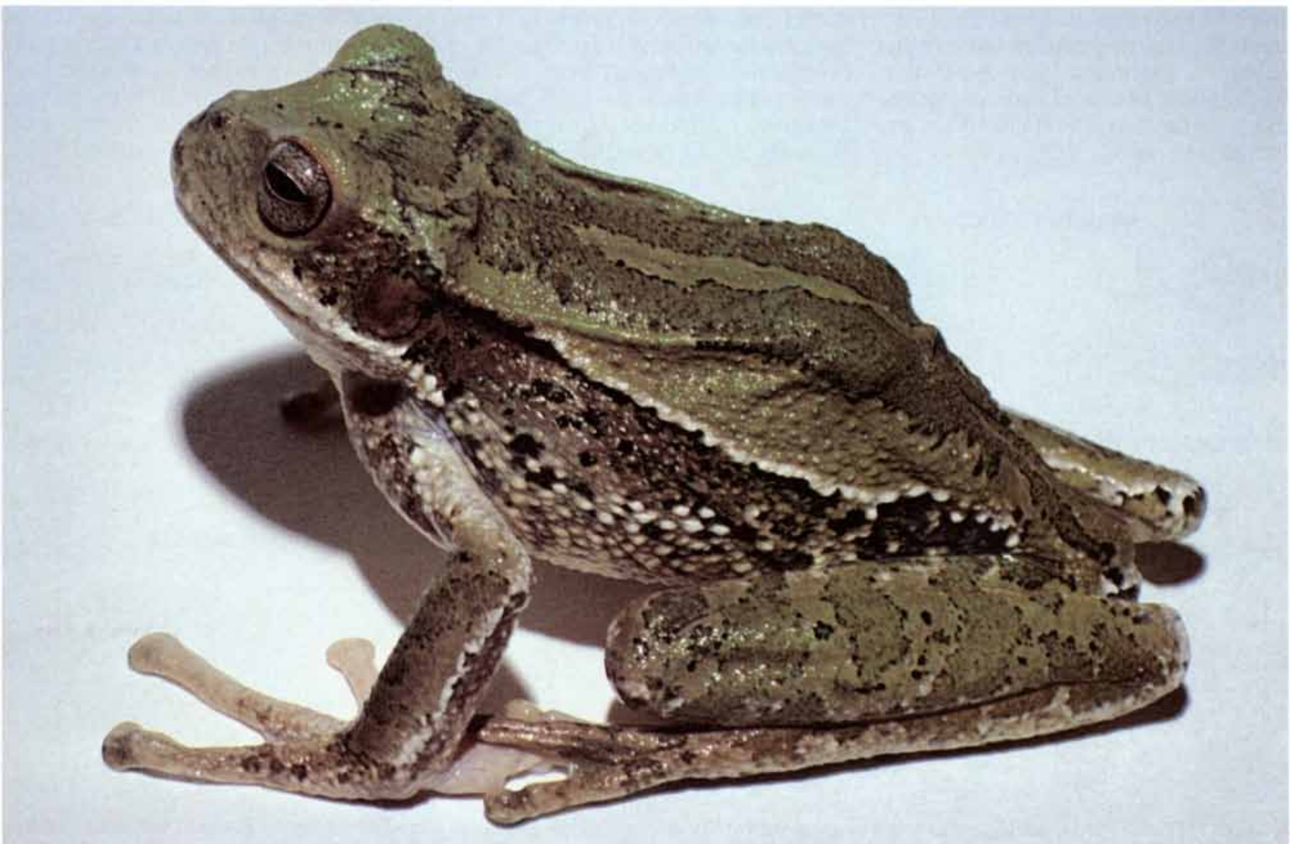
In 1972, when I began to teach at the Pontifical Catholic University of Ecuador in Quito, the unusual reproductive features of egg-brooding frogs were known only in outline. For me these frogs were no more than a textbook curiosity. Soon after moving to Quito, however, I discovered that one species of marsupial frog, *Gastrotheca riobambae,* lived in the gardens of the university. I began studying the frog and its reproductive mode, and I searched for specimens of other egg-brooding species.

G. riobambae is one of the few species of egg-brooding frogs (or any other frogs, for that matter) that have conquered the high-mountain environment of Quito, which lies nearly 3,000 meters above sea level. Most egg-brooding frogs inhabit the canopy of tropical forests in lowland Ecuador and other parts of South and Central America. In studying these species I relied on museum specimens and on the help of William E. Duellman, Linda Trueb, John Simmons and other workers from the Museum of Natural History of the University of Kansas, who collected living specimens of the frogs and brought them to my laboratory.

The tropical rain forest, where egg-brooding frogs are the most diverse, is probably also where they evolved. It is easy to see how selective pressures there might have favored the evolution of egg brooding. The large number of frog species in the tropical forest creates intense competition for breeding sites. (In just three square kilometers of forest near Santa Cecilia, Ecuador, Duellman and his student Martha L. Crump found 81 species of frogs, about as many as live in all North America!) In addition, numerous predators threaten frog eggs and larvae exposed in pools and streams.

As a result many tropical frog species have adopted new sites for their eggs. Some species deposit them on leaves above pools and streams; rain later washes the newly hatched tadpoles into the water. Other frogs hide their eggs under forest-floor litter; the eggs hatch directly into small frogs. The more elaborate device of incubating eggs on the mother's back evolved in one subfamily of the Hylidae, or tree frogs—the Hemiphractinae.

It is an ancient development. Brian E. Scalan and Linda R. Maxson of the University of Illinois at Urbana-Champaign, working with Duellman, have traced the evolution of egg-brooding tree frogs back in time by comparing the amino acid sequences of certain proteins in different species. Provided the protein sequences have changed at a constant rate in the course of evolution, the difference between any two sequences should be proportional to the evolutionary distance between the two organisms—the time since they last shared a common ancestor. The protein comparisons suggest that the lineages began to diverge between 40 and 80 million years ago. Presumably egg brooding had already arisen in the ancestral lineage.



POUCHFUL OF DEVELOPING EMBRYOS has drastically swollen a marsupial frog from Ecuador (*top*), a green specimen of *Gastrotheca riobambae.* After giving birth to some 218 tadpoles through the pouch opening, low on her back, the frog has reverted to a normal size (*bottom*). The photographs are by Charles W. Myers of the American Museum of Natural History.

By now there is much diversity among the seven genera—including about 60 species—of egg-brooding frogs. Three of the genera (*Stefania, Cryptobatrachus* and *Hemiphractus*) lack pouches; the eggs adhere to the back of the mother, where they develop and hatch as small frogs. Four other genera (*Fritziana, Flectonotus, Gastrotheca* and *Amphignathodon*) are true marsupial frogs, equipped with pouches. I shall focus on reproductive features of these pouched species.

Variations in the design of the pouch in different genera suggest that it may have originated from longitudinal folds of the skin on the back. In some species of *Fritziana,* for example, the developing embryos are walled in by the folds but are otherwise exposed. In other species of *Fritziana* and in the genus *Flectonotus* the folds have become flaps that meet at the midline of the back, enclosing the embryos in a pouch derived from two lateral compartments. The pouch has grown more elaborate in *Amphignathodon* and many species of *Gastrotheca:* it still has two lateral chambers, but its opening is now limited to an aperture low on the back, just above the cloaca (the combined reproductive and excretory opening). The high-mountain species of *Gastrotheca* (including *G. riobambae*) have what may be the most advanced form of pouch, in which the two lateral chambers are joined into one.

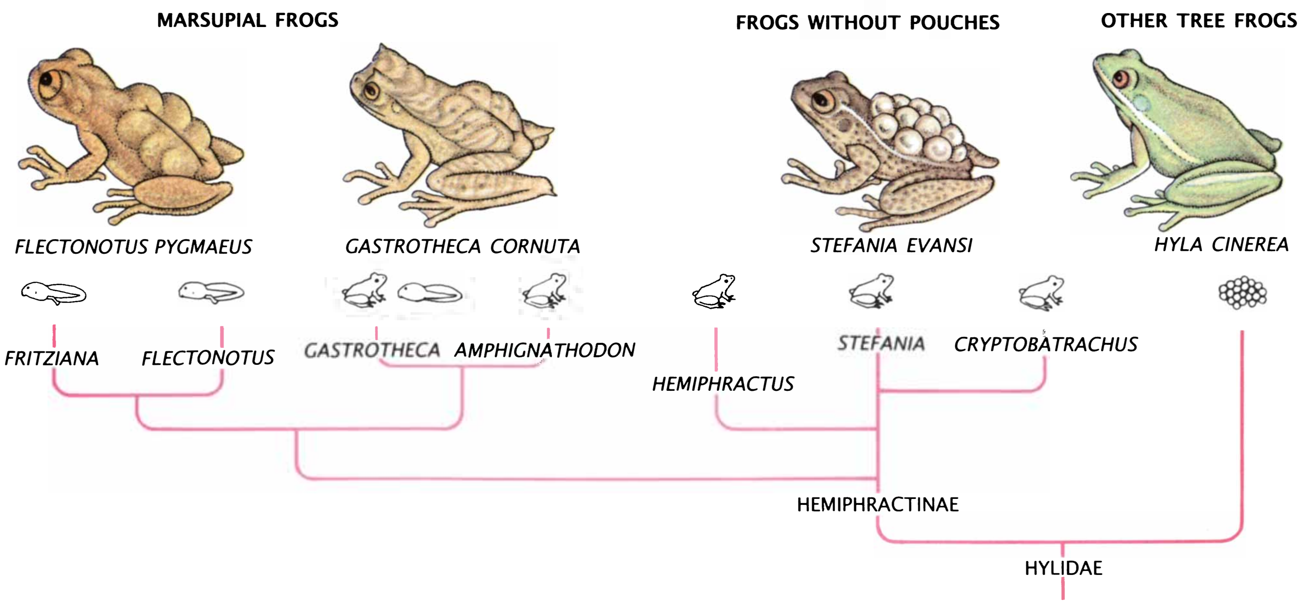
The evolution of the pouch brought new reproductive behaviors. Frogs copulate while the female is laying eggs, which are fertilized as they emerge from her cloaca. In most frogs the encounter takes place in the water, where the eggs will be left. Marsupial frogs mate on land, but they face a new problem: how to get the eggs from the cloaca into the pouch.

In the three or four species for which mating behavior has been observed, the male does the transferring. During the sexual encounter in *G. riobambae,* for example, the male climbs onto the back of the female and opens the pouch with his hind feet. As the eggs emerge from the female's cloaca, one or two of them per minute, the male catches each egg with his feet and maneuvers it into the pouch; the egg is fertilized on the way by sperm the male has deposited on the female's back. The process continues until 130 eggs, on the average, have been stowed in the pouch. The female then presses her back against a stone to work the pouchful of eggs into several even layers.

In many marsupial species several months pass before the eggs hatch. In order to emerge from the pouch the young—advanced tadpoles in *Fritziana, Flectonotus* and a few species of *Gastrotheca,* and small frogs in other cases—again need parental help. A female of *G. riobambae,* for instance, deposits newborn tadpoles in pools or puddles by clinging to a stone with her front legs while twisting her hind limbs to reach the pouch opening. She extracts the tadpoles one by one with her toes, taking two or more days to give birth to a full clutch.

In species that give birth to fully formed froglets, the process need not occur near water. While they are in the pouch the embryos of these frogs develop rudimentary versions of many features normally seen in tadpoles. By the time of birth their metamorphosis is complete: they have bypassed the aquatic stage. Presumably many of these tree frogs can play out their life cycle without ever descending from the forest canopy.

The protected environment of the pouch allows embryonic development to proceed much more slowly in marsupial frogs than in species whose eggs are exposed and vulnerable. Incubation in *G. riobambae,* for example, lasts between 100 and 120 days, in contrast to the day and a half it takes for aquatic eggs to hatch in *Xenopus laevis* (the African clawed frog), a species that is a good point of comparison because it has been thoroughly studied. How is the egg sustained through its long incubation? A frog embryo that hatches into a free-living tadpole within a few days does not need special mechanisms for exchanging gases, fluids and other substances with the environment. In a



EGG-BROODING FROGS, the Hemiphractinae, are a subfamily of the tree frogs, the family Hylidae. Based on comparisons of the egg-brooding frogs' proteins and reproductive features, William E. Duellman of the Museum of Natural History of the University of Kansas and Richard J. Wassersug of Dalhousie University in Nova Scotia proposed the evolutionary relationships shown here. Three of the seven genera lack pouches; the developing embryos simply adhere to the mother's back. The genera are also distinguished by whether the embryos are born as fully formed froglets or as advanced tadpoles (*small drawings*). In one genus—*Gastrotheca*—the form at birth varies among the individual species. Frogs are not drawn to scale.

UNIQUE REPRODUCTIVE BEHAVIOR accompanies egg brooding. During the sexual encounter in one pouched species, *Gastrotheca riobambae,* the female lays eggs while the male clasps her from behind and inserts his toes into the pouch (*left*). As the eggs emerge one by one from the female's uplifted cloaca, the male maneuvers them into the pouch; the eggs are fertilized by sperm he has previously deposited on the female's back. Over several hours, 100 or more eggs are transferred. Embryonic development takes more than three months, leaving the female so swollen that her movements are hampered (*middle*). She gives birth in the water, inserting the toes of her hind limbs into the pouch to help release the tadpoles (*right*).

marsupial frog, however, one would expect both mother and embryo to be equipped for such exchanges.

Between breeding seasons, when the pouch is empty, its interior resembles ordinary frog skin. During incubation, however, the walls develop a thin lining that is richly supplied with blood vessels. This pouch lining—which is sloughed off after the eggs hatch—conforms to the shape of the eggs, forming an individual chamber for each one. Within each egg chamber the embryo is surrounded by a vascular membrane of its own: a specialized set of gills called bell gills.

Bell gills, unknown in other frog species, develop after several weeks of incubation. Beginning as films of tissue extending from each side of the embryo's head, in some species they merge and envelop the embryo completely, forming a fluid-filled sac in which the embryo floats. At birth the bell gills are either resorbed or shed.

Thus the mother and the embryos in her pouch are in intimate physiological contact. Only the egg's capsule of jelly separates the lining of the egg chamber from the embryo's bell gills, and in marsupial frogs the jelly layer is exceptionally thin. Whereas several millimeters of jelly surrounds the eggs of other frogs, the jelly on marsupial-frog eggs can be 1,000 times thinner—as thin as several microns.

In several respects this arrangement resembles pregnancy in mammals. The fluid-filled chamber in which the embryo develops is reminiscent of the sac of amniotic fluid that surrounds mammalian embryos. What is more important, the intimate contact between the blood supplies of the mother and the embryo through the juxtaposed membranes recalls the role of the mammalian placenta, the organ that couples the circulation of the mother and the embryo during pregnancy.

The placenta enables mother and embryo to exchange gases, fluids, nutrients and wastes. Presumably, similar traffic takes place between the two membranes in marsupial frogs. To see just what the embryos gain from the mother, Bertha Escobar and I weighed embryos taken from specimens of *G. riobambae* at various stages of incubation. The wet weight of the embryos increased threefold over the course of incubation, but their dry weight stayed relatively constant. It appeared that they had gained water and perhaps exchanged gases with the mother, but the nutrients sustaining development must have come mostly from the embryos' own stores of yolk.

What hormonal triggers prepare the pouch for incubation? In answering this question my colleagues Carmen de Albuja, Margarita Campos and I came on further parallels between incubation in egg-brooding frogs and pregnancy in mammals. In mammals several hormones take part in readying the uterus for pregnancy. Hormones called gonadotropins are released by the pituitary gland and stimulate the ovarian follicle, a capsule of cells that surrounds the growing oocyte (the cell that gives rise to the egg). In response the follicle liberates hormones of its own: estrogens and progesterone, which cause the lining of the uterus to thicken in preparation for receiving a fertilized egg.

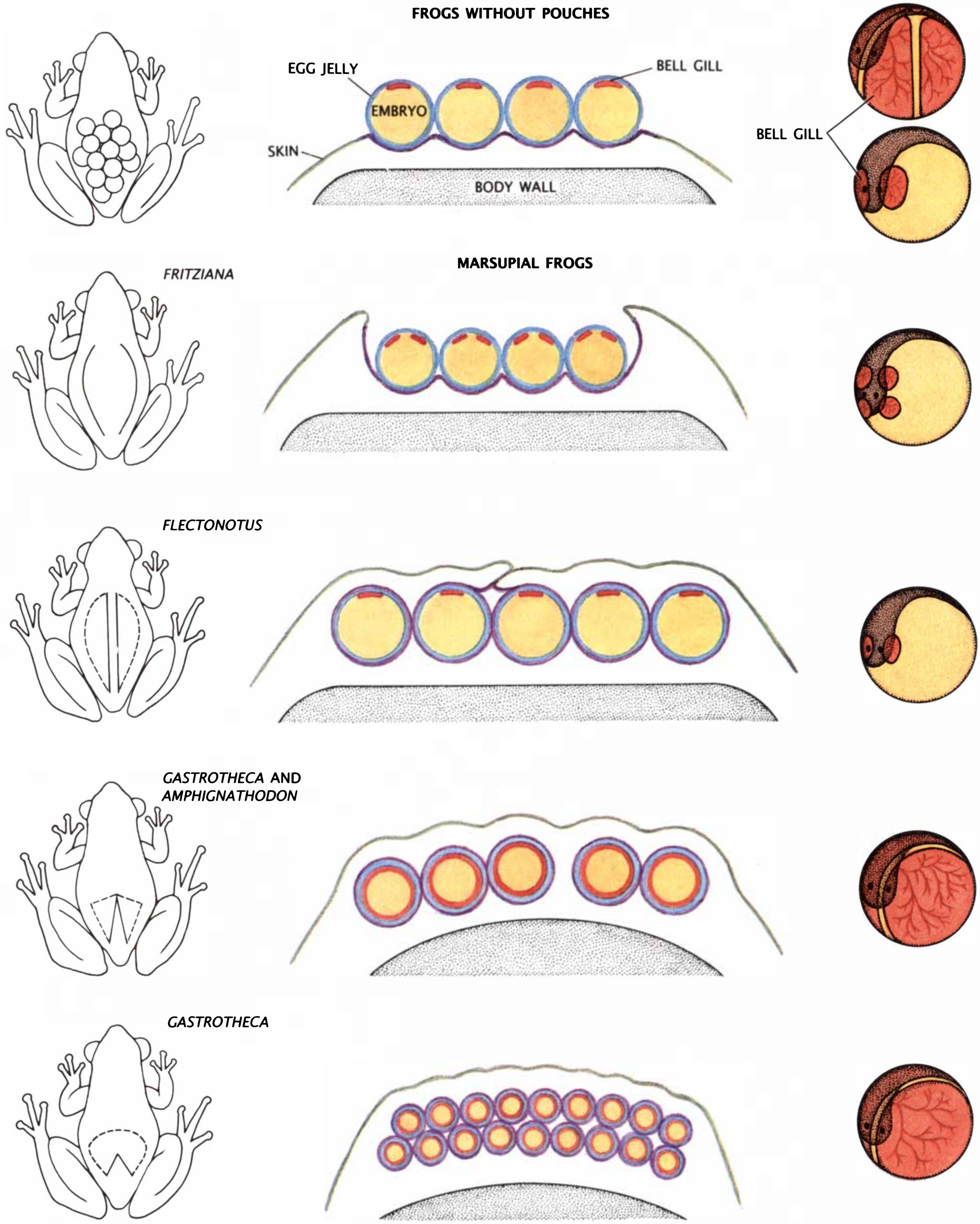
In most frogs a similar two-stage hormonal mechanism exists, but its main effect is to trigger the growth and maturation of oocytes. Gonadotropins stimulate follicles in the ovary to release estrogens and progesterone; the progesterone induces the series of cellular changes that transforms an oocyte into an egg. In egg-brooding frogs, we found, this hormonal apparatus has taken on the additional function of preparing the pouch for incubation.

One requisite preliminary can be seen in *G. riobambae* and other marsupial species as the breeding season approaches. Most of the year the opening of the pouch is dilated, but as the female prepares for reproduction the aperture constricts. The pouch remains closed through incubation, enabling it to retain the embryos until the tadpoles or frogs are born.

The hormonal control of pouch closure can be dissected by injecting individual hormones and observing their effects. If a frog is given an injection of gonadotropin, the pouch will close only if the ovaries contain large follicles. In contrast, a single injection of progesterone will cause the pouch to close within 24 hours, regardless of the follicle size. Apparently it is progesterone from the follicles that is directly responsible for triggering pouch closure; gonadotropins merely stimulate progesterone's release.

The development of the egg chambers, which envelop each embryo in a vascular membrane, also seems to be driven by progesterone. If plastic beads with the same diameter as embryos are placed in the pouch of a female when she is not incubating, the beads are ordinarily expelled within a few hours. If the frog is given progesterone at the same time as the beads are placed in her pouch, however, the pouch closes over the beads and they become encased in vascular egg chambers. The false embryos are retained for as long as a week.

By the time of incubation the follicles have already ruptured and released their eggs. What, then, is the source of the progesterone that fos-

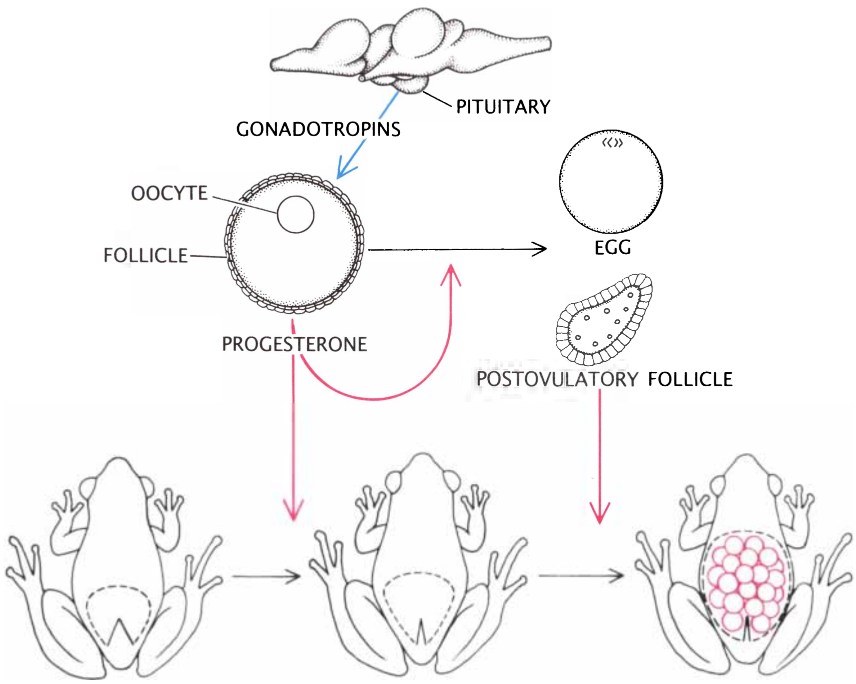


MOTHER AND EMBRYOS of egg-brooding frogs are in intimate contact. The shape of the pouch or the area of the back to which the eggs adhere is shown for each genus (*left*); vertical sections (*middle*) reveal the relation between the mother and the embryos. A richly vascular membrane (*purple*) enfolds each egg, separated only by a thin layer of egg jelly (*blue*) from the embryo's own vascular membranes, called bell gills (*red*). The form of the gills varies; in some genera they cover the embryo completely (*right*). In marsupial frogs the close communication between mother's tissues and embryo's bell gills enables them to exchange gases and fluids, just as mother and embryo communicate through the placenta in mammals.

ters the growth of egg chambers? In the early stages of pregnancy in mammals the empty follicle continues to make progesterone even after the release of the egg, thereby stimulating the development of the uterine wall. In most frogs, postovulatory follicles remain in the ovaries for less than a week and are not thought to have a hormonal role. Yet in *G. riobambae* empty follicles stay in the ovaries for more than a month, throughout the first third of incubation.

