SCIENTIFIC AMERICAN

COMPUTER POKER

$1.50

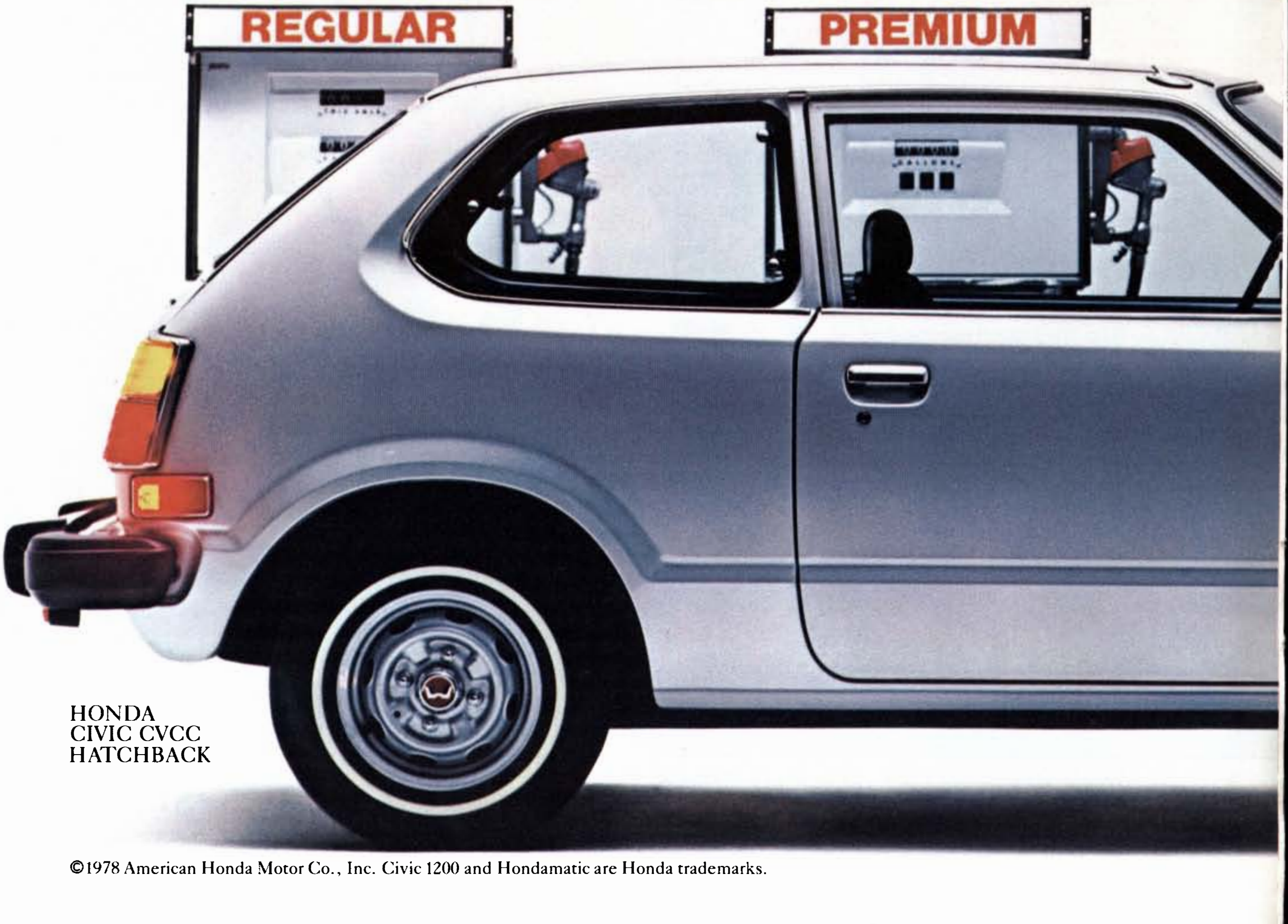
July 1978

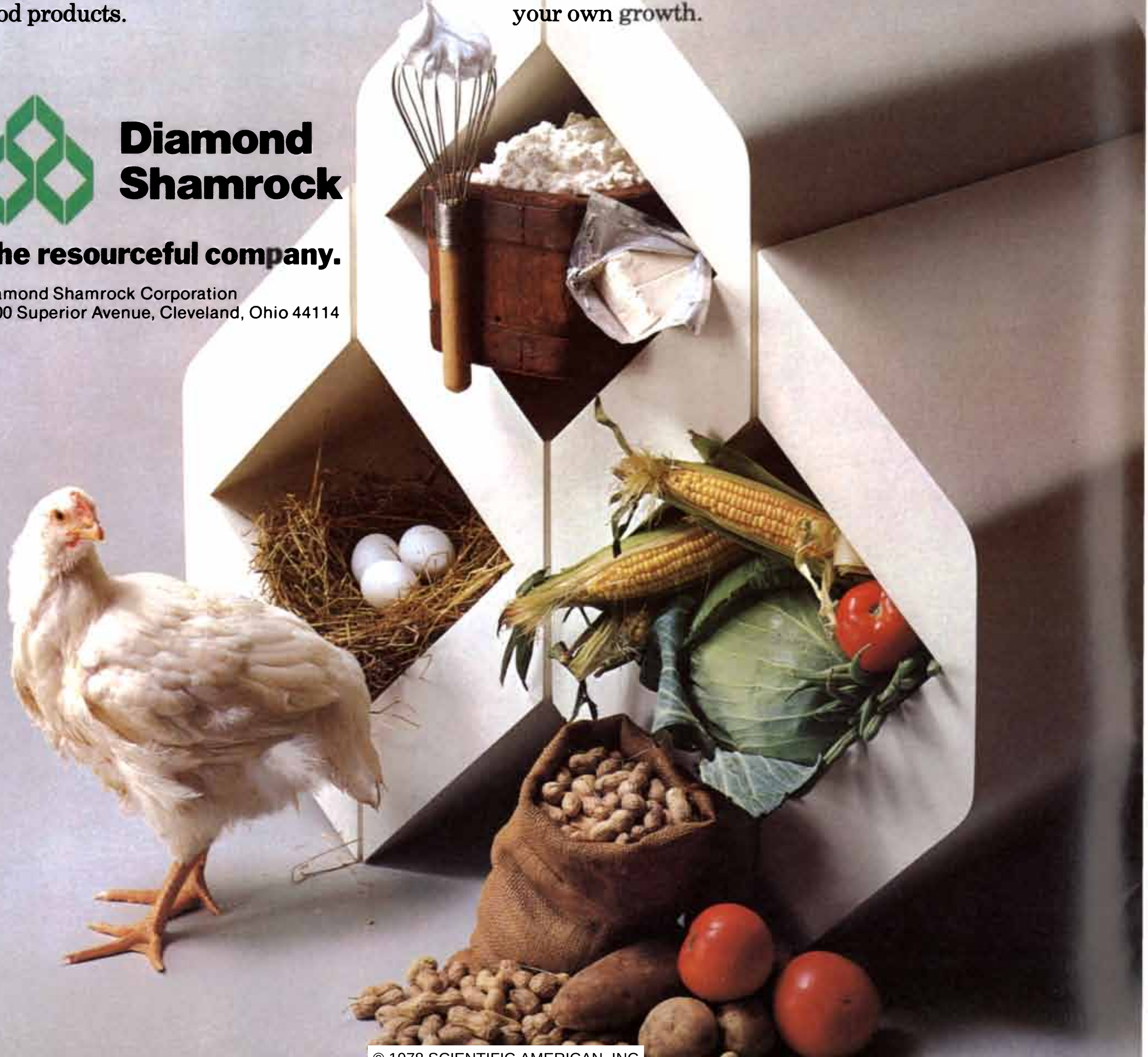
Gasoline made simple.

A lot of the new cars tend to be choosy when it comes to fuel. Most of them have catalytic converters, which means they are designed to accept nothing but unleaded gas.

Hondas are different. All our new cars – the Civic® 1200, the Civic CVCC®, and the Honda Accord® – operate without a catalytic converter. So you can take your pick of regular or unleaded gasoline. As for premium, there is no benefit and it's a waste of money.

Of course, whichever gas you use, you can go a long way on it in a Honda. Our best mileage car, the Civic CVCC 5-Speed, was rated at 47 mpg for highway driving, 37 mpg city, according to EPA estimates.



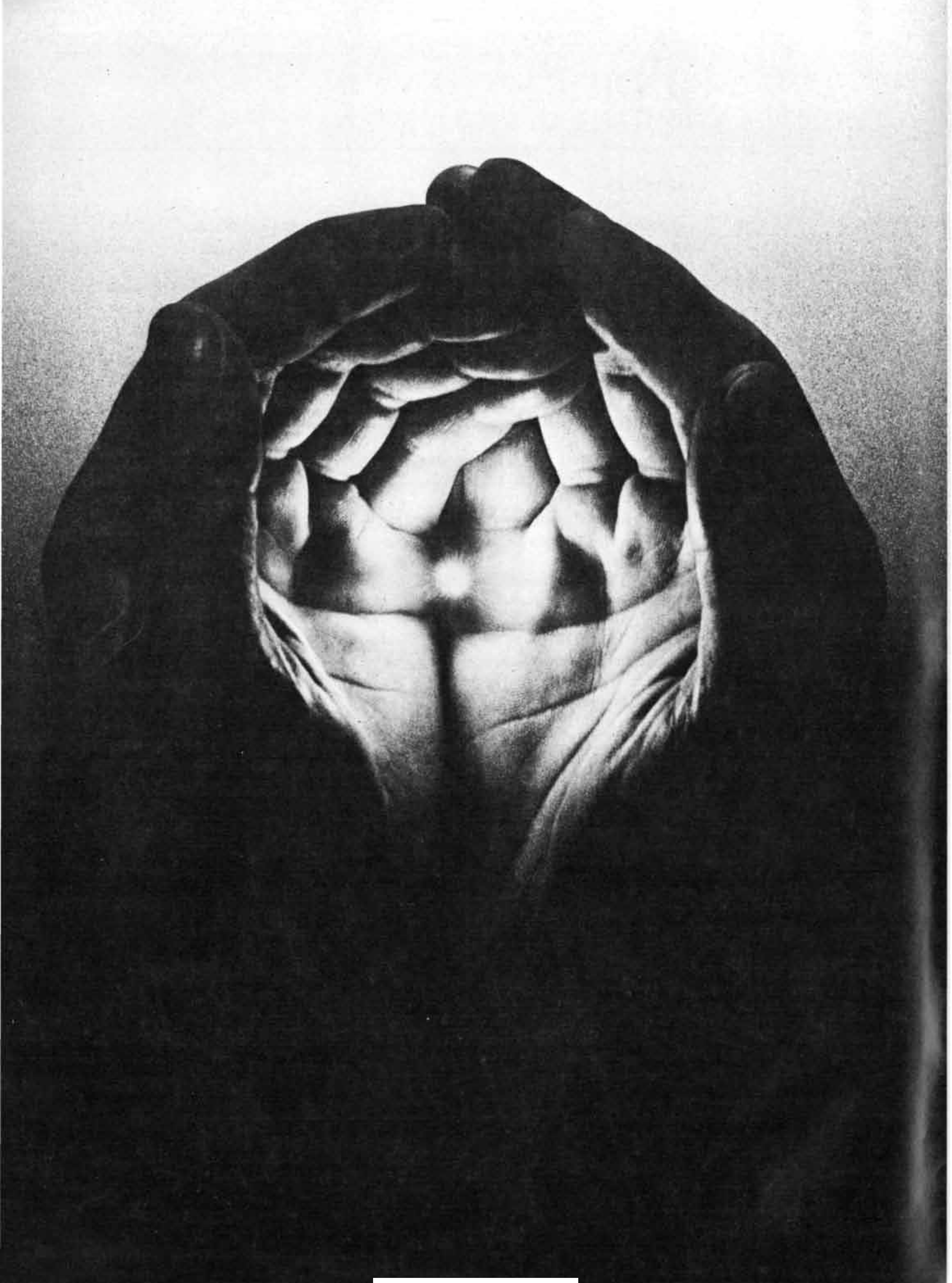


Established 1845 SCIENTIFIC AMERICAN July 1978 Volume 239 Number 1

ARTICLES

DEPARTMENTS

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. SECOND-CLASS POSTAGE PAID AT NEW YORK, N.Y., AND AT ADDITIONAL MAILING OFFICES. AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $18 PER YEAR, U.S., ITS POSSESSIONS AND CANADA; $22 PER YEAR, ALL OTHER COUNTRIES.



THE COVER

The photograph on the cover shows a poker hand (a king-high straight) displayed on the screen of a computer-terminal cathode-ray tube. Computers are now being programmed to play poker, and such a terminal can be used to monitor games in which some or all of the players are computer programs (see "Computer Poker," by Nicholas V. Findler, page 144). Poker is a particularly interesting game because it models many real decision-making situations that involve risk and uncertainty, such as economic planning, political campaigning and war. In creating poker-playing programs computer scientists hope to come to a better understanding of decision-making and thus to learn how computers can be used to solve problems usually thought to require human judgment.

THE ILLUSTRATIONS

Cover photograph by Ralph Morse

Page	Source
18–24	Ilil Arbel
30	Collection of the Library of the New York Botanical Garden
40–48	Dan Todd
49	Ford Research and Engineering Center
51	Dennis Bray, Medical Research Council Cell Biophysics Unit
52	Bunji Tagawa
53	Paul H. Patterson, Harvard Medical School
54	Bunji Tagawa
55	Paul H. Patterson, Harvard Medical School
56	Bunji Tagawa
57	Story C. Landis, Harvard Medical School (*top and bottom left*); Bunji Tagawa (*bottom right*)
58–59	Bunji Tagawa
61	Ilil Arbel
62	Meridian Airmaps Ltd.
63–68	Ilil Arbel
69	Eugene Loh, University of Utah (*top*); Ilil Arbel (*bottom*)
70	Ilil Arbel
87–90	U.S. Air Force
91	© 1969 National Geographic Society (*top*); National Aeronautics and Space Administration (*bottom*)
92	Thomas A. Croft, SRI International (*top*); Allen Beechel (*bottom*)
93	Allen Beechel
94	U.S. Air Force
95	U.S. Air Force (*top*), Bureau of the Census (*bottom*)
96–98	U.S. Air Force
103–108	Patricia J. Wynne
110	Randolf Menzel, Free University of Berlin (*top*); Patricia J. Wynne (*bottom*)
113	Amicon Corporation
114–128	George V. Kelvin
133–135	Alan D. Iselin
136	Cynthia Carey, University of Colorado
138–142	Alan D. Iselin
145	Ralph Morse
146–150	Jerome Kuhl
153	Fritz Goro
154	Michael Goodman
156	Michael Goodman (*top*), Fritz Goro (*bottom*)
158	Michael Goodman

LETTERS

Sirs:

Richard A. Muller's article "The Cosmic Background Radiation and the New Aether Drift" [SCIENTIFIC AMERICAN, May] provides a lucid and enjoyable account of one of the most important measurements in observational cosmology. It is now possible for astronomers to take account of our sun's peculiar motion within the Hubble expansion and more accurately construct the Hubble diagram (the logarithmic plot of red shift against brightness). Already this is yielding results. Observations by my colleagues and me show that one of the nearer clusters of galaxies, the Virgo cluster, has a peculiar velocity of some 655 ± 99 kilometers per second. In other words, the Virgo cluster is moving away from us by this velocity in addition to its Hubble expansion velocity, which by the latest calculations should be about 847 kilometers per second.

The peculiar velocity of the galaxy measured by Muller and his co-workers George F. Smoot and Marc V. Gorenstein, the peculiar velocity of the Rubin-Ford sample of galaxies and the peculiar velocity of the Virgo cluster are consistent with the expected motions of clusters of galaxies within bound superclusters of galaxies. This is a point that would have been worth mentioning in Muller's article. A recent suggestion made by Michael Rowan-Robinson in *Nature* (November 3, 1977) considers the possibility that an even larger-scale motion may be taking place: the rapid motion of entire superclusters relative to one another. Clearly additional observations of "the new aether drift" in the Southern Hemisphere, which Muller plans to carry out, and of additional cluster peculiar velocities will do much to help map out the fine structure of the Hubble diagram.

DONALD H. GUDEHUS

Department of Physics
Los Angeles City College
Los Angeles

Sirs:

R. Haven Wiley's Wyoming sage grouse males may gather impressive harems, but they do not "practice the most extreme polygyny known among birds" ["The Lek Mating System of the Sage Grouse," by R. Haven Wiley, Jr.; SCIENTIFIC AMERICAN, May]. The polygyny champs are still the wild-turkey brothers of Texas, described by C. Robert Watts and Allen W. Stokes in "The Social Order of Turkeys" [SCIENTIFIC AMERICAN, June, 1971]. Watts and Stokes reported no more than six wild-turkey cocks out of 170 accounted for all of the matings in four display groups, or a fecundation index of .035. Wiley reports that "at least 90 percent of the copulations [of sage grouse] are participated in by no more than 10 percent of the males present." Even assuming the 10 percent do all the impregnating, the male sage grouse get the much less impressive fecundation index of .100, certainly nothing to grouse about if you happen to be in that 10 percent but nearly three times less polygynous than their indefatigable cousins.

FRANK J. STECH, PH.D.

Mathtech
Bethesda, Md.

Sirs:

"Mathematical Games" in the May issue of *Scientific American* contains an interesting account of the work of Eric Temple Bell. Another interesting fact, which is not mentioned, is that Bell's first published article on mathematics appeared in *Scientific American* in 1916.

The editors at that time had offered a prize for the best explanation of Einstein's theory, written in language that would be understandable to a person of average intelligence. Bell's contribution won an honorable mention.

ALAN WATTON, SR.

Seattle

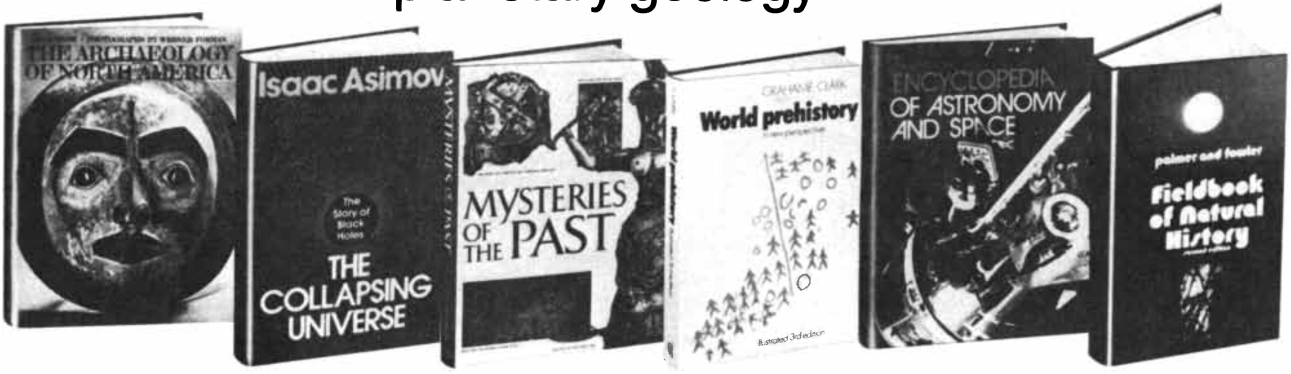
Sirs:

A few musical errors slipped past Martin Gardner's critical eye when he took up "Mozart's palindromic and invertible canon" in his report on fractal curves and "one-over-*f*" fluctuations ["Mathematical Games"; SCIENTIFIC AMERICAN, April].

Mozart scholars now agree that the canon is almost certainly not by Mozart, even though publishers have issued it under his name. For more than 40 years the compilers of the authoritative Köchel catalogue of Mozart's compositions have relegated it to the appendix of doubtful attributions, where along with three other pieces of a similar character, it bears the catalogue number K. Anh. C 10.16. We have no evidence that the piece goes back any further than the last century.

The piece is not for two singers but for two violins. Singers cannot produce the simultaneous notes of the chords in the second measure (and elsewhere), and the ranges of the parts are quite impractical. To perform the piece the two players begin from opposite ends of the sheet of music and arrive at a result that falls far below the standard of Mozart's authentic canons and other *jeux d'esprit*. The two parts combine for long stretches of parallel octaves, they rarely

The discovery of a great lost civilization . . . black holes, quasars, and the universe . . . the secret life of animals . . . planetary geology—



Sooner or later all the secrets of the natural world are revealed to the members of the

Natural Science Book Club

(Publishers' Prices shown)

40500-2. COPAN: HOME OF THE MAYAN GODS. *Francis Robicsek.* A volume of startling beauty, outsized and illustrated with full-color photographs of the monumental center of the ancient Maya. *Counts as 2 of your 3 books.* **$30.00**

54995. THE ILLUSTRATED ENCYCLOPEDIA OF ARCHAEOLOGY. *Glyn Daniel.* The significant sites, cultures, civilizations, terms, tools, and techniques packed into one all-inclusive volume. 1000 entries. More than 100 illustrations. **$17.95**

69305-2. PLANETARY GEOLOGY. *Nicholas M. Short, Ph.D.* Beautifully summarizes the data gleaned in the last few years from successful unmanned space probes, to the Moon, Mars, Venus, Jupiter and Mercury. *Counts as 2 of your 3 books.* **$21.95**

49267. FRONTIERS OF ASTROPHYSICS. *Edited by Eugene Avrett.* From recent solar research, to neutron stars, supernovae and intergalactic matter. **$20.00**

EXTRAORDINARY VALUE!

00470-3. VAN NOSTRAND'S SCIENTIFIC ENCYCLOPEDIA. *Edited by Douglas M. Considine.* Nearly 200 experts have contributed to the thoroughly revised and expanded Fifth Edition of the most authoritative single-volume source of scientific information ever assembled. Over 2,500 illustrations. *Counts as 3 of your 3 books.* **$67.50**

77050. SCIENTISTS CONFRONT VELIKOVSKY. *Edited by Donald Goldsmith.* Carl Sagan and others take a level look at the controversial views of Immanuel Velikovsky. **$8.95**

79715. SPACE AND TIME IN THE MODERN UNIVERSE. *P.C.W. Davies.* Survey of this century's discoveries in physics and chemistry, focusing on the changing ideas of space and time as applied to astronomy and cosmology. **$13.95**

Take any 3 books (values to $67.50) for only 99¢ each

if you will join now for a trial period and agree to take 3 more books—at handsome discounts—over the next 12 months

34210. THE ARCHAEOLOGY OF NORTH AMERICA. *Dean Snow.* Prehistoric Indian cultures, from the crossings into this continent to the 20th century death of the last stone-age man. 195 photos, charts and maps. **$18.95**

55000. THE ILLUSTRATED ENCYCLOPEDIA OF ASTRONOMY AND SPACE. Covers the full range of astronomical knowledge and space exploration. **$16.95**

62920-2. MONUMENTS OF CIVILIZATION—THE ANDES. *Roberto Magni and Enrico Guidoni.* Breathtaking tour of the early South American civilizations. Over 100 full-color photographs. Outsized. *Counts as 2 of your 3 books.* **$25.00**

63340-2. MYSTERIES OF THE PAST. Beautifully illustrated volume explores the riddles of past civilizations. *Counts as 2 of your 3 books.* **$34.95**

MEMBERSHIP BENEFITS • In addition to getting three books for only 99¢ each when you join, you keep saving substantially on the books you buy. • If you continue membership past the trial period, you will be eligible for our Bonus Book Plan, with savings of at least 70% off publishers' prices. • At 3-4 week intervals (15 times per year) you will receive the Book Club News, describing the coming Main Selection and Alternate Selections, together with a dated reply card. In addition, up to 4 times a year, you may receive offers of special selections, always at substantial discounts. • If you want the Main Selection, or the special selection, do nothing and it will be sent to you automatically. • If you prefer another selection, or no book at all, simply indicate your choice on the card, and return it by the date specified. • You will have at least 10 days to decide. If because of late mail delivery of the News, you should receive a book you do not want, we guarantee return postage.

48601. THE FIRST THREE MINUTES/THE COLLAPSING UNIVERSE. First, a particle physicist analyzes the first few minutes following the Big Bang. Then Isaac Asimov investigates such frontiers of astronomy as black holes, pulsars, and quasars. *The 2 count as one book.* **$17.90**

55547. INSIDE THE ANIMAL WORLD: An Encyclopedia of Animal Behavior. *Maurice & Robert Burton.* An elegantly illustrated, comprehensive survey of every facet of animal activity. **$17.50**

87890-2. WORLD PREHISTORY IN NEW PERSPECTIVE. *Grahame Clark.* A stimulating word and picture tour of man's pivotal rise from the stone age to the dawn of civilization. Examines Romans, Greeks, Babylonians, Hittites and many others. *Counts as 2 of your 3 books.* **$29.95**

48475. FIELDBOOK OF NATURAL HISTORY. *E. Laurence Palmer and H. Seymour Fowler.* The most authoritative and comprehensive guide to natural history ever published. With over 800 illustrations, it explores all major plant and animal species and most rocks and minerals. **$17.50**

Natural Science Book Club 4-A4E
Riverside, New Jersey 08075

Please accept my application for membership and send me the three volumes indicated, billing me only 99¢ each. I agree to purchase at least three additional Selections or Alternates during the first 12 months I am a member, under the club plan described in this ad. Savings range up to 30% and occasionally even more. My membership is cancelable any time after I buy these three books. A shipping and handling charge is added to all shipments.

3 books for 99¢ each
Indicate by number the books you want.

A few expensive books (noted in book descriptions) count as more than one choice.

Name ______________________

Address ______________________

City ______________________

State ____________ Zip ____________
(Offer good in continental U.S. and Canada only. Prices slightly higher in Canada.)

achieve even the most rudimentary rhythmic or directional independence and their harmony consists of little more than the most elementary writing in parallel thirds. This little counterfeit is not nearly as interesting as Mr. Gardner's columns.

IRVING GODT

Associate Professor of Music History
Music Department
Indiana University of Pennsylvania
Indiana, Pa.

Sirs:

Jearl Walker's article on kites in "The Amateur Scientist" [SCIENTIFIC AMERICAN, February] reminded me about the remarkable ones made on the island of Bali in Indonesia. Numerous competing clubs (of rice farmers and even of state police!) construct and fly huge fish-shaped bowed kites that often measure more than 20 by 20 feet. At the bottom is attached another large lateral bow strung with a thick strip of rattan or inner tube. This aeolian harp vibrates so strongly in the wind that it produces a complex sound with a powerful low fundamental. During the peak of the season the fleets of kites that fill the sky sound like thrumming propellers.

ANDREW TOTH

Los Angeles

Scientific American, July, 1978; Vol. 239, No. 1. Published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Gerard Piel, president; George S. Conn, vice-president; Dennis Flanagan, vice-president; C. John Kirby, vice-president; Donald H. Miller, Jr., vice-president and secretary; Arlene Wright, assistant treasurer.

Editorial correspondence should be addressed to The Editors, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Manuscripts are submitted at the author's risk and will not be returned unless accompanied by postage.

Advertising correspondence should be addressed to C. John Kirby, Advertising Director, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017.

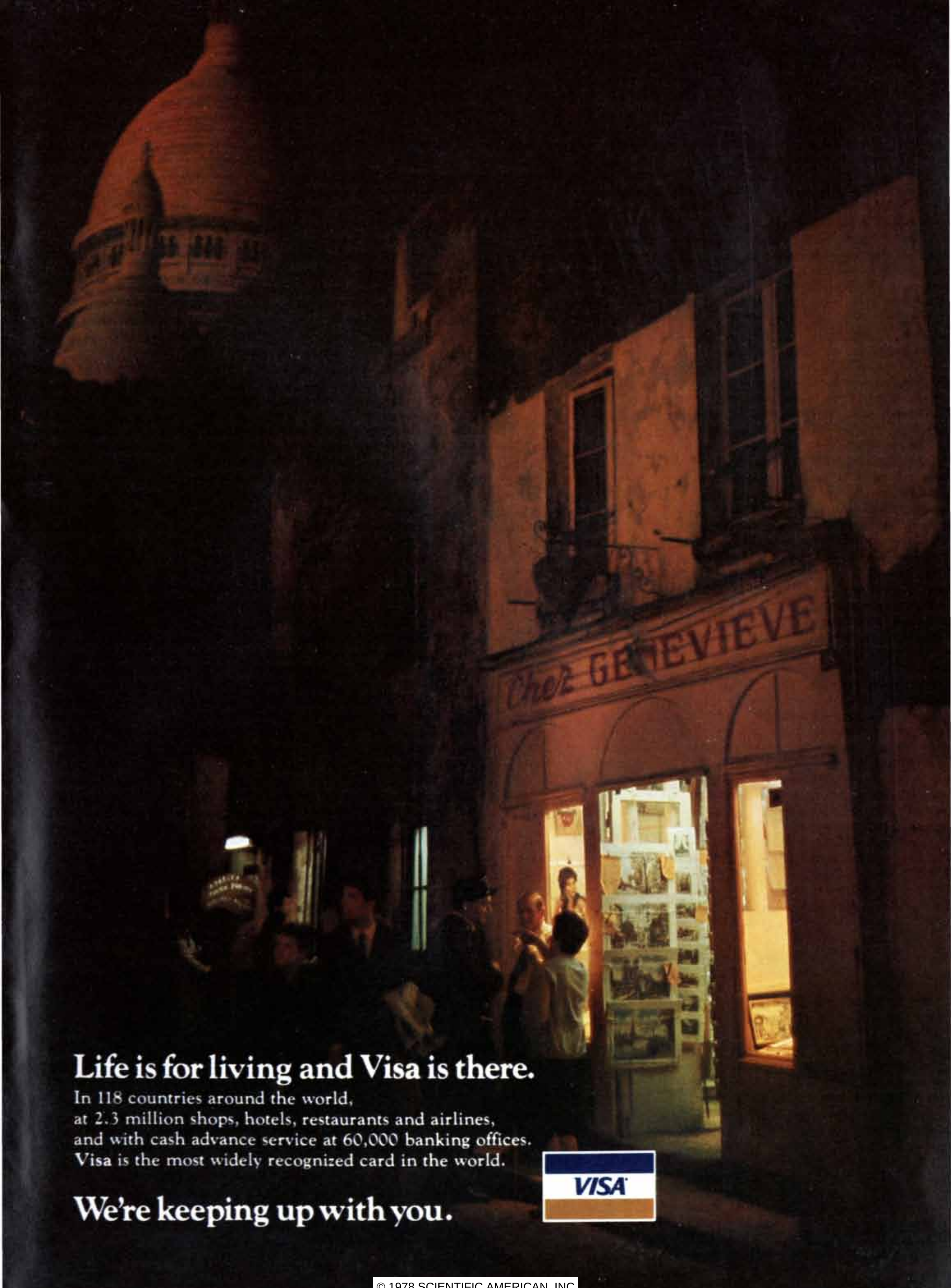
Offprint correspondence and orders should be addressed to W. H. Freeman and Company, 660 Market Street, San Francisco, Calif. 94104. For each offprint ordered please enclose 40 cents.

Subscription correspondence should be addressed to Subscription Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. For change of address, notify us at least four weeks in advance. Send both old and new addresses and enclose an address imprint from a recent issue. (Date of last issue on your subscription is shown at upper right-hand corner of each month's mailing label.)

Name ____________________

New Address ____________________

Old Address ____________________


Life is for living and Visa is there.
In 118 countries around the world,
at 2.3 million shops, hotels, restaurants and airlines,
and with cash advance service at 60,000 banking offices.
Visa is the most widely recognized card in the world.
We're keeping up with you.
VISA

Who do you call to help speed up the firemen speeding to a fire?

Recently, the San Francisco Fire Department came to PRC with a very intriguing problem.

Fires.

Could PRC develop a dispatch system to help stop fires faster?

On the average, an incipient fire doubles in size almost *eight times* per minute.

So every minute saved can help save property—and lives.

Ask 5,000 technical experts and you'll get an answer.

To a challenge of this complexity, PRC brings uncommon resources—the unique strength and professional skills of more than 5,000 experts in 45 countries around the world.

In San Francisco, a PRC project team analyzed the Fire Department's requirements, then developed an advanced computer dispatch system that maximizes situation control and actually eliminates up to five manual steps—saving 30 to 90 seconds in response time to every fire.

The largest diversified professional services company in the world.

In fact, for every problem you can think of, Fire damage can accelerate at blistering speed. A computerized dispatch system developed by PRC now helps the San Francisco Fire Department save critical response time to fires, accidents—more than 1,000 life-or-death resuscitation emergencies every year.

chances are there's a PRC team of specialists already working on answers.

Planning a dam system in the Philippines.

Monitoring Space Shuttle facilities at Cape Kennedy.

Designing a criminal justice data system for New Zealand.

Developing the first national real estate computer service in the U.S.

The list of PRC projects goes on. And on.

Planning, engineering, architecture. Management consulting. Information services. We're here to help.

The point is, there's no problem too technical, too complex or too big for Planning Research Corporation to handle.

Come to think of it, what could PRC do for you? Planning Research Corporation, 1850 K Street, N.W., Washington, D.C. 20006. Call (202) 293-4700.

Call

prc
Planning Research Corporation

NEW CHEVY MONZA.

THE ONLY ECONOMY CAR NAMED AFTER A RACE TRACK.

'78 Chevy Monza 2+2 Hatchback Coupe.

Sporty. Monza, Italy. One of the fastest race tracks in the world. And like its namesake, Chevy Monza is a thrill to drive. To take out on the open road and enjoy. To take into a corner and come out impressed.

Well equipped. Every Monza sports such standard equipment as bucket seats, white stripe tires, bumper rub strips, carpeting, and more. Things that are likely to cost extra on some other cars.

Low priced. With all kinds of standard equipment and exciting moves, you might think that Monza is going to be priced way over your head. Well, think again, because a new Chevy Monza 2+2 Hatchback Coupe isn't an expensive car, it's just expensive looking. Probably well within your budget. What's more, Monza is quite economical to operate. It has EPA mileage ratings that would impress even the most frugal driver.

EPA ESTIMATES: 34 MPG HWY/24 MPG CITY. With standard 151 CID engine and 4-Speed manual transmission. California EPA estimates 31 mpg hwy, 21 city. Your mileage will vary depending on how and where you drive, the condition of your car and available equipment. Monza is equipped with GM-built engines produced by various divisions. See your dealer for details.

Chevrolet

SEE WHAT'S NEW TODAY IN A CHEVROLET.

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

JULY, 1928: "The difference between the Schrödinger atom and the Bohr atom is, to begin with, one of distribution of the electricity. In the Bohr atom there was a positive charge on the central nucleus and negative charges on each planet or electron. The electric charge was localized in spots. But in Schrödinger's atom the charge is spread everywhere throughout the volume of a little sphere of atomic dimensions. Again, the electrons in Bohr's atom were in rapid motion in their orbits, while in the Schrödinger atom the electric charge does not move about. It does, however, change its intensity at different points in the sphere at different instants of time. This fluctuation in the strength of the electric charge sets up light waves in the surrounding space."

"Captain George Hubert Wilkins literally leaped to world fame when he flew across the top of the world from Alaska to Spitsbergen in 21 hours. Captain Wilkins is an Australian. Back of this epochal flight is a long period of service and training in exploration, first as second in command under the explorer Stefansson in the Arctic, then as second in command with the British Imperial Antarctic Expedition of 1920–21, later as naturalist with Shackleton in 1921–22 and in 1923–25 as leader of an expedition for the British Museum. His recent flight, together with evidence provided by some previous deep soundings he had made in the Arctic, practically destroys the force of the old Harris theory of a large Arctic land mass and reinforces the Nansen theory that the Arctic is a great deep ocean basin, similar in this characteristic to those of the other great oceans of the world."

"A growing recognition of the right of the citizen to be protected against offenses to his senses of sight, smell and hearing will mark the future of city administration. Already we have provisions designed to secure for the individual his proper share of light and air. But we think the most serious trespass against the comfort of dwellers in cities, and particularly in those of great size, is the matter of noise, and for ear-splitting, nerve-shattering din we know of nothing, not even the roar of an elevated railway, to compare with the racket of the riveting hammer. We make no complaint against the pneumatic riveter as such; it is one of the most ingenious and efficient of the many tools that have contributed to the great architectural and engineering achievements of the present age. But the time has come when the infernal racket of the machines should be prohibited on buildings that are being erected in districts that are already crowded with office or residential buildings. A substitute for the pneumatic riveter is available in electric welding."

"Although various methods of transmitting and receiving images of moving objects by radio (television) have been described at length in this publication and others, there is one point that must be stressed strongly: beware of any smooth-tongued salesman who attempts to inveigle you into investing in a television-promotion scheme. Although there are several reliable companies assiduously bent on developing worthwhile methods of television, there are (and there will be more) companies not so reliable that are more interested in moving the contents of your pocketbook than they are in moving pictures by radio. Any scheme for television development should be thoroughly investigated before investing. And the investigation should include not only the members of the company but the television system as well. The services of someone well versed in radio will be needed here, but the end will be worth the trouble."

SCIENTIFIC AMERICAN

JULY, 1878: "Mr. Edison is said to have obtained very satisfactory results with a telephone constructed upon the general plan set forth in Gray's caveat, *i.e.,* a variable resistance controlled by the vibrations of a diaphragm. Edison made the discovery that plumbago possessed the curious property of altering its electrical resistance in proportion to the pressure to which it is subjected, and availed himself of this discovery in the construction of his telephone. More recently the same experimenter is said to have obtained still better results by the use of carbon in the form of lampblack, from the smoke of an ordinary hydrocarbon lamp, compressed into a cylindrical button."

"By means of a series of cameras standing one foot apart and operated by electricity a California photographer, Mr. E. Muybridge, has succeeded in taking negatives of every phase of a trotting horse's action while making a complete stride. In this way it becomes possible to study the successive positions of a horse's body, legs and feet while it is going at full speed. The horse photographed was Mr. Leland Stanford's trotter Occident, which was traveling at a 2:24 gait, with a stride of 18 feet six inches. The photographs show the trotter's feet to be all off the ground together twice during the making of the stride, contrary to the assertions of the authorities hitherto accepted."

"The Whitehead torpedo consists of a cigar-shaped case of thin steel built in sections well screwed together, about 17 feet long from end to end, with each section 15 inches in diameter at the widest. The first compartment, at the head, contains the charge of the gun cotton, to be fired by the forcing of a roughened pin into a cap of fulminate, on the torpedo coming in contact with anything after it has been set in motion. The second compartment contains Mr. Whitehead's great secret contrivance, which gives the operator control over the machine, so that he can make it run at any required depth under water. The next section of the torpedo is the reservoir for compressed air, the motive power by which it runs under water. Then comes the machinery, and last the screw and the rudders. The depth having been set, and the amount of pressure in the reservoir for the required speed (which works up to 12-1/2 knots), the torpedo is run into the impulse tube. That tube is an affair very much resembling a telescope; it is forced out by compressed air and, pushing the tail of the fish, gives it a good start on its journey clear of the ship."

"It is known that nearly all the dark lines of the visible solar spectrum correspond with bright lines of the spectra of metallic vapors. M. Cornu has been able to extend this study further into the ultra-violet end of the spectrum and the dark lines it contains, and he has been led to make an attempt at quantitative analysis of the elements of the sun's absorbent layer. He finds that the vapor of iron is the most abundant. We are led to the idea that if all the bodies of the solar system—and perhaps all sidereal bodies—have a common origin, this would be revealed by the presence in each of iron in considerable proportion. Our globe appears to favor this idea. Its mean density, which is 5.5, is nearly double the mean density of the elements forming the superficial crust. We must therefore suppose that the central part of the earth is formed of matter much denser than stony materials, namely of metallic masses. If we think of the directive force of the magnetic needle on different points of the globe, and the symmetry of all these positions to certain great circles of the terrestrial sphere, it seems highly probable that the metallic masses of the center of the earth are formed to a great extent of metallic iron."

THE AUTHORS

DAVID GORDON WILSON ("Alternative Automobile Engines") is professor of mechanical engineering at the Massachusetts Institute of Technology. Born in England, he was educated at the University of Birmingham and the University of Nottingham, where he received his Ph.D. in mechanical engineering in 1953. After working for a time in private industry as a gas-turbine engineer, he came to the U.S. in 1955 on a Commonwealth Fund Fellowship for two years of postdoctoral study at M.I.T. and Harvard University. He then returned to England, where he worked on the design of gas turbines for two more years before taking a position as senior lecturer at the Nigerian College of Arts, Science and Technology. From 1960 to 1966 he was back in England again, serving as vice-president and technical director of the Northern Research and Engineering Corporation. In this period he directed research projects on centrifugal fluid machines, radial-inflow turbines and compact heat exchangers. He joined the M.I.T. faculty in 1966 and became a U.S. citizen in 1972. In the past few years he has been involved primarily in studies of solid-waste management, energy-conservation methods and nonpolluting transportation alternatives.

PAUL H. PATTERSON, DAVID D. POTTER and EDWIN J. FURSHPAN ("The Chemical Differentiation of Nerve Cells") are neurobiologists on the faculty of the Harvard Medical School. Patterson, who has been at Harvard for eight years, acquired his Ph.D. in biochemistry from Johns Hopkins University. Potter and Furshpan came to Harvard in 1959, following a period as postdoctoral fellows in the department of biophysics at University College London. Potter's doctorate, in biology, is from Harvard; Furshpan's, in animal physiology, is from the California Institute of Technology. The work described in their article was done with Dennis Bray, Robert B. Campenot, Linda L. Y. Chun, Philippa Claude, Karen Fischer, Story C. Landis, Peter R. MacLeish, Doreen McDowell, Richard E. Mains, Colin A. Nurse, Kunihiko Obata, Paul H. O'Lague, Louis F. Reichardt, Patricia A. Walicke and Michel Weber.

JOHN LINSLEY ("The Highest-Energy Cosmic Rays") is research professor of physics at the University of New Mexico. He obtained his undergraduate and graduate degrees at the University of Minnesota, where he participated in experiments designed to study the composition of primary cosmic rays by a variety of techniques, all involving the use of high-altitude balloons. In 1954 he joined a group at the Massachusetts Institute of Technology, headed by Bruno Rossi, to work on the Agassiz Station air-shower array, the first such installation capable of determining both the energy and the direction of cosmic-ray air showers. Four years later he and Livio Scarsi built the first giant air-shower array, modeled after the Agassiz installation, at Volcano Ranch, near Albuquerque, N.M. Since then Linsley and a succession of co-workers have used this facility to study many aspects of giant air showers and the high-energy cosmic rays that generate them.

THOMAS A. CROFT ("Nighttime Images of the Earth from Space") is on the staff of SRI International (formerly the Stanford Research Institute). After obtaining his master's degree from Dartmouth College, Croft became a Navy fighter pilot and later worked as an aerospace engineer. The latter experience convinced him that his interests lay in research; he therefore went back to school, acquiring his Ph.D. from Stanford University in 1965. He continued to work at Stanford until 1976, when he took up his present position. His research was concentrated at first in the field of ionospheric radio transmission, with a special emphasis on digital computer simulations. He has also done solar-wind experiments and was a member of both the Pioneer Venus and the Voyager radio-science teams. His interest in nocturnal satellite imagery was triggered during a recent radar study he conducted for the Air Force, when the then-classified images of the Defense Meteorological Satellite Program were supplied to him as reference material.

RANDOLF MENZEL and JOCHEN ERBER ("Learning and Memory in Bees") are with the Institute of Animal Physiology of the Free University of Berlin, where Menzel is professor and Erber assistant professor. Both went to Berlin in 1976 after spending a year doing research at the Australian National University in Canberra. Before that they were both at the Technical University of Darmstadt, Menzel as assistant professor and Erber as a graduate student. Menzel studied zoology, chemistry and physics at the University of Frankfurt and the University of Tübingen; he obtained his Ph.D. in zoology from Tübingen in 1967. Erber, who began his studies in electrical engineering at Darmstadt, went on to obtain his Ph.D. in zoology there.

NICHOLAS V. FINDLER ("Computer Poker") is professor of computer science at the State University of New York at Buffalo. He holds an undergraduate degree in electrical engineering and a doctorate in mathematical physics, both from the Budapest University for Technical Sciences. He left Hungary in 1956, and after a short stay in Vienna he lived and worked in Australia until 1963, when he moved to the U.S. He has been at his present post since 1966, except for a sabbatical year in 1972–73 when he was a Senior Fulbright Scholar at the Technical University of Vienna. Concerning the project reported in this article, he writes, he is "indebted to many poker players, ranging from suckers to experts, for their ideas and criticism. Our research team, affectionately called the poker group, has included Heinz Klein, John Menig, Zachary Levine, Channing Johnson, Jean Rachlin, Gary Higgins, William Gould, Alex Kowal, Jeff Lesinski, David Reed, Steve Feuerstein, Paul Bunting, David Ziffer, John Prieur, André van Tilborg, John Doughtie, Charles Pearson, George Sicherman, Terry Roy, Steve Hagler, Danny Kolis, João Martins and Ernesto Morgado."

HARRY P. GREGOR and CHARLES D. GREGOR ("Synthetic-Membrane Technology") are father and son. The senior Gregor is professor of chemical engineering and applied chemistry at Columbia University. His doctoral work was done at the University of Minnesota, where he came into contact with refugee scientists from Germany, particularly Karl Sollner, under whose guidance he made the first "permselective" membranes. He has been involved ever since in the synthesis, characterization and utilization of various types of membranes, and he holds numerous patents on membranes and membrane processes. His son Charles is a graduate of Carleton College and has a master's degree from Columbia. He spent the past year in Jerusalem studying the Talmud, and he is now at the State University of New York at Stony Brook, where he is working toward a doctorate in physical-organic chemistry.

WILLIAM A. CALDER III ("The Kiwi") is professor of ecology and evolutionary biology at the University of Arizona. A graduate of the University of Georgia, he obtained his Ph.D. from Duke University in 1966. He taught at Duke and at the Virginia Polytechnic Institute before moving to Arizona in 1969. In addition to his university teaching and research tasks he is a life member and trustee of the Rocky Mountain Biological Laboratory. In 1977 he was visiting professor at the University of New South Wales in Australia, a sabbatical, he writes, "spent largely with the curious eggs, incubation and metabolism of New Zealand's kiwis, birds unique among birds in every respect, but known to Americans as merely a trade-design on shoe-polish cans."

MATHEMATICAL GAMES

On Charles Sanders Peirce: philosopher and gamesman

by Martin Gardner

I could make the whole matter clear to you as the noonday sun, if it were not that you are wedded to the theory that you can't understand mathematics!
—from a letter of Charles Sanders Peirce to William James

Most of the famous philosophers of the past had little talent for mathematics, but there are some notable exceptions. Descartes, Leibniz, Pascal, Whitehead, Russell—they are as eminent in the history of mathematics as in the history of philosophy. To this small, select band belongs Charles Sanders Peirce (1839–1914), scientist, mathematician, logician and the founder of pragmatism. In the opinion of many he was America's greatest philosopher.

Peirce was trained in mathematics by his father, Benjamin Peirce, the leading U.S. mathematician of his day, but of the two Charles was by far the more original. His contributions to logic, the foundations of mathematics and scientific method, decision theory and probability theory were enormous. It is remarkable how many later developments he anticipated. At a time when the infinitesimals of early calculus were in disrepute Peirce insisted on their usefulness, a view only recently vindicated by the invention of nonstandard analysis. At a time when determinism dominated physics Peirce's doctrine of "tychism" maintained that pure chance—events undetermined by prior causes—are basic to the universe. This is now essential to standard quantum mechanics. Even Peirce's notion that natural laws are "habits" acquired by a growing universe is no longer as eccentric as it once seemed. There are respectable models of oscillating universes in which random events create a different set of constants at each bounce. As each cosmos explodes it develops laws, some of which change as the universe ages.

Peirce's influence on William James, a longtime friend whom he adored, was much greater than the other way around. The basic idea of pragmatism, including the word itself, was introduced by Peirce in a popular magazine article. James picked up the word and enlarged on Peirce's suggestions in a series of brilliant lectures that became the book *Pragmatism*. Peirce later became so annoyed by what he considered James's reckless exaggerations that he changed his word to "pragmaticism," a term so ugly, he declared, that no one would kidnap it.

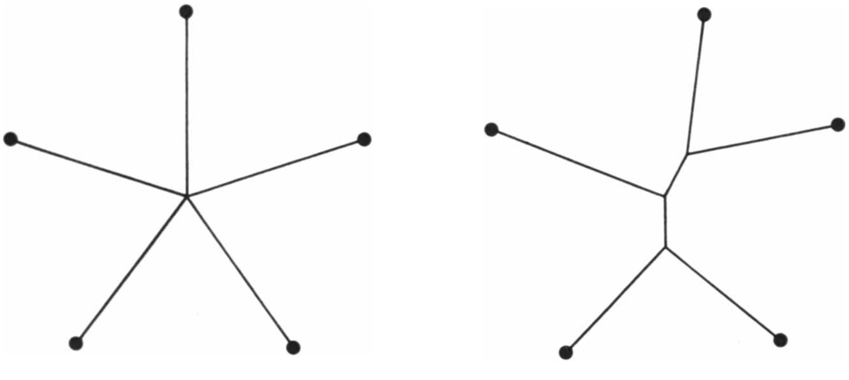
Like all creative mathematicians Peirce enjoyed mathematics hugely as a form of intellectual play. As a child he had had an intense interest in chess problems, puzzles, mathematical card tricks and secret codes. This sense of amusement runs through all his mature writings. He even coined the word "musement" for a mental state of free, unrestrained speculation, not quite as dreamy as reverie, in which the mind engages in "pure play" with ideas. Such a state of mind, he maintained, is the first stage in inventing a good scientific hypothesis. One meditates on all the relevant data, then pushes them around in one's head to form new combinations (like pushing chess pieces to solve a chess problem) until comes that mysterious flash of insight.

In a little-known paper titled "A Neglected Argument for the Reality of God" Peirce argued that "musement" is not only a road to theism but also the only road. It is a leap comparable to the scientist's conjecture, although it is one of the heart rather than of the head. It is not testable, but for those who make it, Peirce wrote, it can be as certain as the belief in one's own existence or the existence of others. It was on such matters of "over belief" (James's term) that Peirce and James agreed.

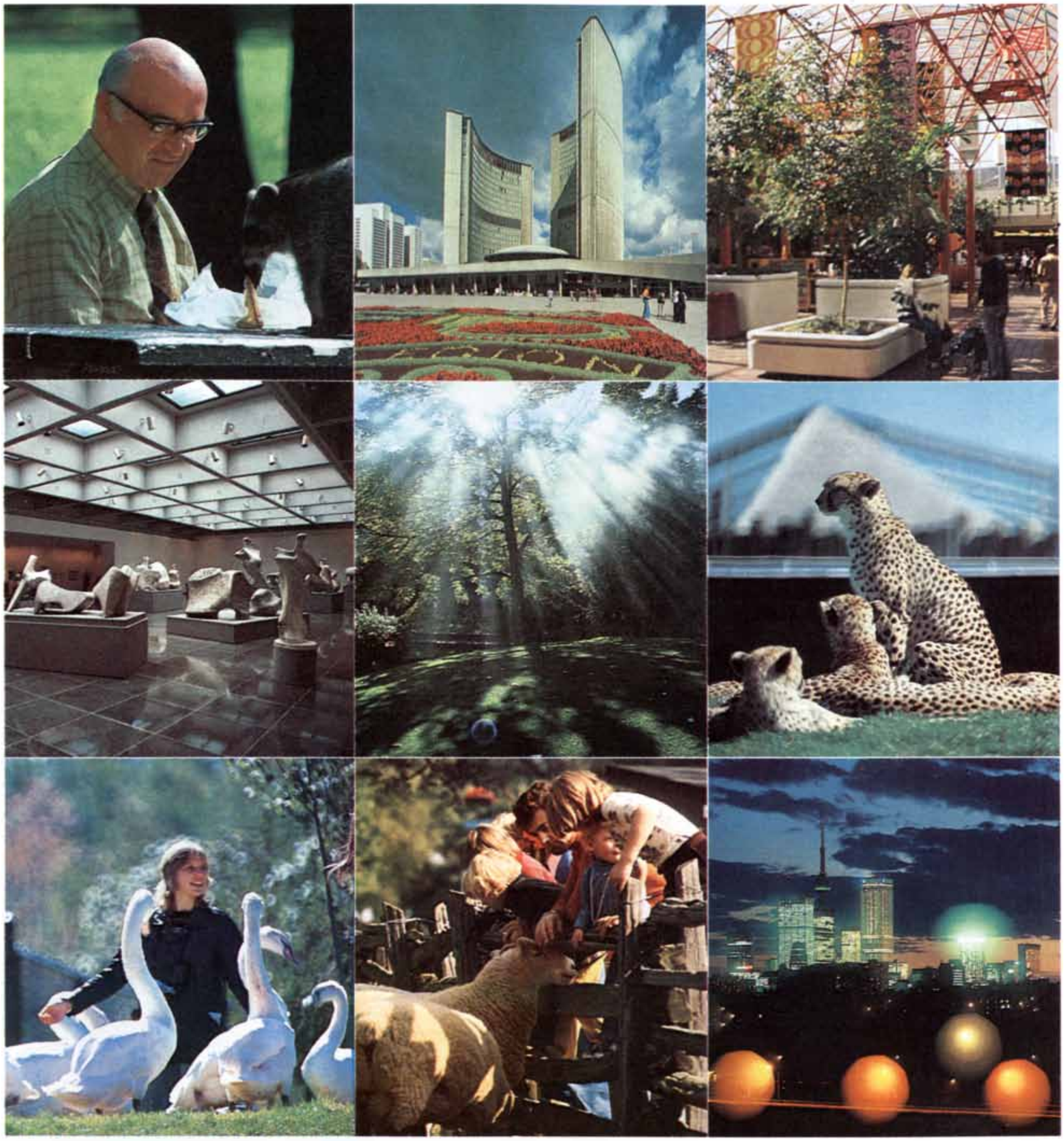
Peirce's recreational approach to mathematics is most evident in his views on how mathematics should be taught to children. He was convinced that the methods then in use produced only dunces. The manuscripts of his three unpublished textbooks are filled with novel ways of using puzzles, games and toys for introducing mathematical concepts. For example, Zeno's paradoxes lead into discussions of the continuum and the limit. Projective geometry and the shadows of a rotating wheel illuminated by a lamp introduce infinity. Peirce recognized—this before 1900!—the great value of elementary topology (he called topology the "easiest, most elementary and most fundamental branch of geometry") in stimulating a child's mathematical imagination. Euler's formula for the skeletons of polyhedrons, knot theory, graph theory, the four-color-map conjecture (which Peirce tried vainly for decades to prove), the Moebius strip—these are only some of the topological topics Peirce used to arouse student interest. He delighted in asking teachers to let him instruct a group of youngsters who detested mathematics and seemed incapable of learning it. He records that in one case, after about 10 lessons, two of his "prize stupids" led the school.

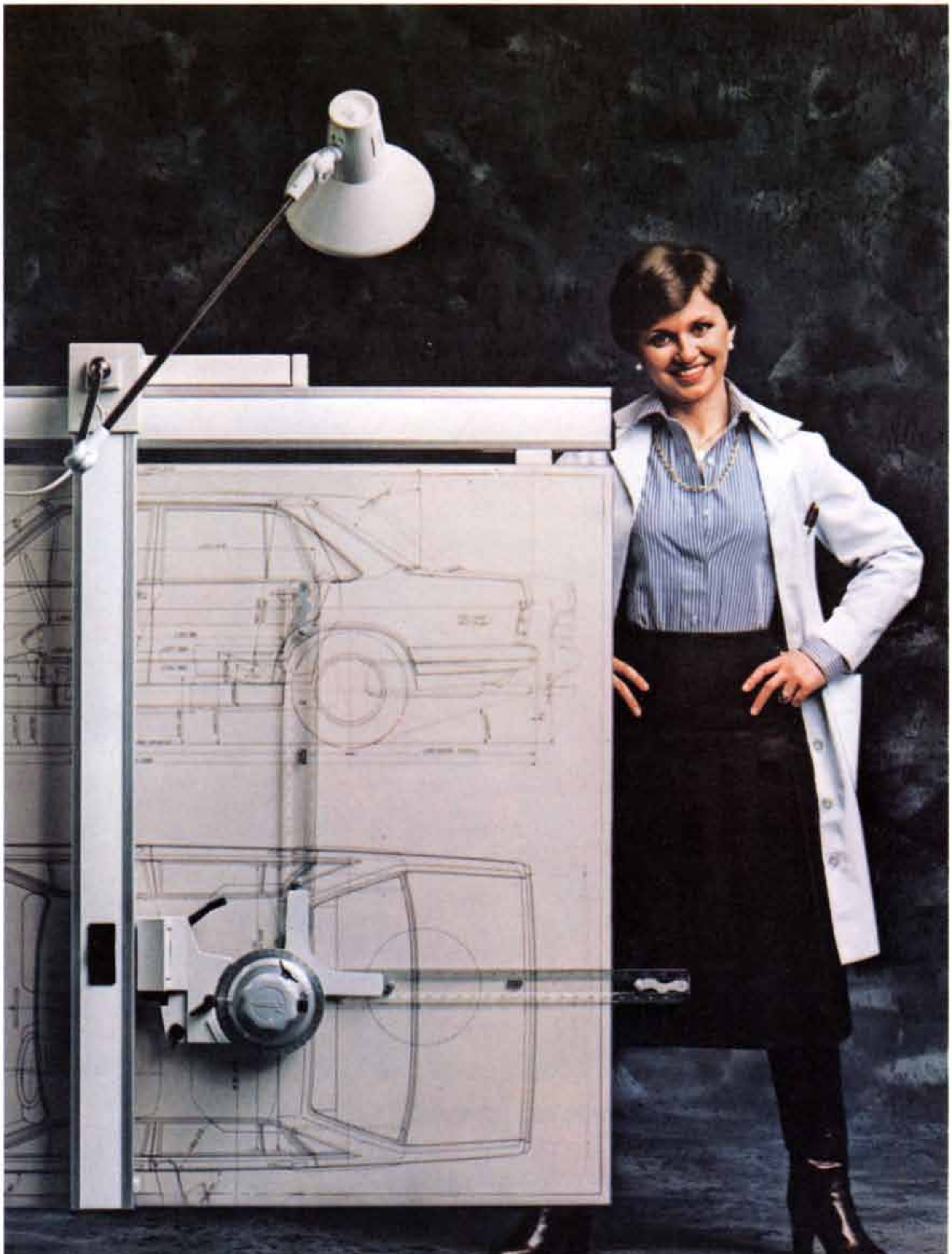
To teach arithmetic Peirce recommended the constant use of counters such as beans, the early introduction of binary notation, the use of 101 cards numbered 0 through 100 and other devices now common in grade school instruction. In one textbook he wanted to insert a cardboard mechanical gadget for doing multiplication. "The objection to inserting this," he jotted in a notebook, "would be that the teachers would not understand the mathematical principle on which it depends, and might therefore be exposed to embarrassing questions."

The use of playing cards is also rec-



How Peirce reduced fifthness to thirdness





"IN MY FAMILY, I PICK THE CAR AND LET MY HUSBAND PICK THE COLOR."

HERMINE FINKENZELLER, AUDI 5000 STRUCTURAL ENGINEER

Isn't it usually the other way round? Finkenzeller: Well, I suppose it's a bit unusual for a woman to be a structural engineer. But I am and it happens that my husband is not. Actually, he's very good at his job, but he's in an entirely different field. I've been designing automotive components for 15 years. These drawings of the Audi 5000 are the ones I did as a member of the engineering team that designed the car.

Was it hard getting started as an automotive engineer? Finkenzeller: In the very beginning, yes. But that was years ago. Today, it's easier for a woman to be accepted as an engineer. Consider how many women chemists and physicists there are now. Things have opened up. I'm respected by the people I work with for what I can contribute as a structural engineer.

But don't men usually know more about cars than women? Finkenzeller: Why do you assume that? Because they played with a little red car, as children, and girls played with dolls? That hardly qualifies a man as an automotive authority. In Europe, some men seem to know technical terms about cars. But a lot of men think in romantic terms about cars. Behind the wheel they see themselves as something they're not, perhaps race drivers. Women don't seem to have this problem. Their attitude toward cars is more rational. A car satisfies their needs, not their fantasies.

How will a woman feel about the Audi 5000? Finkenzeller: Good. They will like its layout. It's probably the most intelligent 15½ feet of engineering on the road today. Women will appreciate that we didn't devote half the car to the power plant. You can't sit people under a hood. My colleagues developed a 5-cylinder gasoline engine that has plenty of power, yet doesn't take up unnecessary space. So the car seats five people very comfortably. It is a big car. As a matter of fact, I believe in your country it's the largest German luxury car for the money, less than $9,000.*

Will men like the Audi 5000? Finkenzeller: Men? Yes. I think they will love its power and handling – and a lot of women will, too. The Audi 5000 may be a rather elegantly conservative car but it's not sedate. It's very fast. With front-wheel drive, it takes corners beautifully, especially for such a big, luxurious car. It's really a lot of fun to drive. People are surprised to find out how responsive the car is. That amuses us. And, of course, delights us, too.

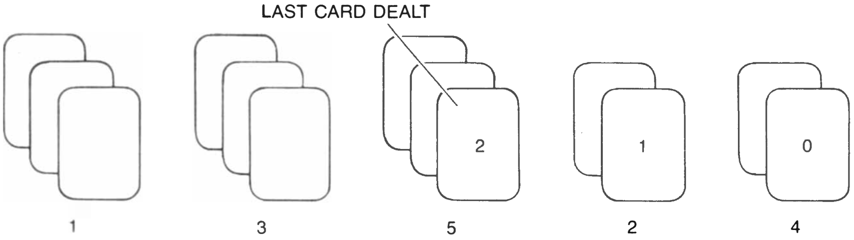
Do you own an Audi 5000? Finkenzeller: No. Not that I wouldn't like to. It's just that my family has no need for a car with all that room. There's only my husband and our one child. What would we do with all the room there is in the Audi 5000? So, instead, we own the Audi Fox, I think that's what you call it in America. It's smaller. But it's also a very nice car. Do you know, I worked on the design of that car, too.

Are you always this sensible? Finkenzeller: Sensible? If you mean logical and precise, I would say yes when it comes to doing my work. In my job, I have to be very precise. But, if what you really want to know is whether I'm ever emotional or even romantic, perhaps you should ask my husband about that.

*Suggested 1978 retail price under $9,000 P.O.E., transp., local taxes, and dealer delivery charges, additional. Come in and test-drive the Audi 5000 at your local Porsche+Audi dealer.

Which Borkum Riff?

For all of the flavor and none of the bite...Borkum Riff.



How to assemble five piles of cards in Peirce's card trick

ommended. "If you will provide yourself, my dear Barbara, with a complete pack of cards with a joker, 53 in all, I will make a little lesson in mathematics go down like castor-oil in milk." Barbara is a character in one of Peirce's unpublished textbooks. She is so named because "Barbara" was the medieval mnemonic name for the syllogism "All *A* is *B*, all *B* is *C*, therefore all *A* is *C*."

In the introduction to another textbook Peirce devotes 15 pages to ticktacktoe! The game is used for showing how a theorem is first guessed, then tested by manipulating diagrams. "Such are the tools," he writes, "with which the mathematician works." Not until recent years, with the huge success of textbooks such as Harold R. Jacobs' *Mathematics, A Human Endeavor,* have some teachers caught up with the proposals in Peirce's unsalable manuscripts.

Like so many other mathematical geniuses—Leibniz and Kepler come to mind—Peirce sometimes became overenthusiastic, almost obsessively so, about some of his creations. This may have been partly the result of his working alone, without the give and take of the classroom or constant discussion with colleagues. Peirce did not get along with most people, and in later years his ill-temper and poverty made him a lonely recluse. James decribed him as a "poor cuss" to whom no university would give a professorship, a "queer being," a "hopeless crank" and a man whose lectures were "flashes of brilliant light relieved against Cimmerian darkness!" In a touching tribute that Peirce wrote after James died, Peirce said, "Who . . . could be of a nature so different from his than I? He so concrete, so living; I a mere table of contents, so abstract, a very snarl of twine."

One of Peirce's two major obsessions was his system of "existential graphs" for diagramming logic. He was on solid ground in seeing the pedagogical value of Venn diagrams for solving syllogisms and more general problems in Boolean algebra, but he wanted to extend such visual aids to every kind of logic, including modal logic. His system grew steadily more complex, relying always on topological properties of the plane. For 20 years he used his curious diagrams as thinking aids, and there is no question that he found them enormously useful. He called them his *chef d'oeuvre,* and believed that if they were "taught to boys and girls before grammar . . . it would aid them through all their lives." If logicians would embrace his method, he wrote, "there would soon be such an advance in logic that every science would feel the benefit of it." Unfortunately no one else found the graphs useful, although it may be too early to give a final verdict. An excellent monograph by Don D. Roberts, *The Existential Graphs of Charles S. Peirce,* was published in 1972.

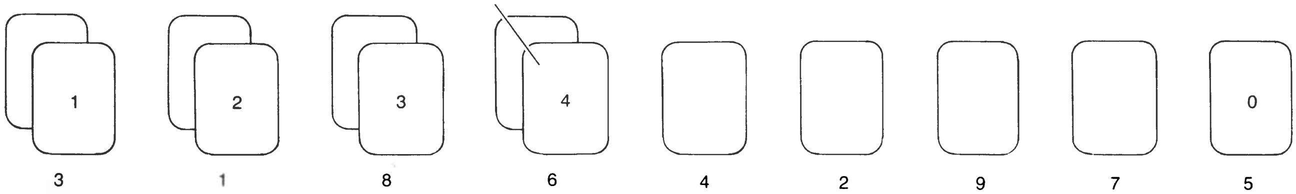
Peirce's other great eccentricity—perhaps I tread on even more dangerous ground in calling it that—was his conviction that in every branch of philosophy the most efficient way to organize concepts is by way of three fundamental categories that he called firstness, secondness and thirdness. Like scientists, philosophers are compelled to classify ideas, and since philosophy is about everything, their schemes often include a list of what they consider the most fundamental categories. Aristotle's 10 categories had such an enormous influence on Western philosophy that it was not until Kant proposed a different set that Aristotle's scheme met serious competition. Kant had 12 categories (in four triads) that he considered essential for describing how human consciousness imposes patterns on the vast, ultimately unknowable sea of being. Since Kant there have been so many different schemes that "category" has become a fuzzy and unfashionable word.

Peirce was firmly persuaded that the most useful of all philosophic tools was the ordering of things into monads, dyads and triads. Firstness considers a thing all by itself, for example redness. Not a red object, just the pure possibility of red: "Redness before anything in the universe was yet red." Secondness considers one thing in relation to another, for example a red apple. It is redness linked to an apple, a "brute fact" of the actual world. Thirdness concerns two things "mediated" by a third, for example an apple falling from a tree. The tree and the apple are linked by the relation "falling from." Our mental concept of a red apple is another thirdness because it involves apple, red and mind. The universe "out there," changing in time, and the inner world of consciousness are equally "real" realms of thirdness.

Peirce applied firstness, secondness and thirdness to every branch of philosophy. There is no need, he argued, to go on to fourthness, fifthness and so on, because in almost every case these higher relations can be reduced to combinations of firstness, secondness and thirdness. On the other hand, genuine thirdness can no more be reduced to secondness than can genuine secondness to firstness. Peirce modeled this notion with a clever bit of graph-theory sleight-of-hand. Let a point represent firstness and the end points of a line segment represent secondness. Thirdness is symbolized by the ends of three line segments meeting at a common point like the map of a forked road. Why not go on to four, five, six and so on by letting more lines join at a point? Because we can always reduce such higher "stars" to thirdness by substituting triadic graphs for the central point as is shown in the illustration on page 18. There is no way this can be done, however, to reduce a triadic graph to one with two end points.

Peirce regarded his three categories as his greatest contribution to philosophy. He denied the charge that he was infatuated by the number three. He admitted having a Hegelian "leaning for 3," but he insisted that this was because thirdness had so many applications. In his youth, he said, he would have considered his categories "crackbrained," but study had convinced him otherwise.

"The most fundamental fact about the number three," Peirce wrote, "is its



How to assemble nine piles in Peirce's card trick generative potency.... So prolific is the triad in forms that one may easily conceive that all the variety and multiplicity of the universe springs from it." He called his graph for the triad "an emblem of fertility in comparison with which the holy phallus of religion's youth is a poor stick indeed."

This is not the place to discuss the fertility of Peirce's categories, and so I content myself with saying that in Peirce's day the only eminent philosopher who shared his enthusiasm for one, two and three was Josiah Royce. James complained that he never could understand Peirce's categories. Among today's philosophers of note I know of only two enthusiasts, Eugene Freeman and Charles Hartshorne. "I believe," Hartshorne has written, "that all things, from atoms to God, are really instances of First, Second, Third, and that no other equally simple doctrine has the power and precision of this one."

Let us turn to something less controversial: a card trick. In the April, 1908, issue of *The Monist* Peirce had an article on "Some Amazing Mazes" that opened with an apt description from Milton's *Paradise Lost* (Book V, 623–624) of a "mystical dance" of angels:

> ... Mazes intricate,
> Eccentric, intervolv'd, yet regular
> Then most, when most irregular
> they seem.