To find out whether these postovulatory follicles are the progesterone source, Grace Sánchez and I removed the ovaries (and with them any postovulatory follicles) from frogs that were incubating eggs. Removal of the ovaries during the first few weeks of incubation resulted in abortion: the partially developed embryos were expelled from the pouch within 15 to 30 days. Removal of the ovaries at a later stage of incubation had no effect. At least in its early stages, then, incubation in egg-brooding frogs seems to have hormonal underpinnings similar to those of pregnancy in mammals.

It may be that the parallels continue later in incubation. Hormonal support for the later stages of pregnancy in mammals comes from the placenta itself, which secretes gonadotropins, progesterone and other hormones. Perhaps the bell gills and pouch chambers in marsupial frogs play an analogous role, by releasing hormones that sustain advanced incubation.



SERIES OF HORMONES regulates reproduction in marsupial frogs, according to work by the author and her colleagues. They believe gonadotropins released by the pituitary gland, at the base of the brain, act in the ovary by stimulating the follicle, the cells surrounding the oocyte (the egg's precursor cell). In response the follicle releases progesterone, which stimulates the maturation of the oocyte and also causes the opening of the pouch to constrict. After the egg has been released from the ovary the empty, postovulatory follicle may continue to make progesterone, which triggers the development of the vascular chambers surrounding the embryos in the pouch. A similar hormonal apparatus readies the uterus for pregnancy in mammals.

I have sketched the adaptations that make it possible for the mother frog to brood the eggs internally. What has this prolonged incubation in the mother's body meant for the eggs of marsupial frogs?

Internal incubation and the leisurely embryonic development it makes possible have one obvious consequence in egg-brooding species: the eggs are exceptionally large and well stocked with yolk, more like the eggs of birds and reptiles than like the eggs of other frogs. The eggs of one egg-brooding species, *Gastrotheca cornuta,* are a centimeter across, and diameters of half a centimeter are common, in contrast to just over a millimeter in *Xenopus,* for example. The stores of yolk, fats and carbohydrates that account for the large size sustain the embryo during its long sojourn in the pouch.

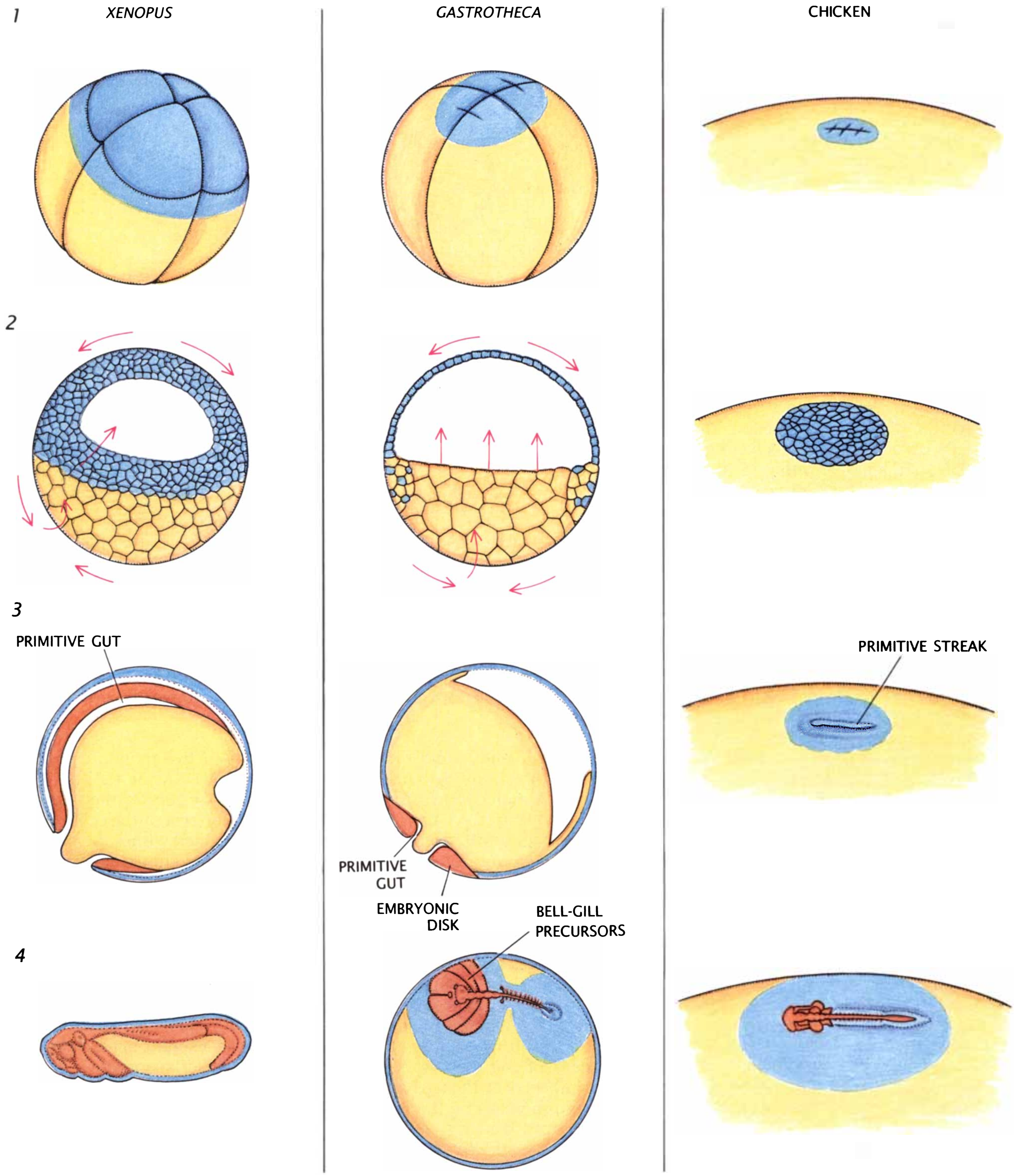
A second contrast between more conventional frog eggs and ones that develop in the mother's pouch is hidden in the molecular machinery of the oocyte, the precursor cell. The rapid development that characterizes most frog embryos demands rapid synthesis of proteins, the building blocks of body tissues. A molecular feature of most frog oocytes in effect gives them a jump on protein synthesis: the presence of large amounts of ribosomal RNA, or rRNA, as well as of other kinds of RNA molecules.

Ribosomal RNA is a key structural element in ribosomes, the cell organelles in which protein synthesis takes place. The large supply of rRNA in an oocyte of most frogs—some 200,000 times the amount in an ordinary cell of the frog's body—means that the cells of the embryo do not need to synthesize it themselves before making proteins, and the vulnerable embryo can develop and hatch quickly. Mammals, in contrast, whose embryo is protected in its mother's body and can develop slowly, store little rRNA in their oocytes.

The high levels of rRNA reflect processes in the frog oocyte's nucleus. Ribosomal RNA, like the other kinds of RNA in cells, is copied from the DNA of genes in the nucleus. Frog oocytes can accumulate high levels of rRNA because their ribosomal genes have been amplified, or duplicated many times. The oocyte of *Xenopus,* for instance, contains about two million copies of the ribosomal DNA, or rDNA. These genes do not reside on the chromosomes with the rest of the cell's genes but are segregated within the nucleus in compartments known as nucleoli—more than 1,000 of them in the nucleus of a full-grown *Xenopus* oocyte.

One measure of the rapid development that the enormous quantity of rRNA accumulated in the oocyte makes possible is the timing of gastrulation. Gastrulation is the stage in embryonic development at which the early embryo, a simple sphere of cells, folds in on itself to produce a new, double-layered sphere surrounding a cavity that is the precursor of the gut. In *Xenopus* gastrulation takes place only 14 hours after fertilization.

In *G. riobambae,* our benchmark egg-brooding frog, the lag is 14 days. I wondered whether this slower development reflects a lower level of ribosomal-gene amplification and a lower content of rRNA. The appearance of *Gastrotheca* oocytes suggested that it does. Michael F. Trendelenburg of the German Cancer Research Center in Heidelberg and I counted fewer than 300 nucleoli in oocytes of *G. riobambae*—perhaps a fourth of the number



EARLY EMBRYONIC DEVELOPMENT in the marsupial frog *Gastrotheca riobambae* has birdlike features. The diagrams, which compare stages in the development of *Gastrotheca* (*middle*) with the same stages in *Xenopus laevis* (*left*), a frog that reproduces conventionally, and in the chicken (*right*), are not all drawn to scale; the *Xenopus* embryo is one third the size of the *Gastrotheca* embryo. In *Xenopus* the initial cell divisions (*1*) produce a hollow sphere of cells, the blastula (*2*), which folds in on itself in a process called gastrulation: large, yolk-laden cells (*yellow*) from one end of the embryo invade the interior, and yolk-poor cells (*blue*) from the other end migrate to cover the embryo's surface. A new hollow, the primitive gut, is thereby created (*3*). The entire embryo goes into forming the body (*4*). In the large, yolk-filled egg of the chicken only a small disk of cells on the surface of the yolk mass undergoes gastrulation and gives rise to the body. In *Gastrotheca* early development departs from the usual frog pattern. The entire embryo takes part in gastrulation, as in *Xenopus*. Probably because of the embryo's large size and high yolk content, however, the cells taking part in later development become concentrated on one side, in a disk surrounding a small primitive gut. As in the chicken, the body takes form on the surface of a yolk mass.

in *Xenopus* oocytes. One would expect the smaller number to signal a lower level of ribosomal-gene amplification and hence a smaller supply of rRNA.

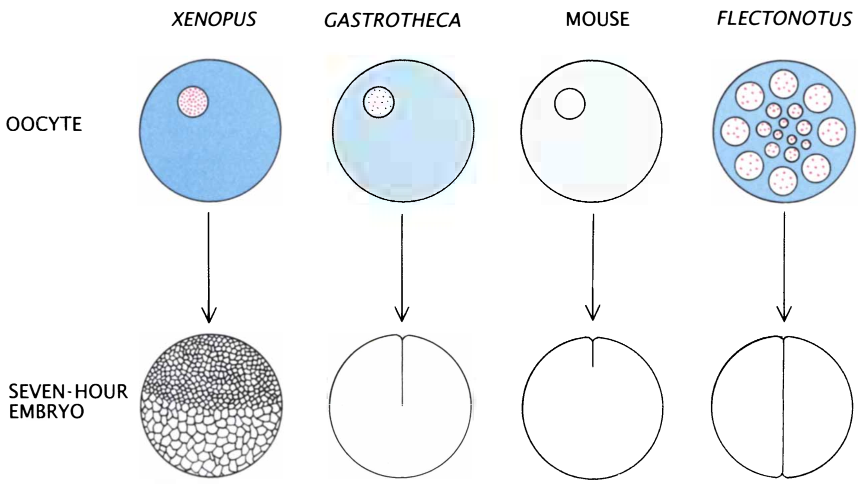
We turned to molecular means to measure the level of ribosomal-gene amplification in *Gastrotheca* oocytes. Christine Dreyer of the Max Planck Institute for Developmental Biology in Tübingen and I exposed the oocytes to a probe consisting of cloned rDNA from *Xenopus.* (Single-strand DNA from one organism can bind to and identify corresponding sequences in the DNA or RNA of other organisms.) The probe detected little rDNA.

To find out whether the low level of gene amplification indeed results in a smaller stock of rRNA, Peter Hausen of Tübingen and I applied a stain specific for RNA to oocyte sections. The staining was faint in *Gastrotheca,* vivid in *Xenopus.* By isolating RNA from the oocytes and then applying an rDNA probe, Herbert Steinbeisser and Ansgar Hofmann of the German Cancer Research Center and I measured the rRNA content directly. We found that *Gastrotheca* oocytes do contain much less rRNA than oocytes of *Xenopus.*

Even at the molecular level, then, certain developmental features of marsupial frogs are not froglike but are like those of mammals. Early in my studies, however, I had discovered a feature of some marsupial-frog oocytes that is practically unknown not only in other frogs but also in other vertebrates generally. In 1976 the late Scott J. Maness of the Rancho Grande Biological Station in Venezuela introduced me to the fascinating marsupial frog *Flectonotus pygmaeus.* No more than an inch long, this inhabitant of the rain-forest canopy incubates between five and 11 eggs in its pouch and gives birth to tadpoles, which it deposits in the pools of rainwater that collect in arboreal plants.

The remarkable aspect of *Flectonotus* was revealed when I opened one of its oocytes under the microscope. Instead of a single nucleus, thousands of spherical structures spilled out. By examining the structures and studying their biochemical activity, my former professor Asa A. Humphries, Jr., then at Emory University, and I positively identified them as active oocyte nuclei. We counted an average of 2,000 nuclei in each oocyte! At the time only one other vertebrate, the tailed frog of North America, was known to have more than a single nucleus per oocyte, and it has just eight. What purpose could so many nuclei serve?

A clue lay in the rapid development



LEVEL OF RIBOSOMAL RNA (*blue*) in the oocyte is correlated with the speed of embryonic development. In *Xenopus*, a frog with aquatic eggs, rRNA is abundant because the ribosomal genes have been amplified (*red*), and the vulnerable embryo develops rapidly. Seven hours after fertilization it has already divided into some 4,000 cells. In the marsupial frog *Gastrotheca*, as in mammals such as the mouse, development can be leisurely because the embryo is protected. The level of rRNA is low, reflecting little or no ribosomal-gene amplification; after seven hours the embryo still has only one or two cells. In some egg-brooding frogs, such as *Flectonotus*, the oocyte has many nuclei, which add more rRNA and speed development.

of *Flectonotus* embryos. The embryos are incubated for only 20 days, and the tadpoles develop into frogs after 11 more days. This accelerated development presumably requires a large stock of rRNA. It is conceivable that *Flectonotus* is descended from a typical egg-brooding species, in which embryonic development had slowed and the amplification of ribosomal genes had been reduced. Faster development might have been favored by *Flectonotus'* small size and shortened life span. The incorporation of thousands of nuclei into each oocyte, each nucleus having limited amplification of the ribosomal genes, was the evolutionary answer to the need for more rRNA. The strategy was effective, as Herbert C. Macgregor of the University of Leicester and I discovered by measuring the overall level of amplification.

It turned out that—unusual as they are in other organisms—multinucleated oocytes are quite common among egg-brooding tree frogs: we found them in 14 of the 36 species we examined. In each case the adaptation presumably arose in response to an environmental pressure that favored faster embryonic development. Yet the feature raises many questions. For example, as a multinucleated oocyte develops in the ovary, the number of nuclei drops precipitously until there remains only one, which will unite with the nucleus of the sperm during fertilization. How is this purge of nuclei controlled? How can one nucleus among thousands be spared?

Whether it develops from an oocyte with one nucleus or many, the egg of a marsupial frog is structurally and molecularly distinctive. One might expect these differences to affect the pattern of embryonic development—the choreography of cell growth and movement that begins once the egg is fertilized and transferred to the pouch. Richard P. Elinson of the University of Toronto and I studied early embryonic development in *Gastrotheca*. Here too we found departures from the usual pattern in frogs and resemblances to the pattern in other animals, specifically birds and reptiles.

Striking anomalies appear at gastrulation, when the primitive embryo folds inward and cells from one side of the sphere migrate around to cover the resulting hollow—the primitive gut. In most frog embryos the primitive gut extends through much of the gastrula. Thus the original body plan and later embryonic development involve the entire embryo. In birds and reptiles, however, the large quantity of yolk that fills the early embryo restricts gastrulation. Only a small disk of cells concentrated on one side of the egg takes part in the movements that establish the primitive gut, and only those cells go on to form the body. The rest of the embryo consists of a large, undivided yolk mass that is used up by the developing embryo.

Elinson and I found that in egg-brooding frogs, as in other frogs, the entire embryo takes part in gastrulation: beginning at one end, small cells containing little yolk work their way around to create a new cell layer covering the entire mass. The infolding that normally takes place during gastrulation is limited, however, and the primitive gut it creates is confined to a small cavity at one end of the embryo. Only the disk of cells surrounding this small primitive gut engages directly in later embryonic development. The developing embryo is splayed on one side of a large yolk mass, like the early embryos of birds and reptiles.

Hundreds of millions of years of evolution separate marsupial frogs from the fully terrestrial animals—mammals, birds and reptiles—with which I have compared them. And yet the frogs have come up with many similar answers to the adaptive challenge of reproducing on land. It is amusing to speculate about the evolutionary potential of egg-brooding tree frogs: Could they in time give rise to a completely new class of terrestrial vertebrates? It would not be the first time amphibians had conquered the land: mammals, birds and reptiles can all trace their ancestry to ancient amphibians.

Yet it is more likely that for many egg-brooding species, at least, the evolutionary adventure is ending. Their habitat, the tropical forest, is vanishing, taking with it many unusual creatures that inhabit the canopy. The less known species of egg-brooding tree frogs may be swept away before their secrets are fully revealed.

FURTHER READING

MULTIPLE NUCLEI DURING EARLY OOGENESIS IN *FLECTONOTUS PYGMAEUS* AND OTHER MARSUPIAL FROGS. Eugenia M. del Pino and A. A. Humphries, Jr., in *The Biological Bulletin*, Vol. 154, No. 2, pages 198–212; April, 1978.

A NOVEL DEVELOPMENT PATTERN FOR FROGS: GASTRULATION PRODUCES AN EMBRYONIC DISK. Eugenia M. del Pino and Richard P. Elinson in *Nature*, Vol. 306, No. 5943, pages 589–591; December 8, 1983.

OOGENESIS IN THE EGG-BROODING FROG *GASTROTHECA RIOBAMBAE* PRODUCES LARGE OOCYTES WITH FEWER NUCLEOLI AND LOW RNA CONTENT IN COMPARISON TO *XENOPUS LAEVIS*. E. M. del Pino, H. Steinbeisser, A. Hofmann, C. Dreyer, M. Campos and M. F. Trendelenburg in *Differentiation*, Vol. 32, pages 24–33; 1986.

For the cost of a morning coffee break, you can break the cycle of poverty for one small child.

It may not seem as though 72¢ will go very far these days. But for a small child like this, it can lead her out of desperate poverty and into a future of hope.

And right now, you can reach out and make that future happen...by sponsoring one child overseas through Foster Parents Plan.

For only 72¢ a day, you can make it possible for your Foster Child and family to have adequate food and clothing, a decent place to live, needed medical care, education... *and the chance for a better life.* Because while you're helping, the child's own family will be working with Foster Parents Plan to develop the skills and resources they need to help themselves.

You'll be amazed at the difference you can make. And you'll see the difference for yourself. In photographs. Detailed progress reports. And heartwarming letters from your Foster Child. You'll be doing so much for that child...and for so little.

Won't you help now? Simply mail the coupon today or call toll-free. It could be the biggest break a child ever got.

Foster Parents Plan®

Help so complete, it touches a child for life.

Founded in 1937, Foster Parents Plan is the leader in combining family assistance programs, community development, and personal communications between Foster Parent and Foster Child. These comprehensive programs provide long-term solutions to the unique problems facing Foster Children and their families. As a non-profit, non-sectarian and non-political organization, we depend on Foster Parents to make our work possible. A copy of our financial report, filed with the New York Department of State, is available on request. Your sponsorship is tax-deductible.

To start helping even faster, call toll-free:

1-800-543-2400

Detach and mail or call toll-free today. (N6)

Send to:
Foster Parents Plan, Inc.
P.O. Box 877, Dept. K1 Providence, RI 02901

I want to become a Foster Parent to:

☐ **The child who has been waiting the longest,** or as indicated:

☐ Boy ☐ Girl ☐ Either

☐ Any age ☐ 6–8 ☐ 9–11 ☐ 12–14

☐ Bolivia	☐ Guatemala	☐ Kenya	☐ Philippines
☐ Burkina Faso	☐ Honduras	☐ Liberia	☐ Sri Lanka
☐ Colombia	☐ India	☐ Mali	☐ Sudan
☐ Ecuador	☐ Indonesia	☐ Nepal	☐ Thailand

☐ Enclosed is a check for $22 for my first month's sponsorship of my Foster Child. Please send me a photograph, case history, and complete Foster Parents Sponsorship Kit.

☐ I am not yet sure if I want to become a Foster Parent, but I am interested. Please send me information about the child I would be sponsoring.

☐ Mr. ☐ Mrs.
☐ Miss ☐ Ms.

Address Apt.#

City State Zip J 366

Optical Fibers in Medicine

Instruments to view internal anatomy plainly, sensors to analyze blood reliably and laser systems to perform surgery inside the body safely—all rely on fiber-optic technologies

by Abraham Katzir

As optical fibers transform communication, they are also revolutionizing imaging, diagnosis and therapy in medicine. These ultrathin, flexible fibers have opened a window into the living tissues of the human body. By inserting optical fibers through natural openings or small incisions and threading them along the body's established pathways, physicians can peer into the airways of the lungs, the folds of the intestine, the chambers of the heart and many other areas that were previously inaccessible. By placing fiber-optic sensors in the bloodstream, physicians can do rapid, reliable chemical analyses at the bedside, in the examining room or in the operating room—analyses that otherwise involve taking blood samples from patients for examination in laboratories. By directing a beam of laser light down an optical fiber, physicians can even perform surgery inside the body, sometimes eliminating the need for invasive procedures in which healthy tissue must be cut through to reach the site of disease.

By delivering laser light through optical fibers, for example, gastroenterologists have cauterized vessels to stop intestinal bleeding, vascular surgeons have begun to vaporize plaque and blood clots in peripheral arteries, and neurosurgeons may soon bond nerves together in the brain and spinal cord. Fiber-optic devices may soon combine diagnosis and treatment—for instance by incorporating both a means of detecting cancer cells and a means of destroying the cancer without damaging nearby healthy tissue.

Many fiber-optic procedures for diagnosis and treatment do not require anesthesia and can be performed safely in a physician's office; the continued development of fiber-optic techniques should therefore reduce the risk and cost of medical care. Fiber-optic procedures may also be applied in cases where invasive surgery is dangerous or impossible, as it sometimes is in the young or the elderly.

ABRAHAM KATZIR is professor of physics and director of the applied physics group at Tel Aviv University in Israel. He received B.Sc., M.Sc. and Ph.D. degrees from the Hebrew University of Jerusalem. He has spent several years as a visiting scientist at the California Institute of Technology, at the AT&T Bell Laboratories and as a visiting professor at Boston University and the Massachusetts Institute of Technology. His group at Tel Aviv develops infrared optical fibers and systems for medical and industrial applications. His father, the late Aharon Katzir-Katchalsky, and S. Lifson wrote "Muscle as a Machine" for SCIENTIFIC AMERICAN in March, 1954. Following his family's tradition, Katzir has been "bridging the gap between physicians and scientists."

The first medical applications of optical fibers were imaging systems, called fiberscopes. Basil I. Hirschowitz and Lawrence Curtis of the University of Michigan School of Medicine constructed the first fiberscope, meant for viewing the stomach and esophagus, in 1957. Since then the devices have been so refined that they can be used to inspect virtually every organ system of the body. Indeed, most of the optical fibers used in medicine are incorporated into fiberscopes. The modern fiberscope consists of two bundles of optical fibers. One, the illuminating bundle, carries light to the tissues, and the other, the imaging bundle, transmits the image to the observer.

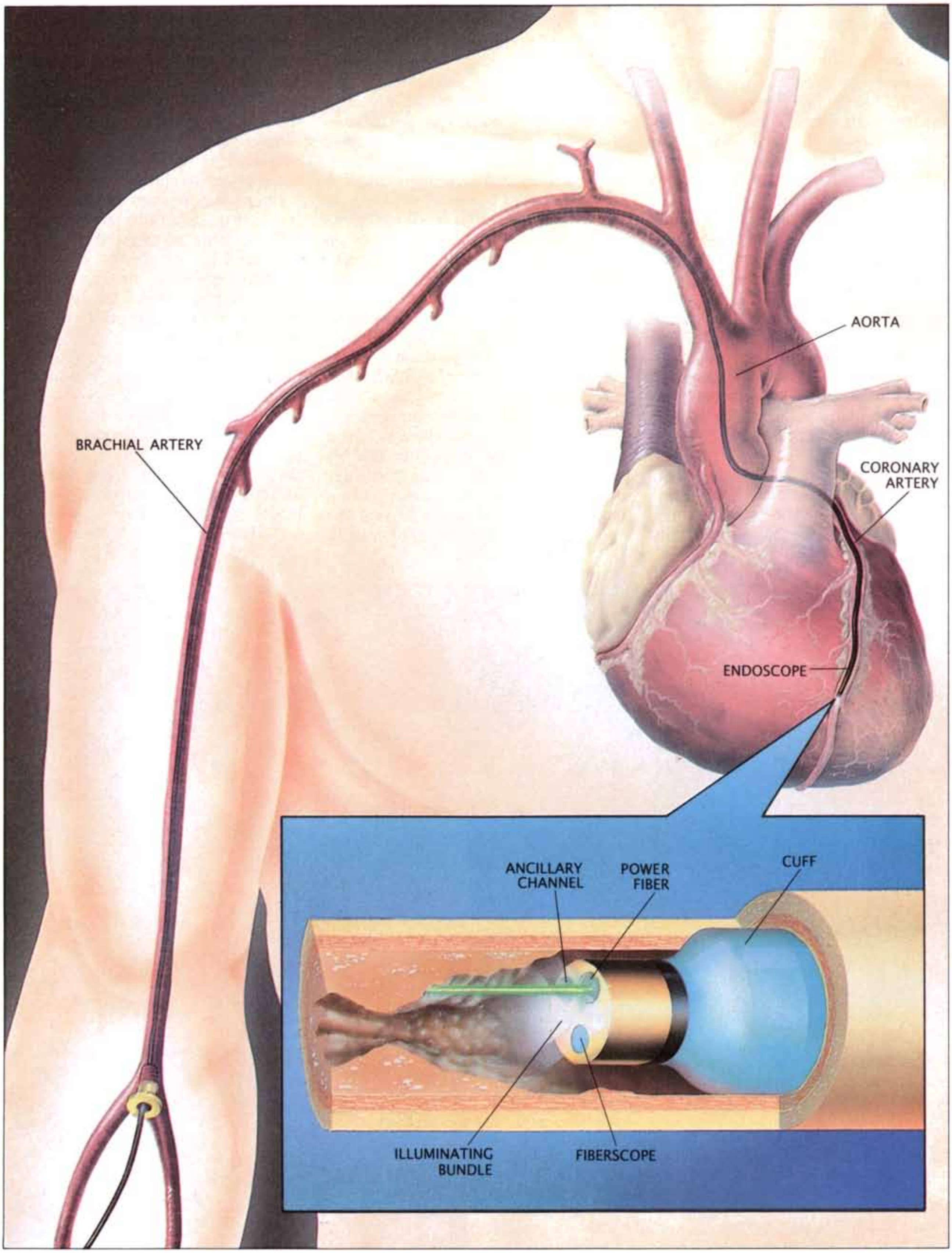
The illuminating bundle is coupled to a high-intensity light source such as a xenon arc lamp. Light enters the cores of the fibers, which are made of high-purity silica glass. Such glass attenuates light 10,000 times less than a windowpane and so can carry light over a distance of many kilometers. Because light tends to stray from the core, the core is wrapped in a cladding that reflects most of the stray light back into the core. This is the mechanism by which light is guided through all optical fibers [see "Communication by Optical Fiber," by J. S. Cook; SCIENTIFIC AMERICAN, November, 1973].

A lens collects the light that is reflected off tissue and focuses it onto the receiving end of the imaging bundle. Each optical fiber in this bundle admits only the reflected light that is aligned with its axis, and so each fiber transmits only a small fraction of the total image. The fibers are glued together just at the ends to allow flexibility and yet avoid scrambling the image. The reconstructed image can be viewed through an eyepiece, recorded by a camera or displayed on a television screen. Since thousands of fibers can be combined in a single bundle with a diameter of less than a millimeter, a fiberscope can carry images that have high spatial resolution and virtually perfect chromaticity.

The illuminating and imaging bundles can easily fit in a catheter a few millimeters in diameter. This fiberscope can then be inserted through orifices in the body and brought to focus on tissues at distances of from five to 100 millimeters from the tip.

Fiberscopes are often part of larger instruments, called endoscopes, that include ancillary channels through which physicians can perform other tasks. Through one channel, for example, fluids can be drawn or water or air injected to clear debris away, thereby improving visibility. Another channel may contain fine wires that can be manipulated to angle the endoscope's tip. A third may make it possible to insert tiny scalpels for cutting tissue or snares for removing it or needles for injecting drugs. Most endoscopes on the market are from .3 meter to 1.2 meters in length and from 2.5 to 15 millimeters in diameter.

With these devices physicians can gain intimate views of the digestive, reproductive, circulatory and respiratory systems. They can remove small



OPTICAL FIBERS can deliver beams of laser light into the coronary arteries to vaporize plaque. One device that may be developed in the near future consists of a fiberscope, inflation cuff and power fiber. A physician would insert the device into the brachial artery of the arm and direct it into the coronary artery. The fiberscope would enable the physician to see plaque or other blockages (*inset at bottom*). The cuff could then be inflated to stop the blood flow temporarily; a beam of laser light sent through the power fiber would vaporize the plaque. Deflating the cuff would then reestablish the flow of blood.

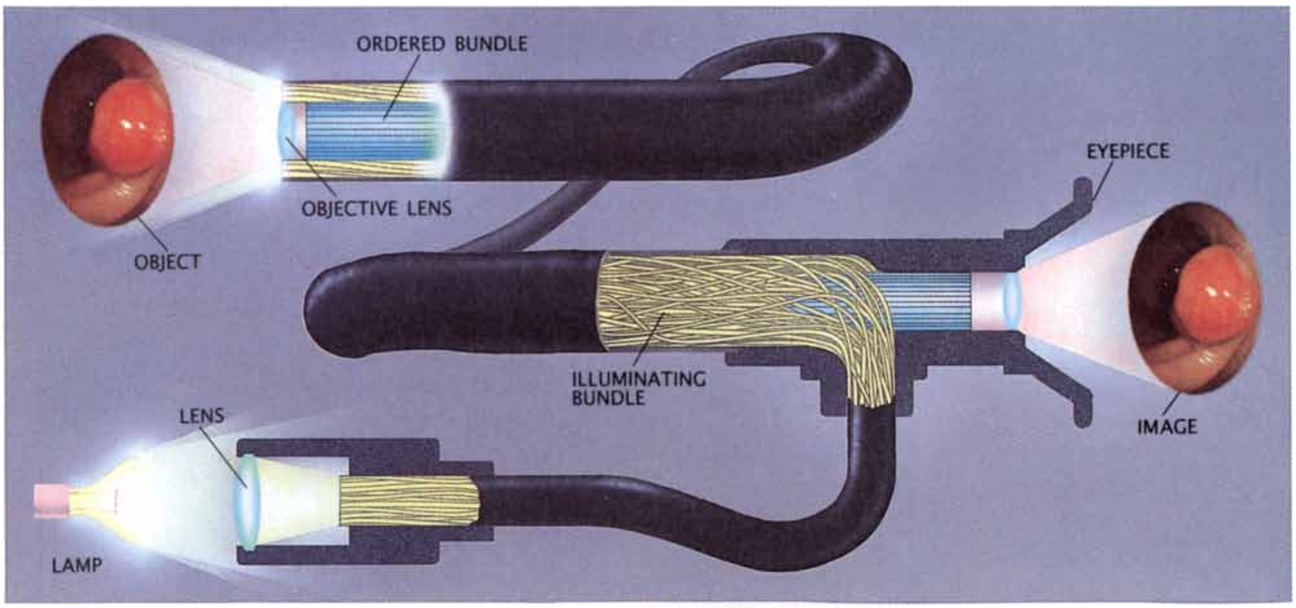
samples of tissue for laboratory analysis and even perform surgery. Peering through fiberscopes, physicians can detect a polyp in the colon, a foreign object in the lungs or a tumor in the esophagus and then remove it with miniature surgical instruments.

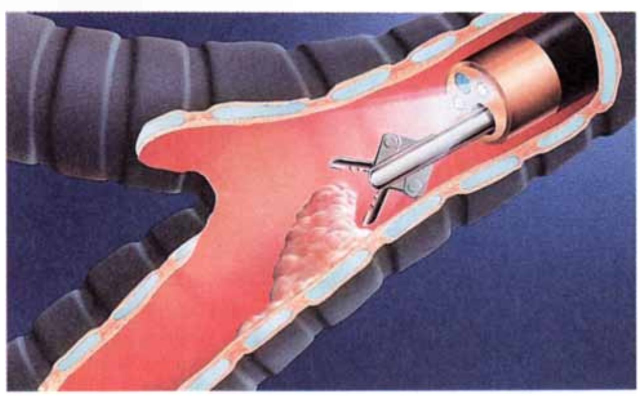
During the past five years the fabrication of ultrathin fibers has led to a reduction in fiberscope diameter and an increase in the number of fibers within the imaging bundle, which in turn has increased resolution. The newest fiberscopes incorporate 10,000 fibers in a bundle less than one millimeter in diameter and can resolve objects 70 microns across. These fiberscopes, inserted through an artery in the arm, can convey images of heart valves as well as obstructions in the coronary arteries, the vessels that carry blood to the heart.

In addition to providing images optical-fiber systems can offer direct and immediate chemical and physical analysis of the blood and other aspects of human physiology. The basic sensing system consists of an optical fiber inserted through a catheter into the body. The external end of the fiber is coupled to a light source and an optical processor, a device for analyzing the reflected light. The other end of the fiber sends light directly into the body or to a miniature sensor called an optode. The reflected light is carried back through the fiber to the processor, which extracts information about physiological conditions from the wavelength and intensity of the reflected light.

In many cases these fiber systems may be more sensitive, reliable and cost-effective than traditional methods, in which fluids are drawn from the body and tested in the laboratory. The fiber systems eliminate delays and reduce the chance of error. Furthermore, the fiber-optic sensors are compatible with body chemistry (for example, they will not provoke a response from the immune system), and in principle the optical fiber itself may be disposable. Fiber systems also appear to be more durable, flexible and potentially safer than the microelectronic devices that have also been developed to collect data from within the body. Some of the fiber-optic devices are already available on the market, whereas many other devices are being tested in clinical trials.

Fiber-optic measurements of blood flow rely on light reflected from blood cells. A fiber is inserted through a catheter into an artery. Low-intensity laser light sent through the fiber strikes red blood cells. When the light scatters off the moving cells, its wavelength changes because of a phenomenon known as the Doppler effect. The faster the cell is moving toward the tip of the fiber, the shorter the wavelength of the scattered light will be. Some of this light will be transmitted back through the fiber. The processor at the other end of the fiber then will determine the difference in wavelength between the laser light and the scattered light and will compute the





FIBERSCOPE (*above*) can convey images of the stomach and many other organs. A lens focuses light from a lamp onto a bundle of optical fibers. The light transmitted through the fibers illuminates a polyp in the stomach. An objective lens focuses the light that reflects off the polyp onto an ordered bundle of fibers. Each fiber carries a fraction of the total image. As the fibers of the ordered bundle release their light from their external end, the image is reformed and may be viewed through an eyepiece. The photographs of the polyp (*insets above*) were taken through fiberscopes at the Mayo Clinic. Fiberscopes are often incorporated into larger devices called endoscopes (*right*), which include channels through which other instruments can be passed. Here tissue in a bronchial tube is illuminated for viewing while the tissue is removed by means of a surgical instrument inserted through the endoscope.

velocity of the blood in the region of the tip.

The first experiments demonstrating this technique were reported almost three decades ago, but only in recent years have clinical trials begun. The instantaneous measurements of blood velocity that a fully developed instrument will provide will help physicians determine whether an adequate blood supply is reaching vital organs.

Optical fibers also make it possible to determine directly the oxygen content of the blood. Hemoglobin (the chemical responsible for oxygen transport in the blood) reflects much more red light when the compound carries oxygen than when it does not. Infrared light, however, is reflected equally from all hemoglobin regardless of its oxygen content. If red and infrared light are transmitted into the blood through optical fibers, the intensity of reflected red light reveals the amount of oxygen-carrying hemoglobin, whereas the intensity of infrared light measures the total amount of hemoglobin. This technique, now applied routinely, helps to reveal the capacity of a patient's blood to carry oxygen or the ability of the heart and lungs to supply oxygen.

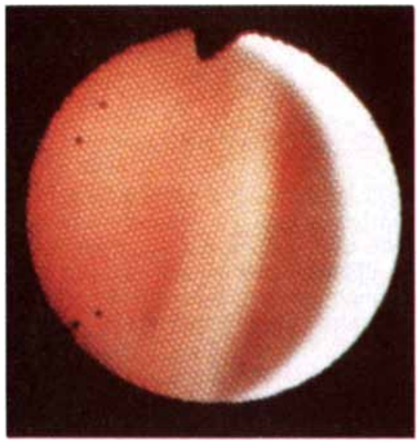
Miniature sensors connected to the ends of optical fibers have made numerous other physiological measurements possible [*see illustration on next page*]. They have been devised to measure pressures in the arteries, bladder, urethra and rectum. A sensor consists of a small tube at the end of an optical fiber, the far end of the tube sealed by a thin reflective membrane. If the pressures inside and outside the tube are equal, the membrane will stay flat, and light conveyed through the fiber will be reflected straight back. If the pressure is greater outside the tube than inside, the membrane bends inward, creating a convex mirror that reflects less light back through the fiber. If the pressure is lower outside the tube than inside it, the membrane bends outward, and the concave surface focuses more light onto the fiber. In several clinical experiments a good correlation was found between the results obtained from such an optical device and those from conventional instruments.

Sensors that measure blood chemistry have also been developed. Ones for measuring the acidity or alkalinity of blood (its *p*H) are based on organic dyes that luminesce under ultraviolet radiation. A change in *p*H causes some of these dyes to luminesce more brightly and the luminescence of other dyes to change color. These dyes, encased in a porous polymer, are attached to the end of a fiber. The polymer allows hydrogen ions to enter but prevents larger ions from interacting with the dye and degrading it. In practice this sensor can measure the *p*H of blood to within .01 *p*H unit.

Sensors of similar design incorporate biomolecules such as enzymes or antibodies. A change in the fluorescence of these biomolecules can reveal the presence or absence of specific chemicals in a rather complex mixture such as blood or biological tissue. A fiber would transmit information from the sensor to a processor, which would interpret the information to provide highly specific measurements of a large variety of chemicals inside the body. These sensors have already been employed to monitor levels of glucose and penicillin and may soon be used to measure therapeutic and metabolic substances, hormones, toxins and microorganisms.

In recent years the most significant application of optical fibers in medicine has been the delivery of laser energy inside the body for surgery and therapy. How laser radiation interacts with human tissue depends on the wavelength and intensity of the radiation. Although a particular laser emits light of a single wavelength, or color, a variety of lasers are now available that generate light throughout the visible, infrared and ultraviolet spectrum. Light is absorbed by tissue to a degree that varies depending on the wavelength and on the tissue's chromophores—coloring agents such as hemoglobin, melanin and keratin. Therefore, a laser of a particular wavelength targets specific tissues and produces specific photochemical reactions.

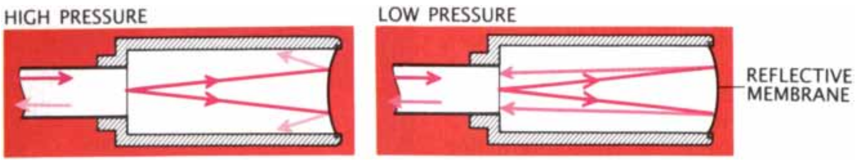
In general low-power lasers cause a gentle, local heating that coagulates blood and causes proteins to congeal. In this way laser light can bond soft tissues together and thereby seal wounds or join blood vessels. High-power lasers ablate tissue, in most cases by boiling water away. Such a beam also cauterizes incisions, so that a minimum amount of blood is lost during operations. Surgical applications of lasers require from 10 to 100 watts of continuous laser power or pulsed laser power that reaches peaks of 10,000 to one million watts. Since this power is discharged onto areas smaller than a square millimeter, the power densities are roughly equal to those a "Star Wars" laser would need for penetrating the skin of a guided missile in boost phase.



RIGHT VENTRICLE of a human heart was photographed through an ultrathin fiberscope, inserted into an artery in the upper arm. The fiberscope, which is less than a millimeter in diameter, was developed at Olympus Corporation in Tokyo.

The challenge of delivering this power into the body through optical fibers has occupied the workers in this field for the past two decades. All optical fibers attenuate light to some extent because of scattering and absorption. Some of the attenuation depends on the intrinsic properties of the fiber material; this intrinsic attenuation varies with the wavelength of light and the total power transmitted. Laser-beam energy can also attenuate as a result of scattering off the fiber's surface or defects inside the fiber. All these effects limit the power levels that can be delivered to the operating site. If the laser input power is increased beyond a critical level, the fiber ends may overheat, melt or vaporize. In recent years many of these difficulties have been overcome by the discovery of new optical-fiber materials and the development of methods for fabricating high-purity fibers from these materials.

In contrast to the silica-glass fibers used in imaging and diagnostic systems, fibers designed for laser surgery are made of rather exotic materials. Quartz fibers have been developed to transmit the green light of argon lasers, the ultraviolet light of excimer lasers and the near-infrared radiation from YAG lasers (named for the neodymium, doped yttrium-aluminum-garnet crystal that is the lasing material). Such fibers have transmitted up to 100 watts of continuous laser energy,



FIBER-OPTIC SENSOR can measure pressures in the arteries, bladder and uterus. The sensor consists of a tube at the end of a fiber, the far end of the tube sealed by a reflective membrane. If the pressure is higher outside the tube than inside (*left*), the membrane bends inward, creating a convex mirror that reflects little light back through the fiber. If the pressure is lower outside the tube than inside (*right*), the membrane bends outward; the concave surface focuses more light onto the fiber.

enough for most surgical purposes. Investigators are still searching, however, for a fiber that can efficiently transmit the radiation of infrared lasers, carbon dioxide lasers in particular [see "Infrared Optical Fibers," by Martin G. Drexhage and Cornelius T. Moynihan; SCIENTIFIC AMERICAN, November, 1988]. Currently the best infrared optical fibers for this application are polycrystalline fibers made from metallic, halide crystals.

The first medical device incorporating lasers and optical fibers was designed in 1973 to control bleeding from peptic ulcers. The device consisted of a fiberscope for observing the procedure and a separate power fiber for delivering the laser energy that cauterized the ulcer. Since then these devices have successfully treated bleeding ulcers in the stomach, intestine and colon. A laser-transmitting fiber has also served for shattering renal stones. The most interesting prospects for the therapeutic use of laser fiber systems, however, lie in the treatment of cardiovascular diseases and localized tumors.

Many of the most devastating diseases of the cardiovascular system occur when arteries become blocked by calcified, fibrous fat deposits, known as atherosclerotic plaque, and by blood clots. If these obstructions cut off circulation, they cause strokes, heart attacks and gangrene in the limbs. To treat a patient with obstructions in the coronary arteries, a physician today might first resort to a method called percutaneous transluminal coronary angioplasty. This technique relies on a special catheter that has a tiny balloon attached to its end. If an artery is only partially occluded, the tip of the catheter can be inserted through the constricted region. The balloon is inflated so that it compresses and reduces the occlusion. This method is not applicable for treating complete occlusions, and beneficial results may only be temporary.

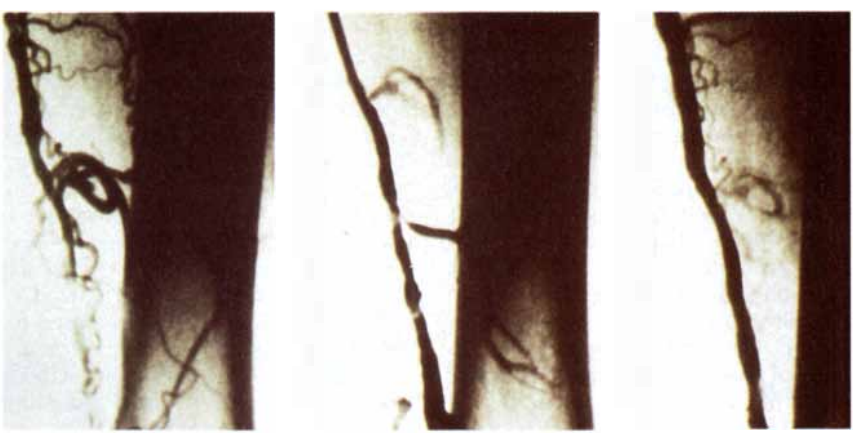
A physician might then recommend a more radical surgical procedure known as a coronary artery bypass: a vein is removed from the leg, the chest cavity is opened to expose the heart and the vein is implanted to carry blood around the blocked artery. Although the bypass operation usually reestablishes blood flow to the heart, it is a traumatic procedure, which involves a lengthy recovery period and great expense.

The development of optical fibers capable of delivering significant laser energy has now made available several new techniques for removing arterial obstructions, known as laser angioplasty. In one set of techniques the end of a fiber is covered with a small metal tip. If the fiber is inserted into an obstructed artery and laser light sent through the fiber, the metal tip will heat up and melt the obstruction. Very careful control of treatment is needed in this technique. Otherwise the hot tip can stick to the arterial wall or even perforate it. Several successful laser-angioplasty instruments are now in use.

An alternative that is potentially more effective but technically more complex is a system in which laser light ablates the plaque directly. The first experimental systems of this type included an argon laser, which emits green light. This laser was chosen because of its reliability and because its light could be efficiently transmitted through conventional quartz fibers. Experiments revealed, however, that the absorbed green light caused extensive thermal damage to neighboring tissue and did not remove plaque efficiently. These problems could be overcome by means of a pulsed ultraviolet or infrared laser beam. Recently Warren S. Grundfest, James S. Forrester and Frank Litvack of the Cedars-Sinai Medical Center in Los Angeles tested a system based on an ultraviolet excimer laser and a quartz fiber. Their system has successfully cleared blockages of the coronary arteries in several patients.

The major problems still to be solved are how to manipulate the laser beam inside the artery and how to distinguish between normal and diseased segments of an artery before using the laser to vaporize tissue. In the future "smart" systems, which will rely on endoscopic fluorescence, will perform this task and will prevent blood vessel perforation.

In the near future my colleagues and I at Tel Aviv University believe that new optical fibers capable of transmitting the infrared radiation from a carbon dioxide laser may lead to a safer, more durable and reliable system. We tested such a system by implanting human arteries that were blocked by atherosclerotic plaque into animal hosts. The fibers were then inserted into the arteries, and a carbon dioxide



X-RAY PHOTOGRAPHS of the femoral arteries show the success of laser and balloon catheter angioplasty. In the X ray at the left, the artery is completely blocked. After laser light is applied to the blockage, only 30 percent of the original occlusion remains (*center*). To clear the remaining blockage, a balloon catheter can be inflated in the artery. When the catheter is removed, normal blood flow is restored (*right*).

TUMOR obstructing the trachea (*left*) is treated by photodynamic therapy. The patient is injected with a dye that the tumor absorbs more readily than healthy tissues do. An optical fiber embedded in the tumor (*center*) conveys laser light, which alters the dye chemically. As is apparent two days later (*right*), the altered dye has killed the tumor, so that it can be removed.

laser beam was used to reopen the arteries.

The state of the art suggests that within a few years such surgical instruments will be refined and incorporated into a device that includes a fiberscope and sensors. This laser endoscope will probably have a diameter of less than two millimeters; the fiberscope will take up about half the space. The device will also carry a power fiber for transmitting infrared, ultraviolet or visible laser light. Fiber-optic sensors would serve for measuring blood pressure, temperature and flow; another channel would allow fluids and gases to be pumped through the endoscope.