"About 1860," Peirce begins, "I cooked up a *mélange* of effects of most of the elementary principles of cyclic arithmetic; and ever since, at the end of some evening's card-play, I have occasionally exhibited it in the form of a 'trick'... with the uniform result of interesting and surprising all the company, albeit their mathematical powers have ranged from a bare sufficiency for an altruistic tolerance of cards up to those of some of the mightiest mathematicians of the age, who assuredly with a little reflection could have unraveled the marvel."

By cyclic arithmetic Peirce meant what is today called congruence arithmetic. Some teachers call it "clock arithmetic" because it is so nicely modeled by a clock. For example, 2 is equal to 14 modulo 12. This means that if you divide 2 and 14 by 12 (the modulus) the remainder in each case is 2. In clock terms at 14 hours past noon the hands of the clock are in the same position as they are two hours from noon.

The first of Peirce's card tricks, reprinted in Volume 4 of his *Collected Papers* as "The First Curiosity," is surely the most complicated and fantastic card trick ever invented. I cannot recommend it for entertaining friends unless they have a passion for number theory, but for a teacher who wants to "motivate" student interest in congruence arithmetic it is superb. There is no way to prove that the trick always works without learning a great deal about "cyclic arithmetic," including a famous theorem of Fermat's about prime numbers.

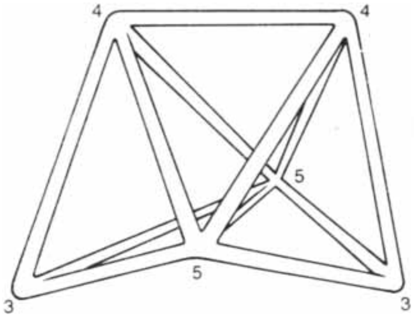
Before reading further, the reader is urged to get a deck of cards and carefully follow the procedure. Remove all the hearts and arrange them in serial order from ace to king, the ace on top of a face-down packet. Do the same with spades, except for the king, which is not used. Thus the spade packet consists of 12 face-down cards from ace on the top to queen on the bottom. Put the red packet face-down on the table. Hold the black packet face-down in one hand.

Deal the black cards face-up onto two piles. (Whenever cards are dealt into piles they are held face-down and dealt face-up, from left to right, starting on the left.) The last card (the queen) is discarded by placing it face-up to one side to form a discard pile. Substitute for it the top card (the ace) of the red packet, putting it face-up on the second pile in place of the discarded queen. Assemble the two piles by picking up the pile farthest to the left and dropping it face-up on the second pile. Turn the black packet (now containing one red card) face-down and repeat exactly the same procedure. This time the red deuce replaces the last black card (the jack). The jack goes face-up on the previously discarded black queen. The procedure is continued until it has been performed 12 times in all. You may be surprised to discover that you now hold an all-red packet, and that the discard pile contains all the black cards. Pick up the remaining king of hearts and add it to the bottom of the face-down red pile.

To make sure you have done all this properly, check the red packet. Held face-down, and reading from the top, the order of cards should be: 7, 8, *J*, 9, 4, *Q*, 6, 10, 3, 5, 2, *A*, *K*. The black packet should be: *Q*, *J*, 9, 5, 10, 7, *A*, 2, 4, 8, 3, 6.

The two packets are correlated in a curious manner. The value of the card at the *n*th position from the top of either packet gives the position from the top of the other packet of a card with the value *n*. For example, where is the jack of spades? Counting the jack as 11, look at the 11th card in the red packet. It is a 2. Check the second card in the black packet. It is the jack of spades. Where is the five of hearts? The fifth card in the black packet tells you. It is a 10. The 10th card in the red packet is the five of hearts.

Before you reveal this remarkable correlation to your audience, however, the red packet is apparently randomized by the following procedure. First allow the packet to be cut as often as anyone wants. Hold it face-down and ask someone to name a number from 1 through 12. Call the number *k*. Deal the cards face-up into *k* piles, then assemble them by starting with any pile designated. The assembled packet can then be cut again



The "other" eight-sided deltahedron

and the procedure repeated as often as you like, either with the same *k* or a different one. One would suppose that cutting, dealing into *k* piles, assembling and cutting, and repeating this routine many times with any *k* requested, would hopelessly mix the red cards. Astonishingly, thanks to the theorems of congruence arithmetic, the correlation of the two packets is conserved!

The only difficult part of the mixing procedure is this. When the *k* piles are assembled, you must do it in a precise manner that depends on the value of *k*. Think of the row of piles as being circular, the last pile adjacent to the first, so that you can count "around" the row in a direction either clockwise or counterclockwise. Note the pile on which the last card was dealt. In your mind call the end pile on the right zero and count to the pile that got the last card. Count clockwise or counterclockwise, whichever is shorter.

Suppose you have dealt the red cards into five face-up piles. The middle pile will get the last card, as is shown in the top illustration on the preceding page. It is second from the right, counting counterclockwise from the right end. This means that you must assemble the cards as follows. Pick up any pile and place it face-up on the second pile to the left, counting counterclockwise. Pick up the enlarged pile and place it on the second pile leftward, and continue until there is a single packet. The numbers under the piles in the illustration show the order of assembly if you start with the first pile on the left.

It is important to remember that in gathering the piles you count positions, not the actual piles. For some *k* the pile getting the last card is adjacent to the rightmost pile. This makes the assembly simple because the piles go on adjacent piles. But if there is a wider separation (as in the case of nine piles where the shortest distance is 4) it takes a bit of experience to assemble the piles rapidly. The bottom illustration on the preceding page shows the order of picking up nine piles if you start with the second pile from the left. In this case you proceed clockwise because the shorter count from the right end to the pile that

There's hardly an audio enthusiast alive who doesn't admire the Nakamichi 1000II.

But at $1,650,* admiring it is about all most people can do.

That's why Pioneer created the new CT-F1000. A cassette deck that offers all the features and performance of the Nakamichi 1000II, but costs almost $1,000 less.

(We realize this is hard to believe, but be patient. The facts bear us out.)

It's a fact that the $600* Pioneer CT-F1000 and the $1,650 Nakamichi 1000II are both honest three headed cassette decks that let you monitor right off the tape as you record.

Both feature separate Dolby systems for the playback and recording heads. So when you're recording with the Dolby on, you can monitor the same way.

And both are filled with all the remarkable features you'd expect to find on cassette decks of this caliber: there's everything from jam-proof solenoid logic controls, to multiplex filters for making cleaner FM recordings, to memories that

It's value.

automatically let you go back to a particular spot on the tape.

The comparison holds up equally well when it comes to performance.

The CT-F1000 and the Nakamichi 1000II both have total harmonic distortion levels of less than 1.5%.

Both have all but conquered the problem of wow and flutter. (An identical 0.05% for each deck.)

And both have signal to noise ratios that are so similar only sophisticated laboratory equipment can tell them apart.

If the incredible value of the CT-F1000 still sounds a bit hard to believe, we suggest you go hear it for yourself at any Pioneer dealer.

Our viewpoint is simple: if you can't hear the difference, why pay the difference?

PIONEER®
We bring it back alive.

got the last card is a clockwise count. As you are practicing you can mark the positions with a row of pennies. After a while you can dispense with the markers.

The assembled packet can always be cut as many times as it is desired and the mixing repeated with a different number of piles. When everyone is satisfied that the red pile has been thoroughly "shuffled," it is necessary to give both the red and black piles a single cut. Cut the red pile to bring the king to the bottom when the pile is face-down. Peirce suggests remarking that since there is no king of spades, you will cut the red pile to bring the king to the bottom "and so render any searching for that card needless." As you do this, note the packet's top card. Suppose it is a 4. The black pile must now be cut so that its ace is fourth from the top. The two piles will then be correlated as before!

To dramatize the correlation Peirce suggests dealing the black cards in a face-down row. Ask anyone to name a red card. Suppose he says the 7 of hearts. Tap the black cards one at a time, counting from 1 to 7, and turn the seventh card face-up. Note its value n. Count to the nth card in the black packet. It will be the 7 of hearts.

"The company never fail to desire to see the thing done again," writes Peirce. After revealing a number of cards to be at the indicated positions, you can repeat the trick by mixing the red cards again with several deals, then adjust the two packets by the required cuts. "If you wish for an explanation of it," Peirce concludes, "the wish shows that you are not thoroughly grounded in cyclic arithmetic." He refers his readers to a book by Richard Dedekind, but adds that on another occasion he may write a little essay on the topic.

This "half promise" he "half redeems," as he puts it, in an article in the July, 1908, issue of *The Monist*, also reprinted in the *Collected Papers*. Peirce's explanation of the trick runs to 58 pages! The essay is pure Peirce, complete with generalizations and formulas, horrendous existential graphs that look like abstract art and delightful digressions on such things as the value of keeping notes on file cards, logic machines, the meaning of continuity, how mind influences matter and the nature of free will and time.

In 1958 Alex Elmsley, a London magician, pointed out in *Ibidem* (a Canadian magic periodical) that in the first phase of Peirce's trick it is not essential that the last black card each time be the one that is replaced with a red card. The card to be replaced can be at any position in the packet. Thus you may allow someone to choose any number n, from 1 through 12, then on each deal you replace the nth card with a red one.

Peirce's writings are now gathered into 13 volumes, six edited by Hartshorne and Paul Weiss, two by Arthur W. Burks and five by Carolyn Eisele. It is a triadic scandal (1) that the bulk of Peirce's mathematical papers were not published until two years ago, when Mrs. Eisele skillfully put them together for *The New Elements of Mathematics*, (2) that these books have received almost no advertising or reviews and (3) that the set costs more than $300.

The illustration on page 24 shows the answer to the first of last month's problems: an eight-sided deltahedron (all faces equilateral triangles) that is not a regular octahedron. The regular octahedron has four edges meeting at each corner. On this solid two corners are meeting spots for three edges, two for four edges and two for five edges.

The second problem was to use Euler's formula, $F + C - E = 2$, to show that no sphere can be covered with a "regular map" of hexagons, each vertex the meeting point of three edges. Assume such a map exists. Each hexagon has six edges and six corners. Therefore if the hexagons did not share corners and edges, there would be six times as many edges as faces. Each corner is shared, however, by three faces; therefore the number of corners in such a map must be $6F/3$. Similarly, each edge is shared by two faces; therefore the number of corners in such a map must be $6F/2$. Substituting these values in Euler's formula gives the equation $F + 6F/3 - 6F/2 = 2$, which simplifies to $F + 2F - 3F = 2$, or $0 = 2$. This contradiction proves the original assumption to be false.

What happens when the above argument is applied to the regular maps formed by the edges of the five Platonic solids? In each case we get a formula that gives F a unique value: 4, 8 and 20 for the tetrahedron, octahedron and icosahedron respectively, 6 for the cube and 12 for the dodecahedron. Since a regular polyhedron cannot have faces with more than six edges, we have proved that no more than five regular solids can exist.

Euler's formula also underlies an elementary proof that there are exactly eight convex deltahedra. See "Deltahedra" in *Excursions into Mathematics*, by Anatole Beck, Michael Bleicher and Donald Crowe (Worth Publishers, 1969), pages 21–26.

In April's column, one of the captions was not quite accurate. The landscape imagined by Benoît Mandelbrot and programmed by Richard F. Voss should have been called a *modified* Brownian surface. I also neglected to say that the magazine's striking cover, showing Mandelbrot's Peano-snowflake curve, was drawn with a computer-graphics program written by Sigmund Handelman and Mark Laff, both members of the staff at the Thomas J. Watson Research Center of the International Business Machines Corporation.

BOOKS

Limestone landscapes, herbals, malaria and the fuel economy of gasoline engines

by Philip Morrison

Morphogenetics of Karst Regions: Variants of Karst Evolution, by László Jakucs. Translated from the Hungarian by B. Balkay. John Wiley & Sons ($40). Limestone topography of every form, from "rounded slopes with their surface fragrance of thyme" to the incredible but quite real rocky pinnacles dear for 1,000 years to the landscape painters of China, has naturally attracted students of the origins of land forms. The great caverns and the underground rivers, the weird dripstones and powerful deep springs all add their subterranean wonders. The prototype of limestone countryside, the region of Kars along the eastern shores of the Adriatic, has given its name to this diverse land-form complex: karst.

Technical but hard to put down, this treatise by a distinguished Hungarian scientist, well translated and well illustrated, approaches the problem with a lifetime's appreciation of its dynamical essence. The central theme is the geologically swift corrosion of soluble rock, mainly of limestones by water rendered acid by dissolved carbon dioxide. But the music of these remarkable landscapes is much richer than that simple motif. Consider gypsum: it is 200 times more soluble than limestone, but the corrosion of it is self-limiting. The pressure of the overburden keeps the low-density form (the anhydrite without water of crystallization) from swelling when it is hydrated. If any fissures form, below a few meters in depth all but the largest ones therefore close up, and so gypsum karsts remain superficial.

That is nothing new, of course, but it will stand as a sign of the subtleties of the world of karst. For Jakucs no mere mechanical definition will do at all. He maintains convincingly that karst is rather the "state of evolution... of a mountain-sized mass of limestone," without excluding the rare examples in rock salt, loess and even granite.

The issue is not verbal. The complex problem has often led its students to paradoxes of oversimplification. In the 1950's the literature centered on the argument that the concentration of indispensable acidifying carbon dioxide in rain and snow and the measured content of dissolved carbonate in river waters were an order of magnitude greater in the cold Arctic than in the subtropics; the Tenana River of Alaska was carefully compared (nearly 10,000 measurements were taken) with Florida's Kissimmee. Tropical and subtropical karsts, however, are patently faster to develop, again by perhaps an order of magnitude, than polar ones.

The reconciliation is now in. The polar world is low in living plants. There most of the carbon dioxide is supplied by the atmosphere; the process of reaction is slow and the runoff takes the solute with it. In the verdant Tropics none of these circumstances holds. The biogenic carbon dioxide of the living soil dominates the supply; the corrosion takes place at the grass roots. Evaporation controls runoff, and the rivers poorly reflect what is going on. The transport of the carbonate is not in a slow horizontal flow but in swift, dramatic vertical lines. The land forms reflect that fact: "natural karst corrosion... is simply the formal imprint upon the soluble parent rock of the phenomena of biological and chemical evolution of the soil." Detailed microstudies of the grass-root carbon dioxide contents support this assertion, and they are translated with other data at least tentatively into a quantitative account of the individual contributions to karst corrosion in five climate zones. The main effect worldwide is the result of biogenic carbon dioxide; the other organic acids of tropical soils are a good second.

The second large topic is the role of more ordinary processes, stream erosion and scouring rather than solution, in fixing the forms of karst. The point is that the karst-forming waters often are controlled by nonlimestone surfaces. A limestone body whose surface is covered by impermeable deposits will develop a normal stream valley until the cover is worn through. The solution processes are then slow, since the attack of the waters is only linear, not areawide as in the purer case. Many deep valleys in karst regions arise from this kind of process, and not usually from the collapse of cavern roofs. This is only the simplest of the cases ingeniously considered. Here the book offers much field detail, particularly from eastern Europe and Cuba, where the author has done most of his fieldwork. (He has been a noted speleologist and enthusiastic cave director since his youth.)

Readers will probably seek a more static account of the land forms in any region of their own interest; Professor Jakucs does not offer any overall synopsis of specific karsts, preferring to deploy his carefully studied examples in the service of the dynamical theory. There is, however, a rich list of references, and two one-page world maps make it clear where your own limestone country can be found. It is easy to idealize the human relationship with limestone; the romance is clear to the cavers, and the richness is manifest in the bluegrass pastures on the downs. But what of the "dwellers inhabiting the rim of the Fatničko Polje in... Yugoslavia, who can never be sure that their crops will ripen or be inundated before the harvest?" The photograph shows a donkey and three dinghies beside the pond; if certain natural underground valves should unhappily close, the farmers must take at once to their boats, to harvest even half-ripe crops. Widen the outlets and swallow holes; never try to close them!

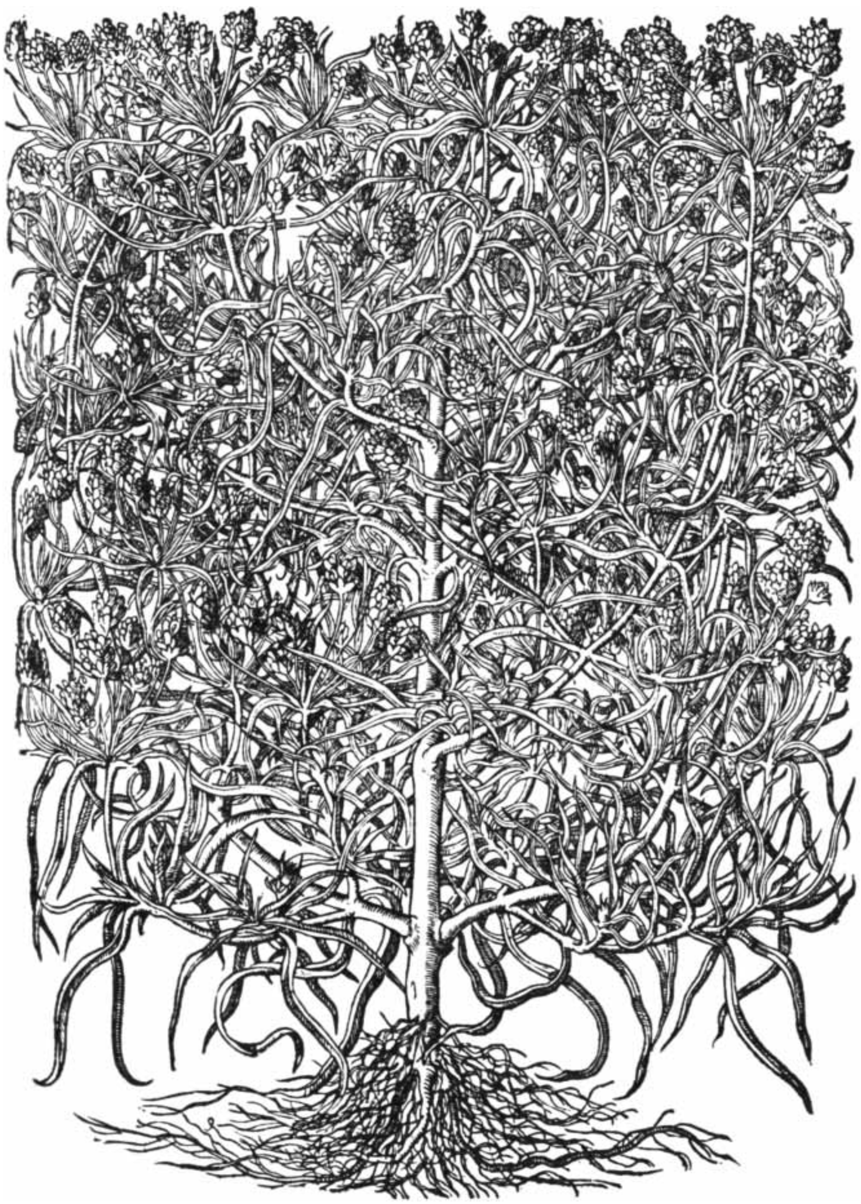
An Illustrated History of the Herbals, by Frank J. Anderson. Columbia University Press ($16.95). Herbs and simples are everywhere known to the wise. Once the written word offered the opportunity to codify, dilate on and spread such valued lore, books on the pharmacy of nature, mostly of growing plants, became staples of the erudite healer. Even in the most ancient scripts—hieroglyph, cuneiform or ideograph—works of the class are well known. The early printed books of Europe naturally included such manuals, illustrated as practical visual aids for untrained gatherers of the powerful plants as well as for the fur-robed physicians and their wealthy patrons. These volumes are the herbals, the topic of this modest book. It is itself a well-illustrated and knowing guide, not to the plants but to the old books of the plants.

More than 30 famous books are described in historical order. Their contents, which come more to reflect the natural world and less the haphazard content of tradition, are characterized. A hundred illustrations sample their figures of plants and animals, of worthy savants and physicians in grave consultation, and the grand allegorical title pages. The first book discussed is rightly the *Materia medica* of the Greek physician Dioscorides, the classical model for most of the genre by way of a swarm of manuscripts in half a dozen tongues. Once the Renaissance scholars realized that many plants praised in that work

did not always reward a long and careful search of French or Flemish fields, the science of botany in a real sense had begun.

One of the latest herbals included is the *Touchstone of Plants* (*Phytobasanos*) by Fabio Collonna, a man who with Galileo was a charter member of the Academy of the Lynxes. His book, printed in 1592 with his own etched designs inside woodcut borders, emphasized botanical information, although he still sought to offer a key to the classical curative herbs. His sharp eye was among the first to recognize that flower parts and seeds reveal the kinship of plants better than leaf shape, "especially if they agree by the taste with the other parts of the plant."

The herbals not only signal the rise of a deeper botany; they also exemplify the growth of printing, of text and image and of the entire bookish world of Western civilization. Anderson is himself a learned botanical bibliophile and translator, and to each brief chapter he has added an interesting set of bibliographic notes that together span the history of the printed book. The encyclopedic if derivative *Buch der Natur* was a runaway best-seller of the 15th century even in manuscript; the Augsburg printers brought out six distinct editions before 1500, illustrated by the best woodblock designers. The first woodcut to show plants for themselves, not as mere ornament or in a landscape, was in that book, and the first cut of animals as well. The marine biology of its 14th-century author is less than secure: we see "the lobsterlike ceruleo, with claws sixty cubits in length," able to pull elephants under water. Yet Conrad von Megenberg also reports his own experiences, true wonders: a white rainbow, a meteor and the ominous use of firearms near Ratisbon in about 1350. By the time of the great Brunfels herbal in 1530 the master of its splendid illustrations, Hans von Weiditz, pictured the plants as he saw them, sometimes with inset drawings of details, sometimes with a broken stem or damaged leaf, just as the specimen was.



PSYLLIUM PLANT appears in a woodcut from a 16th-century herbal by Pier Andrea Mattioli. Woodcut is reproduced in *An Illustrated History of the Herbals,* by Frank J. Anderson.

Most of the figures sampled here are about half scale. The reader lucky enough to have access to a rare-book library will seek out the full-sized herbals. One might find, for example, the first representation of maize in the big *De Historia Stirpium,* by Leonhart Fuchs. Some specialists prefer the octavo versions of Fuchs, portable field guides from the 16th century, judging that in full folio the line of the 500 cuts is too light and thin. Anderson marks out a delightful garden path into this thicket of erudition. One misses bibiographic references to some fine modern facsimiles that are alluded to a few times. Perhaps they are still mainly the dream of the bibliophile. (In November, 1976, a volume of facsimile woodcuts from Pier Andrea Mattioli's sumptuous herbal of the 1550's was reviewed in these pages.)

RADIATION EXPOSURE FROM CONSUMER PRODUCTS AND MISCELLANEOUS SOURCES, NCRP Report No. 56 ($4). NATURAL BACKGROUND RADIATION IN THE UNITED STATES, NCRP Report No. 45 ($5). National Council on Radiation Protection and Measurements, P.O. Box 30175, Washington, D.C. 20014. Ionizing radiation ye shall have with you always. The issue, carcinogenic if not burning, is how much radiation. These two reports, each the product of an identified committee of experts, aim to provide an essential basis for the evaluation of such concerns. The publisher, the National Council on Radiation Protection and Measurements, is a not-for-profit organizing body chartered by Congress in 1964 and charged broadly to keep an appraising eye on national radiation exposure. Some 30 NCRP committees are at work on specific questions from microwaves to mammography, and the yield has been two or three reports per year. The reports are not themselves research papers; they are critical summaries of the research literature, seeking to appraise current results and their uncertainties, and they are written for readers who have scientific background but are not specialists in health physics. They offer plenty of tables and graphs, often a glossary of terms and, when the title so indicates, descriptions of current techniques of measurement.

These two reports together sketch a picture of the baseline radiation dos-

"Control Data PLATO® increased

"PLATO computer-based education will save Merrell-National over 12,000 classroom hours on just one drug product training program alone."

Gary J. Wilson, Sales Training Manager
Merrell-National Laboratories
Division of Richardson-Merrell Inc.

"Today's modern ethical drugs demand sales people who are better informed, more technically oriented and better able to relate vital product information to help doctors do a better job. To meet the challenge, Merrell-National Laboratories has initiated a program to improve the quality of pharmaceutical representative training.

"At the heart of this program is the Control Data PLATO system, a truly versatile and cost-effective approach to computer-based

For further information, write on your business letterhead to Control Data Education Company, HQN11I, P.O. Box 0, Minneapolis, MN 55440. Or call 612/853-7600.

our training efficiency 167%."

education. Unlike programmed instruction that simply feeds out information, PLATO individualized instruction creates a constant dialogue with the learner...increases retention, compresses training time.

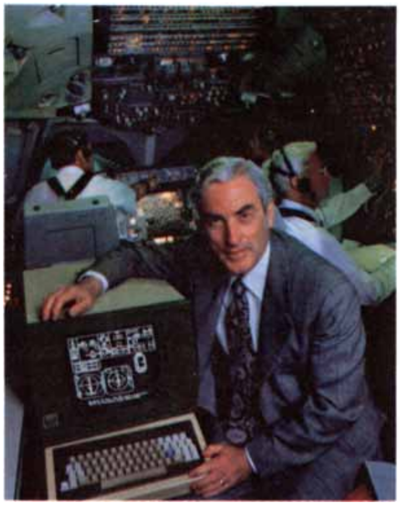
"For the initial phase of our program, we selected an anti-depressant that is a mainstay of the Merrell line. With the PLATO system, we increased training efficiency over 167%...reducing a 4 day session for new representatives to a 1½ day course with increased comprehension. That's a savings of 20 classroom hours per student. When we retrain all of our representatives, we estimate we will save over 12,000 meeting room hours.

"The PLATO system allows us to decentralize our training. Control Data maintains PLATO-equipped Learning Centers throughout the U.S. These centers are specifically designed for individualized instruction. Our people work in the privacy of study carrels at a pace that is most comfortable for them. Not only are we

producing better trained personnel this way, but we're saving the cost of lodging, food and the miscellaneous expenses of taking people away from their jobs for extended in-plant training."

Control Data PLATO is solving the tough training problems in many other industries

By simulating cockpit systems on a PLATO terminal, American Airlines, for a portion of crew training, has cut a full day off its training schedule...and is saving thousands of dollars on crew training.



Dr. Robert Houston, Director of Technical Training Support, American Airlines.

It costs about $4,000 an hour to use an actual airplane for pilot training. A flight simulator brings the cost down to $400 an hour. The cost of training with the PLATO system is less than $20 an hour...and the added convenience and quality comes as a bonus.

Some other business examples:

- **Electric Utilities** are using PLATO simulations for plant personnel training...to guard against human failure as well as mechanical failure.
- **A Heavy Manufacturer** is using the PLATO system to upgrade the product knowledge level and marketing techniques of its entire distributor organization.
- **Banks** are using PLATO-equipped Learning Centers to train tellers in banking procedures.
- **Detroit Auto Makers** are using PLATO's graphic design and structural analysis capabilities to build more space into smaller cars.

The PLATO system is more than an education and training tool. It is a means of managing information and simulating real-life situations. You don't identify PLATO advantages by what it has done for others. You do it by examining what the PLATO system can do to solve your own unique training problems.

And we'd like to help you do just that, with an interview or private demonstration. To get started, write on your business letterhead to Control Data Education Co., HQN11I, P.O. Box 0, Minneapolis, MN 55440. Or call 612/853-7600.

CONTROL DATA CORPORATION

More than a computer company

age expected today for the U.S. population, leaving out of consideration the effluents and wastes of nuclear power plants or the terrible outcome of nuclear weapons explosions on target. (The earlier report, published at the end of 1975, includes the status of weapons fallout from atmospheric weapons testing, since 1963 reduced in rate by a factor of 20 or 30 under the partial test ban.) The later and smaller report adds a study of a wide variety of man-made sources, some of them expected and some surprising.

The natural-background study includes estimates not very different in scope or result from the earlier work of a United Nations expert committee that was reviewed here in 1973. Once again a reader is impressed by the diversity of the arguments required for analysis. It goes without saying that the worker is dealing with a large population of complex organisms exposed externally and internally, tissue by tissue, to a variety of chemical elements with distinct nuclear properties, but that is only the beginning. Indoor living in part shields the dose delivered from rocks and the soil; the study here takes into account the numbers of people at work, in school and at home and estimates that the mean whole-body dose is cut by 20 percent. Frame houses are low in radiation from gamma-ray sources, whereas masonry structures are higher; new buildings, even multistory ones, are built of chemically processed materials, such as plastic, steel and aluminum, from which natural radioactive mineral grains have been preferentially removed.

A survey has shown that the average American spends about 5 percent of his time outdoors at work and play. Outdoors the eternal sources abide: cosmic rays from the sky above, mainly from the galaxy, and natural radioactive elements from the earth below, both thorium and potassium outweighing uranium in dose. Heaven and earth contribute about alike; at higher altitudes in the rockier mountainous regions both increase, twofold or so in the highest towns. Conversely, the alluvial coastal lowlands show half the dose. Granite masonry, gypsum board and unburned radon in the natural gas that fuels our ranges are major man-placed contributors, although they add little either to the external dose or to the expected intake from the natural radon leak out of the soil.

Only travel through the stratosphere or space imposes intimacy with the cosmic rays, but intimacy with minerals is less rare. The dwellers on the active black sands of the shore of Kerala in India, rich in heavy elements, have no significant American counterpart. Everyone knows the gloomy story of the people in uranium country who built with concrete blocks incorporating uranium-mine tailings. Quite unexpected is the result of a different mineral intimacy: false teeth. The porcelains in about half of these prosthetics contain feldspars with a high potassium content, a contribution amounting to 100 times that of the teeth they replace. A full set of such dentures—there may be 19 million chewing today—might double the total natural dose to the wearer, at least in the mouth. For 50 years a trace of uranium salt has been added as well to achieve the natural look; nothing has been found to replace its contribution to tooth color and fluorescence. The uranium content is rigidly controlled, and manufacturers have been using low-activity (and cheaper) depleted uranium for a decade. This effect would appear largest of all, a 500-fold increase over the background dose, but the alpha particles that account for it can penetrate only a paper-thin layer, and probably do not reach significant living tissue. Pessimists might enjoy learning that some rose-tinted spectacles (thorium) irradiate the eye.

Tobacco smoke brings in natural polonium through a long chain of events beginning with the structure of the leaf. Whether localized spots of deposition at the bronchial forks are really the source of the smoker's risk of tumor is not known; the calculated dose is 100 times the natural background, if the spots are as bad as they can be.

X-ray leakage from electron microscopes, gamma rays from radium-dial pocket watches (internationally discouraged) and X rays from unshielded high-voltage vacuum switches in power substations are typical sources of significant doses to special groups of people. "Select groups" of habitual air passengers may be exposed to as much as double the common lot as they ride with the million-a-year lightly wrapped air shipments of radiopharmaceuticals. A long list of other products and practices is examined and judged to be of little risk to the public at large.

There comes through all this detail a reassuring sense of thoughtful consideration. The work is evidently worth the real effort. We shall await the important reports in progress. Once again the scrupulous concerns of peace seem merely fastidious in the face of the stored megatons of gross thermonuclear war.

MOSQUITOES, MALARIA AND MAN: A HISTORY OF THE HOSTILITIES SINCE 1880, by Gordon Harrison. E. P. Dutton ($15.95). "The remarkable thing about them is undoubtedly the pigment. I have never seen this before in any mosquito (I have now examined hundreds—or a thousand)." Beneath the oil-immersion lens was the stomach dissected out of a big brown mosquito with dappled wings, the last insect of 10 gorged on the blood of the malaria patient Husein Khan. In the dark, hot little office of the British army hospital in Begumpett 80 years ago Surgeon Major Ronald Ross of the Indian Medical Service peered down the tube. Ross was a gifted, arrogant man of 40, half dilettante and half tireless devotee. He was convinced of eventual fame, whether through the new algebra he felt he had discovered or through his romantic verse and the novels he dreamed of writing or even through the bacteriology he had studied for a couple of months at St. Bartholomew's. The mosquito with the dappled wings did it. Ross wrote then in stiff, elevated lines: "This day relenting God/ Hath placed within my hand/ A wondrous thing.... Seeking His secret deeds/ With tears and toiling breath,/ I find thy cunning seeds,/ O million-murdering Death."

Ross's pigment was the black spoor of the one-celled parasites that dwell for a while within the red blood cells of their host, eating out the rich oxidizing protein; it had for 40 years been recognized as a degraded hemoglobin found copiously in the white cells of malarial blood. His oracular verse did not much exaggerate. Within a few years the complex life cycle of the protozoon infection, so long with us that the relationship antedates our species itself, was mainly unraveled.

More than a third of this chronicle, critical and close to the sources and at the same time perceptively and excitingly written, tells of the puzzling out of malaria. The hero is Ross, isolated at his work and yet not quite alone. During the five years of his main quest he was linked to the world of science by one monthly journal and also by a steady exchange of letters with his celebrated Harley Street sponsor and mentor, Patrick Manson. It fell to a powerful Italian group (Ross and Giovanni Grassi were to engage in a lifelong polemic over credit) to complete the structure Ross had fully laid out in India: final proof that mosquitoes of the genus *Anopheles* were the vectors for human malaria and that the *Plasmodium* parasite followed the very path Ross had inferred for it, although he could demonstrate it fully only in the related disease of birds.

The burden of the book, however, is not this indispensable intelligence coup, but rather a critique of the subsequent long campaign waged in the field. The theaters included such marshy worlds as the Roman countryside, classically held after nightfall by the enemy, and also those distant lands where Europeans ruled over tropical empires. Ross himself attacked *Anopheles* in Freetown in Africa but failed. The town was too big, too flat, too filthy, too poor. Only a cool hill station for the Civil Service could be kept secure for the settlers, above both the protozoa and the Africans. There were other tropical defeats not unlike

LEICA
R3

the loss at Freetown. A brilliant victory finally was won by William Gorgas, a U.S. Army doctor, in the Canal Zone. Draining, poisoning, screening, swatting, Gorgas' little army made the zone heathful in 10 years, but the price was high in dollars per acre and in the military control required over human behavior. In Italy too, under Fascism, heavy investment in rural drainage and resettlement improved the Pontine Marshes, and the ancient disease ebbed.

General development in rural Europe and America had a lasting effect. Small changes in all the parameters of spread make a vast difference to malaria. It is an improbable disease with exquisite adaptations, and details of agricultural landscape and human habits can tip the balance, even if often no one can say just how. This is true above all along the malarial margins in the temperate world; in the long, wet, warm seasons of their tropical homelands, the swift generations of well-adapted vector and parasite survive little changes to burst out in epidemic proportions whenever the defense slackens or weather favors the delicate larvae. So it was in Ceylon in 1934. After years of weather unfavorable for breeding, both monsoons failed. The swift streams shrank to a myriad of puddles. Before the rains came again in late 1935 a third of the people had fallen ill and 80,000 were dead. The mathematical modelers could explain it, but that was little comfort.

The disease made a desperate sortie to Brazil in the late 1930's. Government and the Rockefeller Foundation (then a worldwide staff for this war) acted well and fast. Pyrethrum sprays were brought against the invading forces, an African species that had been carried by ship to the shoulder of Brazil. By 1941 the invaders were done for; the body count gives the very date of the last single larva put to death. A good army had won, but it won against a foreign foe: a long-domesticated strain constrained to dwell near man, not a homeland guerrilla able to vanish into the wild hiding places among the birds and the beasts of the jungle. Endemic malaria remains.

The ultimate weapon, the cheap, long-lasting and toxic DDT, offered the dream of a final solution (from which sub-Saharan Africa was from the start excepted). From Naples in 1944 to India and Sri Lanka by the late 1960's, malaria retreated before the attack of these big battalions wherever DDT was tried. Jeeps and shoulder sprays were the arsenal; two grams of DDT per square meter of wall in every hut, every year, was the standing order. The soldiers were numerous, and they required but did not always have persistence, competent officers, inspectors and alert patrols to locate covert refuges. Eradication was the battle cry. This bold, oversimple tactic has failed at many points, not least in the growing tolerance of the cunning enemy to the new poison. Malaria is weakened today but indomitable, resurgent particularly in Asia "where the battle had seemed almost won." In India there were perhaps 75 million cases in 1950, about one in 100 of them fatal. The number fell by a factor of a full 1,000 in one decade of hard campaigning, but by 1976 the number of cases climbed back toward 10 million. The World Health Organization changed the overconfident name of its Malaria Eradication Division in 1973; assault gave way to control. That was the old wisdom and is again the new: like our adversaries, we must bend and adapt. But we alone can understand.

This war will not end in unconditional surrender. A thousand million generations of selection make that unlikely. Rather, *Homo sapiens* needs to protect itself against *Plasmodium* for a long time to come, year after year pitting its minds and organizations, with a gene or so every 10,000 years, against the labile codes of the "million murdering" foe.

FUEL ECONOMY OF THE GASOLINE ENGINE, edited by D. R. Blackmore and A. Thomas. John Wiley & Sons ($22.50). The three-wheeled special vehicle that won the Thornton mileage marathon in 1976 set a pretty unlikely standard of economy: 1,141—yes, one thousand one hundred and forty-one—miles to the gallon. (The gallon was the sturdy British one, to be sure, and a U.S. gallon would have fallen short of a 1,000-mile trip by 50 miles.)

How can a gallon be stretched so far? That is the topic, in full engineering detail, of this unusual book. Its authors are a dozen specialists from Shell's Thornton Research Center in Chester, England. The chapters discuss chiefly the usual gasoline engines, with attention to emissions, mixtures, additives and engine-cycle parameters; they follow the energy to its point of application, discussing adjustments of timing and ignition and the losses from lubricants in engine, transmission and axle and from the rolling tire.

That mileage marathon has been held at the Thornton center for a decade. (There is an even older marathon at another Shell laboratory in the U.S.) Slightly modified small production cars yield about 120 miles per U.S. gallon at an average speed of 30 miles per hour; lower mean speeds in the power-burst, engine-off coasting cycle allow semistock cars to do three times as well. Estimates of the possibilities for increasing mileage suggest that a 50 percent improvement is practical even for compact European cars, mostly from engine design, gasoline changes, improvements in losses and accessories. The formula for required power makes it plain that speed exacts a cost (quadratically in the air-drag term), and that weight costs too, linearly for rolling resistance and acceleration. Engine efficiency rises with the compression ratio, but then the energy consumption shifts to the refinery, which needs more crude oil and more capital to make higher-octane fuel. The trade-off is complex, since the motorist who pays taxes on the gallon will always want higher mileage. High compression, high efficiency and lower octane number is the only way to please both parties, and design is moving in that direction. The trick is to meet emission requirements as well.

The complex problem comes down to a control problem, and here the Diesel engine enters as a top competitor. At wide-open throttle an engine is quite efficient. To reduce power at a given speed the Diesel controls simply reduce the fuel flow, changing the air-to-fuel ratio; they do not throttle the airflow. In the ordinary gasoline engine the controls keep a more or less fixed air-to-fuel ratio but throttle the overall airflow. The newer versions, the "lean burn" gasoline engines, are intermediate; they promise a good solution overall (a match for the Diesel, according to these engineers). The future of all this may turn out to be in microelectronic chips that make possible the electronic adjustment of emission and mixture controls. Cold starts, short trips, too thick cold oil—therein lie the many practical engineering possibilities. The subject is discussed in more detail elsewhere in this issue of *Scientific American* [see "Alternative Automobile Engines," by David Gordon Wilson, page 39].

How do you measure mileage? As one might guess, the most relevant ways (on the road at uncontrolled speeds) are the least reproducible. A variety of trade-offs exists, with a practical competition between controlled-cycle test tracks and a car set rolling in the laboratory on a high-quality chassis dynamometer, which must simulate airflow, drag and normal inertial effects. A good road-track test is repeatable with a standard deviation of only 2 percent, but the laboratory test is free from the effects of local weather.

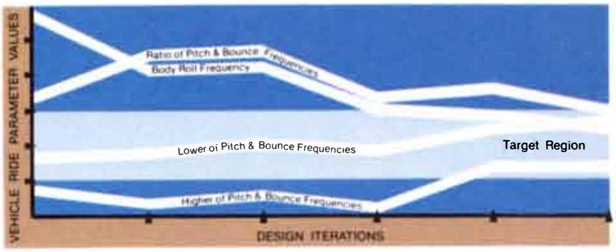
The "preposterous petrol parer" was a sports bicycle with an outrigger wheel, ball-bearing hubs and high-pressure racing tires. Its engine was a 49-cubic-centimeter Honda air-cooled, four-stroke single cylinder with a handmade special carburetor, with no throttle, with the magneto removed and with the engine insulated to stay hot! There were about 10 power bursts per lap, each lasting some five seconds and followed by a minute or so of coasting. The average speed was 10 miles an hour, and the vehicle weighed 111 pounds (without the driver). It would be an unsettling way to commute, but the economy is certainly outstanding.

Before computers, engineering designs were evaluated the hard way. They were built, tested, modified, retested, remodified — and so on.

Computers improved the process dramatically, enabling much of the evaluation to be done with mathematical models rather than costly prototypes. Great for the evaluator.

But that didn't help the designer. He still had the root problem of juggling parameter values to find the right combination. And because of the endless possibilities, he could only hope that his final, intuitive design was close to optimal.

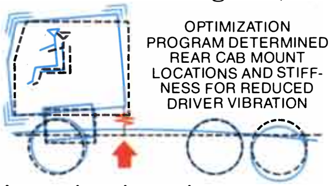
OPTIMIZATION PROGRAM AUTOMATICALLY FUNNELS DESIGN INTO TARGET REGION



Now he can do more than hope. Engineering mechanics specialists here at the General Motors Research Laboratories have devised computer concepts to aid the designer. Through these concepts, depending on how tight his constraints are and how well he defines the problem, a designer can:

- Automatically converge on the solution, if one exists.
- Or finding no solution, determine the trade-offs in relaxing the constraints.

The concepts, built on the theories of mathematical optimization and artificial intelligence, have been embodied in a new computer program that can be applied to almost any design problem.



OPTIMIZATION PROGRAM DETERMINED REAR CAB MOUNT LOCATIONS AND STIFFNESS FOR REDUCED DRIVER VIBRATION

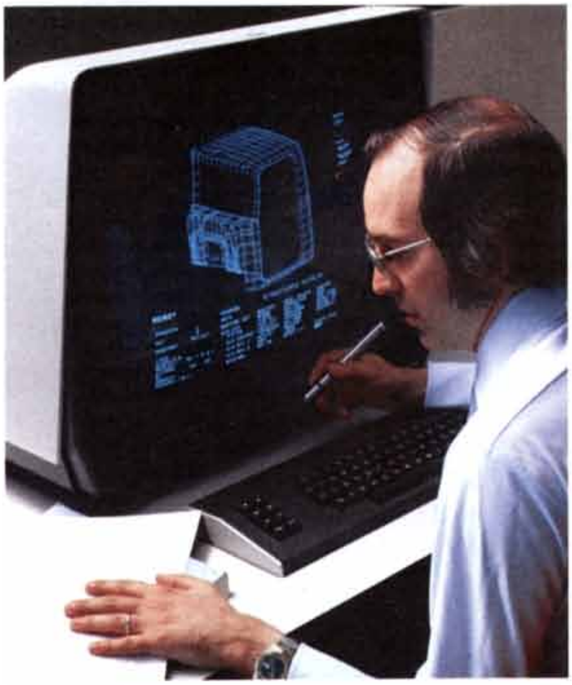
For example, we're using it to improve truck ride, increase bearing capacity, and lighten vehicle body and chassis structures.

Each application requires a computerized mathematical model of the system and fixed design goals. The "optimizer" does the rest. It shows how to alter the design to achieve those goals, cutting what could be weeks of iterative analysis down to a day or two.

We're automating the design process . . . minimizing the art, maximizing the science . . . to meet the demands of an increasingly complex world.

If you have a Ph.D. in engineering or the physical, mathematical or biomedical sciences, we invite you to check a number of current openings at the General Motors Research Laboratories. Please write GMR Personnel, Dept. 710. An Equal Opportunity Employer.

Who says designing can't be a science?



General Motors Research Laboratories
Warren, Michigan 48090

Established 1845 SCIENTIFIC AMERICAN July 1978 Volume 239 Number 1

Alternative Automobile Engines

Requirements for cleaner and more efficient engines have stimulated a search for alternatives to the conventional spark-ignition engine. So far the defects of the alternative engines are clearer than the virtues

by David Gordon Wilson

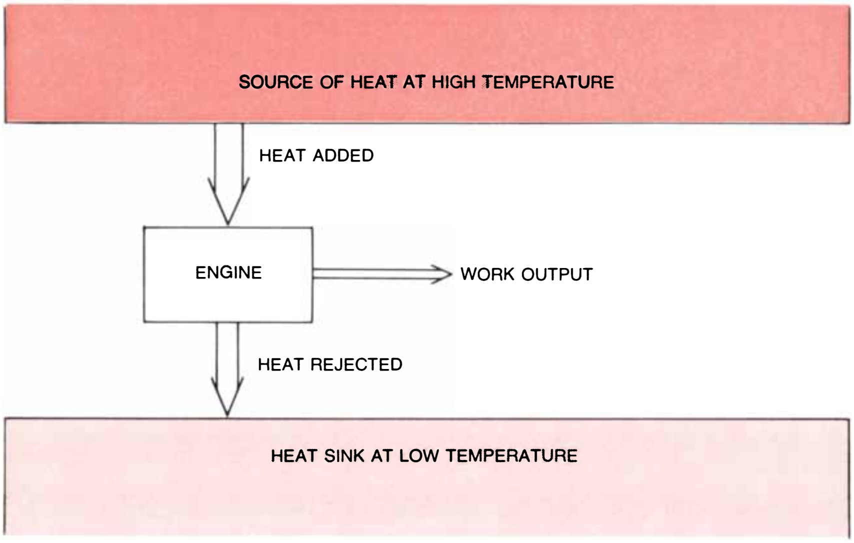
The conventional spark-ignition automobile engine has inherent virtues that become more apparent when alternative engines are pursued. These virtues include a respectable efficiency (particularly under partial load), ease of starting, acceptable emissions (with controls) and a negligible requirement for expensive fabrication materials (and hence a low manufacturing cost). The national concern with energy conservation and air pollution has nonetheless focused Congressional attention on the automobile engine as Americans have known it since early in the century. The 125 million cars and light trucks on the road account for about 16 percent of all the energy used in the U.S. and for a third of all the liquid fuel consumed. The Energy Policy and Conservation Act of 1975 mandated the fuel-economy labeling of all cars and light trucks sold in the U.S. and laid down a schedule of minimum fuel-economy standards that must be met by each manufacturer's "fleet" for each model year until 1985, when the fleet averages must reach 27.5 miles per gallon. The composite city-highway fuel economy of the average 1978-model automobile is 19.6 miles per gallon, up from a low point of 14 miles per gallon in 1974; the low value was attributable to engine resettings adopted that year to meet Federal exhaust-emission standards. Fuel economy advanced in succeeding years with the installation, at some cost, of catalytic converters in the exhaust system of most cars. Unless the scheduled standards are relaxed most 1980 vehicles will require still more costly emission controls.

The propulsion systems most often mentioned as being potentially more promising than the spark-ignition (Otto) engine for meeting or exceeding the Federal requirements in regard to fuel economy and emissions are the compression-ignition (Diesel) engine, the steam (Rankine) engine, the gas turbine (Brayton engine), the Stirling engine and battery-electric drives. All have a potential for substantial improvement over their present state of development that would make them extremely attractive contenders for future mass production, although the battery-electric drive could have only a restricted duty. Perhaps the principal problems facing Government decision makers who want to encourage the development of an improved engine are the missionary fervor with which partisans back one particular engine and the oversimple criteria by which engines tend to be judged. The choice of a new engine is in fact extremely complex. After reviewing the advantages and disadvantages of the various candidate engines I shall suggest the kind of Government policy I believe most likely to identify, quickly and economically, the engine or engines (if any exist) that can supplant the 102-year-old invention of Nikolaus August Otto.

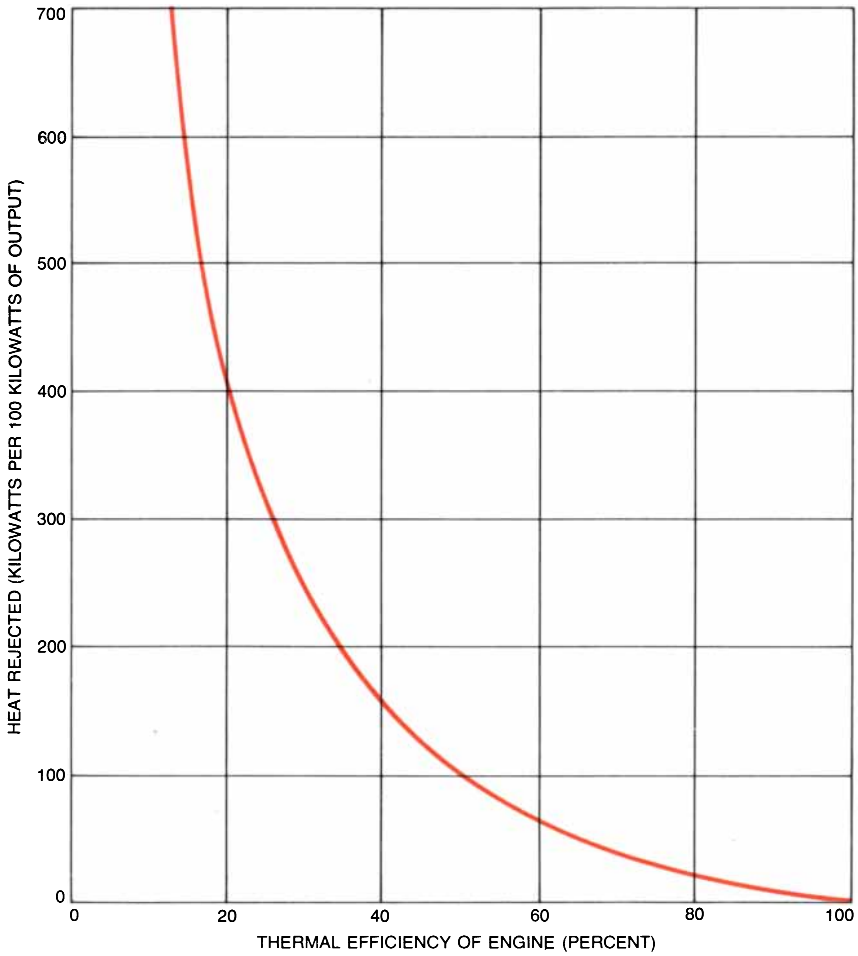
After the problem of air pollution caused by internal-combustion engines was clearly connected with the emission of hydrocarbons, carbon monoxide and various oxides of nitrogen (collectively designated NO_x), and after it became clear that curtailing such emissions was not going to be easy, there was a rash of enthusiasm for the steam engine as an alternative power plant that would remove all the emission problems of the internal-combustion engine. I remember an enthusiast telling me of a steam motor about the size of a football capable of delivering 75 kilowatts (100 horsepower). I was impressed. I asked him how large the boiler would have to be to supply such a motor, and he told me it was about the same size. I relayed this information to a senior colleague who always goes directly to the fundamentally important question. "Ask him," he suggested, "how large the condenser is." That, of course, was the key question. Let me explain why.

Almost all transportation energy comes from heat engines. These are devices that produce mechanical work (shaft power) from a flow of heat provided by a heat source. They convert some of the heat to work, and they must reject some of it to a heat sink: their surroundings. Internal-combustion engines, steam turbines, gas turbines and the like depend on a heat flow supplied by high-temperature gases generated by the combustion of fuel. Most water turbines (and all sailing vessels) get their energy from atmospheric processes driven by the temperature differences that result when radiation from the sun is absorbed by the earth and its atmosphere at one energy level and is radiated back into space at a lower energy level. The only exceptions to the universal application of heat engines for transportation purposes involve muscle power. A basically similar direct-energy-conversion system, the fuel cell, is useful in long space missions. Neither muscle power nor the fuel cell is going to supply a significant fraction of our future highway transportation. We can therefore concentrate on heat engines.

In any heat engine the heat flows from a source at high temperature into a device capable of converting a certain fraction of the heat into mechanical work while rejecting the balance into a sink whose temperature is lower than that of the source. In a steady-state engine the law of conservation of energy (the first law of thermodynamics) requires that the energy flowing out of an engine equal the energy flowing into it. In other words the rate of heat addition



FIRST LAW OF THERMODYNAMICS requires that the energy flow out of a heat engine equal the energy flow into it. Only part of the output can be mechanical work. The balance is heat rejected into a heat sink. Thermal efficiency is defined as work output divided by heat input. According to the principle enunciated by Nicolas Léonard Sadi Carnot maximum thermal efficiency attainable by any heat engine is calculated by subtracting from 1 the ratio of absolute temperature at which heat is rejected to absolute temperature at which heat is added.



AMOUNT OF HEAT REJECTED by a heat engine climbs steeply as the efficiency of the engine falls. The problem of the rejection of heat is a serious one in a closed-cycle engine of moderate to low efficiency, such as a steam engine, because the heat must be removed through a condenser, which adds size, weight and cost to the system. In open-cycle engines, such as the conventional spark-ignition gasoline engine, waste heat is rejected directly to the atmosphere.

is equal to the power output plus the rate of heat rejection. The thermal efficiency of an engine can be defined as the work output divided by the heat input. In order to reduce energy consumption in motor vehicles it is necessary to raise the thermal efficiency of the engine, particularly the efficiency at the low-power conditions in which an automobile engine operates for most of its running life.

For a given value of efficiency transportation energy can be saved only by reducing the required work output. Much of the energy required to propel a vehicle is proportional to the total mass of the vehicle. Energy is required, for example, to accelerate the vehicle from traffic stops and slowdowns, to climb hills and to overcome the rolling resistance of the tires. Air drag is significant only in sustained high-speed driving, and only a small portion of the country's automotive fuel is consumed in this way. Therefore a primary requirement of automotive engines, in addition to high thermal efficiency, is that they be light, and a secondary requirement is that they should not add to the total air drag of the vehicle.

The heat-rejection process in some engines adds both mass and drag. We can rearrange the energy-conservation equation to show that the heat rejected per unit of work output is an inverse function of the efficiency of the engine: heat rejected per unit of work output equals 1 divided by the thermal efficiency minus 1. The equation tells us that if we have a 100-kilowatt engine of 40 percent efficiency, a level that can be approached by a Diesel or Stirling engine, 150 kilowatts of heat have to be rejected. If the efficiency is only 20 percent, a value more typical of low-efficiency internal-combustion or steam engines, 400 kilowatts of heat must be rejected for the same power output. The need to reject heat in such quantities can exact a severe penalty. It means, for example, that a steam engine must carry a big condenser through which a large airflow must be induced.

Nature seems to be punishing the weak and inefficient and rewarding the strong and efficient. Engines such as the spark-ignition engine, the Diesel engine and the open-cycle gas turbine, however, do not suffer a weight or drag penalty for heat rejection because they make use of the working fluid (air) to oxidize the fuel within the engine. That is why they are called internal-combustion engines. The working fluid loses energy and must be discharged at the end of each cycle (hence the term open cycle), and a new charge must be introduced. Therefore with such engines the heat-rejection process is accomplished by the air. (The heat that is rejected in the "radiator" of a piston engine represents a mechanical problem rather than a thermodynamic one: the lubricating oil and certain critical components must be kept reasonably cool. Engine efficiencies would improve if materials and lubricants that did not require cooling of the cylinder could be used.)

To summarize the argument so far, one can say that in addition to the obvious fuel-saving benefit from the attainment of high efficiency in any type of vehicle engine there is a less obvious benefit that applies only to closed-cycle external-combustion engines, because the amount of heat which must be rejected is greatly reduced as the thermal efficiency increases. In a closed-cycle engine, for example a condensing steam engine, the working fluid (water in the case of a steam engine) is sealed within the engine and is heated by combustion external to the fluid. That is why the working fluid must be cooled in a physical device rather than being discharged hot to the atmosphere as it is in an open-cycle engine.

A corollary of the second law of thermodynamics is that the maximum possible efficiency of a heat engine depends only on the ratio of temperatures at which heat is added and heat is rejected. (The simple equation is: maximum thermal efficiency equals 1 minus 1 divided by the ratio of those two temperatures.) If high efficiency is to be achieved in an engine, the temperature ratio must be high. This condition, however, is necessary but not sufficient. An engine with a low temperature ratio cannot have a high efficiency. An engine with a high temperature ratio will have a high thermal efficiency only if it exploits internal processes that do not introduce any of a large number of thermodynamic losses.

Let us compare the processes and components of typical heat engines. All the engines rely on four basic processes: pumping or compression, heat addition, expansion and heat rejection. In internal-combustion, positive-displacement engines the first three processes operate in sequence in a single component. Because they are open-cycle engines the fourth process, heat rejection, operates in the atmosphere.

In the vapor-cycle engine (which may work on steam or on more exotic fluids such as mercury or Freon) and in the gas turbine the four processes operate continuously in separate components. One component is continuously pumping or compressing, another component is continuously heating the fluid and a third component is continuously expanding the fluid. In the case of the closed-cycle gas turbine and the steam engine there is a fourth component, the cooler or condenser, which is also in continuous operation. These components can be optimized for each process, so that the component efficiencies are higher than they are for the sequential processes operating in the single component (the cylinder) of an internal-combustion engine. Some engines also have a regenerator to transfer heat from the hot low-pressure working fluid to the fluid in the cooler, higher-pressure part of the cycle.

The Stirling engine in its most typical modern form is an external-combustion, piston-cylinder engine that has a small quantity of confined hydrogen as its working fluid. Invented in 1816 by Robert Stirling, a Scottish minister, it is a device of considerable thermodynamic sophistication, and it was the first heat engine to exploit the principle of regeneration. Its operation can best be understood with the aid of the illustration on pages 44 and 45. For the present purpose, however, it is not necessary to linger over its subtleties.

The division of engines into those in which different processes operate sequentially in a single component and those in which the four processes operate continuously in separate components has an important bearing on the peak potential thermal efficiency. The sequence of processes in an internal-combustion piston engine is so rapid that very high gas temperatures can be tolerated (of the order of 2,500 degrees Celsius). The walls of the combustion chamber remain at a lower temperature, one that can be safely accommodated by cast iron or aluminum parts and by a fairly ordinary lubricating oil.

In the continuous-process engines, on the other hand, the walls of the combustion chamber and of the expander inlet are continuously scrubbed by gases at high temperature. The maximum allowable temperature is thus limited to what the available (or affordable) materials can withstand. In closed-cycle external-combustion engines, such as the steam engine, the closed-cycle gas turbine and the Stirling engine, the heat added must be transferred through heater walls (for example boiler tubes). The walls must therefore be at a higher temperature than the upper temperature of the working fluid, and the thermal deterioration of the wall materials becomes an absolute limitation. In open-cycle internal-combustion continous-process engines, such as the open-cycle gas turbine, the

	TEMPERATURE RATIO	COMPONENT EFFICIENCIES	EFFICIENCY RELATIVE TO MAXIMUM THERMODYNAMIC EFFICIENCY	HEAT-REJECTION DEVICE (THERMODYNAMIC)
SPARK-IGNITION (OTTO) ENGINE	HIGH	FAIR	POOR	NONE
COMPRESSION-IGNITION (DIESEL) ENGINE	HIGH	FAIR	POOR	NONE
VAPOR-CYCLE (RANKINE) ENGINE	LOW	GOOD	FAIR	LARGE
STIRLING ENGINE	MODERATE	GOOD	VERY GOOD	MODERATE SIZE
OPEN BRAYTON-CYCLE ENGINE (GAS TURBINE)	MODERATE, POTENTIALLY HIGH	VERY GOOD	GOOD	NONE
CLOSED BRAYTON-CYCLE ENGINE	MODERATE	VERY GOOD	GOOD	MODERATE SIZE

HEAT ENGINES SUITABLE FOR AUTOMOBILES exhibit different virtues and defects. The ratio of heat addition to heat rejection establishes the maximum thermodynamic efficiency of the system (*first column*). The highest heat-addition temperatures are obtained when fuel is mixed with air and ignited in an Otto or Diesel engine. The efficiency with which the energy released is thereafter converted into mechanical work depends on the efficiency of the engine's components (*second column*). An engine that has efficient components comes closer to achieving its maximum thermodynamic efficiency than does an engine with less efficient components (*third column*). Additional penalty is levied on closed-cycle engines because they require heat-rejection devices such as heat exchangers (*last column*).

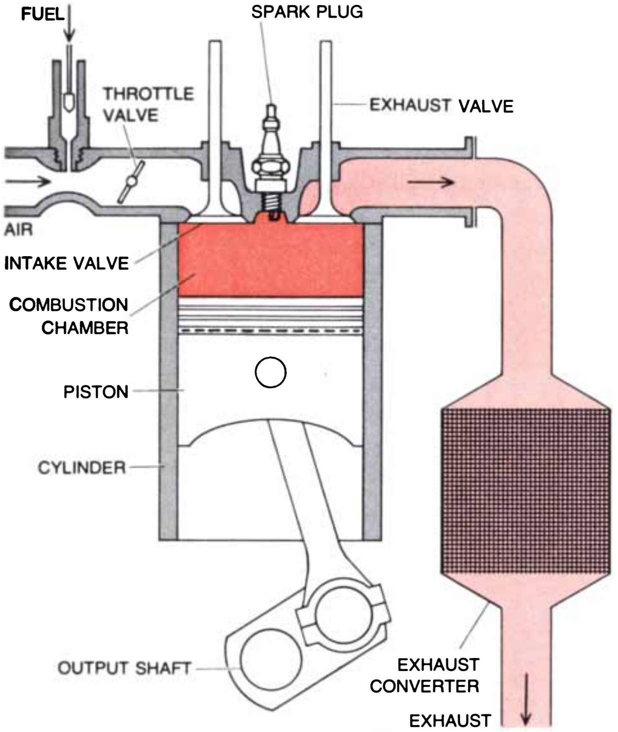
gas temperature can be several hundred degrees above the maximum permissible wall temperature because the heat is released within the stream rather than at the other side of a confining wall.

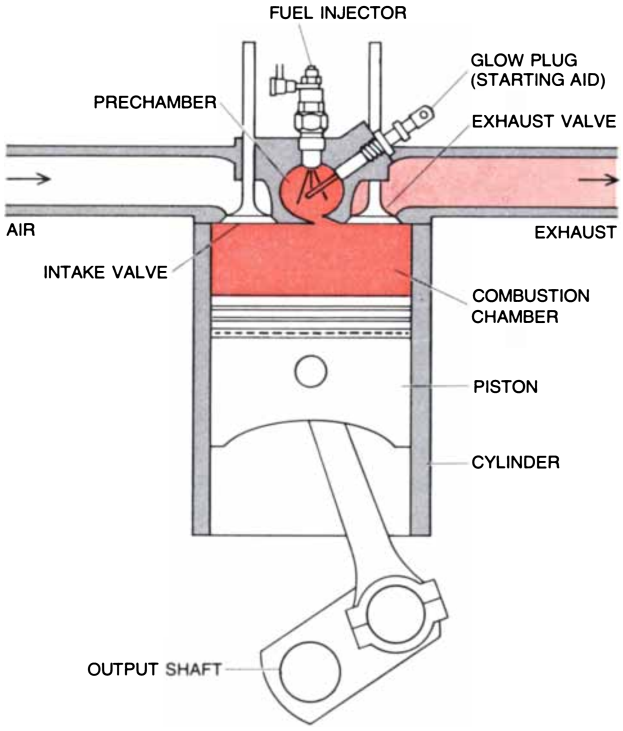
Much effort is being expended to cool the confining walls in the heat-addition and expansion parts of a gas turbine so that even higher gas temperatures can be exploited. In addition, ceramics are being experimented with as materials for gas-turbine construction, not merely in the stationary parts such as combustors and nozzle vanes but also in the rotor. Turbine-inlet temperatures in advanced commercial aircraft are now in the range between 1,250 degrees and 1,350 degrees C., and in military aircraft the temperatures are at least 100 or 200 degrees higher. On test beds turbine-inlet temperatures of well over 1,650 degrees C. have been achieved for at least a decade. Automotive gas turbines use the atmosphere for heat rejection, so that the effective heat-rejection temperature is about 25 degrees C. The ratio of temperatures of heat addition to heat rejection (based on absolute temperatures, that is, temperatures measured in degrees above absolute zero) should be at least five to one, making it possible for the engines to attain a very high thermal efficiency when they are under full load.

Unfortunately vapor-cycle engines cannot use such high input temperatures, not only because the heat has to be transferred through highly stressed heater walls but also because possible working fluids have peculiar properties at high temperatures. At temperatures above 600 degrees C. steam becomes almost corrosive, causing steels to adsorb hydrogen and become brittle. In the steam turbines of stationary power plants a maximum temperature of 566 degrees C. has been almost universally adopted for the past several decades following several disastrous failures in turbines operating at higher temperatures. Some experimental automobile turbine engines, however, have been operated at temperatures substantially above 566 degrees through the use of rather expensive materials. The best organic fluids that have been developed for vapor cycles decompose when they are heated above 400 degrees or so. Mercury vapor at a peak temperature of about 650 degrees was employed in some two-fluid cycles in the 1930's, and at a temperature of 705 degrees in recent spacecraft units. Mercury, however, gets one into extraordinary problems, for example the problem of promoting the wetting of condenser surfaces in order to achieve high efficiency.

Nature has therefore been unkind to vapor cycles. It has limited the temperature at which heat may be added, which thereby reduces the potential peak efficiency, which in turn enormously increases the amount of heat that must be rejected. In large stationary power plants, however, thermodynamic efficiencies as high as 44 percent can be achieved by reducing thermodynamic losses to a minimum and achieving the lowest possible heat-rejection temperature, notwithstanding the low limit that has to be placed on the heat-addition temperature. The condenser, although it is expensive, does not have to be penalizingly large as long as cold seawater or river water are available. Otherwise enormous cooling towers must be installed to reject the heat. Current nuclear power stations are constrained to use input steam temperatures even lower than those in fossil-fuel-fired plants, with the result that heat-rejection costs in nuclear plants are very large even when the condensers can be cooled by seawater or river water.

In the small engines required for vehicles one cannot resort to the special measures taken in large stationary plants to reduce thermodynamic losses, such as having a large number of turbine stages, reheating steam during expansion and heating feedwater by bleeding off a little steam from the turbine. As a result the maximum obtainable efficien-





CONVENTIONAL (OTTO) ENGINE requires a spark plug to ignite a mixture of fuel and air. The energy contained in the expanding gases is absorbed by a piston and transferred to an output shaft. On the return stroke the piston drives out the spent gases, preparing the cylinder for the next charge of fuel. Most new automobiles are equipped with catalytic converters to reduce the exhaust emissions.

DIESEL ENGINE dispenses with a spark plug and ignites the fuel mixture in the cylinder by the heat of compression. The compression ratio of a typical Diesel engine is about 20 to one, or more than twice the compression ratio of most spark-ignition engines. A glow plug facilitates starting. Except for oxides of nitrogen (NO_x), Diesels release a smaller quantity of pollutants than spark-ignition engines do.

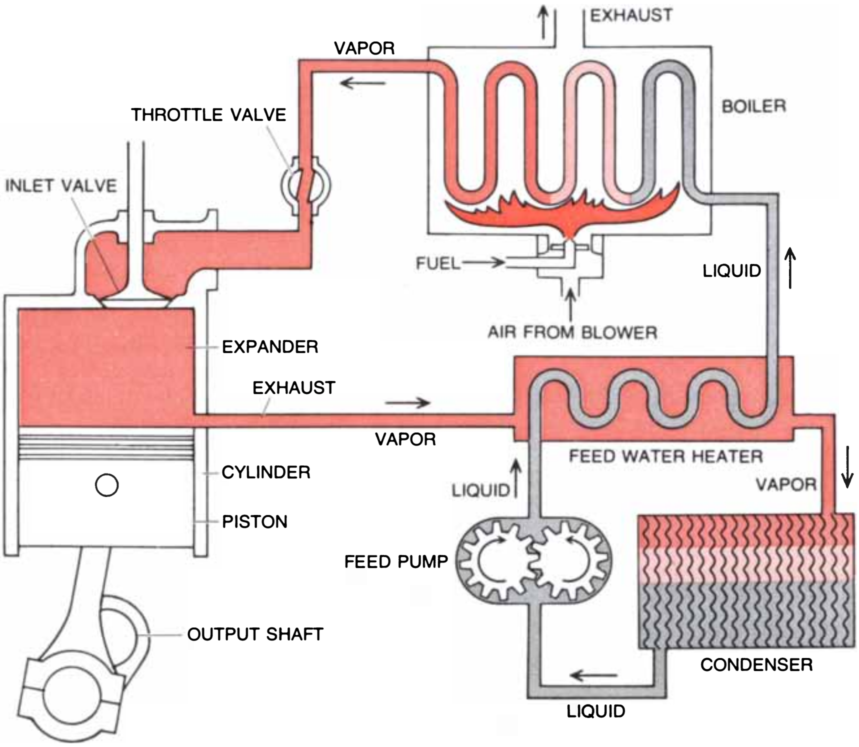
cy drops to about 30 percent. Even this value implies that the temperature of heat rejection approach the temperature attainable in large power plants, namely, about 35 to 40 degrees C. To attain this level with a moderate air temperature of 25 degrees C. (77 degrees Fahrenheit), however, would require an extremely large airflow and such a large condenser that it would dominate any vehicle.

Accordingly the automobile-engine designer reduces the airflow requirements by allowing a larger temperature rise for the air, which means the condensing temperature has to be increased, the thermal efficiency deteriorates further and the amount of heat to be rejected increases further still. The condenser does get smaller, however. A condenser of acceptable size, somewhere near the minimum, is usually reached at a heat-rejection temperature of between 95 degrees and 120 degrees C., at which point the overall thermal efficiency, allowing for the power deductions necessary to drive the condenser fan and another fan to push air into the combustion chamber, drops for current steam-vapor cycles to between 22 and 24 percent.

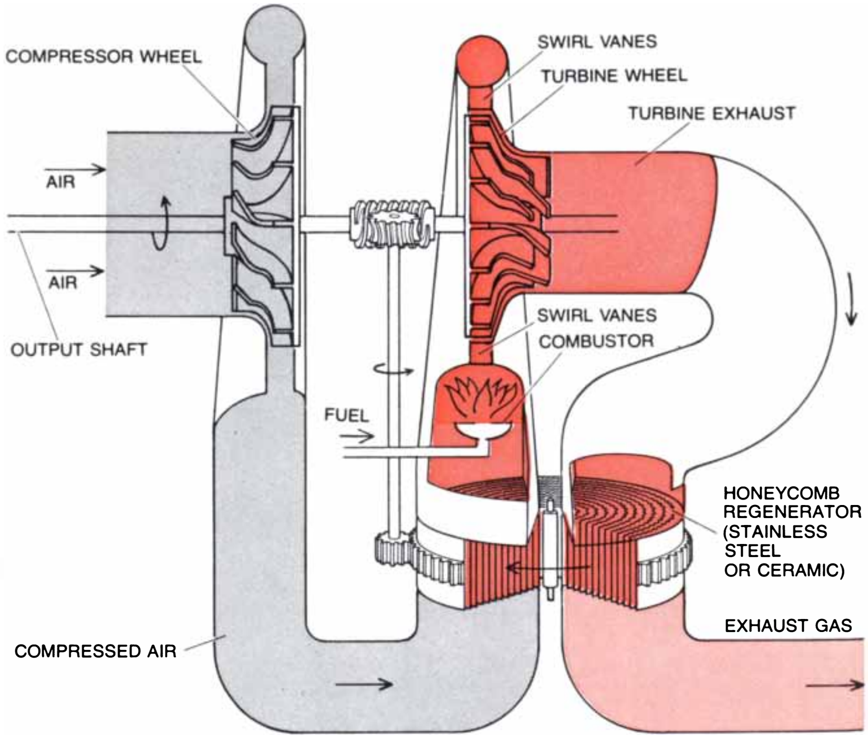
Marginal reductions in condenser size can be achieved with better fins on the tube or plate surfaces of the condenser. Any substantial reduction in the size of the condenser, however, will call for either a new ice age, making it possible to have lower heat-rejection temperatures with higher efficiencies and lower quantities of heat rejected, or much higher steam temperatures at the expander inlet, through the exploitation of materials that are not subject to being embrittled by hydrogen.

So far I have discussed the limitations to power-plant thermal efficiency only in terms of the overall ratio of heat addition to heat rejection, and the additional penalty imposed on closed-cycle engines of having to raise the heat-rejection temperature when the engine efficiency falls below some critical value, roughly 30 percent. The efficiencies of the separate engine components, that is, the efficiency with which each process operates, may differ greatly from one engine to another and from one cycle to another. In the case of the steam engine, pumping, the first of the sequential processes basic to all heat engines, absorbs only a small fraction of the engine's total output because the fluid being pumped (forced into the boiler) is in the liquid state. Thus it does not matter very much whether the pump is 90 percent efficient or only 75 percent; the amount of "negative work" the pump is called on to perform is small.

It is quite otherwise in the case of the gas turbine, where the fluid being pumped is a gas and the amount of negative work required is very large. Charles Dickens' Mr. Micawber defined misery



VAPOR-CYCLE (RANKINE) ENGINE transfers heat from combustion of fuel to a closed system containing a condensable working fluid, water in the case of a steam engine. Working fluid is vaporized and superheated in a boiler, expanded to produce power and condensed for reuse. Pistons, which provide positive displacement, are preferred tc turbines for automobiles.



GAS TURBINE (BRAYTON ENGINE) is driven by the expansion of hot gases through a turbine wheel. A portion of the output of the turbine is used to compress the inlet air that is fed into the engine's combustor. In configurations under development for automobiles the exhaust gases from the turbine pass through a honeycomb regenerator that acts to heat the incoming air.

as a state of having an income 5 percent less than necessary expenditures and happiness as a state of having an income 5 percent greater than expenditures. Early gas-turbine designs were poor because the compressors (the pumps) absorbed more power than the expanders produced. The net output of the gas turbine is the difference between these two large and often similar quantities. For this reason I call it a Micawber engine: one that is very sensitive to component efficiencies.

The first gas turbine with which I was involved was capable of only sustaining itself and could produce no useful power. Under the circumstances an improvement of 1 percent in the efficiency of either the expander or the compressor would have raised the overall efficiency from virtually zero to some finite value, thus making the engine infinitely sensitive to the efficiency of its components. One way to increase the expander output of a gas turbine, of course, is to increase the temperature of the inlet gas. Such an increase will also raise the overall thermal efficiency of the engine, provided the heat-rejection temperature remains unchanged.

As gas turbines become more efficient by virtue of the increase in their expander-inlet temperatures the difference between the power output of the expander and the power-input requirements of the compressor increases and the sensitivity of the thermodynamic efficiency of the engine to the efficiency of its components decreases. In addition, the separate components have been developed to achieve high individual efficiencies.

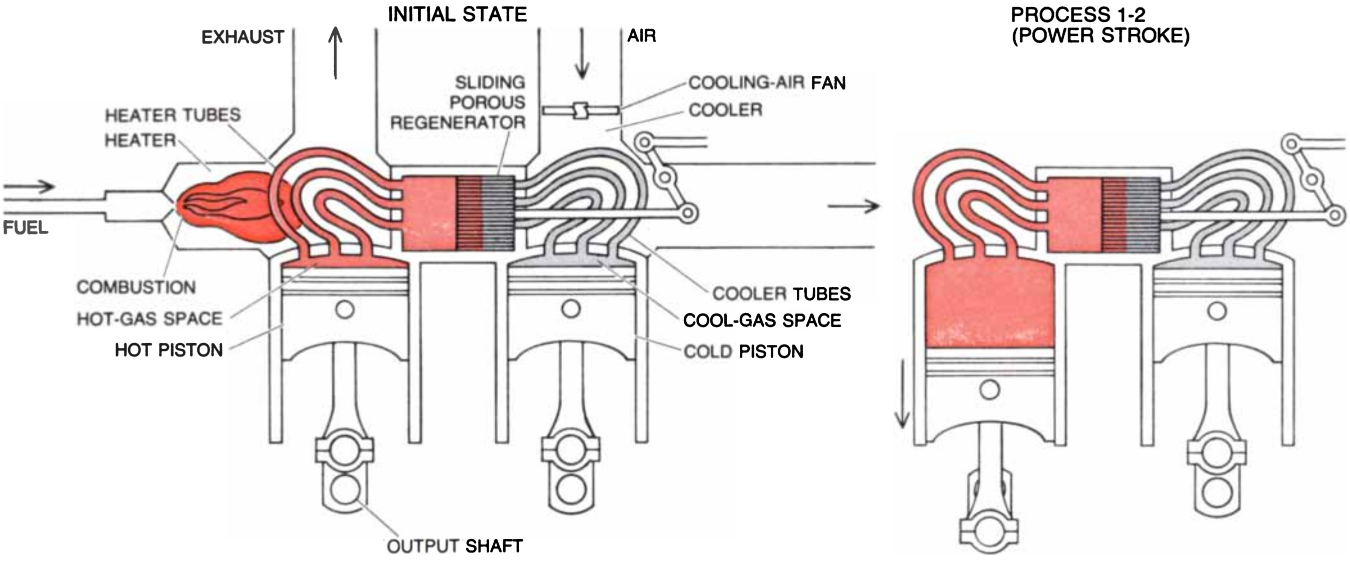
A spark-ignition piston engine conceals within itself an expander that is producing much more power than the advertised engine horsepower and a compressor (the same piston and cylinder) that is absorbing most of the difference between the expander power and the actual engine power output. Since piston engines can tolerate high temperatures, however, they are not very sensitive to the efficiencies of their separate processes. This is just as well, because the compromises involved in making one component—a cylinder-piston combination—serve three functions significantly reduce the efficiencies of the individual processes.

The thermal efficiency of an engine operating at full load is not the entire story in comparisons of engines. Most automobile engines spend most of their running lives operating at power levels below 25 percent of full power. The conventional gas turbine in particular suffers a sharp drop in efficiency under partial loads. In order to reduce the power in a normal open-cycle gas turbine one reduces the fuel flow, which reduces the overall temperature ratio and the pressure of the gas leaving the compressor. The compressor efficiency can drop off sharply from the design point, particularly if the compressor is designed for high pressure ratios. Therefore high-pressure-ratio open-cycle gas turbines have poor part-load efficiencies. The gas turbine can, however, be made in many alternative forms, some of which have an acceptable efficiency under partial load.

All other possible alternative engines behave fairly well at low load insofar as efficiency is concerned. The efficiencies of spark-ignition and compression-ignition engines do decrease at low load, but not unacceptably. The efficiencies of Stirling engines, and of some configurations of vapor-cycle engines and closed-cycle gas-turbine engines, can actually increase at low loads. This is partly because heat exchangers normally become more efficient when they handle smaller mass flows at lower velocities, partly because components generally show lower pressure losses when the velocities are reduced and partly because at low loads parasitic losses from fans for the condenser or cooler can be eliminated.

The initial impulse for finding a new engine for the automobile was the recognition of the large quantities of pollutants that the engines of only a decade ago were emitting. In the conventional engine many of the pollutants are formed as a consequence of the intermittent combustion process, with its rapid chilling of the combustion products. All the continuous-combustion engines emit far smaller quantities of pollutants than the spark-ignition engine emits in the absence of control devices. Experimental models of vapor-cycle engines, gas-turbine engines and Stirling engines have surpassed the most stringent emission requirements yet established.

Although the Diesel engine is also an intermittent-combustion engine, it has the advantage of operating with excess air, so that its carbon monoxide emission is negligible. Moreover, hydrocarbons are normally a small constituent of Diesel exhaust. Oxides of nitrogen are, however, formed in hot flames and are



STIRLING-CYCLE ENGINE, invented in 1816 by Robert Stirling, was the first heat engine to take advantage of the principle of regeneration, which makes it possible to recapture heat that would otherwise be wasted. The Stirling engine has many possible configurations; the schematic (and hypothetical) one shown here was selected to clarify the engine's principle of operation. The heat of combustion is absorbed continuously by a gas, usually hydrogen, held in a sealed system. In the initial state the hot-gas space is filled with hot gas at maximum pressure. Both pistons are at their innermost positions. The regenerator is at the right, or cold, side of the engine. In the power stroke, process 1–2, the hot piston is driven downward by expanding high-pressure gas, delivering mechanical energy to a flywheel and the output shaft. The heater tubes continue to heat the expanding gas. In the transfer stroke, process 2–3, the hot and cold pistons move

"frozen" by rapid chilling. The Diesel engine cannot at present meet the stringent NO_x standards mandated for 1980 and beyond. There are also unsubstantiated suspicions that Diesel-engine exhaust particulates carry carcinogens.

In view of the difficulty of optimizing all the efficiency and emission requirements in a single engine, hybrid propulsion systems have their advocates. I shall define a hybrid engine as one consisting of a small engine running steadily combined with some form of energy storage to supply supplemental power for meeting normal varying load requirements. Such hybrids appear attractive because a large, conventional spark-ignition engine is both inefficient and at its worst in generating pollutants when it is running at low power. Accordingly most hybrid-engine concepts have involved a small steady-running spark-ignition or Diesel engine, whose emissions are easily controlled, coupled with flywheel, battery or compressed-air energy storage.

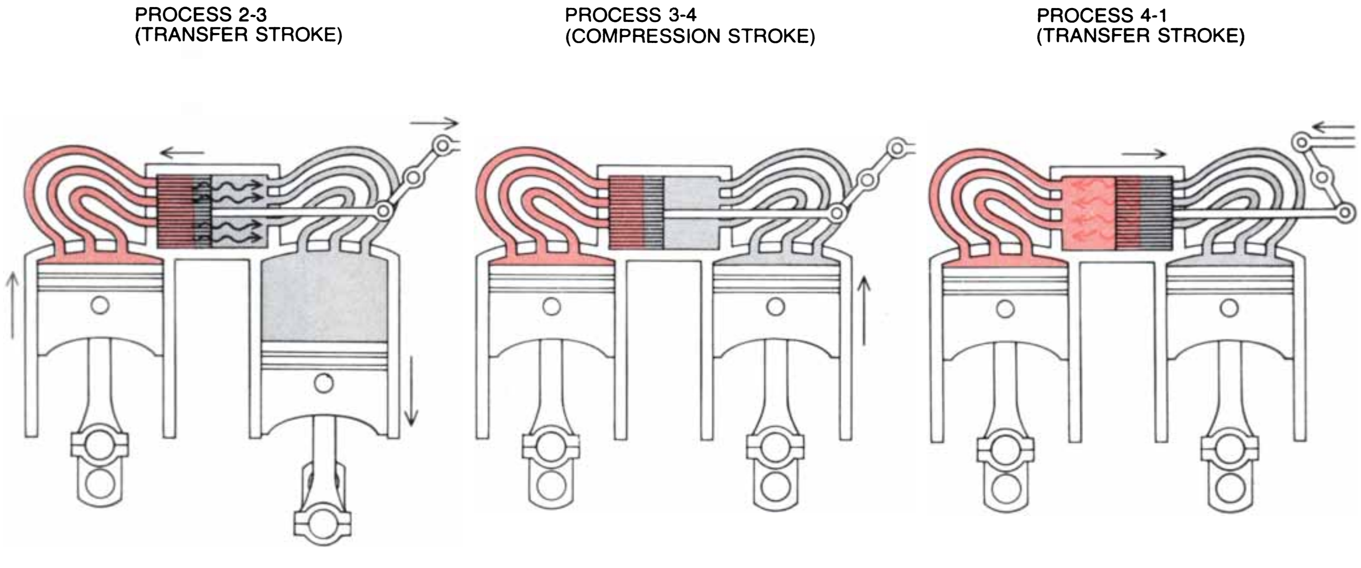
Although such systems often look good on paper, they usually turn out to be heavy, complex and expensive. They can be designed to accelerate an automobile satisfactorily under normal conditions, but they intrinsically cannot match the performance of today's cars on long climbs or in other circumstances requiring high and sustained engine output, as for example when the car is towing a trailer.

Finally we must consider the case for the all-electric vehicle, which is inherently emission-free, at point of service if not at the point of original power generation. Automobiles and light trucks powered by storage batteries and electric motors labor under two substantial disadvantages. The first is that in comparison with present-day gasoline engines, storage batteries and electric motors are both very heavy for their respective tasks, namely energy storage and power generation. Batteries much lighter and more durable than today's can undoubtedly be developed, but the combined mass of the batteries, the motor and the control mechanism still seems likely to exceed the weight, in kilograms per kilowatt, that can be achieved by fuel-burning power systems. Such a weight penalty implies an energy-loss penalty even with equal power-plant efficiency.

The second and perhaps crucial disadvantage is that the overall efficiency of electric propulsion, measured from the fuel consumed in the central power plant to the power delivered on the road, probably does not exceed the efficiency attainable with any of the alternative propulsion systems, and in fact it may be a good deal lower. To compute the efficiency of power delivered to the wheels of an electric vehicle one must multiply the sequential efficiencies of five separate processes: the thermal efficiency of the central power plant, the efficiency of the electric-power grid that delivers energy to the individual consumer, the efficiency of the battery charger, the efficiency of the battery charge-discharge cycle and the efficiency of the motor and controller. If one takes the highest plausible efficiency for each step in the chain, the sequential product is $.40 \times .95 \times .95 \times .85 \times .75$, or 23 percent. Actually the average efficiency of fossil-fuel steam generating plants in the U.S. is only 33 percent, not 40 percent, and the thermal efficiency of nuclear power plants is about 1 percent lower. If the sequential product of the five steps is calculated for a more typical set of values, the net efficiency of all-electric propulsion drops to about 12 percent.

Thermal efficiency, of course, is not the only measure of value. Liquid fuels are an increasingly scarce and costly commodity. Battery-electric vehicles can make use of energy derived from coal or nuclear fuels (although for some people the latter alternative is no attraction). In addition, most batteries can be charged during the night, when they will improve the load factor of power generation and thus the overall generating efficiency. Whether or not such advantages will be sufficient to compensate for the disadvantages of electric propulsion is not clear at present. In any case the comparisons I have made apply only to vehicles that can be substituted directly for today's all-purpose automobile. Battery-electric cars suitable only for low-speed local commuting and shopping trips would require smaller batteries, carry a lower weight penalty and therefore would be more attractive for such purposes.

One critical factor often overlooked in evaluating all-electric propulsion systems is the enormous load that would be placed on the electric-utility industry if a large fraction of all vehicles on the road were electrically powered. The nation's 125 million cars and small trucks now consume half as many energy units

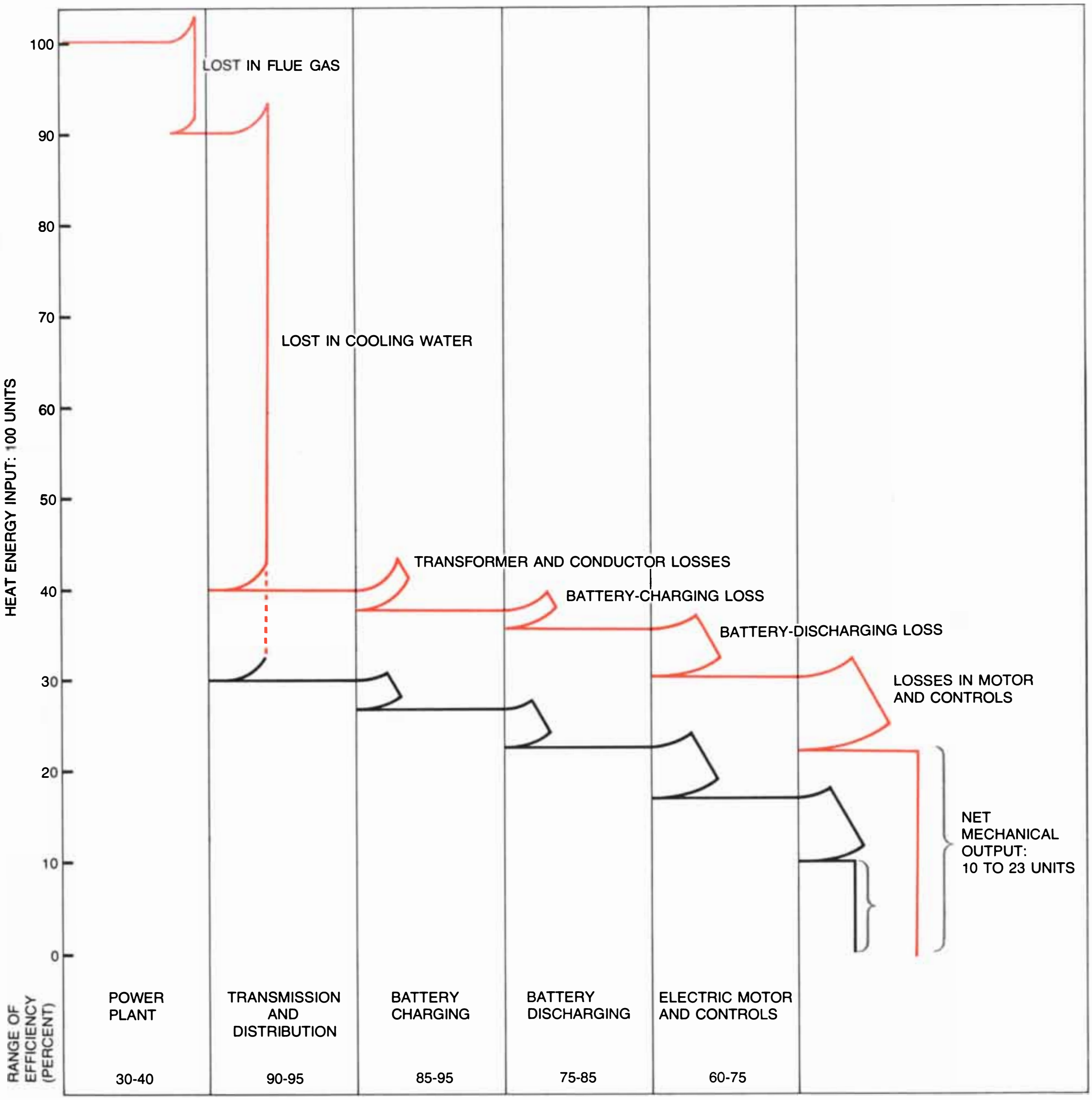


in opposite directions as the regenerator travels to the left. Essentially all the gas is forced to pass through the regenerator, which "soaks up" and stores much of the heat contained in the gas. Since the gas falls in temperature at constant volume, its pressure falls. In the compression stroke, process 3–4, the cold piston travels upward, compressing the gas as its temperature is held constant by the tubes of the cooler. In the second transfer stroke, process 4–1, the pistons remain at their innermost positions as the regenerator moves to the right. As the gas passes through the regenerator it picks up the stored heat, therefore increasing in temperature at a constant volume. In consequence the pressure of the gas rises, creating the conditions for the next power stroke and a repetition of the cycle. The regenerator provides the means for capturing the heat that is continuously entering the system during the interval between the power strokes.

as the nation's electric-power plants. In other terms roughly a fifth of all the new generating capacity planned between now and the year 1990 would be needed to provide the power for even 50 million vehicles with present-day performance and efficiency.

For completeness I should again mention the fuel cell, a device in which hydrogen or some hydrocarbon compound is converted directly into electricity. It is true that fuel cells can operate with a net efficiency of 40 to 50 percent and could deliver their output directly to a motor and controller for an overall thermal efficiency of at least 25 percent and perhaps as high as 40 percent. Cost and durability aside, however, all current projections show fuel-cell power systems to be so heavy that they must be considered out of the running for future vehicles.

Let us now sum up. As we have seen, the venerable spark-ignition engine is difficult to beat except with respect to exhaust emissions. Moreover, the en-



EFFICIENCY OF ELECTRIC PROPULSION for automobiles depends heavily on the efficiency with which the heat units in the original fuel are converted into electricity at the central power plant and transmitted to the consumer before he can begin to charge the batteries in the automobile itself. The colored curve represents about the best that can be achieved at each step along the way with known technology. Out of 100 energy units, only 38 reach the battery charger and 23 finally emerge in the form of mechanical energy available to drive the automobile. If each step is performed with efficiencies at the lower end of the typical range, as few as 10 units may emerge in mechanical form. The chart may be misleading if one concludes that condenser losses are the most significant ones. Actually condenser losses are of low-temperature heat that has little work-producing potential ("availability"). The flue-gas losses are of a somewhat higher-temperature heat and therefore represent a higher loss in availability. All the other losses are of pure work energy and are therefore thermodynamically very costly. The chart is based on an analysis that was conducted by Paul D. Agarwal of the General Motors Corporation.

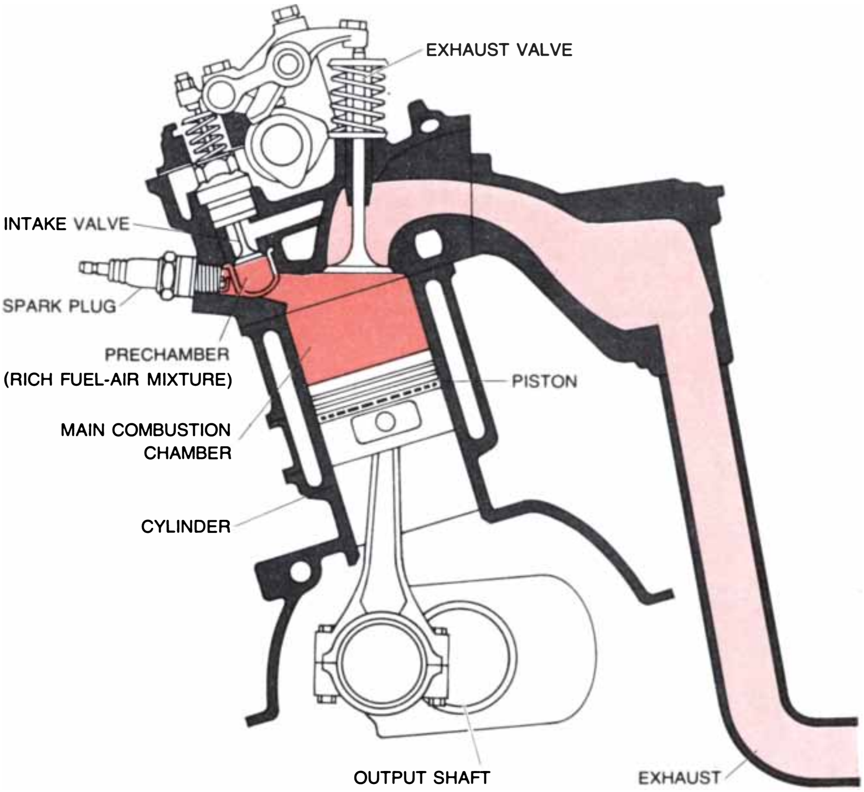
gine has been substantially cleaned up, although (with one or two exceptions) the task has called for catalytic devices that tend to degenerate with service. Otherwise the conventional engine is light in weight, acceptably efficient and extremely cheap to build. At least until recently the manufacturing cost of spark-ignition automobile engines was well under $2 per horsepower.

A promising alternative to catalysts for reducing emissions from the spark-ignition engine is a modification of the combustion chamber and fuel-injection system that results in a two-stage combustion process. A very rich fuel-air mixture is ignited first and is then used to burn a very lean mixture. In this way the maximum combustion temperature is lowered, with a consequent reduction in the production of NO_x emissions. Engines of this type are known as stratified-charge engines. One built by Honda, the Honda CVCC, has been in commercial production since 1974, and others are under development. Whether or not such engines will win out over more conventional engines modified to burn very lean, uniform mixtures reliably cannot be predicted.

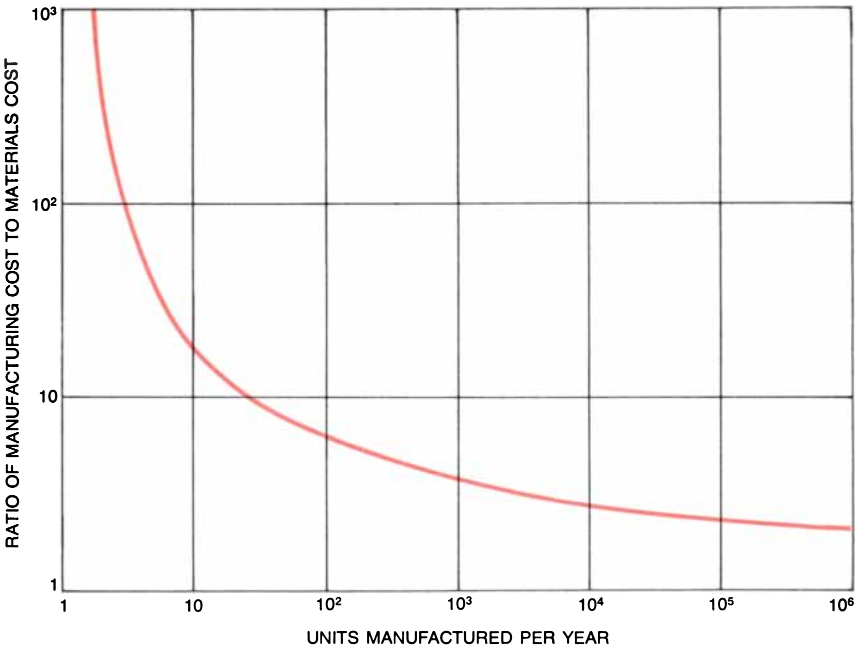
The leading alternative to the spark-ignition engine is the compression-ignition engine, the Diesel, available in the Mercedes and Peugeot for many years, and now being offered in the Volkswagen, Oldsmobile and Cadillac. The present Diesel engines are less attractive than the spark-ignition engine in three major respects. They are heavier for a given power, they are more expensive (as much as twice as costly per horsepower) and they appear unable to meet the stringent NO_x levels due to be enforced in 1980. Lesser drawbacks are hard starting in cold weather, a smaller range of operating speeds, more noise and an unpleasant exhaust odor, with the possibility of some degree of toxicity remaining as a threat.

Nevertheless, the Diesel engine can be improved. One method, used in many highway trucks and introduced in passenger cars by General Motors, is to couple the Diesel with a small turbine operating on the exhaust gas from the Diesel to run a compressor that raises the pressure of the air entering the engine's combustion chambers. Such an arrangement lowers the compression ratio required in the Diesel, reduces the overall engine weight for a given output and probably reduces the overall cost.

Another suggestion is to redesign the automobile Diesel for a two-stroke cycle, again as is used in many trucks and buses, so that fuel is injected on every revolution of the crankshaft instead of every two revolutions, as it is in the conventional design. The advantage for small engines seems questionable, however. Also under study are devices to swirl the air more intensely in the com-



STRATIFIED-CHARGE ENGINE meets exhaust-emissions standards without requiring an exhaust-gas recirculation system or a catalytic converter. The example shown here, the Honda CVCC, has a main chamber and a prechamber for each cylinder. A rich fuel-air mixture enters the prechamber at the same time that a very lean mixture enters the main chamber through a separate intake valve (not visible because it is located behind the exhaust valve). The spark plug fires the rich mixture, which in turn ignites the lean main mixture. This results in a stable slow burn. Peak temperatures are low enough to minimize formation of nitrogen oxides and mean temperature is high enough to limit release of carbon monoxide and hydrocarbons.



ENGINE MANUFACTURING COSTS are closely bound to the cost of construction materials and drop sharply with increasing volume. In effect the costliness of the materials sets the total cost of the engine. In comparison with alternative power systems the materials used in today's automobile engine are cheap. The cost of present engines is about $2 per horsepower.

bustion chamber at the point of fuel injection with the goal of reducing exhaust emissions and odor, together with the somewhat irritating Diesel "knock" and the maximum pressure needed for ignition. The net result may be a reduction in engine weight at a small cost in efficiency. It remains to be seen whether these stratagems or others will make the Diesel a fully competitive automobile engine.

Although no major manufacturer is pursuing the steam engine, it always has its devoted followers. The engines under development work with steam rather than organic vapors simply because the higher temperatures attainable with steam yield the highest efficiencies and the smallest condenser sizes. Efficiencies nonetheless remain low, and the only real hope of major improvement is to incorporate rather expensive materials that would make it possible to raise the steam temperature. There is finally the troublesome problem of the freezing of the water in cold weather, for which no entirely satisfactory solution has yet been devised.

The Stirling-cycle engine is under joint development by the Ford Motor Company and N. V. Philips Gloeilampenfabrieken in the Netherlands. Hydrogen is the preferred working fluid because it transfers heat better than any other gas and offers a minimum of fluid friction. The part-load efficiency of some experimental Stirling engines is higher (40 percent) than that of any present competitor. The engine has the further advantage, shared by steam and gas turbines, of being able to operate on almost any fuel. It is very quiet, is the

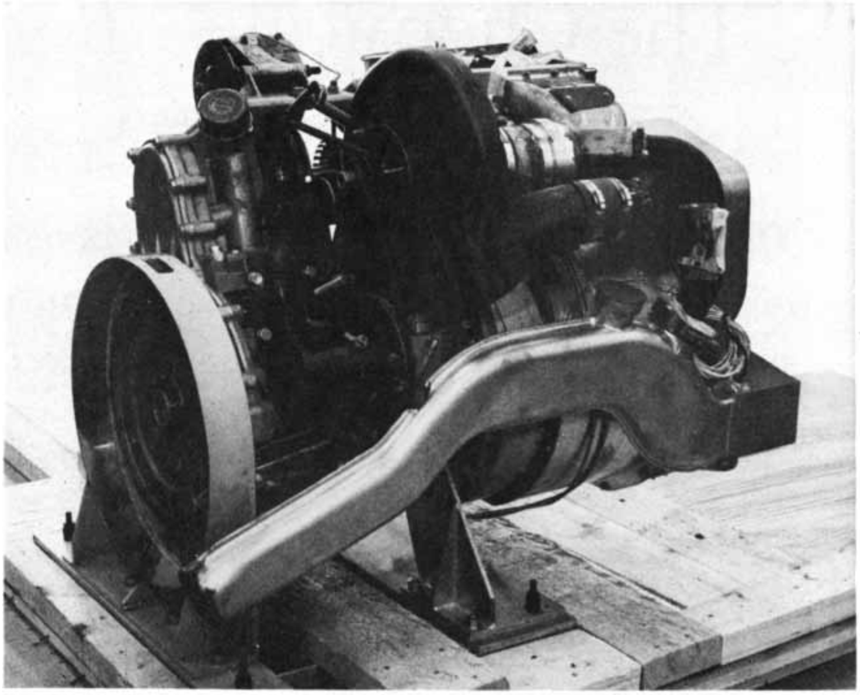
	ENGINE EFFICIENCIES (PERCENT)			POWER/MASS RATIO KILOWATT/KILOGRAM	EMISSIONS	MANUFACTURING COST
	FULL LOAD	25 PERCENT LOAD	10 PERCENT LOAD			
SPARK-IGNITION (OTTO) ENGINE	26 / 30	18 / 27	15 / 24	.3 / .8	ACCEPTABLE WITH CONTROLS; LIKELY TO DEGENERATE	VERY LOW
COMPRESSION-IGNITION (DIESEL) ENGINE	26 / 36	20 / 35	18 / 32	.2 / .5	HIGH IN NITROGEN OXIDES (NO_X)	MODERATE
VAPOR-CYCLE (RANKINE) ENGINE	20 / 30	18 / 26	15 / 22	.2 / .5	VERY GOOD; NOT LIKELY TO DEGENERATE	HIGH
STIRLING-CYCLE ENGINE	30 / 42	30 / 40	28 / 38	.2 / .5	VERY GOOD; NOT LIKELY TO DEGENERATE	HIGH
OPEN BRAYTON-CYCLE ENGINE (GAS TURBINE)	25 / 44	10 / 30	8 / 25	.6 / 1	VERY GOOD; NOT LIKELY TO DEGENERATE	MODERATE; POTENTIALLY LOW (CERAMIC)
CLOSED BRAYTON-CYCLE ENGINE	22 / 36	22 / 36	20 / 34	.2 / 1	VERY GOOD; NOT LIKELY TO DEGENERATE	VERY HIGH

PROPOSED POWER PLANTS may excel the present-day engine in efficiency, particularly under partial loads, but their projected manufacturing costs are high, with the possible exception of the gas turbine. Their biggest present virtue (the Diesel excepted) is low exhaust emissions. Efficiencies are given as a range in which the lower figure represents the efficiency attainable with present technology and the upper figure represents what may be attained in advanced engines 10 years from now. Power-to-mass ratios are given on the same basis. All figures for the closed Brayton-cycle engine are highly speculative since it has never been built in an automobile configuration.

easiest of all alternative engines to start and has no freezing problem. Because present experimental models need a substantial amount of high-temperature metal they would be expensive to manufacture. There is hope that alternative ceramic components can be developed. Since the Stirling engine works on a gas cycle, it has the potential of operating at still higher temperatures and pressures than it does at present, which would make it even more efficient and lighter.

The lightest available engines are the gas turbines. With the operating temperatures already in prospect the thermal efficiency of the gas turbine can be the highest of all alternative engines, perhaps as high as 55 percent. On the other hand, the open-cycle gas turbine has probably the lowest efficiency under part load. The part-load efficiency improves with the addition of a heat exchanger and the lowering of the pressure ratio. If the cycle is completely closed, the part-load efficiency is second only to that of the Stirling engine. Such closed-cycle turbines, which have been seriously proposed but not yet built in an automobile configuration, would have a lower thermal efficiency than an open-cycle turbine because heat would have to be transferred through highly stressed heater walls and because a heat-rejection cooler would be needed that would raise the heat-rejection temperature above that of an open-cycle turbine. My own opinion is that the Stirling engine offers a better prospect for automobiles than the closed-cycle gas turbine.

The low-pressure-ratio gas turbine with a heat exchanger (or regenerator) is the best gas-turbine prospect for automobiles and has received the most development effort, notably by the Chrysler Corporation. In its present form it uses substantial amounts of high-temperature materials and would thus be expensive. Efforts are being made to incorporate ceramics for the hot parts of the engine, with the possibility of achieving turbine-inlet temperatures of 1,650 degrees C. Simultaneously costs would decrease because ceramics should cost less than $1 per kilogram, whereas heat-resistant alloys can cost more than $20 per kilogram. The manufacturing costs would decrease proportionately, because when mass production is achieved, the costs of manufacturing any engine fall to only a small multiple of the cost of materials. The efficiency of such a gas turbine under partial load should exceed that of the conventional spark-ignition engine; hence the turbine would be hard to beat for automobile applications and many others.



EXPERIMENTAL STIRLING ENGINE has been developed jointly by the Ford Motor Company and N. V. Philips Gloeilampenfabrieken in the Netherlands. Because the Stirling engine burns fuel continuously under optimum conditions, a feature it shares with steam engines and gas turbines, its exhaust is low in pollutants. The Stirling engine is potentially more efficient under partial load than any competitor, an important consideration for highway vehicles.

I have presented the argument that the choice of an alternative to the now-dominant spark-ignition engine is highly complex. If the Government were to throw its support behind just one or two alternative approaches, the action would be not only unfair but also probably counterproductive because of the inhibiting effect on the alternative engines not supported. What, then, should be the Government's role?

The Government has two principal choices when it is faced with activities that involve what economists refer to as large external costs, that is, costs that must be borne not by the individual consumer but by the society as a whole. The sticker price of an automobile that pollutes the air everyone breathes does not include an assessment for the multifarious consequences of such pollution. That cost is externalized. Similarly it can be argued that there is an externalized cost when automobile drivers in effect "destroy" a large proportion of a nonreplaceable resource such as petroleum.

In such situations the Government has the choice of regulating or of taxing. Regulation is clearly preferable when the total costs of regulating are acceptably small, or when the costs of taxing are unacceptably high. Thus if the Government had set easily attainable emission limits and fuel-consumption goals, the costs of regulation would have been acceptable. The emission limits actually legislated, however, are quite stringent and the fuel-economy goals are quite high. In such circumstances the costs of the alternative approach, taxation, would almost certainly have been less. The Government could have put a graduated tax on the measured emissions of each make and model of car and an energy tax increasing the price of fuel, or possibly a comparable energy-thrift tax on each model's measured fuel consumption.

Such a policy would have internalized, in a rough but approximate way, the external costs of pollution and the wasteful consumption of a disappearing resource. Most consumers would presumably choose cars with the lowest taxes consistent with their transportation requirements. Manufacturers, freed from the uncertainties of changing regulations, would presumably develop those engines most attractive to consumers. The incentives offered to the buyer and the seller would tend to drive the engine-development effort toward a social optimum. To complete the feedback loop the proceeds of the taxes should be redistributed directly to the people who now bear the cost of the externalities, which is essentially the entire population. In other words, the taxes should be rebated equally to all legal adult residents of the U.S.

The logistics of automobile production prohibit the introduction of any new engine on a large scale before the 1990's. It is my view that the optimum way of choosing a future engine, or engines, is through the modified free market created by taxation, rather than through regulation. For the present we can only wait to see what will emerge from the path actually selected.

The Chemical Differentiation of Nerve Cells

The development of the intricate network of nerve cells that makes up the nervous system requires each cell to "choose" a transmitter substance appropriate to its specific connections with other cells

by Paul H. Patterson, David D. Potter and Edwin J. Furshpan

The flow of information through the nervous system is determined by the particular pattern of connections made by the neurons, or nerve cells, with one another and with other body tissues such as muscles and glands. How this vastly complex network is established during the development of the organism is a basic and intriguing question. In the embryo the initially separated neurons make contact partly by migrating and partly by sending out the thin extensions of the cell body called axons and dendrites. These fibers usually branch repeatedly and may project to distant parts of the nervous system or leave the nervous system to innervate effector tissues. The site of communication between two neurons or between a neuron and an effector cell is the synapse. Here the axon terminal of the innervating neuron comes very close to the target cell but does not quite touch it; the intervening gap is filled with fluid and is called the synaptic cleft. On the arrival of a nerve impulse molecules of a neurotransmitter chemical are released from the terminal, travel across the synaptic cleft and combine with highly specific receptor proteins on the surface of the target cell. The activation of the receptors by the neurotransmitter triggers an electrical or biochemical response.

The fact that both chemistry and topology are involved in the architecture of the nervous system provides the system with an order of complexity beyond that of the man-made circuitry of computers. More than a dozen different substances are thought to be neurotransmitters, and they can be either excitatory (making the target cell more active) or inhibitory (preventing the target cell from becoming active). How do neurons in the developing nervous system "decide" which neurotransmitter to secrete in accordance with their location and function? And how is the synapse formed so that there is an appropriate match between the neurotransmitter secreted by the innervating neuron and the receptors present on the target cell? The investigation of these questions has involved the collaboration of electrophysiologists, biochemists and electron microscopists in several laboratories.

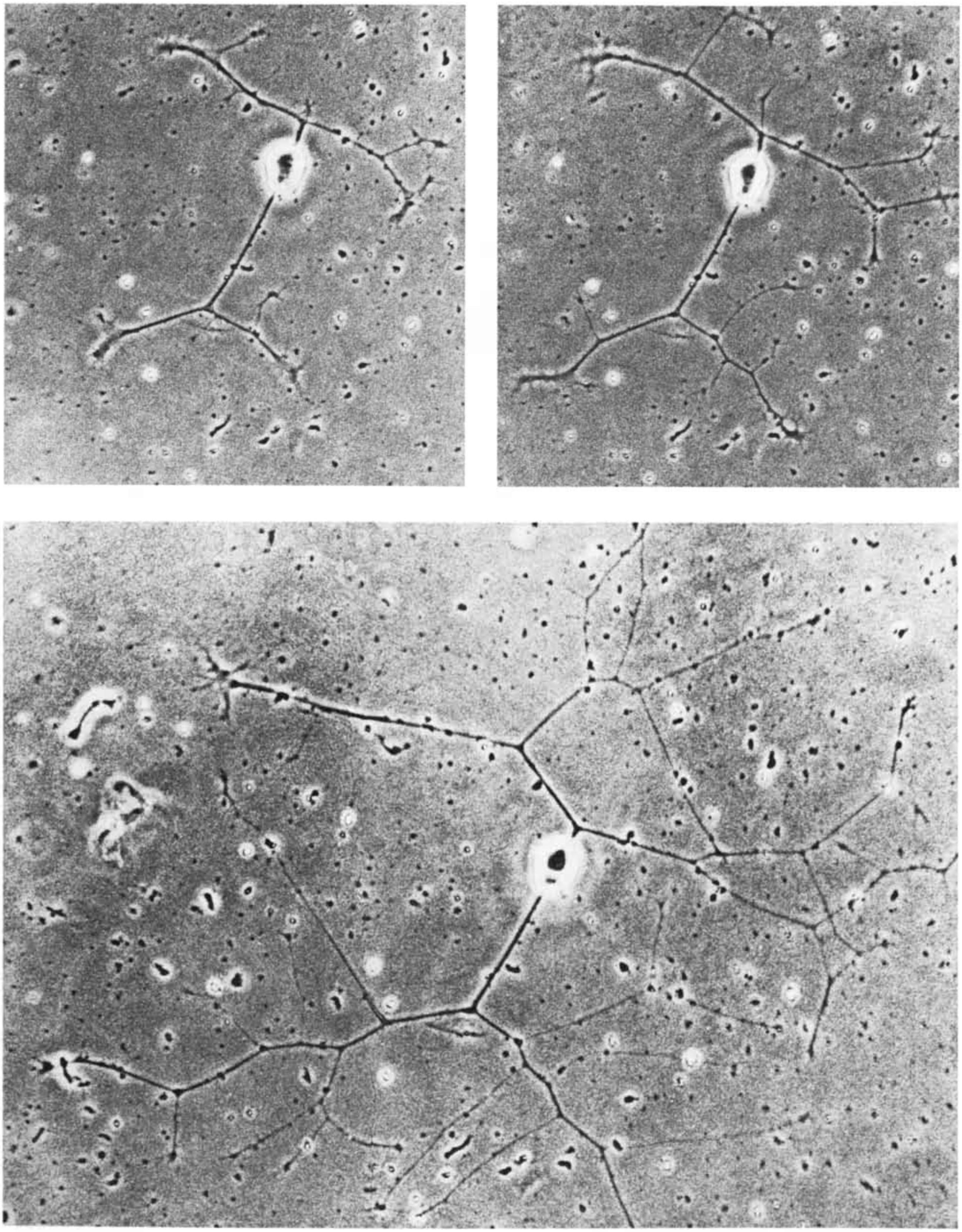
In our laboratory at the Harvard Medical School we and 14 colleagues have approached the general problem of how neurons differentiate chemically by examining the development of the autonomic nervous system, which regulates the activities of the circulatory system and a variety of organs and glands. The advantage of this system for such work is that its peripheral portions are readily accessible to the investigator. Moreover, the autonomic nervous system's chemical organization is relatively simple: most of its constituent nerve cells secrete either acetylcholine or norepinephrine (noradrenalin); they are respectively termed cholinergic or adrenergic. It seems reasonable to assume that the basic mechanisms involved in the decision of an autonomic neuron to become either cholinergic or adrenergic are analogous to those that determine which of several possible neurotransmitters will be secreted by a neuron in the brain.

There are two major types of pathway in the autonomic nervous system: sympathetic and parasympathetic. These two subsystems often have antagonistic actions on the same organ. For example, the muscle cells of the heart, which contract spontaneously and rhythmically, are innervated by both sympathetic and parasympathetic nerves. In situations of stress the sympathetic nerves secrete norepinephrine, which makes the heart beat faster and more strongly. In periods of calm or inactivity the parasympathetic nerves secrete acetylcholine, which makes the heart beat slower.

Both autonomic pathways innervating the heart begin with neurons in the brain or spinal cord. These central neurons do not project directly to the heart muscle but make excitatory cholinergic synapses on neurons in a ganglion, a way station outside the spinal cord; the ganglionic neurons in turn innervate the heart muscle. The sympathetic pathway therefore consists of a cholinergic central neuron and an adrenergic ganglionic neuron, both of which are excitatory; the parasympathetic pathway consists of a cholinergic central neuron and a cholinergic ganglionic neuron, the former excitatory and the latter inhibitory. (The ability of acetylcholine to excite at the first synapse and inhibit at the second synapse is due to differences in the chemistry of the target cells.) Clearly the normal control of the heart depends on the ability of the neurons in the two autonomic pathways to secrete the correct neurotransmitter. For example, if the ganglionic sympathetic neurons secreted acetylcholine instead of norepinephrine, the heart would be inappropriately inhibited.

What determines whether a given neuron in an autonomic pathway will become cholinergic or adrenergic? In investigating this question we and our colleagues have done most of our experiments on immature neurons obtained from sympathetic ganglia of newborn rats and grown in tissue culture. This preparation was chosen for three reasons. First, the adult sympathetic ganglion contains both adrenergic neurons and a smaller number of cholinergic neurons (which innervate certain blood vessels and sweat glands), so that the differentiation of the two types of neuron can be studied in the same system. Second, as has been demonstrated by Rita Levi-Montalcini and her colleagues at Washington University, the normal development of sympathetic neurons in the intact animal depends on the presence of a particular protein named nerve-growth factor (NGF). When NGF is added to the culture medium, sympathetic neurons can be grown in culture.

Dennis Bray, who is now at the Medi-

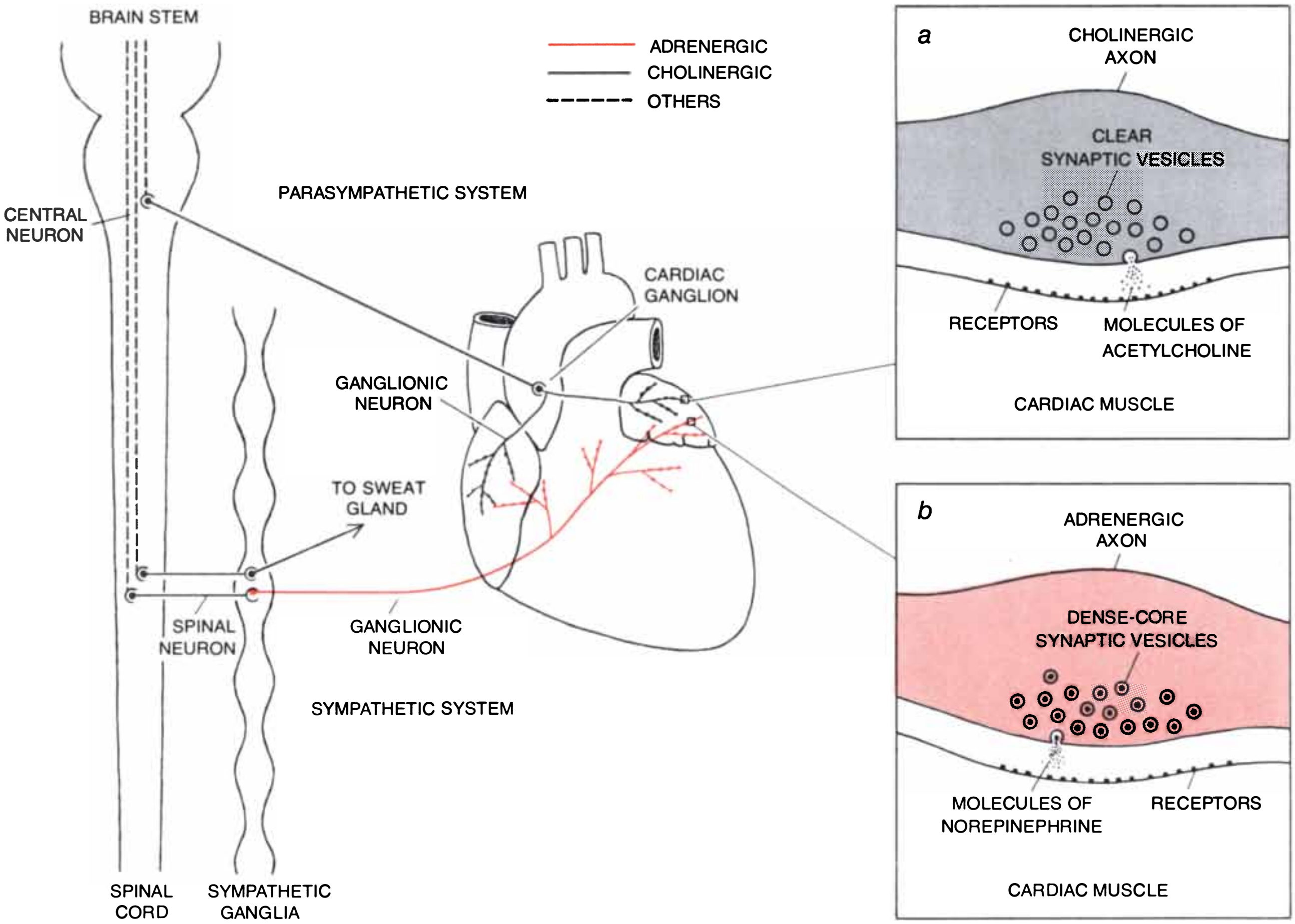


CULTURED NEURON, or nerve cell, obtained from the sympathetic ganglion of a newborn rat sprouts fine branching extensions when incubated in the presence of a specific protein called nerve-growth factor (NGF). The micrograph at the top left shows the cell after 240 minutes of incubation. The one at the top right shows it after 300 minutes, and the one at the bottom shows it after 500 minutes. The extensions grow continuously and eventually form a meshwork over the surface of the culture dish. They may also form synapses—sites of communication—with other neurons or with cultured muscle cells. The advantage of studying neurons in cell culture is that their environment can be carefully controlled. Micrographs were made by Dennis Bray of Medical Research Council Cell Biophysics Unit in London.

cal Research Council Cell Biophysics Unit in London, later discovered specific culture conditions in which sympathetic neurons of the newborn rat survive but the non-neuronal cells of the ganglion die off. In this instance the sympathetic neurons develop in the virtual absence of any other cell type. Third, when neurons are removed from the sympathetic ganglia of the newborn rat, they are still immature in many basic respects: most of them have not made synaptic connections and are at an early stage in the differentiation of neurotransmitter metabolism.

The advantage of studying the development of neurons in culture rather than in the intact animal is that the investigator has better control over the cellular and fluid environment of the neurons and can thereby more readily investigate the factors crucial for the induction or control of particular cellular functions. Of course, the cultured neurons must be supplied with nutrients that they would normally receive from the blood, such as amino acids, sugars and vitamins. The proper balance of oxygen and carbon dioxide, normally maintained by the lungs, is accomplished by gas exchange between the culture medium and the controlled environment of the culture incubator. The removal of toxic wastes, normally achieved by the kidney, is assumed by the periodic replacement of the culture medium. Because the isolated cells have no immunological defenses against microorganisms, the instruments, dishes and mediums must be sterile. Finally, in order to enhance the adherence of the axons and dendrites to the substrate, the plastic surface of the culture dish is coated with a thin film of collagen, a natural constituent of the intercellular matrix.

In the presence of NGF, but not in its absence, the sympathetic neurons survive and grow extensions that continue to elongate for many weeks. Eventually the surface of the dish becomes covered with a dense meshwork of axons and dendrites. Surprisingly the neurons that are grown in this highly artificial environment in the absence of non-neuronal cells acquire many of the characteristics of normal sympathetic neurons and do so on a time course similar to the one observed in the intact animal. The cultured neurons generate impulses, possess surface receptors for acetylcholine (the transmitter released by the central neurons that normally innervate them) and develop the ability to manufacture norepinephrine from its amino acid precursor, tyrosine. Indeed, most if not all of the neurons in sympathetic ganglia manufacture norepinephrine before the rat is born. These findings suggest that before the neurons were removed from the animal they received developmental signals that instructed them to utilize norepinephrine as a transmitter. In the presence of NGF, which allows the cells to survive in culture, they simply read out the appropriate genetic program.



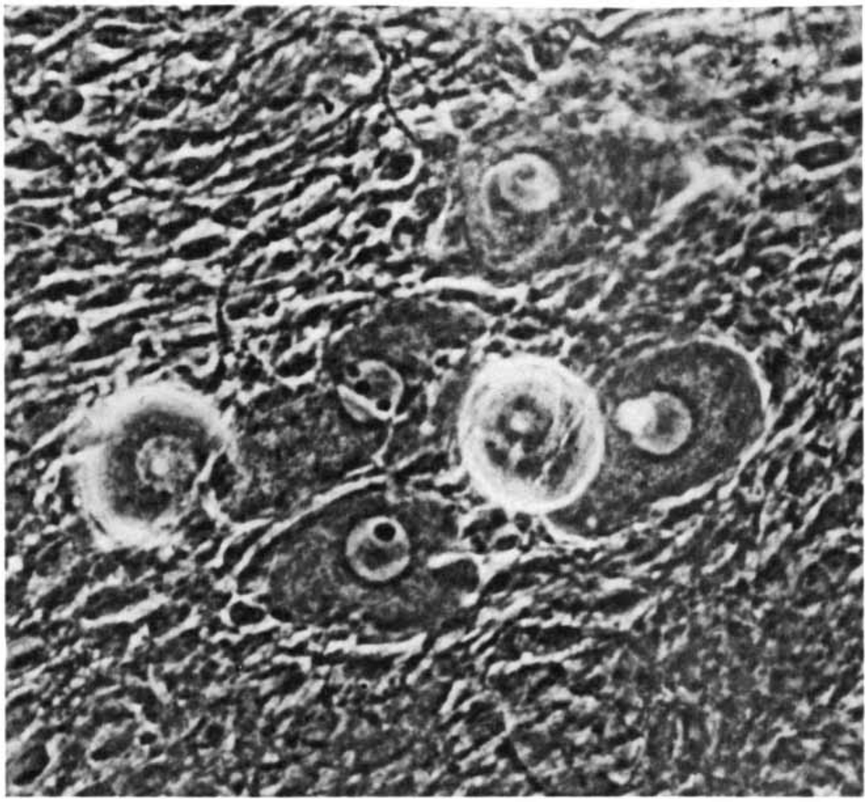
TWO AUTONOMIC PATHWAYS with antagonistic actions control the frequency of the heart beat. When the body is under stress and the heart beats faster, the sympathetic system is active. Reflex pathways excite sympathetic neurons in the spinal cord; these spinal neurons project to neurons in a ganglion, a way station outside the spinal cord. When the ganglionic neurons are excited by the spinal input, they make the heart beat faster by secreting norepinephrine onto the cardiac muscle cells. In contrast, when the body is relaxed and the heart beats more slowly, the parasympathetic system is active: reflex pathways excite parasympathetic neurons in the medulla, which project to target neurons in the cardiac ganglion of the heart. The ganglionic neurons in turn make the heart beat more slowly by secreting acetylcholine onto the cardiac muscle cells. The secretion of neurotransmitter takes place at varicosities: small swellings at intervals along the norepinephrine-secreting (adrenergic) and acetylcholine-secreting (cholinergic) axons, as shown in insets at far right.

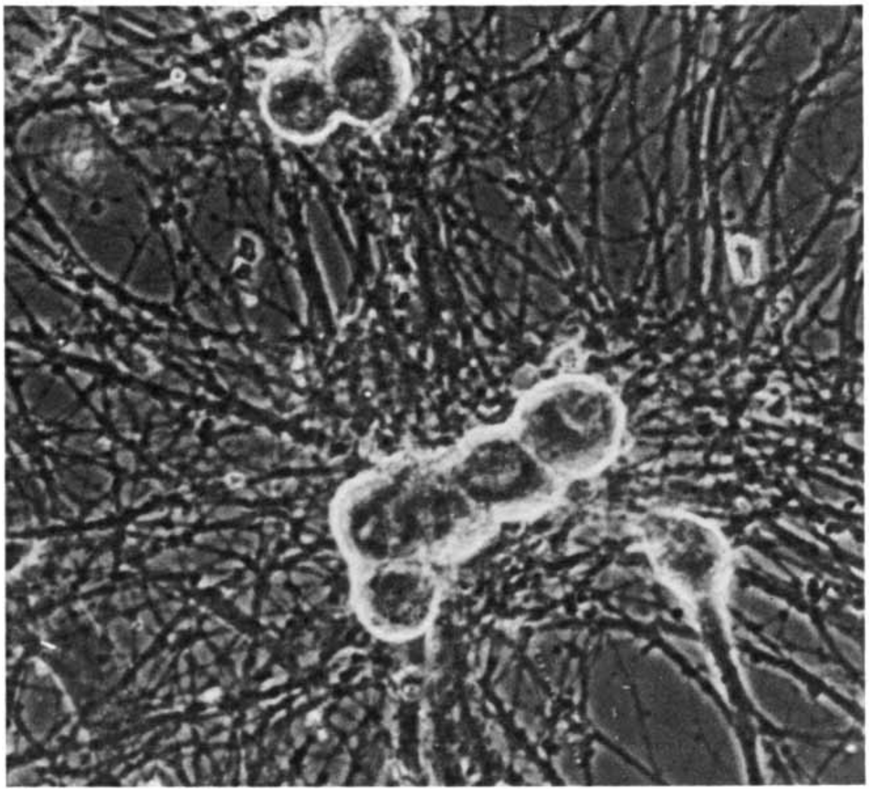
In the absence of non-neuronal cells the cultured sympathetic neurons also form synapses with one another. As has been shown by Richard P. Bunge and his colleagues at the Washington University School of Medicine, these synapses structurally resemble those made by adult adrenergic neurons in that the norepinephrine is stored in tiny vesicles, or sacs, that have dense cores when they are stained with permanganate for electron microscopy. Although norepinephrine is released and taken up at these synapses, the neurotransmitter does not have any detectable effect on the electrical activity of the target neuron, meaning that either the cultured neurons do not possess receptors for norepinephrine or, if receptors are present, they fail to evoke a significant electrical signal. The existence of such electrically silent synapses is quite unusual, and few other instances are known.

When the sympathetic neurons are cultured together with non-neuronal cells from the ganglion, striking differences in the chemical differentiation of the neurons are observed. Instead of expressing only adrenergic functions, as they do in the absence of the non-neuronal cells, the mixed cultures manufacture and accumulate high levels of acetylcholine, indicating that many of the neurons have become cholinergic. Electrophysiological investigation has revealed that a substantial fraction of the neurons also form cholinergic synapses with their neighbors, thereby substituting for the normal cholinergic input from the central neurons, which are absent from the culture. Unlike the electrically silent adrenergic synapses formed in the absence of non-neuronal cells, these cholinergic synapses are electrically active and excitatory.

Is the effect of the non-neuronal cells on the choice of transmitter exerted by way of direct contact with the sympathetic neurons, or rather by the release of some chemical factor into the culture medium? To answer this question we grew the neurons in one set of dishes and the non-neuronal cells in another set. Every two days the medium that had been "conditioned" by incubating it with the non-neuronal cells was transferred to the neuronal cultures. We found that treatment with conditioned medium was sufficient to induce cholinergic properties in the sympathetic neurons: the cells synthesized and accumulated acetylcholine and secreted it at functional cholinergic synapses with one another.

The effect of conditioned medium was dose-dependent: the greater the amount that was added to the growth medium of the neurons, the more acetylcholine the cells made and the more likely they were to form cholinergic synapses with one another. At the same

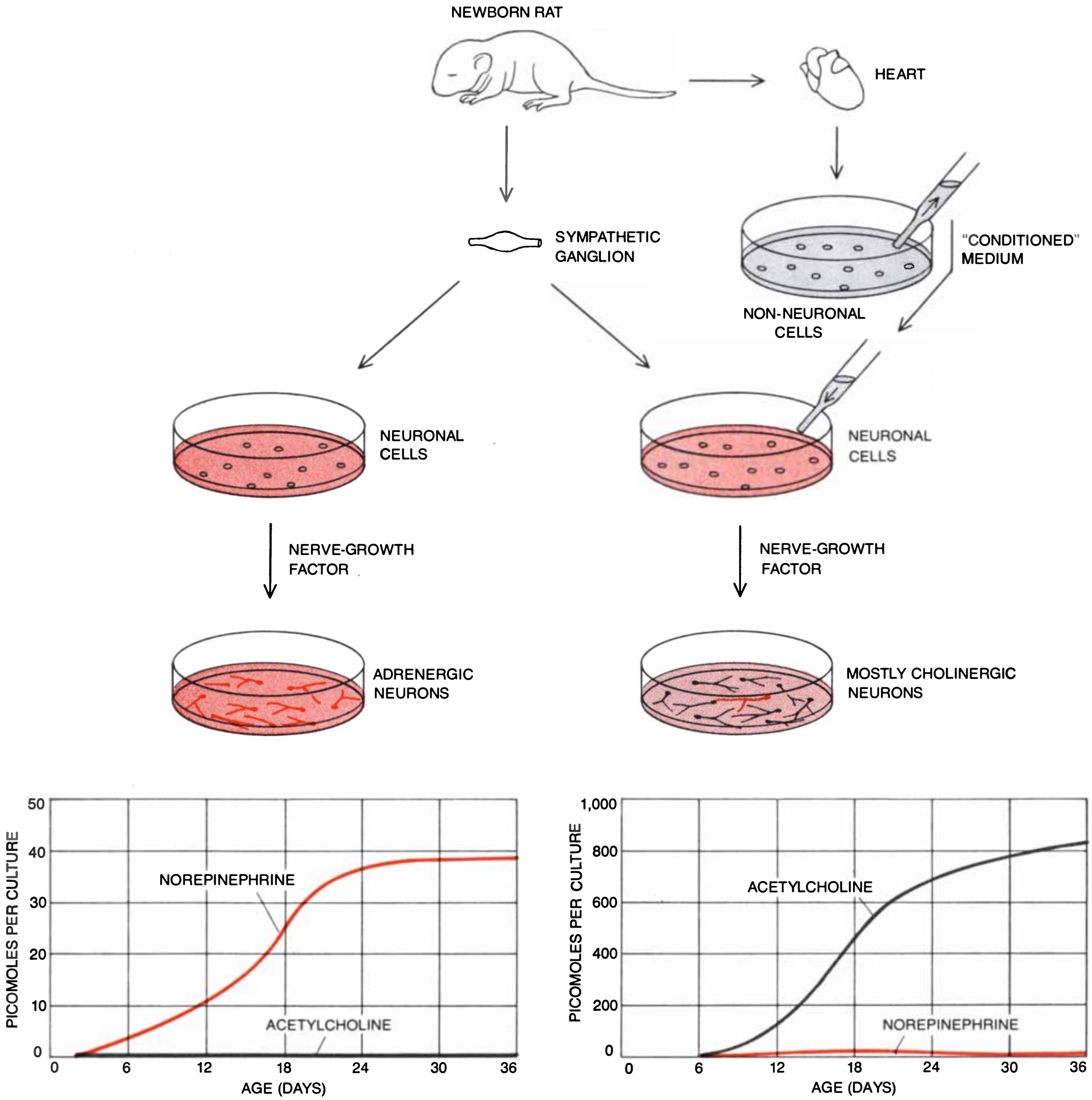




IMMATURE NEURONS obtained from sympathetic ganglia of a newborn rat can be cultured in tissue-culture dishes, as is shown in these micrographs of living cultures made with the phase-contrast microscope. In standard medium the non-neuronal supporting cells of the ganglion also survive and divide, eventually forming a continuous layer around and beneath the neurons (*top*). Cultures that contain only neurons (*bottom*) can be prepared by growing the ganglionic cells in a medium in which the neurons survive but the non-neuronal cells do not. One such medium contains an analogue of a nucleic acid precursor, cytosine arabinoside, which is toxic when it is incorporated into the nucleic acid molecules of the dividing non-neuronal cells but does not affect the neurons, which do not divide but merely grow larger.

time the neurons made less norepinephrine and fewer adrenergic synapses. Thus in the population of neurons as a whole the expression of cholinergic functions was roughly reciprocal to the expression of adrenergic functions. Moreover, the differentiation of cholinergic metabolism in the presence of conditioned medium had a time course similar to that of adrenergic metabolism in the absence of conditioned medium. We have recently obtained evidence that the active ingredient of conditioned medium is a large molecule, and attempts to purify it are now under way.

How does this developmental factor released by non-neuronal cells influence the choice of neurotransmitter? One possibility is that at the outset there are two populations of neurons, one predestined to become adrenergic and the other predestined to become cholinergic; in that case the conditioning factor would enhance the survival and growth of the cholinergic population while diminishing that of the adrenergic population. This hypothesis is unlikely, how-

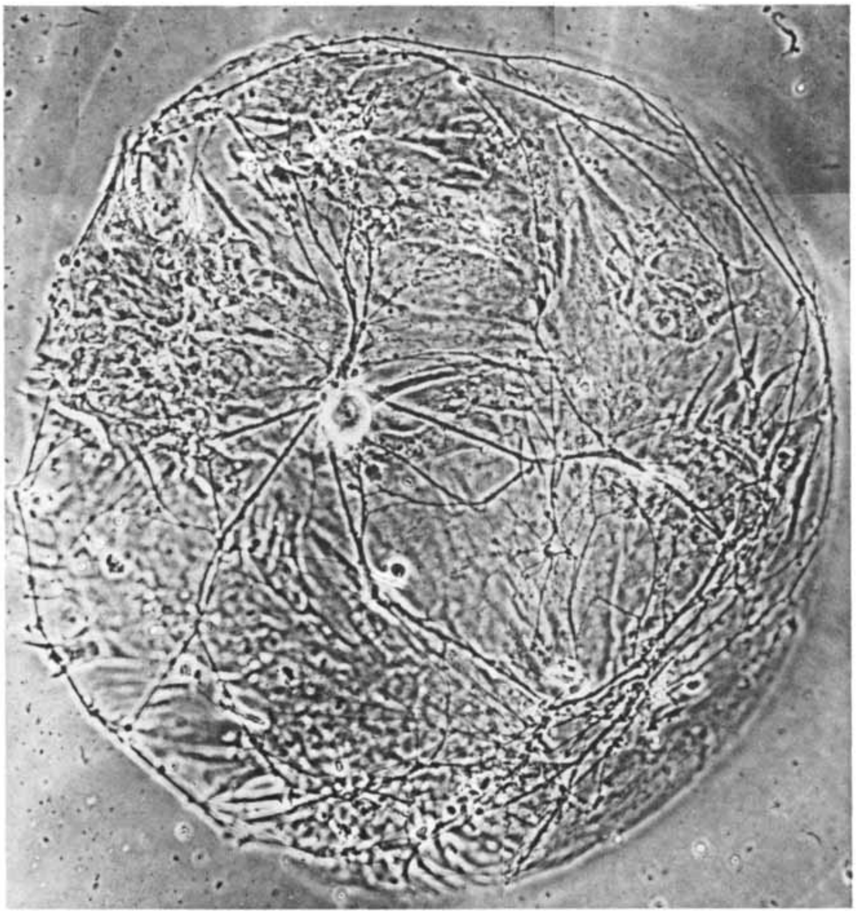


EFFECT OF CULTURE ENVIRONMENT on the chemical differentiation of sympathetic neurons is striking. If immature neurons are obtained from the sympathetic ganglia of newborn rats and grown in pure cultures, nearly all of them manufacture and secrete norepinephrine. If, however, the immature sympathetic neurons are cultured together with non-neuronal cells, or if they are treated with culture medium that has been "conditioned" by incubating it with non-neuronal cells, a large majority of the neurons will manufacture and secrete acetylcholine. The graphs at the bottom show the expression of cholinergic and adrenergic function in the presence and absence of conditioned medium. The results are expressed as picomoles (10^{-12} mole) of transmitter synthesized and accumulated in the neuronal culture during the four-hour incubation period. In the absence of conditioned medium (*left*) the ability to make norepinephrine rises over a three-week time course and the ability to make acetylcholine remains negligible. Conversely, in sister cultures grown in 62 percent conditioned medium (*right*), the ability to make norepinephrine rises at first but then declines to a low level as ability to make acetylcholine rises. Thus in the presence of high concentrations of conditioned medium the expression of the two neurotransmitters is reciprocal.

ever, since conditioned medium has no effect either on the survival or growth of cultured neurons. A second possibility is that the conditioned medium induces the expression of cholinergic properties in a predetermined cholinergic population that otherwise survives but does not express any transmitter metabolism. This hypothesis was also ruled out by the finding that in the absence of conditioned medium virtually no neurons were "silent" with respect to transmitter synthesis.