The endoscope could be inserted into the coronary artery, just as a regular catheter is. The physician would look at the blockage and measure pressure and blood flow. The blood flow would be stopped with a balloon; clear saline solution injected through the irrigation channel would clear away the remaining blood. A beam of laser light would then be sent through the power fiber to vaporize the blockage. Gases released from vaporization of the plaque would be pumped out through the aspiration channel or removed by the body. The fiber-optic sensors, connected to a computerized processor, would monitor the procedure and warn of overheating of the fiber tip or ablation of a healthy blood vessel. At the completion of the angioplasty the balloon would be deflated and the blood flow measured to confirm that the coronary artery had been reopened.

Another emerging application is the use of optical fibers for the detection and treatment of small malignant tumors. A diagnostic method known as fluorescence endoscopy has been very successful in detecting lung tumors that are too small to be found by computerized axial tomographic (CAT) scans or by chest X rays. If hematoporphyrin derivative, a dye that fluoresces red under ultraviolet radiation, is injected into a patient, over the course of several days, tumors will absorb much more hematoporphyrin derivative than healthy tissues. If the suspect region is illuminated with a suitable ultraviolet source, such as a quartz fiber coupled to a krypton laser, malignant tissue reveals itself by emitting red light. To detect the red light, a filter that transmits it and blocks reflected ultraviolet light is attached to the imaging bundle.

If the tissue is bathed in sufficiently intense red light instead of ultraviolet light, the results are strikingly different. Hematoporphyrin derivative absorbs red light strongly. The absorbed energy causes a series of photochemical reactions that kills the malignant tissues that have been saturated with hematoporphyrin derivative. A gold-vapor laser can provide the high-intensity red light, which is transmitted through a quartz fiber and delivered directly inside the tumor. This light then selectively destroys cancer cells. Such photodynamic therapy is currently undergoing clinical trials [*see illustration above*].

In the near future these diagnostic and surgical systems may be combined into a single laser endoscope for treating tumors. The device would include a fiberscope, a quartz fiber for transmitting ultraviolet excitation light and another quartz fiber for transmitting red light for photodynamic therapy. The endoscope would be inserted through a natural orifice or through the skin to reach a suspected tumor. Ultraviolet laser light would be transmitted through the first fiber, and the fiberscope image would be viewed through a red-transmitting filter. Malignancies would be identified by their red emission, and high-intensity red light would be sent through the second quartz fiber to destroy them. The quartz fiber might be inserted directly into a large tumor. After a few weeks (or after several treatments), the procedure might be repeated to ensure that tumor growth had been arrested.

Flexible fiberscopes, optical sensors and laser-power delivery systems open a new era of minimally invasive medical procedures. Laser endoscopes will include a fiberscope complete with an imaging bundle and illumination light guides, optical fibers coupled to diagnostic sensors, a power fiber for delivering laser radiation and a channel for injecting liquids and aspirating liquids or gases. The clarity of a high-definition television, the precision of a medical laboratory and the skill of a surgical team wrapped into an instrument that can fit through the smallest vessels in the body: this is the future of optical fibers in medicine.

FURTHER READING

FIBER OPTICS: THEORY AND PRACTICE. W. B. Allen. Plenum Press, 1973.
HANDBOOK OF FIBER OPTICS: THEORY AND APPLICATIONS. Edited by H. F. Wolf. Granada Publishing, Inc., 1979.
LASERS IN SURGERY. John A. Dixon in *Current Problems in Surgery,* Vol. 21, No. 9; September, 1984.
OPTICAL FIBERS IN MEDICINE IV. In *Proceedings of the Society of Photo-Optical Instrumentation Engineers,* edited by A. Katzir, Vol. 1067; June, 1989.

HENRY NORRIS RUSSELL (1877–1957) was born in Oyster Bay, N.Y., and after the age of 12 spent most of his life in Princeton. The "Dean of American Astronomers," he is remembered as the co-discoverer of the Hertzsprung-Russell diagram, which relates the brightness of stars to their colors, and of Russell-Saunders coupling, which describes two-electron interactions in atomic spectra. After 1919, when the Indian physicist Meghnad N. Saha successfully developed a theory of ionization based on quantum mechanics, Russell devoted his energy to gaining a physically rational understanding of stellar spectra.

Henry Norris Russell

One of the leading astronomers of his generation, Russell understood the need to place astronomy on a firm theoretical foundation; in doing so he helped to create modern astrophysics

by David H. DeVorkin

In July of 1923 George Ellery Hale received the following report from one of his Mount Wilson Observatory staff about what had become a welcome summer ritual:

"Henry Norris Russell arrived, 'sailed in high,' and . . . with plenty of oil in his crankcase. The talking became a solo and continued unabated during his stay. He gave us three or four talks a week on spectral series applied to atomic structure. . . . Most of his time for a while before he left was devoted to working out the titanium series, according to the selections made in the furnace classification. The complexity proved greater than he expected, and he is still at the job, but the fundamental sorting out was made and it became clear what the character of the multiplets is . . ."

This was Henry Norris Russell in his prime: a dynamo of ideas and suggestions on how to incorporate modern physics into spectroscopic astronomy. He was the Russell of the Hertzsprung-Russell diagram, of the Russell method for eclipsing binaries, of the Russell-Saunders coupling for two-electron spectra and of Russell, Dugan and Stewart's *Astronomy: Analysis of Stellar Spectra,* which helped to train two generations of astronomers. And he was the Russell who for more than 40 years contributed a monthly astronomy column to the pages of SCIENTIFIC AMERICAN.

A nervous bundle of energy, Russell was brought to Mount Wilson by Hale to exploit the store of information gathered there on laboratory and celestial spectra and to help Hale's staff inform their observational studies with the explanatory powers of modern physics. Hale hoped that Russell and other physicists would fill the gap between the laboratory bench and the observatory dome. How closing that gap became Russell's raison d'être provides insight not only into his professional life but also into a science in transition.

In 1889 the 12-year-old Russell, eldest son of a minister in Oyster Bay, Long Island, was sent to live with his maternal aunt in Princeton to take advantage of the good schools there. From his mother, Eliza Norris, he had inherited a flair for mathematics and puzzle-solving and a keen sense of duty. Armed also with total recall he graduated from Princeton Preparatory School at 16 and moved on to Princeton University, where he studied mathematics and astronomy. As an undergraduate he came under the influence of Charles A. Young, a pioneer solar spectroscopist, and prepared a senior thesis on the visual classification of stellar spectra. When he went on to graduate work, his tutors included Young, the mathematician Henry B. Fine and the astronomer E. O. Lovett. This blend of mathematics and observation led to a timely doctoral thesis: a mathematical study of how Mars gravitationally perturbed the orbit of a recently discovered asteroid called Eros; the analysis would lead to a refined value for the distance between the earth and the sun.

While still a graduate student Russell established his hallmark: an acute ability to ferret out new computational techniques. He devised ways to solve for the masses of stars in visual binary systems and to determine the densities of variable stars of the Algol type, then thought by most astronomers to be a special type of eclipsing binary system. Russell soon came to realize that others were in the race too: Hendrikus J. Zwiers beat him to the visual binary technique by almost two years, and Alexander Roberts suggested the Algol technique at about the same time Russell did.

His first brushes with competition served only to push Russell harder. After completing his thesis in 1900 he broke down from overwork and recuperated during the next year by touring France. In 1902 he went on to do postdoctoral work at the University of Cambridge, where his ambitious program of theoretical and observational work on the trigonometric parallaxes of stars was again interrupted by serious illness—perhaps a nervous collapse. Leaving his work unfinished, he returned to Princeton in 1905, took a faculty post and spent the next five years attempting to find order among various observed and calculated characteristics of stars: their intrinsic brightness, colors, masses, densities and spectra.

For this study he relied heavily on star spectra provided by Edward C. Pickering, director of the Harvard College Observatory, whose army of female assistants—in particular Annie J. Cannon—was responsible for assembling the largest collection of stellar spectra and apparent brightnesses in the world. Pickering suggested that Russell compare the absolute brightness he could derive from the parallaxes of his Cambridge stars to their spectra in the Harvard collection.

The eventual outcome of this investigation was the famous Hertzsprung-Russell diagram. (The Danish astronomer Ejnar Hertzsprung devised it independently between 1908 and 1910.)

DAVID H. DEVORKIN is curator of the history of astronomy at the National Air and Space Museum of the Smithsonian Institution in Washington, D.C. He holds degrees in astronomy and astrophysics from the University of California, Los Angeles, San Diego State University and Yale University and a 1978 Ph.D. in the history of science from the University of Leicester. His primary interests are the development of modern astrophysics and the origins of space science. DeVorkin's latest book is the *Race to the Stratosphere: Manned Scientific Ballooning in America,* to be released this month by Springer-Verlag. The author acknowledges permission from the Princeton University Library, manuscript division, to quote correspondence from the Henry Norris Russell papers.

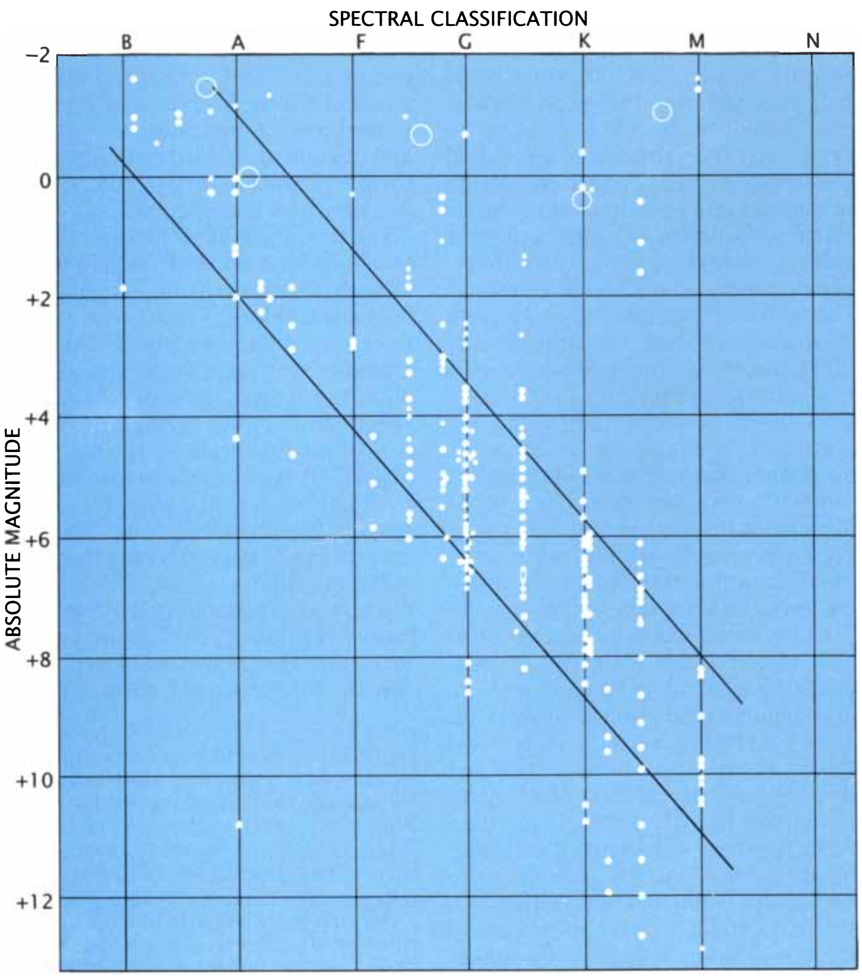
This diagram, published by Russell in 1914 [*see illustration below*], shows how almost all stars can be classified according to their brightness and color and has played a crucial role in guiding theorists in their search for a theory of stellar evolution.

During the same years Russell himself was developing a theory of stellar evolution (his giant and dwarf theory) based on an earlier hypothesis of Norman Lockyer. According to the theory a star begins its life as a vastly extended cloud of gas (a red giant), contracts and heats under self-gravitation to a critical point at which it no longer behaves like a perfect gas. It then cools (as a dense dwarf star) as it contracts further and spends the rest of its life cooling and contracting to oblivion. One must remember that in those days nuclear reactions were unknown and the theory was based simply on gravitational attraction and the kinetic theory of heat. Astronomers also had little evidence to show that stars do behave as perfect gases. The connection between modern physics and the stars was still weak.

The years 1911 and 1912 saw yet another contribution from Russell: the first quick and efficient method of deriving the orbital parameters and physical characteristics of eclipsing binary stars—later known as the Russell method for eclipsing binaries.

For his work in stellar astronomy Russell was made a full professor at Princeton in 1910, and he remained there the rest of his life. Although he shied away from politics and built no institutions, by World War I he had become a major force in American astronomy, able to dictate research agendas and direct the professional lives of workers far beyond the confines of Princeton. One early student, Harlow Shapley, became director of the Harvard College Observatory in 1921. Hale, Edwin B. Frost and Otto Struve, as successive editors of the *Astrophysical Journal,* turned to Russell for advice on what was acceptable for publication in the journal.

Although Russell's students were few, they included the very best. Donald H. Menzel led efforts to understand the outer layers of the sun and the physics of gaseous nebulae and became director of the Harvard College Observatory in 1954. Lyman Spitzer, Jr., pioneered the application of plasma physics to the stars and became director of the Princeton University Observatory on Russell's retirement in 1947. Indeed, in surveying the American astronomical landscape of the 1930's, Struve found that the only place one could go to learn theoretical astrophysics was Princeton, at the side of Henry Norris Russell.

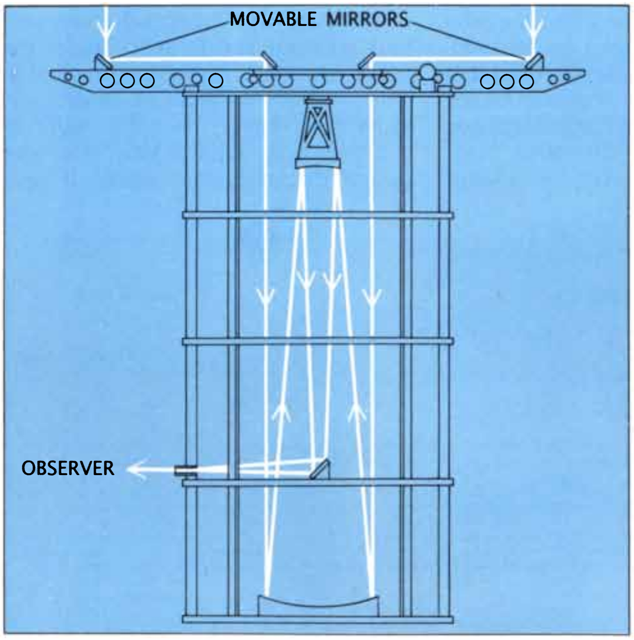


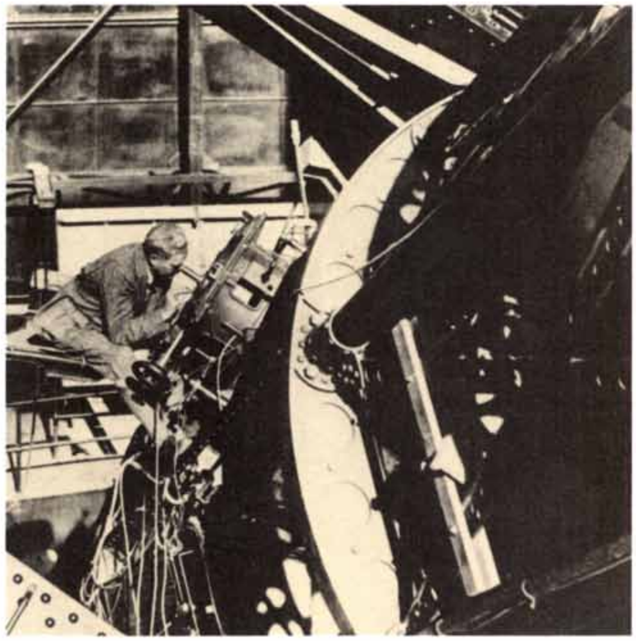
HERTZSPRUNG-RUSSELL DIAGRAM was developed independently by Ejnar Hertzsprung of Denmark and Russell between 1908 and 1913. In the H-R diagram the absolute brightness, or magnitude, of stars (*vertical axis*) is plotted against their spectral classification (*horizontal axis*), which normally ranges from blue (*O*) to red (*M*). Most stars fall along the so-called main sequence, running diagonally from the upper left to the lower right. Red giants fall in the upper right, and white dwarfs cluster at the bottom left. Much of 20th-century astrophysics has been devoted to providing a theoretical interpretation for the diagram. The diagram shown here is Russell's first published version, which appeared in *Popular Astronomy* in 1914. The circles represent various confidence levels for stellar parallaxes.

World War I forced Russell to turn his energies to aerial navigation and sound-ranging, but he sought diversion in planning the future of American astronomy. The country's strength lay in observational work; it boasted the world's largest telescopes and the clearest known skies. The Harvard College Observatory, Lick Observatory and Yerkes Observatory in Wisconsin were deeply committed to star surveys that had been instituted before the turn of the century.

Russell understood, however, that America lagged behind Europe in theoretical astrophysics, and he began to question the seemingly endless cataloguing projects that mapped the heavens in ever-greater detail without adequate theoretical underpinnings. Centuries earlier Newton's laws had served to describe the orbital motion of celestial bodies. What about the laws of radiation physics and quantum mechanics forged by Planck, Bohr and Einstein, which might describe the structure of stars?

Russell and others thought such questions had to be faced before more observational programs were planned. In 1916 he openly questioned the life work of his old friend and patron Pickering: "It is too often still the case





MICHELSON INTERFEROMETER was mounted on top of the 100-inch telescope on Mount Wilson in 1920. Light from a star is reflected off two movable mirrors at each end of the 20-foot-long girder; the two light beams are brought together at the focus of the telescope, where they combine, or interfere. Moving the mirrors can cause the interference fringes to vanish; the separation of the mirrors at that point is a measure of the star's angular size. With this technique the resolving power of the telescope is made equivalent to that of a telescope with a mirror 20 feet in diameter. The instrument was constructed under the supervision of Albert A. Michelson, and in late 1920 Francis G. Pease (shown here) made the first determination of the angular size of a star—Betelgeuse—with the mirrors set by hand by a night assistant sitting on the beam.

that the routine work is initiated too early, before the methods are fully perfected." In 1917 Hale asked Russell to prepare the first of several research surveys for the newly established National Research Council. In his survey Russell argued that many areas in modern astronomy could be strengthened by a closer link to theoretical physics. And again, in 1917, he wrote to Pickering about the ongoing work of Pickering's assistant:

"To be quite frank it seems to me that Miss Cannon has been more concerned with what the spectra *look like* than what they *mean.* I do not think that this fact diminishes the service she has rendered to astronomy; on the contrary, her strict attention to the facts, disregarding the current theories, has given her a peculiar aptitude for her great work."

Pickering, believing that Harvard's observational program deliberations had not been adequately represented in the NRC, shot back that perhaps, in the best interests of astronomical progress, it would be good "for the dreamer to suggest to the practical man what facts he wants." The argument never erupted further. When Pickering died in 1919 Russell paid tribute to him, saying that "it was the spectroscopic information [Pickering] sent me for my parallax stars—a free gift to a young and unknown student—that started me [on the] trail that led to the whole giant and dwarf theory. I do not believe there was ever a more generous man of science."

In spite of the tribute Russell remained unshaken in his conviction that the direction of astronomy had to change. As he had written to Pickering in 1917, "It seems to me that present-day astronomy is like an army advancing with two wings, one along the line of routine observation and the other along that of investigation of principles. If the wings are not in constant touch with one another, the army will not get far." It became Russell's self-appointed task to lead American astronomy's theoretical wing and transform it into modern astrophysics.

In spite of his rhetoric, in the years that immediately followed World War I, Russell was drifting. Between 1914 and 1919 he had published 26 astronomical papers in 15 categories: binary stars, the orbit of the moon, stellar energy and evolution, stellar magnitudes and masses, parallaxes, variable stars and more. The work was good, but it certainly did not serve as the appointed signpost to the future.

Russell's 1917 NRC report shows he understood that the key to applying the revelations of Planck, Bohr and Einstein to the structure and evolution of stars lay in the analysis of stellar spectra, and yet he was stymied in his attempts to do so. At that time astronomers could identify the chemical elements producing certain absorption or emission lines only by comparing them with laboratory standards. They could classify stars according to their spectra but only in a qualitative way—by the absence or presence of lines and by the strength of the lines. The physical mechanisms behind absorption or emission features were poorly understood. Russell and others acknowledged that a difference in stellar temperatures was the primary cause for differing spectra and stellar composition probably only a secondary cause, but there were strong detractors from this view. There was as yet no physical argument for why changes in temperature created different stellar spectral features.

Astronomers were on slightly firmer ground regarding continuum spectra: the continuous, rainbowlike background on which a star's absorption or emission lines are superposed. Largely as the result of excellent colorimetry performed in Germany, there was some confidence that the amount of radiation a star emitted in a given wavelength band (its color index) bore some resemblance to Planck's famous black-body spectrum, which is the spectrum of a theoretical object whose radiative properties depend solely on temperature. But the link was as yet quite weak; moreover, many stars did not seem to behave at all like black bodies.

Russell was not the only one who recognized that a better understanding of stellar spectra was crucial. His chief counterpart in England, Arthur Stanley Eddington at Cambridge, was also thinking along the same lines. The Planck formula gives the amount of radiation that an object emits per unit area into a unit solid angle—its surface brightness. To confirm that a star behaves like a black body, one must then know the solid angle it subtends—or, equivalently, its apparent diameter. Conversely, if one

12 SCIENTIFIC AMERICAN July 2, 1921

The Heavens for July, 1921

What a Study of Atoms and Electrons Tells Us of the Stars

By Henry Norris Russell, Ph.D.

IT is becoming more and more evident, as both sciences advance, that the astronomy of the future will be intimately associated with and dependent upon the concepts and the results of physics, and especially of that branch of physics which deals with the constitution and properties of atoms. Our knowledge within the latter field has been very greatly extended within the last decade, and many astronomical observations which before were puzzling have thereby been explained.

This is particularly true in the realm of spectroscopy. The main facts regarding the emission of light by hot bodies, and by hot gases in particular, have been known for many years; but it is only recently that we have even begun to have an idea of the processes taking place inside the atoms of the gas, which are involved.

For example, when the vapor of a given element, such as calcium or iron, is confined in a heated tube or "furnace" and observed through the end of the tube, the spectrum of the light which it emits shows certain bright lines. If the temperature is raised these lines grow stronger and new lines appear in addition. When the same metal is brought into an electric arc (which is hotter, and also subject to direct electrical action), more lines appear; while a yet more advanced stage may be reached by passing a powerful spark, fed by a source of current of high tension, between two bits of the metal; and in the spectrum from this lines may be found which were not to be observed at any of the lower stages of temperature.

Extensive studies have been made of these phenomena, and long lists of "furnace" and "spark" lines compiled, with important astronomical applications. But the physical explanation, from the atomic standpoint, lagged behind, and came only with the application of the modern quantum theory, which has been remarkably successful.

Why Are the Spectral Lines?

We have good reason to believe that an atom of any element consists of a central, and very small, nucleus, carrying a positive electrical charge, surrounded by a number of negatively charged electrons, which under the system of forces acting between them and the nucleus arrange themselves automatically in a definite pattern, probably consisting of several concentric shells or layers, at least in the heavier atoms. In the hydrogen atom there is but one electron; in helium two; in oxygen eight; in sodium eleven; in iron twenty-six; and so on up to 82 for lead and 92 for uranium. The inner electrons are held by very powerful forces, and are hard to dislodge; but a few of the outermost are relatively easy to displace, and it is these which are concerned in the chemical affinity between atoms of different sorts, and also in the production of the radiation of the visible spectrum. To pull one of these electrons away from the rest of the atom, or as it is called to ionize the atom, demands a certain expenditure of energy; and this produces an absorption of light by the gas of which this atom is a part. When some other free electron comes near to the ionized atom, it will be attracted to it (provided it does not go by too fast); and, in falling back, a corresponding amount of energy will be emitted in the form of light radiated by the gas.

Recent research has shown that this is but part of the story. There appear to be many different positions in which the electron can stop, short of being pulled clear away from the atom. The farther out it gets the more energy is required to raise it—the greatest amount of all corresponding to the complete removal of the electron, or the ionization of the atom.

Now when an electron changes from one of these states to another, light is absorbed, if it is pulled up to a "higher level" nearer the outside of the atom, or emitted if it drops to a "lower level"; and this light consists of vibrations at a perfectly definite rate, giving a sharp line in the spectrum. The most remarkable feature remains to be mentioned. The number of light vibrations per second is exactly proportional to the amount of energy which is required to pull the electron up from one position to the other, or is liberated when it comes back. The reason for this famous "quantum relation"—and indeed the reason why the various possible positions for the electron should exist at all—remains still a mystery, which is regarded by the ablest physicists as one of the hardest problems of science. But the fact has been tested in so many ways that no doubt remains.

When the spectra of the elements are studied from this standpoint it is found that the furnace lines correspond (in the case of absorption) to the raising of the electron from the very lowest "level" at which it normally is situated in the undisturbed atom to various higher levels; while the arc lines, in general, correspond to the raising of the electron from one of these higher levels to another. When light is emitted we have to do with an electron falling back over one of the same intervals.

The enhanced lines correspond to still another process. After one electron has been taken clear out of the atom, it is often possible, by a greater force, to pull a second electron out, and doubly ionize the atom. In this process too there are various possible "levels" between which the second electron may shift, and a corresponding set of lines, all quite different from the furnace or arc lines. It is even possible that an atom may lose a third or actually a fourth electron, and there is reason to suppose that some spectral lines, produced only in very violent sparks, are of this origin.

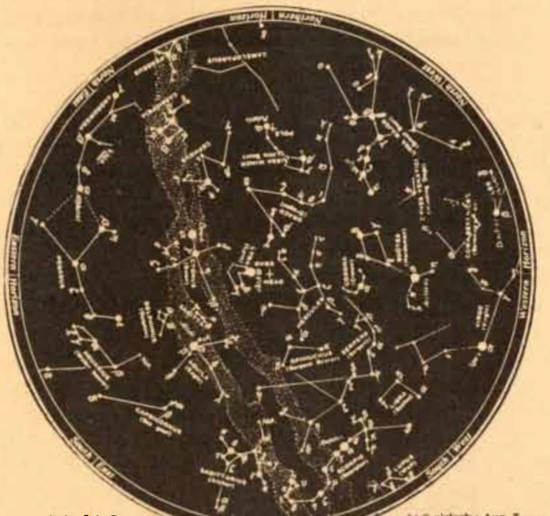
What It Means to the Astronomer

With these ideas in mind it is very easy to see why the furnace lines are characteristic of the red stars, like Betelguese; the arc lines of yellow stars, like the sun; and the spark lines of very white stars, like Sirius. In the hot atmospheres of the stars, the atoms collide and jostle one another. The red stars are the coolest, and the collisions are the least violent, so that most of the atoms are in their undisturbed condition, and absorb only the flame lines. In the hotter atmosphere of the sun many of the atoms are jostled so that the electrons within them are raised to higher "levels" and are in a position to be raised further, with absorption of the arc lines. Finally, in the still hotter white stars most of the atoms are completely ionized, and are therefore ready to have a second electron removed, with absorption of the light corresponding to the spark lines. For some elements, such as calcium, this process occurs with relative ease; hence the spark lines of calcium—the great H and K lines in the violet—appear strongly in the sun. Helium on the other hand is the most difficult of all the elements to ionize; and the amount of energy required even to lift an electron from the lowest "level" to the next above is so great that the corresponding light vibrations are exceedingly rapid, and lie so far in the ultra-violet that all ordinarily transparent substances are opaque for them. The visible lines of helium correspond to a lifting of an electron from the second, or even a higher level to one still above, and can only be produced in an atom which has already been violently jostled, so as to throw the electron up to the second "level." This explains why the absorption lines of helium are found only in the very hot stars, like those in Orion. Spark lines of helium, corresponding to the loss of a second electron, are known; but these are found only in a very few stars which, from other evidence as well, we have reason to believe to be the hottest in the heavens.

Many beautiful applications of this theory have recently been worked out by an Indian physicist, Dr. Megh Nad Saha, of the University of Calcutta. Much of the foregoing discussion is adapted from his work, and one more instance of it may be given. The dark lines of sodium are strong in the solar spectrum. Those of potassium are present, but weak. The rare alkali metals, rubidium and caesium, show many strong lines but these do not appear in the sun at all. This has long been a puzzle, but Dr. Saha has given the solution.

Laboratory experiments have shown that it is fairly easy to remove an electron from a sodium atom, easier to get one out of a potassium atom, and still easier for rubidium and caesium. To get a second electron away from any of these atoms, after the first is gone, is however very difficult. Calculation shows that, in the sun's atmosphere, sodium vapor should be largely ionized, with however a considerable percentage remaining un-ionized atoms, which still retain one electron that may be removed by the action of light, with absorption of the well-known sodium lines. For potassium, almost all the atoms are ionized, leaving very few in a position to produce the absorption lines. Rubidium and caesium, still easier to ionize, would be completely ionized, leaving no atoms at all in a position to produce the absorption lines which are so conspicuous under the less extreme conditions of our laboratories. Hence the weakness of the potassium lines, and the absence of those of the other elements, is completely explained.

When more laboratory work has been done (largely by electrical methods) on these matters, it probably will be possible to calculate with fair precision the temperatures of the atmospheres of the various types of stars, simply from a knowledge of the degree to which the various sorts of atoms in them are ionized, as indicated by the lines in their spectra.



At 11 o'clock: July 7.
At 10½ o'clock: July 14.
At 10 o'clock: July 22.
At 9½ o'clock: July 30.
At 9 o'clock: Aug. 7.
At 8½ o'clock: Aug. 14.
At 8 o'clock: Aug. 22.

The hours given are in Standard Time. When local summer time is in effect, they must be made one hour later: 12 o'clock on July 7, etc.

NIGHT SKY: JULY AND AUGUST

The Heavens

At our hour of observation Vega is almost overhead. Cygnus is high in the east, and Aquila in the southeast, a little lower. Below it lie Capricornus and Aquarius, and to the right, due south, is Sagittarius, with Scorpio to the west of it, and Ophiuchus above the latter. Bootes is the most conspicuous western constellation, with Corona above it and Hercules almost overhead. Ursa Major is in the northwest, Ursa Minor and Draco in the north, Cassiopeia and Cepheus in the northeast, and Pegasus has just risen in the east.

The Planets

Mercury is an evening star at the beginning of the

(*Continued on page 16*)

assumes that a star behaves like a black body and its surface brightness is known, one can predict its apparent diameter.

In 1920 Eddington remarked in his presidential address to the British Association for the Advancement of Science that "probably the greatest need of stellar astronomy at the present day, in order to make sure that our theoretical deductions are starting out on the right lines, is some means of measuring the apparent angular diameter of stars." He then went on to estimate from its color index the temperature and surface brightness of Betelgeuse and showed that if it behaved like a black body it should subtend .051 second of arc, or 1/36,000th the angular diameter of the moon.

Confirmation came quickly. Both Eddington and Russell knew that at Mount Wilson a radical new astronomical tool was then being built under the care of Albert A. Michelson that could measure close double-star separations or even the angular diameter of stars: a giant 20-foot optical interferometer designed to sit atop the newly inaugurated 100-inch Hooker telescope [*see illustration on page 129*]. Hale, then at Pasadena, had barely put down his copy of the September 2 issue of *Nature*, where Eddington's address had been reprinted, when he took up his pen to inform Michelson that at last they had an acceptable prediction for the angular diameter of a star. By December the Michelson interferometer was put into operation and Eddington's prediction confirmed. This was the first observational check that stars do indeed behave according to the laws of modern physics.

Russell had also made a prediction of Betelgeuse's diameter, which appeared in print in late December. It came close to the observed value, but Eddington's prediction was closer. Although Michelson later acknowledged Russell's prediction, Russell knew from this episode that the future of astronomy would largely center on the methods of the Cambridge school, which was at the cutting edge of applying physics to the study of stars.

Another event of December, 1920, changed everything. That month a copy of the October issue of the *Philosophical Magazine* reached Russell. There an obscure young Indian physicist in Calcutta, Meghnad N. Saha, boldly linked the ionization potential of a chemical element (the energy needed to detach an electron from the nucleus) and its degree of ionization to the temperature and pressure of the surrounding environment. Russell quickly recognized that this was the master key to stellar spectra he was looking for. To Walter S. Adams, one of his Mount Wilson collaborators, Russell wrote: "I believe that within a few years we may utilize knowledge of ionizing potentials, and so on, to obtain numerical determinations of stellar temperatures from spectroscopic data."

Working within the framework of the Bohr atom, Saha had applied the concepts of thermodynamic equilibrium and thermal ionization to stellar atmospheres. According to the Bohr model a photon impinging on an atom may excite an electron from one energy level to another; in the process the photon is absorbed. Conversely an electron falling from a higher energy level to a lower one emits a photon. Saha saw how these concepts could help illuminate the behavior of atoms in stellar atmospheres. In particular he showed why spectral lines of a given element are stronger at some temperatures and weaker at others.

For example, in stellar atmospheres much below 4,000 degrees Kelvin, most hydrogen atoms have their electrons in the lowest energy state. (Such stars fall into classes *K* and *M* in the familiar Harvard sequence *O, B, A, F, G, K, M,* where *O* stars are the bluest and *M* are the reddest.) Photons are unable to excite transitions between higher levels and be absorbed in the process. Therefore, absorption features corresponding to the transitions between the higher levels are absent. At about 10,000 degrees Kelvin transitions do take place between the higher levels, and absorption lines (the "Balmer series") dominate the visible spectrum. In much hotter stars (class *A*) most of the hydrogen is ionized—or exists in high states of excitation—and any transitions are above visible frequencies and hydrogen absorption features weaken again. High pressure tends to reduce the amount of ionization, and so stellar pressure also influences absorption and emission features.

Saha's October paper on stellar spectra was followed in rapid succession by two others while he was still in Calcutta. He then traveled to England, hoping to work at Cambridge, but was befriended only by Alfred Fowler of the Imperial College of Science and Technology in London. A fourth paper appeared after he had joined Fowler's laboratory. It established, as the above discussion indicates, that the *O, B, A, F, G, K, M* classification represented not only a sequence in color from blue to red but a sequence of absolute temperatures as well, *O* being the hottest and *M* being the coolest.

While Saha was in London he complained, "We have practically no laboratory data to guide us." Russell also felt the acute lack of high-quality spectroscopic and ionization data. Unlike Saha, however, he knew where to get it: from George Ellery Hale and his Mount Wilson staff.

Already in his December letter to Adams, Russell had identified ways to exploit Saha's revelations. Adams had long suspected that pressure differences in stellar atmospheres could influence the appearance of their spectra. His own method of spectroscopic parallaxes, which he developed with Anton Kohlshutter and in which certain spectral features could serve as distance indicators, hinted at the role of pressure in stellar spectra; it was now placed on a firm theoretical footing by Saha's theory. Russell pointed this out and directed Adams to his own vast store of stellar spectra at Mount Wilson, suggesting that he look for molecular hydrogen in very cool stars.

The following summer Russell himself traveled to Mount Wilson as a summer research associate to apply Saha's theory to stars. For the next two decades Russell left Princeton at least once, sometimes twice, a year for the Pasadena offices of the Mount Wilson Observatory. En route he often stopped at other observatories to advise on research, raid their plate vaults for data and lecture on a wide range of subjects. Although he maintained many of his earlier research interests, such as binary and variable stars, he added the new goal of pushing toward a theory of stellar spectra. During his first summer at Mount Wilson, he told a Berkeley audience that with Saha's theory "astronomy, physics and chemistry now had an atomic model for the emission and absorption of radiation." Russell had already used the store of spectroscopic images of the sun at Mount Wilson to show that some of Saha's predictions were correct and that overall the theory was able to explain the behavior of the elements in the solar atmosphere. He added:

"This is but a single illustration of the immense possibilities of the new field of investigation which opens up before us. A vast deal of work must be done before it is even prospected—much less worked out, and the astronomer, the physicist, and the chemist must combine in the attack.... It is not too bold to hope that, within a few years, science may find itself in

MEGHNAD N. SAHA, the fifth child of a shopkeeper, was born in what is now Bangladesh in 1893. Saha displayed talent in his chosen disciplines of mathematics and physics; he was also aggressive and often found himself at odds with his superiors because of his political views and personality conflicts. Nevertheless, he received a D.Sc. in 1918 from the Calcutta University of Science and soon turned to astrophysical problems, exploiting what he had learned from teaching thermodynamics and quantum mechanics. His series of papers, written in Calcutta and England between 1919 and 1921, applied the concept of thermal ionization to stellar atmospheres and paved the way for an understanding of stellar spectra; they are sometimes considered the beginning of modern astrophysics.

In spite of the importance of his work, Saha was not invited by his European and American colleagues to collaborate with them in refining and extending the theory. He returned to India and the University of Calcutta but shortly thereafter became a professor of physics at the University of Allahabad. During the next two decades, he became a central figure in the creation of the National Institute of Sciences of India, the Indian Physical Society and the Indian Science News Organization. After moving back to the University of Calcutta in 1938, he established an institute for the study of nuclear physics and argued for the necessity of scientifically informed national planning. He became highly visible in areas ranging from flood control and economic planning to the reform of the Indian calendar and the peaceful uses of atomic energy. In 1951 he was elected to the Indian Parliament as an independent and remained a parliamentarian until his death in 1956.

possession of a rational theory of stellar spectra, and, at the same time, of much additional knowledge concerning the constitution of atoms."

Russell was not the only one to see these possibilities. Saha himself pleaded for the support necessary to continue his work, and he wrote to Hale asking for the very things provided to Russell. At the same time Ralph H. Fowler and Arthur Milne in England also recognized the potential of what Saha had done and in the next few years worked to complete the theory. They noted that Saha's formula for describing the Harvard spectral classes as a sequence of absolute temperatures did not properly account for the fact that more than one element was present in a stellar atmosphere, and they rectified the omission. At the same time they refined the role of pressure.

Many others based their own work on Saha's, but it was at Mount Wilson, under the coordination of Russell and Hale's staff, that the general attack on spectra took place. As a favor to Hale, Russell answered Saha's letter, outlining the planned agenda for Mount Wilson. He assured Saha that they were going to follow his lead. Saha, however, was not invited to the party.

Soon after verifying Saha's predictions Russell moved in the direction Fowler and Milne were taking to refine the theory itself, partly to try to explain the spectroscopic anomalies unaccounted for by Saha's original version. There were two puzzles at first: the unexpected behavior of barium and the persistence of hydrogen in the spectra of all stars.

Russell found that barium was more highly ionized than sodium in the solar spectrum, which was strange because the two elements have the same ionization potential. The barium puzzle led Russell deep into the realm of physical theory to sort out the spectra of the alkaline earth elements, which include barium.

The alkaline earths distinguish themselves by having two valence, or outer, electrons instead of one. With F. A. Saunders, Russell derived a refined model for the structure of atoms in which two electrons participated in the generation of spectral lines. The rules for two-electron interaction that Russell invented during the collaboration is now called the Russell-Saunders coupling; with it the spectra of barium and the alkaline earths were explained. Even before completing this work Russell took the next step—to examine spectra from atoms with three valence electrons. He chose titanium and found yet another transition rule. Russell was elated; not only was his physics revealing how stars worked, but stellar spectra also could serve as tools to probe the mysteries of the atom.

Although barium and titanium were great successes, hydrogen remained a thorn in Russell's side. At the time astronomers believed that no one element dominated in stellar atmospheres, which were thought to be gaseous admixtures of generally heavy elements—in particular iron. Moreover, Eddington's theory of stellar structure required that the average molecular weight of the gas had to be much higher than that of hydrogen. Yet hydrogen persisted in virtually all stellar spectra.

The hydrogen puzzle led Russell and the Mount Wilson astronomers to recalibrate the solar spectral wavelengths against laboratory standards. At Princeton Russell and his indefatigable assistant Charlotte E. Moore calibrated the strengths of the solar spectral features against a new theory, developed by Russell and others, that gave the various line strengths in terms of the relative concentrations of the elements in the sun's atmosphere.

Still, the hydrogen puzzle was not resolved. In spite of the spectroscopic evidence Russell was skeptical that hydrogen dominated all stellar atmospheres. While he scratched his head over hydrogen, Russell had sent Menzel to Harvard to exploit its incomparable plate vault of stellar spectra. At the same time a young astronomer named Cecilia Payne had arrived fresh from Cambridge and Eddington, armed with insights from Fowler and Milne. She too planned to explore atomic structure with the help of the Harvard spectra and determine the elements present in stars better than Saha had been able to do. Payne succeeded in all of this in her monumental 1925 doctoral thesis, becoming the first person to recognize that hydrogen was by far the most abundant element in the atmosphere of stars.

Her conclusion was not easy to swallow. It threatened Eddington's theory of stellar structure, which prompted Russell to suggest to Payne that her result was clearly impossible. Payne dutifully followed Russell's guidance in the published form of her thesis. Privately she stuck to her conclusions.

The issue of hydrogen abundance raised its head again and again and plagued Russell's efforts to deny it through 1928, even though many of his colleagues had begun to believe that Payne's original conclusion was correct. Russell marshalled all of his forces at Princeton and Mount Wilson to make an assault on hydrogen. Finally, in a masterly 1929 paper that he referred to as his "reconnaissance of new territory," Russell gathered together all the spectroscopic evidence he and Moore had collected and declared that stellar atmospheres were, after all, mostly hydrogen.

During his final assault on hydrogen Russell was aware that Albrecht Unsöld—a student of the theoretical physicist Arnold Sommerfeld and completely conversant with the most modern forms of quantum theory—was able to derive absolute-abundance information from spectroscopic line profiles, something Russell had never tried to do. Unsöld was also confirming that stellar atmospheres were composed chiefly of hydrogen, and Russell knew that the young German's techniques were far more powerful than his own.

Unsöld's mastery of the new quantum physics represented the future of stellar-atmosphere studies. He and a host of other Europeans would ultimately refine Russell's first crude estimate of the relative abundances in the solar atmosphere. Russell was delighted with these extensions of his work, but it was apparent that the new "quantum mechanics" were rapidly taking over. Even in the mid-1920's, as Russell worked to find a theory of multiplet spectra (spectra produced by closely clustered atomic energy levels), Sommerfeld, H. Hönel and R. de L. Kronig were hot on the trail—and in fact they beat him into print.

Such theoretical races made Russell aware of the army of European physicists who were then attacking atomic structure. Russell knew he was outnumbered and to a great extent outdistanced. In part this stemmed from an aversion to the direction in which physics was heading. Russell never felt comfortable with the complex mathematical formalism of quantum mechanics and was always happier with what he called the astronomical model of the Bohr atom, right down to its metaphor of "spin." Much later Russell still referred to Heisenberg's uncertainty principle as a "Principle of Limited Measurability," following Max Born. Neither was he ever comfortable with the wave-particle duality of matter, although he was willing to apply either model to "practical problems."

After the mid-1920's Russell preferred to let others handle developments in theory. He continued to admire the power and generality of the Bohr model and its ability to provide rules for the calculation of atomic spectra. Along with so many other spectroscopic physicists of his generation, Russell contented himself with puzzle-solving: applying the Bohr model to spectra in order to unravel the structure of atoms.

Russell's role as a pioneer in quantitative astrophysics was that of one who pointed the way. He was a transition figure who never made the transition fully himself; his students did, however, and much of Russell's lasting influence came from them, astronomers such as Spitzer and Menzel and those they trained in turn.

Russell's many roles prompted Shapley to knight him with the title "Dean of American Astronomers." In Russell was found an unusual blend of two classic scientific styles: the hedgehog and the fox—the deep versus the broad. He was sympathetic to the need for programmatic observation but uncomfortable with vast projects uninformed by theory. Russell often saw through the fog to suggest fruitful lines of research that others might carry out. As Cecilia Payne-Gaposchkin once said late in life, "Henry Norris Russell knew a good thing when he saw it." By taking advantage of these "good things" and also conveying to others the need for long-term systematic enquiry informed by physical theory, Russell accrued wide influence. In a survey of astronomers made in 1946, Russell was listed most frequently as an especially stimulating teacher; citations of his work still averaged about 50 per year in the 1960's and 1970's, long after his death at 79 in 1957.

FURTHER READING

THE HISTORY OF ASTRONOMICAL SPECTROSCOPY. Donald H. Menzel in *Annals of the New York Academy of Sciences,* Vol. 198, pages 225–244; 1972.

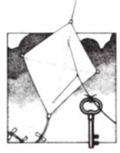
ASTRONOMERS TAKE UP THE STELLAR ENERGY PROBLEM. Karl Hufbauer in *Historical Studies in the Physical Sciences,* Vol. 11, Part 2, pages 277–303; March, 1981.

QUANTUM PHYSICS AND THE STARS. David H. DeVorkin and Ralph Kenat in *Journal for the History of Astronomy,* Vol. 14, No. 40, Part 2, pages 102–132; June, 1983 and Vol. 14, No. 41, Part 3, pages 180–222; October, 1983.

STELLAR EVOLUTION AND THE ORIGINS OF THE HERTZSPRUNG-RUSSELL DIAGRAM IN EARLY ASTROPHYSICS. David H. DeVorkin in *Astrophysics and Twentieth-Century Astronomy to 1950,* edited by Owen Gingerich. Cambridge University Press, 1984.

THE AMATEUR SCIENTIST

How to stop worrying about vibration and make holograms viewable in white light



by Jearl Walker

A hologram is a kind of photograph, but unlike a normal photograph it creates an illusion of depth and also allows the viewer to see the imaged object from various points of view. Indeed, the image can be so realistic that it appears to be the object itself. There are various ways to make a hologram, in all of which film is exposed for several seconds to laser light that has been scattered from the object. Many of the methods are notoriously sensitive to vibrations during the long exposure, so that holographers must often go to great lengths to steady their apparatus.

Recently Roland M. Bagby of the University of Tennessee and Laurie Wright of the Eastman Dental Hospital in London sent me a report on how to make holograms without any fuss. Their rig was designed by Wright and the late Brian Keane of the Royal Sussex County Hospital in England and subsequently modified by Bagby. It is based on a technique for making holograms that was invented in 1962 by Yuri N. Denisyuk of the Soviet Union. The apparatus is so small and sturdy and the technique is so insensitive to vibration that Bagby can load the rig into his car, drive to a local school and then set it up on an ordinary table and begin making holograms within minutes. As a bonus, these holograms can be viewed in white light from an incandescent bulb as well as in light from a laser.

In order to appreciate the Denisyuk technique you need to know the basic principles behind a hologram. Its illusion of depth and faithful representation of perspective stem from the fact that it is a record of an interference pattern created by two beams of light during the exposure. One beam, called the object beam, was scattered from the object, whereas the other beam—the reference beam—was not.

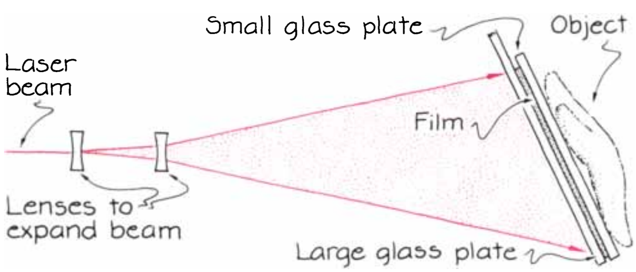
The beams must originate from the same source (these days it is a laser) so that there is a fixed phase difference between them when they reach the film. Phase refers to the state of a light wave as it passes a chosen point. The wave is in a certain phase when a "crest" passes and in an opposite phase when a "trough" passes. If two waves of the same wavelength pass the point, their phase difference is a measure of how closely their states match: the waves are completely in phase if they are in the same state and completely out of phase if they are in opposite states.