A third explanation for the effect of conditioned medium is the most consistent with our observations. According to this hypothesis the neurons that express adrenergic properties at the outset are still "plastic" with respect to neurotransmitter choice for a considerable period after birth and can become cholinergic under the influence of conditioned medium. This concept implies that the active ingredient of conditioned medium determines the choice of transmitter and the type of synapses made by a sympathetic neuron without affecting whether the neuron survives or how extensive its axon and dendrites are.

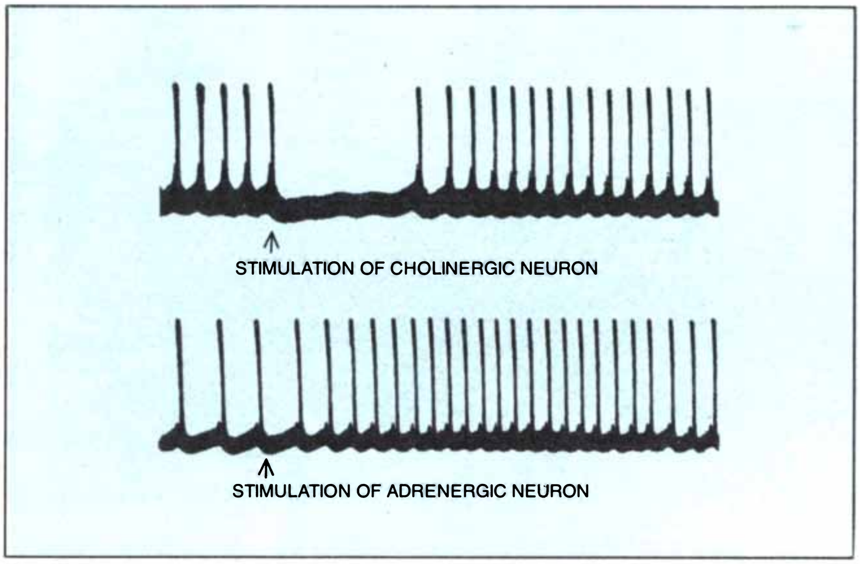
The reciprocity in the expression of adrenergic and cholinergic functions can be ascribed to a "competition" of some kind within each neuron between the prenatal instruction to become adrenergic and the new instruction to become cholinergic. If such a competition exists, one may ask if an individual neuron can express both transmitter systems simultaneously, at least for a short period. We attempted to answer the question by growing single neurons on small beds of heart-muscle cells for about two weeks. The neuron and an adjacent heart cell in such microcultures were impaled with microelectrodes to monitor their electrical activity; in that way the transmitter choice of the neuron could be assayed by triggering the release of the transmitter and observing its effect on the heart cells, which "beat" spontaneously and rhythmically in culture. Unlike the cultured sympathetic neurons, the heart cells possess both cholinergic receptors and electrically active adrenergic receptors; thus a slowing or stopping of the beating indicates the secretion of acetylcholine, whereas an increased frequency of the beating indicates the secretion of norepinephrine. Further evidence for the identity of the transmitter can be obtained by observing the effects of certain drugs that compete specifically with the natural transmitter for binding to the receptors on the heart cells. For example, atropine blocks cardiac acetylcholine receptors, whereas propranolol blocks cardiac norepinephrine receptors.



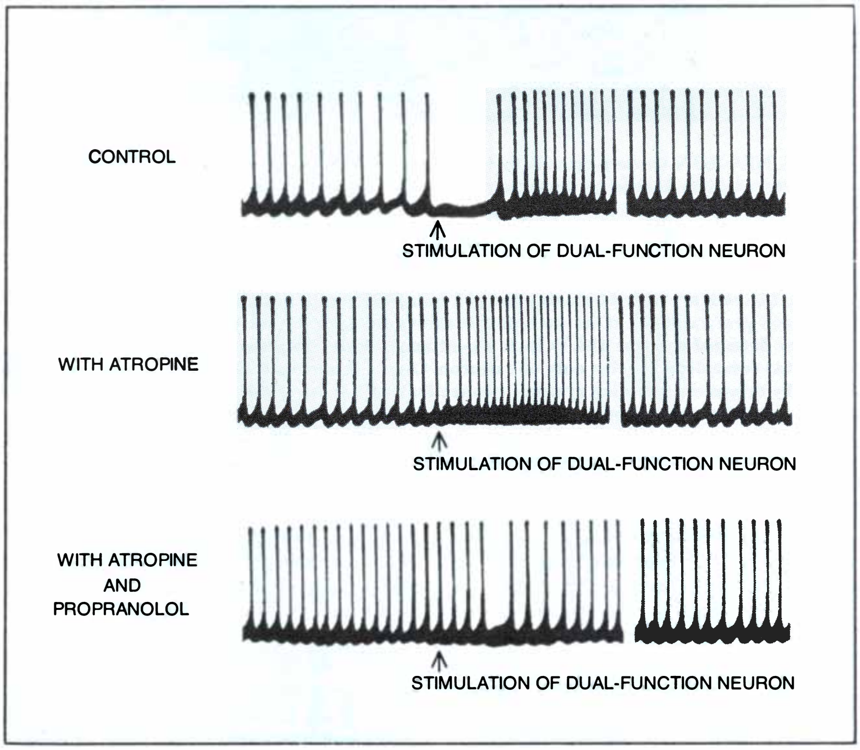
SINGLE NEURON (*center*) extends its branches over a background layer of heart cells in a microculture system devised by the authors and their colleagues at the Harvard Medical School. The preparation is useful for studying the differentiation of sympathetic neurons with respect to neurotransmitter choice: the transmitter secreted by the neuron can be identified by its effect on the cardiac muscle cells, which contract spontaneously and rhythmically in culture. The first step in preparing the microculture is to make a suspension of heart cells from the newborn rat. The cells are plated onto tiny disks of collagen about .5 millimeter in diameter to which they attach preferentially. (One such disk is shown here.) A day or two later, after the heart cells have flattened over the disk to form a continuous layer, the neurons are plated at low density so that some disks receive only one neuron. Once the neuron is attached, it sends out long threadlike extensions, forming synaptic connections with some of the cardiac muscle cells. The neurotransmitter secreted at these synapses is either acetylcholine, which slows the beating of muscle cells, or norepinephrine, which accelerates it. Some of the neurons secrete acetylcholine, others secrete norepinephrine and at a certain stage in differentiation some secrete both.

Working with this technique we have identified three types of neurons in the two-week-old microcultures. The first type is adrenergic: it excites the heart cells and the effect is blocked by propranolol. The second type is cholinergic: it inhibits the heart cells and the effect is blocked by atropine. The third type of neuron exhibits both cholinergic and adrenergic activity: stimulation of the neuron inhibits the heart with an atropine-sensitive mechanism and then speeds up the heart with a propranolol-sensitive mechanism. Since only one neuron is present in each microculture, it is clear that the same cell mediates both effects. When the microcultures are examined with the electron microscope, numerous dense-core synaptic vesicles are seen in the adrenergic neurons, clear vesicles are seen in the cholinergic neurons and a few dense-core vesicles combined with a large majority of clear vesicles are seen in the dual-function cells. An obvious advantage of studying cultures that contain only one neuron is that it is possible to make an unambiguous correlation between the structure and function of the cell.

These findings establish that a single neuron can express both transmitter systems simultaneously at an immature stage. Dual function may seem to be a novel concept, but developing neurons have not previously been investigated in ways that might reveal this behavior. In retrospect dual function is a logical intermediate step in the conversion of an adrenergic neuron to a cholinergic neuron under the influence of conditioned medium. Even if there is no temporal overlap between the synthesis of enzymes and other components involved in manufacture and release of the two neurotransmitters, it is reasonable to assume that the enzymes and synaptic ves-



FREQUENCY OF CONTRACTION of the cardiac muscle cells in a microculture containing a single neuron reveals the identity of the neurotransmitter secreted by that neuron. In a typical experiment one microelectrode inserted into the neuron stimulates the release of transmitter as a second microelectrode inserted into an innervated muscle cell monitors the cell's contractions. Trace at the top shows the effect of stimulating a cholinergic neuron: the spontaneous contractions of the cardiac muscle cell cease temporarily. Bottom trace shows effect of stimulating an adrenergic neuron: the contraction frequency of the innervated muscle cell increases.



DUAL-FUNCTION NEURON present in a 13-day-old microculture displays a combination of cholinergic and adrenergic properties. In the trace at the top stimulation of the neuron causes an initial cessation of the beating of the innervated cardiac muscle cells (mediated by acetylcholine) followed by a resumption of beating at an enhanced rate (mediated by norepinephrine). Both transmitters appear to be secreted at once; the temporal separation of inhibition and excitation may be due to the fact that the effects of norepinephrine on the heart cells come on more slowly and persist longer than those of acetylcholine. In the lower two traces drugs that specifically block either cholinergic or adrenergic transmission have been added to culture medium to verify dual function. Middle trace shows effect of atropine, which competes with acetylcholine for its receptor: the inhibition is removed but the excitation remains intact. Bottom trace shows effect of adding propranolol, which competes with norepinephrine for its receptor, to the atropine solution: both the excitatory and inhibitory effects of the neuron are blocked.

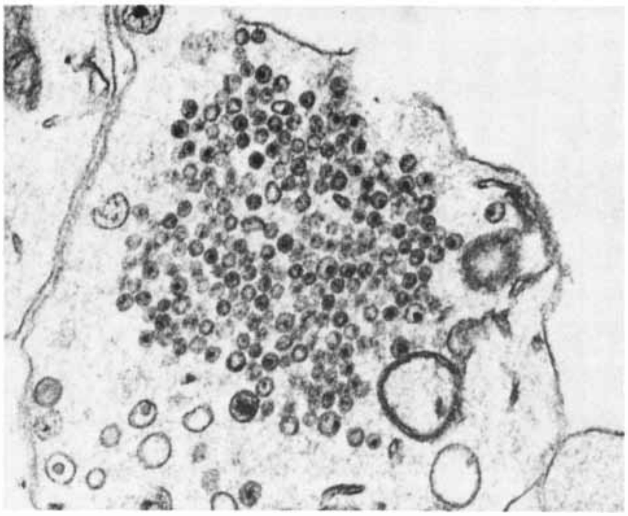
icles involved in adrenergic transmission would continue to function for a while after their synthesis had ceased. The precise duration of the dual-function state is not yet known.

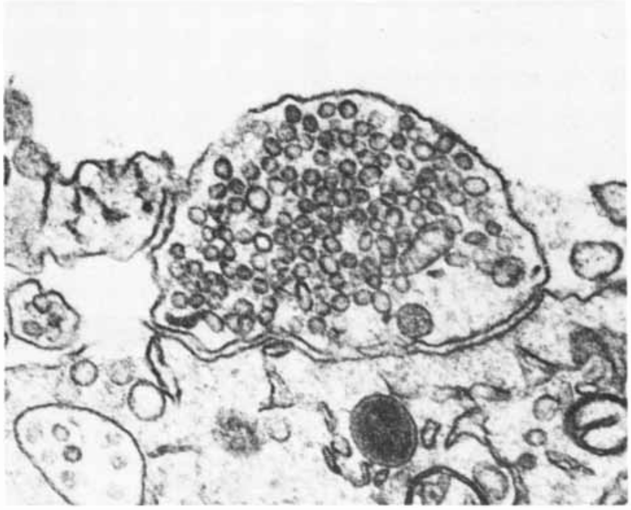
After about four weeks the single neurons in microculture have grown so large that biochemical assays can be made on them. The cells are incubated in a mixture of radioactive tyrosine and choline for eight to 12 hours; then the amount of norepinephrine and acetylcholine that have been synthesized from these precursor molecules is determined. In the absence of conditioned medium virtually all the neurons make detectable quantities only of norepinephrine. In the presence of heart cells, however, a substantial majority of the neurons make only acetylcholine. Under no circumstances is there a significant number of "silent" or dual-function neurons after four weeks in culture.

These findings indicate that most neurons are adrenergic at the time they are put into culture but are susceptible to a "flip-flop" control mechanism that determines their ultimate choice of transmitter. The duration of the transition period during which dual function may be expressed is not known, but by four or five weeks after birth virtually all the neurons have differentiated into a state in which only one transmitter system is expressed to a significant degree. We are currently following the status of single neurons over a period of time by intermittently testing their effects on heart cells. In this way we hope to observe the transition from adrenergic to cholinergic behavior directly and to determine its time course.

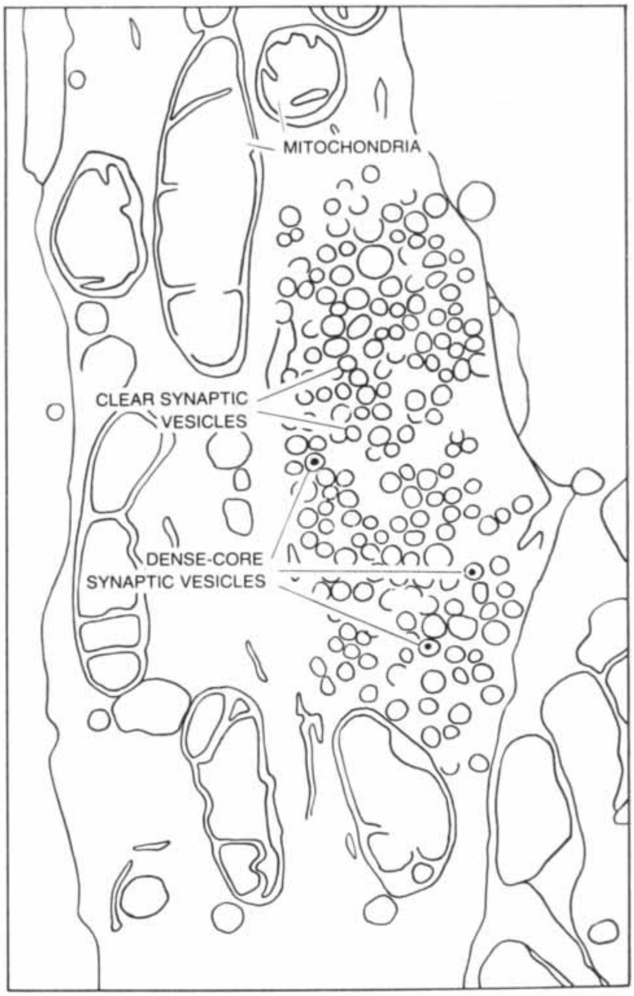
When in the life of the culture are the neurons most sensitive to the action of conditioned medium, and how reversible are its effects? To answer these questions conditioned medium was added to cultures of sympathetic neurons for a 10-day "pulse" at various stages of cell maturity. We found that the responsiveness of the cells to the pulse of conditioned medium declined rapidly with the increasing age of the cells, reaching a very low level when the pulse was applied between days 40 and 50. Caryl Hill and Ian Hendry of the Australian National University and C. David Ross, Mary Johnson and Bunge of the Washington University School of Medicine obtained similar results in intact animals. When small pieces of sympathetic ganglia were taken from rats of various ages and placed in culture for a standard period, the expression of cholinergic function declined progressively with increasing age of the animal from which the explant was taken, reaching a minimum in the ganglia of adult rats.

One of the most intriguing aspects of the influence of non-neuronal cells on the choice of neurotransmitter is the dis-

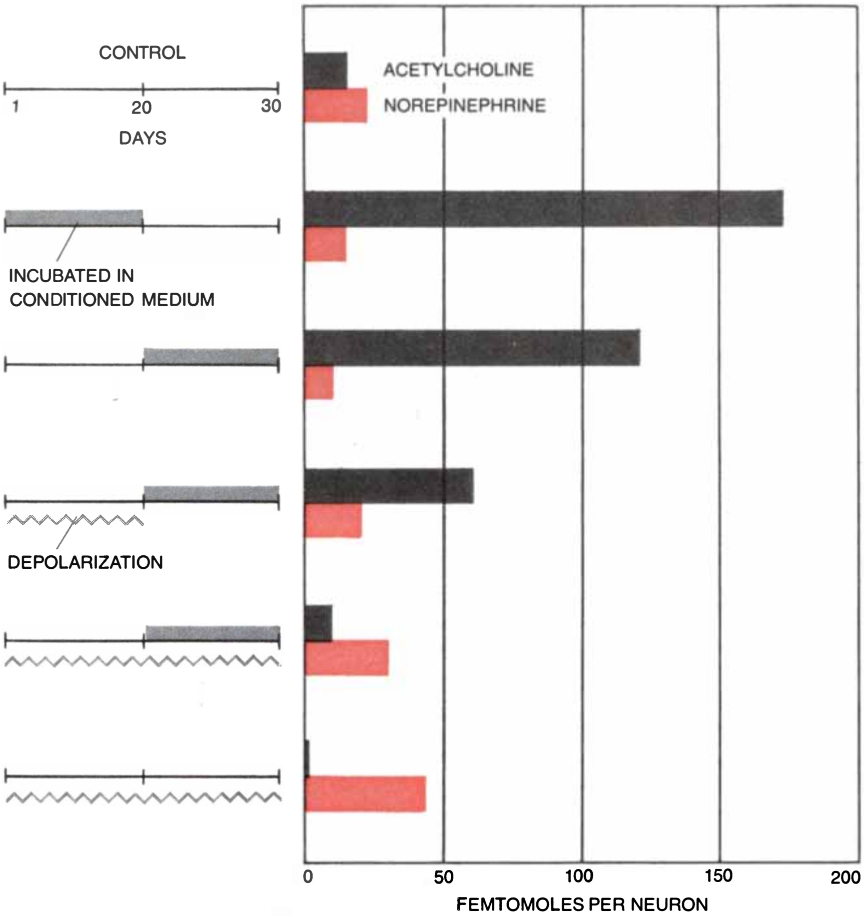




SYNAPTIC VARICOSITIES along the axons of cultured cholinergic or adrenergic neurons display characteristic differences in their fine structure. The molecules of neurotransmitter are stored in tiny sacs called synaptic vesicles. When the varicosities of adrenergic neurons are stained with permanganate and examined in the electron microscope, they contain synaptic vesicles with small, electron-dense cores (*left*). The varicosities of cholinergic neurons, on the other hand, have vesicles with clear interiors (*right*). These two electron micrographs and the one below, which enlarge the structures 44,000 diameters, were made by Story C. Landis at the Harvard Medical School.



DUAL-FUNCTION VARICOSITY contains a few dense-core vesicles characteristic of an adrenergic neuron (*see map*) together with a large majority of clear vesicles characteristic of a cholinergic neuron. Thus the fine structure of the varicosity is correlated with its combined inhibitory and excitatory effects. The dual-function state is thought to be a transient phase during the conversion of an adrenergic neuron to a cholinergic one under influence of conditioned medium. The precise duration of dual function, however, is not yet known.



EFFECT OF ELECTRICAL ACTIVITY on the chemical differentiation of sympathetic neurons is to reduce their ability to become cholinergic when they are exposed to conditioned medium. In the intact ganglion activity results from excitatory innervation from spinal neurons. The excitatory influence of the spinal neurons can be imitated in cell culture by treating the sympathetic neurons with high concentrations of potassium ion, which depolarizes them (reverses the voltage across their outer membrane). In the experiment shown here, potassium or conditioned medium was added on day 0 or day 20, and each transmitter manufactured by the neurons was assayed on day 30. (One femtomole equals 10^{-15} mole.) Depolarization of the cells with potassium largely blocked the effect of conditioned medium. Therefore the minority of neurons in the intact sympathetic ganglion that end up cholinergic may not become electrically active until they have undergone cholinergic induction by the non-neuronal cells.

tribution of effectiveness among the various non-neuronal cells of the body. If this property were properly localized, it might act as a specific determinant of transmitter choice. To investigate the matter we studied the effect of various tissues from the newborn rat on transmitter choice. All the non-neuronal cells were able to induce cholinergic function, but there were clear quantitative differences in effectiveness: cells from skeletal muscle were most effective, cells from the heart were intermediate and cells from the liver were least effective. A conspicuous feature of this order was its relation to the amount of cholinergic innervation received by the target tissue: skeletal muscle fibers receive only cholinergic input, heart-muscle cells a mixture of cholinergic and adrenergic input, and liver cells, when they are innervated at all, only adrenergic input. This finding suggested that during development target cells release chemical factors that influence the transmitter metabolism of the neurons that innervate them.

Other observations did not, however, appear to be encompassed by this simple hypothesis. Several types of cells that are effective medium conditioners are not normally innervated, in the sense of receiving synapses. Among these cells are fibroblasts (connective-tissue cells), a line of glial tumor cells serially propagated in culture (glial cells are the supporting cells of neurons in the brain and spinal cord) and the non-neuronal cells in sympathetic ganglia. The ability of ganglionic non-neuronal cells to induce cholinergic differentiation seemed particularly paradoxical. During the first three weeks after birth the non-neuronal cells in the ganglion surround each neuron closely, yet most of the neurons in the intact ganglia end up adrenergic. Why is it that non-neuronal cells or conditioned medium can cause almost all sympathetic neurons to become cholinergic in culture, yet in the body only a few sympathetic neurons become cholinergic even in the presence of an abundance of non-neuronal cells?

This apparent paradox indicated that additional variables were involved in the choice of transmitter. The sympathetic neurons in the body were exposed to the cholinergic signal provided by the non-neuronal cells, yet most of them were prevented from responding to the signal by some factor that was missing from the culture medium. One possible candidate for this factor was NGF, since it is critical for the survival and growth of adrenergic sympathetic neurons both in the intact animal and in culture. Moreover, when neuronal cultures are exposed to high concentrations of NGF, there is an increase in the manufacture of specifically adrenergic components, such as norepinephrine, relative to the manufacture of nonspecific cell components such as protein and lipid. It turns out, however, that NGF has the same potentiating effect on cholinergic differentiation. Thus with respect to transmitter production NGF is permissive rather than instructive: it stimulates the growth and differentiation of immature sympathetic neurons along either the adrenergic path or the cholinergic one but does not influence which path is taken. In contrast to NGF the cholinergic factor secreted by non-neuronal cells does not affect neuronal survival or growth but does instruct the neurons with respect to neurotransmitter choice.

Another candidate for the factor in the intact animal that prevented the majority of sympathetic neurons from becoming cholinergic was the normal excitatory input from the central neurons, which were absent from our cultures. Ira Black, Hendry and Leslie Iversen at the Medical Research Council Laboratory of Molecular Biology in Cambridge, England, had found that if the spinal input to the sympathetic ganglia of young mice is cut, further development of the adrenergic metabolism is reduced. These observations raised the possibility that electrical activity imposed on the ganglionic neurons during the first week after birth plays a role in determining the choice of transmitter. Although the innervating central neurons are not present in the cell culture, we were able to mimic their excitatory effect on the ganglionic neurons (in which the membrane of the neuron is depolarized, that is, the voltage across the membrane is reversed). This was done by raising the potassium concentration of the medium, by adding the drug veratridine, which causes an influx of sodium ions into the neurons, or by stimulating the cells electrically once per second for many days.

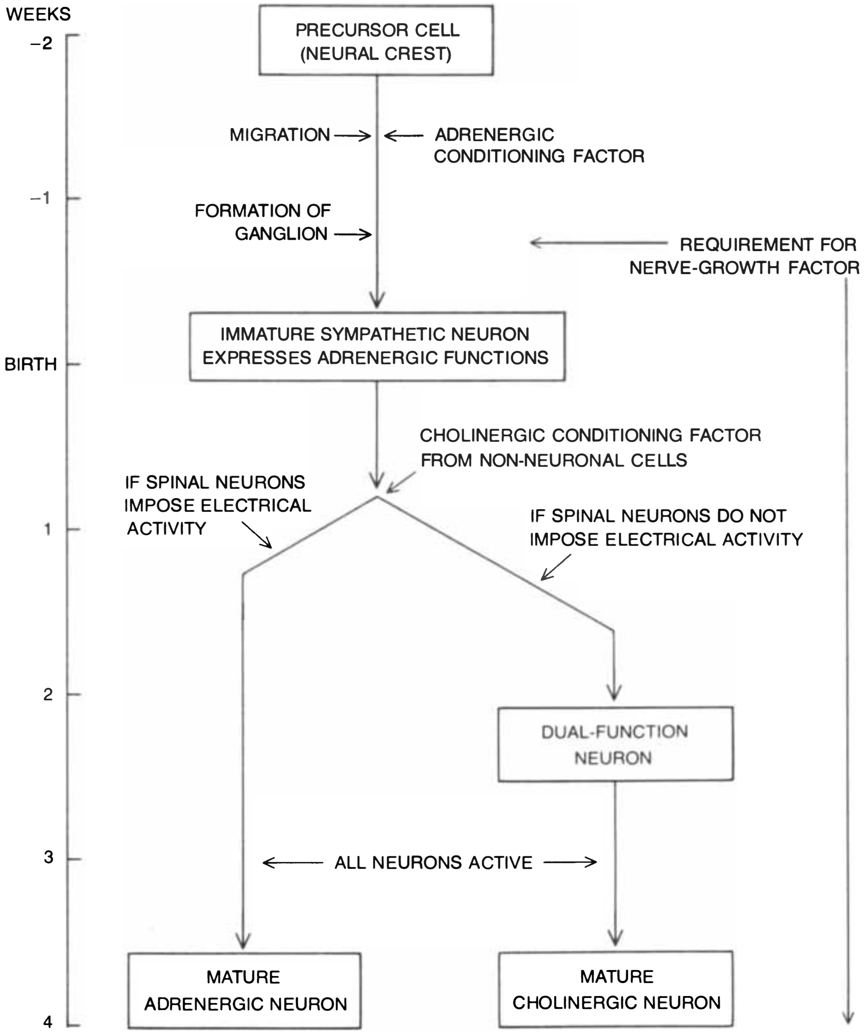
When mass neuronal cultures were depolarized either in the presence of conditioned medium or for seven to 10 days before the addition of conditioned medium, the neurons remained primarily adrenergic. Indeed, depolarization depressed the ratio of acetylcholine synthesis to norepinephrine synthesis as much as 300-fold compared with cultures that simply received conditioned medium. These changes occurred without a significant alteration in the survival of neurons, suggesting that the neurons, which would have become cholinergic in response to conditioned medium, now remained adrenergic. It was as if the depolarization of the sympathetic neurons stabilized their prenatal instruction to become adrenergic and curtailed their plasticity with respect to transmitter choice.

Electrical activity is accompanied by the entry of calcium ions into neurons. Because the entry of calcium ions is important in controlling the secretion of neurotransmitters and probably other cellular functions, the question arose whether the effects of depolarization on transmitter choice could be diminished by preventing the influx of calcium. This was done by raising the concentration of magnesium ions in the medium or by adding the drug D600, which selectively blocks the movement of calcium ions across the cell membrane. Indeed, when calcium entry was blocked, depolarization no longer prevented the induction of cholinergic properties by conditioned medium. Thus the effect of depolarization in blocking the effect of conditioned medium on transmitter choice appears to be mediated through calcium ions, although the precise mechanism is not yet understood.

These observations raise the possibility that in the intact ganglia the majority of neurons are preserved in their prenatal adrenergic condition by electrical input from the spinal cord that begins during the first week after birth. A corollary hypothesis is that the neurons in the ganglia that are destined to become cholinergic (those that innervate certain blood vessels and sweat glands) acquire their electrical input only after they have been influenced by the non-neuronal cells. Inherent in such a mechanism is an interesting possibility. Perhaps the selective formation of synapses between the central neurons and the ganglionic neurons not only establishes the circuitry of the autonomic pathways but also determines the choice of transmitter that is appropriate to the circuitry. This concept emphasizes the potential importance of neuronal activity in the chemical differentiation of the nervous system.

The experiments described in this article leave little doubt that developing sympathetic neurons are at least transiently plastic with respect to their choice of transmitter. A similar conclusion has recently been reached by Nicole Le Douarin and her colleagues at the Institute of Embryology at Nogent-sur-Marne in France, who are doing ingenious transplantation experiments in bird embryos. They find that the decision of immature neurons to become adrenergic or cholinergic can be altered if the cells are transplanted to new sites in the embryo. For example, a precursor population that normally gives rise to cholinergic neurons will yield adrenergic neurons if it is transplanted to a region where adrenergic neurons normally arise. In sum, both in culture and in intact embryos, this important neuronal decision is not entirely preprogrammed but can be influenced by other cells, including neurons and non-neuronal cells.

We have discussed here the control of just one decision in the development of a single class of nerve cells. The development of the intricate synaptic network of the adult nervous system presumably involves almost innumerable decisions, many of which also depend on cellular interactions. Disentangling these interactions is likely to require a combination of research approaches, including the study of development in the embryo and in simplified and controlled culture environments. Ultimately one would like to be able to specify everything that must be added to a culture dish to cause an embryonic neuron to display its entire range of behavior from the time cell division ceases to senescence.



FLOW CHART represents a hypothesis about the developmental steps involved in the chemical differentiation of a sympathetic neuron. According to this hypothesis the environmental influences that affect the cell's choice of neurotransmitter are exerted both by non-neuronal cells (which release developmental factors) and by other neurons (whose electrical activity modifies the neuron's response to the factors). Time axis is approximate. Since development of the neurons is not synchronized, they would traverse pathway at somewhat different times.

The Highest-Energy Cosmic Rays

A small number of fast particles from space are more energetic than the particles produced by the largest man-made accelerators. They are revealed by the showers of other particles they create in the air

by John Linsley

A steady rain of particles moving at nearly the speed of light falls on the earth at all times and from all directions. These particles, the cosmic rays, have attracted the attention of investigators for more than six decades. If one were to study cosmic rays on the surface of the moon, where an atmosphere and the influence of the earth's magnetic field are absent, one would find that they are primarily nuclei of the more abundant elements, those ranging in atomic weight from the lightest, hydrogen, to the group that includes iron (atomic weight 56). In addition there would be a few nuclei of heavier elements, some electrons and positrons, a few gamma rays and the ever-present but elusive neutrinos. These otherwise familiar particles are given the name cosmic rays because they are found shooting through outer space at nearly the speed of light, as fast as the particles in the beam of a synchrotron or other man-made accelerators. Indeed, some cosmic-ray particles have an energy far higher than that of the fast particles generated by the largest accelerating machines yet built. It is these highest-energy cosmic rays that are the most puzzling. Where do they come from? How are they accelerated to such energies?

In spite of the fact that the highest-energy cosmic rays are subatomic particles, the smallest objects known to exist, their kinetic energies are so great that they can be equated with objects in everyday life. The most energetic cosmic rays observed so far have had energies of two or three times 10^{20} electron volts, which equals the kinetic energy of a very well hit tennis ball.

Cosmic-ray research in the region of highest energies, say upward of 10^{15} electron volts, has a special character because in trying to answer questions that belong to astrophysics those doing the research continually encounter unanswered questions that belong to a different branch of science now called high-energy physics. This used to be the state of affairs in cosmic-ray research at energies lower than 10^{15} electron volts. It took a long time, following the 1912 balloon flight that carried Victor F. Hess to his discovery of cosmic radiation, to find an answer to the basic question: What are cosmic rays?

At first physicists tried to find the answer by using the same techniques they were accustomed to using for identifying the fast particles emitted by radioactive materials or those generated by the early "atom smashers." They encountered difficulties. They found, for example, that when cosmic rays entered the atmosphere they practically never came quietly to rest, which is what they would have done if their initial energies had been as low as those of the fast particles the investigators were accustomed to identifying. They found instead when they followed the paths of what they came to call primary cosmic rays that as the particles penetrated matter nearly all of the paths led to the site of a violent collision between the cosmic-ray particle and an atomic nucleus. Leaving the site were the paths of many fast subatomic particles named secondary cosmic rays.

The complex phenomena associated with collisions of this kind proved to be so interesting and to have such obvious importance for gaining an understanding of nature that the study of such phenomena was increasingly pursued for its own sake. When machines were developed that were capable of artificially accelerating electrons and protons to energies as great as those of at least some cosmic rays, the study of these phenomena achieved independence as the new branch of science called high-energy physics.

It goes without saying that if it is possible to do a given experiment in high-energy physics by using artificially accelerated particles, then it is much better to do the experiment that way than by using cosmic rays. At present virtually everything that happens in collisions between cosmic rays and target nuclei can be duplicated at accelerator laboratories up to cosmic-ray energies of about 10^{12} electron volts. Over the next decade new machines are planned that will push the figure upward to about 10^{15} electron volts. The cost of building them will be very high, however, and the new machines may well be the last of their kind.

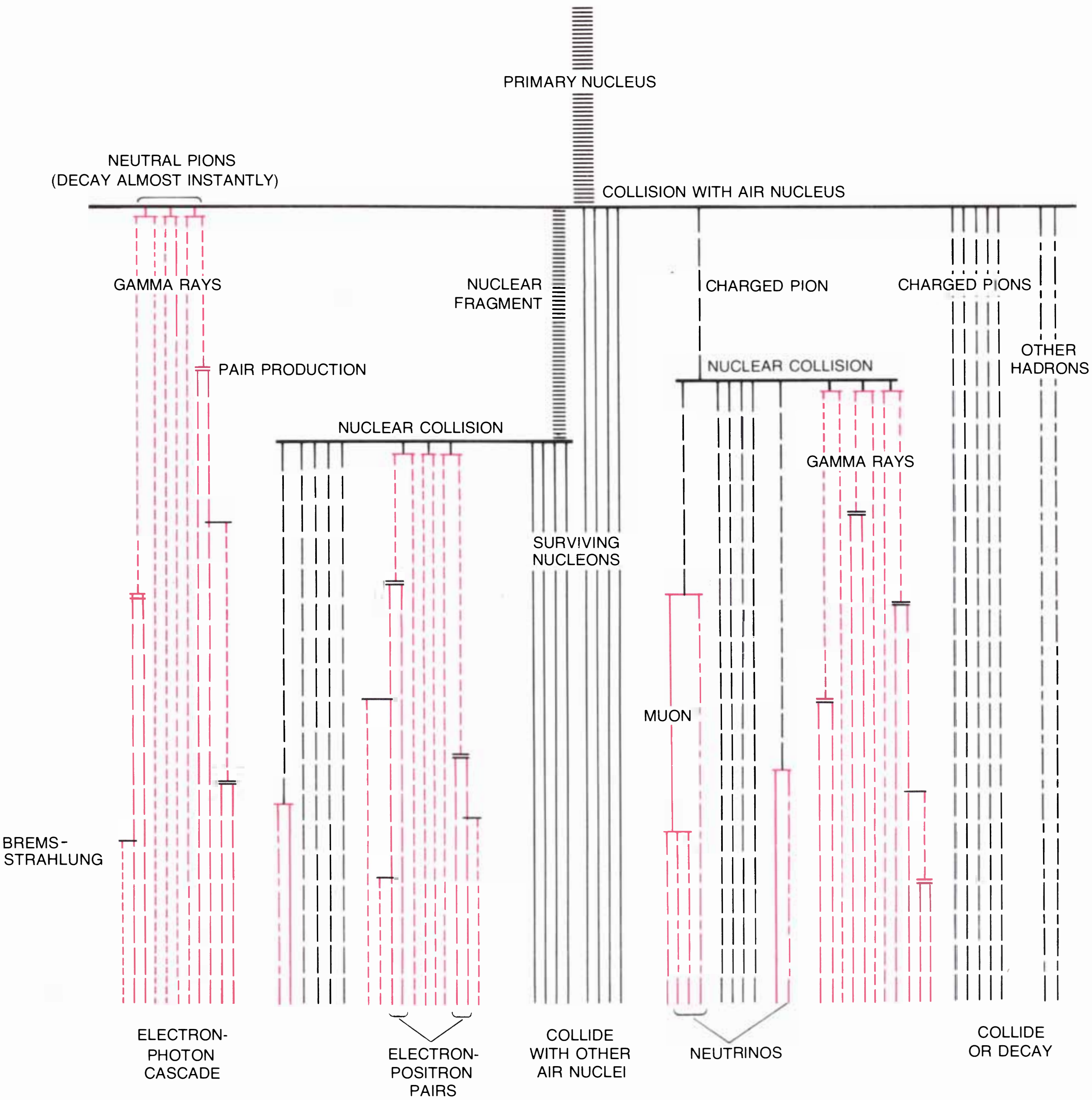
The fact that collisions such as those between cosmic rays and air nuclei can now be studied using artificially accelerated particles has changed the character of cosmic-ray experiments in the energy domain below 10^{12} electron volts. The experimentalists profit by using the accurate, detailed information published by high-energy physicists to calculate how their own instruments will respond to cosmic radiation. Or they simply transport their instruments to an accelerator laboratory and calibrate them by placing them in a beam of particles identical to the cosmic-ray component they intend to measure. On the other hand, hardly any cosmic-ray experiments in the energy domain below 10^{12} electron volts have any longer the goal of acquiring new information about high-energy collisions.

Experiments in the energy domain between 10^{12} and 10^{15} electron volts have an intermediate character. Above 10^{15} electron volts, in the domain of the highest-energy cosmic rays, they still have the character that all cosmic-ray experiments used to have. Working in this energy domain one may be, perforce or by intention, an astrophysicist one day and a high-energy physicist the next.

Regardless of viewpoint, the same two major difficulties stand in the way. One is that the earth's atmosphere tends to prevent the primary cosmic rays from reaching ground level. The probability that an incoming proton, for example, will reach sea level without striking an air nucleus is only about one in a million. The secondary cosmic rays produced in the collisions with air nuclei tend to be unstable particles that decay into others. As a result the cosmic radiation that is present at ground level consists almost entirely of the decay products: muons, neutrinos and gamma rays, together with electrons that are the byproducts of the muons and gamma rays. If, as an astrophysicist, one seeks to study the primary radiation directly, one must send detectors above the atmosphere with balloons, rockets and artificial satellites. If, as a high-energy physicist, one seeks to study collisions between high-energy nuclei and the atomic nuclei in a target, one may find it necessary to do the same thing, or at least to set up shop on a high mountain.

A still greater difficulty is that the intensity of cosmic radiation is very low. Even in the most favorable energy band (from 10^9 to 10^{10} electron volts) and in the most favorable location (above the earth's atmosphere) there is only enough intensity to produce about 10 hits per minute on an area of one square centimeter. This rate is so low that in spite of the high energy of the individual cosmic-ray particles the total energy they deliver to such a target, per unit time, is no more than is received in the form of starlight. Moreover, the intensity of what I am calling here the highest-energy cosmic rays is a great deal lower. The fraction of primary cosmic rays with energies greater than 10^{15} electron volts is so small that in order to capture one particle per hour one would need a collecting device at least 10 meters on a side. Of course, regardless of size there is no device that could literally capture such a particle with its energy intact.

How then was it possible for Pierre Auger and his colleagues at the École



AIR SHOWER is created when a primary cosmic-ray nucleus such as an alpha particle (a helium nucleus) first enters the earth's atmosphere some 20 kilometers above the surface and collides with the nucleus of an atom in the atmosphere. This highly schematic diagram shows the first stages in the development of the shower. The lines in color represent leptons: particles such as photons (which include gamma rays), neutrinos, electrons, positrons and muons. The lines in black represent hadrons: heavier particles such as nucleons (protons and neutrons), compound nuclei and pions. The particles produced in the two secondary collisions shown make up the second generation of particles. The particles they produce, if they collide with other particles before they decay, make up the third generation, and so forth. Hadrons can create both leptons and more hadrons, but leptons can create only more leptons. "Pair production" refers to creation of electron-positron pair by a photon. "Bremsstrahlung" (braking radiation) refers to photons created by the deceleration of other particles.

Normale Supérieure in Paris to deduce, as they did in 1938, that there exist some primary cosmic rays with energy as high as 10^{15} electron volts, from data obtained with a few Geiger tubes having a combined area of less than one square meter? They did it by showing that the simultaneous counter discharges they observed were caused by a previously unknown cosmic-ray phenomenon, and by demonstrating some of the essential features of the phenomenon. They attributed the simultaneous discharges to showers of cosmic-ray particles incident on their equipment from the air above it. From the fact that simultaneous counts occurred even when their Geiger tubes were separated by many tens of meters they deduced that the showers were extensive, blanketing an area of some hundreds of square meters, and that the number of fast particles in the showers was large, perhaps tens of thousands.

The showers discovered by Auger and his colleagues are called extensive air showers or simply air showers. One should not, however, read too much into the name. Unlike showers of rain or confetti, cosmic-ray showers have a highly complex but well-defined structure, and they are over and done with in the wink of an eye. By detecting air showers rather than the primary particles that produce them one can overcome both of the difficulties described above. The equipment for studying the showers can be located at ground level; it does not have to be carried above the atmosphere. And the equipment does not have to completely cover a large area. Because air showers are extensive they can be detected at a distance. A large area can be made sensitive to the impact of air showers by covering it with an array of small, widely spaced fast-particle detectors, and data can be obtained at a reasonable rate in spite of the very low intensity.

The year before Auger and his colleagues discovered air showers a theoretical explanation of similar cosmic-ray showers produced in solid targets had been developed almost simultaneously by Homi J. Bhabha and Walter H. Heitler in England and by J. F. Carlson and J. Robert Oppenheimer in the U.S. In keeping with that interpretation air showers were first thought to be simple electron-photon cascades initiated high in the atmosphere by primary cosmic-ray electrons. A large body of evidence now shows that nearly all of the primary cosmic rays that give rise to air showers are atomic nuclei, and that air showers are a good deal more complex than was first imagined.

At the core of an air shower is a nuclear cascade involving a chain of nuclear interactions. It begins with a collision between two atomic nuclei, the primary cosmic-ray particle and the nucleus of an air atom. Many but not all of the secondary cosmic rays produced by that collision subsequently undergo similar collisions farther along the axis of the cascade, producing secondaries of their own, and so forth. In each generation of collisions, however, some of the secondary cosmic rays decay, notably the pions.

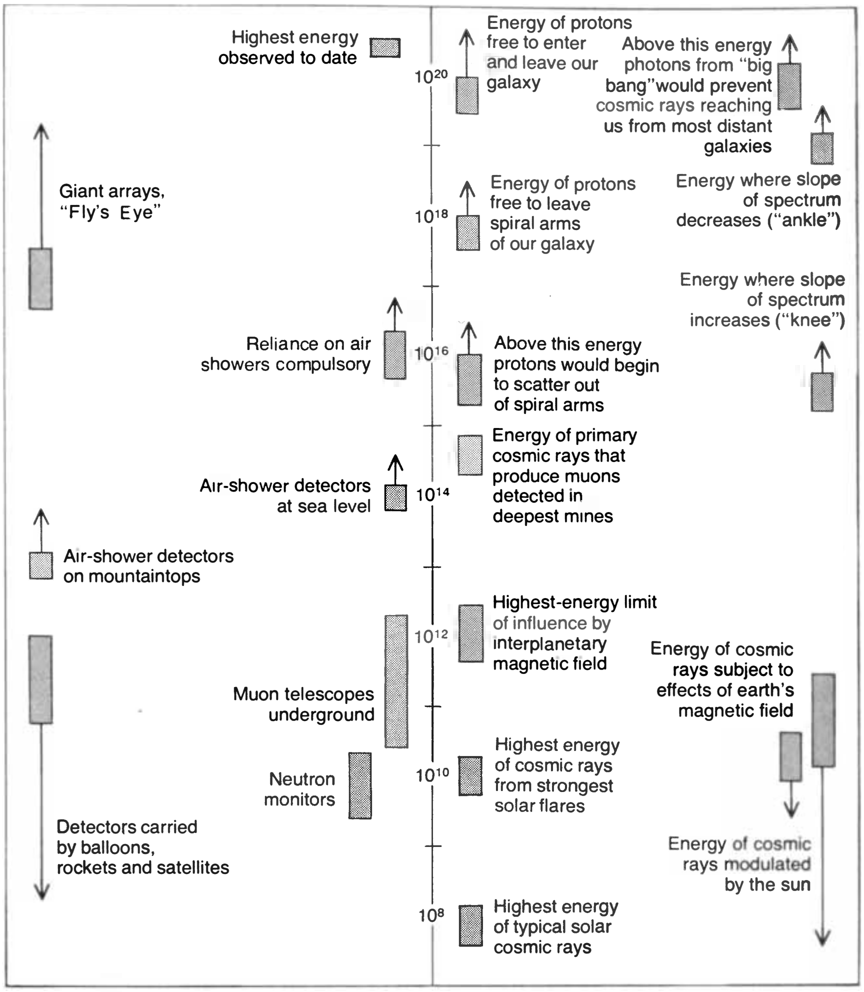
The charged pions produced in the early stages of the cascade usually collide with the nuclei of air atoms rather than decaying, because their lifetimes are so long when allowance is made for time dilation according to the special theory of relativity. Those produced in the later stages of cascades usually decay because their energies are lower and the relativistic effect is not as great. The particles that result from charged-pion decay are muons and neutrinos. The neutral pions are so unstable that they decay immediately, giving rise to pairs

AIR SHOWERS ARE DETECTED by a giant array of detectors mapped by the dots on this aerial photograph of Haverah Park outside Leeds in England. The diameter of the roughly hexagonal area covered by the detectors is about four kilometers. Each detector is made up of numerous sealed tanks of very pure water with light sensors. The sensors (photomultiplier tubes) respond to Čerenkov light generated by secondary cosmic rays penetrating the water. The combined area of the 32 detectors indicated here is 500 square meters.

of high-energy photons. These photons, which receive a substantial fraction of the energy released in each collision, initiate electron-photon cascades. To summarize, a cosmic-ray air shower consists of a cascade of nucleons (protons and neutrons) at its core, together with two groups of fast subatomic particles arising from decay processes: a muon-neutrino component and an electron-photon component.

The electron-photon component is made up of a large number of electron-photon cascades. Such a cascade also involves a chain of collisions. They are collisions, however, in which a high-energy electron or photon interacts with the electric field of an atom. The strong short-range force that binds together nucleons to form atomic nuclei does not play any role. In the early stages of an electron-photon cascade the number of fast particles increases because on the one hand a high-energy photon can materialize when it encounters a strong electric field, forming two particles, an electron and a positron, and on the other hand a high-energy electron or positron encountering a strong electric field can radiate a substantial fraction of its energy in the form of a high-energy photon. As the number of fast particles increases, however, the average energy per particle has to decrease. Eventually dissipative processes come into play and the cascade dies out.

An air shower can be visualized by constructing a 1:20,000 scale model consisting of a disk half a centimeter in diameter cut from thick paper and strung on a thread one meter long. The thread represents the shower axis, a prolongation of the path of the primary particle. If the thread is held vertical, the upper end will correspond to a typical height of the first collision (assuming that the lower end is at sea level). To represent the brief existence of an air shower one should imagine the disk sliding down the thread at the speed of light.



ENERGY OF COSMIC RAYS that have been observed span an enormous range from 10^8 to 10^{20} electron volts (*right*). The energy range and the limitations of different types of detectors must be tailored to investigate different types of cosmic rays (*left*). The length of the rectangles for both the particles and the detectors indicates the typical energy range. The arrows pointing up indicate the range is a lower limit; arrows pointing down indicate the range is an upper limit.

The disk represents the region in which electrons are to be found. Electrons are by far the most abundant charged particles in air showers. The core region in which the nucleon cascade takes place is only as large as the hole for the thread, a few meters in diameter in real air showers. The disk is a fairly good representation of the longitudinal distribution of the electrons. In the central region of real air showers the longitudinal spread of the electrons is only about one meter. At right angles to the axis it represents the situation as being simpler than it is. Unlike the density of the paper, which is constant out to a certain radius and then drops suddenly to zero, the density of shower electrons decreases smoothly and continuously at increasing distances from the axis, without any observed cutoff. The disk in the model represents the region containing half of the electrons. In real air showers the diameter of this region is about 100 meters regardless of the energy of the particle that initiated the shower. The density of shower particles, however, increases with increasing primary energy at all distances from the shower axis, so that even as far from the axis as a kilometer the density of the particles in very high-energy showers is great enough for the showers to be detected with counters of modest size.

If one could follow an air shower in the direction of its motion and watch it develop, one would see that the number of particles increases rapidly at first, reaches a maximum where the creation of new particles is balanced by absorption and then decreases. The number of shower particles at the stage called maximum development, that is, the maximum number of particles in a given air shower, is proportional to the energy of the primary cosmic-ray particle that generated the shower. For example, a primary cosmic ray with an energy of 10^{15} electron volts will generate an air shower containing half a million particles at maximum development. Such comparatively low-energy showers reach maximum development high in the atmosphere. A shower of that energy, however, can still be easily detected at sea level if the primary cosmic ray was traveling in a nearly vertical direction when it entered the atmosphere. The air showers produced by the most energetic cosmic rays yet observed, those with energies equal to two or three times 10^{20} electron volts, appear to reach their maximum development when they have penetrated about eight-tenths of the atmosphere, or to an altitude of about 2,000 meters above sea level.

The experiments of Auger and his colleagues opened a door to studying the highest-energy cosmic rays by showing that air showers provide a means of compensating for the rapid decrease of cosmic-ray intensity with increasing energy. In those experiments, however, the showers were merely detected, not measured individually. It is self-evident that from the isolated fact that discharges occurred at the same moment in widely spaced Geiger tubes one cannot determine how many cosmic rays were involved or find their directions. One cannot even be sure that the coincident discharges were caused by an air shower, since unrelated discharges are occasionally simultaneous by chance. The conclusions reached by Auger and his colleagues, namely that air showers exist and that some of them have energies as high as 10^{15} electron volts, depended on the statistical analysis of many events, aided by the newly developed theory of electron-photon cascades.

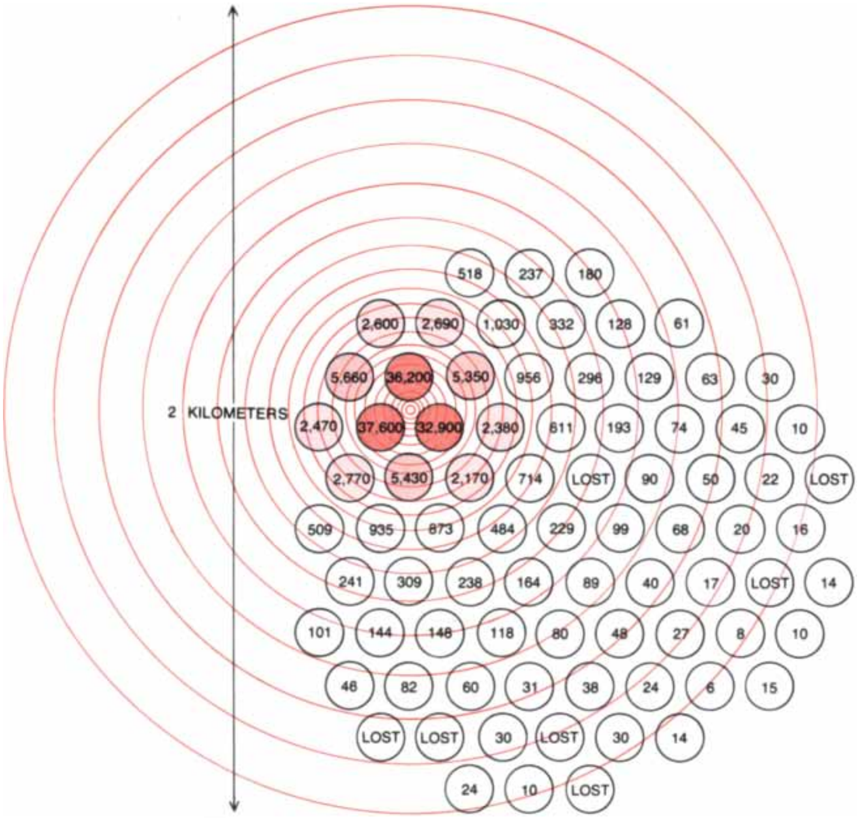
In 1948 an important advance in the study of air showers was made by Robert W. Williams of the Massachusetts Institute of Technology. Instead of Geiger tubes he used ionization chambers, whose signals he recorded by photographing oscilloscope traces. Working on Mount Evans in Colorado, he placed his detectors at the center and vertexes of an equilateral triangle 12 meters on a side. When a sufficiently large air shower struck inside his array of detectors or close outside it, the fast particles produced signals in all of the ionization chambers at practically the same time, so that the event could be detected in the usual way. The oscilloscope traces photographed by Williams revealed more to him, however, than the fact that a shower had struck his apparatus. He knew that the amplitude of the signal from each ionization chamber was proportional to the number of particles that had passed through it. With that information Williams was able to locate the axis of the shower and to calculate the number of particles in the shower.

Five years later, in 1953, the methodology of air-shower observations was further advanced by Pietro Bassi of M.I.T., working with George W. Clark and Bruno Rossi. Instead of ionization chambers they used scintillation counters, which were spaced 30 meters apart and were connected to an electronic recording system with a very fast response. It was they, working with this apparatus, who showed that the particles near the axis of an air shower move in a thin flat disk. They also succeeded in calculating the directions of motion of individual showers from the time delays they observed between pairs of detectors on base lines at right angles to each other.

The scintillation counters used by Bassi and his colleagues are like ionization chambers in that the amplitude of the output signal is proportional to the number of fast particles that produced the signal. That feature was not used in the experiment I have just described, but the possibility of using it in a subsequent experiment was not overlooked. Almost at once a larger group formed at M.I.T. under Rossi's leadership. Its goal was to combine the new direction-measuring technique with the method used earlier by Williams for determining the number of particles in showers. By 1956 Rossi's group was measuring cosmic-ray air showers with an array of 15 much larger scintillation counters outside Boston at the Agassiz Station of the Harvard College Observatory. The area of the array was about a sixth of a square kilometer. It could measure the number of particles in showers with an accuracy of about 20 percent and it could measure shower directions to within a few degrees. The Williams-Bassi technique was soon adopted by almost all other groups that were studying air showers.

The most exciting result from the Agassiz array was the discovery of evidence, provided by an unusually large air shower recorded in 1957, that the energy of primary cosmic rays can be as great as 5×10^{18} electron volts. This discovery created a stir in the astrophysical community because of its bearing on the problem of cosmic-ray origin. Just a few years earlier air-shower evidence had been decisive in disproving a proposal advanced in 1949 by R. D. Richtmeyer and Edward Teller: that all cosmic rays might come from the sun. Richtmeyer and Teller argued that the observed isotropy of cosmic rays (that is, the fact that their intensity far from the earth's influence appeared to be the same in all directions) could be accounted for by the action of a hypothetical magnetic field that confines them (by bending their paths) within some large volume containing the sun. They conceded that it would not be possible to confine particles with energies much greater than 10^{16} electron volts without assuming too large a volume or too strong a magnetic field. Almost immediately other workers pointed to evidence from air showers that there were many primary cosmic rays with energies greater than 10^{16} electron volts, so that the sun could not be the source of all cosmic rays.

Although the solar-origin hypothesis is no longer believed, arguments similar



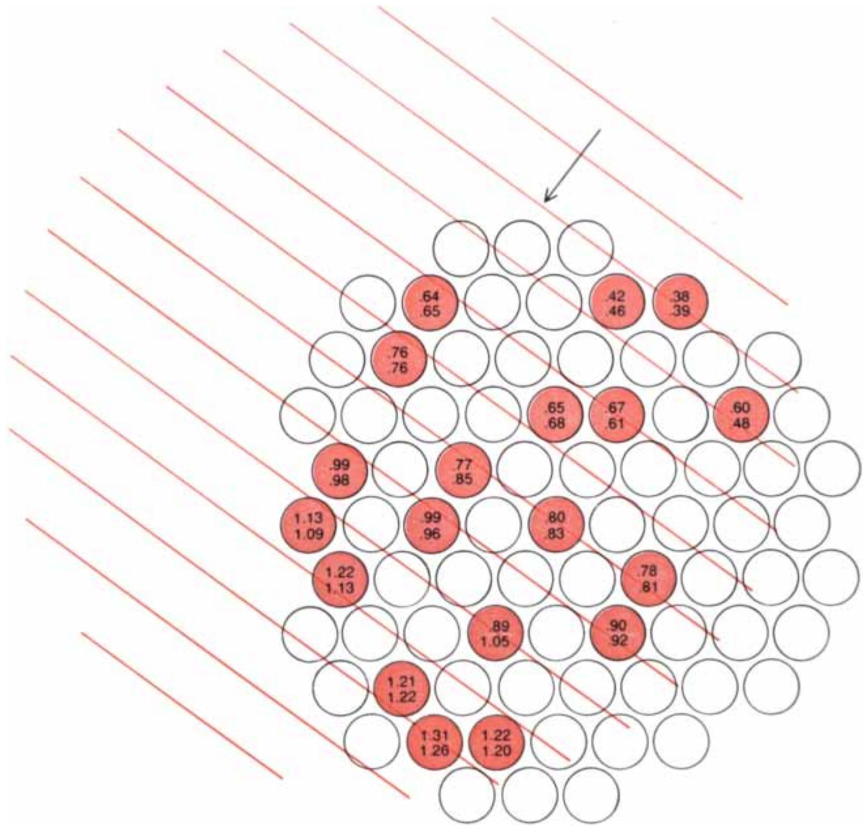
INTENSITY MAP of a very large air shower produced by a cosmic ray with an energy of 10^{19} electron volts was recorded by an array of detectors at Volcano Ranch in New Mexico in March, 1973. The concentric lines in color represent contours of equal intensity measured in terms of number of particles per square meter. The shower axis was nearly vertical, so that the contour lines are nearly circular. The black circles represent scintillation counters one meter in diameter that are located 150 meters apart on level terrain. (The detectors are not drawn to scale; if they had been, each detector would be smaller than the period at the end of this sentence.) The number inside each detector represents the number of particles from the shower measured at that point. The largest signals, received by the detectors in color, show the axis of shower passed within a few meters of the center of a triangle formed by adjacent detectors.

to those used by Richtmeyer and Teller continue to be adduced when alternatives to the solar-origin hypothesis are discussed. For example, the notion that cosmic rays are confined by the action of a magnetic field of some kind can be applied to the entire galaxy. According to the galactic-origin hypothesis there is a critical cosmic-ray energy analogous to the critical energy of 10^{16} electron volts predicted by Richtmeyer and Teller's theory. The magnetic field of the galaxy appears to be somewhat weaker than the field hypothesized by Richtmeyer and Teller, but the volume of the galaxy is a great deal larger than the volume they proposed. Hence the critical cosmic-ray energy for galactic confinement is higher than 10^{16} electron volts by several orders of magnitude.

The new cosmic-ray energy record set by the largest air shower detected with the Agassiz array did not exceed the "galactic limit," but for defenders of the galactic-origin theory it came uncomfortably close. Almost as soon as the Agassiz array went into operation work began at Cornell University on a similar array that would have almost five times as much sensitive area. At M.I.T. plans were made for eventually relocating the Agassiz detectors in Bolivia to form an array with four times the original area. The results coming in from the Agassiz experiment, however, confirmed that the intensity of cosmic radiation falls off very rapidly with increasing energy. Arrays that were only four or five times larger than the Agassiz array would not resolve the question of whether some cosmic rays are so energetic that they cannot be held within the galaxy.

As a new member of Rossi's group I had been amusing myself privately with grandiose plans for the largest array I thought realistic to build and operate. It would be a giant array, not five times larger than the one at Agassiz Station but 50 times larger. I was not the only one in our group with such dreams, but I was the only one free enough at the time to pursue his plan. I had the encouragement of Rossi, who saw to it that I was supplied with all of the material resources that were needed. Above all, it was through Rossi that I acquired an invaluable colleague, Livio Scarsi, who was visiting M.I.T. on leave from the University of Milan. By 1959 Scarsi and I had built our giant array at Volcano Ranch near Albuquerque, N.M.

The ice had been broken: giant arrays were seen to be a realistic possibility. Within a few years a group headed by J. G. Wilson of the University of Leeds brought a second giant array into operation at Haverah Park near Leeds. Since then two more giant arrays have been set up, one in Australia by a group at the University of Sydney led by C. B. A. McCusker and the other in Siberia near



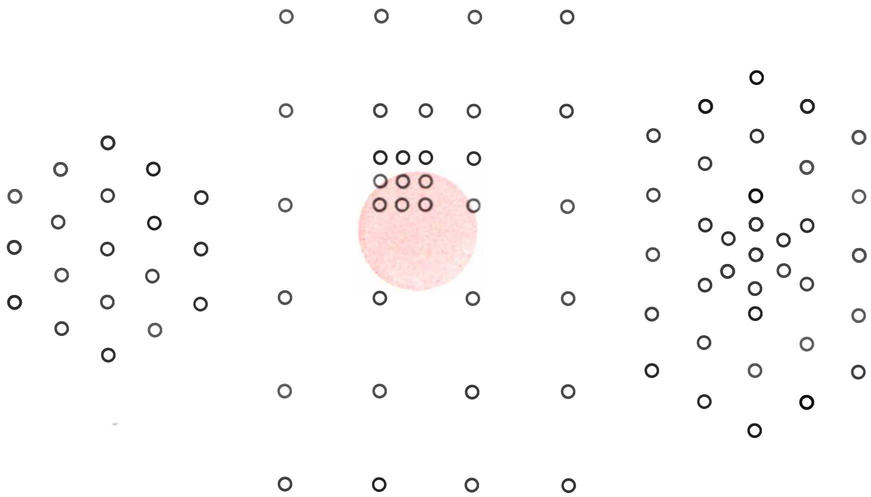
ARRIVAL-TIME MAP of the same large air shower was also recorded by some of the detectors in the Volcano Ranch array. (To reduce costs not all of the detectors were equipped to measure arrival time as well as intensity.) The diagonal lines in color show where the plane of the shower disk intersected the plane of the ground at successive instants. Because the shower axis was nearly vertical the shower plane was nearly horizontal, and the line of its intersection with the ground swept across the array of detectors in less than a microsecond (a millionth of a second). The orientation of the shower axis was calculated from the measured times of arrival by fitting a theoretical plane to them. Agreement between theoretical arrival times (*bottom numbers*) and measured ones (*top numbers*) signifies shower direction was determined to accuracy of a few degrees. All of the times are in microseconds from an arbitrary starting point.

the city of Yakutsk, just a few degrees south of the Arctic Circle, by a group at the Institute of Cosmophysical Research and Aeronomy led by Yu. G. Shafer.

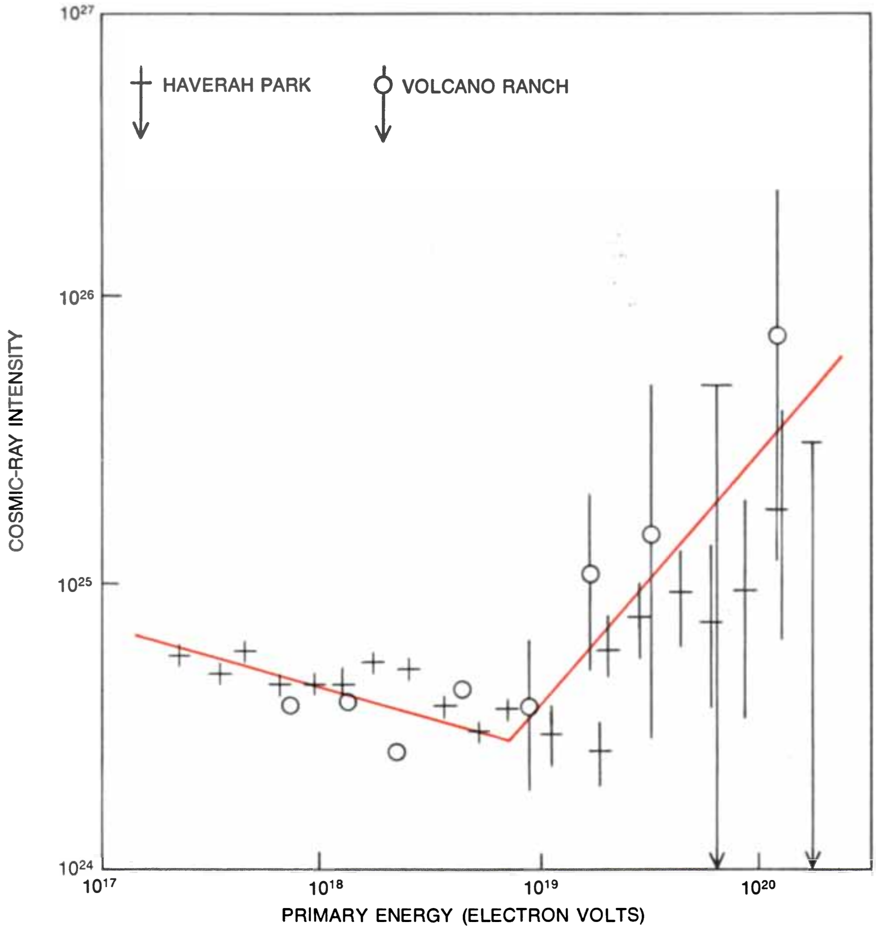
The notion that cosmic rays are confined within galaxies had been threatened in 1957 by the most energetic air shower observed with the Agassiz array. In 1962 it was dealt a crushing blow by the largest air shower Scarsi and I observed at Volcano Ranch, a shower whose energy was estimated to be 10^{20} electron volts. When Scarsi and I started our project, we had not expected to detect a shower with such a high energy as that. Equally unexpected was our discovery that there were more than the predicted number of primary particles at all energies above a few times 10^{18} electron volts. At those high energies the cosmic-ray spectrum became flatter: there was an "ankle" in it. (The energy spectrum of cosmic rays is a curve that describes how the cosmic-ray intensity changes as one proceeds up or down the energy scale.)

The flattening of the spectrum that was indicated by the data from Volcano Ranch was a result no one had expected. In fact, the proponents of the galactic-origin theory had predicted just the opposite. They predicted that the cosmic-ray energy spectrum would become steeper at energies approaching the critical energy for galactic confinement. Could it be that above some energy the loss of cosmic rays from our galaxy is more than compensated by cosmic rays entering the galaxy from intergalactic space? If so, what is the source of the cosmic rays that flow inward? Are the quasars or any other objects known to astronomy capable of filling the vast reaches of space between galaxies with such remarkably energetic subatomic particles?

The 10^{20} electron-volt event observed at Volcano Ranch soon took on added significance. In 1965 Arno A. Penzias and Robert W. Wilson of Bell Laboratories discovered a peculiar electromagnetic radiation, at microwave frequencies, that apparently fills the entire universe. It is sometimes called the relict



GIANT ARRAYS, three of which are diagrammed here, are usually named for their location. There is no one best way for arranging the detectors in an array. From the left to the right are diagrams of the arrays at Volcano Ranch, at Sydney in Australia and at Yakutsk in the U.S.S.R. The circle in color, representing an air shower with an effective diameter of two kilometers, indicates the scale of the arrays. The Volcano Ranch array no longer exists in the form shown. After four years of operation (1959 through 1963) it was taken apart and reassembled to form the smaller but denser array portrayed in the illustrations on the preceding two pages.



ENERGY SPECTRUM of the most energetic primary cosmic rays has been determined by two of the giant arrays. The vertical axis represents the cosmic-ray intensity in terms of the Hillas function (differential intensity multiplied by the third power of the energy). If the spectrum had the form of a simple inverse-cube power law, the experimental points would fit a horizontal straight line. Actually the spectrum appears to be somewhat steeper than an inverse-cube law up to an energy of 10^{19} electron volts, and it is flatter than an inverse cube above that energy. The portion of energy spectrum where spectrum changes shape is called the ankle.

radiation because it is believed to be the radiation left over from the "big bang" with which the universe began. Learning of the discovery by Penzias and Wilson, Kenneth Greisen of Cornell and G. T. Zatsepin of the P. N. Lebedev Physics Institute in Moscow realized independently that encounters between cosmic rays and photons of the relict radiation would severely drain the energy of the cosmic rays above some energy threshold that they could easily calculate. They showed that if the sources of cosmic rays are uniformly distributed throughout the universe, the consequence of these encounters would be an abrupt steepening of the cosmic-ray spectrum at about 6×10^{19} electron volts. The intensity of cosmic rays with energy as great as 10^{20} electron volts would be so low that 10^{20}-electron-volt showers would be impossible to detect.