When the waves are completely in phase, they are said to interfere constructively, and the point is brightly illuminated owing to the alignment of crests with crests and troughs with troughs. When the waves are completely out of phase, they interfere destructively, and the point is dark because of the complete misalignment. If the waves are long and continuous, their phase difference stays the same as they continue to pass through the point, and so does the level of illumination there. The direction in which the beams travel does not matter: they can be moving in the same direction, in opposite directions or at an angle to each other.

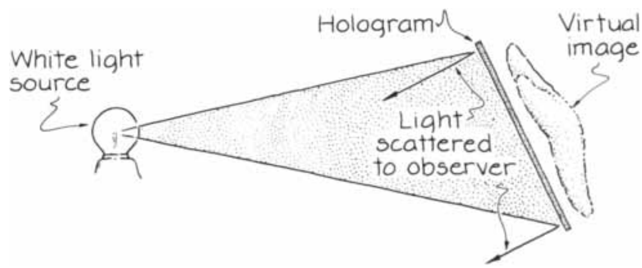
In a hologram an interference pattern is produced in the emulsion of a piece of film. The beams begin completely in phase because they come from the same source, but since the object beam undergoes scattering, they arrive at the emulsion with a variety of phase differences. At some points in the emulsion bright illumination activates the grains of silver and at other points darkness leaves the grains unchanged. When the film is developed, the altered grains are opaque whereas the unaltered grains are transparent. The film, now a hologram, is filled with tiny dark lines and transparent ones, a record of the original interference pattern. In certain processes, including the Denisyuk technique, the film is bleached to brighten the hologram. All the lines are then transparent, but they differ in their index of refraction and hence still provide a record of the pattern.

If the hologram is illuminated by a beam identical with the reference beam, the light scatters from the arrangement of lines and reconstructs the object beam. When you view the hologram from the appropriate angle and intercept some of the scattered light, you perceive an image of the object. As you shift your view somewhat you intercept a different part of the scattered light and gain a different perspective on the object.

The earliest holograms were made without benefit of laser, and so their images were dim and murky. Today holographic images are brighter and sharper because of lasers, better film and better ways of exposing the film. Commonly a laser beam is split by a half-silvered mirror into two beams,



Arrangement for making holograms by the Denisyuk technique



How to view the hologram in white light

which are then successively reflected by other mirrors until they reach the film. Along the way one of them is scattered from an object and becomes the object beam. The other one, the reference beam, is sent into the film on the same side as the object beam but along a different path.

When holograms are made in this way, the rig must be carefully isolated from vibrations in order to maintain a consistent interference pattern on the film during the long exposure. If something along the path of either the object or the reference beam is jostled, the phase difference of the waves reaching each point on the film shifts and the recording of the interference pattern is washed out.

The technique usually has another disadvantage: the holograms must be viewed in the same type of light that exposed the film. The requirement excludes viewing a hologram in white light from a bulb, because white light consists of many wavelengths. The scattering the waves undergo in the hologram depends on wavelength; when there are many wavelengths, you intercept a number of scattered patterns yielding such a jumble of images that nothing is recognizable.

One way around the problem is to expose the film so that the finished hologram selects out only one wavelength to create an image. In this technique the object and reference beams are sent into the film from opposite sides; the film has a thick emulsion, so that many layers of constructive and destructive interference, separated by half a wavelength, span the emulsion. Information about the object is still recorded in the lateral variation of the interference pattern, but now the wavelength of the light is recorded in the spacing of the layers.

After the film is developed it is illuminated with a beam of white light along the path previously taken by the reference beam, and you view it from the light-source side. Although many wavelengths enter the hologram, the only light that is scattered in your direction is light whose wavelength matches that of the original reference beam. The selective scattering results from the half-wavelength separation between the embedded layers: light of the "proper" wavelength is strongly backscattered by the arrangement, whereas light with any other wavelength is not.

The scattering can be thought of as a form of reflection, and so this kind of hologram is called a reflection hologram. When you intercept some of the scattered light, you perceive an image on the far side of the hologram. It is a "virtual" image constructed by your visual system, which mentally extrapolates the rays your eyes receive back to their apparent origin. If you placed a card at the apparent position of the image and looked at the card directly rather than through the hologram, you would not see the image.

A Denisyuk hologram is a reflection hologram, but the laser light is not split into two beams by a half-silvered mirror. Instead it is spread by one or two lenses and then sent directly through a tilted, transparent film to reach the object, which sits immediately behind the film. Some of the light, acting as the object beam, scatters back to the film and interferes with the oncoming light, which acts as the reference beam. When the film is developed, you can view the hologram by shining white light onto it along the same tilted path taken by the initial laser beam. Note, incidentally, the advantage of the tilted orientation of the film during the exposure. If the laser beam had been perpendicular to the film, you would have to hold the white-light source directly in front of your face in order to view the hologram instead of off to one side.

The Denisyuk technique is particularly convenient because the rig requires little isolation from vibration. The object and the film are next to each other; if one wiggles, the other wiggles almost in unison, and so the interference pattern within the film is largely unaffected. If you position the object farther away from the film, the advantage is lost and the rig needs to be isolated from vibrations.

You can make a hologram with the Denisyuk technique by building the rig designed by Wright and Keane. It is shown on the next page and the parts it calls for are listed on page 137. The parts marked with an asterisk can be bought from the Mode Corporation, P.O. Box 1697, San Leandro, Calif. 94577. (For an extra 50 cents per piece the tubing will be cut to specified sizes; otherwise order it in 10-foot lengths.) The optical devices can be bought from the Edmund Scientific Co., 101 East Gloucester Pike, Barrington, N.J. 08007. The precise size and design of the parts are not critical, and Bagby and Wright suggest that readers may enjoy improvising.

Construct the main frame of the rig from the tubing and joints and the inserts that connect them. Use a rubber or plastic mallet to make the connections, but do not ram the pieces hard. Lay the frame on the plywood and test it for stability; if it wobbles, adjust the joints until it is stable, and then fasten it to the plywood with the shelf supports and screws.

A U-shaped mount that will support the film is made with three shorter pieces of tubing. (Hold the mount inside one end of the main frame to be sure there is a clearance on each side of about an eighth of an inch.) Outside the bottom section of the mount attach an identical length of tubing with bolts. Run the bolts through holes drilled in both pieces; either thread the holes in one of the sections to hold the bolts or secure the bolts with nuts. The extra section of tubing forms a narrow shelf to support whatever is to be photographed.

Find and mark the balance points of the mount and then prop it upright in the end of the main frame about three-quarters of an inch above the bottom tube of the frame. Mark the heights of the mount's balance points on the vertical tubes of the frame, remove the mount and then drill holes a quarter of an inch in diameter through the frame at the marks. Also drill threaded holes through the balance points on the mount. Return the mount to the end of the main frame, run a bolt through each hole in the frame, add washers to separate the mount from the frame and then turn the bolts into the threaded holes in the mount. There should be enough washers so that the mount can be easily rotated about the bolts but is kept in place, once positioned, by friction from the washers. Press the cladding channel (a rubberized track that holds the pane in place in some windows) onto the inside of the mount and slide a piece of plate glass into the channel. Later the film will rest against the glass. (You may be able to simplify the entire rig. Wright has made one of wood.)

The optical bench should be between one foot and two feet in length, which may necessitate your cutting a standard bench. Two pin holders are mounted on the bench to hold the pins that are screwed into the lens holders. You can either buy commercial pins and lens holders or hold the lenses with household broom clips, screw the clips into wood dowels and then insert the dowels into the pin holders. (A homemade optical bench might be substituted to reduce the cost of the rig.) The lenses are plano-concave or double concave, with short focal lengths of from −15 to −30 millimeters. Before final assembly the plywood, the bench and everything on it except the lenses should be sprayed with flat black paint to eliminate stray light during the exposure.

Choose a sturdy table to hold the rig. To help isolate the rig from vibrations, stand the legs of the table in coffee cans partially filled with some compliant material such as vermiculite. (If vibrations later prove to be a problem, you may have to mount the table on inflated inner tubes.) Put the bench and optical equipment into the rig, cut a white sheet of paper to the size of the film (about four by five inches) and lay the paper on the glass in the mount, which is tilted with its top part toward the lenses. Then adjust the lenses so that they are aligned with the center of the paper. Turn on the laser and adjust its height and the height and horizontal position of the lenses until the beam spreads uniformly over the paper. (*Never look into a laser beam,* and take great care not to allow any bright reflection of it to reach your eyes.) When the optical alignment is satisfactory, secure the bench to the plywood and mark the locations of the pin holders.

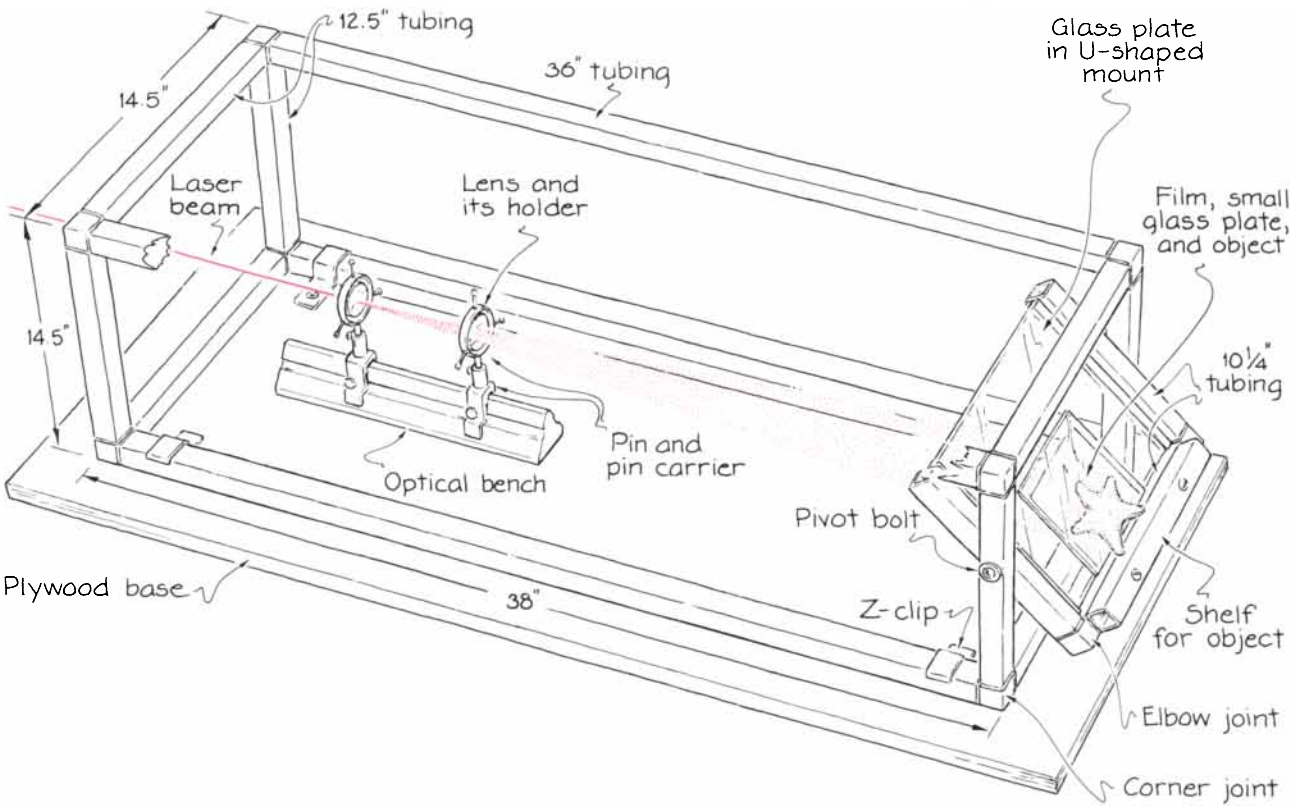
Bagby and Wright say any helium-neon laser will do; those that emit polarized light and have an output power of at least five milliwatts work best. (The weaker the laser is, the longer the exposures must be, and long exposures can make vibrations a problem after all.) The emulsion should be thicker than six microns, transparent to light on both sides (ask for film with a no-antihalation, or "NAH," backing) and sensitive to the red laser light: say Agfa film type 8E75HD NAH, Kodak spectroscopic film type 649-F or a Kodak high-resolution plate.

To process the film you will need a fine-grain, high-contrast developer such as Kodak type D-19 and also a bleach mixture. The developer can be reused if it is stored in a cool place in a brown or opaque plastic bottle. The bleach is needed to brighten a hologram; without it, reflection holograms are disappointingly dark. Bagby and Wright sent recipes for two alternative bleach mixtures. To make one of them, add 25 grams each of potassium bromide and potassium ferricyanide to about 900 milliliters of water (distilled water is best). Stir until the powders completely dissolve, add enough water to bring the volume to 1,000 milliliters and then cautiously add 10 milliliters of concentrated sulfuric acid. (Do all of this in a sink whose appearance is unimportant, and run tap water into it so that any spilled acid is diluted before it gets to the pipes. And whenever you are handling a bleach mixture, be sure to wear safety goggles and laboratory gloves.)

The second bleach mixture yields even brighter holograms, but it also may shift the color of the image. (The shift is no problem when the hologram is viewed in white light, but if it is viewed in the original laser light, the shift may dim or eliminate the image.) The mixture is prepared by adding 30 grams each of potassium bromide and ferric sulfate to 900 milliliters of water, stirring and then adding enough water to bring the mixture to 1,000 milliliters.

You will also need some absolute (100 percent) methanol, a green safety light and a hair dryer, preferably one in which the heat and air speed can be controlled separately. The green safety light enables you to see while developing the film. The methanol serves to dry a hologram quickly, but you must be careful not to breathe it or bring it near any flame or spark, which could ignite it. The hair dryer is used in the final drying stage.

Now you are ready to make a Denisyuk hologram. With the room lights out and the laser on, recheck the beam alignment by placing a white sheet of paper on the plate glass in the mount, which should be tilted between 30 and 45 degrees from the vertical. On the paper lay a piece of plate glass somewhat larger than the film you will be using. Mark the positions of the left



The rig devised by Laurie Wright and Brian Keane for generating Denisyuk holograms and right edges of the second glass with masking tape on the larger glass and then slip the paper out from the assembly.

Block the laser light with cardboard positioned in front of the laser and then slide the film between the two pieces of glass, maneuvering it so that it fits between the tapes. The emulsion side of the film should face the laser. Place the object to be photographed on the top layer of glass and wait for several minutes to allow any vibrations to damp out. Lift the cardboard slightly, wait again for about 30 seconds for any new vibrations to damp out and then lift it completely to expose the film.

The proper length of the exposure depends on the film, the strength of the laser beam, the size of the film and the reflectivity of the object, and so it requires experimentation. If the beam from a five-milliwatt laser is spread over film that measures four by five inches, and if the object has moderate reflectivity, the exposure might require about five seconds. To stop the exposure put the cardboard back in front of the laser. Then retrieve the film and put it in a lightproof compartment until it can be developed.

Develop the film in a room illuminated only by the green safety light. Wearing safety gloves, slip the film into the developer, making sure the emulsion side of the film faces up so that it will not be scratched on the bottom of the container. Swirl the film until it is quite dark; that can take from 30 seconds to two minutes, and getting it right will require some experimentation. Then bathe the film under running water for two minutes before putting it into one of the bleach mixtures. If the film does not soon become more transparent, it was overexposed, overdeveloped or processed with bleach that is too old.

If the film does clear, bathe it again in running water for two minutes. If the tap water is hard, rinse the film in distilled water to eliminate any deposits. Next, blot it with a soft paper towel enough to remove any clinging water but not enough to dry it fully. To complete the drying, submerge the film in methanol for about two minutes. (If the methanol bath gains too much water, the developed hologram will be murky.) After this submersion, work quickly: lift the film, allow the fluid to drain from it and lay it on a soft, dry paper towel with the emulsion side up. Gently apply another paper towel to the emulsion side and immediately complete the drying with the hair dryer, setting it for both heat and air. (Keep the dryer away from the methanol fumes in case it has any internal sparking.) You then have a hologram that can be viewed in white light, provided that the light source is small: a flashlight or a slide projector without its front barrel will serve, but a fluorescent tube will not.

If an object you want to photograph does not balance well on the narrow shelf of the U mount, you can mount it and the film horizontally on a piece of plate glass positioned to straddle the top of the rig. The glass should be a quarter of an inch thick and measure 12 by 14 inches. To reflect the laser beam up to the glass you will need a mirror, which should measure eight by 10 inches and have its reflecting layer on the front surface. Angle the top of the U mount away from the laser, place the mirror on it and adjust everything until the laser light spreads uniformly over a sheet of paper placed on the horizontal glass. Then follow the exposure procedure outlined above.

You may well wonder why all holograms are not made the Denisyuk way. The fact is that the illusion of depth is often weaker in a Denisyuk hologram than it is in a hologram made with a split-beam method. The weakness stems from the fact that a laser does not emit a single, continuous wave but rather a succession of continuous waves, none longer than about the length of the laser itself. The phase changes abruptly and randomly when one wave leaves off and another begins, and so if there is to be consistent interference between two beams of light when the film is exposed, the beams must come from the same wave. In the Denisyuk setup a wave essentially folds back on itself when it scatters from some point on an object. If the point is near the film, the returning part of the wave can interfere with the oncoming part of the same wave at the film, but if the point is too far away, the returning part of the wave meets an oncoming part of another wave; the phase difference between the two waves is unpredictable, and during the exposure there is no consistent interference at the film. Denisyuk holograms therefore record the nearby points of an object well enough but not the distant points.

Although the Denisyuk method is the simplest way to make holograms, it is still challenging; you may want to consult the references listed below for further advice. In addition, Bagby has volunteered to answer questions addressed to him at the Department of Zoology, University of Tennessee, Knoxville, Tenn. 37996-0810.

1" square black tubing	* 4 — sections, 36" long
	* 8 — sections, 12½" long
	* 4 — sections, 10¼" long
Other hardware supplies	* 8 — 4:1 corner joints
	* 2 — 1:1 elbow joints
	* 2 — packages of inserts
	* 2 — 10" lengths of ¼" wide cladding channel
	4 — 1" recess Z-clip shelf supports
	4 — round-head self-tapping screws, ¾" long
	1 — sheet 1" plywood, 17" by 44" or longer
	1 — ¼" thick plate glass, 10" by 10¼"
	2 — ¼" thick glass sheets, about film size
	6 — steel washers, ¼" inside diameter
	2 — steel bolts, ¼" by 20 thread, 1¾" to 2" long
Optical supplies	1 — optical bench
	2 — lens holders
	2 — pin carriers, 30-mm. width
	2 — mounting pins to fit pin carriers
	2 — plano-concave or double-concave lenses

Supplies required for building the rig

FURTHER READING

HOMEGROWN HOLOGRAPHY. George Dowbenko. American Photographic Book Publishing Co., Inc., 1978.

LASERS AND HOLOGRAPHY: AN INTRODUCTION TO COHERENT OPTICS. Winston E. Kock. Dover Publications, Inc., 1981.

HANDBOOK OF HOLOGRAPHY: MAKING HOLOGRAMS THE EASY WAY. Fred Unterseher, Jeannene Hansen and Bob Schlesinger. Ross Books, 1982.

COMPUTER RECREATIONS

Simulated Evolution: wherein bugs learn to hunt bacteria

by A. K. Dewdney

"For those, like me, who are not mathematicians, the computer can be a powerful friend to the imagination."
—RICHARD DAWKINS,
The Blind Watchmaker

On the muddy bottom of a stagnant pool of water a number of protozoa creep about, feeding on the bacteria that slowly rain down on them. The protozoa all look alike, but their behavior shows important differences. Some of them move erratically in search of bacteria and consequently eat little. Others move with more purpose, following a search pattern that seems almost methodical; they find plenty to eat. Such microscopic worlds have a fascination all their own, but this particular scene has special significance: the methodical protozoa evolved from their erratic cohorts in the space of only one hour!

As some readers may already have guessed, such a scene is not viewed through a microscope but on the display screen of a computer. It is generated by a program called Simulated Evolution that was written by Michael Palmiter, a high school teacher from Temple City, Calif. Tiny white protozoan creatures, which Palmiter calls bugs, crawl about on the screen, gobbling up purple bacteria. As generations of bugs pass by, one can watch new feeding behaviors evolve.

Richard Dawkins of the University of Oxford has also looked for insights into evolution by investigating programs that attempt to simulate its various aspects. One such program, written by Dawkins himself, was this department's subject more than a year ago [see SCIENTIFIC AMERICAN, February, 1988]. Dawkins' program displays biomorphs: computer-generated forms that sometimes resemble living creatures. They evolve by a process of "artificial selection": the computer operator arbitrarily selects one of nine possible variant forms of the current biomorph as the basis for future generations of biomorphs.

The biomorphs that emerge from Dawkins' program can be bizarre and amusing—and sometimes even lifelike—but they cannot be said to have evolved naturally, that is, under internal selective pressures. Yet Dawkins thinks it ought to be possible to write a computer program that mimics natural selection. Computer-generated species having such "evolvability" would radiate increasingly complex forms, which selection would pare to a manageable number. Moreover, the surviving descendants would then have to be capable of evolving in new ways that were completely unavailable to ancestors.

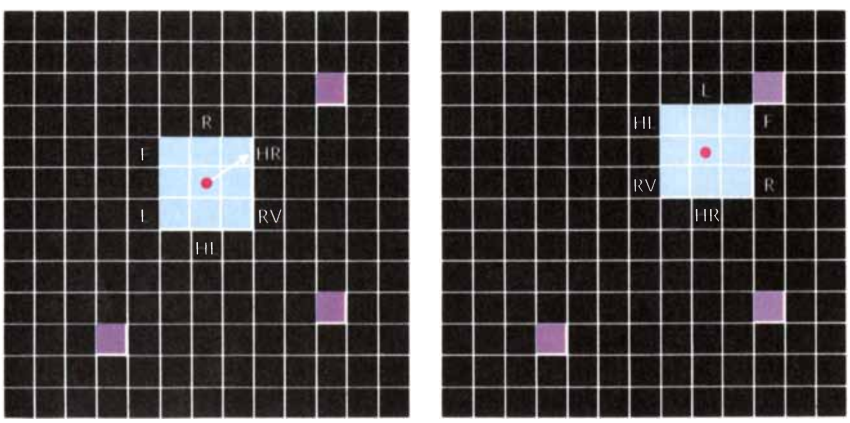
Palmiter's protozoan bugs certainly move us one step closer to Dawkins' goal. As a glance at the illustration on the opposite page reveals, the bugs (the white blips) live within a rectangle on which bacteria (the purple blips) are continually being deposited. The bugs pursue a life dominated by moving and feeding on bacteria. Each bacterium eaten by a bug provides the bug with 40 units of energy, which is enough to make 40 moves. In places where the feeding is rich, a bug may readily acquire 1,500 units of energy in a few minutes. If that happens, however, a strange mechanism kicks in: all further eating does not benefit the bug until its energy level falls below 1,500 units.

On the other hand, it may happen that a bug finds very little to eat over an extended period. In such a case the bug's energy reserves could gradually drop to zero. At that point it would appear to sit morosely for a few cycles, as though pondering its end, and then wink out like a small light.

A bug's success in finding food depends, of course, on the relative abundance of bacteria in its immediate neighborhood. Because the bacteria are deposited more or less uniformly within the white rectangle, if localized feeding has depleted the bacteria in one place, they are bound to be plentiful elsewhere. Some bugs appear to get to the areas of relative abundance quicker than others. It all depends on the moves a bug makes—its search pattern, so to speak.

The Darwinian scenario of the program Simulated Evolution, albeit abstract, hinges on the "genes" that govern the way a bug moves. These particular genes probably do not exist in real protozoa, but Palmiter's bugs have six of them. They are labeled *F, R, HR, RV, HL* and *L* for Forward, Right, Hard Right, Reverse, Hard Left and Left. (All directions are expressed from the bug's point of view. Normal turns amount to 60 degrees in one or the other direction, whereas hard turns are 120 degrees.)