Could Scarsi and I have been mistaken about the energy of our largest shower? It seems we were not. As the other giant arrays came into operation they too recorded air showers with energies as high as ours and even higher. The ankle in the spectrum has also been confirmed. What picture of cosmic-ray origin do these results suggest, and what other kinds of information can be given by the giant arrays?

It is generally believed that practically all cosmic rays in the spectrum below the ankle are generated by sources inside our galaxy. Moreover, it is attractive to suppose that these cosmic rays are confined to the galaxy, because in that case the hypothetical sources of the cosmic rays need not be as powerful as they would otherwise be. If cosmic rays are confined, their intensity in interstellar space builds up until it is balanced by leakage into intergalactic space (losses inside the galaxy being unimportant). Because the supposed confinement is brought about by magnetic fields it is not equally effective for cosmic rays with different energies. Low-energy cosmic rays are confined very effectively, those with energies approaching the critical energy have an increasing tendency to leak out of the galaxy and those above the critical energy are not confined at all.

There is no conflict between the notion of confinement and the two-component theory of cosmic-ray origin, according to which cosmic rays below some energy are predominantly galactic in origin whereas those above it are predominantly extragalactic. Many cosmic-ray theorists were attracted to the two-component model when air-shower experiments showed that the cosmic-ray spectrum, which had been expected to steepen at energies approaching the critical energy, instead became flatter. Other theorists, however, continued to defend the hypothesis that all cosmic rays are generated inside the galaxy. There is more than one way, they pointed out, to compensate for the loss of high-energy cosmic rays to intergalactic space. Little is known about the natural accelerators that generate the cosmic radiation; perhaps the cosmic-ray spectrum flattens because the efficiency of the accelerators increases dramatically above some energy, more than compensating for whatever increases there may be in leakage from the galaxy. Thus the shape of the cosmic-ray energy spectrum is not crucial for deciding between the galactic theory of origin and the two-component theory; one must look to some other kind of evidence in order to make a choice. Measurements of air-shower directions provide the kind of evidence that is needed.

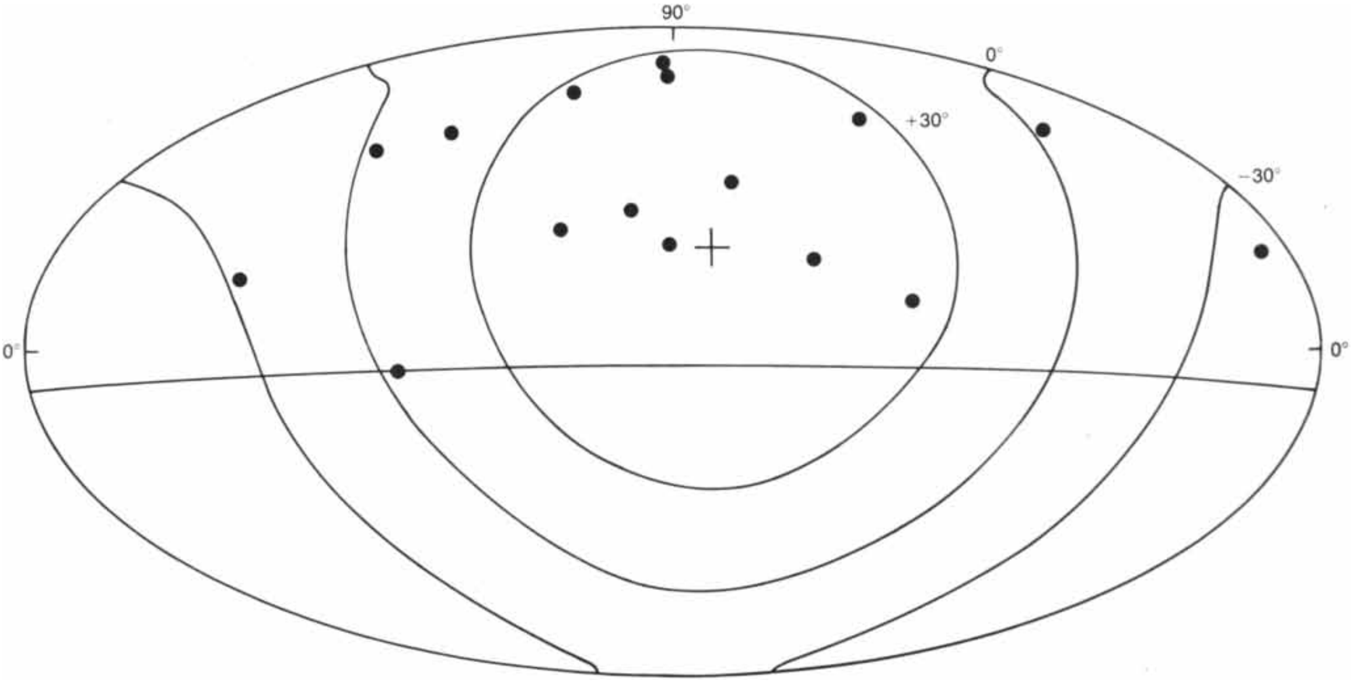
The confinement hypothesis accounts for the fact that cosmic rays with low and moderately high energies are nearly isotropic: they come from all directions in space with almost equal intensities. Above the critical energy, however, cosmic rays are not confined at all, and they cannot even be deflected very much by the galactic magnetic field; therefore their directions of arrival must indicate where their sources are located. If even the highest-energy cosmic rays originate within the galaxy, then most of them should arrive from the plane of the Milky Way and very few from directions more or less perpendicular to that plane. If they originate in active galaxies outside our own, it may eventually be possible to pinpoint the most important cosmic-ray-emitting galaxies.

It is exciting to think about what may be revealed when more data have accumulated, particularly at energies above 6×10^{19} electron volts. It seems virtually certain that the energy spectrum will show some kind of structure that can be related to the Greisen-Zatsepin effect, thus providing a check on the accuracy of present energy measurements in this domain, which are based on the detailed analysis of air-shower structure. The Greisen-Zatsepin effect also has the potential of providing an independent astronomical distance scale covering the range from 30 million to one billion light-years.

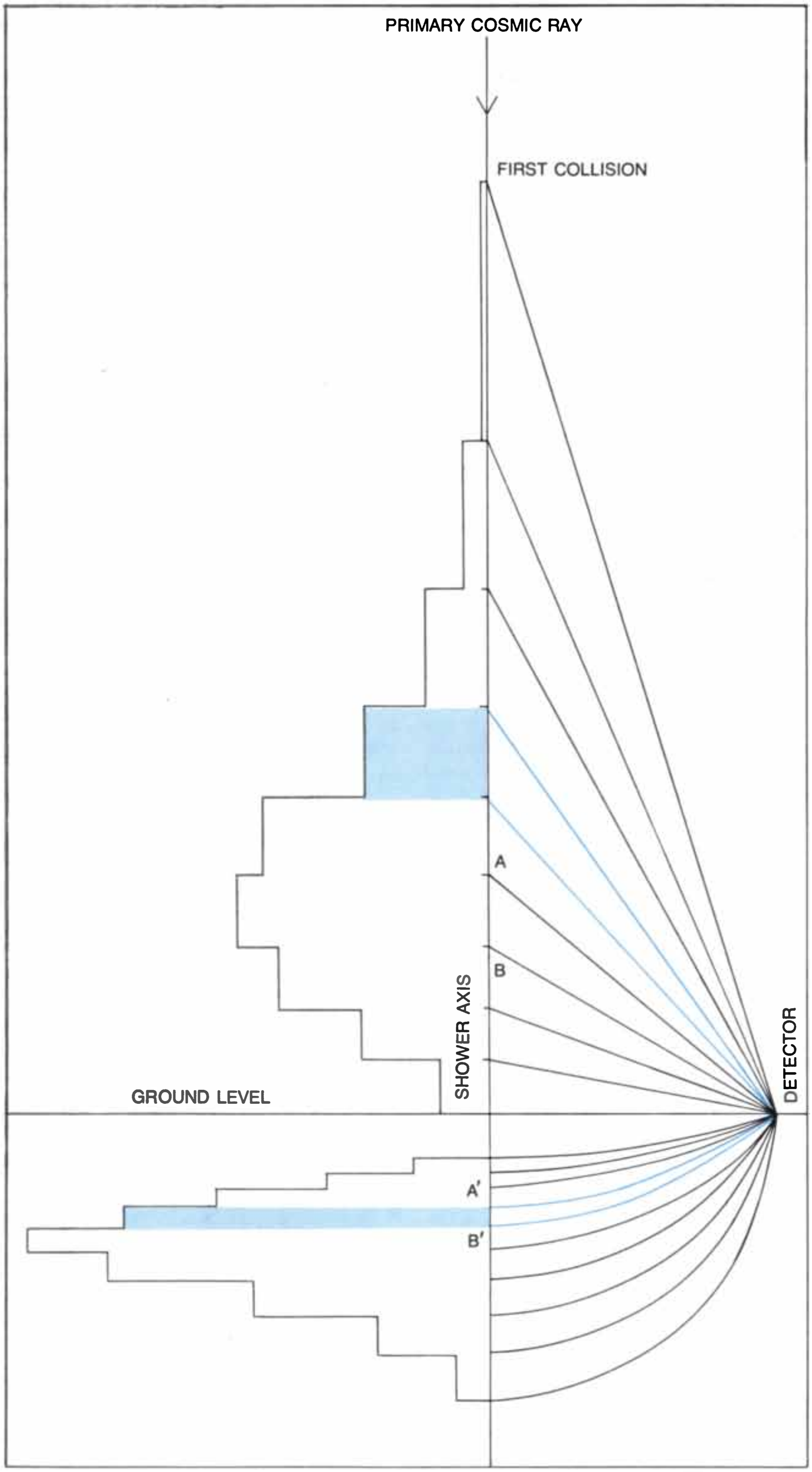
In order to be so optimistic, however, I have had to assume that a major problem is going to be solved, a problem that I have not yet given the attention it deserves. Underlying my comments about the interpretation of existing data, particularly the data on cosmic-ray directions of arrival, is a well-known formula for calculating the curvature of a charged particle's path when it moves perpendicular to a magnetic field. If the particle is highly relativistic, that is, if it is moving at a velocity very close to the velocity of light, the radius of curvature is proportional to the ratio of its total energy to the strength of the field. In the case of cosmic rays, air-shower measurements give the total energy of the particle and astronomical data provide an estimate of the strength of the magnetic field. But the radius of curvature is also inversely proportional to the electric charge of the moving particle, which for low-energy cosmic rays is likely to vary from 1 (for protons, or hydrogen nuclei) to 26 (for iron nuclei).

Thus if the critical energy for galactic confinement of cosmic-ray protons is 2.6×10^{19} electron volts according to a certain model, the critical energy for confinement of cosmic-ray iron nuclei according to the same model will be 26 times less, or only 10^{18} electron volts. It therefore makes a great deal of difference whether on the one hand the highest-energy cosmic rays are protons or iron nuclei or on the other they have a composition unlike either of those alternatives. Discovering the composition of the particles that initiate large air showers is a major unsolved problem in research on the highest-energy cosmic rays.

How can one tell what primary particles produce the air showers that are observed? There is no question that the lower-energy particles that give rise to small air showers are virtually all atomic nuclei. That is certain because low-energy primary cosmic rays have been studied in some detail with detectors above the earth's atmosphere. Studies of this type are impracticable at energies above 10^{15} electron volts; the structure of air showers furnishes evidence,



ARRIVAL DIRECTIONS of the 16 most energetic showers registered by the Haverah Park array since 1963 provide the single most convincing piece of evidence that the most energetic cosmic rays originate outside our galaxy. Each shower had an energy greater than 5×10^{19} electron volts. The individual directions are shown on an Aitoff equal-area projection of the celestial sphere. The portion of the sky that is above the line running across the middle of the projection drifts over the array and is scanned by the array once every 24 sidereal hours. The cross above the center represents the galactic north pole, which is tilted some 60 degrees with respect to the axis of the earth. Also shown are the parallels of galactic latitude of −30 degrees, 0 degrees and +30 degrees. If the 16 energetic cosmic rays had originated inside the galaxy, far fewer of them should have arrived from directions that were within 30 degrees of the galactic north pole.



LONGITUDINAL DEVELOPMENT of an air shower is reflected by the shape of the signal pulse produced in a detector that is located to the side of the shower axis. The longitudinal profile of a shower can be measured by detecting the radiation emitted by the shower as it passes through the atmosphere; such a profile is a rich source of information about the particle that created the shower. The principle is illustrated under three simplifying assumptions: first, that the region from which detected particles are emitted has a negligible size; second, that it moves along the shower axis at the speed of light, and third, that the detected radiation travels in straight lines also at the speed of light. The histogram above ground level is the emission profile, representing relative contributions to the light signal from successive intervals along the shower axis. The lower histogram is the corresponding arrival-time profile. Since all particles are traveling at the same speed, time intervals are proportional to distances. This approach is used with muon detectors and with light detectors for receiving atmospheric Čerenkov light.

however, that at least a great majority of the cosmic rays that give rise to large air showers are also atomic nuclei. Although it is possible that some of the highest-energy cosmic rays are subatomic particles less massive than nuclei, there is no evidence that this is actually the case.

Methods intended for determining what fraction of large air showers are produced by the atomic nuclei of different elements are based on various ways of measuring two properties of an individual shower: the fraction of muons in the shower at a specified stage of development and the shape of the shower's longitudinal profile (the curve that describes the growth and eventual decline in the number of shower particles). There are strong reasons for believing that each of these properties will vary in a distinctive fashion, depending on the Lorentz factor of the primary particle that initiated the shower. The Lorentz factor of a moving particle is the ratio of its total energy to its rest-mass energy. The greater the Lorentz factor of the primary particle is, the smaller the fraction of muons in the shower and the broader the profile of the shower.

The fraction of muons in a shower can be measured by means of a detector array of the Williams-Bassi type, which measures the number of particles of all kinds, together with a fast-particle detector shielded by enough concrete, lead or earth to absorb all electrons and photons, leaving only muons. Methods of directly recording individual shower profiles are a recent development to which I shall return. In the past it has only been possible to look for differences in profile between showers of the same energy by looking for differences in some property that is correlated with the profile.

If one examines these properties for all the sizes of showers that have been observed, one can tell that as the total energy of the shower increases the Lorentz factor roughly keeps pace. Therefore one can be certain that there is no radical change in the mass of the primary particles that generate showers of different sizes.

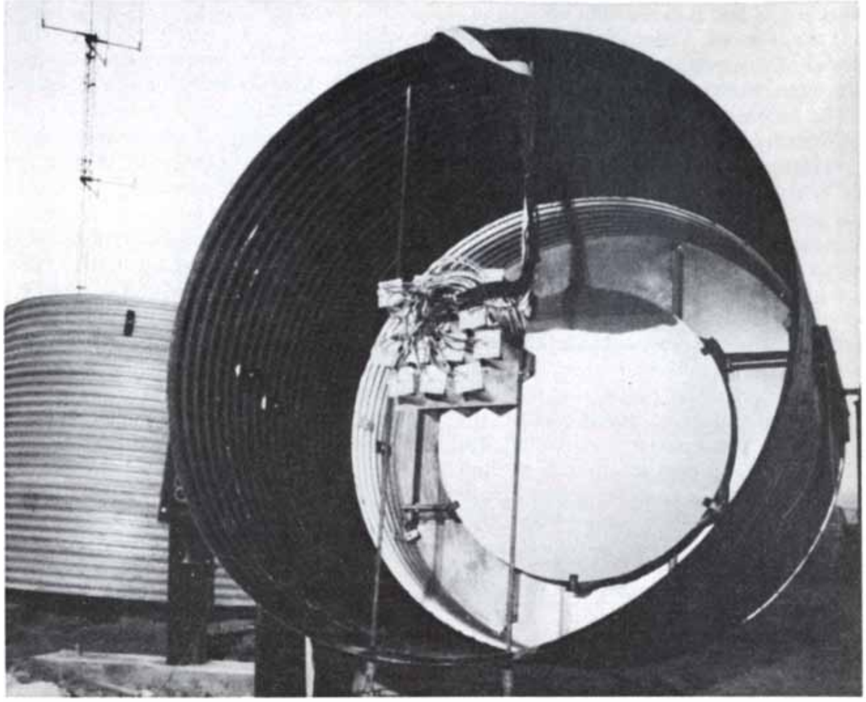
The theory relating the two properties to the Lorentz factor of the primary particle is not exact, however, and neither are our measurements. With current methods one is not able to distinguish between different nuclei with the same energy, not even between a proton and an iron nucleus. The problem is a serious one. Even if one knows the energy of a cosmic ray and the direction from which it entered the atmosphere, how can one go back through interstellar space, retracing its magnetically curved path, if its charge is not known to better than a factor of 26? It should not be forgotten that above an energy of 10^{15} electron volts cosmic rays may continue indefinitely to provide the only means for investigators to study nuclear interactions. How can one interpret what happens in collisions between particles at those energies without knowing what particles make up the incident beam?

Several new experiments are currently under way to directly measure the profile of showers as they develop in the atmosphere. Recall the model of a small disk sliding on a thread. The particles within the disk generate light. The air-shower disk is therefore a kind of fireball, which grows in brightness as the number of particles increases, reaches maximum brightness as a shower reaches maximum development and then begins to fade. The newest methods for measuring air-shower profiles are conceptually equivalent to photographing the moving fireball at night by means of the light it emits. The record of the air shower provided by these methods is equivalent to a sequence of photographic frames.

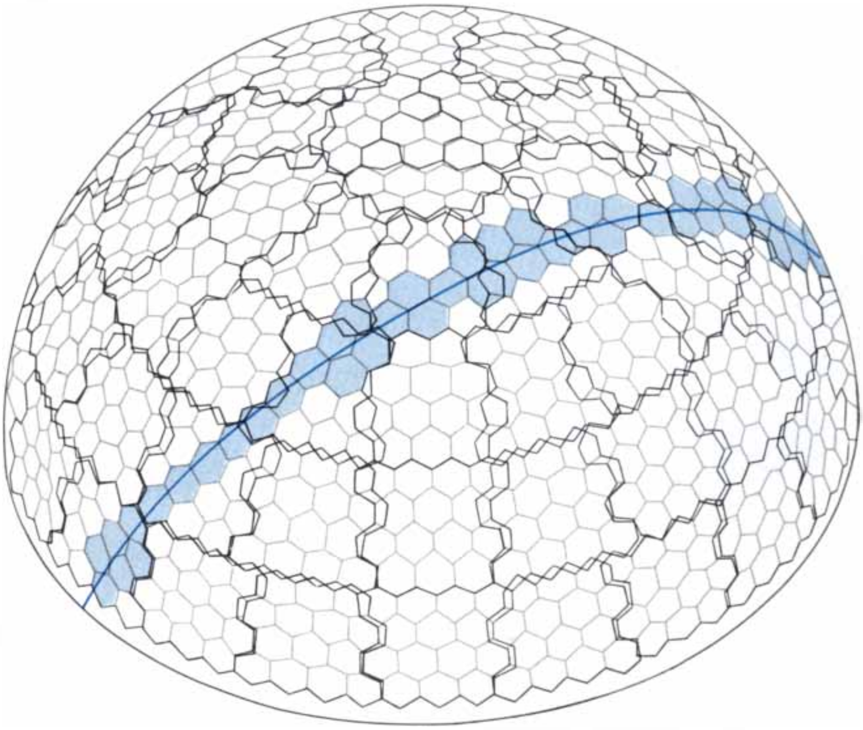
The light from the fireball consists of two components corresponding to two different processes by which fast particles can produce light. One component, which in this case is much the more intense, is beamed forward like the light from a locomotive headlamp. The less intense component is emitted equally in all directions like the light from a meteor. The forward-beamed light is produced by the Čerenkov effect and is called atmospheric Čerenkov light. The light that is emitted isotropically is produced by fluorescence. Some of the new experiments use the atmospheric Čerenkov light produced by air showers; others use the fluorescent light.

The technique for photographing air showers by means of atmospheric Čerenkov light has already been developed and is in use at two of the giant arrays, the one in England and the one in Siberia. The best camera angle for obtaining intense exposures results when the shower approaches the camera head-on. The fireball moves, however, with virtually the same velocity as the light it emits, so that the angle between the shower axis and the camera's line of sight cannot be too small or it becomes impossible to resolve the shower's profile. The best compromise between signal intensity and resolving power is given by showers that strike several hundred meters from the camera. The light collector employed in an experiment of this type has little resemblance to an ordinary motion-picture camera. The collector has no optical components. The successive frames are superposed spatially; they are resolved in the time domain.

After the shower trajectories have been calculated in the usual way from data provided by the giant array, the apparent spatial profile can be derived from the time profile by a simple graphical method. Deriving the true profile,



ONE ELEMENT of the new Fly's Eye air-shower detector, under construction at Little Granite Mountain on the Utah salt flats, consists of a mirror 1.5 meters in diameter. The entire detector will be made up of 67 such mirrors, each with either 12 or 14 photomultiplier tubes mounted in its focal plane. (The mirror shown here has 14.) Each photomultiplier tube will detect and record faint light emitted by cosmic-ray showers in atmosphere. When all 67 mirrors are completed, they will be assembled to form composite "eye" with 880 photomultiplier tubes.



FLY'S EYE DETECTOR, when completed, will monitor the sky for cosmic-ray air showers. Each slightly overlapping polygon, outlined in black, represents the field of view of one of the 67 mirrors. Each polygon is made up of either 12 or 14 hexagons, each representing the six-degree field of view of a photomultiplier tube. An air shower passing over the Fly's Eye will emit faint light that will activate the photomultiplier through whose field of view it passes (*color*). Each photomultiplier tube will have its own electronic channel for recording time and intensity of signals it receives. A computer will reconstruct the shower profile from the signals.

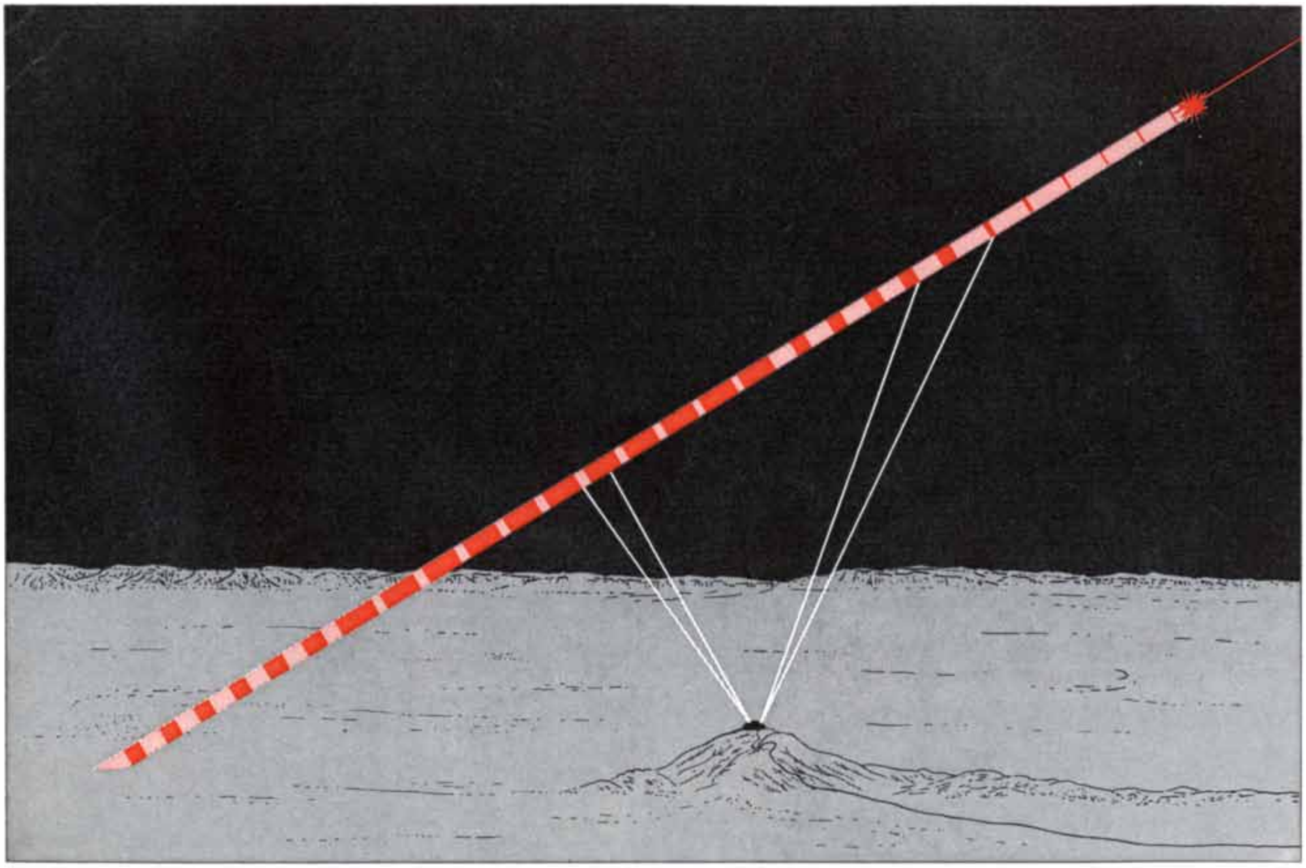
however, is much more difficult, requiring the use of a detailed theoretical model of nucleon cascades, with attendant uncertainty in the results.

The technique for photographing air showers by means of atmospheric fluorescent light is now being developed in the U.S. and Japan. One advantage of this approach is that the air showers are photographed from the side, the ideal camera angle for good spatial resolution. Another advantage is that it should be possible by this method to detect large air showers at a considerable distance, perhaps tens of kilometers. The camera required for registering the fluorescent light must be much more elaborate than the one for registering the Čerenkov light; on the other hand, the shower trajectories can be derived from an analysis of the shower photographs. Hence the atmospheric-fluorescence camera is an independent instrument, one that does not require supplementary information from an adjacent detector array of the Williams-Bassi type. Lastly, and perhaps most significantly, only minor corrections are needed in order to derive the true profile of a shower from photographs of its fluorescent light.

The great potential advantages of detecting air showers by means of atmospheric fluorescence have been understood at least since 1962. So have the difficulties, chief among which is the extremely low intensity of the fluorescent light. To achieve even limited success one would need numerous large, efficient light collectors in a nearly ideal location, with clean air, a favorable weather pattern and as little background light as possible. Several unsuccessful attempts to build a practical system were made between 1965 and 1974. Two years ago a group at the University of Utah led by Haven E. Bergeson brought three new highly sensitive prototype light collectors to Volcano Ranch and carried out successful tests in conjunction with the Volcano Ranch array. A complete device, consisting of 67 light collectors whose field of view is divided among 880 photomultiplier tubes, is now under construction. As collectors are finished they are being installed on Little Granite Mountain at Dugway Proving Grounds, about 100 miles west of Salt Lake City.

The Utah device has been named the Fly's Eye. Each photomultiplier tube monitors its own narrow hexagonal field of view, just as each element in the compound eye of a fly does. For ease of operation the collectors are being mounted separately, each in a rotating housing of its own. After nightfall, in good weather, the collectors will automatically swing into position so that together they will view the entire sky. Signals from the photomultiplier tubes will be scrutinized by a small on-line computer. The "signature" of an air shower is a fireball moving in a straight line at the speed of light. When the computer recognizes an event with that signature, it will command a high-speed motion picture to be recorded, in code, for later analysis. From the record so obtained the shower track can be reconstructed in all three spatial dimensions. The reconstruction will at once give the energy and the arrival direction of the primary cosmic ray that gave rise to the shower.

If the performance of the Fly's Eye measures up to expectations, the device will accumulate data on the most energetic cosmic rays at 10 times the rate of all the existing giant arrays combined. The directly measured shower profiles it will also provide should play an important role in resolving the question of composition and thus overcoming one of the last remaining obstacles to discovering the origin of the highest-energy cosmic rays.



TRAJECTORY OF AN AIR SHOWER can be calculated by data that will be recorded by the Fly's Eye. Here the completed detector is depicted. Since the light from air showers is so faint, it will be detectable only at night. Air-shower fireballs move in a straight line at virtually the speed of light. The four lines extending from the path of the air shower (*color*) to the detector represent the paths by which light from the shower will reach two separate pairs of adjacent photomultiplier tubes in the detector. The time delay between signals from adjacent photomultipliers will vary in a distinctive fashion depending on geometrical relation between the detector and the path of shower.

ANWYL, HAHN, and SCOTT

Too many words per dollar. Far too many words.

We use thoroughly up-to-date marketing methods. We make products to please the widest possible spectrum of humanity and place them in humanity's hand with an efficiency that only the most modern techniques can achieve. That is our policy. To this we are dedicated. That is the image we wish to present. And those three employees are not helping us do so.

They see themselves as missionaries for photography, a realm with vague boundaries. They market no equipment with nice $50,000 price tags. They market what an accountant calls "supplies," yet they encourage a heart-to-heart chat before placing an order for products such as the following, announced during the past 12 months.

Kodak special product, type 129-01, for combination of autoradiography with transmission electron micrography (TEM). 4-oz. bottle containing solidified emulsion shipped on a Monday or Tuesday directly from Rochester to the user. Under proper safelight conditions a few grams in distilled water warmed to liquefy, then allowed to cool. Slide dipped in bearing specimen section on collodion over conducting layer of carbon. Monolayer of silver halide crystals thus laid down, where previous product required centrifuging. Now, with just right dilution, monolayer turns out 85 nm thick with little overlap of crystals. Thickness and uniformity judged by interference color. Slide then left wrapped in refrigerator for up to a month while radioisotope exposes silver halide grains. No inert gas needed anymore because latent image decay is less than with previous product. Then develop, fix, wash, strip off collodion mount for TEM; see over what organelles blackening has occurred. More than twice as sensitive as previous product. Works best with iodine-125 or tritium. Less background blackening. "Special" in product name warns that progress in improving characteristics and quality control has not yet leveled off.

The more electrons used in painting a TEM image on photographic material the less grainy. Possibility of artifacts from broiling of specimen sets limit. Another limit set by need for negative not to approach uniform density all over. There we can help. Kodak electron image plate designed with this in mind. Glass plates don't outgas much to contaminate EM vacuum system, but neither do they bounce when dropped nor store as compactly as film. Estar film base is almost as good as glass for outgassing, but manufacturing problems arise in trying to duplicate photographic characteristics between film and plates. This nearest film equivalent dubbed "SO-163" while working on these problems. Now close enough to drop the "SO" business. Five sheet-film sizes generally in stock at headquarters.

From "Special" status, now Kodak spectroscopic plates, type IIIa-F, have emerged into the bright sunshine of our confidence. Literally, however, must be handled in total darkness. No Kodak commercial product is more sensitive to light. Statement requires qualification two ways: 1) 2- to 4-month delivery time for 36 plates (minimum order) still leaves it "commercial," 2) "sensitive" to weak light for long exposure rather than "fast" in freezing action by short exposure. "Speed" in sense of fewer hours instead of fewer microseconds can be further enhanced by pretreatment with hydrogen gas, in which case user had better take precautions against worse troubles than insufficient emulsion speed. This product not aimed at the impulse buyer. Very high resolving power, rather uniform sensitivity from 250 to 690 nm.

Replaces Kodachrome II professional film (type A) in 135-36 for lecture-grade color slides. One more notch of improvement in granularity, sharpness, and color quality with tungsten (3400 K) illumination—as long as one does not need the higher speeds of Kodak Ektachrome films and likes an elaborately equipped plant to do the processing. Some people simply prefer Kodachrome films to Ektachrome films unless minutes count. (You can process latter yourself in 37 minutes in a small tank if you can't turn to your own photo department or a local photographer who offers E-6 processing. About as complicated as histological staining. Scott claims that benefit of swift access with Ektachrome films exacts no price whatever in quality. Man noted for candor.)

By far our most versatile b&w film because 2 to 4 range of gamma attainable with conventional developers. Therefore several more specialized films we have been making no longer needed. Since more control now in user's hands, not recommended for users who shun darkrooms for fear of either boredom or feelings of inadequacy. Longtime favorites Kodak Panatomic-X film (γ limited to 1, sensitive to λ625 nm) and Kodak high contrast copy film (sensitive to λ625 nm) not discontinued. But now this Kodak technical pan film (Estar-AH base) SO-115 gives reasonably flat response to λ680 nm. Available also in 150-ft length and 4" x 5". Arousing too much customer enthusiasm to be held down long to "SO" status.

If we gave a phone number here, the temptation to phone first would be too strong and chances of getting through to whichever of these three gents can best answer your questions would decline further. It is therefore sparing of phone tolls and patience to see how many of those questions are answered on data sheets you will get by writing Kodak, Scientific and Technical Photography, Rochester, N.Y. 14650. Then phone if you still need answers. One or the other of the above gentlemen will be available in the flesh August 1 through 9 at the International Congress of Electron Microscopy in Toronto.

Kodak

SCIENCE AND THE CITIZEN

The Dossier Problem

The advantages of a Federal computer network that would link the thousands of computers operated by Government agencies, with attendant benefits in economy and efficiency, were recognized by Congress as early as 1965. Public concern about the ability of computer systems to provide adequate protection for personal information about U.S. citizens has been reflected in Congressional action, however, and no such network has been established. In the mid-1960's a proposal to set up a National Data Center was rejected, as were a joint General Services Administration–Department of Agriculture project called the Federal Information Network in 1975 and a computerized Tax Administration System early this year. The number of computers in Government agencies grew from 2,412 in 1965 to 11,328 by last January; the detailed information stored for those computers, particularly for those operated by such agencies as the Internal Revenue Service, the Veterans Administration and the Social Security Administration, would constitute a formidable dossier on each of millions of Americans if it were centrally accessible. It is not centrally accessible: an I.R.S. computer cannot, for example, call up a war veteran's history of psychiatric treatment. And the Privacy Act of 1974 specifically precludes most interagency sharing of personally identifiable information without the individual's consent.

The economy of scale and the efficiency of shared data bases and programs implied by computer networks nonetheless remain a temptation. In a recent report to Congress the General Accounting Office considers the extent to which computer networks might be made adequately protective of private information. The G.A.O. holds that even in the absence of any large networks personal information is not now sufficiently protected against misuse, either by authorized users of a computer system who are untrustworthy or dishonest or by "malicious penetrators" of the system. Most computers in operation today were designed when the security issue was not adequately recognized. More important, even now there are no formal criteria for security adequate to guide designers and operators of computer systems.

The security problem increases with the size of systems and their interconnection. The advent of microelectronic computer elements and minicomputers has made it feasible to "dedicate" a small system to sensitive information and thus reduce the risk. In many cases, however, that is not practical or economic: computer hardware, data and communications need to be shared. In such cases it is the skilled programmer who can do the most damage with the most impunity, and to date no operating system or mechanism can "provide protection from skilled programmer penetration attacks." It is therefore necessary to isolate the system from the programmer. One way is to take away the programming capability of people with access to computer terminals, limiting them to "transaction processing," and at the same time to keep programmers away from the terminal and subject all program changes to approval by an independent testing group. Another way is to separate particularly sensitive data from other data. When data must be shared between two systems, security can be improved by making the output of one system available—on tape, for example—as the input to another system. When a number of computers are linked by a shared communications system, unauthorized access can be prevented by specific network controls. If these are considered inadequate for a given degree of protection, computer cryptography can be introduced: the data are automatically enciphered before transmission and deciphered at the point of reception. This method is expensive, however, and is recommended only for special cases.

A Government agency seeking to acquire a computer network cannot merely specify various security-related features in its request for proposals from contractors, the G.A.O. points out. It is necessary to decide on a particular level of security and spell it out formally: what information is to be protected, how it is to be recognized, who should have access to it and how those people are to be recognized. The formal statement of requirements can be represented by a mathematical model that specifies a computer program; presented with that program, individual vendors can develop the proper algorithms and computer-language instructions to carry it out with their particular products. At each stage the extent to which the security requirements are met can be verified mathematically or by testing.

The immediate need, particularly in view of the mandates of the Privacy Act, is a set of guidelines for establishing security levels, for procurement and for developing adequately protective systems. The G.A.O. calls on the Office of Management and Budget (which is responsible for overseeing the implementation of the Privacy Act) to provide such guidelines.

Esse Est Percipi

A singular feature of quantum mechanics is that the theory not only describes physical systems, such as the motion of an elementary particle, but also specifies what an observer can know about that motion. A recent series of experiments addresses an even more perplexing issue arising from the theory: the possibility that the observer's knowledge or ignorance may have some influence on the state of the particle. The experiments seem to call into question the philosophical conviction that the world exists and has fixed properties independent of any observer. In order to reconcile the experimental results with common sense, it may be necessary to postulate that a quantum-mechanical system simply has no defined state until that state has been determined experimentally.

The experiments in question are variants of one proposed in 1935 by Albert Einstein, Boris Podolsky and Nathan Rosen, and reformulated by David Bohm in 1952. Two protons are to be brought together in a configuration called the singlet state, then allowed to separate. A distinctive characteristic of the singlet state is the orientation of the intrinsic spin angular momentum that each proton must carry; the spins must point in opposite directions. If the spin of one proton is pointing "up" then the other must point "down." It must be emphasized, however, that the quantum-mechanical description of the singlet state says nothing about the actual spin of either particle.

In the experiment three pieces of apparatus are to be lined up along the trajectory of each proton. First is a telescope, which merely monitors the passage of protons and emits a signal if a suitable pair is detected on the two paths. Next is a "spin filter," which allows only protons with some specified orientation to pass; for example, it might be set so as to block all protons except those with the spin pointing up. Each path terminates in a detector, which registers one count whenever a proton reaches it by passing through the preceding spin filter.

Suppose that a proton has been noted in each of the telescopes and that the two particles are in flight toward their respective spin filters and detectors. Within the context of quantum mechanics it is impossible, even in principle, to

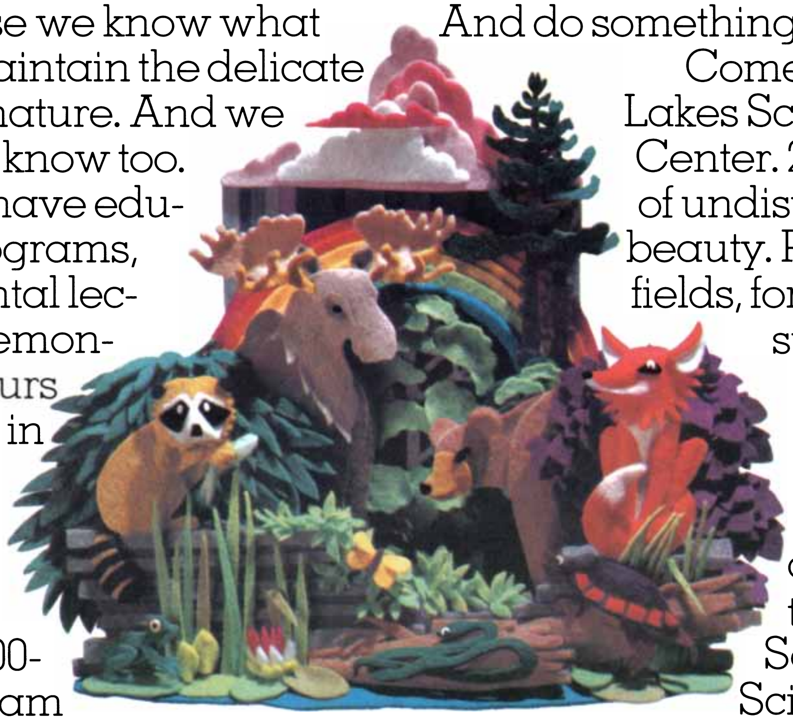
Some nice words about poison ivy, blackflies and snakes in the grass.

At Squam Lakes Science Center we have some good things to say about plants and animals most people don't think too much of.

Because we know what it takes to maintain the delicate balance of nature. And we want you to know too.

So we have educational programs, environmental lectures and demonstrations. Tours that put you in touch with nature. And more.

Walk through a 100-year-old steam powered sawmill. Spend time in an authentic blacksmith shop.



Or step into the future with our operating solar and wind energy exhibits.

Be aware of the environment. Know the environment. And do something about it.

Come visit Squam Lakes Science Center. 200 acres of undisturbed beauty. Ponds, fields, forests and streams.

To become a member or make a contribution write: Squam Lakes Science Center, Box 146, Holderness, N.H. 03245. After all, it's only natural.

Squam Lakes Science Center
Come lead the wild life in New Hampshire.

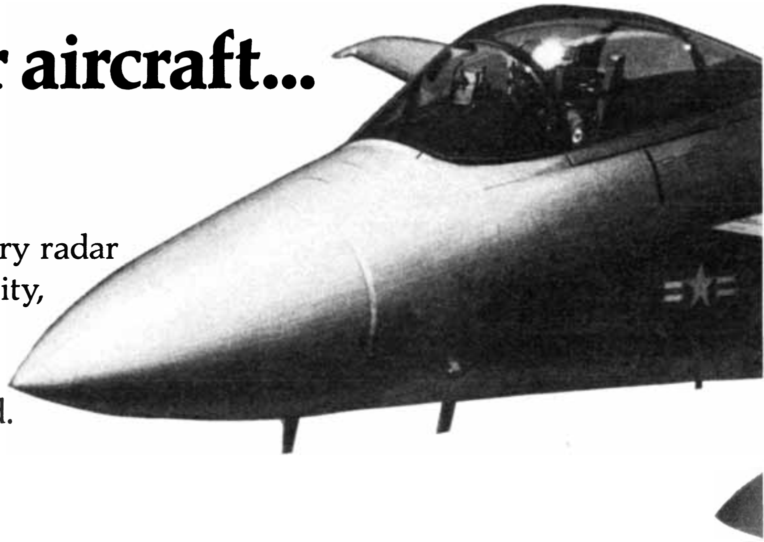
predict the spin of either proton; each one must be assumed to have a 50-50 chance of being up or down. In one arrangement of the experiment the two spin filters are oriented in the same way (up, say), but one filter-and-detector set is much closer to the origin than the other. When a proton reaches the first filter and detector its spin becomes known: if it registers in the detector then it must have been able to pass through the filter, and so it must be spin-up. As long as the second proton remains in flight its spin is formally undetermined; the experimenter, however, can make up his mind about the second proton as soon as the first one is detected. If the first one was spin-up, then according to the rules for the singlet state the second must be spin-down. The experimenter knows that when the second detector is reached, the proton will not register.

It is often said that according to quantum mechanics that the properties of a particle cannot be measured without disturbing the particle in some way. The spin of the second proton, however, seems to have been determined by an experiment that does not interact at all with that particle. The only measurement was made on another proton in another place, perhaps a quite distant one. The only apparent change resulting from that measurement was in the observer's knowledge of the system, and yet it seems also to convert the second proton from an indeterminate state to a fully determined one.

Einstein and his colleagues proposed this experiment in an attempt to demonstrate the inadequacy of quantum mechanics. It is well known that Einstein never endorsed the quantum theory, but in the 1935 paper written with Podolsky and Rosen he did not dispute the accuracy of the theory's predictions; he argued only that the theory was incomplete and should be regarded as provisional. The spin of both protons must be fully determined at all times, he maintained, and it is ambiguous in the quantum-mechanical description only because there are certain "hidden variables" whose value is unknown to the experimenter. If all the initial conditions of the experiment could be accounted for, including the hidden variables, the statistical expressions of the quantum theory would remain valid, but they could be replaced by a more fundamental, deterministic theory.

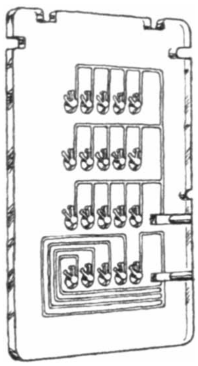
For some 30 years hidden-variables theories were investigated only sporadically, but they remained a possible successor to quantum mechanics. In 1965, however, John S. Bell of the European Organization for Nuclear Research (CERN) proved a theorem showing that no deterministic theory of the kind envisioned by Einstein could reproduce all

You'd expect McDonnell Douglas to build the world's best fighter aircraft...

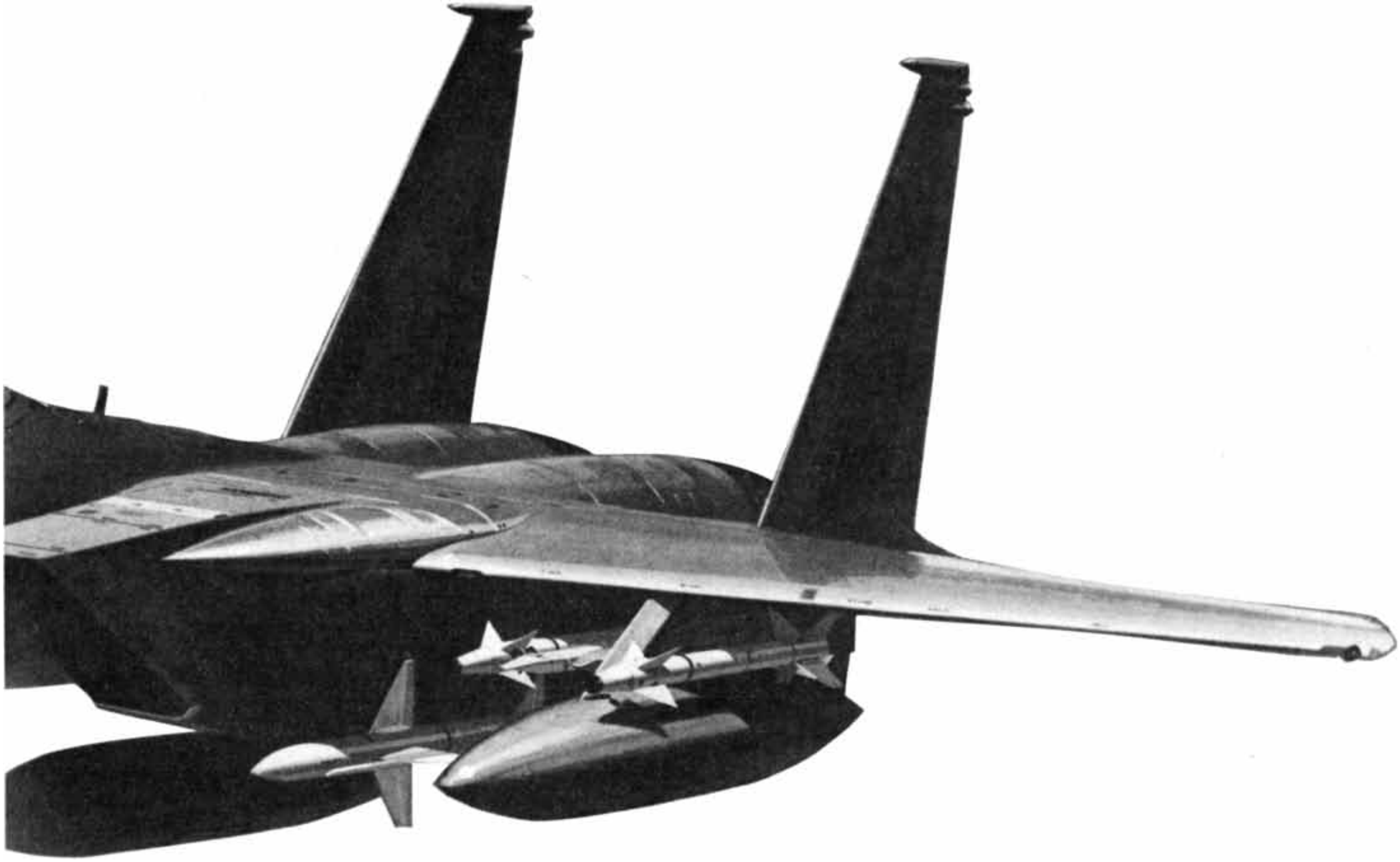


The F-15 Eagle. The peace keeper. America's air ace. With extraordinary radar and electronics, more maneuverability, it is awesome enough to prevent attack, powerful enough to win. It's the best fighter aircraft in the world. And it's keeping watch over you.

but we also help your doctor in diagnosis...



Speaking of looking after you, our Vitek Automicrobic System (AMS)® takes the process one step further. AMS is a health care system we developed to test for and identify microbes for hospital labs—in only one day or less. But more, AMS can actually test and evaluate the performance of antibiotics on the particular microbes and recommend the most efficient treatment to your physician. And, AMS delivers a complete record of the test, monitored hourly, automatically at the end of its four- to thirteen-hour cycle. Giving the attending physician what he wants to know—a new resource from space age technology to help your doctor help your health.



and make computers that talk...

Which brings us to Mavis, a computer that can talk to you and tell you what you want to know. Developed from our aircraft technology research, Mavis is a communication system we are making available to phone companies so they can make more information available to you. With the Mavis system, you can dial a predetermined number and Mavis computers can give you—in a natural sounding human voice—sports scores, the latest stock quotes, ski conditions, weather reports, just about anything you want to know by phone in your home or office.

At McDonnell Douglas, we've used our expertise in aircraft and aerospace technology to develop products and systems with real down-to-Earth benefits for you. Products that hopefully will make your life a little better, easier and even longer. These are just a few. To find out more about us and what we're doing, just write McDonnell Douglas, Box 14526, St. Louis, MO 63178.

MCDONNELL DOUGLAS

We bring technology to life.

EQUAL OPPORTUNITY IN PROFESSIONAL CAREERS. SEND RESUME: BOX 14526, ST LOUIS, MO 63178 the statistical results of quantum mechanics. Bell required such theories to meet just one additional condition (which had also been required by Einstein): they must have the property called locality, which forbids action at a distance. Bell's proof showed that the correlation of the spins expected in quantum mechanics could arise in a deterministic theory only if there were communication between the two detectors. That communication would sometimes have to be accomplished at a speed faster than the speed of light, and hence it would violate not only the locality condition but also the special theory of relativity.

Einstein intended his argument to be understood as a "thought experiment" in which ideal apparatus would be employed. Such imaginary procedures were important in the development of quantum mechanics, but most of them would be impossible to carry out in practice. The Einstein-Podolsky-Rosen experiment, however, has been adapted through certain compromises to the real laboratory. The inefficiencies of real spin filters and of real detectors make necessary additional assumptions that weaken the argument somewhat, but none of the new assumptions seems implausible.

At least seven such experiments have now been completed. Only one of them employed proton pairs; the rest employed photons, or quanta of electromagnetic radiation, which have a spin whose orientation is observed as polarization. Pairs of either protons or photons can readily be created in the singlet state. Photons have been favored mainly because their spin can more easily be analyzed.

The results of the seven experiments are not unequivocal: five support the predictions of quantum mechanics and two disagree. All of the findings are evaluated in a recent review article by John F. Clauser of the Lawrence Livermore Laboratory of the University of California and Abner Shimony of Boston University. They present arguments for attributing the two anomalous results to systematic errors in the experiments. None of the seven experiments, they point out, could be considered definitive by itself, but the weight of evidence favors the quantum-mechanical viewpoint.

If the predictions of quantum mechanics are confirmed, as is expected, what are the consequences for physics and philosophy? One way to resolve the conflict is to abandon the locality principle and allow instantaneous action at a distance, but most physicists would dismiss that possibility out of hand. The alternative is to renounce the possibility of constructing a deterministic and objective theory of atomic and subatomic physics. The spin of the second proton actually is indeterminate—the proton itself does not "know" which way it spins—until the experimenter makes the first measurement; thereafter the spin is fully determined.

There have been earlier formulations of quantum mechanics in which some role is given to the mind of the observer. The great conceptual difficulty of such theories is that there is no apparent reason for a human agency to intrude in a description of inanimate particle motion. Indeed, if a complete regress into solipsism is to be avoided, then there is excellent reason to believe particles obeyed the same quantum-mechanical rules before any human observer was present.

It must be emphasized that nowhere in quantum mechanics can the observer's knowledge alter the outcome of a real measurement. The second spin may be indeterminate until it is decided by a distant measurement, but it could not be definitely up at first and later definitely down. There is no hint of "mind over matter." This constraint on the influence of the observer suggests what may be the least discomfiting interpretation of the experiments. The interpretation has been described by Bell as a "retreat to the Copenhagen orthodoxy of quantum mechanics," set forth by Niels Bohr in his reply to Einstein, Podolsky and Rosen. Bohr maintained that the quantum theory specifies only those quantities in an experiment that are actually observed; until the experimenter has decided what quantities to measure and has acted on that decision, no experiment has been carried out. For Bohr the question of what happens before a measurement is made was simply out of bounds. In this way paradox is avoided, although a certain small realm is permanently excluded from inquiry.

Voices in the Light

For several years the Bell telephone system in the U.S. has been steadily developing and testing the capability of transmitting messages by light waves instead of the electrical and radio waves that are now standard. The basic motivation of the effort is that the capacity of a given communication channel increases with the frequency of the signal, and the frequency of light waves is about 1,000 times the frequency of the shortest radio waves. In 1976 the Bell system started testing a prototype light-wave setup in Atlanta, and last year it opened a 1.5-mile light-wave network in Chicago under typical operating conditions. These test systems have been so successful that the Bell network now plans to have its first light-wave communications link in regular service by the end of 1980.

The results of the Chicago test are described in *Bell Laboratories Record*. The light source is either a long-life aluminum gallium arsenide laser or a high-radiance light-emitting diode; they are interchangeable in the Chicago system. The light is carried on hair-thin glass fibers of high transparency. The optical cable installed for the network contains 96 such light guides. A single pair of light guides can carry 672 simultaneous two-way conversations, compared with 24 carried on a copper-wire pair.

The lasers have proved to be highly reliable. Through the middle of March the system had experienced only one laser failure in more than 100,000 hours of cumulative operation of all the lasers. (Bell Laboratories recently announced a new generation of small solid-state lasers with a projected average lifetime of a million hours, or about 100 years.) Moreover, during the first 10 months less than one second per day contained a transmission error, which means that the system was error-free 99.999 percent of the time (measured in one-second intervals). Another measure of reliability is the amount of time that a telecommunication system is out of service. The usual short-distance transmission system is designed so as to ensure that it is "down" no more than two hours per year. The light-wave system in Chicago had only 20 seconds of "down time" in its first 10 months. The results of the Chicago test, according to the *Record,* confirm the prospect that light-wave communication "has the potential of providing significant advantages to telecommunications."

Microscopy by Ultrasound

Progress in ultrasonic imaging technology, which began a decade or so ago with macroscopic applications primarily in materials science and biomedical research, is apparently on the verge of providing a versatile new tool: an ultrashort-wave acoustic microscope with a resolving power comparable to that of existing optical microscopes. One type of acoustic microscope, the Sonomicroscope, developed and manufactured by Sonoscan, Inc., is already on the market; at a frequency of 100 megahertz (100 million cycles per second) it has a resolving power of 20 micrometers, and it is said to be capable of operating at higher frequencies and correspondingly higher resolutions.

More advanced types of acoustic microscope are currently in the research stage. One of the leading experimental groups working on the problem of increasing the operating frequency and hence improving the resolving power of

Setting the Standard in Biochemistry

Biochemistry

Lubert Stryer, Stanford University

"This must be one of the best biochemistry books on the market. It is certainly the easiest to read and study. The text is lucid and digestible. The figures complement the text in a way which is not normally found in books at this level . . . For me, difficult concepts in biochemistry (those which I've always avoided) were made so simple it was a pleasure revisiting tricky grounds . . . " *—Bios*

"One can open this book almost at any page and discover a brilliant example of artistic expression used to illuminate scientific concepts. But in addition, the author has an attractive style of writing; and most important, he has something new to say." *—Biochemical Education*

"Professor Lubert Stryer has succeeded in producing an introductory biochemistry textbook that will probably be the standard text for life science, biochemistry, and medical students for many years to come."
—Yale Journal of Biology and Medicine

1975, 877 pages, 626 illustrations, 55 tables,
ISBN 0-7167-0174-X, $22.00

The Stryer slide series is a teaching aid that will enhance the student's understanding of biochemistry.

Molecular Architecture

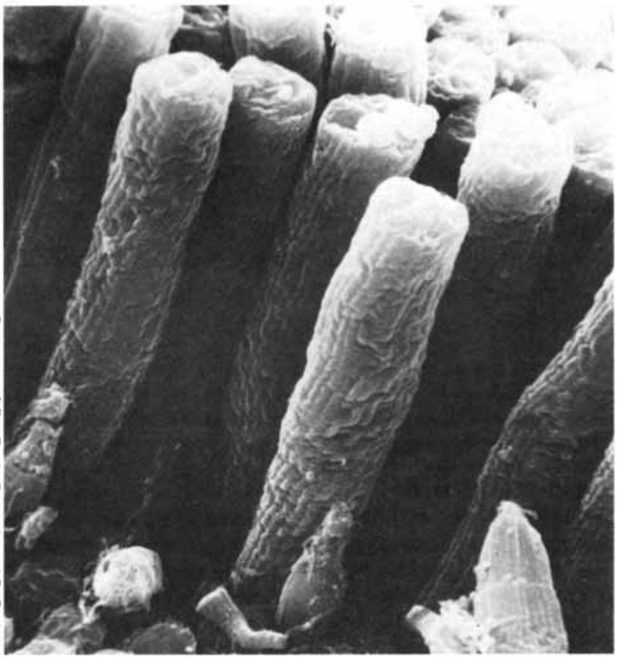
This collection of 50 full-color 35mm slides, selected from the illustrations in Lubert Stryer's *Biochemistry,* is organized to illustrate the theme of molecular architecture. Price $50.00

Energy Generation and Storage

A new addition, these 90 full-color 35mm slides illustrate processes related to the generation and storage of metabolic energy, including glycolysis, the citric acid cycle, photosynthesis, glycogen metabolism, fatty acid metabolism, the pentose phosphate pathway, amino acid degradation, and the urea cycle. Price $90.00.

Special offer to adopters

Each department that adopts 50 or more copies of Stryer's *Biochemistry* is entitled to receive both sets of slides at no cost.



Courtesy of Dr. Deric Bownds and S. Carlson

Physical Biochemistry

Applications to Biochemistry and Molecular Biology

David Freifelder, Brandeis University

With clear descriptions, a wealth of practical examples, and a de-emphasis of mathematical derivations, David Freifelder describes the wide range of techniques employed in the modern study of the physical properties of biological macromolecules, including separation and identification, hydrodynamic methods, and spectroscopy. Readers will appreciate the review of filtration techniques, pH measurement, and other basic laboratory procedures. Observational methods such as light and electron microscopy and autoradiography are also included.

Readers will come away from this book able to read, understand, and judge the work reported in contemporary scientific literature. They will also gain a better understanding of general principles, since many of the techniques discussed were originally used to establish major principles of modern biochemistry.

1976, 570 pages, 304 illustrations, 28 tables,
cloth: ISBN 0-7167-0560-5, $16.95
paper: ISBN 0-7167-0559-1, $10.95

W. H. Freeman and Company
660 Market Street, San Francisco, CA 94104
58 Kings Road, Reading, England RG1 3AA

such devices is headed by Calvin F. Quate of Stanford University. In a recent article in *Mosaic*, a publication of the National Science Foundation, Quate is quoted as saying that within a year his group expects to have pictures made with sound waves at "wavelengths equal to optical wavelengths," which would correspond to an acoustic frequency of some three gigahertz (three billion cycles per second). "When that instrument is operating," Quate adds, "all the fine detail now available in optical micrographs will also be available in acoustic micrographs."

The Stanford group's efforts to develop a satisfactory acoustic microscope started in the 1960's with a suggestion made by Rudolf Kompfner. In its current form the Stanford device, designed by Quate in collaboration with Ross A. Lemons and others, is described in the *Mosaic* article as being essentially "a scanning instrument that moves the specimen through a focused acoustic beam. Two short sapphire rods form the key elements of the instrument. They serve to focus the beam on a liquid cell holding the target specimen and to collect the acoustic energy scattered by the object. A thin, piezoelectric film of zinc oxide is fixed to the outer end of each sapphire rod, while each rod's innermost face is ground to a concave spherical lens. Electrical energy is converted into acoustic energy by the piezoelectric film, transmitted the length of the rod to the lens and focused by the lens on the specimen, which is held in its liquid cell between the concave rod ends by surface tension. The concave surfaces act as positive lenses, focusing the incoming sound beam in the liquid to a narrow 'waist' created where the focal points of the input and output lenses meet.... In theory, the waist can be made smaller and smaller, sharpening the focus more and more, simply by going to higher and higher sonic frequencies."

Because the field of view of this highly focused beam is negligibly small, the article goes on, "the object...is scanned mechanically by moving it through the waist of the beam. The scanned image is constructed and recorded, line by line, and displayed on a television screen." The main disadvantage of such a system, according to Quate, is that "it takes time to scan across a specimen.... If it is a static object, this may not matter, but if it is a living specimen and moves, then this could lead to some problems. However, the scanning time is short enough that we can still get good images from living cells."

The actual and potential applications of high-resolution acoustic microscopes range from fine-scale materials analysis to the study of biological tissues and organs. The principal advantage of acous-

This Polaroid SX-70 photograph by Ansel Adams was exhibited at a major New York gallery.

Ansel Adams made the image, "Cherubs, Mission San Xavier del Bac." It was exhibited at an Adams retrospective at New York's Light Gallery. Adams used the finest instant photographic system available: Polaroid's® SX-70® Land camera, the world's only folding single-lens reflex camera, and SX-70 film. This system is used by millions of photographers everywhere. With its unique metallized dyes, SX-70 film has the qualities important to photographers: vivid colors, sharp detail and permanence. Polaroid instant photography, selected by Ansel Adams for the expression of his art.

tic microscopy over optical microscopy for such applications, according to Quate, arises from the fact that whereas the optical microscope depends for its imaging function on the distribution of refractive indexes in the object, the acoustic microscope "depends for its contrast on the distribution of the mechanical qualities of the object: its density, elasticity and viscosity." These properties, he adds, "are far more vital to the function of living tissue than is the optical refractive index."

Shortly before his recent death Kompfner remarked: "The acoustic microscope is an important instrument that will eventually take its place with the electron microscope and the light microscope in exploring almost anything that lends itself to microscopical study.... This opinion is based not so much on the resolution that we can now get, which is quite respectable; it is based rather on the fact that we are looking at quite different properties of matter."

Sweet Hemoglobin

Diabetes mellitus is a disorder of carbohydrate metabolism that afflicts some 3 percent of the population. Usually inherited but sometimes acquired, it results from the insufficient manufacture and secretion of the hormone insulin by the islet cells of the pancreas. The lack of this important messenger chemical renders the body incapable of utilizing the sugar glucose as an energy source or of storing excess glucose as glycogen (animal starch). As a result, although glucose can accumulate in the blood to toxic levels, the tissues are unable to consume enough of it to meet their energy needs. Diabetes therefore creates the paradoxical situation of starvation in the midst of plenty.

Although diabetes has no definitive cure, the ideal therapy—replacement of the missing hormone—has been available since Frederick G. Banting and Charles H. Best of the University of Toronto Medical School isolated insulin from the pancreas in 1922. Nevertheless, the difficulty of monitoring glucose levels in diabetic patients and the almost ubiquitous long-term complications of the disease continue to challenge clinicians. Diabetes of many years' duration may be accompanied by the accelerated formation of atherosclerotic plaques in the large blood vessels, by the thickening of the basement membrane in the smaller vessels, by pathological changes in such tissues as the kidney and the retina and by an increased susceptibility to infection and gangrene.

The possible cause of these diverse consequences has recently been illuminated by studies of a minor form of hemoglobin, the oxygen-carrying protein of red blood cells. Called hemoglobin A_{1c}, it is identical in structure to hemoglobin A (the principal hemoglobin in adults) except for the presence of two molecules of modified glucose that are covalently bound to two of the four polypeptide chains in the protein. The finding that although hemoglobin A_{1c} comprises about 5.5 percent of the total hemoglobin of normal people, it is elevated to as much as 17 percent in patients with diabetes has kindled interest in its possible clinical significance. Writing in *Annals of Internal Medicine,* Charles M. Peterson and Robert L. Jones of Rockefeller University discuss the relevance of the modified hemoglobin to the pathophysiology and control of diabetes.

Hemoglobin A_{1c} is thought to be formed by the addition of glucose to preexisting molecules of hemoglobin A, a modification that occurs either spontaneously or through an enzyme-catalyzed step. The rate of conversion of hemoglobin A to hemoglobin A_{1c} is closely correlated with the concentration of glucose to which red blood cells are exposed, suggesting that the percentage of modified hemoglobin might provide a good indication of blood-glucose levels. Moreover, the amount of hemoglobin A_{1c} reflects the mean concentration of glucose over the preceding weeks or months. Unlike the glucose assays presently in use, which are susceptible to short-term fluctuations such as those that follow a meal, changes in glucose levels of less than a week's duration fail to evoke a significant change in the amount of modified hemoglobin.

Periodic evaluation of hemoglobin A_{1c} levels could therefore provide a more objective assessment of carbohydrate control in individual diabetic patients than is currently possible, enabling the physician to better evaluate treatment regimes. An obstacle to the widespread application of this approach is the lack of a simple assay for hemoglobin A_{1c} that would be suitable for use in the clinical laboratory.

The high levels of hemoglobin A_{1c} found in diabetics also provide the clinical investigator with a conceptual framework for understanding the relation between elevated blood glucose and the long-term complications of diabetes. The fact that glucose can become covalently bound to previously manufactured hemoglobin, causing a change in its affinity for oxygen, suggests that other proteins and enzymes in the body might be structurally and functionally modified in the presence of high levels of blood sugar. Such modifications, Peterson and Jones propose, could lead to the diverse changes associated with the late stages of diabetes.

Nighttime Images of the Earth from Space

An unusual aspect of the earth is revealed in pictures recorded at midnight by U.S. Air Force weather satellites. The brightest lights on the dark side of the planet are giant waste-gas flares

by Thomas A. Croft

When the earth is viewed from space during the day, cloud patterns and geographic features dominate the scene. The evidence of man's presence, when it is visible at all, constitutes a minor modification of the picture. At night, however, particularly at midnight on a moonless night, the darkened image of the earth sparkles with the bright lights of man's creations. In addition to the transient brilliance of such natural phenomena as lightning, forest fires and the aurora, the nocturnal side of the earth is girdled with complex arrays of city lights and more diffuse bands of agricultural fires. By far the brightest of man's works, however, are the gas flares found in association with some of the world's major oil fields.

This unusual aspect of the earth has become apparent only recently as an outgrowth of the U.S. Air Force's Defense Meteorological Satellite Program. The primary mission of the spacecraft put into orbit in the course of this program is to provide imagery on which to base short-range cloud-cover forecasts. At first the pictures that were relayed electronically back to the earth for this purpose were routinely discarded after a day's use, but the waste of potential scientific data was soon recognized and steps were taken to save the better auroral images. These pictures were finally made public in 1973. Provisions were then made for the establishment of an archive of the original films at the University of Wisconsin and for the distribution of the auroral images on microfilm to interested investigators through the facilities of the Department of Commerce. Most of the pictures that accompany this article were obtained in this way from the Wisconsin archive.

Characteristics of the System

The Air Force meteorological satellite is not the only system capable of supplying nighttime images of the earth from space. The three Landsat spacecraft launched by the National Aeronautics and Space Administration in 1972, 1975 and 1978 could also operate at night, although they were not designed for that purpose and have seldom been called on to make nighttime exposures. The Landsat system offers the advantage of four-color imagery and a finer ground resolution than the Air Force meteorological satellite can achieve, but the latter system is much more sensitive to faint light sources. For a study of the brighter light sources the two systems are complementary: the Air Force satellite is well suited for conducting a wide-ranging survey of the entire earth, whereas the Landsat system can provide high-resolution color pictures of specific areas selected from the survey.

The spacecraft of both of these U.S. programs are in near-circular orbits at altitudes of about 800 or 900 kilometers. The orbits are designed to miss the earth's poles by about nine degrees, making the plane of the orbit precess around the earth at the same rate as that with which the earth moves around the sun. For photographic missions such sun-synchronous orbits offer an important advantage: the time of day (or night) below the camera does not change with the passing of the seasons, and hence shifts in lighting are minimized. The practical effect of such a scheme is that the images are recorded in the form of successive strips that run approximately north-south, with each strip representing one orbit along the track of the satellite. The time of day in each area photographed is unchanged from strip to strip. The overlap between adjacent strips increases at high latitudes: at 90 degrees north and south latitude the poles appear in each strip; they are photographed alternately every 102 minutes.

One of the most striking things about many of the images obtained with this system is the light source: moonlight at midnight. A similar but more detailed picture of the earth is recorded by the satellite in daylight on the other half of its orbit. The satellite program thereby provides cloud-cover pictures anywhere on the earth every 12 hours or so, during periods when moonlight falls on the night side of the earth (about half the time). There are several other types of sensor on each of these satellites. One device makes simultaneous infrared images, relying on thermal radiation from the earth and from the clouds. The resulting films show the relative temperatures of different clouds, and they provide a means for mapping cloud cover even on moonless nights.

In images made in full moonlight, when the satellite's cameras are operated at reduced sensitivity, many man-made lights and fires are visible, but only the gas flares are bright enough to stand out clearly [*see illustration on pages 88 and 89*]. As the moonlight decreases, however, the pictures undergo a transformation much like the change that takes place in the evening sky when the sun sets and the stars come out, except that here the cities and the fires come out! For example, consider the satellite image on the opposite page. To the north the boot of Italy stands out clearly in the lingering moonlight, and it is easy to locate many of the major cities of Europe. The brightest spots in this strip are gas flares coincident with known oil fields in Algeria, Libya and Nigeria. The fairly uniform band of small lights sprinkled across Africa south of the Sahara appears to originate with fires set deliberately to clear agricultural or pastoral land. Many similar fire patterns are seen elsewhere in the world.

Several shortcomings of the Air Force meteorological-satellite system are apparent in this picture. The details near the edges are quite fuzzy, owing to the low altitude of the spacecraft, which is only about a third as high as this image is wide. A redesigned scanning device now flying on the Air Force meteorological satellite contains features that improve the resolution of such images near the edges.

The imaging system that produced these pictures works as follows. The image originates in an optical telescope that scans the earth's surface in an "across track" direction with a constant angular resolution. (In nonpolar regions the scan is roughly east-west.) At the focal plane of the telescope is a light sensor, which resembles a photographer's light meter except that the field of view is only two-tenths of a degree. The edges of the scene are far from the satellite and are viewed obliquely, so that the telescope's angular field of view covers a surface area some six or seven times larger there than the area covered at the nadir (that is, with the camera pointing straight down). Data from this system are sent to the ground in the form of a stream of binary numbers representing the variations in brightness recorded by the scanning sensor. Later, on the earth, the numbers are converted back into light spots and the images are reassembled on film in a conventional maplike form. In the process some approximate corrections are made for distortions, which can result from a variety of sources. (The orbit is not exactly circular, the earth is not exactly spherical, the spacecraft is not always at the same altitude, and so forth.)

Some of the larger gas flares are so bright that the internal reflections within the body of the telescope become visible; this effect accounts for the rings around the gas flares in some of the images. Similar artifacts appear in astronomical telescopes, which are also designed to detect objects that are very small and yet very bright in relation to their surroundings. Except in conditions of extreme contrast, such as the contrast that exists between the bright flares and their dark surroundings, this artifact causes no problem for the imaging system of the satellite.

From a comparison of such images with their infrared counterparts, which are exactly the same size and easy to match by overlaying the two films on a light table, one can see that the city

THREE MAJOR LIGHT SOURCES associated with human activities are visible in this nighttime satellite image, made recently in the course of the Air Force's Defense Meteorological Satellite Program. The numerous bright spots arrayed across the upper third of this picture are the city lights of Europe. The larger isolated lights near the middle and bottom arise from gas flares at oil fields in Algeria, Libya and Nigeria. (The rings around the brighter flares are caused by internal reflections within the body of the telescope.) The uniform band of smaller lights scattered across Africa south of the Sahara appears to originate with agricultural and pastoral fires.

lights penetrate thin clouds with little degradation. Moderately heavy clouds diffuse the image, whereas very heavy clouds can completely block the light.

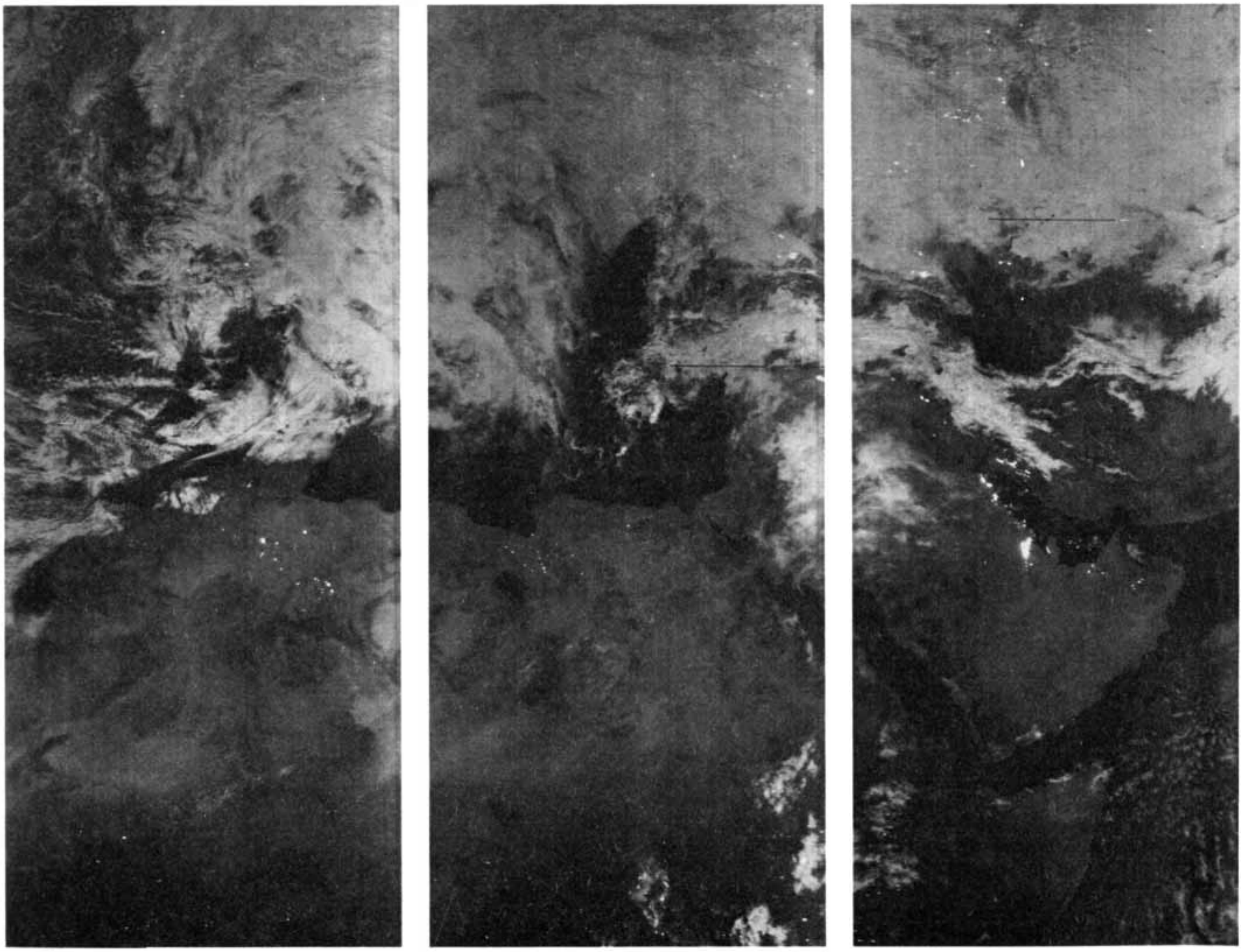
Gas Flares

The nighttime satellite images show bright gas flares in many parts of the world, but by far the greatest concentration of them is in the vicinity of the Persian Gulf. Even in full moonlight this area is aglow with numerous large flares [*see bottom illustration on page 90*]. The burning of waste gas in the Persian Gulf oil fields has been a spectacular sight in that part of the world for decades and was reported by aviators flying over the region as early as 1933. The practice of venting unwanted gas from oil wells to the air without burning goes back even further; more than 100 years ago an article in *Scientific American* reported such waste from early oil wells in Pennsylvania; the writer estimated that one well released a million cubic feet of gas per hour and had been doing so "for years."

Since gas flares are the most conspicuous features of these nighttime images, they warrant closer attention. The imaging system of the Air Force meteorological satellite is designed to measure only the average brightness over each of its picture elements, called pixels, which are roughly three kilometers on a side. (The exact size varies from place to place on the image, and from sensor to sensor.) The contribution of a single flame to this average brightness must therefore depend not only on its intrinsic brightness but also on the percentage of the area of the pixel that is occupied by the flare's image. In other words, the flares must be either very bright or very big to show up so clearly on the images. It turns out that they are both. The large areas covered by the flares result from the practice of spreading the fire laterally to ensure complete combustion. Usually this is done by dispensing the gas from a series of standpipes placed in a row 50 or more meters long [*see top illustration on page 91*].

In an effort to analyze more closely two of the larger gas flares in Algeria seen in the illustration on the preceding page, I obtained Landsat photographs of the same region made during the day [*see bottom illustration on page 91*]. Two flames are clearly visible in several of these photographs, with a north-south separation of about 11 kilometers. In terms of the Air Force satellite's imaging system the two flares are three pixels apart.

By scrutinizing enlarged transparencies of the Landsat photographs under a magnifying glass I was able to determine that the red flames are some 500 meters across, much larger than the flares usually seen in ground-level photographs. In daytime satellite imagery such gas



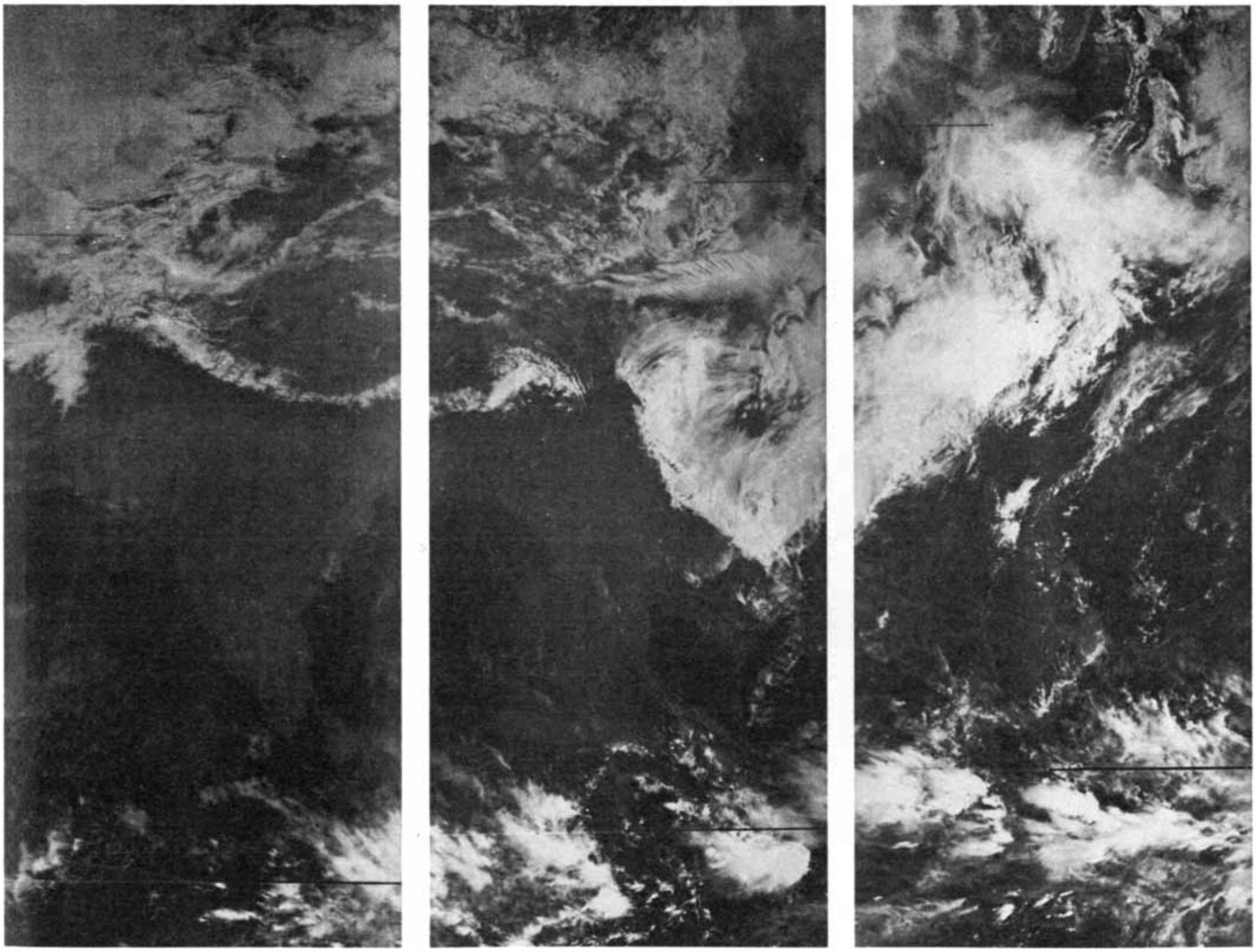
MOONLIT PANORAMA of an expanse of earth stretching from northwestern Africa to southeastern Asia is composed of parts of six images obtained on six successive north-south orbits by the Air Force meteorological satellite. All of the pictures were made at local midnight on February 6, 1974. The primary illumination was provided by reflected light from the moon. The overlap between adjacent strips

flares are easy to spot, since their smoke trails extend many tens of kilometers downwind. In an attempt to avoid giving such graphic evidence of pollution, the more modern gas flares are equipped with "smokeless" burners. This practice is no doubt justified from the standpoint of appearances, but in view of the concern about the rising level of carbon dioxide in the atmosphere, mankind might be better off in the long run if some of the carbon particles in the smoke were left unburned.

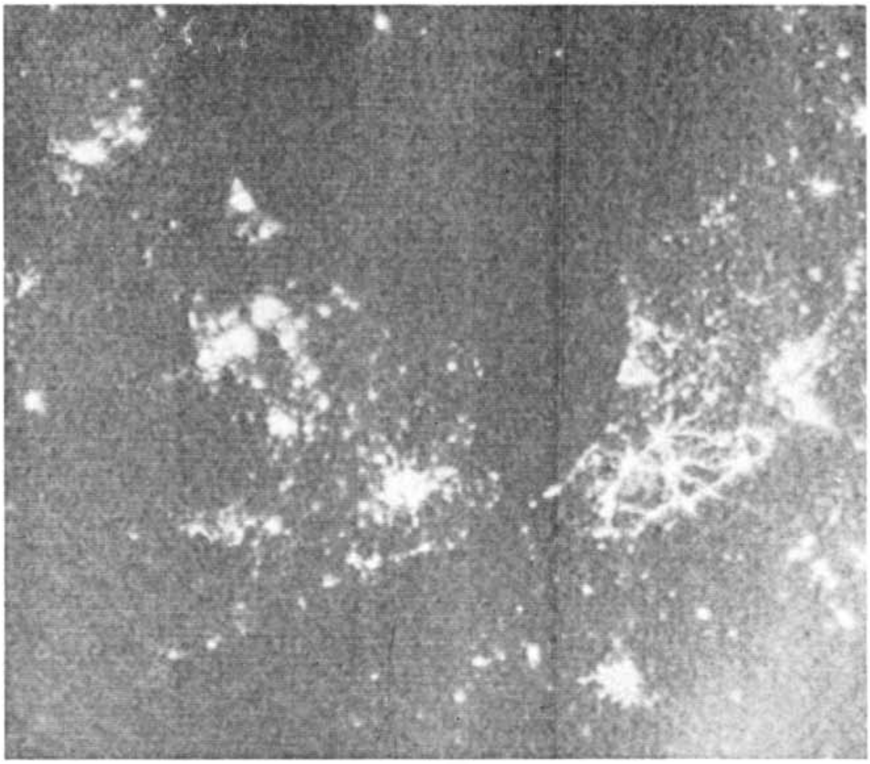
A more precise estimate of the lateral extent of these flares can be obtained from nighttime Landsat images. For this purpose one of the Landsat spacecraft was activated over Algeria on the night of January 31, 1976, at the request of the U.S. Geological Survey. An extreme enlargement of one of the flares photographed on this occasion is shown in the top illustration on page 92. The flare corresponds to the larger, more southerly one shown in the daytime Landsat picture at the bottom of page 91.

To understand what this particular nighttime image shows, it is necessary to know something about how it was made. As the Landsat satellite passes overhead, its multispectral scanning telescope looks downward by means of a 45-degree mirror that rocks back and forth to scan the scene from left to right (that is, in the across-track direction). Patterns of light falling on the focal plane of the telescope are conveyed to electronic sensors by way of tiny glass light pipes, each of which has a square cross section corresponding to a square 79 meters on a side on the earth's surface at the nadir. (At the edge of the image, 93 kilometers from the projection of the satellite's track on the earth, each light pipe covers a slightly larger area, so that the stated resolving power of the system is usually rounded off to 80 meters.) The electronic sensors are "read" once for every 57 meters of the scan; hence the Landsat pixels are rectangles measuring 57 meters by 79 meters at the nadir. A large bright object can affect two or three consecutive sensor readings, partly because the light pipes are larger than the 57-meter pixel and partly because of a lag in the electronics. (A few months ago the Russian delegates to the United Nations suggested that 50-meter pixels should be the smallest allowed; anything finer would be considered a violation of a nation's right to privacy.)

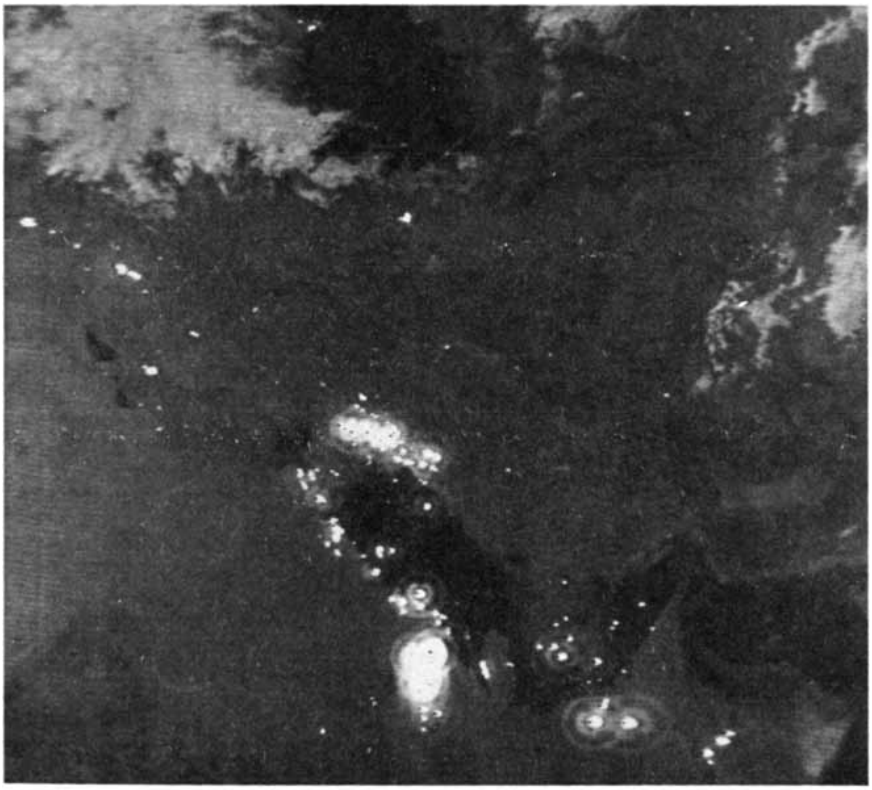
Within each pixel the Landsat imaging system obtains four separate sensor readings, one for each of four spectral ranges, or colors. (The term "color" is used loosely here; two of the ranges are actually in the infrared and near-infrared regions of the spectrum and are therefore invisible to the human eye.) For each color and for each pixel the Landsat sends to the ground a separate



increases at high latitudes. In this case the strips were cropped at top and bottom; the originals include both poles in each strip. Although some city lights and agricultural fires are visible in such brightly moonlit scenes, only the gas flares are bright enough to stand out clearly. An enlargement of the bright cluster of waste-gas flares in the Persian Gulf area is shown in the bottom illustration on the next page.



CITY LIGHTS of the countries bordering the English Channel and the North Sea show up with unusual clarity in this enlarged Air Force satellite image, in spite of the comparatively poor contrast. The picture was made on a moonless night in May, 1977, when the satellite's sensors were operating at peak sensitivity. The British Isles are at left. Network of lights at center right is in Belgium. Well-lighted night spot near the edge of the image at lower right is Paris.



GAS FLARES are concentrated conspicuously in the Persian Gulf area, as evidenced by this enlarged moonlit image recorded by the Air Force satellite's optical system. Dark area at top is Caspian Sea; bright spot on promontory is Baku in the U.S.S.R.; larger spot due south of the Caspian is Tehran, Iran; similar spot near large lakes at center left is Baghdad, Iraq.

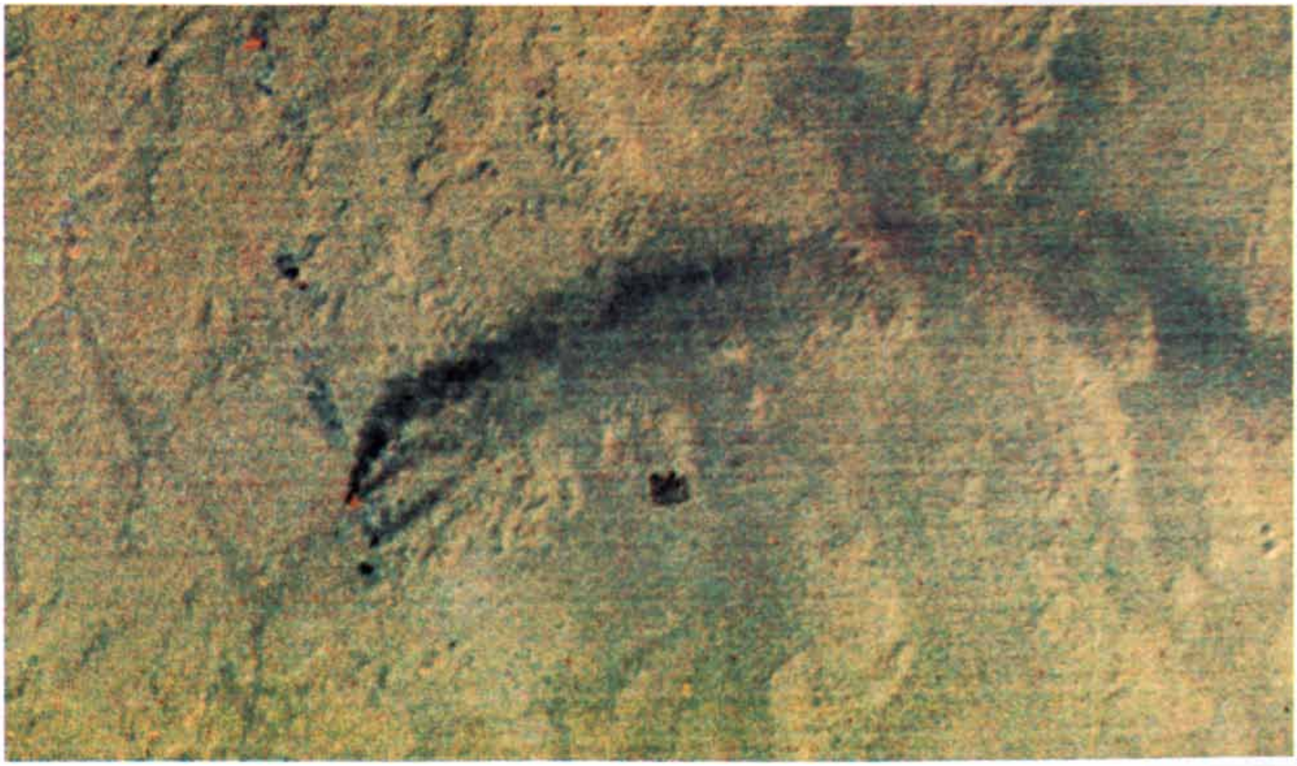
binary number coded to represent the power level of the light. Since there are only 35 illuminated pixels in the particular enlargement under consideration here, it was feasible to produce each small colored rectangle individually by ordinary photographic means. To make the infrared and near-infrared data visible I adopted a "false color" method of reproduction, shifting each digital power level toward the shorter-wavelength end of the spectrum by about a fifth of a micrometer. Thus the original infrared reading is reproduced as red, the near-infrared as green and the red as blue. (There was practically no green in the original image.)

A standard Landsat picture is made up of 7,581,600 pixels, each with four separate spectral readings. It follows that a single daytime image represents more than 30 million numbers sent down from space. No simple photographic method could be practical for reassembling such a color image on film. In fact the system now in operation for making Landsat films does not even use light. Instead the film is held in a vacuum and exposed to direct bombardment by an electron beam. This approach enables one to expose very fine lines on the film, with each line corresponding to an across-track string of contiguous pixels. Color composites are then made by mechanically superposing the black-and-white films, each of which represents one of the original spectral bands. The reconstruction of the Air Force satellite program's images is done in much the same way, except that a light beam is used instead of an electron beam.

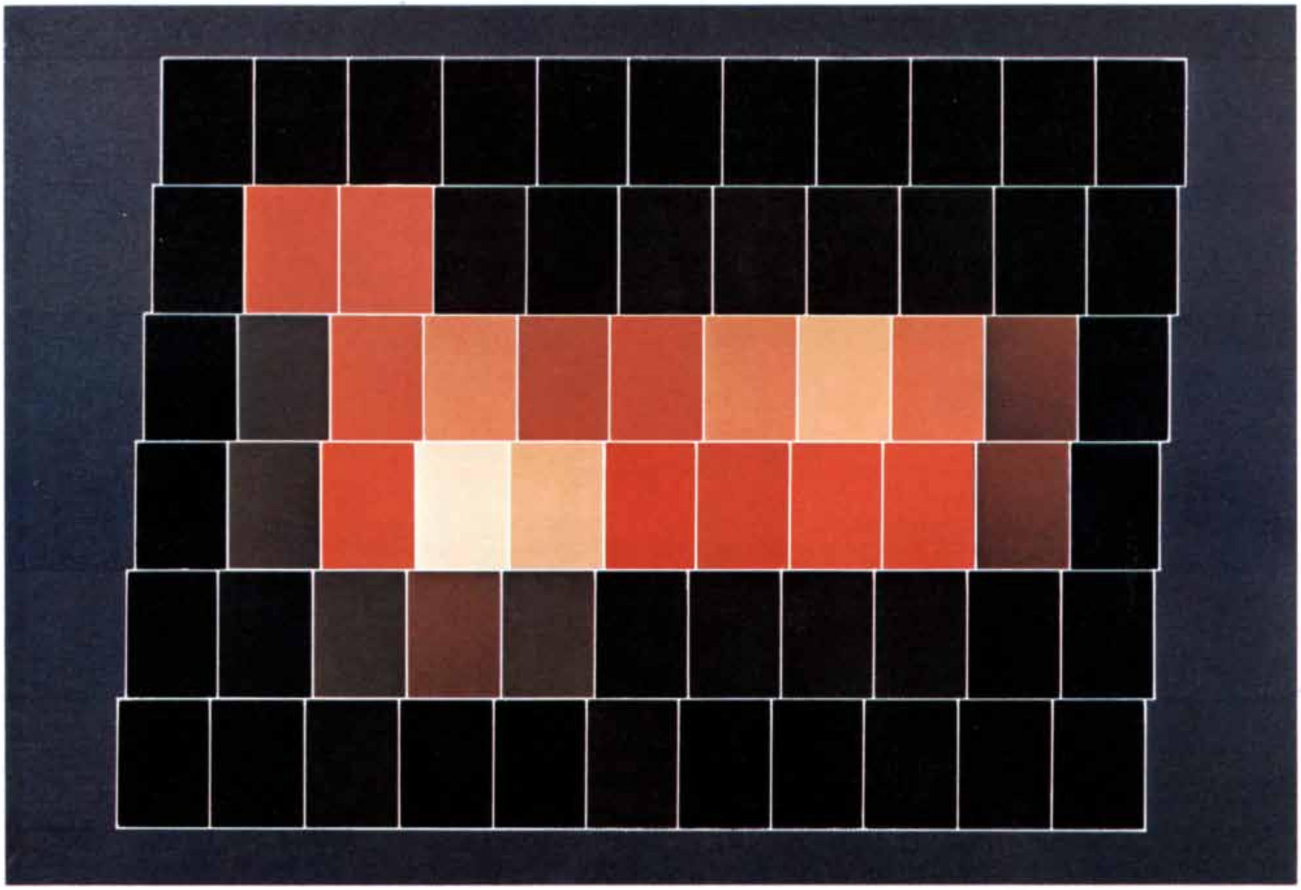
With this understanding of the basis of the Landsat imagery in mind, let us reexamine the top picture on page 92 and try to visualize the flame that produced this light. Many of the pixels along the border of the image reveal only a weak light, which is what one would expect if only a wisp of flame intruded into that rectangle. Although the overall image is about 500 meters long, it appears that it might be caused by two flares 200 meters apart. On the other hand, the dark center might be smoke obscuring part of the flame. (The flames of this same fire in the daytime Landsat image did not appear to be separated.) The matter could probably be resolved by examining several more Landsat images of the same site.

From a comparison of the larger-scale Air Force images with published charts of the world's principal oil and gas fields it is clear that many such flares coincide with oil fields but few coincide with gas fields. This is not surprising; a gas well can simply be capped if the gas cannot be sold, whereas an oil well unavoidably releases gas as the oil is extracted from the ground. Oil often reaches the surface in the form of a

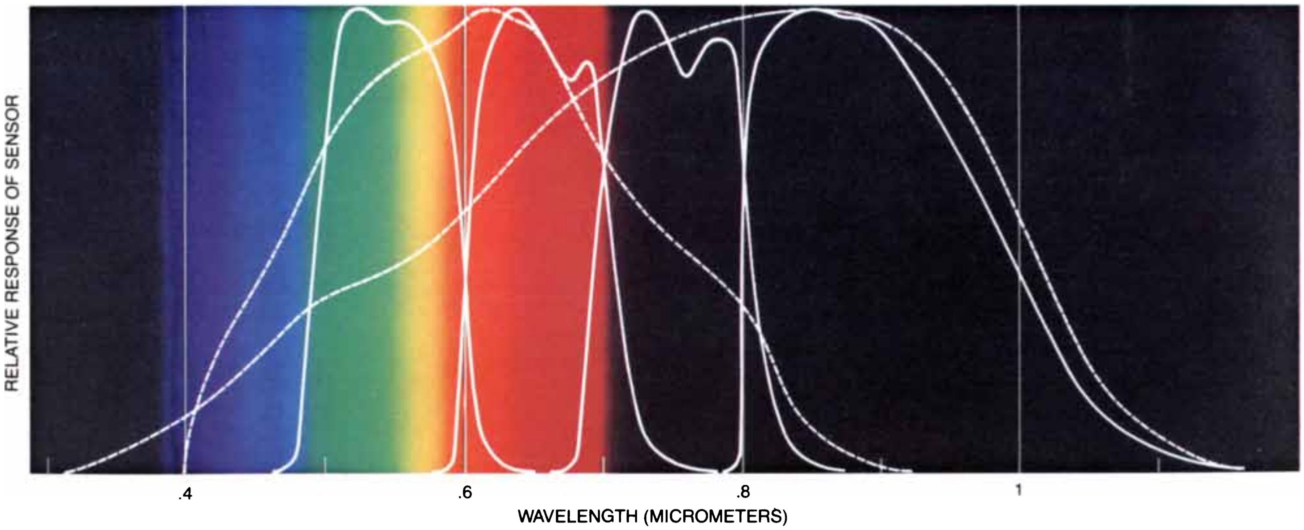
GAS FLARES ARE SPREAD over a large area in remote regions in order to ensure complete combustion. This photograph, made from the vantage point of a nearby drilling rig, shows such a flame at an oil field in Kuwait. The standpipes used to dispense the gas are roughly eight feet tall. The photograph was made by David F. Cupp and appeared originally in the May 1969 issue of *National Geographic.*



FOUR-COLOR SATELLITE IMAGE of two large gas flares in Algeria was obtained in daylight by the *Landsat 1* spacecraft, operated by the National Aeronautics and Space Administration. The flames of the burning gas are seen here as red splotches at the head of long smoke plumes. A greatly magnified nighttime Landsat image of the more southerly flame of the two appears at the top of the next page.



BLOWUP of an extremely small portion of a nighttime Landsat image containing one of the Algerian gas flares shown in the bottom illustration on the preceding page was made by the author in an effort to estimate the lateral extent of the flare. Every Landsat image is composed of millions of tiny picture elements, called pixels, that correspond to rectangular areas on the earth's surface 79 meters long and 57 meters wide. Within each pixel the Landsat imaging system obtains four separate sensor readings, one for each of four spectral ranges. In this case there were only 35 illuminated pixels in the enlarged image, so that it was feasible to produce each small colored rectangle individually by ordinary photographic means. A "false color" method of reproduction was adopted in order to shift the normally invisible infrared and near-infrared ranges of the Landsat data into the visible part of the spectrum. The resulting image suggests that the overall length of the gas flare was about 500 meters, although the possibility remains that it was composed of two flares 200 meters apart.



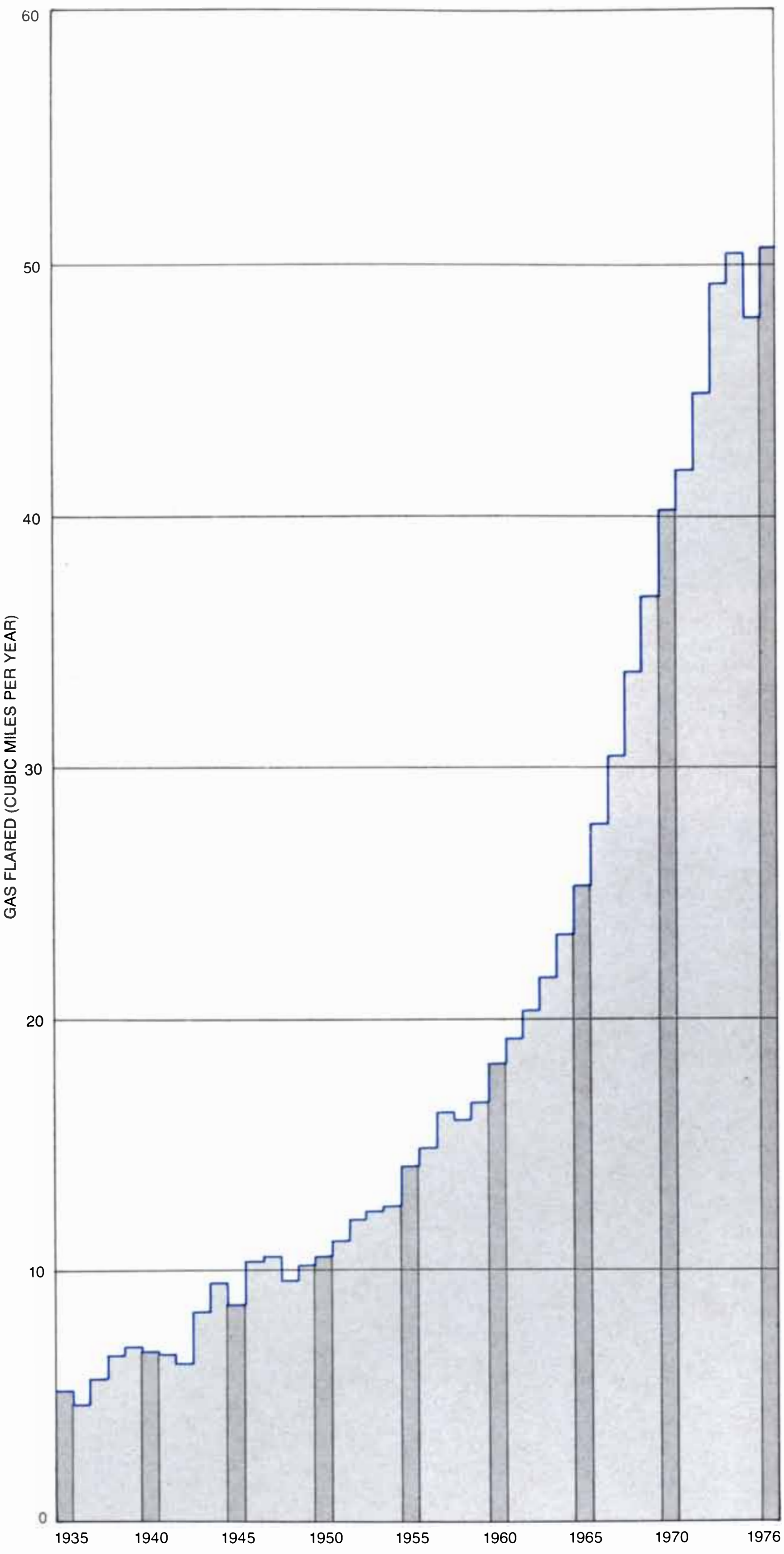
SPECTRAL RANGES of the Landsat multispectral scanning system and two versions of the sensor carried on board the Air Force meteorological satellites are compared here. The six curves are all shown with the same peak value, although in actuality both the old Air Force sensor (represented by the broken white curve on the right) and the new Air Force sensor (represented by the broken white curve on the left) are more sensitive than any of the Landsat sensors (represented by the four solid white curves). The band of colors running across the graph from top to bottom in the wavelength range from approximately .4 to .7 micrometer corresponds to the visible portion of the electromagnetic spectrum. A comparison of the curves of the satellite sensors with this visible range shows why both spacecraft are particularly effective at detecting fires: much of the energy from fires is transmitted at wavelengths to which the human eye is insensitive.

frothy mixture with various gases, and the oil must settle before it can be piped away. The associated gas must be disposed of in some way, and in remote areas it may be more economical to burn the gas than to transport it to a buyer. Government policymakers might well ponder the significance of these pictures in setting domestic price controls on natural gas. The artificial depression of gas prices in the U.S. and other developed countries can be viewed as an encouragement to such waste in the remote oil fields of the less-developed countries. Presumably the economic dilemma of the policymakers in the controlled economies of the world is similar, since large gas flares can be seen in association with oil fields in the U.S.S.R. and other Communist countries.

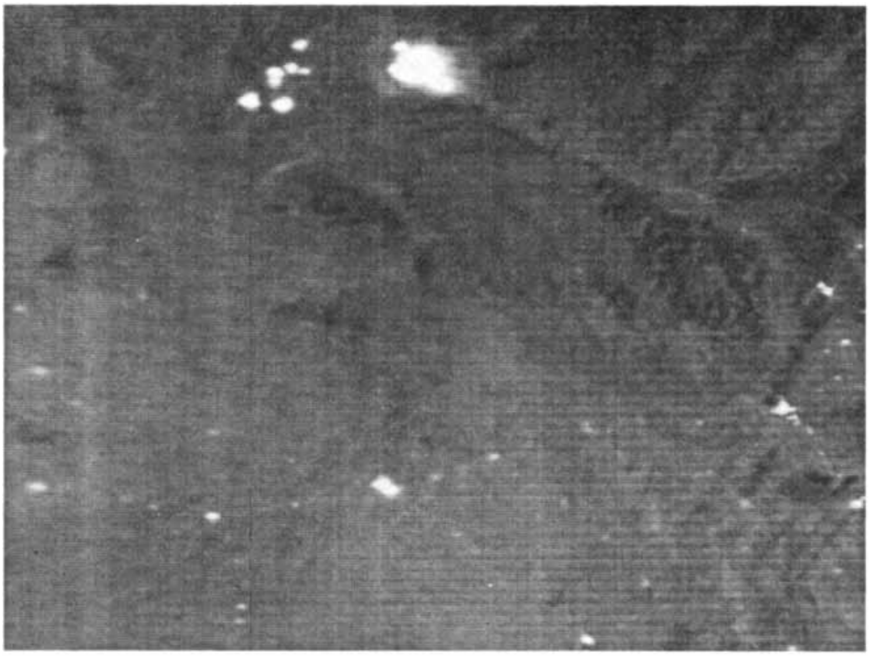
Since mankind has been venting or burning such gas for more than 100 years, it is only natural to ask how much has been wasted in this way. The answer is difficult to determine, since the parties most directly involved have little motivation to be frank about the details. In spite of these difficulties Ralph M. Rotty of the Oak Ridge Institute for Energy Analysis has compiled annual estimates of the total world volume of flared gas from 1935 through 1976 [*see illustration at right*]. His figures show that the amount of gas flared has grown at a steady pace that exceeds the growth rate of a simple exponential model, at least prior to the recent rise in oil prices. By adding these yearly estimates, and extending the numbers forward to 1978 at a fixed rate and backward to the discovery of oil in 1859 at an exponential rate, I have arrived at a total volume of wasted gas (at atmospheric pressure) of 4,200 cubic kilometers, or roughly 1,000 cubic miles. Rotty estimates that in recent years flare gas accounts for about 3 percent of all the hydrocarbons burned by man. The nighttime satellite observations I have described here might offer an independent means for reappraising this estimate of the rate at which gas is currently being flared in the world.

Japanese Lanterns

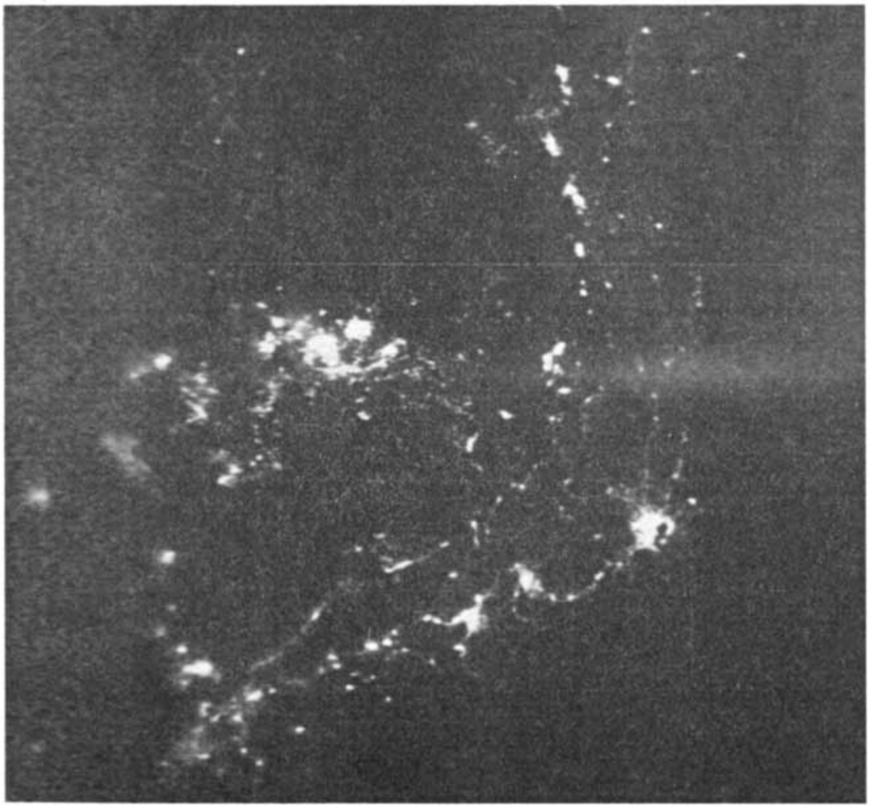
One of the strangest sights on the dark side of the earth is seen in the satellite image at the bottom of the next page. On the right is Japan, one of the world's brightest nations at night, owing to its dense population and advanced state of development. On the left, in the middle of the Sea of Japan, an area that is normally black on such a moonless night, there is a great swirling mass of bright lights. The lights seem to be arrayed in an orderly fashion, and their pattern changes from night to night. It has long been known that this region is subject to severe thunderstorms accompanied by extraordinary displays of light-



ANNUAL ESTIMATES of the total volume of gas (at atmospheric pressure) flared in the world's oil fields between 1935 and 1976 were compiled by Ralph M. Rotty of the Oak Ridge Institute for Energy Analysis. Using rough approximations to carry these figures backward to the discovery of oil in 1859 and forward to the present, the author has arrived at a grand total of 1,000 cubic miles for the volume of gas wasted in this way in a little more than a century.



GIANT GAS FLARE IN SIBERIA is seen in this nighttime satellite image, made in January, 1975. The flare, one of the biggest yet detected by the author in his examination of the Air Force pictures, is located near Surgut in northern Siberia. The large oil field known to exist in this remote region apparently lacks a gas pipeline to places where more productive use could be made of the waste gas. The route of the Trans-Siberian railroad is marked by the regularly spaced large and small cities strung out like a necklace across the bottom of the picture.



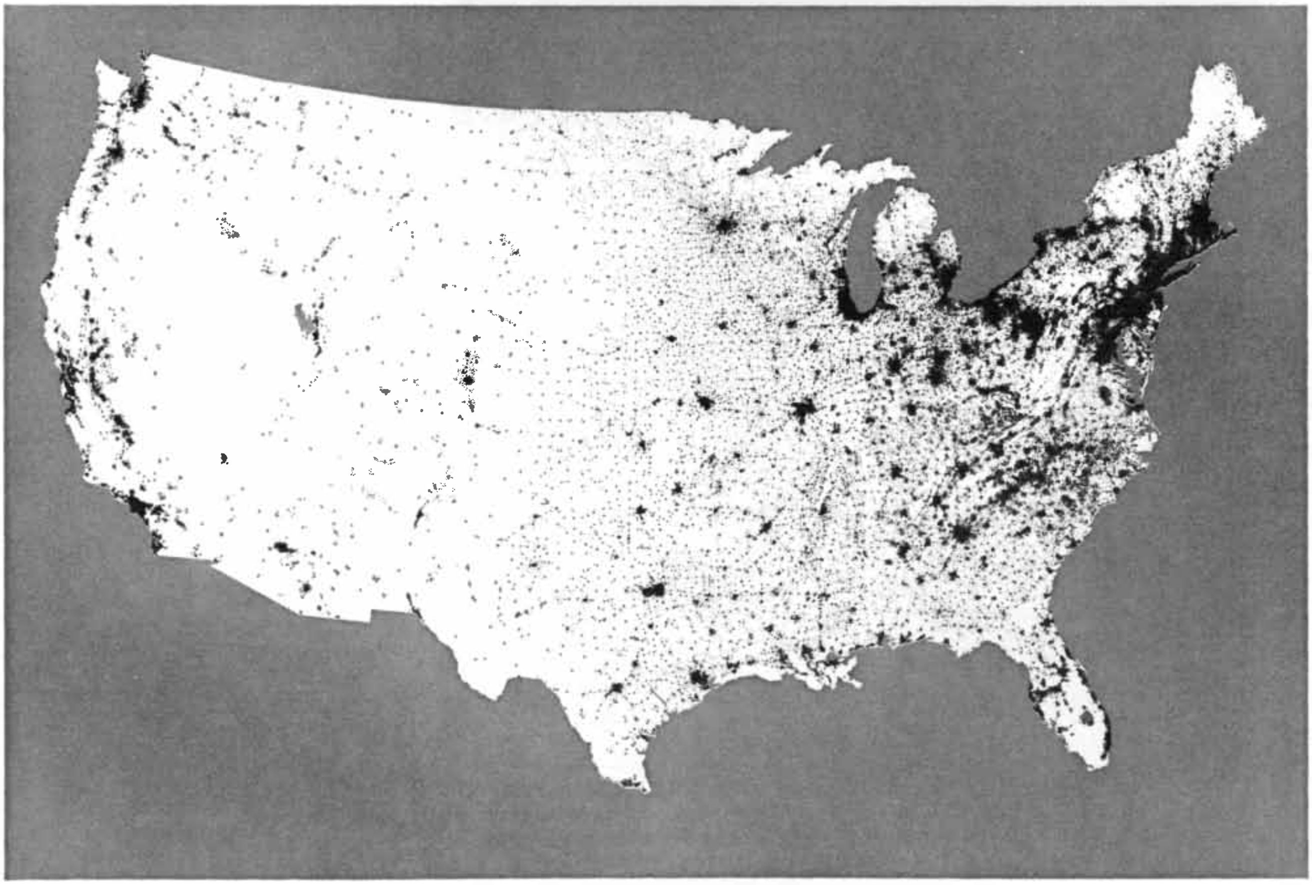
BRIGHT LIGHTS IN THE SEA OF JAPAN appear at left in this Air Force satellite image, made on a moonless night in June, 1975. The lights coincide with the known position of the Japanese squid-fishing fleet at this time. The several thousand boats of this fleet are strung with incandescent bulbs totaling hundreds of millions of watts, which serve to attract the squid to the surface. Japan itself is sharply delineated by city lights of its densely populated coastal regions.

ning. In response to my inquiry, however, the Japanese Meteorological Agency reported that there were no major thunderstorms in the Sea of Japan on the night this image was recorded. A second inquiry brought a more satisfactory answer. Scientists at the Japanese government's fisheries bureau sent me charts showing the location of the squid-fishing fleet on the date in question. The match between the charts and the satellite images was unmistakable.

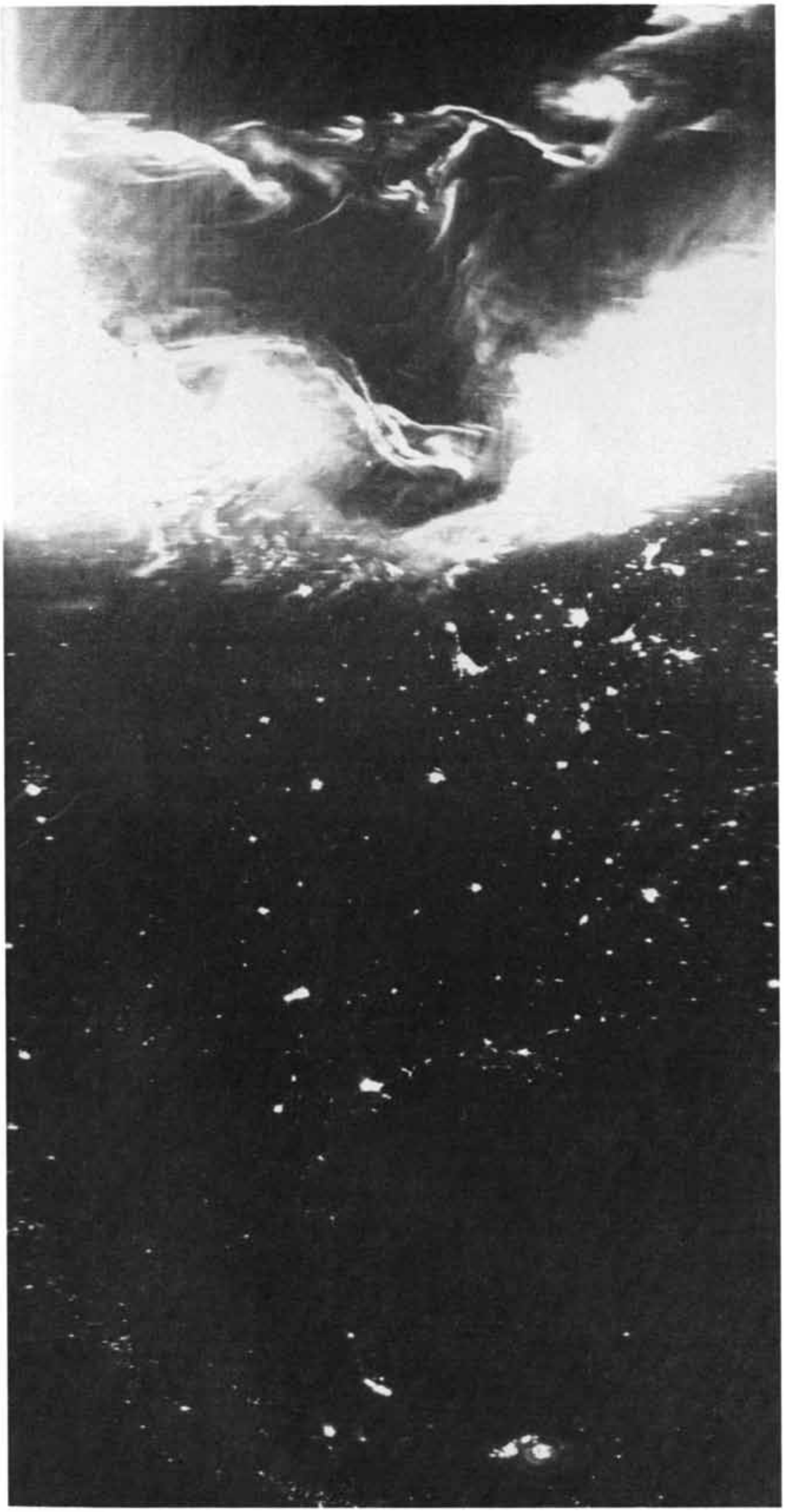
I was informed that during the squid-fishing season vast fleets of both Korean and Japanese boats operate in the Sea of Japan. The Japanese fleet alone is composed of some 2,600 boats, each weighing between 60 and 100 tons and each carrying as many as 50 incandescent lamps with an average power of 3,500 watts per lamp. That adds up to a total of more than 400 million watts of electric light! The lamps are used to attract the squid to the surface, where they are netted. Each boat strings its bulbs in a double row about two or three meters above the deck. Half the bulbs have no shades, and the other half have only small shades. On the average only about half of the Japanese fleet fishes at any one time, so the total emitted light might be only on the order of 200 million watts. Even so, it must seem like daylight out there. Moreover, in addition to these boats there is the smaller Korean squid-fishing fleet and a lighted Japanese fleet in shallower waters seeking another fish: the saury. (A separate Japanese squid-fishing fleet reportedly works the waters near New Zealand, but so far I have not found evidence of their activities in the Air Force weather-satellite imagery.)

One can get a better feel for the magnitude of this wattage by comparing it to the total power sent skyward by the U.S. population in the form of visible light. A decade ago this output was measured from space and was found to be about 40 million watts—roughly a fifth of a watt per person. Making allowance for the low efficiency of the fishermen's incandescent lamps compared with the combination of lights used in the U.S. at night, I estimate that the sum of all the light detected by the Air Force satellite from the U.S. is only two or three times greater than that detected from the squid-fishing fleet in the Sea of Japan.

To be sure, the city lights of any country, including the U.S., are much more dispersed. In fact, their pattern rather accurately reflects the distribution of population. That correspondence can be readily appreciated by comparing the montage of three nighttime satellite images of the U.S. shown at the top of the opposite page with the computer-assisted census map of the U.S. reproduced at the bottom of the same page. There are only a few discrepancies; for example, two lights appear east of Salt Lake City



POPULATION DISTRIBUTION of the U.S. is accurately reflected by the pattern of city lights revealed in the montage of three nighttime satellite images at top. A computer-plotted population map, obtained from the Bureau of the Census, is reproduced at bottom.



BRILLIANT AURORA extending across Canada as far south as the Great Lakes is the most conspicuous feature of this picture, recorded by the Air Force meteorological satellite on the night of April 18, 1974. City lights of U.S. occupy middle portion of this strip. Smaller lights scattered along the western side of Mexico at lower left are agricultural fires. The large bright gas flare at lower right is associated with the Reforma oil field on the eastern coast of Mexico.

that have no counterparts on the census map.

Agricultural and Natural Fires

In many cultivated parts of the earth vegetable matter is regularly burned as part of the seasonal cycle of growth. I have already pointed out the band of agricultural fires that runs across central Africa south of the Sahara. Another good example closer to home is seen in the satellite image at the left. Since the lights scattered all along the western side of Mexico are absent at other times of the year, I surmise that they must be also caused by agricultural burning. The gas flares in this image, on the other hand, show little variation in the course of the year. The bright gas flares visible in this strip are associated with the Reforma oil field on the eastern coast of Mexico. (This particular strip also illustrates beautifully how far south the aurora can appear at times.)

Similar agricultural fires can be seen in many other parts of the world. The agricultural or pastoral origin of these fires can usually be determined by the fact that they disappear in certain seasons. Daytime Landsat images are useful for studying the broad band of agricultural burning across central Africa, since the burned areas appear blackened on subsequent passes. The total area of land cleared in this way has been measured by means of the Landsat images.

It appears that the less-developed countries are the only regions where there is much agricultural burning. Presumably the practice is related to the lack of powered machines to prepare the land for planting; as a result fire is used instead to help clear the residue of the previous season's crop. Some crops of course require fire even in circumstances where other alternatives exist. Nevertheless, it is somewhat surprising to see the extent to which fire is still used in countries that lack modern agricultural machinery. A century ago similar fires were set to help clear the vast forests that once covered the central U.S.

Natural fires are also seen in the Air Force satellite imagery, for example in northwestern Australia [*see illustration on page 98*]. Unlike the fires in cultivated lands, these fires are very large. Two aspects of the fires make them appear to be accidental: they are in remote regions not generally cultivated, and they are so large that men could probably control them only with difficulty, if at all. (On the other hand, it has been reported that the aborigines once burned the scrub in this area to maintain grazing lands for kangaroos; perhaps the practice is still followed on a large scale.) Apparently these large fires are destructive to wildlife; I have received an inquiry from the Australian Wildlife Service expressing

Radio Shack's personal computer system? This ad just might make you a believer.

You can't beat the 4K system at $599

TRS-80 "Breakthru"
- TRS-80 microcomputer
- 12" video display
- Professional keyboard
- Power supply
- Cassette tape recorder
- 4K RAM, Level-I BASIC
- 232-page manual
- 2 game cassettes

...or the step-up 16K system at $899

TRS-80 "Sweet 16"
- Above, except includes 16K RAM

...or the fast 4K/printer system at $1198

TRS-80 "Educator"
- Above, except includes 4K RAM and screen printer

...or the Level-II 16K/printer/disk system at $2385

TRS-80 "Professional"
- Above, except includes 16K RAM, disk drive, expansion interface, and Level-II BASIC

So how are you gonna beat the system that does this much for this little? No way!

...The amazing new 32K/Level-II/2-disk/ line printer system at $3874

TRS-80 "Business"
- Above, except includes 32K RAM, line printer, and two disk drives

Get details and order now at Radio Shack stores and dealers in the USA, Canada, UK, Australia, Belgium, Holland, France, Japan. Write Radio Shack, Division of Tandy Corporation, Dept. C-073, 1400 One Tandy Center, Fort Worth, Texas 76102. Ask for Catalog TRS-80.

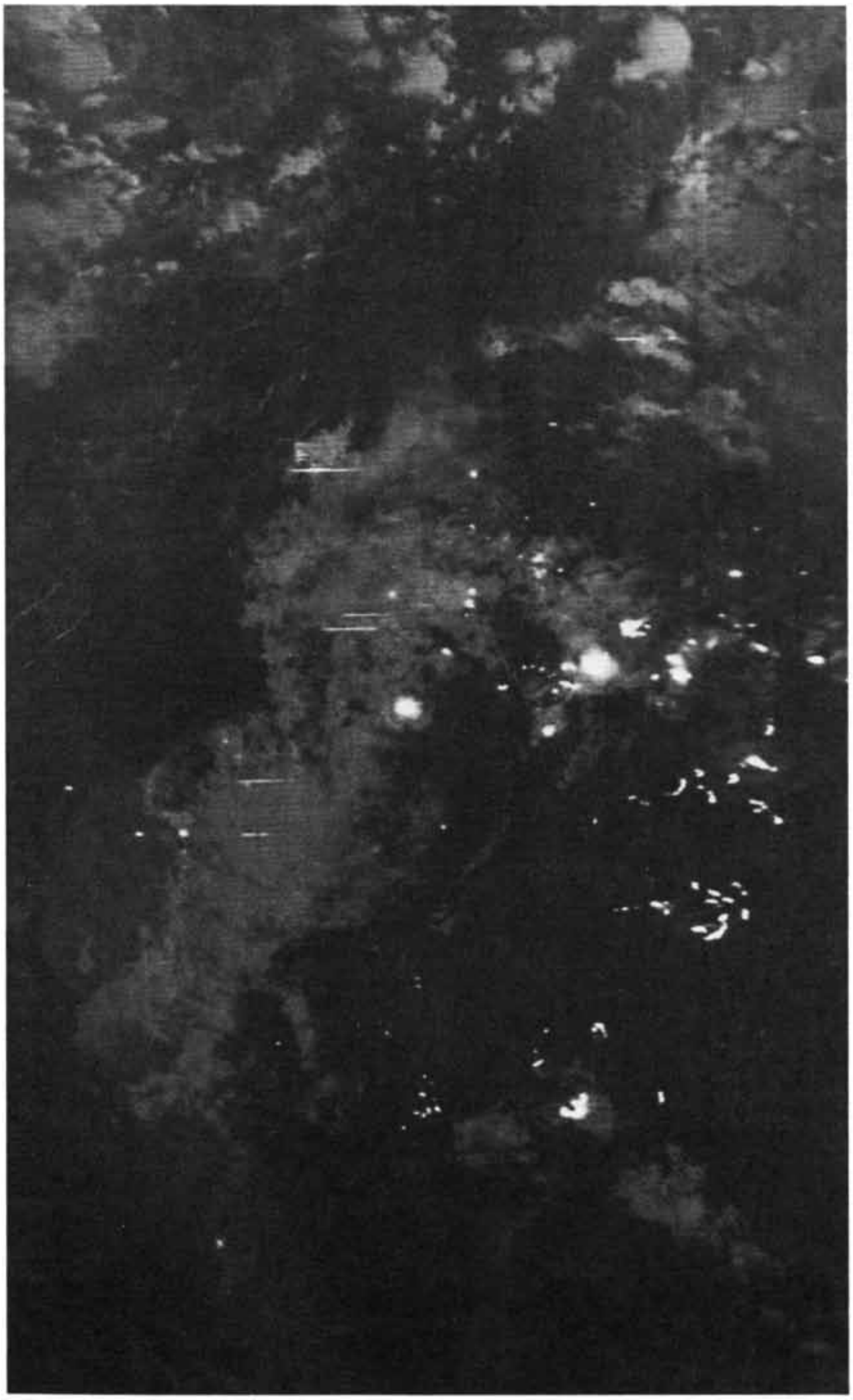
Prices May Vary at Individual Stores and Dealers

Radio Shack®

The biggest name in little computers

interest in the use of such satellite images to monitor the recurrent fires in the region. It is clear that such a project is both feasible and, given the concurrence of the observed country, desirable as a means for assessing the damage caused by large natural fires. Except for times of heavy clouds, the fires could be charted daily.

Another civilization watching the earth from space would see no evidence of an energy shortage here; indeed, nighttime viewing would reveal great waste, indicating a glut of energy. The waste arises in part from the unwillingness of the developed countries (particularly those that are well lighted at night) to depend on energy imports because of a perceived threat to their national security and economic well-being resulting from any such dependence.



NATURAL FIRES appear to be responsible for the large bright spots seen in this nighttime satellite image of western Australia. These fires are much larger than the controlled fires seen in cultivated regions. Short bright horizontal streaks visible above clouds nearby are caused by lightning. Southwestern corner of Australia is clearly illuminated by moonlight at lower left.

The city lights of the developed world and the gas flares of the less-developed world are a comparatively recent event in history. Only a century ago nocturnal images of the earth from space would have revealed neither cities nor flares and far less agricultural burning. What will the dark side of the earth look like in another 100 years? I would venture to suggest that the inhabitable regions will look much like the U.S. does today, covered with city lights but with no gas flares. Gas and oil will probably be highly prized as chemical stocks, and they will be used as fuels only in certain systems at premium prices.

Why do the operators of oil fields not stop the waste of flare gas now by installing liquefaction machinery? The answer to this obvious question apparently has three parts. First, billions of dollars are in fact being invested in gas-liquefaction equipment and in the ships to carry the liquefied gas. The speed and scale of this effort have been hampered by statements from the leaders of the developed countries indicating their aversion to the purchase of such liquefied gas for fear of becoming too dependent on it. Here again balance of payments problem and the concern for national security indirectly result in the burning of more waste gas.

A second factor also plays an important role: present-day gas-liquefaction machinery is inherently economical only when the plant is large. Small flaring centers must be connected to large centers or their gas probably cannot be saved economically. A small but economical gas-liquefaction plant needs to be designed. Perhaps these satellite images will reveal that the small flares are numerous enough to warrant a substantial investment in the development of such plants.

The third factor standing in the way of the productive use of flare gas is simply that many flares currently dispose of highly corrosive waste gas for which no economic use has yet been found.

Concerning the agricultural fires, it seems likely that this kind of burning on such a large scale removes valuable soil nutrients, casting them to the winds. Although fire is known to be necessary for the survival of some ecological systems, it seems too great a coincidence that the regions of poorest agricultural productivity are also the regions of greatest agricultural burning. Perhaps the use of fire to prepare the land for planting made more sense when people were not so numerous. Alternatives to burning that conserve soil nutrients could clearly have worldwide significance in view of the scope of the practice, as revealed in these satellite images.

SOMETIMES A PRODUCT JUST KEEPS GETTING BETTER AND BETTER.

The Kodak Carousel projector is a good example of such a product. Tried, proved, and improved over time.

It uses gravity to gently drop each slide in place. It's whisper-quiet, too.

It features a curved-field projection lens. For sharp edge-to-edge projection.

It uses the Transvue slide tray that lets you see the number of the slide shown, even in the dark.

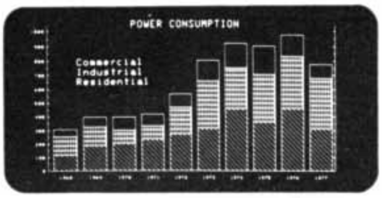
It comes in many models. So you can choose just those features you want. Such as remote control. Automatic focus. Even automatic slide changing at timed intervals. And there is a wide variety of accessories made for it. Everything from special-purpose lenses, trays, stack loader to special "presentation aids."

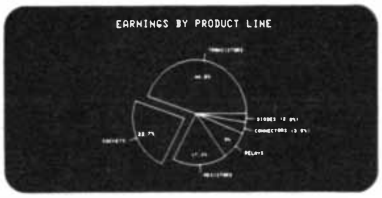
See the Kodak Carousel projectors at your photo dealer's. From less than $123. And enjoy the projector that keeps getting better.

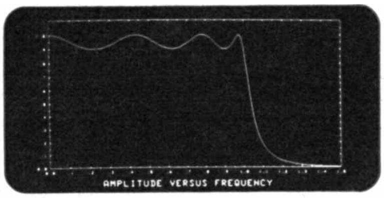
Price is subject to change without notice.

KODAK CAROUSEL PROJECTORS

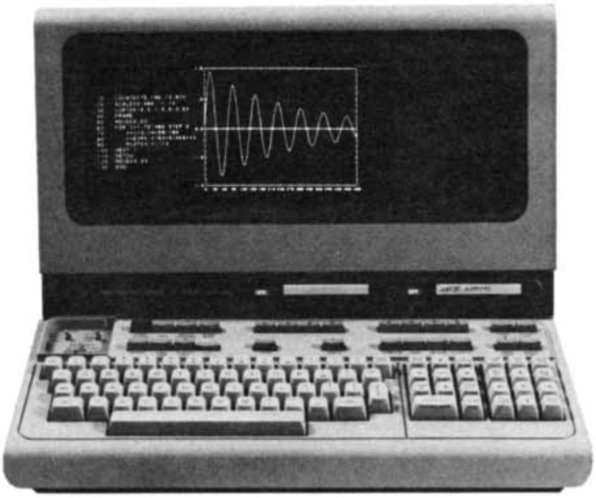
HP measurement and computer advances







With its multiple automatic plotting capability, the HP 2647 can display tabular data in formats of your choice, even if you have no programming knowledge. A simple menu helps you key in data parameters, after which a single keystroke plots the data automatically. Charts and graphs like these can be created without CPU software help.



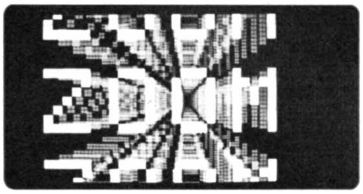
Meet HP's new intelligent graphics terminal. It's programmable, and it speaks BASIC.

The HP 2647 brings a new level of intelligence to our growing family of terminals. It combines user programmability in BASIC with sophisticated graphics capabilities to help you process and visualize complex data or detail an idea from concept to execution—with or without the assistance of a CPU depending on the complexity of the task.

By virtue of its microprocessor intelligence, plus 220K bytes of local storage on dual tape cartridges, the HP 2647 terminal is able to come up with some bright ideas:

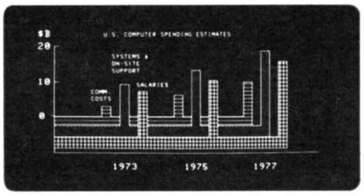
- Tabular data can be entered, the desired type of graph selected (bar, pie, X-Y Cartesian, or logarithmic), and the terminal will automatically label, scale, and plot the data for you.
- A simple BASIC program in the terminal can reformat or post-process data from a CPU into an appropriate graphic form, or check the accuracy of data prior to transmission to the CPU data base.
- Processing loads can be effectively shared or distributed between terminal and CPU; for example, repetitive tasks can be performed by the terminal leaving the CPU free to perform complex tasks using its full power.

Other graphics capabilities of the HP 2647 include:



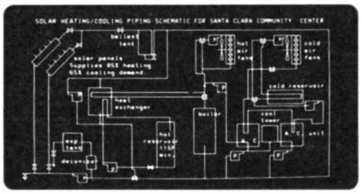
Zoom and pan

Magnify any area of the display up to 16 times (16×) to investigate or modify dense or complex areas. Pan in any direction for close-up viewing of the magnified display, without affecting display memory.



User-definable area shading

You can specify any pattern that an 8-by 8-bit cell can generate to make clear visual definitions of separate areas with similar shapes, as in bar charts and mechanical or architectural drawings.



Independent display memories

Alphanumeric and graphic data can be shown independently or at the same time, as in this process control schematic. And because alphanumeric and graphic displays each have separate memories and cursor controls, you can change one without disturbing the other.

Price of the HP 2647 starts at $8300*.

extend your possibilities.