On any given move the bug heads in a direction chosen by lottery: the program picks one of the six possible directions from a kind of mathemati-



A bug's turns are relative to its current direction

cal hat. If the program chooses *L*, for example, the bug makes a 60-degree turn to the left. The probability that a particular direction will be chosen is given by a value assigned to the corresponding gene. Hence the higher a gene's value, the greater its contribution to the bug's overall pattern of movement. If, for example, a bug has a large *L* value in relation to the other five gene values, the bug will spend a lot of time veering to the left.

Every possible combination of gene values results in a different general pattern of movement, and whatever a bug's genetic makeup may be, the bug is stuck with it for life. It can only hope (to be somewhat anthropomorphic) that its offspring will do better.

After a bug has made 800 moves, it becomes "mature" and is ready to reproduce. It does so only if it also happens to be "strong," that is, if it has 1,000 or more units of energy stored under its electric-white membrane. Paramecia undergo a process called conjugation when they reproduce, but the bugs fission: a strong mature bug splits into two new ones, each with half the energy of its parent. When that happens, the new bugs inherit the movement genes of the parent but with a small difference. The value of one of the genes in each offspring is increased or decreased slightly.

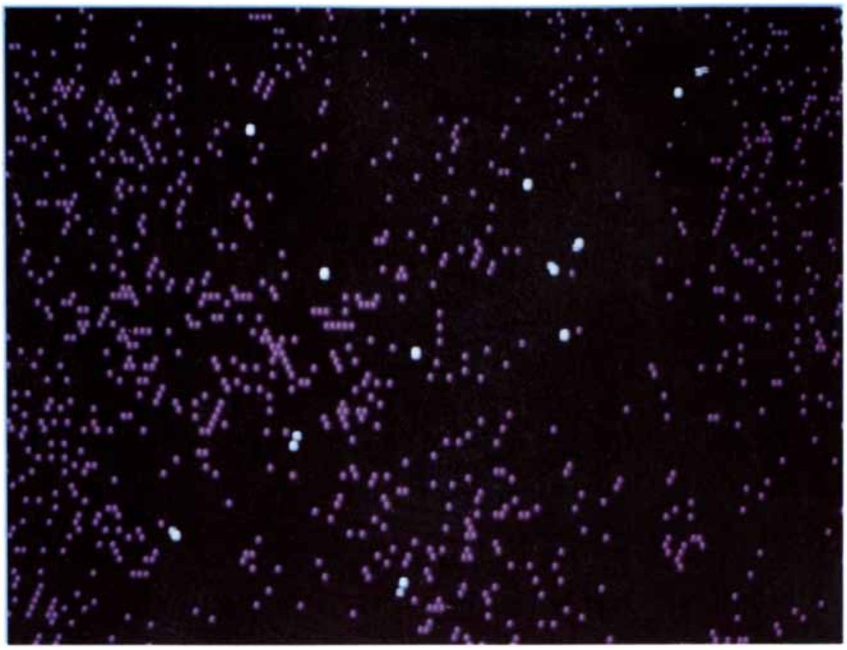
Suppose, for example, a strong mature bug has the gene values $F=3$, $R=2$, $HR=0$, $RV=-2$, $HL=0$ and $L=1$. Its two offspring, labeled *A* and *B*, might inherit the following, mutated forms of that genetic makeup:

A: $F=4$ $R=2$ $HR=0$ $RV=-2$ $HL=0$ $L=1$
B: $F=3$ $R=2$ $HR=0$ $RV=-2$ $HL=-1$ $L=1$

As can be seen, in offspring *A* the *F* value has been incremented by 1, and in offspring *B* the *HL* value has been decremented by 1.

How will the offspring differ from their parent? Offspring *A* will have a slightly greater tendency to move forward than its parent did, whereas offspring *B* will have a slightly lesser tendency to make hard-left turns. Such small shifts in tendencies are barely perceptible on the computer screen to a trained observer.

In its simplest mode, Simulated Evolution starts out by endowing 10 bugs with a random genetic structure, which causes the great majority of them to jitter from side to side in an unpredictable manner. As a rule such "jitterbugs" exhibit a high death rate. They simpy tend to eat up most of the food in their immediate vicinity and then jiggle themselves into starvation on barren ground. Nevertheless, some do survive.



"Jitterbugs" slowly evolve into "cruisers"

Generation succeeds generation every minute or so. This miniature life-and-death struggle makes for absorbing viewing, but the drama is greatly heightened after several minutes, when the viewer becomes aware that some of the bugs have begun to behave differently. They do not jitter; they bobble. Then, a few minutes later, there are bugs that tumble. After 20 minutes or more one can see bugs that glide—at least for short distances. These bugs appear to do much better than their jittery ancestors. Indeed, they proliferate before one's very eyes for precisely that reason.

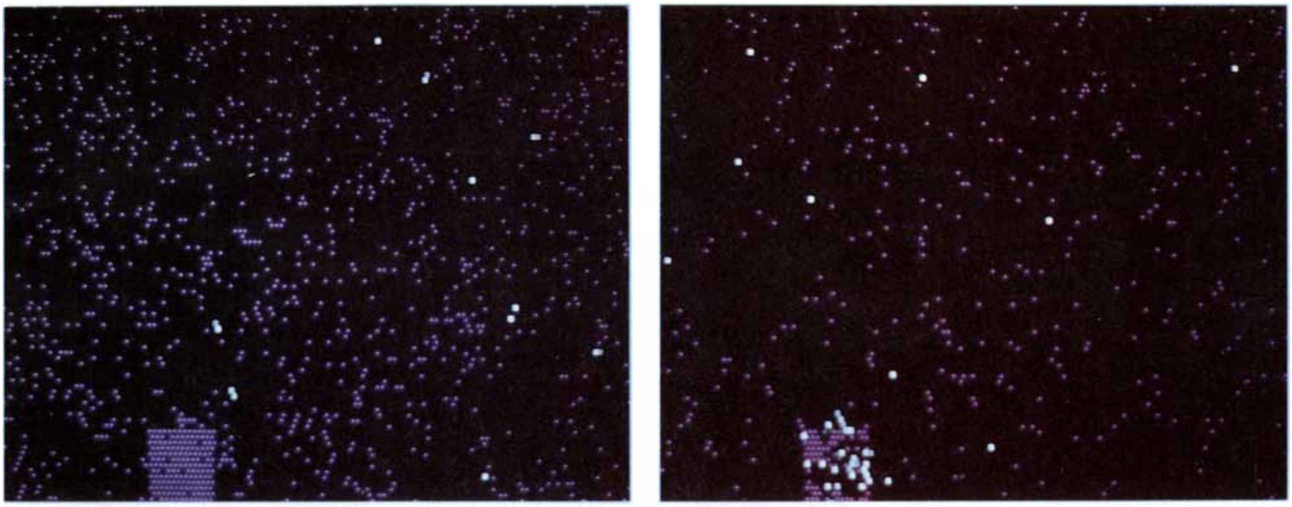
In due course "cruiser" bugs develop that move forward most of the time but turn every now and then. This means they are almost always moving toward denser populations of delicious purple bacteria. Once the behavior is established in just a few individual bugs, it comes to dominate the entire population, since the cruisers end up gathering the lion's share of the food.

Although the cruisers constitute a species of sorts, there nonetheless is still some variation within the cruiser population. For example, some cruisers turn more often to the right than to the left whereas others favor left turns. There are also occasional setbacks, of course. Some cruisers spawn maladapted descendants. A common genetically transmitted disease is the "twirlies," wherein a bug makes too many turns in one direction. Such unfortunate creatures usually die without having known the joy of fission.

It is interesting enough to watch the cruiser species emerge, but Palmiter's program offers more. What if there is variation in the environment? Will more than one species evolve? That question is answered by running Simulated Evolution in a mode in which the screen looks much the same except for a particularly rich patch of bacteria in the lower left-hand corner. The bacteria in that patch are replenished at a much higher rate than normal [*see illustration on next page*]. Palmiter calls that bountiful area the Garden of Eden.

As generations of bugs come and go, the cruisers evolve as before. But within the Garden of Eden something quite different happens. A few lucky jitterbugs that have stumbled into the bacterial banquet are promptly rewarded for their lack of an organized feeding method. Jiggle as they will, food continues to surround them.

As food becomes scarcer in the Garden of Eden, however, a subtle environmental pressure begins to operate. Jiggling and jittering soon are no longer viable strategies. That is when the twirlers make their appearance. What normally is a disastrous genetic defect is actually an advantage in an overpopulated Garden of Eden. Indeed, in the course of time those bugs with a strong tendency to turn in one direc-



***The Garden of Eden* (left) *fosters the evolution of "twirlers"* (right)**

tion predominate in the garden. The reason is obvious. A bug that turns frequently in the same direction, say to the right, will tend to remain in the Garden of Eden longer than its jittery ancestors.

Within a few hours at most the Garden of Eden is populated almost exclusively by highly specialized twirlers that might as well be called nervous orbiters. They follow a specific orbit for many cycles and then suddenly move just one square away and repeat the orbit, sweeping up bacteria with each shift.

Is the Simulated Evolution program a valid model of biological evolution? Only in a very limited sense. It shows how an environment can favor certain variations in offspring, leading ultimately to the formation of new species. But that is as far as the similarities go. Once one or two stable bug species have emerged, nothing else happens. What would it take to realize Dawkins' dream of an indefinitely continuing computer-generated evolution? Perhaps nothing less than a miniature universe inside the computer.

Readers who would like to study the subject can order a copy of Simulated Evolution for $39.95 from Life Science Associates, a small educational-software company. The address is 1 Fenimore Road, Bayport, N.Y. 11705. The program runs on IBM PC and compatible computers, and it comes with an elaborate manual. For those relatively advanced programmers who prefer to write their own version of Simulated Evolution, I shall now describe BUGS, my name for a simplified version of the program.

A BUGS bug can be represented by a small square that has three pixels on a side. The six directions in which such a bug moves can then be illustrated as they are on page 138. A simple table specifies how the coordinates of a bug's central pixel change depending on the direction in which the bug is heading. The table contains two six-element arrays, *xmove* and *ymove*.

dir	0	1	2	3	4	5
xmove	0	2	2	0	-2	-2
ymove	2	1	-1	-2	-1	1

The direction in which a bug is heading (with respect to the computer screen) is given by the value of the variable *dir*. The corresponding numbers in *xmove* and *ymove* indicate by how many pixels, in the horizontal and vertical directions respectively, the bug must accordingly be shifted on the screen in a single move. If, for example, a bug is headed in direction 2, it must be shifted to the right by two pixels and down by one pixel, since $xmove(2) = 2$ and $ymove(2) = -1$.

BUGS determines the direction of motion for each of the creatures in its charge by consulting a formula based on each bug's genetic code, which is contained in a two-dimensional array called *gene*. The element $gene(k,j)$ contains the jth gene value of the kth bug. In the formula each gene value is made an exponent of 2 in order to avoid having to deal with negative numbers. The probability that a bug will move in direction d is then found by dividing 2 raised to the d-gene value by the sum of 2 raised to each of the gene values. For example, the probability that the bug will next turn hard left is the result of dividing 2^{HL} by the sum $2^F + 2^R + 2^{HR} + 2^{RV} + 2^{HL} + 2^L$.

In this manner BUGS calculates six probabilities for each of the possible moves. When the probabilities are added up, the result naturally is 1. One can think of the probabilities as six different ranges that together span a number line extending from 0 to 1. In other words, if the probabilities for the six different directions of motion are represented by p_0 through p_5, then range 0 consists of the interval from 0 to p_0, range 1 consists of the interval from p_0 to $p_0 + p_1$, range 2 consists of the interval from $p_0 + p_1$ to $p_0 + p_1 + p_2$ and so on.

In each cycle BUGS then determines a new value for the direction of motion for a particular bug by selecting a random number between 0 and 1, seeing in which range it has happened to fall and assigning the range number to the variable *turn*. In this scheme, then, *turn* equals 0 for *F*, 1 for *R*, 2 for *HR*, 3 for *RV*, 4 for *HL* and 5 for *L*.

A few simple statements complete the motion algorithm:

$$dir \leftarrow dir + turn \pmod 6$$
$$bugx(k) \leftarrow bugx(k) + xmove(dir)$$
$$bugy(k) \leftarrow bugy(k) + ymove(dir)$$

The arrays $bugx(k)$ and $bugy(k)$ contain the kth bug's current coordinates.

In the first line the current direction *dir* is changed by adding the result of the turning lottery embodied in the variable *turn*. Addition must be modular. For example, if $dir = 5$ (which means the bug is heading up and to the left) and $turn = 2$ (which means it needs to turn hard right), the new value for *dir* will be $5 + 2 \pmod 6 = 1$, and the bug's next move is up and to the right.

BUGS must move all bugs according to this formula, at each step checking whether a bug has hit a barrier or landed on a bacterium. In addition it must keep a record of each bug's age and energy supply in order to determine whether a particular bug should be extinguished or allowed to fission. When a bug is ready to fission, the program merely replaces the old bug with two new ones at the same location. These inherit the old bug's gene values except that a randomly selected gene value is increased by a certain amount in one offspring and another randomly selected gene value is decreased by the same amount in the other.

This description of BUGS will be enough for some to try their hand at writing the program. Those who find the description a bit spare may order a more detailed algorithmic outline from me, enclosing a check or money order for $2 to cover costs.

Enthusiasm for the fractal-generating program SLO GRO, which I described in last December's issue, did not grow slowly. A hefty bag of mail hinted at the continuing interest in fractals in any shape or form. The program can be described scientifically as a simulator of the diffusion-limited aggregation, something we see in the formation of certain minerals, electrolytic plating of metals and even in the accumulation of soot.

The SLO GRO recipe was sufficiently simple for many readers to follow, and many in fact did so. The basic algorithm involves the injection of a randomly walking "particle" into a circle from a random point on the circle's circumference. When the particle comes in contact with a stationary fellow particle, it too ceases to move and thus produces an aggregation of particles. The program was easy to write but was somewhat painful for certain people to watch. Why should they spend their time watching a point of light jittering for what seemed forever? As a result several readers thought of changes in the algorithm that speeded its operation.

Edward H. Kidera IV of Columbia, Md., achieved a definite speedup by starting with a small circle and steadily increasing its radius as the aggregation grew. A number of readers also made suggestions for speeding up the test for contact with a crowd of fellow particles. The test involved comparing the particle's neighboring pixels with the recorded positions of every particle in the growing crowd.

Ronald C. Read of the University of Waterloo in Ontario made the following suggestion on this very point. "For those who use BASIC (as I'm sure many of your readers do) there is a much easier way. That is to use the POINT command of BASIC in order to tell whether the pixel in question has been given a color. In effect, then, one is using the screen as a storage device."

Most impatient of all was William H. Pratt of State College, Pa. Why make the particle wander randomly at all? Why not just give it a random position next to the growth itself? Pratt was dismayed, however, to find that his growth looked nothing like last December's illustrations. It certainly was ragged about the edges but more solid—a different creature altogether.

Pratt was unknowingly playing with what is known as Richardson's growth model, a favorite research tool of a group of mathematicians called "the particle mafia." These investigators, some of whom are based at the University of Wisconsin at Madison, have been studying a great variety of growth models for more than a decade. I hope to report on a recent visit to Madison in a future column.

In January this department featured people puzzles: logic puzzles that can be solved only by thinking about what other people are thinking. An entire class of such puzzles was represented by three philosophers who awoke from an afternoon slumber under a tree. Each philosopher noted that the foreheads of the other two had apparently been befouled by a bird. Only in the course of the ensuing laughter did the wisest of them realize that his own forehead was decorated. How did he make the deduction?

It had not occurred to me, as it did to James D. Klein of College Place, Wash., that there is a two-philosophers puzzle of sorts. Klein tested his own children with the story of a pair of workmen who fall from a scaffold onto the ground. The fall does not hurt either of them, but it does dirty the face of one. Why did the workman with the clean face rush to wash up while the one with the dirty face merely went back to work? Klein writes, "It is interesting to hear them think out loud and watch their eyes as the solution dawns."

Another people puzzle was borrowed from Dennis Shasha's book *The Puzzling Adventures of Dr. Ecco.* In this conundrum two 19th-century generals whose armies are separated by a ridge of land decide to coordinate their attack on the enemy by sending messages by carrier pigeon. But what message to use? If the first general sends the message "Attack at dawn," he must wait for a return message from the second general to confirm that he has received it. What if one of the pigeons never makes it to the other side? And even if both pigeons arrive at their destinations, how does the second general know that his confirmation has been received? An infinite regress of messages appears to be inescapable.

The generals' predicament reminded Warner Clements of Beverly Hills, Calif., of a little-known off-Broadway play that involved a would-be double agent shuttling back and forth between two hostile nations. It begins when the agent learns that country *A* has broken the secret military code of country *B*. The agent goes to *B* in order to sell that country's intelligence officers the information. "We already know that," the officers say. The agent is at first discouraged but then realizes he can sell that information to the intelligence officers of country *A*. They in turn reply, "We know the *B*'s have broken our code. We have been sending them false information!" The agent rushes back to country *B*: "Do you realize the *A*'s know you have broken their code?" "Oh yes," reply the *B* officers. The agent returns to the *A*'s to apprise them of the situation, and so on. How long might the agent have to continue the back-and-forth journeys, bearing an ever lengthening message about what the other side knows? Although in this puzzle the two military factions are not coordinating but competing, it makes the solution no easier—there not being one.

More down-to-earth people puzzles involved real people in everyday situations like those studied by the late Erving Goffman, a sociologist. I asked for examples and received several, including one from P. M. Cambeen of Muiden in the Netherlands. During World War II an officer in the German force occupying Holland expressed to a resident his puzzlement at Dutch people's attitudes. He was told: "The Dutch have three virtues. They are intelligent, loyal and pro-Nazi. Any given Dutch person, however, has only two of these virtues and the opposite of the third." While the logical implications of his statement were being worked out by the officer, the wit had enough time to get away.

FURTHER READING

THE BLIND WATCHMAKER. Richard Dawkins. W. W. Norton & Company, 1987.

ARTIFICIAL LIFE: THE PROCEEDINGS OF AN INTERDISCIPLINARY WORKSHOP ON THE SYNTHESIS AND SIMULATION OF LIVING SYSTEMS HELD SEPTEMBER, 1987, IN LOS ALAMOS, NEW MEXICO. Edited by Christopher G. Langton. Addison-Wesley Publishing Company, Inc., 1989.

BOOKS

The big chill, the war god's sister, keys to secrets, the galaxies in color

by Philip Morrison

The Little Ice Age, by Jean M. Grove. Methuen & Co. Ltd. ($144).

A Lady from Philadelphia, seen in an 1899 photograph wearing the long, full skirt and wide-brimmed hat prescribed for proper outdoor guests at the Glacier House Hotel nearby, can be seen to have taken a few blows with her ice axe at the dusty tongue tip of the receding glacier. She was Mary Vaux, the amateur who along with her family systematically measured and photographed five glaciers in the Canadian Rockies during 25 years of summer holidays, the first New World reconnaissance of the shrinking ice.

The Alpine glaciers had drawn travelers and artists for centuries. Recently a Zurich scholar, H. J. Zumbühl, combed the collections of half of the continent to find 300 representations of the Lower Grindelwald glacier made between 1640 and 1900. He dated them all scrupulously and established the precise viewpoint of each artist. ("Some masterpieces turned out to be quite unsuitable.") The ice front was found to have moved, every few decades, several hundred meters toward or away from a roughly constant position. Since 1850 the Grindelwald has been in overall retreat; it lies now about 1.5 kilometers farther upvalley than it did in the oldest records.

The rich chronicles of the region of Mont Blanc noted the early advance of the ice sometime around 1550. E. Le Roy Ladurie, in his pioneering history of climate published in 1971, was the first to sift for hard fact the plentiful 17th-century French accounts of the "impetuosity of a great horrible glacier." Some of the old tales are embellished, but some remain compelling. In 1690 the frugal peasants of Chamonix even underwrote the expenses of a visit from the bishop of Geneva so that he might exorcise the grim intruder from their pastures and farmyards. The ice reverently withdrew—but only for a few years.

It is the particular virtue of this comprehensive and absorbing study that Jean Grove, a Cambridge geographer caught up for 20 years in the web of climate and history, has brought the same fresh and informed comparative examination to the work of the geologists that we recognize in her first-rate historical scholarship. Plainly this topic invites such unity: the geologists need all the help they can get, once they seek to study the most recent of the Recent, quick bites out of geologic time. How could physical dating methods name individual peak years of cold weather?

The impressive long runs of some single, half-capricious if conspicuous correlation, such as the annual date of the leafing out of the beeches or of the first cherry blossoms, are scarcely to be trusted. They depend on special attributes of the weather. A wider mark is the persistent failure of many crops. The winter of 1695 was long and hard, the coldest year of a cold decade; glaciers grew and famine took a major toll in Finland, Estonia, Norway and Scotland. The year 1771 was the worst in a run of wet and snowy summers in central Europe, again a time of famine and the beginning of a decade that witnessed a rapid advance of the Swiss glaciers. The year 1816 was without a real summer on both sides of the North Atlantic; corn was mostly lost, hay yielded a small crop but wheat and rye were fair to normal. Governments and farmers both moved in that year and afterward to surmount what J. D. Post called the "last great subsistence crisis in the Western world."

Let earth science speak. It does not claim year-by-year time resolution. Even the annual count of tree rings (say in the wood fragments found buried in the gravel banks, called moraines, that receding glacier tongues abandon in their slow retreat) cannot give a tight date. Except for the rare case when an erect tree was sheared off in full growth by encroaching ice, a tree might have lived on as ice crept over it, or it might have died years earlier from the cold breath.

Two new dating schemes have been applied to the recent past. One scheme is based on the growth of lichens on bare rock. The sizes of the patches must be calibrated on the spot, taking account of lichen species and rock surface. The other measure rests on the weathering of rock fragments, say of sandstone. The depth of the textural changes that constitute the weathered rind of the fresh rock also needs to be calibrated. These clever new methods seem less than surefire. But error is demonstrable too when carbon-14 decay is small (as it is over times measured in centuries, only a few percent of the mean lifetime of decay). The activity of old wood has been counted, tree ring by tree ring, to show substantial deviations in the indicated carbon dates by as much as 50 or 100 years over the past few centuries, even allowing for the known magnetic-pole shift that affects the cosmic-ray source of radiocarbon. Perhaps the cause is a variation in cosmic-ray flux that depends on solar activity.

Even if geologists are fuzzy timers compared with the careful bureaucrats and the artists, their reach in space and time is breathtaking. Mostly by dating glacial deposits from samples of wood and soil found within them, geologists have inferred glacial growth and waste from Patagonia to Kanchenjunga, from Mount Kenya to the Southern Alps of New Zealand. It is pretty clear that the Little Ice Age was worldwide. The shrinking and extending glaciers kept pace north and south, given a decade or two of lack of synchronicity. The same kind of data show that the change was not single or smooth but was modulated, although never so much that the northern sea ice withdrew as far as it had in the balmy times of the Norse voyages to Greenland. (Leif the Lucky was just that.) Most glaciers have grown and dwindled by turn a few dozen times in the 10,000 years since the great ice cap left Wisconsin, probably because the mean global temperature has varied up and down by three or four degrees Fahrenheit.

The book opens with a chapter on methods. Then it recounts the significant observations of the Icelanders, those keen naturalists and eloquent poets, during 1,000 years of struggle against both winter's ice and volcanic fire. The chapters widen in geographic scope, first to Norway, then to all of mountainous Europe and finally to

the entire glacial world (except for the Antarctic ice cap, a particularly slow-responding mass of ice). Three chapters discuss causes, consequences and the long run of 10 millenniums. Nearly 2,000 works are referenced, although the 1986 volume by Friedrich Röthlisberger, who has carried out the most elegant worldwide moraine studies, came just too late.