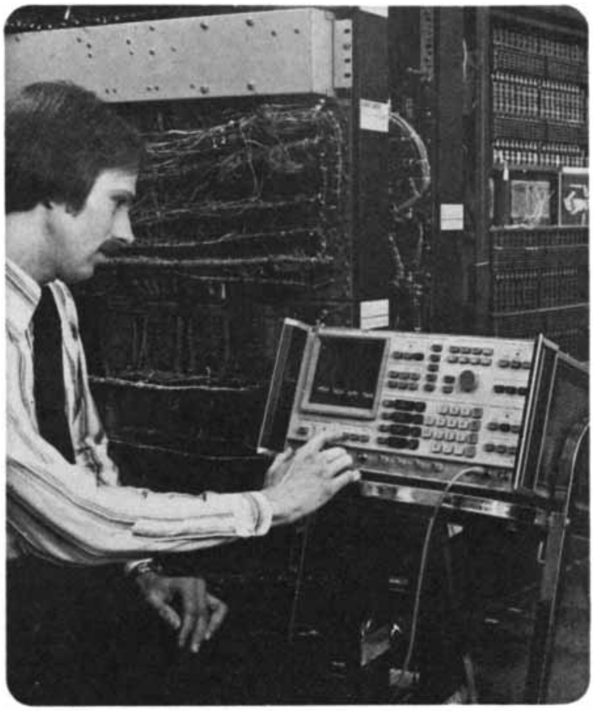
This new HP spectrum analyzer makes difficult measurements easy in the 20 Hz to 40 MHz range.

The addition of two technologies to that of spectrum analysis — frequency synthesis and microprocessor control — enhances the performance and simplifies the operation of the HP 3585 spectrum analyzer.

By portraying a signal's properties in the frequency domain, the standard spectrum analyzer can help measure linear and nonlinear circuit performance, distortion, modulation, frequency response, and many other properties. And while its spectral displays offer good qualitative information, the amplitude measurements derived from these displays are generally inaccurate. The HP 3585 combines synthesizer and microprocessor technologies to overcome this limitation, and to achieve some other significant benefits.

Measurement performance.
With a synthesized local oscillator based on a new type of phase-locked loop, the HP 3585 can enter center frequency and span settings with an 0.1 Hz resolution and $\pm 1 \times 10^{\pm 7}$ per month stability overthe analyzer's entire range of 20 Hz to 40.1 MHz. This frequency precision and stability make it possible to use the narrowest resolution bandwidth, 3 Hz, for close-in analysis even at 40 MHz. Microprocessor control (an example of HP's on-going NMOS II microcircuit technology) provides ± 0.5 decibel accuracy over most of the -135 to $+30$ dBm amplitude range.

Ease of operation.
Microprocessor control also gives the HP 3585 user keyboard control of those functions with variable parameters such as center frequency, span, and reference level. All the information that defines the displayed spectrum is displayed alphanumerically at top and bottom of the CRT for quick interpretation or permanent record. Advanced microprocessor software gives the user the choice of either variable analog or precise digital control where appropriate.



Programmability.
By connecting the HP 3585 to a computing controller such as the HP 9825 via the HP Interface Bus (HP-IB), measurement problems can be solved that would be impenetrable to analyzers alone. Three modes of operation in particular make this possible:

- Keyboard programming. Every front panel key on the 3585 can be programmed through the HP-IB.
- Analyzer as terminal. The controller can be used to place text on the analyzer's CRT, and the analyzer's keys can be used for input to a remote controller.
- Displayed manipulation. By transferring traces to, and loading traces from, the controller, analyzer results can be further processed, then redisplayed on the analyzer screen.

Those involved with audio, TV, telecommunications, and radio development will find the HP 3585 spectrum analyzer well worth its price of $17,500*.

HEWLETT PACKARD

1502 Page Mill Road, Palo Alto, California 94304

For assistance call: Washington (301) 948-6370, Chicago (312) 255-9800, Atlanta (404) 955-1500, Los Angeles (213) 970-7500

Mail to: Hewlett-Packard, 1502 Page Mill Road, Palo Alto, CA 94304.
Please send me further information on

( ) HP 2647 intelligent graphics terminal
( ) HP 3585 spectrum analyzer

Name ____________________
Company ____________________
Address ____________________
City ____________ State ________ Zip ________

*Domestic U.S. prices only.

00847

Learning and Memory in Bees

A bee is able to learn quickly and to remember for long periods the color and the odor of flowers that yield nectar or pollen. Now the neural basis of this programmed behavior is being elucidated

by Randolf Menzel and Jochen Erber

If you watch a single bee as it moves from flower to flower, you will see that it always chooses the same type of flower and pays no attention to nearby flowers of different color, shape and odor that are attracting other bees. Clearly the bee could not follow this pattern if it were not capable of two things: distinguishing one type of flower from another and learning what flowers offer nectar or pollen and so are worth visiting. In other words, the bee has a memory and an ability to learn.

We and our colleagues at the Free University of Berlin have been investigating these abilities of the bee in the hope of elucidating the cellular basis of learning and memory. Our studies are both behavioral (we watch bees in various experimental situations in order to determine the extent of their behavioral repertory) and physiological (we explore the anatomy and function of the bee's brain in an effort to ascertain what parts of it are involved in learning and memory). We find that even though the bee is a tightly programmed behavioral machine, its mechanisms of learning and memory are quite similar to the ones that appear to operate in much more highly evolved organisms, including man.

Although it was obvious to many careful observers in the past (the written comments range from Aristotle to Charles Darwin) that bees can see the colors and smell the odors of flowers and can remember what type of flower they have visited previously, the experimental proof had to wait until 1910, when Karl von Frisch at the University of Munich conducted a simple test that stands as one of the few really crucial and decisive experiments in the study of animal behavior. He attracted foraging bees to a table by offering a honey solution. When enough bees were visiting the place regularly, the dish containing the honey was placed on a piece of blue cardboard, so that the bees saw the blue color as they approached the table and as they sucked the honey.

Now von Frisch was ready for his test. He surrounded the blue cardboard with other pieces of cardboard of the same size but ranging in color from white through various shades of gray to black. On each piece of cardboard, including the blue one, was an empty dish. The bees, which had been conditioned to expect that they would find food at the table, continued looking, and nearly all of them confined their search to the blue cardboard.

The experiment provided strong support for both of the postulates we have mentioned. The bees were distinguishing the blue color from all the other colors, which meant that they could see blue as a color and also that they had learned to take the color (in that situation) as a signal for food. To confirm the second finding it was necessary to show that the bees had not been attracted spontaneously to the color. Von Frisch did so in various ways, as others have after him. Indeed, his experiment has been repeated with a variety of colors, odors and black-and-white patterns.

The behavioral aspects of the bee's performance are therefore well known. Surprisingly, the learning process involved in all these experiments was left unstudied until about 10 years ago. We shall show that the process has several interesting features that make it attractive for an approach toward a neuronal, or nerve-cellular, analysis of memory formation in the bee's brain.

Let us ask a first question. How much precision do bees show in selecting the type of flower from which they collect nectar and pollen? This question can be answered quickly and accurately without the necessity of following a bee on its flight from flower to flower. When a bee makes one of its periodic returns to the hive, it carries a record of its choices in its two pollen lumps (one on each hind leg). By inspecting the contents of the pollen lumps of individual bees one finds that the pollen brought back by a bee from a single excursion is entirely or almost entirely from a single type of flower. Since a bee visits as many as 500 flowers on a single excursion, the degree of precision in the bee's choice is high indeed.

This precision of response, together with the fact that bees can easily be trained to respond to various stimuli, makes the insects particularly suitable for the experiments we had in mind. It is also helpful that all the worker bees from one hive are sisters and therefore are closely related genetically and that bees are highly motivated to search for food, so that each bee makes a large number of decisions (from 1,500 to 2,000 per day).

To prepare for a direct approach to the neuronal events underlying learning and memory we first examined the learning process in detail with behavioral methods. Here we concentrated on experiments in which bees were trained to respond to different colors. (Other workers have employed different cues. For example, Martin Lindauer and his colleagues at the University of Würzburg trained bees either to odors or to a daily time schedule, and Rüdiger Wehner at the University of Zurich trained bees to discriminate among black-and-white patterns.)

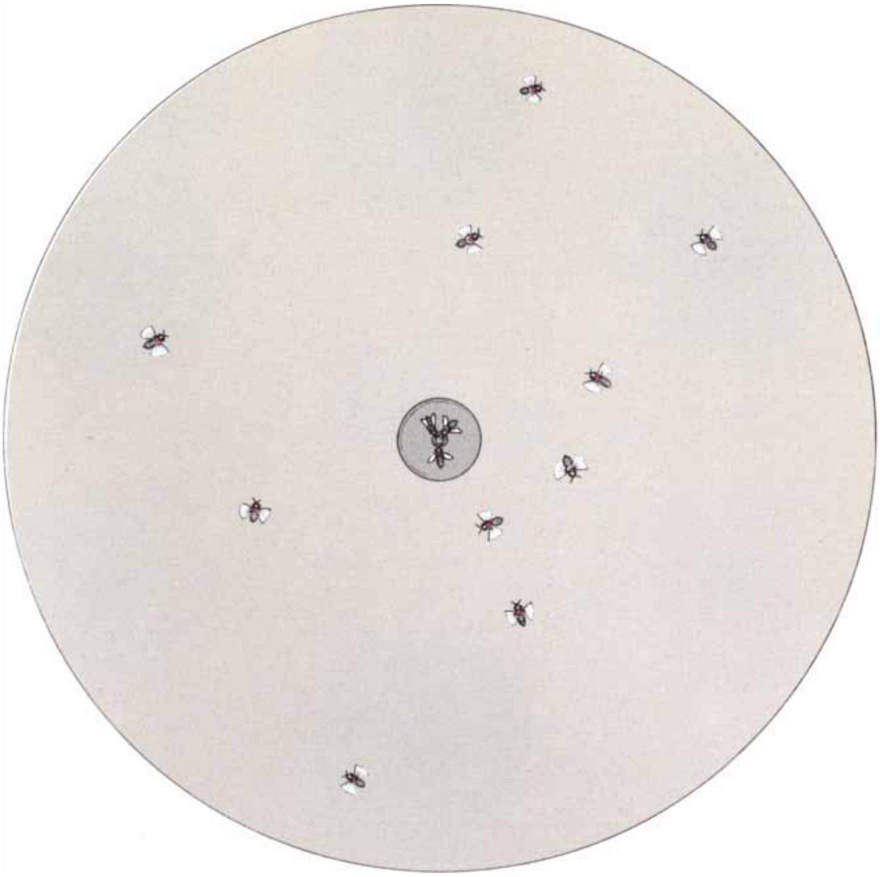
Our arrangement for color training involved an apparatus in which we could employ monochromatic light of adjustable intensity to illuminate two ground-glass disks. Individually marked bees are fed on an unilluminated glass disk in the center of the gray training table. By means of the "waggle dance" elucidated by von Frisch the bees inform their hive mates of this source of food. A newcomer arriving in response to this information is allowed to visit an unilluminated disk three times in order to become familiar with the arrangement. After three visits that yield food a bee is motivated to search for food even when none is offered. During the first three visits of a new bee all other bees are removed from the apparatus and kept in a cage, so that only the experimental bee flies back and forth between the hive and the apparatus.

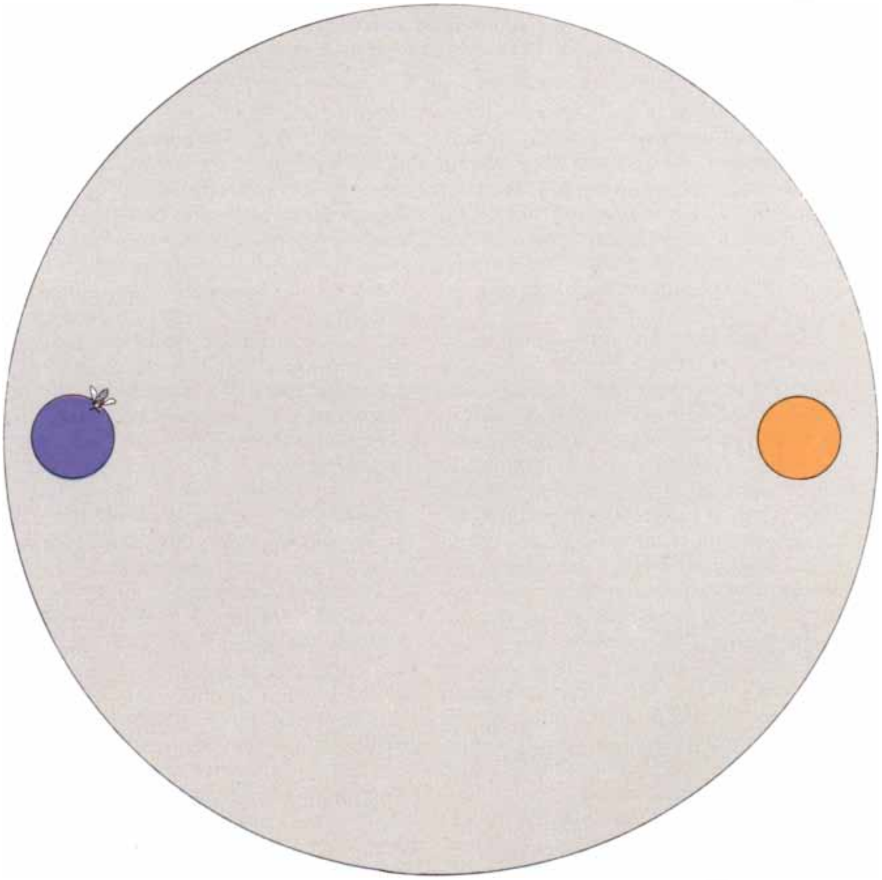
Our first test evaluates the spontaneous-choice preference of the bee. For the first time the experimental bee is confronted with two glass disks of different colors. Neither disk bears food, and each one is equidistant from the center (the former feeding place). The bee approaches the disks, lands on them and searches for food. If the frequency of choice is not equal for both colors, the intensity of one color or the other is adjusted until the bee displays no preference for either one.

After a test period of a few minutes the bee is rewarded with food on one of the two colors. When it comes back, a test is performed in the same way as in the spontaneous-preference trials. Then it is rewarded again, and so on. From this procedure we obtain learning curves for individual bees and therefore are enabled to evaluate quantitatively the factors influencing the learning behavior of bees.

Colors are learned within a few trials. Violet (at wavelengths of from 400 to 420 nanometers) is learned fastest; only one reward is needed to produce a high frequency of correct choices. Bluish-green (490 nanometers) is learned slowest. The characteristic differences of the learning curves are independent of such factors as the intensity of the two colors (if the intensity is at least half a logarithmic unit above the bee's perceptual threshold), the wavelength of the alternative color, the type of pretraining, the age of the test bee, the time (of day or season) and the weather.

One would expect the dependence of learning on wavelength to show a close correlation with the color-vision system of bees. It does not. The bee's color-vision system is trichromatic, with receptors maximally sensitive to ultraviolet radiation (350 nanometers), blue light (440 nanometers) and green light (540 nanometers). Ultraviolet is the brightest and most saturated color to bees, as was shown 20 years ago by Karl Daumer of the University of Munich. As one would expect from such a system, the bee's ability to distinguish colors is best for violet and bluish-green, since the sensitivity of the receptors is





COLOR TRAINING of bees was done by the authors with an apparatus that enabled them to illuminate ground-glass disks (seven centimeters in diameter) with monochromatic light of adjustable intensity. At first individually marked bees are fed (*top*) on an unilluminated disk in the center of the gray training table. With their "waggle dance" they inform other bees in the hive of the food source. A single responding bee is fed at the central disk, tested for its spontaneous choice of two colors equidistant from the central disk (*bottom*) and then rewarded with food on the central disk, now illuminated with one of the two colors. Even a brief and small reward causes the bee to return to that color consistently.

COLOR LEARNING by bees is quite rapid for some colors and somewhat slower for others, although a bee can learn any color that is visible to bees. In this set of drawings the flower of the wild rose is represented at the center as the human eye sees it and then, reading clockwise from the top, in order of the bee's ability to associate a color quickly with a reward of food. The color learned fastest is a deeply saturated violet, then blue, ultraviolet (*shown as gray*), green and orange. Bluish-green, the normal background color for bees, is learned slowest.

most pronounced in that region and they also overlap in sensitivity there. Indeed, this hypothesis was proved to be correct by Otto von Helversen of the University of Freiburg. Such sensitivities are necessary for a precise identification of the predominantly violet and blue flowers, which are specialized in attracting insects for cross-fertilization.

Although bees are most sensitive to ultraviolet radiation, it is not the color that they learn best. To the bee ultraviolet radiation is sky radiation, and it seems not to be expected as a pure color at food sources. Mixed with long-wavelength light to form "bee purple" it is learned much faster. Bluish-green is the background color and also is not expected as a food marker. The interactive evolution of flowers and their insect pollinators has led not only to an adaptation of flower colors and bee color perception but also to pre-fixed learning dispositions in bees. This interpretation is supported by Lindauer's studies on the odor learning of bees; naturally occurring odors are learned faster than artificial and uncommon ones.

We note in this context the recent tendency among experimental psychologists to appreciate the value of the ethological viewpoint, which is that behavior is best studied in relation to the animal's natural environment and evolution. Color and odor learning in bees appear to be examples of behavior that can most fruitfully be examined in this way. As an ethologist von Frisch intuitively interpreted his results along this line 40 years ago. Today one can quantitatively separate the design of the bee's perceptual system (the behavioral program for gathering context-specific information) and the neural programs for processing this information into stored memory.

How stable is the bee's memory for food signals? Anecdotal reports from beekeepers suggest that it is highly stable and also durable. One of the first quantitative observations was made in 1963 by Lindauer, who found that bees arrived spontaneously at a feeding place after five months of absence during the winter. To obtain a still more quantitative evaluation we trained bees to blue or yellow with from one reward to three, kept them in individual mesh cages in the hive for varying periods of time and then tested their choice behavior. After one reward the frequency of a given choice stayed high for several hours, but it declined on the next day and reached the spontaneous-choice level in from five to seven days. After three rewards, however, we found no reduction in correct choices during the period of about 13 days in which a bee can survive in a cage. It is apparent that bees need only a few rewards to establish a highly stable memory.

The situation is different if the bees are allowed to learn anew. They will switch to a new food source marked with a different color, but the speed of the switch depends on the number of rewards they received at the color that is now unrewarded. In this situation one can observe a curious behavior well known from studies of vertebrates. The animal makes the change more rapidly if it has had a long period of training (25 rewards or more) to the initial color. Another finding is that if two signals carry rewards at different rates, the animal chooses them according to their reward probabilities. (This effect has been studied in detail by Bernhard Schnetter of the University of Würzburg.) Another example is simultaneous training to more than one color. The bee can learn two colors simultaneously and can distinguish them from other colors. If more than two colors are rewarded simultaneously, however, the bees stop distinguishing colors. All of this reveals something of the design logic in their learning programs.

Such complex and long-lasting learning effects are interesting from the viewpoints of ethology and comparative psychology, but they are not suited for an analysis of the neuronal events underlying behavior. We therefore concentrated on analyzing the neuronal processes during and after a single association of a color or an odor signal and a reward. This turned out to be a fortunate decision, because suddenly new perspectives on the learning process appeared.

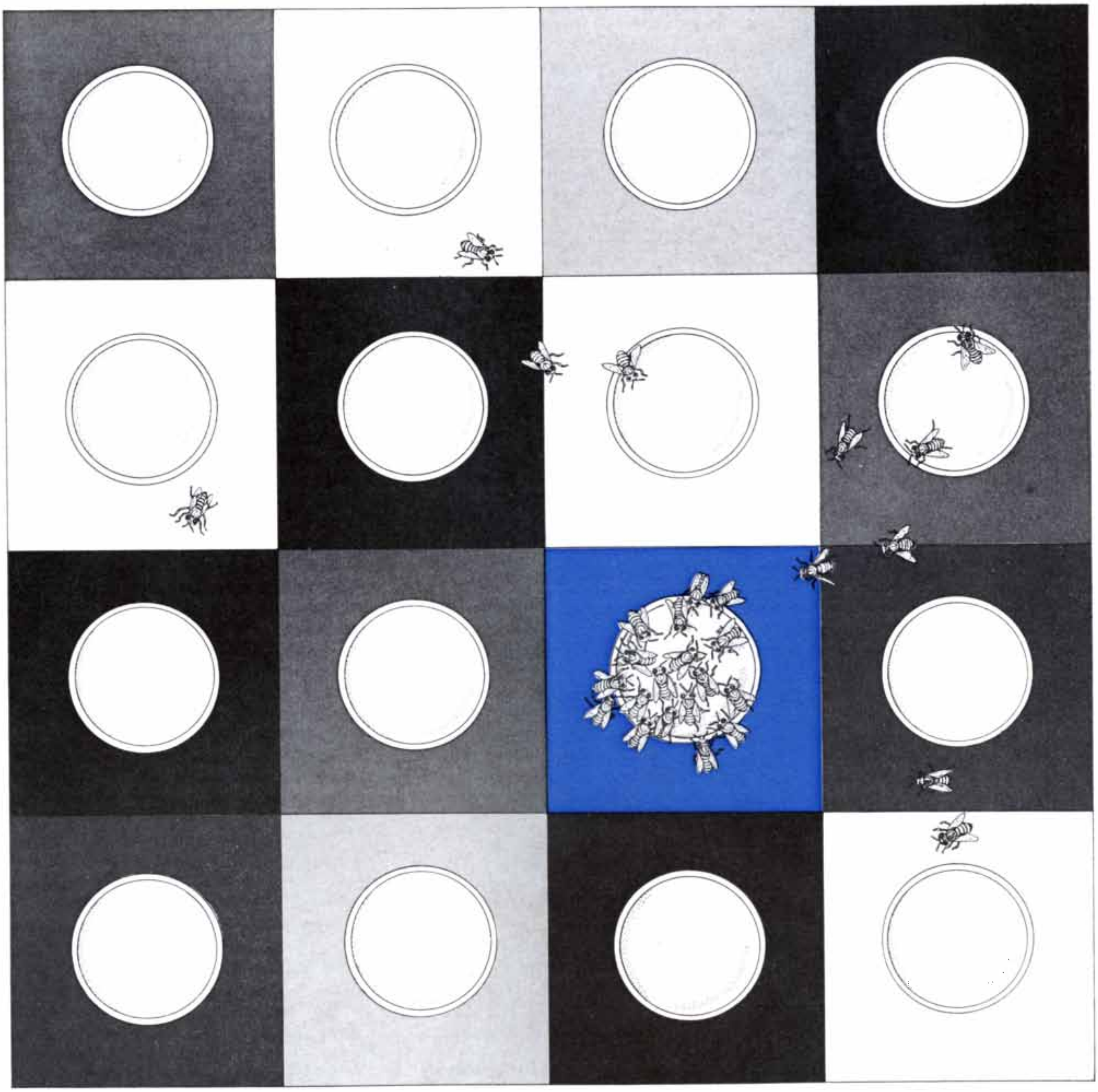
We first established a most useful result: After a single reward the response level is independent of the quantity of reward, provided that certain threshold limitations are observed. Even a reward that persists for only 100 milliseconds and has a low concentration of sucrose is sufficient to produce a significant change in response.

Another interesting result appeared when we trained bees with a single short reward (two, five or 10 seconds). Bees tested immediately after the reward displayed a high rate of correct choices. Over the next two minutes, however, the score dropped drastically, reached a minimum at between two and three minutes and then rose again to about the initial level.

This kind of relation between memory formation and time appears in other animals and in man. The physiological mechanisms underlying the association between a signal and a reward (or a punishment) need time (seconds, minutes and even hours) and most probably go through phases in which the neural substrate for the memory is different. In the typical case the first storage is in a sensory memory, from which the information may quickly disappear if it is not reinforced. On reinforcement the information goes into short-term storage, which is reflected in the high level of correct responses made by a bee immediately after one reward or more. The rise after a subsequent decline represents a phase termed consolidation, wherein the information is transferred to long-term storage.

For the bee, as one might expect, the phases of learning are nicely adapted to the natural learning situation. Most flowers offer such small quantities of nectar that a bee will suck for only a few seconds and will usually approach the next flower within another few seconds. If the flowers are widely distributed, the bee may need up to a minute to find the next similar flower. The initially high rate of correct choices indicates a short-term mechanism that assures the bee of being able to find a similar flower even after very short rewards. The longer-lasting consolidation phase seems to be adapted to the bee's flight between the food-collecting area and the hive. When the bee comes back from the hive, it again makes a high level of correct choices, indicating that the information is in long-term storage.

How long is a color signal stored



COLOR TEST made by Karl von Frisch in 1910 established the ability of bees to associate colors with food and to remember a color thus learned. He put a dish containing a honey solution on a table. Bees were attracted to it. Then he put the dish on a piece of blue cardboard, so that blue was the color the bees saw as they approached the dish and sucked the honey. Next von Frisch surrounded the blue cardboard with other cards of the same size but colored white, black or shades of gray. Each card, including the blue one, bore an empty dish. The bees had been conditioned to expect food at the table, and so they continued to come. Most of them looked only on the blue card.

in the bee's sensory memory? The answer lies in ascertaining the conditions under which a color will still be associated with food even though the color is no longer visible. In our tests of this matter we have found that a color signal perceived by a bee within two seconds before the start of the reward is learned as well as if the signal is present through the entire period of approach and sucking. If a color signal is switched on after the bee starts sucking, it is not associated with food.

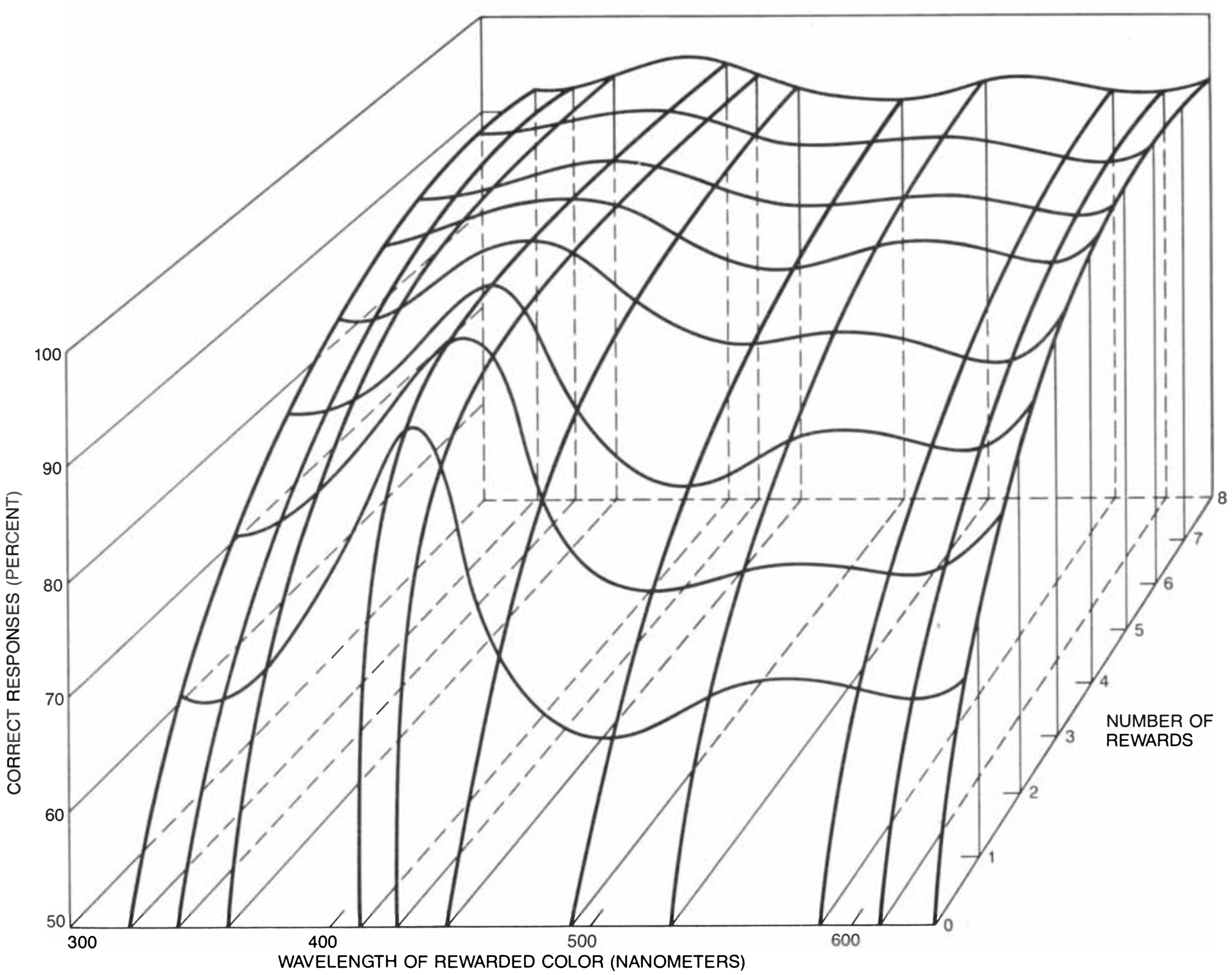
Elisabeth Opfinger observed in 1931, when she was a student working with von Frisch, that bees learn a color only if they see it during the approach to a flower; the colors seen during sucking and flying away are not learned. Now we know that a color signal must be present within three seconds before and about .5 second after the start of sucking to be learned. Since a single sucking period of two seconds is enough for the bee to show a highly significant level of learning, the total initial learning period (sensory memory and the association period plus the duration of the reward) appears to be from four to five seconds. This is a convenient situation for exploring the temporal phases of a storage of memory initiated by a single reward.

Do the short-term and long-term phases reflect two physiologically different storage mechanisms? We have treated bees with various procedures that are known to block neural activity in the brain. All of them prevent the formation of short-term memory. (The bees thus treated behave after recovery in exactly the same way as untreated bees. Moreover, control experiments excluded the possibility that the experimental procedures caused a negative reinforcement.)

We have found a correlation between the initial period of declining memory and the period of susceptibility to the blocking procedures. The results tell us that during an initial period of a few minutes the storage of memory is based on physiological mechanisms that are different from the ones involved in later storage. Most probably the early physiological mechanisms are orderly neuronal activities, since our treatments block such activities.

The first step in associative learning obviously is a mechanism of fleeting storage of sensory information, the phase we have called sensory storage. If the information is not reinforced promptly, it disappears. Reinforcement of the visual signal depends on a second sensory system: the sugar receptors on the bee's tongue. If the receptors are stimulated by food, the next phase of the learning program (storage of the information in a memory system) is set in motion.

It is plausible to assume that these initial and rapid processes of channeling and selecting information are based on the electrical activity of particular nerve cells. Electrophysiological recordings might help to identify the cells. But



LEARNING CURVES indicate the speed with which bees learned different colors. The curves have been averaged from the performance of many bees. The average of correct responses rises with the number of food rewards. In these experiments the alternative to the rewarded color was the complementary color: for ultraviolet at 361 nanometers it was bluish-green at 494; for violet at 428 nanometers it was orange at 590, and for blue it was "bee purple," a mixture of 80 percent orange and 20 percent ultraviolet (*curve farthest to the right*). The learning is quickest for violet at 408 and 418 nanometers and slowest for bluish-green at 494 nanometers. Trained to violet, bees reach an accuracy level of 90 percent after only one reward; for bluish-green they need about five rewards to reach the same level.

where in the bee's brain should one insert a recording electrode? Although the bee's brain is relatively small (about one cubic millimeter in size), it contains something like 850,000 nerve cells—too many for a random searching strategy. We have therefore sought instead to localize the areas of the brain that are involved in the initial storage processes.

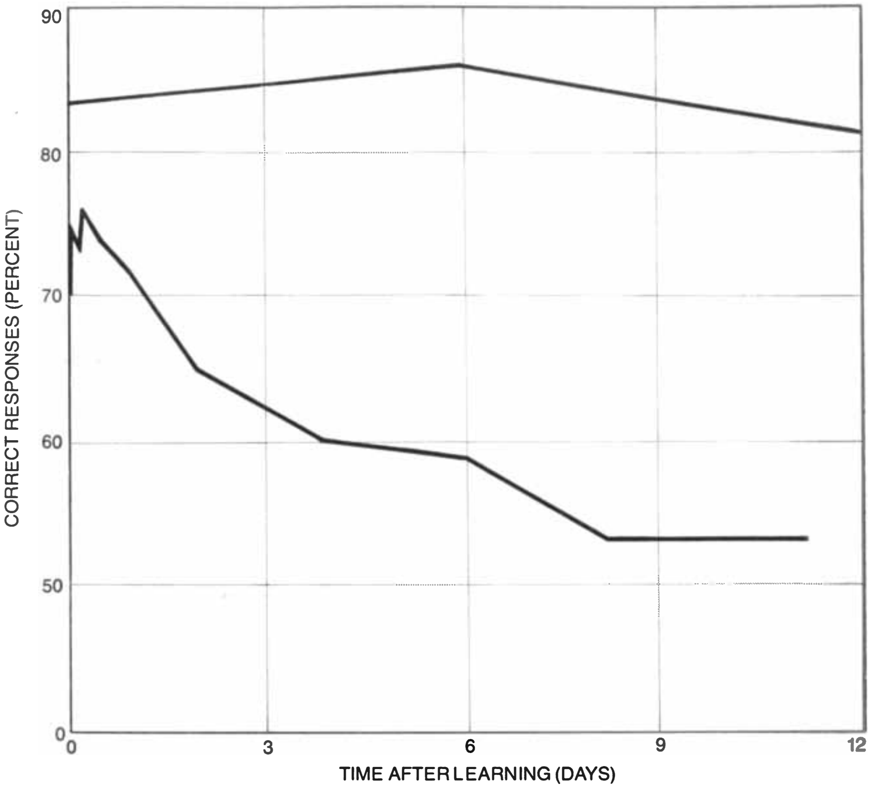
Our approach is through the reflex response of the proboscis to sugar-solution stimuli. A bee extends its proboscis as soon as an antenna is touched with sugar solution. We fix a bee in a small metal tube and present a whiff of odor just before touching the antenna with sugar solution and allowing the bee to lick the solution with its proboscis for a few seconds. If the odor is then presented without a sugar solution, the bee will almost invariably extend its proboscis. Only one conditioning trial is needed to establish a high level of responses.

We found that we could expose the brain during this process without unduly affecting the learning behavior. Then we could reversibly block neuronal activity in small areas of the brain by means of a thin cooled needle. The blocked area is as small as 250 micrometers in width and 150 in depth.

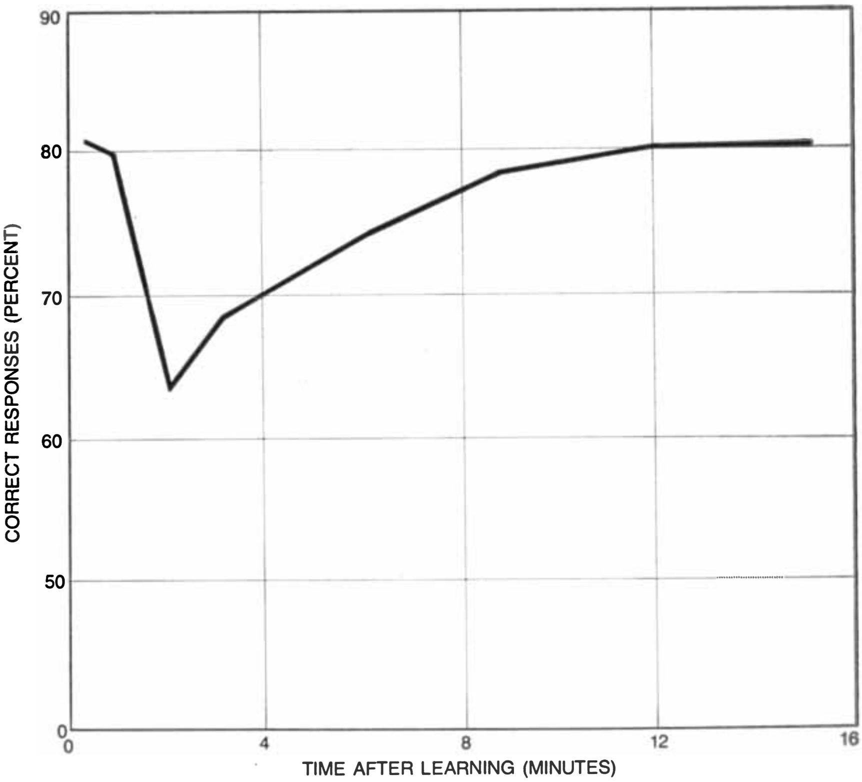
Through this procedure our interest came to focus on the corpora pedunculata, also known as the mushroom bodies. There are two of them, each densely packed with neuropils. A mushroom body has four major substructures: the caplike calyx, the stalk, the frontal alpha lobe and the bottom beta lobe. Its main input comes from another structure, the antennal lobe. Histologists of the 19th century observed that mushroom bodies are particularly prominent in social insects and inferred that the highest brain functions (learning, memory and complex social behavior) are probably located there.

When we impaired parts of the antennal lobes and mushroom bodies by cooling them, we found a correlation with what we had observed by blocking neuronal activity in freely flying bees. The effect of impairment decreases with increasing time between trial and treatment. The susceptibility to impairment decreases fastest in the antennal lobe and slowest in the calyx.

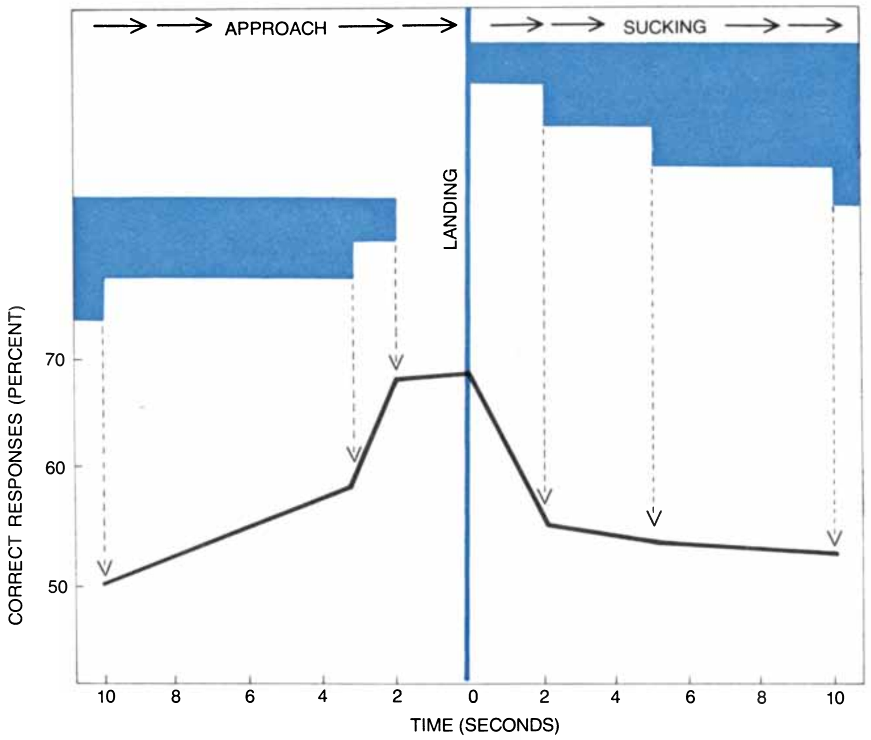
The results indicate that the antennal lobes and parts of the mushroom bodies are involved in the processing of information into the memory, although to different extents and over different periods of time. The antennal lobes are involved only during a short initial period of about three minutes. The calyx is involved longer. Since the memory can be shown to be impaired even seven minutes after cooling of the calyxes, they are probably the essential structures in the transfer of information from the short-term memory to the long-



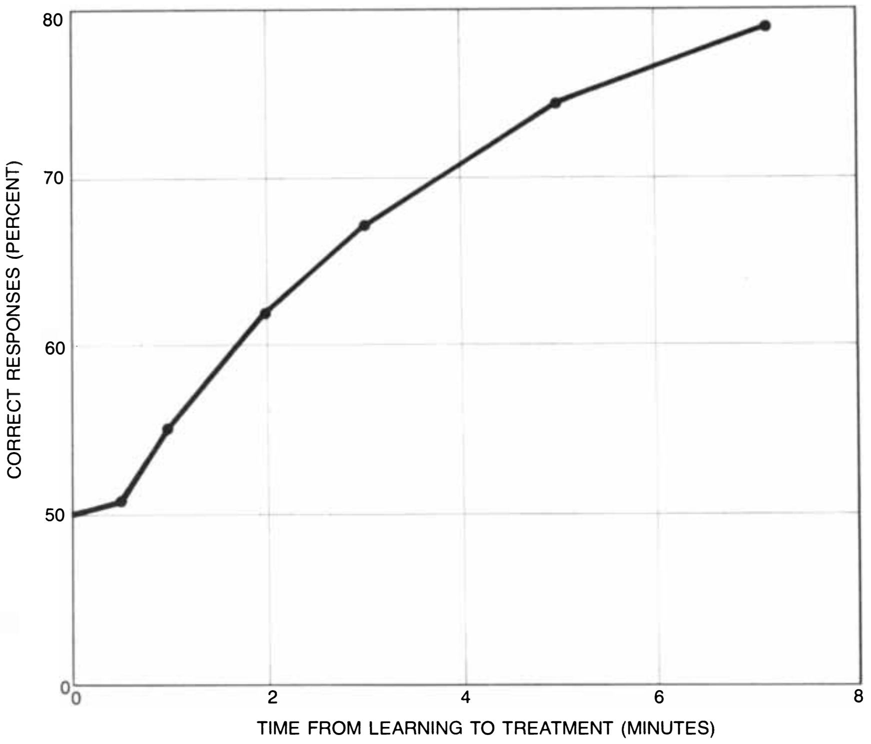
STABILITY OF MEMORY in bees is charted. If bees are rewarded three times on a particular color and then kept in the hive for up to two weeks to exclude new learning, they keep a high level of response afterward (*upper curve*). With only one reward the response level rises at first, a period called the consolidation phase. One day later, however, the response begins to drop; after about six days it reaches the level of spontaneous choice. The reward disk was blue.



CHOICE BEHAVIOR of bees following a one-trial learning period showed this pattern. Bees tested immediately after learning choose the rewarded color with high accuracy. Two minutes later the number of correct responses falls but rises again thereafter in the consolidation phase.



ASSOCIATION PERIOD between a color and a food reward is well defined and quite brief. A color, which in these tests was blue, seen by the bee more than four seconds before landing or more than two seconds after landing and beginning to suck is not associated with the reward. A color signal switched off two seconds before the start of the reward is learned. The experiments suggest a sensory memory that persists for as long as three seconds. Information lodged there must be reinforced with food if it is to be processed into short- and long-term memory.



DISRUPTIVE PROCEDURE produced the effects charted here on early formation of memory in bees. After a bee's spontaneous choice of colors had been determined, the bee was rewarded on blue light for 30 seconds. Then its brain was chilled either immediately or after various intervals indicated by the circles. The curve shows the average percentage of correct responses after the recovery of bees thus treated. Performance of bees chilled after seven minutes was about normal. Earlier chilling evidently interfered with formation of short-term memory.

term one. The alpha lobes, which are connected to the calyxes by thousands of nerve fibers, are involved too.

The one-trial learning situation in our experiments enabled us to follow the time course of the transcription from short-term memory to long-term. The time courses that we found for various brain structures make it most unlikely that the electrochemical and biochemical processes involved in the transcription follow a simple dynamic within a confined synaptic area. Since blocking neural activity both in freely flying color-trained bees and in fixed odor-trained bees had similar results over time, it appears that the central nervous processes underlying the early storage of memory are similar for both color and odor.

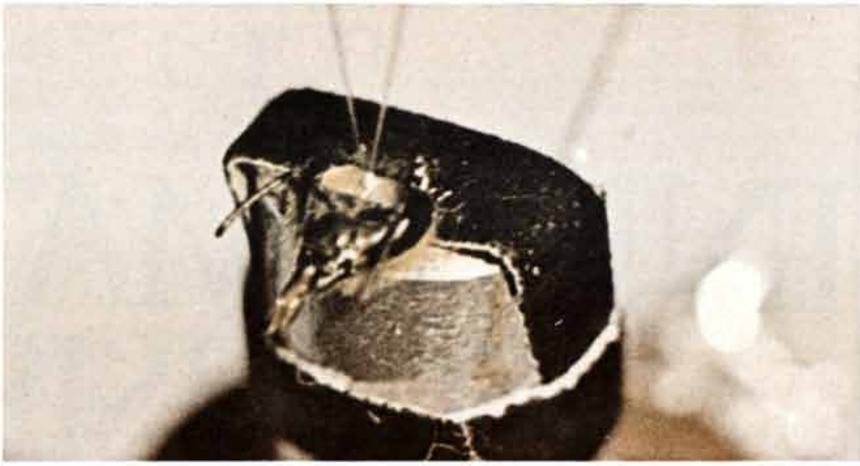
We have identified particular areas of the brain that are involved in the formation of memory of odors. Whether the same areas are correlated with visual learning is an open question. In any event the time courses of learning of both olfactory and visual information are quite similar. Electrophysiological recordings performed on our immobilized odor-conditioned bees should therefore elucidate also the general mechanisms of associative learning that are at work in freely flying bees.

It is not easy to make recordings from single neurons in the bee's brain because the neurons are usually very thin. Nevertheless, we have found it possible to make such recordings while the insect is free to move its proboscis and can be conditioned to odor. Our electrophysiological studies are only beginning, but a few general findings are available.

The majority of the neurons in the mushroom body receive inputs from more than one sensory system. For example, a neuron might respond to a light flash with inhibition (firing less), to sugar water on the antenna with excitation (firing more) and to a mixed odor with excitation. We have observed many different combinations of responses to different stimuli.

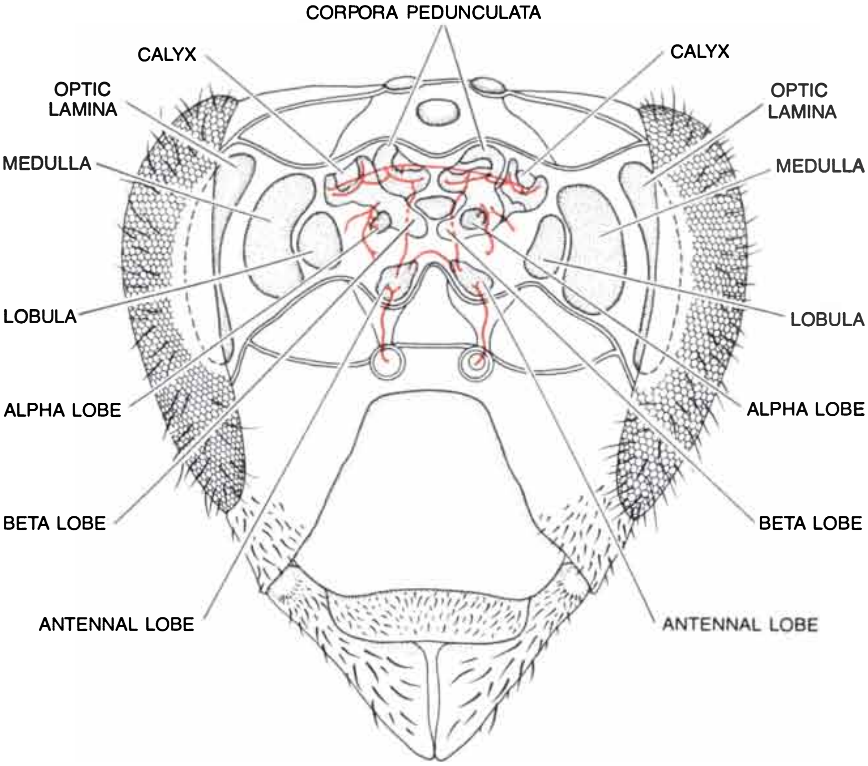
Only a few of these neurons change their responsiveness during conditioning. The ones that do were already responding to odor and sugar before conditioning. The change is usually a rise of the spontaneous-discharge rate of the neuron accompanied by an increase in sensitivity to stimuli of odor and sugar water. In some neurons we found a close correlation between the number of conditionings and the change in responsiveness. By intracellular staining of neurons we are trying to identify the particular cells that learn. Most of them turn out to be in the area around the mushroom bodies, reinforcing the hypothesis that this part of the brain is highly involved in the processing and storing of olfactory information.

It is apparent from the work that we and others have done with bees that associative learning in the bee follows



ODOR CONDITIONING of bees offers another way of testing the formation of memory. The bee is fixed in a tube and its antenna is touched with a drop of sucrose solution. The bee extends its proboscis in an unconditioned reflex. If a whiff of odor is presented shortly before the antenna is touched, and the bee is allowed to taste sugar with its extended proboscis, the odor becomes a conditioned stimulus. After a single conditioning, that is, one taste of sucrose, most bees extend the proboscis even when the whiff of odor is presented without a reward of food.

all the major learning steps that have been established for vertebrate animals. Moreover, bees have sensory, short-term and long-term memories quite similar to those found in vertebrate species. The neural strategies underlying learning and memory storage in the bee also do not differ basically from what is known about vertebrates. With work on the bee, therefore, we should have a good chance of unraveling a few basic rules of associative learning and memory that are also relevant to complex vertebrate neural systems.



BEE'S BRAIN is portrayed schematically in a frontal view. The lamina, medulla and lobula mainly process visual information. The midbrain incorporates the two corpora pedunculata, or mushroom bodies, which are composed of densely packed and highly ordered interneurons. The regions of the mushroom bodies are the calyxes, the alpha lobes and the beta lobes. Two other structures, the antennal lobes, are the primary sensory centers for inputs from the antenna. Major nerve tracts handling information flow between mushroom bodies and the antennal lobes are shown in color. Bee's cubic-millimeter brain has some 850,000 nerve cells.

The Ultimate Tax Shelter

by
TED NICHOLAS

Tax experts are now referring to a small, privately owned corporation as "The Ultimate Tax Shelter." This is especially true since the passage of the Tax Reform Act of 1976. This law makes most former tax shelters either obsolete, or of little advantage. Investments affected include real estate, oil and gas drilling, cattle feeding, movies, etc. These former tax shelters have lost their attractiveness. Aside from that, these tax shelters required a large investment. Only a small segment of the population could benefit from them.

I've written a book showing how you can form your own corporation. I've taken all the mystery out of it. Thousands of people have already used the system for incorporation described in the book. I'll describe how you may obtain it without risk and with a valuable free bonus.

A corporation can be formed by anyone at surprisingly low cost. And the government encourages people to incorporate, which is a little known fact. The government has recognized the important role of small business in our country. Through favorable legislation incorporating a small business, hobby, or sideline is perfectly legal and ethical. There are numerous tax laws favorable to corporate owners. Some of them are remarkable in this age of ever-increasing taxation. Everyone of us needs all the tax shelter we can get!

Here are just a few of the advantages of having my book on incorporating. You can limit your personal liability. All that is at stake is the money you have invested. This amount can be zero to a few hundred or even a few thousand dollars. Your home, furniture, car, savings, or other possessions are not at risk. You can raise capital and still keep control of your business. You can put aside up to 25% of your income tax free. If you desire, you may wish to set up a non-profit corporation or operate a corporation anonymously. You will save from $300 to $1,000 simply by using the handy tear-out forms included in the book. All the things you need: certificate of incorporation, minutes, by-laws, etc., including complete instructions.

There are still other advantages. Your own corporation enables you to more easily maintain continuity and facilitate transfer of ownership. Tax free fringe benefits can be arranged. You can set up your health and life insurance and other programs for you and your family wherein they are tax deductible. Another very important option available to you through incorporation is a medical reimbursement plan (MRP). Under an MRP, all medical, dental, pharmaceutical expenses for you and your family can become tax deductible to the corporation. An unincorporated person must exclude the first 3% of family's medical expenses from a personal tax return. For an individual earning $20,000 the first $600 are not deductible.

Retirement plans, and pension and profit-sharing arrangements can be set up for you with far greater benefits than those available to self-employed individuals.

A word of caution. Incorporating may not be for you right now. However, my book will help you decide whether or not a corporation is for you now or in the future. I review all the advantages and disadvantages in depth. This choice is yours after learning all the options. If you do decide to incorporate, it can be done by mail quickly and within 48 hours. You never have to leave the privacy of your home.

I'll also reveal to you some startling facts. Why lawyers often charge substantial fees for incorporating when often they prefer not to, and why two-thirds of the New York and American Stock Exchange companies incorporate in Delaware.

You may wonder how others have successfully used the book. Not only a small unincorporated business, but enjoyable hobbies, part time businesses, and even existing jobs have been set up as full fledged corporations. You don't have to have a big business going to benefit. In fact, not many people realize some very important facts. There are 30,000 new businesses formed in the U.S. each and every month. 98% of them are small businesses; often just one individual working from home.

To gain all the advantages of incorporating, it doesn't matter where you live, your age, race, or sex. All that counts is your ideas. If you are looking for some new ideas, I believe my book will stimulate you in that area. I do know many small businessmen, housewives, hobbyists, engineers, and lawyers who have acted on the suggestions in my book. A woman who was my former secretary is incorporated. She is now grossing over $30,000 working from her home by providing a secretarial service to me and other local businesses. She works her own hours and has all the corporate advantages.

I briefly mentioned that you can start with no capital whatsoever. I know it can be done, since I have formed 18 companies of my own, and I began each one of them with nothing. Beginning at age 22, I incorporated my first company which was a candy manufacturing concern. Without credit or experience, I raised $96,000. From that starting point grew a chain of 30 stores. I'm proud of the fact that at age 29 I was selected by a group of businessmen as one of the outstanding businessmen in the nation. As a result of this award, I received an invitation to personally meet with the President of the United States.

I wrote my book, **How To Form Your Own Corporation Without A Lawyer For Under $50**, because I felt that many more people than otherwise would could become the President of their own corporations. As it has turned out, a very high proportion of all the corporations formed in America each month, at the present time are using my book to incorporate.

Just picture yourself in the position of President of your own corporation. My book gives you all the information you need to make your decision. Let me help you make your business dreams come true.

As a bonus for ordering my book now, I'll send you absolutely free a portfolio of valuable information. It's called "The Incom Plan" and normally sells for $9.95. It describes a unique plan that shows you how to convert most any job into your own corporation. You'll increase your take-home pay by up to 25% without an increase in salary or even changing jobs in many cases. If you are an employer, learn how to operate your business with independent contractors rather than employees. This means that you'll have no payroll records or withholding taxes to worry about. And you'll be complying with all I.R.S. guidelines. "The Incom Plan" includes forms, examples and sample letter agreements to make it possible.

I personally guarantee your purchase in the fairest way I know. If you feel my book is not all that I've described, return it undamaged within a full year and I'll promptly refund your money with no questions asked. If you should decide to return it, you may keep the bonus Incom Portfolio for your trouble.

To get your copy, write the words "Corporation Book and Bonus" on a plain sheet of paper, along with your name and address. Enclose a check for $14.95. There is no sales tax on your order and your purchase price is tax deductible. Mail your order to me, Ted Nicholas, c/o Enterprise Publishing Co., Dept. SA-83C , 1000 Oakfield Lane, Wilmington, Delaware 19810. C1016

Synthetic-Membrane Technology

Thin films of cross-linked polymers can separate molecules according to size, charge or other properties. The synthetic membranes can be applied to such industrial tasks as waste treatment and desalination

by Harry P. Gregor and Charles D. Gregor

A surprising number of industrial processes can be characterized as simply the separation of one thing from another. In mining, for example, pure metal must be extracted from the ore in which it is found. The separation of salt from seawater yields two commodities: the salt and potable water. In the treatment of sewage and of many industrial effluents, wastes must be concentrated for efficient disposal by separating them from a fluid stream.

A straightforward approach to each of these tasks would be to construct a barrier that allows particles or molecules of one kind to pass but that excludes all others. Such a selective barrier can be created in the form of a synthetic membrane made from a polymer. Membranes of this kind have been known for many years, but only recently have certain impediments to their commercial exploitation been overcome.

The synthetic membranes can discriminate between substances on the basis of several properties. In the process called ultrafiltration molecules are separated according to their size, in electrodialysis according to their electric charge, in solvent extraction according to their differing affinities for two dissimilar solvents. Membranes that are permeable almost exclusively to water are employed in the process called reverse osmosis. With an arrangement of several membranes a solution of a salt in water can be decomposed into an acid and a base, and by binding an enzyme to a membrane chemical reactions can be promoted with efficient separation of reactants and products.

Perhaps the most straightforward membrane technology is ultrafiltration, in which dissolved molecules and suspended particles are sorted out primarily according to whether or not they are small enough to pass through pores in the membrane. In ordinary filtration a liquid is forced under pressure through a porous medium, such as a sheet of paper or a bed of packed sand. Ultrafiltration is not fundamentally different, except that the pores are roughly 1,000 times smaller than those in an ordinary filter, so that not only particulate matter but also selected molecules can be removed and concentrated. In addition, some ultrafiltration membranes tend to exclude electrically charged molecules, regardless of their size.

Almost all suspended particles have sizes in the range of 10^{-6} meter, or one micrometer. Dust particles are rarely smaller than one micrometer in diameter; bacteria are generally from five to 10 micrometers in their largest dimension. The pores of an ultrafiltration membrane are on the scale of 10^{-9} meter, or one nanometer; average pore diameters range from less than a nanometer to about 10 nanometers. Hence virtually all suspended particles are rejected by ultrafiltration, including bacteria. So are viruses, which have a minimum size of some 20 nanometers.

Ultrafiltration is in essence molecular filtration; it discriminates between molecules of different sizes, and roughly speaking between those with different molecular weights. Water, with a molecular weight of 18 and an effective diameter of about .2 nanometer, can pass freely through all ultrafiltration membranes. Simple sugars, such as glucose (which has a molecular weight of 180 and a size of about one-half nanometer), can also pass through most ultrafiltration membranes. Proteins and other large biological molecules, on the other hand, are generally excluded, since they often have a molecular weight approaching 100,000 and a size of several nanometers. By choosing a membrane with carefully controlled pore sizes it is even possible to separate molecules with modest differences in size, such as sucrose (molecular weight 342) and cyanocobalamin, or vitamin B-12 (molecular weight 1,300).

An ultrafiltration membrane is made, or "cast," from a solution of long-chain polymers, such as cellulose acetate dissolved in acetone. The solution is viscous (rather like honey), and it is spread with a knife edge so as to form a thin layer. Some of the solvent is then allowed to evaporate, with the result that the polymer starts to come out of solution, creating a thin, semisolid matrix. Finally the membrane is plunged into another solvent, such as water, that rapidly precipitates all of the remaining polymer. It is this quenching that forms the pores in the membrane, as the rapid precipitation leads to the clumping or coagulation of the polymer. The size of the pores is determined largely by the time allowed for partial drying and by the choice of solvents. Although the pores are never entirely uniform, their average diameter can be controlled rather closely.

Membranes cast in this way are quite thin, generally from five to 30 micrometers. Since they are required to withstand pressure, mechanical support must be provided by a substrate with macroscopic pores. A membrane can be cast on a glass plate, then peeled off and mounted on a suitable substrate. Today, however, it is more common to cast the membrane directly on a porous support, such as plastic, paper or cloth.

The first synthetic ultrafiltration apparatus was probably one made by the German chemist and biologist Moritz Traube in 1870. Traube precipitated a gel of cupric ferrocyanide within the macroscopic pores of an unglazed flowerpot. By the 1930's polymer ultrafiltration membranes of various pore sizes had been developed. In that era work of particular importance was done by William J. Elford, who cast membranes of cellulose acetate and cellulose nitrate, and who also elucidated most of the principles of ultrafiltration.

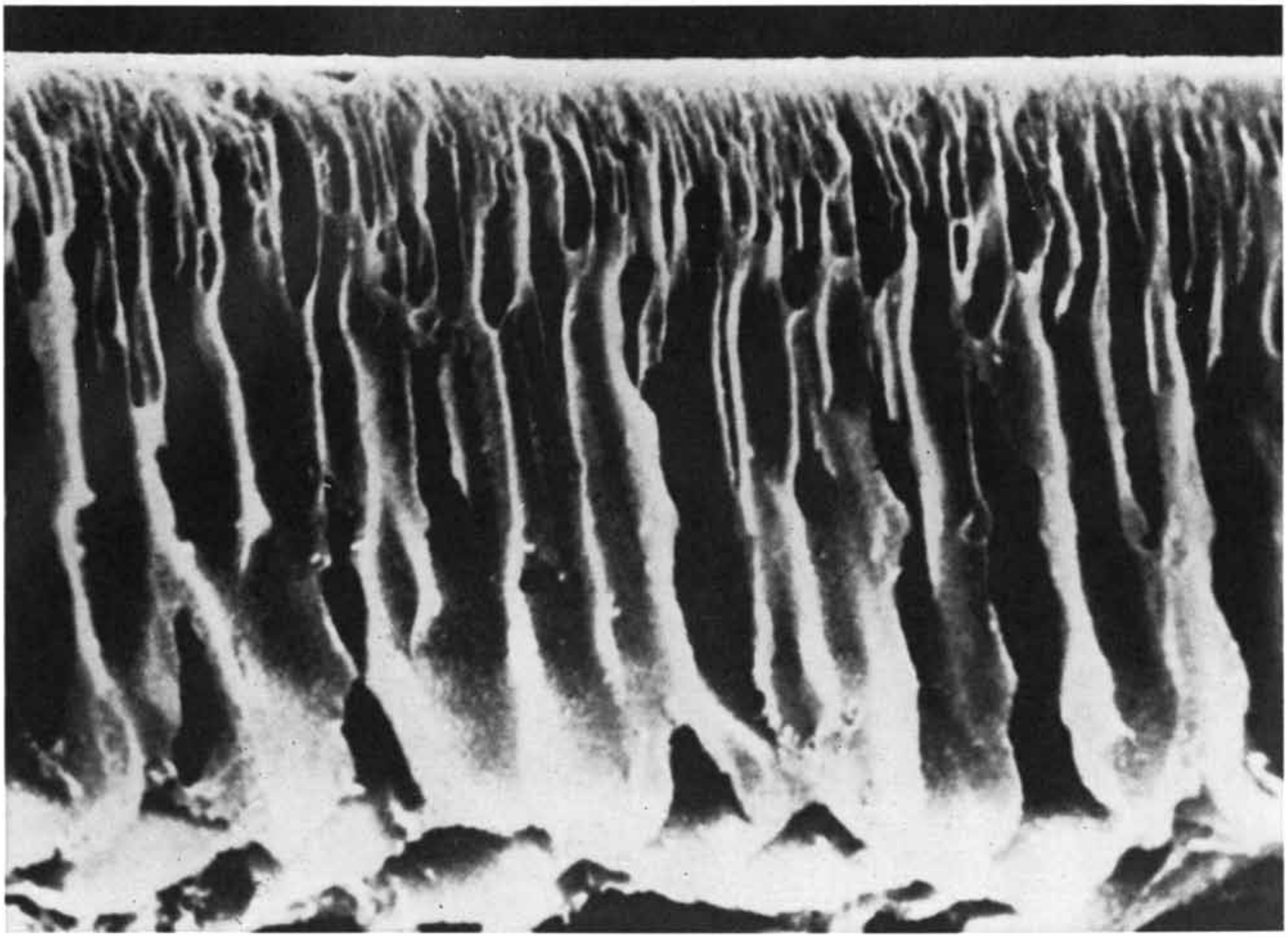
The theory of ultrafiltration can be illustrated by a simple model in which a membrane is interposed between two liquids, one consisting of pure water and the other of water with some substance dissolved in it. It can be assumed initially that the membrane is freely permeable to water and completely impermeable to the solute. The magnitude and direction of flow through the membrane are then governed by a combination of four factors.

The first of these factors is an osmotic pressure that tends to drive water from the side of the membrane in contact with pure water across the membrane and into the solution. The origin of the osmotic pressure can be explained rigorously only by introducing the laws of thermodynamics and in particular the concept of entropy, but it can be understood intuitively as a tendency for the solvent (in this case water) to equalize its concentration throughout the volume, just as if the membrane were not present. Since the concentration of water is greater in pure water than it is in a solution, the water molecules tend to pass through the membrane into the solution. The magnitude of the osmotic pressure is equal to the difference in the concentration of water on each side of the membrane.

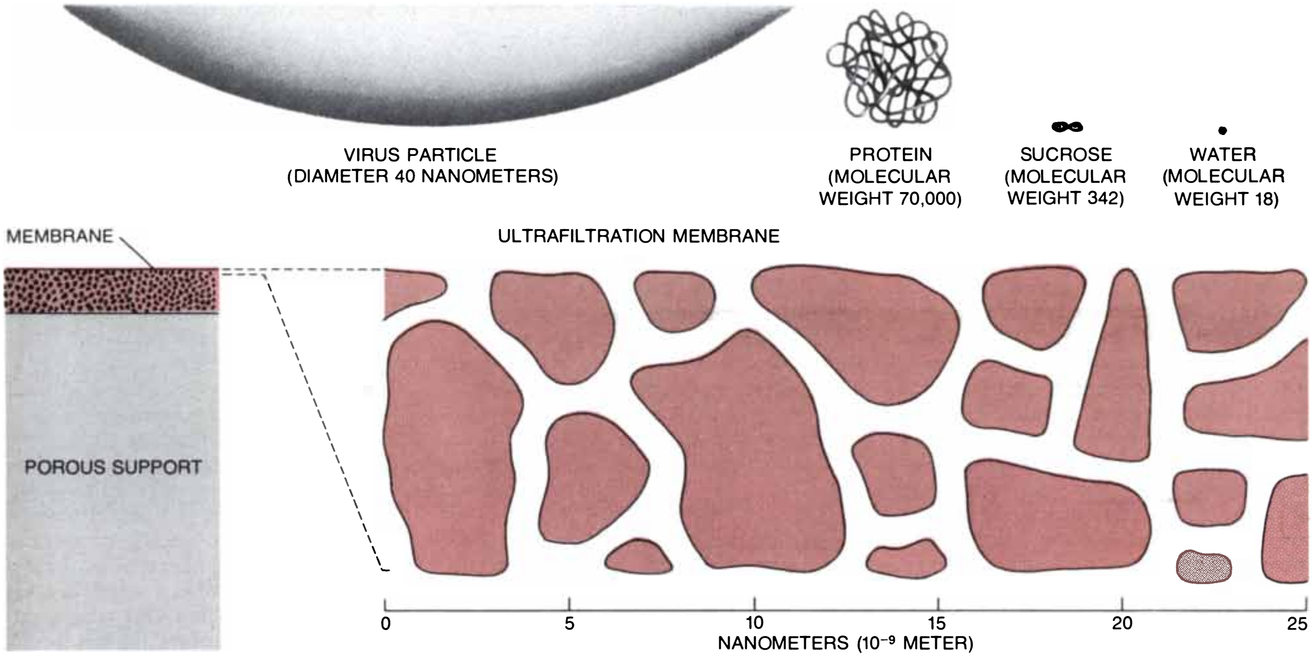
If the membrane system were left to itself, pure water would flow so as to dilute the solution, the opposite of the effect desired. It is therefore necessary to oppose the osmotic pressure with a counterpressure applied to the solution. If the applied pressure exceeds the osmotic pressure water will flow out of the solution, thus leaving behind a greater concentration of solute. The rate of flow is proportional to the difference between the two pressures. The rate is also affected by two other factors: it is directly proportional to the permeability of the membrane to water and inversely proportional to the membrane thickness. Because even a moderately concentrated solution can generate an osmotic pressure of a few hundred pounds per square inch, high applied pressures are often required in ultrafiltration.

A molecule that is too large to fit through the pores of a membrane is rejected absolutely by ultrafiltration, but even smaller molecules, those to which the membrane is permeable, can be concentrated to some extent. The flux of such solutes is determined predominantly by the process of diffusion, which describes the random movement of molecules and which is independent of pressure. Molecules of a permeant solute diffuse across the membrane in both directions, but the net movement is toward the region of lower concentration. The magnitude of the flux is directly proportional to the difference in concentration and inversely proportional to the membrane thickness. Since the movement of the solvent is also inversely proportional to the membrane thickness, the effective rejection of a permeant solute does not depend on the thickness of the membrane. It varies only with the applied pressure, since higher pressure increases the flow of water without affecting the flow of solute.