Will the carbon dioxide from our forest clearings and engines rout the glaciers, flood cities and plains, shift croplands? Is the rapid recession of the ice since 1850—it remitted a little in the 1970's—really a harbinger? (Its beginning came rather too early to have reflected any major industrial effect.) Or will another Little Ice Age return instead on the wings of solar variation, volcanic sulfates or subtler feedback systems based on ocean upwelling or snow and ice reflectivity? We do not yet know; one hot summer, like one bishop, does not end the prospect of icy centuries. "It might be wise not to assume that a return to Little Ice Age conditions is entirely out of the question."

THE GREAT TEMPLE OF THE AZTECS: TREASURES OF TENOCHTITLAN, by Eduardo Matos Moctezuma, translated by Doris Heyden. Thames and Hudson Inc. ($29.95).

Time after time rejected as neighbors, the Mexica Aztecs under Tenoch wandered with their strange gods to an island set in the swampy lake of a lonely mountain valley. There all prospered, gods and men alike. It was the year A.D. 1325. A great temple was soon raised to honor both the fierce god of sun and warfare and the quieter deity of rain and sustenance. Six times during the next two centuries the twin stepped pyramids that simulate mountains were enlarged anew in dressed and sculptured stone as the Mexican state extended its sway, not by widening its borders but by extorting rich tribute from its neighbors in exchange for peace.

In 1519 the "tall but accursed" temple of white-plastered masonry, about 20 years into its seventh stage, stood at the junction of three resplendent causeways over the waters. The lake was busy with the laden canoes and luxuriant gardens of 200,000 well-ordered subjects. The bicentennial of the temple was at hand, but it never came. Vengeful war and terrible plague reached the god of war first, the attack speared by a few iron-hard Spaniards leading their many Indian allies. The gods fell along with their temple, itself leveled stone by stone.

In a couple of decades the throngs would gather as they always had in the wide temple precincts. But no longer did they watch the bloodied priests fling dismembered bodies down the war god's steep steps; they viewed and took their parts in Christian plays and rites that ended in mass baptisms. The cathedral that had replaced the towers was built of their very stones, reset "by an abundance of people . . . singing and shouting" as they hauled. "It is customary . . . that the Indians cover their own expenses; they find the materials and pay the stonecutters . . . and if they do not bring their own food, they fast." The National Palace rose on the site of Motecuhzoma II's own sumptuous dwelling; his zoo with its jaguars and foxes, its rattlesnakes tenderly reared in feather-filled bins, his aviary of iridescent birds, his fragrant gardens and the busy workshops of his subtle artisans were all gone, with hardly a trace aboveground.

At first the vanquished gods concealed themselves only a little. "The unmistakable image or idol was hidden behind some masonry work . . . or within an altar." Soon the Spanish displaced the central square of their city a little to the south to abandon the unholy altars. Bit by bit the lake itself was drained; the new broad roofs and paved streets spread out as though to impede the soaking rains of renewal. If the gods were not dead, they were not to be seen among all those figures of saints and the Virgin.

In 1790 the time came for practical men to dig into the plaza, to bring it water pipes and paving. From the underworld the old gods suddenly reappeared: first of all was an eight-foot statue of Coatlicue, the mother goddess, "She of the Serpent Skirt," cat-clawed and bedecked with carved human hearts and hands. She was quickly reinterred. That prince of Enlightenment travelers, Alexander von Humboldt, wrote: "The professors, at that time Dominican priests, did not want to exhibit this idol to the Mexican youth so they buried it again in one of the halls" at the university. On his visit in 1803 Humboldt pleaded for a look at the dazzling antiquities. For him they dug fearsome Coatlicue up from her new, shallow grave.

A map of a few central blocks of the old city shows, along with the subway station, the parking garages, the cathedral and the mint, no fewer than 10 sites where since 1790 signs of godhead have been excavated, usually by inadvertence, although in this century also by the probing archaeologists. Professor Matos, now director of the renowned National Museum of Anthropology and a leader in both fieldwork and interpretation, was in the early stages of a project to consider excavation of the Great Temple area when early in 1978 he heard the great news: "Workers from the Electric Light Company" digging at an intersection in the heart of the city had come on a large carved stone. They put their tools aside for the day and telephoned the salvage archaeologists at the National Institute of Anthropology and History. Four days' work disclosed a 10-foot disk that bore in low relief the representation of a dismembered female nude. In a photograph we see the author taking casts from that newfound red disk where it lay, six feet below the level of the busy streets.

Most of a city block was then laid bare, only a small part of the temple precinct reported by the chroniclers to have held 78 structures in its 65 acres. We see it here in good photographs: details of the frogs and serpents in stone that adorn the broad terraces, and overall views of walls and tiers and interspersed layers of rubble fill that was set between successive stages of the edifice as it grew. The first level has not been reached; it is still down there in the waters, for the lake has not vanished but—like its gods—has gone entirely underground. The seventh level was cleared almost completely by the conquerors. No more than a layer of the platform a meter thick remained along one side. We are given a line drawing that re-creates the Templo Mayor standing among the buildings of today, a cutaway of the stages that are there superposed and a careful plan showing the present exposure of each level. The intricate finds are made reasonably clear; this is not a monograph for the expert but a brilliant summary. Even a general reader could wish for more detail.

The temple is filled with offerings; some were held in special enclosures, but even when they were simply buried loose in the fill, their axial locations and their spatially ordered contents identify them. In the later stages of the structure, when its growth was fastest, four-fifths of the finds are exotic, given in tribute or brought by the Aztecs from some other state. Estimates based on Spanish chronicles suggest that a major part, perhaps half or even more, of the gross national product of the Mexica came as tribute borne by porters from afar, tens of thousands of men on the road to or from the valley at any given time. The rest was the handiwork of the artisans of the Valley of Mexico and the fruits of their verdant and varied gardens, grown in beds patiently raised above the shallow waters of the lake.

Among the treasures, unexpected but glorious, is a pair of life-size figures made of joined pottery. They are Eagle warriors in their regalia: talons at the ankles, plumed wings along the arms and a helmet like a raptor's open beak, out of which the elite soldier looks in all dignity. Three superb conch shells sculptured in hard stone at a scale greatly enlarged from life are found among much more from the hands of sculptor and lapidary.

Written records and manuscript drawings remain in plenty from the post-Conquest years; these wonderful chronicles retell the myths and describe the customs of the Aztec world, even to royal history. The witness of the spade confirms and extends the old tradition. In a brilliant chapter Matos demonstrates how here men first created gods in their own images—a war god to ennoble war for this warrior state—and then reenacted their own myths in ritual, with living (and indeed dying) actors to carry out the high ceremonies of the temple.

The goddess of the find of 1978 is the war god's sister. In myth he was once obliged to decapitate her and to throw her dismembered body down the sacred mountain slope. That powerful relief was found just where it belonged, on the platform at the base of the steep stairs of the war god's pyramid, where many a prisoner of war was sacrificed to recapitulate the conflict of the godly siblings.

The terrible and beautiful gods are no longer in hiding. In 1965 a plaque was installed north of the city center to mark the very spot of the final heroic stand of Aztec power. In 1521, it recalls, "Tlatelolco fell to Hernán Cortés. That was neither a triumph nor a defeat, it was a painful birth of the *mestizo* people who are the Mexico of today." This absorbing volume is more than an attractive archaeological report; it is a work that touches the springs of a vivid national history.

CODES AND CRYPTOGRAPHY, by Dominic Welsh. Oxford University Press ($59.95; paperbound, $29.95).

In a world where innumerable long strings of characters flash by, beset by both random and cleverly contrived noise, the old and subtle art of coding has become a powerful and imaginative branch of applied mathematics. An Oxford mathematician has prepared this up-to-date, brief, lucid but by no means simple textbook on the topic. Aimed at well-prepared undergraduates, it is not for the nonmathematical, although a determined novice might seek specific answers here. A reader will enjoy proofs formal and informal, helpful examples and knowing summaries of the bold conjectures and instructive missteps that fill this lively literature. (Some modern algebra and elementary probability and a good deal of purposeful thinking are prerequisites.)

Among the earliest papers "to treat cryptography as a piece of formal mathematics" were a few published in about 1930 by L. S. Hill. They took some of the clever rearrangements of the traditional cipher into the wide domain of linear transformations. But the present stream of thought owes most to the seminal postwar work of Claude Shannon on communication theory. Indeed, most of the first half of this book is a fine review of Shannon's remarkable results, based on his careful—here it is even axiomatized—definition of informational entropy and what follows from it. Real channels have noise; we see proved the limits and opportunities of encoding to minimize error. The best of codes will not beat the ideal channel capacity, although there are clever ways to draw near to it even through the noise. The space of error-correcting codes and its algebra and geometry are examined briefly but closely, and we end up with natural language sources and their empirical redundancies.

The string DM QASCJDGF... is not English but a string of idly random letters, a zero-order approximation. By going to the second-order approximation, letters chosen from a book to follow two chosen predecessors, we are likely to recognize HE AREAT BEIS HEDE... as being drawn from English (Runyon), SENECTOR VCI QUAE... as Latin (Cicero) and MAITAIS DU VEILLECAL... as French. In a few pages we meet the most familiar ciphers and confront them. Shannon defined the now classical approach: a cryptosystem is a triple set of strings—messages *M*, keys *K* and ciphertexts *C*, with the two-way mappings or algorithms that link them. Breaking a cipher system is finding the right key to read any ciphertext, assuming that the mapping scheme is known.

It is the space of the unknown keys that we fix on. Perfect secrecy is defined as a system such that the statistical entropy of a message without its ciphertext is no different from that of the same message with the ciphertext. For this definition it is enough (and it is necessary as well) that there be at least as many keys as there are messages! Enter the one-time pad of the secret agent (actually it was devised for teleprinters in 1926, long before the theory, by G. S. Vernam). In this system each message is given its own individual random key and hence encodes into an entirely random ciphertext, correlate what you will. It is alleged that this scheme is still in use for the very highest levels of secret communication; if it is, practice honors the beautiful theory.

In the final chapter the idea returns in its new, rather flawed computer-based avatar. We know no fast way to generate all those truly random sequences, keys for the one-time pad. Our computers send us pseudorandom strings all the time, dubbed random but in fact generally cryptographically insecure. For instance, the parameters of the most popular iterative generator that forms the courtesy "random number" in your computer from a seed number can be recovered by reasonable computation from a rather short list of the early numbers it generates. Maybe the idea of entropy itself is due for generalization.

The break between the classical theory and the current one came in the 1970's, when the theory of computational complexity emerged. Given a string of n bits, how long does some computation with that string take? Of course, it depends on n. A useful classification separates all calculations into two: those whose completion requires a number of repeated basic operations that grows no faster than some finite power of n and those whose working out requires actions more numerous than that, effectively exponential in n. Anyone could multiply two n-bit numbers in about n-squared operations with the familiar schoolboy long multiplication, paper, pencil and much patience. By first forming several smaller numbers out of the data you can do much better. It looks as though a safe lower limit is n operations (you must at least take account of all the data) and in 1971 Schönhage and Strassen told us how to do it by their *Schnelle Multiplikation* in about $(n\log n)(\log(\log n))$ steps, which for a million-digit number beats the long scheme by a factor of 100,000—only tenfold worse than the lower limit.

The celebrated codes with public key emerged. Computer networks encourage sending confidential messages in vast numbers to hordes of authorized recipients, say routine transfers of funds to banks everywhere. If you distribute one-time keys, you will work very long hours. If you distribute

SCIENTIFIC AMERICAN

CORRESPONDENCE

Offprints of more than 1,000 selected articles from earlier issues of this magazine, listed in an annual catalogue, are available at $1.25 each. Correspondence, orders and requests for the catalogue should be addressed to W. H. Freeman and Company, 4419 West 1980 South, Salt Lake City, Utah 84104. Offprints adopted for classroom use may be ordered direct or through a college bookstore. Sets of 10 or more Offprints are collated by the publisher and are delivered as sets to bookstores.

Photocopying rights are hereby granted by Scientific American, Inc., to libraries and others registered with the Copyright Clearance Center (CCC) to photocopy articles in this issue of SCIENTIFIC AMERICAN for the flat fee of $1.25 per copy of each article or any part thereof. Such clearance does not extend to the photocopying of articles for promotion or other commercial purposes. Correspondence and payment should be addressed to Copyright Clearance Center, Inc., 21 Congress Street, Salem, Mass. 01970. Specify CCC Reference Number ISSN 0036-8733/89. $1.25 + 0.00.

Editorial correspondence should be addressed to The Editors, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Manuscripts are submitted at the authors' risk and will not be returned unless accompanied by postage.

Advertising correspondence should be addressed to Advertising Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017.

Address subscription correspondence to Subscription Manager, SCIENTIFIC AMERICAN, P.O. Box 3187, Harlan, IA. 51593. Telephone inquiries: 1-800-333-1199, U.S. only; other 515-247-7631/32. The date of the last issue on subscriptions appears on each month's mailing label. For change of address notify us at least four weeks in advance. Please send your old address (if convenient, on a mailing label of a recent issue) as well as the new one.

Name ______________________

New Address

Street ______________________

City ______________________

State and ZIP ______________________

Old Address

Street ______________________

City ______________________

State and ZIP ______________________

many trickier, shorter keys, you face eventual leakage. The neat solution is to publish a partial key *K* for everyone, along with a published scheme for applying it. The secret of the message is kept because each user holds on to a private key *L*, without which the public one is useless. It must be so that given *M* and *K*, the cipher *C* is easy to compute. Then it must also be that given *C*, *K* and *L*, the message *M* is easy to form, although if all you know are *C* and *K*, *M* takes exponentially long. There have to be huge numbers of key pairs *K* and *L* as well, so that no one with a CRAY can simply exhaust the list. Everything is made public except each particular private key *L*.

The realization offered for this by half a dozen workers in the 1970's is based on the long-tested conjecture that there is no easy way to factor numbers that are the product of two large primes—no way that is not exponential in *n*. (There are related schemes that work, and plausible unrelated proposals later shown to be doubtful.) These clever schemes lead to outrageously number-crunching ciphers. They would be absurd without digital computers; they suit the age.

Indeed, the present legal standard for the encryption of data is one such scheme. It incorporates a 56-bit key, which was adopted against the objections of the pioneers of the method. That key length is too short to be fully safe against brute-force factor searches, which may be practical for the basementful of supercomputers at the National Security Agency. The Federal agencies kept straight faces and fingers crossed when the standard was adopted, but we are told that the NSA is not prepared to approve the marginal algorithm again when it comes up soon for its 10-year review. (Indeed, in late 1988 a number 100 decimal digits long was factored into two primes by a worldwide parallel-processing combine, taking the free slack time of 400 mainframes for about one month!)

Sharp reasoning and helpful introductions, profound issues and some history, accounts of state-of-the-art computation, problems and their solutions, a fine list of references—and politics too; this is a mathematics text of unusual vivacity recommended to any who are prepared to enjoy it.

THE COLOR ATLAS OF GALAXIES, by James D. Wray. Cambridge University Press ($79.50).

A generation has grown up since Allan Sandage, whose insightful foreward graces this book, compiled his own wonderful atlas, celebrating and illustrating the Hubble classification of galaxies by their form. Those big dramatic pages in black and white are unforgettable. Then Halton Arp issued his gallery of rogues—galaxies awry, nose to nose, even interlaced—evidence in black and white that galaxies are not always loners but often evolve strongly through mutual tidal interaction. For the past decade and more a variety of galaxies have been shown to us in striking color, gaudy whorls suited to Versailles under the old regime, fascinating but a little confusing.

Here is a richly coherent "museum collection of bright galaxies" in color. It presents images of some 600 galaxies over the entire sky, north and south. The couple of hundred largest galaxies, reckoning by apparent diameter on the sky, are here complete (except for the four biggest and closest of them all, the two Magellanic clouds and the two familiar spirals of the Local Group). The rest sample the next-nearest systems of our supercluster and go beyond it to include a few more-distant examples chosen for intrinsic interest. James Wray has made these portraits over some dozen years of painstaking work at the big telescopes of the McDonald Observatory in Texas and at a couple of Chilean mountain stations. He kept scrupulous control over the color and surface brightness of his subjects by his virtuoso use of image intensifiers and filters to expose photographic color separations in three bands, giving much weight to the ultraviolet. The three distinct black-and-white images are combined by an elaborate dye-transfer technique into one colored image for reproduction.

Wray's collection gleams here, each hue and form against a dark sky. Some of the exciting photographs span a full page and others are postage-stamp size, but they all present comparable visual data, the color values virtually independent of brightness. That is not usually the case for the showy color reproductions we see, in which bright centers always bleach out to an overexposed white.

The eye cannot see these hues in the sky: ultraviolet is beyond human color range and galactic light is in general too faint for visible color. But the photographic standardization of color chosen uncovers much deeper matters. Massive newly formed stars dominate the blue knots; much older populations of cooler stars stand out in smooth yellow; here and there fluorescent gas clouds glow in violet; regions appear green either because sunlike stars are green in this system or because short-lived blue stars are superposed on staid old boroughs of yellow. (A single picture element may often sample the merged light of thousands of stars.) Dust commonly obscures and reddens; the forms of spiral arms, straight bars, thin disks, wide halos and brilliant cores hint at their dynamical stories. Galaxy by galaxy, color and form testify to evolution. For each image (along with a careful record of technique and scale) Wray has added a caption, keyed to the details of the image, in which he often outlines the story each galaxy seems to tell.

The work is opened to the beginner by text and photographs that clarify the procedure. One bright blue knot is resolved into its dozens of constituent stars within the Andromeda galaxy. Individual bright blue young stars at the hot end of the main sequence dominate, although the redder giants and supergiants show clearly. Overall the light is blue. Only when the blue stars have spent their energies, some 10 million years from now, will the knot fade, to turn first green and finally yellow, lighted then by many faint stars at the yellow-red end of the main sequence and by some bright red giants. A single revolution of a typical galaxy is ordinarily a couple of hundred million earth years long, and so blue regions are fleeting. They can form, of course, at any time from the gas that is starstuff, wherever and whenever the self-gravitating gas becomes dense enough to collapse to give birth to a star.

Normality among galaxies is marked by a bulging central yellow lens long populated by old stars. Many galaxies show nothing more. In them no new stars form. When there is a disk as well, it is often dusty and gassy, studded by newly formed blue knots marking out spiral lanes winding far from the center. There are plenty of such normal galaxies. But what shall we make of fresh blue knots scattered quite randomly over a faintly spiraling yellowish disk? The brightest ultraviolet galaxy in the entire atlas shows only blue starlight; it is currently "one single mass of star formation."

We know rather well how most stars evolve; the evolution of the galaxies is what we have to seek to understand in the next decades. This splendid compilation will challenge our proposals and illuminate our gains for a long time; in a decade when the digital image is gaining wide astronomical use, it is one of the most ambitious and attractive offerings of the past century of analog imagery.

ESSAY

How we can regain our competitive edge

by Simon Ramo

Every informed American knows that we have lost technological preeminence. Rival nations can equal our achievements in space, electronics, biotechnology and computer hardware and software. Indeed, there are many fields in which the U.S. has now been surpassed. Entire domestic industries have died, bested by offshore competitors. This is an extremely dangerous development. If we do not excel in technology, our economy will falter and our standard of living will fall.

Is the situation certain to improve? Hardly. We are investing a smaller percentage of the gross national product in civilian research and development and in capital equipment than other countries are. Our industrial managers and government officials are preoccupied with short-term concerns.

The education workers and managers alike will need if they are to be effective in an increasingly technological society receives scant attention. Test scores rate our schoolchildren's education in science and mathematics well behind that of virtually all industrial nations. The fraction of our university students studying engineering has fallen to seven in 1,000; the figure for Japan is 40. With technological advance increasingly tied to physics expertise, we have fewer physicists now than we did 10 years ago and are producing far fewer Ph.D.'s in engineering and the physical sciences. Such a trend points to disaster. Without technological leadership industry can have no competitive edge. A labor force that lacks enough intellectual flexibility to shift from one product to another is fatally handicapped, given the high obsolescence rate and short lifetime of high-technology goods and services.

No one segment of our country has a monopoly on the blame. Government, industry, educational institutions and labor organizations have all contributed to the loss of technological superiority. Yet regaining technological leadership is not impossible. New policies and programs can be devised that build on America's special strengths. Let me illustrate with one example, that of technological entrepreneurship.

America has been virtually unique in marrying risk capital to technological advance. Entrepreneurship is in our blood. Successful entrepreneurs become national heroes. We have thousands of venture-capital firms, a broad national network to tap money from large and small investors, insurance reserves, pension funds and bank trusts, a ubiquitous brokerage system for marketing the equities of small companies and a stock exchange for trading them.

Neither Japan nor Western Europe has anything comparable. Japan, Inc.—the short name for Japan's tight integrating of industry, government, banking and labor to pursue chosen technological objectives—supports their large corporations, not their entrepreneurs. An innovator who leaves the "establishment" to start a new company is regarded as an eccentric, not a hero. (Soichiro Honda told me he was pressured severely to stick to motorcycles and stay away from automobiles, which had been assigned to other, larger companies!)

The mature technological companies of the U.S. will continue to employ the bulk of our engineers; the vigor of these enterprises is essential to our economy, many facets of which can be handled only by huge entities. Yet every strength of a big corporation is offset by a weakness. With bigness come caution, risk avoidance and laborious procedures—in other words, bureaucracy, the mortal enemy of creativity. Opportunities are missed and wide gaps are left.

Note that during the first half of the century of electricity (which began in the 1880's) U.S. entrepreneurship created the world's most powerful electrical industry. Yet in the second half, now ending, the developed companies missed the computer, the greatest electrical advance of the age. Today's leading American computer companies either did not yet exist or were not electrical when, halfway through the century, the computer surfaced as a product of unprecedented potential.

The semiconductor device, the most significant invention since the wheel, replaced the vacuum tube. But not one of the top 10 vacuum-tube manufacturers became a top-10 semiconductor producer. New companies developed the area; the existing giants all failed to grasp the opportunities. (I do not count the prebreakup AT&T, which as a regulated utility could produce semiconductors only for its own use.)

Do the executives of typical large American corporations think entrepreneurially, as their founders did? Finding themselves with a major cash surplus, do present corporate leaders risk it on a long-range, speculative advanced-technology program that will lower short-term profits? More likely, to judge from recent behavior, they buy another company in an unrelated field, such as financial services. Several factors even make such behavior rational: stock-price pressure, the size of the investment, the competitiveness of short-term alternatives and the uncertain fiscal and monetary environment.

Both the entrepreneur and the financial backer need the right environment: low real-interest rates, low cost of capital, low inflation, low taxes on long-term capital gains, and high savings rates. No laws of economics need to be broken, no one's lifestyle needs to be crimped to realize these conditions. Both Japan and Germany have achieved them all.

Where will the ideas and entrepreneurs come from? The universities could stimulate a great rise in technological entrepreneurism. In the main, U.S. engineering schools teach only the underlying science and analytical tools. Yet real-world engineering involves much more. It includes designing products and systems that can be manufactured and implemented in a cost-effective way—products that either fill an old need better or create a strong new market, products that are globally competitive and can generate a proper return on invested capital. To become players in the technological entrepreneurship game, graduates must emerge with their creativity and innovative powers turned on, endowed with entrepreneurial zeal, anxious to compete. Government and industry will have to provide the strongest support if the universities are to meet this challenge.

The Bush Administration has shown indications of serious interest in the creation of an economic and political climate that would encourage high-technology entrepreneurism. Such an initiative should be among its highest priorities.

SIMON RAMO, a founder (he is the *R*) of TRW, Inc., is the author of *The Business of Science,* recently published by Hill and Wang.


"White Label"
DEWAR'S
BLENDED SCOTCH WHISKY
John Dewar & Sons Ltd
PERTH
SCOTLAND
Stay the course, and always
maintain a firm grip on the business at hand.
In Scotland, that's life in the fast lane.
The good things in life stay that way.
DEWAR'S.
"White Label".
never varies.
Authentic.
The Dewar Highlander