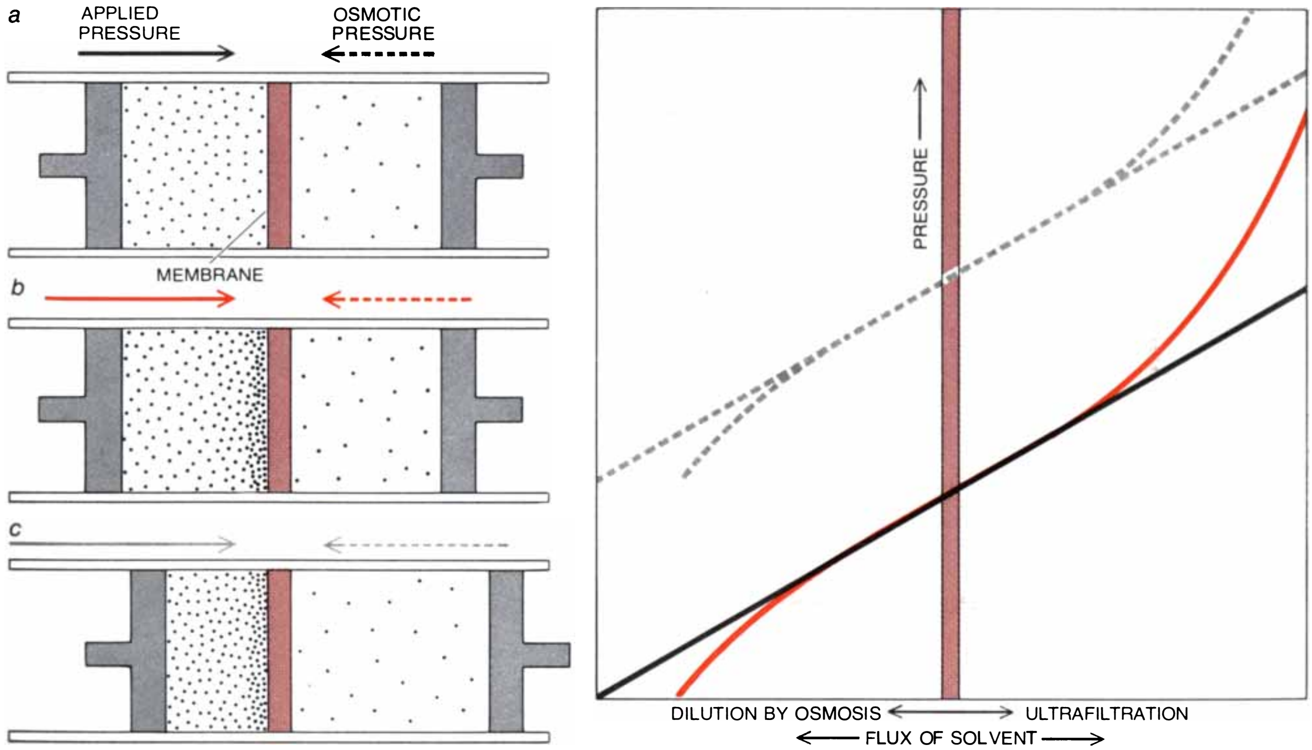
This analysis has assumed that the concentration both of water and of solutes is uniform throughout the volumes on each side of the membrane. Actually that condition cannot be met. Immediately adjacent to the membrane there is always a boundary layer of stagnant, unstirred solution where water is constantly being extracted and where the rejected solutes therefore pile up in unusually high concentration. This effect,



POROUS MEMBRANE is seen in a cross section magnified some 500 times in a scanning electron micrograph. The membrane was formed from a solution of polymers that was spread in a thin layer and allowed to dry partially, then plunged into water. The water quickly extracted the remaining solvent and thereby coagulated the polymer, creating the pores that appear in the micrograph as vertical channels. Finer pores near the top surface of the membrane were created by allowing the surface layers to dry more thoroughly, so that a "skin" formed on the polymer solution. The membrane, which was made by the Amicon Corporation, is designed for employment in ultrafiltration, a process in which not only suspended particulate matter but also many dissolved molecules can be removed from a fluid.



ULTRAFILTRATION MEMBRANE has pores of molecular dimensions. The membrane itself is a film some five to 30 micrometers thick cast on a much thicker porous substrate, such as paper or cloth, which provides mechanical support. Membranes can be fabricated with pores having an average diameter ranging from one nanometer to 10 nanometers (one nanometer being a billionth of a meter). The membrane shown in schematic cross section here (*bottom right*) has pores about one nanometer in diameter. Water readily passes through the pores, and so do many small molecules, such as the sugar sucrose. Most proteins are excluded, however, as are the much larger viruses and other microorganisms. An ultrafiltration membrane of the appropriate pore size can separate molecules such as sugars and proteins.



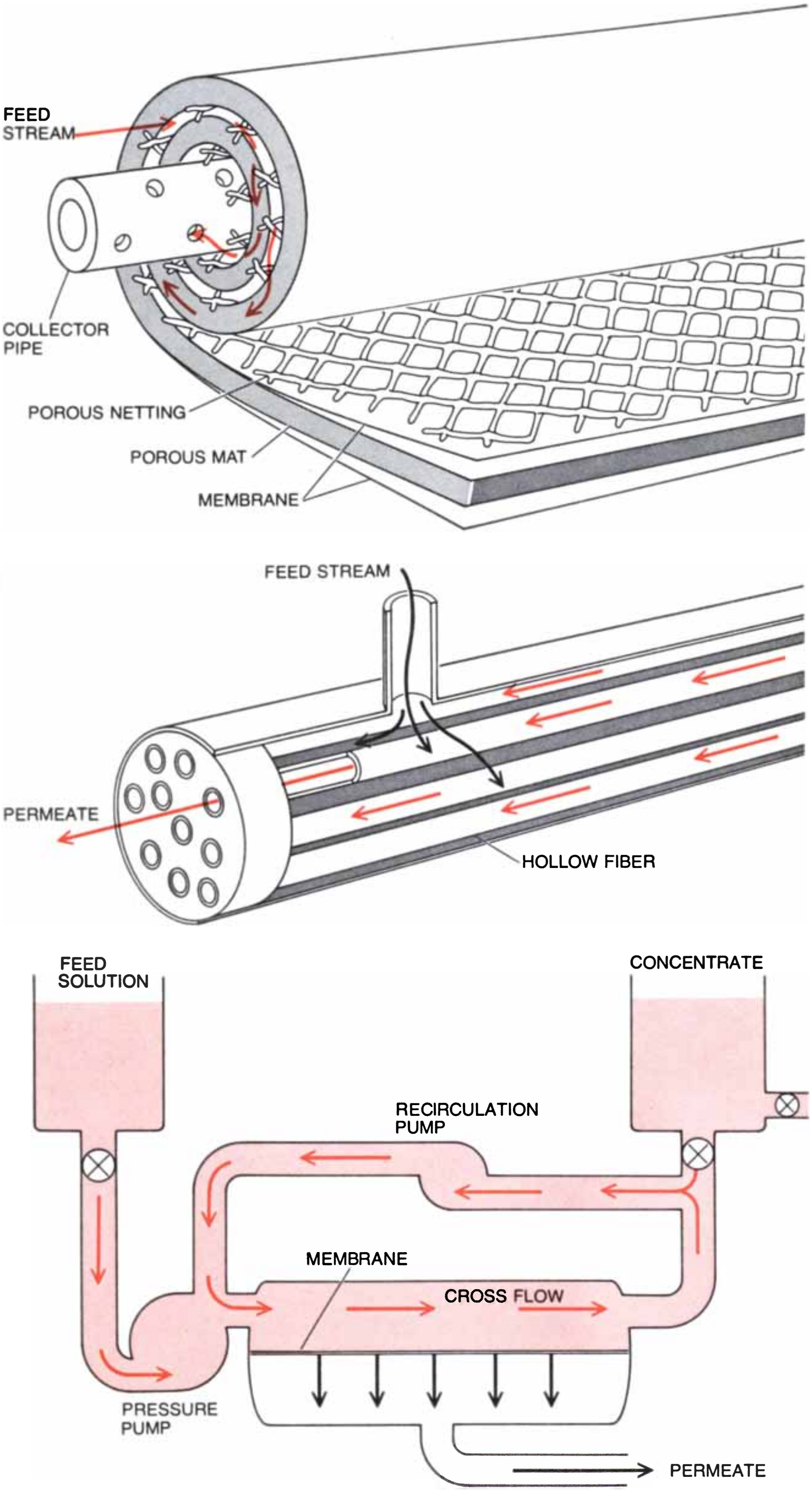
FLOW OF WATER or of another solvent through an ultrafiltration membrane depends in part on intrinsic properties of the membrane and in part on pressure. A difference in concentration generates an osmotic pressure that tends to drive pure water through the membrane and into the more concentrated solution, a flow in the direction opposite to that desired. If filtration is to take place, this osmotic back pressure must be opposed by a greater applied pressure. In the ideal case (*a, black curve*) the flux across the membrane is simply proportional to the difference between the applied pressure and the osmotic pressure. In practice any appreciable flow causes a buildup of solute molecules along the concentrate surface of the membrane (*b, colored curve*); this condition, called concentration polarization, increases the osmotic back pressure and therefore increases the applied pressure required to maintain a given flow. Moreover, if the feed solution is not continuously diluted, its concentration rises as filtering proceeds (*c, gray curves*), further increasing the osmotic pressure.

which is called concentration polarization, increases the osmotic pressure opposing filtration, since the difference in solvent concentration "perceived" is that of the fluids nearest the membrane. The degree of concentration polarization increases in proportion to the flux of solvent, and so it is also proportional to applied pressure. As a result there is effectively a maximum rate of flow per unit area for any membrane; beyond that maximum any further increase in applied pressure yields no greater flux but only results in raising the osmotic back pressure. Concentration polarization can also give rise to other undesirable phenomena such as the precipitation of solids at the membrane surface, and so it is avoided by limiting the applied pressure.

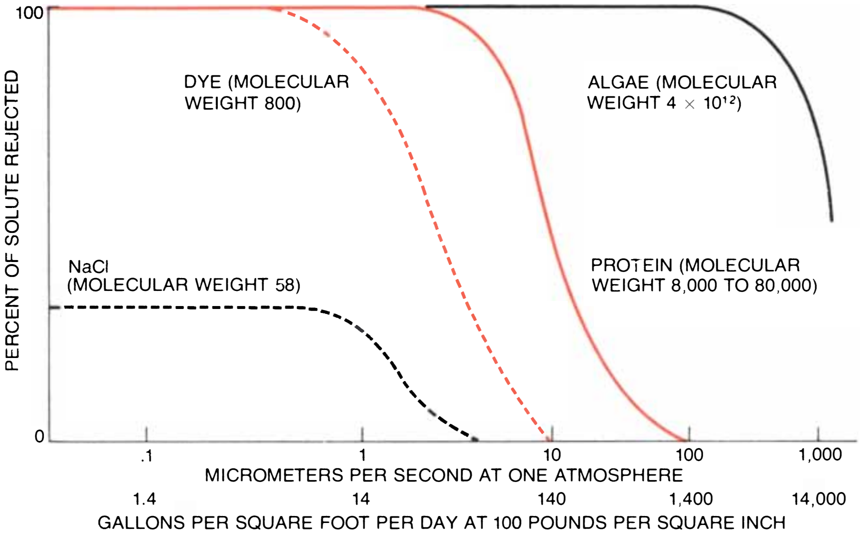
The flux through an ultrafiltration membrane can be measured in terms of the volume transmitted by a unit area of membrane in a given period and at a given applied pressure, or in terms of the velocity imparted to the stream by a given pressure differential. As might be expected, the flow increases as the pores become larger. The practical unit is gallons per square foot per day at an applied pressure of 100 pounds per square inch; the metric unit is micrometers per second at a pressure difference of one atmosphere.

Directly measuring the size of the pores in a membrane would be a difficult procedure, but the size can be estimated by measuring the rate of flow under standard conditions, which is determined largely by pore size. The measured flux can be calibrated by observing the rejection of molecules of various sizes by a membrane with a known rate of flow. Typical test molecules might be beef serum albumin (a protein of molecular weight 75,000), the dye erythrosin (molecular weight 800) and sucrose (molecular weight 342). A "tight" membrane, with a flux of one micrometer per second at one atmosphere (or equivalently 14 gallons per square foot per day at 100 pounds per square inch) rejects the larger molecules almost completely and excludes some of the sugar. A membrane of average flux, approximately three micrometers per second at one atmosphere, rejects the dye and the protein but allows most of the sugar to pass. A membrane characterized as "fast," with a flux of from 10 to 20 micrometers per second, rejects only the proteins, and one that is very fast (greater than 50 micrometers per second) allows even the protein to pass but excludes particulate matter such as microorganisms.

In ordinary filtration the feed stream with its suspended and dissolved materials is simply pumped through the filtering medium, which is often a bed of packed sand. Finely divided particles in the stream adhere to the filter and to



ULTRAFILTRATION APPARATUS must provide for a flow of concentrate solution across the face of the membrane in order to minimize fouling and to reduce concentration polarization. In one configuration (*top*) two membranes separated by a porous mat are wound with a spacer of plastic netting around a porous pipe. The solution is pumped through the spacer, and the permeate, after crossing the membrane, spirals through the mat to the pipe at the center of the "jelly roll," where it is collected. Synthetic membranes can also be spun into fine hollow fibers (*middle*) with an inside diameter of some 40 micrometers. The fibers are arranged in a bundle and the feed solution is pumped around them, with the permeate flowing out the ends of the fibers. In any ultrafiltration cell a pump supplies the pressure that drives the filtration. A second pump is often employed to recirculate the concentrate and to provide a cross flow.



PERMEABILITY OF MEMBRANES to molecules of various sizes is correlated with pore size, and that in turn is measured by the rate at which water flows through the membrane under a standard pressure. The tightest membranes allow substantial volumes of sodium chloride to pass, and the loosest ones block algal cells. Between these extremes some membranes reject only large molecules, such as proteins; others exclude molecules of moderate size such as dyes.

each other, forming a layer of dirt, which has come to be known by its German name schmutzdecke. Because the schmutzdecke soon becomes impermeable, the filter must periodically be either replaced or backwashed to maintain a satisfactory rate of flow. In ultrafiltration a different technique is employed. The feed solution is circulated continuously across the surface of the membrane, a procedure that not only helps to avoid fouling but also helps to reduce concentration polarization.

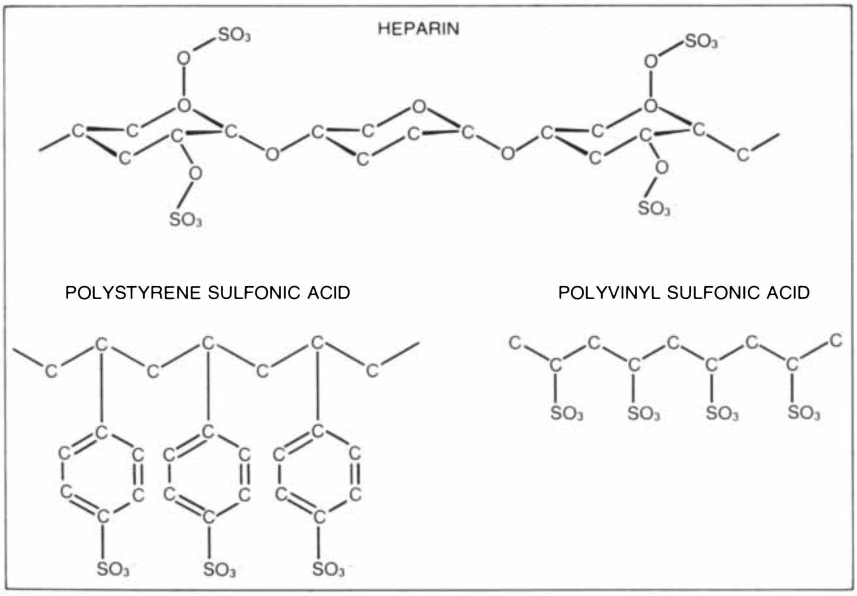
Ultrafiltration apparatus can be assembled in several ways. For example, parallel sheets of membrane can be mounted in a frame, or a porous pipe can be lined with a membrane. One efficient configuration is a "jelly roll" made by winding membranes and spacers around a porous pipe that serves as a collector for the permeate. Membranes are cemented to each side of a sheet of porous paper or cloth, then the double membrane is wound with a spacer of plastic netting around the collection pipe. The feed solution is pumped through the space between the layers of double membrane, and the permeate spirals through the porous medium to the pipe at the center. The design is substantially cheaper than the frame-and-plate or membrane-lined-pipe filters, and it ensures a strong crosscurrent along the surface of the membrane on the concentrate side.

In another favored configuration the membrane is spun in the form of a hollow fiber, usually no thicker than a human hair. Because the fibers have thick walls compared to their overall dimensions they can usually withstand pressures of as much as 1,500 pounds per square inch. The cost is about the same as it is for the spiral-wound construction, but the hollow-fiber geometry is susceptible to fouling. If the feed stream is passed through the hollow core the fibers tend to become clogged with particulate matter, and if it is circulated around the outside of the fibers the cross flow is inadequate.

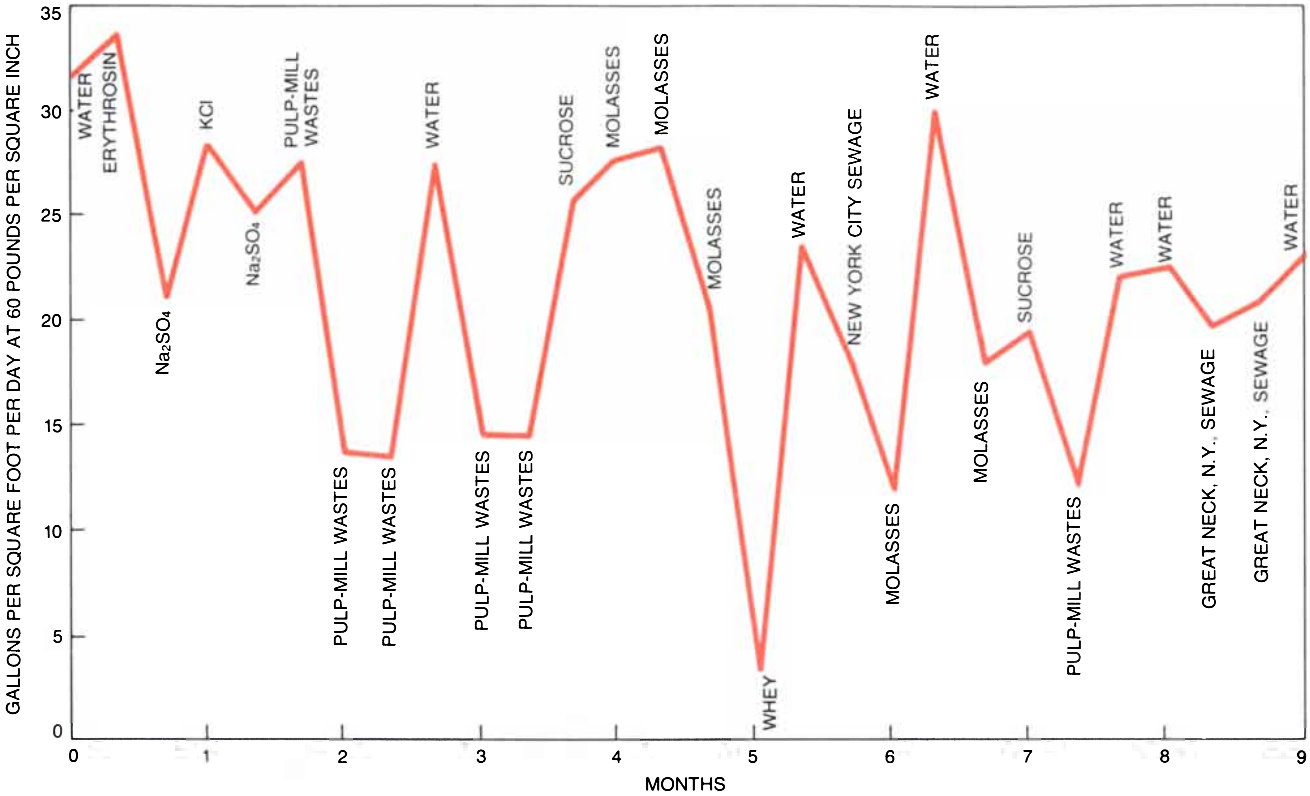
Ultrafiltration was developed in the 1920's, but for 40 years the process was rarely employed outside the laboratory, chiefly because of a single problem: the fouling of membranes. When a membrane was exposed to a feed stream containing suspended solids the particles adhered to the surface and clogged the pores. The rate of flow declined rapidly and within minutes or hours the filter became inoperable.

In about 1960 one of us (Harry P. Gregor) began a systematic study of membrane fouling with the aim of preparing membranes that would resist it. The most serious fouling problems, it soon became apparent, were caused by materials such as clays, oily particles and proteins, all of which have large surface areas that are hydrophobic, or in other words repel water. When a hydrophobic substance is in an aqueous environment, it can reduce its total energy by reducing the area exposed to the water; two hydrophobic particles, for example, tend to clump together, expelling the water from the space between them and thereby reducing their exposed surface area. In the same way such a particle can be held to the surface of a membrane by the elimination of repulsive interactions with the surrounding water. Most of the fouling materials also bear a negative electric charge, and hydrogen bonding involving these charges could also contribute to fouling. In this kind of bonding the slight positive charge of a hydrogen atom at the surface of the membrane attracts a negatively charged group in the fouling particle.

One obvious way to avoid hydrophobic interactions is to create a membrane that is extremely hydrophilic, one that has a very strong affinity for water. Such a material remains wetted even in the presence of hydrophobic particles, so



NONFOULING POLYMERS incorporate in their structure the sulfonate group (SO_3^-), which gives them an extreme affinity for water and thereby prevents most foreign matter from adhering. Most fouling particles are hydrophobic, or in other words repel water; most of them also bear a negative electric charge. Heparin is a biological antifouling agent that is thought to keep the blood vessels wetted. The active chemical groups in heparin are also sulfonates.



RESISTANCE TO FOULING of a membrane cast from a sulfonate polymer was evaluated by continuously exposing a single specimen membrane to a series of test liquids over a period of months. Several of the fluids, such as the wastes from pulp and paper manufacturing, molasses, whey and sewage, represent a severe test of filtration efficiency and would foul ordinary membranes in minutes or hours. The sulfonate membrane remained clean even at the end of the test period. The flow rates indicated were measured at the end of each test. Variations in the rate are probably caused less by fouling of the pores than by differences in the osmotic back pressure induced by each solution.

that the particles cannot adhere to the surface by excluding water.

The most hydrophilic of known polymers are those bearing on their surface the sulfonate group, SO_3^-. Like the sulfuric acid from which they are derived, sulfonate polymers have an enormous affinity for water. Each sulfonate group is ordinarily surrounded by many water molecules, which cannot be displaced by hydrophobic particles. The sulfonate group is also incapable of hydrogen bonding; indeed, its negative charge repels the negatively charged fouling particles. The negative charge of the sulfonate is also maintained even in the presence of a large concentration of hydrogen ions (that is, at low *p*H), because the sulfonate is the conjugate base of a very strong acid. Moreover, the sulfonate group resists deactivation by positive ions, such as those of the heavy metals. Even the lead and barium salts of sulfonate polymers remain quite soluble in water, indicating that the sulfonate groups are still hydrated, whereas most other lead and barium salts are almost completely insoluble.

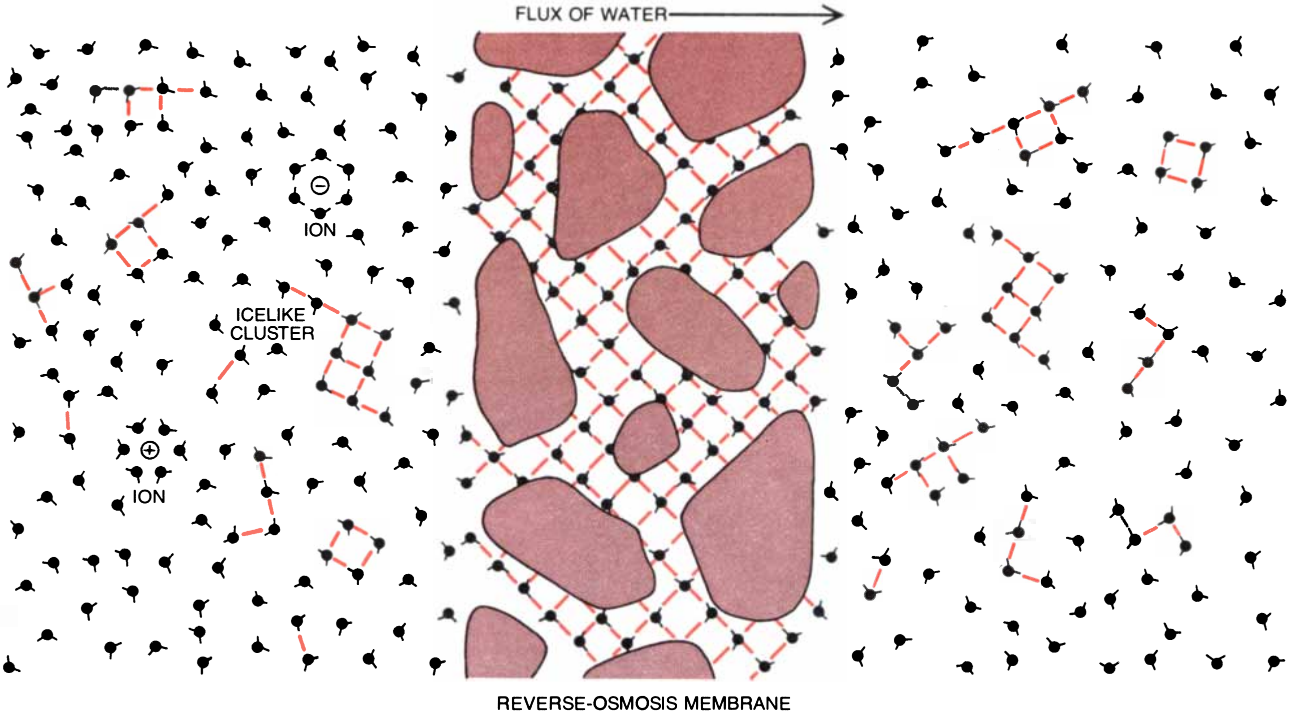
The choice of sulfonate polymers as a material for nonfouling membranes was encouraged by an example from nature: the colloidal substance heparin, secreted from the lining of the blood vessels. The function of heparin is thought to be that of keeping the walls of the blood vessels wetted and unfouled. In structure it is a sulfonate polymer.

Procedures were developed for casting membranes of styrene-based polymers with sulfonate groups exposed at the membrane surface and within the pores. In general these procedures involve forming a thin film of the sulfonate polymer, then adding a chemical agent that creates crosslinks among the long-chain molecules.

When the sulfonate membranes were tested with feed solutions that usually cause severe fouling, the results proved to be remarkable. Standard testing procedure had been to mount a membrane in the test apparatus and then to operate it with a single solution until its performance became seriously degraded. A new membrane was then installed for the next test solution. With the sulfonate membranes it was possible to expose a single specimen membrane to many fouling liquids in succession over a period of several months with negligible loss in filtering efficiency. One membrane, for example, was tested for nine months with a sequence of liquids that included salt solutions, dyes, wastes from pulp and paper mills, molasses, whey from cheese manufacture and sewage. The flux across the membrane varied with each solution, but the variation was probably caused not by fouling but by differences in the osmotic back pressure. At the end of the test sequence the membrane was still clean and glistening.

The resistance to fouling exhibited by sulfonic acid polymers may make them useful in fields other than membrane technology. One of us (Harry P. Gregor) has prepared screens and pipes coated with such polymers and has shown that they remain comparatively clean even under adverse conditions. Such a nonfouling coating might inhibit the adhesion of marine microorganisms on apparatus submerged for long periods in the sea. One potential application of the coatings might be on devices that would tap the energy of temperature gradients in the sea. The fouling of heat-transfer surfaces in such devices substantially reduces their efficiency.

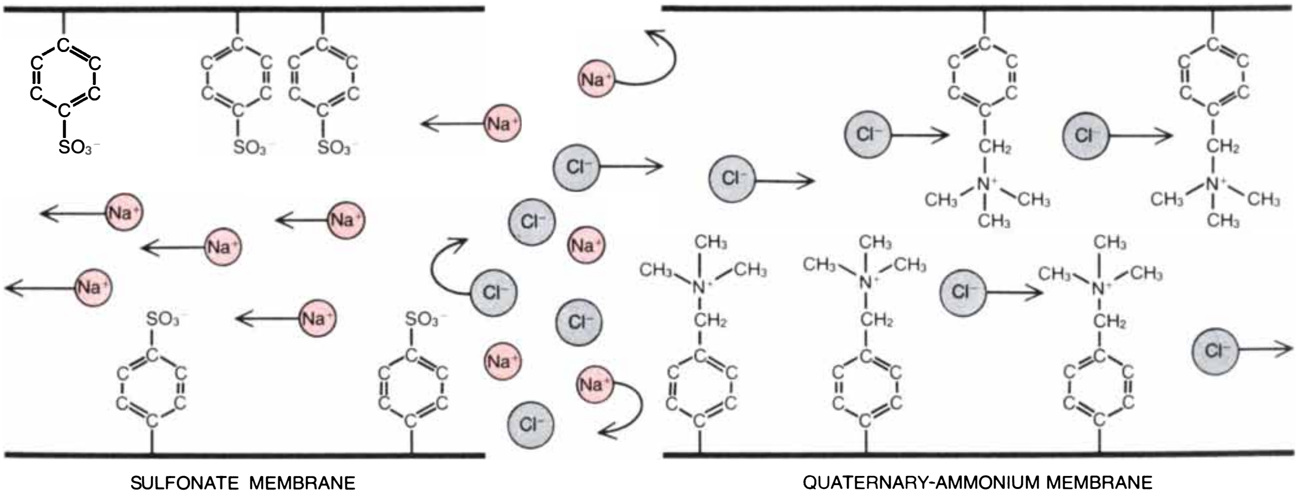
In addition to providing resistance to fouling, the fixed charges of the sulfonate groups confer another property on membranes: enhanced rejection of salts. Uncharged membranes, even those with quite small pores, have only limited ability to exclude salts, which exist in solution as negative and positive ions. In the sulfonate membranes, however, the fixed negative charges generate an electric field that repels ions of the same charge and thereby excludes them from the membrane pores. Since the electrical



REVERSE OSMOSIS is similar to ultrafiltration in that water is pumped through a membrane under pressure as solutes are excluded, but the mechanism is quite different. The reverse-osmosis membrane does not have definable pores, only spaces between polymer fibers where a small volume of water can be taken up. There the water assumes an icelike state (*color*) in which the molecules have an orderly arrangement and are held together by hydrogen bonds. (Even in the bulk liquid small icelike clusters are constantly forming and disintegrating.) The geometry of the icelike state is actually tetrahedral, with each oxygen atom surrounded by four others at equal distances, but it is shown here projected onto a plane. Under pressure water molecules join the icelike configuration at the concentrate side of the membrane (*left*) and melt away at the other side (*right*), so that there is an effective flow of water through the membrane. Other molecules and particles are rejected, including not only those that are too large to fit through the membrane but also small molecules that cannot conform to the icelike structure. Ions in particular are excluded because they are shielded by a layer of water that would disrupt the icelike lattice.

neutrality of the solution must be maintained, positive ions are also blocked.

The rejection of salts by charged ultrafiltration membranes is effective as long as the salt solution is comparatively dilute. In a concentrated salt solution there are enough positive ions available to surround and shield the negative charge of the sulfonate groups, allowing both kinds of ion to pass. Because of such shielding, ultrafiltration is not a practical technology for the desalination of seawater, which has about 35 grams of salt per liter. With brackish water, however, which contains about five grams of salt per liter, an 80 percent reduction in salinity can be achieved. What is more, divalent metal ions, such as the calcium ion Ca^{++} and the magnesium ion Mg^{++}, are preferentially removed by charged films. These ions



CHARGED MEMBRANES discriminate between ions in solution according to their electric charge. Single pores of two charged membranes are here shown schematically. The negatively charged membrane (*left*) has sulfonate groups along the walls of the pores and resembles a nonfouling ultrafiltration membrane except that the density of charged groups is greater. The sulfonate membrane allows positive ions, such as those of sodium (Na^+), to pass, but it blocks negative ions, such as those of chlorine (Cl^-). The positively charged membrane (*right*) has the opposite selectivity. The positive charges are provided by a derivative of the quaternary ammonium ion (NH_4^+).

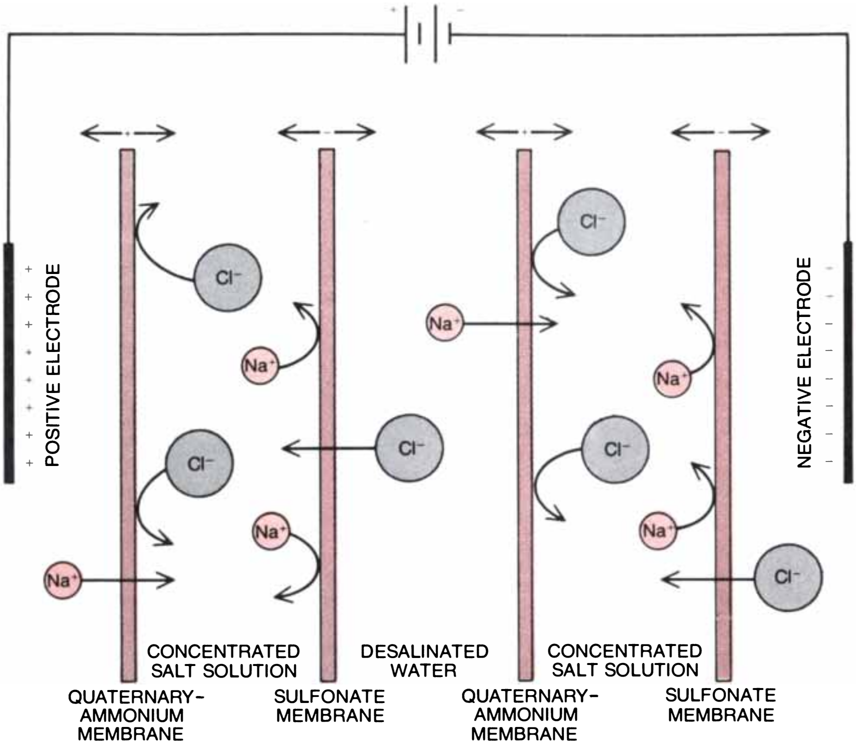
make the major contribution to water "hardness," and so the membranes can be employed for water softening.

If the pores of a cast polymer membrane are made progressively smaller, a state is eventually reached where the membrane is permeable to water but to almost nothing else. Water can be extracted from a solution by driving it through such a membrane under pressure, a process called reverse osmosis. In procedure reverse osmosis would seem to be nothing more than ultrafiltration at extremely small scale, but in fact its mechanism is quite different. One indication of that difference is that even molecules as small as water or smaller do not pass through the membrane. Indeed, the membrane has no definable pores, and even water does not flow through it in the usual sense; instead the water exists within the membrane structure in what is called an icelike state.

In liquid water about half the molecules at any moment are in clusters that have the same orderly structure as a crystal of ice, a fact first demonstrated in 1933 by J. D. Bernal and R. H. Fowler. In the clusters each water molecule is placed so that the oxygen atom occupies the vertex of a tetrahedron and so that a hydrogen bond connects each pair of water molecules. In ice this stable structure extends over long distances, but in the liquid state the icelike clusters generally include only a few molecules each, and they are constantly forming and disintegrating.

In a reverse-osmosis membrane a small volume of water is taken up in the spaces between the membrane fibers, where it assumes the icelike configuration. When pressure is applied across the membrane, molecules on the high-pressure side are incorporated into the icelike structure, replacing molecules that "melt" away on the other side. Dissolved substances, with very few exceptions, cannot conform to the icelike structure and so they are excluded. Ions in particular are surrounded in solution by water molecules aligned so as to shield the ionic charge, and the ions therefore cannot be made to fit the icelike matrix. The few exceptions are molecules capable of hydrogen bonding with water, such as methanol and urea; significantly, those substances can penetrate not only reverse-osmosis membranes but also ordinary ice.

Reverse osmosis has been known for more than a century, but because the permeability of such tight membranes is extremely small the flux of water through the membrane was always too slow to be practical. The solution to this problem was to make the membrane thinner: as was pointed out above, the flux of solvent is inversely proportional to the membrane thickness, but thickness has no influence on the rejection of solutes. Casting membranes thin enough to give a useful flux did not seem to be feasible, however.

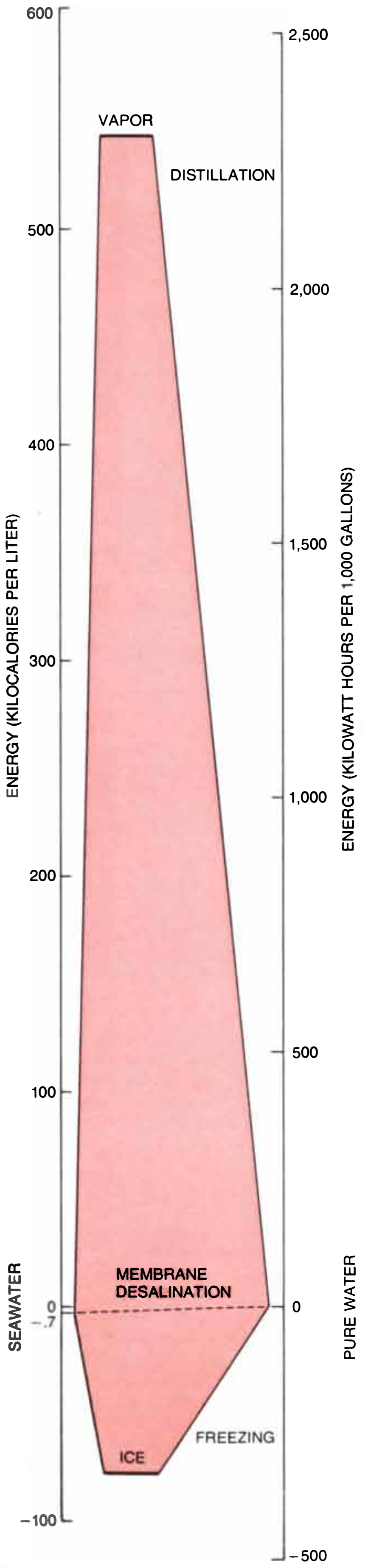


ELECTRODIALYSIS employs fixed-charge membranes to extract pure water from a salt solution. The process is driven not by applied pressure (as in ultrafiltration and reverse osmosis) but by an electric field. The membranes are identified at the top according to the ionic charge they allow to pass. Chloride ions are attracted by the positive electrode and therefore move to the left, crossing any sulfonate membranes in their path but being blocked by quaternary-ammonium ones. In the same way sodium ions move to the right until they are stopped by a sulfonate membrane. The result is that pure water accumulates in the central cavity and the salt solution is concentrated in the spaces lying between the two pairs of peripheral membranes.

In 1960 Sidney Loeb of the University of California at Los Angeles (and now of Ben Gurion University in Israel) discovered a method of casting membranes that are both thin enough and tight enough for practical reverse osmosis. His method takes advantage of a phenomenon observed when a membrane is cast by allowing a solvent to evaporate: after an initial period of drying a skin appears because the solution exposed at the surface dries faster than the layers underneath. The formation of such a skin is ordinarily avoided, so as to create a membrane of uniform porosity. Loeb deliberately created a skin, then plunged the membrane into water to coagulate the polymers in the deeper layers that had not yet completely dried. The result was a membrane with an extremely thin film of tight, coherent polymer over a much thicker layer of highly porous material. Such membranes are strong enough to withstand high pressure, but their active skin is only about 1 percent as thick as that of an ordinary ultrafiltration membrane. As a result the flux of water is 100 times greater.

A major application of reverse osmosis is the extraction of pure water from salt solutions, particularly brackish water and seawater. Unlike ultrafiltration membranes, the reverse-osmosis membranes do not become permeable to ions at high salt concentrations, but there is nonetheless a limit on the extent to which the feed solution can be concentrated. For a practical flux of water the net pressure across the membrane must be at least a few hundred pounds per square inch. When the concentration of the salt solution reaches about 3 percent the osmotic pressure amounts to some 400 pounds per square inch, and concentration polarization can effectively raise it to about 600 pounds per square inch. The cost of pumps and of other equipment limits the applied pressure to about 1,000 pounds per square inch, so that further concentration is impractical. Pure water can be extracted from seawater by reverse osmosis if the rejected salt can be infinitely diluted by a very large reservoir (such as the ocean). If the salts must be concentrated for consumption or disposal, however, reverse osmosis must be combined with other processes that are not limited in efficiency by the concentration of the feed solution.

Desalination by reverse osmosis pro-



ENERGY REQUIRED for the desalination of seawater is in theory independent of the means of extraction, but in practice differences in efficiency introduce large differences in the energy requirement. The irreducible quantity of energy needed to recover a liter of pure water from a large volume of seawater is .7 kilocalorie. In distillation the energy must initially be raised to 540 kilocalories, but with a perfect heat exchanger all that energy except the .7 kilocalorie could be recovered. Actually much of it is lost. Similarly, removing the salt by freezing requires an investment of −79 kilocalories, only some of which can be recovered. Because membrane desalination is not a thermal process there is no similar requirement for an energy investment, and the efficiency is quite high. The energy consumption is about three kilocalories per liter, or roughly four times the theoretical minimum.

vides an illustration of a property common to many membrane processes: they require less energy than comparable thermal means of separation. To extract one liter of pure water from an unlimited volume of seawater requires a minimum of .7 kilocalorie; this is the thermodynamic, or theoretical, minimum energy, and it is a limiting value no matter how the separation is performed. The thermodynamic minimum can be approached only when all the stages of a process are reversible, that is, when they operate near equilibrium. Evaporation and condensation can be carried out in a way that is very nearly reversible by performing both processes at the same temperature and pressure. Under such equilibrium conditions, however, thermal processes are exceedingly slow and require enormous equipment. The speed can be increased by operating with a large temperature difference, but that departure from equilibrium and from reversibility has a large cost in energy. Extracting a liter of water by distillation, for example, calls for some 540 kilocalories, only some of which can be recovered in a heat exchanger. Membranes, on the other hand, function quite well near equilibrium. The liter of water can be separated by reverse osmosis for an investment of some three kilocalories, or only about four times the minimum energy.

The force driving both ultrafiltration and reverse osmosis is the applied pressure, but the separation of molecules by membranes can also be powered by other forces. In one such process, called electrodialysis, energy is supplied by an electric field applied across a membrane or a set of membranes. In dialysis (as distinguished from osmosis) it is the solutes that move across the membrane instead of the solvent; in electrodialysis the solutes are ions and the membranes are selectively permeable to those ions bearing either a positive or a negative charge.

In 1923 the German biochemist Leonor Michaelis showed that membranes made from dried nitrocellulose are selectively permeable to positive ions, but he did not explain the mechanism of this effect. In 1929 Karl Sollner began the first systematic study relating the structure of membranes to their functioning. With his students Charles Carr and one of us (Harry P. Gregor) he fabricated the first membranes designed not only for high selectivity but also for optimum permeability.

The technology of electrodialysis is now well developed and membranes for it are manufactured in commercial quantities. The membranes have extremely fine pores (about one nanometer in diameter), and they are made from cross-linked polymers with a dense concentration of fixed charges. The negatively charged membranes are made from polystyrene with sulfonate groups attached, and they therefore reject all negatively charged ions. They are similar to sulfonate ultrafiltration membranes but have a much lower permeability to water and a much denser concentration of fixed charges. Because of the large number of charged groups the membranes are highly selective even in salt solutions as concentrated as 175 grams per liter.

The corresponding positively charged membranes are constructed on a framework of the polymer polyvinyl benzyl chloride, and the necessary charge is supplied by groups derived from the quaternary ammonium ion (NH_4^+), such as trimethylammonium. As would be expected, these membranes exhibit the opposite selectivity: they reject positive ions and pass negative ones. The sulfonate membranes, like the similar ones employed in ultrafiltration, are highly resistant to fouling, but the quaternary-ammonium ones are not. The latter foul severely if they are exposed to fluids containing natural products, and for electrodialysis it is therefore sometimes necessary to pretreat the solution by ultrafiltration.

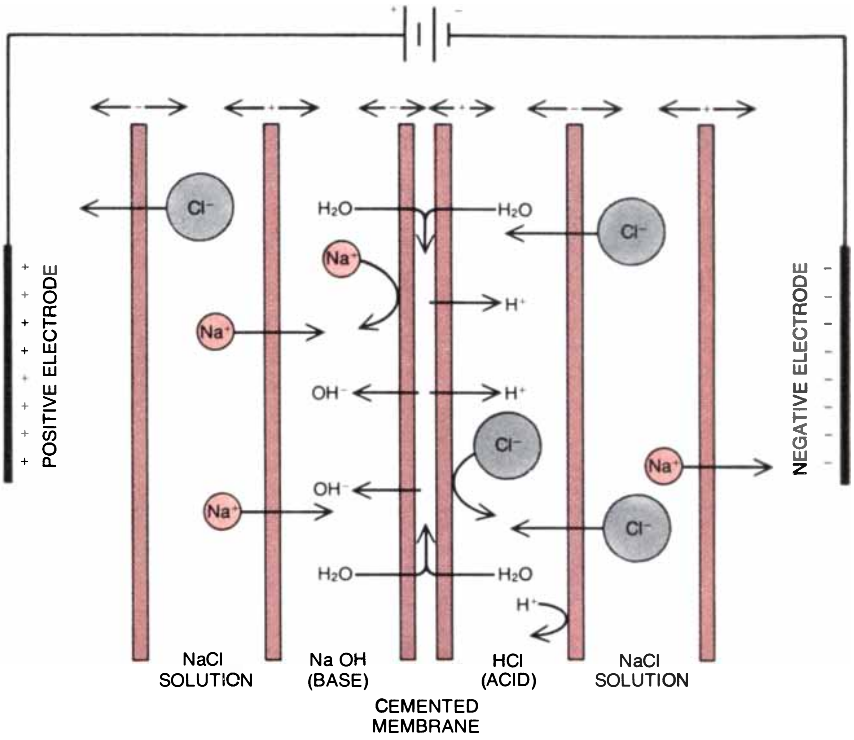
In an electrodialysis cell, membranes of alternating polarity are arranged in a stack and an electric field is applied across the entire stack. Positive ions are attracted by the negative electrode and can cross a sulfonate membrane to reach it; they are then stopped, however, by the next membrane in the stack, which is a quaternary-ammonium one. Negative ions move in the opposite direction until they are stopped by a sulfonate membrane. The result is that alternating cavities between membranes fill with purified water and with concentrated salt solution.

In Japan most table salt is extracted from seawater by a method that includes a stage of electrodialysis. After magnesium and sulfate ions have been removed by another membrane process electrodialysis concentrates the sodium chloride solution about sixfold, to 17 percent. The remaining water is removed by evaporation. The energy requirement for the electrodialysis is about 300 kilowatt hours per ton of salt, and an additional 250 kilowatt hours is consumed in evaporating the remainder of the water. If the entire work of concentration were done by evaporation, the energy requirement would be 2,100 kilowatt hours per ton, so that the hybrid technique consumes only about a fourth the energy.

In 1913 Albrecht Bethe (a physiologist and the father of the physicist Hans Bethe) mounted an animal membrane between two solutions and passed an electric current through the system in order to observe the resulting movement of water molecules. He was surprised to discover that the solution on one side of the membrane became slightly acidic and the solution on the other side slightly basic. He established that the effect resulted from the splitting of water molecules into hydrogen ions (H^+) and hydroxyl ions (OH^-). The phenomenon discovered by Bethe remained a laboratory curiosity until it was observed that commercial electrodialysis membranes also split appreciable quantities of water. It then became apparent that the membranes could be employed to make an acid and a base from a salt.

It is important to distinguish water splitting from the superficially similar process of electrolysis, in which the products are molecular hydrogen (H_2) and oxygen (O_2). The decomposition of the water molecule into its component atoms has an irreducible energy requirement, whereas the dissociation of the molecule into ions takes place spontaneously, without an input of energy. In pure water about one molecule in 10^7 is ionized at any moment. The electric current required in the membrane water-splitting technique serves primarily to separate the ions after they have formed, although it has recently been found that the electric field does slightly enhance ionization.

A water-splitting cell can be constructed from membranes quite similar to those employed in electrodialysis. Two membranes of opposite polarity are cemented back to back and these are surrounded by two more membranes on each side, again of alternating polarity. The entire cell is immersed in a salt solution, such as sodium chloride, and an electric current is passed through the system. Hydrogen and hydroxyl ions forming within the space between the cemented membranes are separated according to charge; they cross one membrane and then are rejected by the next one, and so combine with sodium and chloride ions trapped in the same intermembrane space. On one side of the double membrane hydrochloric acid (HCl) is formed and on the other side sodium hydroxide (NaOH).



WATER-SPLITTING CELL employs fixed-charge membranes to prepare an acid and a base from a salt solution. The two membranes in the center are cemented together but have a small space between them that water can enter. As in any water, a small fraction of these water molecules (about 10^{-7}) ionize spontaneously. The ions are then separated by the electric field, hydrogen ions (H^+) moving to the right and hydroxyl ions (OH^-) to the left. At the same time sodium ions (Na^+) are concentrated in the compartment to the left of the cemented membranes and chloride ions (Cl^-) are concentrated to the right. Recombination of the ions from water with the ions from the salt yields sodium hydroxide and hydrochloric acid.

A problem of early water-splitting membranes was the desiccation of the double membrane as the hydrogen and hydroxyl ions themselves and also water molecules surrounding them were drawn out of the interior space. At the high current densities required for industrial applications the membranes soon became inoperable. Recently one of us (Harry P. Gregor), together with Kenneth Brennen and Bruce Benjamin, has developed methods for avoiding desiccation by the careful joining of selected membranes.

The technology of water splitting is still under development, but it has a number of promising applications. In the mining of metals, for example, acid and base might be produced together at the mine site. The acid could then be employed in extracting the metal from the ore, while the base would be reserved for neutralizing waste liquids before they are discharged.

A modification of the water-splitting technique dissociates a salt formed from a weak acid and a weak base into its component gases. The salt ammonium bisulfide (NH_4HS), for example, is a byproduct of petroleum refining, formed when sulfur is removed from crude oil by treatment with ammonia gas. A water-splitting membrane could break down this waste material, yielding valuable ammonia and hydrogen sulfide.

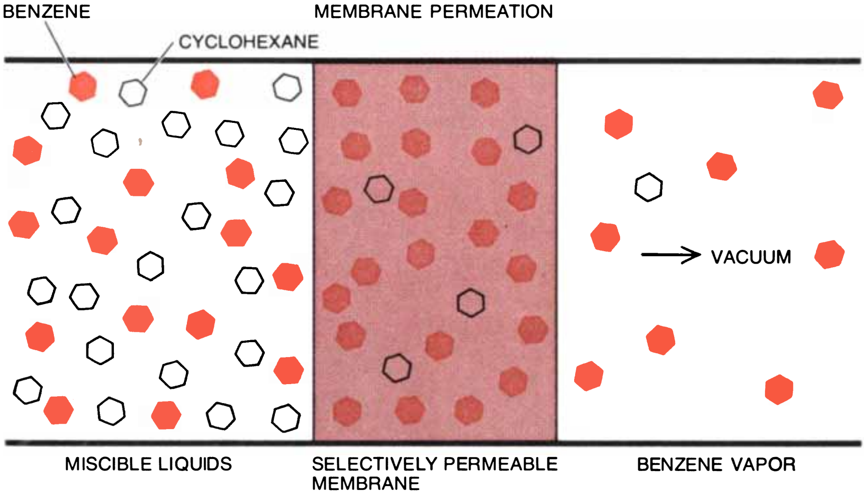
In addition to size and electric charge, a third criterion for discriminating between molecules is their solubility in various solvents. Differences in solubility are exploited in the process called solvent extraction. In its simplest form solvent extraction consists in purifying a substance dissolved in water by shaking the solution with another solvent, such as oil, that is not miscible with water but in which the substance is more soluble.

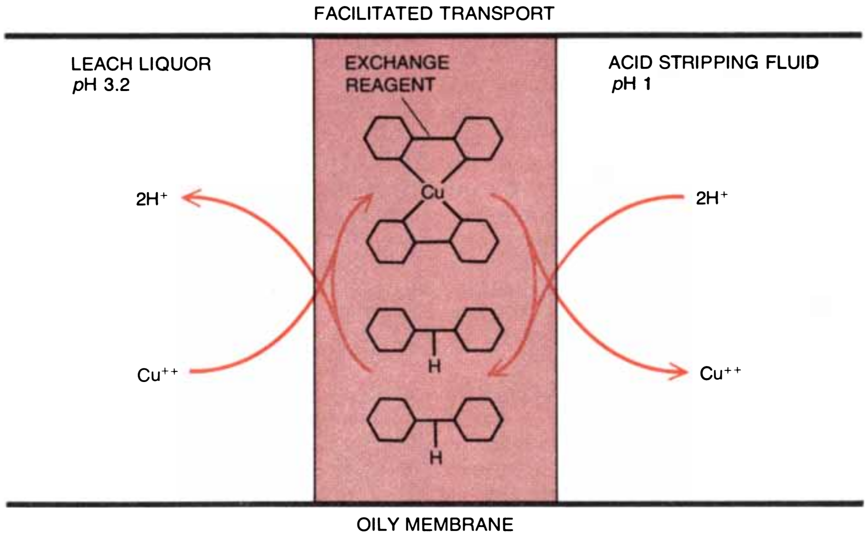
Solvent extraction across membranes has been known at least since 1913, when Fritz Haber and Reinhardt Beutner showed that a thin film of oil could be employed as a membrane in two kinds of extraction process. In one process, now called membrane permeation, a mixture of two miscible liquids, such as benzene and cyclohexane, is placed on one side of a membrane and a vacuum is drawn on the other side. If the membrane is more permeable to one component, such as benzene, the vapor appearing on the vacuum side is largely benzene. Unlike distillation, membrane permeation requires only one component of a mixture to be vaporized, thereby reducing the energy required.

Another form of solvent extraction

devised by Haber and Beutner, called facilitated transport, resembles in some respects a mechanism presumed to operate in the active transport of molecules across biological membranes. Aqueous solutions, one containing the substance to be isolated, are placed on each side of an oily membrane. A mobile molecule within the membrane then forms a complex with the selected species; the complex diffuses across the membrane and dissociates on the opposite side.

Lawrence Brandlein and one of us (Harry P. Gregor) have shown that copper can be extracted by this method from the solution in which it is leached from the ore. The copper is present with other metals in a weakly acidic solution, which is placed on one side of a membrane impregnated with a kerosene solution. None of the metal ions is soluble in kerosene, and so the membrane is impermeable to them. The copper ions are taken up into the membrane by an organic reagent that selectively forms a kerosene-soluble complex with copper but not with other metals. At the other side of the membrane the copper is stripped from the complex by a strong acid. In this way the copper is extracted from the leach liquor and isolated from the contaminating metals in one operation. Solvent extraction across membranes might also be applied to the reprocessing of nuclear fuels.





SOLVENT EXTRACTION across a membrane separates molecules according to their different affinities for a solvent or for a specialized transport molecule within the membrane. In one form of solvent extraction, called membrane permeation (*top*), a mixture of two miscible liquids, such as benzene and cyclohexane, is placed on one side of a membrane that absorbs one component of the mixture more readily than the other. A vacuum on the other side of the membrane then draws through mainly molecules of the more permeant component, in this case benzene. A second process, called facilitated transport (*bottom*), can be employed to extract copper from the acid that leaches the metal from the ore. The membrane employed is impermeable to copper (and to other metal ions) but includes an exchange reagent that selectively forms a complex with copper. The complex forms at the concentrate surface of the membrane and diffuses to the other side, where the copper is stripped away by a stronger acid, regenerating the reagent.

In one remaining membrane process the selective capabilities of the membrane are not exploited; the membrane serves primarily as an inert matrix on which enzymes are mounted. In 1954 Nikolaus Grubhofer showed that an enzyme could be chemically coupled to a synthetic matrix without loss of activity. Subsequent work by Ephraim Katchalsky and Georg Manecke led to the development of chemical reactors in which enzymes are bonded to porous, insoluble beads. (Katchalsky, who later took the name Katzir, has just retired as president of Israel.) Such reactors are now in commercial service (for example converting glucose into invert sugar) but they have certain deficiencies. Bonding the enzyme to the beads usually requires the simultaneous presence of both the enzyme and a corrosive coupling agent, with the result that much of the enzyme is often destroyed. Moreover, because reactants and products circulate only by diffusion, the process is rather slow.

Paul Rauf and one of us (Harry P. Gregor) have shown that these difficulties can be avoided by coupling the enzyme to a membrane with pores about five times as large as the enzyme molecules. The membrane is prepared by pumping a coupling agent, such as cyanogen bromide (CNBr) through the pores, forming an unstable intermediate capable of binding the enzyme. After the excess coupling agent has been washed out the enzyme is introduced under pressure. By controlling the pressure and the reaction time the pores can be coated with a dense layer of enzyme, which preserves as much as 80 percent of its native activity. The enzyme is essentially spot-welded to the membrane, and the enzymatic activity of a particular molecule is affected only if the weld happens to involve a position near the active site.

In one experiment a membrane was impregnated with the enzyme chymotrypsin, which normally digests proteins but is also capable of decomposing an ester into an organic acid and an alcohol. The rate of conversion for esters was found to rise with the feed pressure, and only at an applied pressure of about 100 pounds per square inch was the catalytic capacity of the enzyme reached.

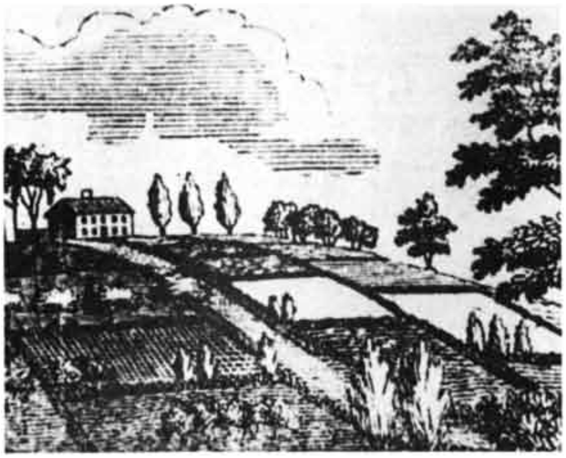
Several applications of synthetic membrane technology have been described above. Among other potential applications three industries that might employ membranes on a particularly large scale can serve as illustrative examples. They are sewage treatment, the food processing industry and the pro-

SCIENTIFIC AMERICAN
Filmstrips

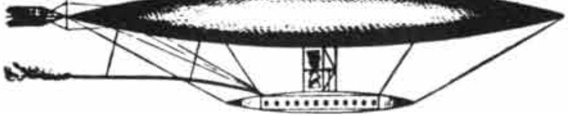
Rufus Porter

A Scientific American

"We have been frequently solicited by enterprising mechanics to engage in the publication of a new scientific paper, for the advancement of several new, curious and useful arts which have recently been discovered and introduced . . ."



"It is not infrequently the case that the productions of self taught artists far surpass in excellence those of regular bred artists who have studied with the most popular masters . . ."



"It may be anticipated that within a few months these aerial machines may be seen soaring in various directions and at different elevations, some above the clouds, and others, like swallows, sailing leisurely just above the surface of the earth."

Rufus Porter—artist, inventor, journalist and founder, in 1845, of SCIENTIFIC AMERICAN—should be better known to his countrymen. His exuberant murals adorn 100 New England homes and are finding their way into museums as treasures of American art. His long list of inventions was crowned before 1850 by his demonstration of a self-propelled flying machine. His SCIENTIFIC AMERICAN spoke for the buoyant rationalism that committed this country to industrial revolution from early in the 19th Century.

The 20-minute filmstrip for DuKane projectors *Rufus Porter, A Scientific American* [$22, shipping included] brings this Yankee Leonardo back again to life. In his own colorful words and pictures it recounts the story of his untiring enterprise and places him in the context of the era in American history he so well personified.

Recommended for High School and Junior High School classes in American History, Social Studies and Art History. Available in filmstrip and tape cassette for DuKane projectors.

SCIENTIFIC AMERICAN, Inc.
415 Madison Avenue, New York, N.Y. 10017, Dept. RP

Please send me ____________ copies of *Rufus Porter, A Scientific American,* a filmstrip presentation for DuKane projectors, at $22 (shipping included). I understand I may return same and receive a full refund if I am not completely satisfied.

Enclosed is my check for $ ____________.
Please include appropriate sales tax.

Name (please print)

Institution (if any)

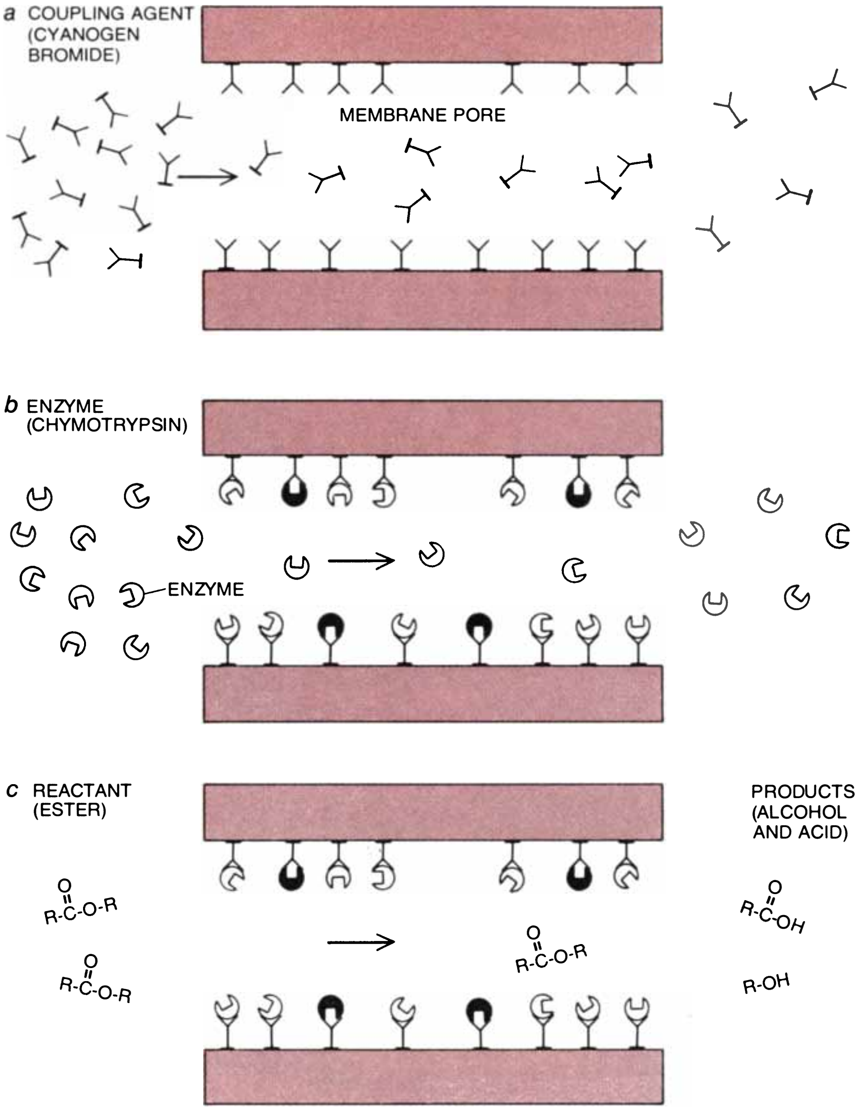
Address

 ate Zip

posed cultivation of green plants as a source of fuel.

Conventional sewage treatment employs a combination of large settling tanks, bacterial cultures and sludge-thickening devices to decontaminate waste water and to concentrate the solid residue. The primary treatment, in which larger particles are allowed to settle, probably cannot be improved by membrane technology. The present secondary treatment, however, could be replaced by ultrafiltration through a nonfouling membrane. Kang Hsu, Sadi Mizrahi and one of us (Harry P. Gregor) have shown that the effluent of such a filtration is of a quality equal to or superior to that of conventional secondary treatment. Indeed, since bacteria and viruses cannot penetrate the membrane, the permeate solution is almost sterile. What is more important, because the solids filtered from the stream are not decomposed, they settle more readily and are more easily concentrated, facilitating final disposal. The concentrated sludge will support its own combustion in an incinerator; also, since it is from 20 to 30 percent protein, it is a potential source of nitrogenous fertilizer.

Because ultrafiltration units are modular, small treatment plants could be



ENZYMATIC REACTOR can be constructed by binding an enzyme within the pores of a polymer membrane. A coupling agent, such as cyanogen bromide (CNBr) is first pumped through the membrane and binds to the walls of the pores (*a*). The enzyme itself, in this case the digestive enzyme chymotrypsin, is then introduced, again under pressure, and binds to the coupling agent (*b*). Some enzyme molecules are deactivated, mainly those that happen to bind to the coupling agent near the active site of the enzyme, but as much as 80 percent of the activity of the enzyme can be preserved. Once the membrane has been prepared a substrate of the enzyme can be pumped through it, the products of the enzymatic reaction appearing on the other side (*c*). Chymotrypsin, for example, splits an ester into an organic acid and an alcohol.

economically feasible. Indeed, the possibility of employing ultrafiltration as an adjunct to septic-tank treatment for individual homes is under consideration. Ultrafiltration might also have special advantages for waste-water treatment in desert regions such as the American Southwest, where salts must be removed along with other contaminants.

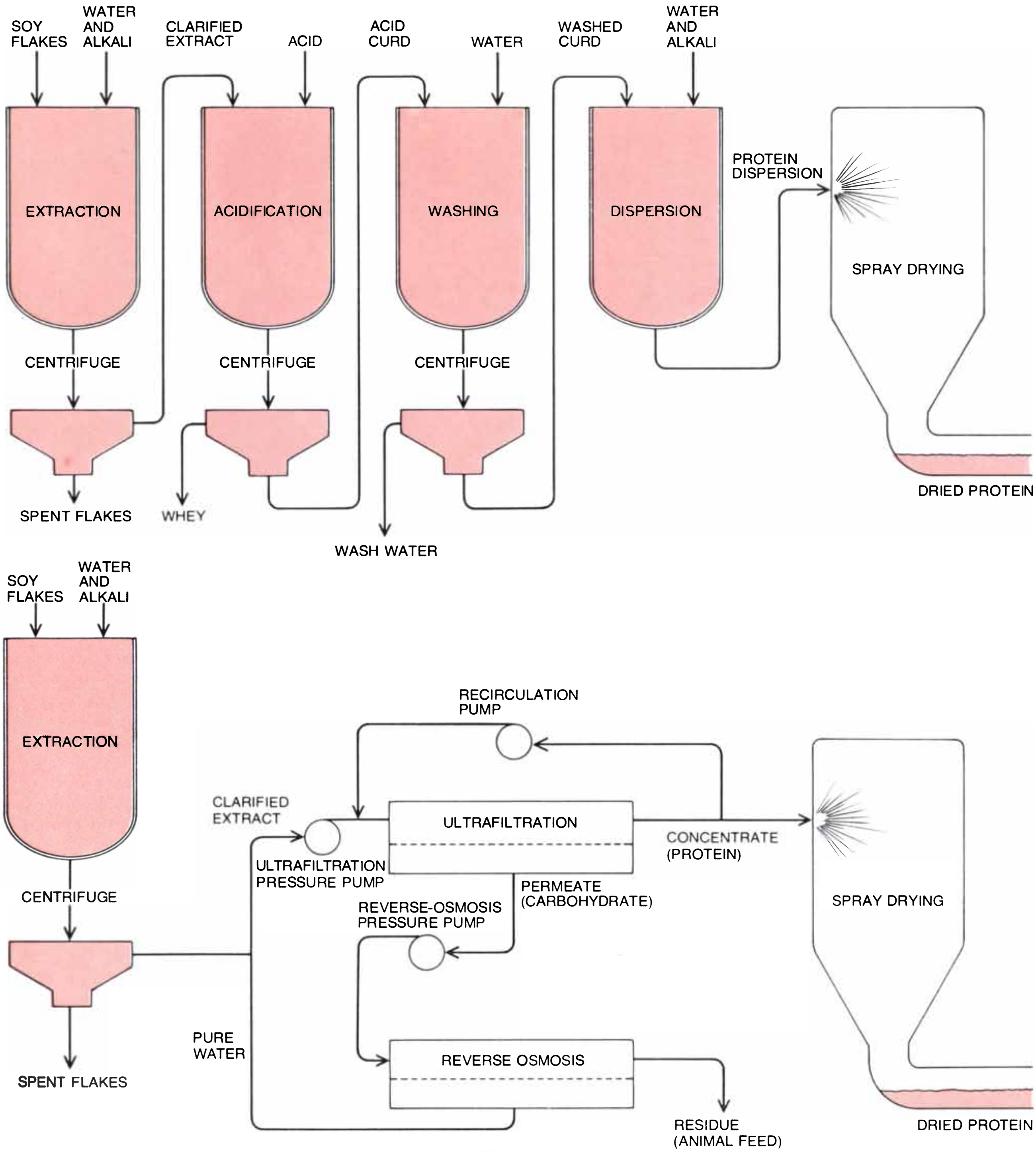
In the manufacture and refining of many foods, such as cheese and soybeans, waste liquids are created that contain large quantities of nutrients, but in concentrations too low to make their recovery economically feasible. If the liquids are simply discarded, however, they can cause severe pollution.

In the manufacture of cottage cheese the waste liquid is whey, which is produced in the U.S. at the rate of about five billion gallons a year. Whey includes about 3 percent of the sugar lactose, 1 percent of the milk protein lactalbumin and smaller amounts of ash and lactic acid. It can be concentrated by evaporation and spray-dried for use as an animal feed, but the energy required for the process makes it economically unattractive. Recent studies have shown, however, that the ultrafiltration of whey yields a concentrate made up of from 10 to 20 percent protein, together with a clear permeate containing lactose, salts and lactic acid. The lactose can be further concentrated by reverse osmosis to about 24 percent, and the pure sugar can then be recovered by evaporation. Alternatively the lactose solution can be passed through an enzymatic reactor that converts it to the simple sugars glucose and galactose, which are more palatable than lactose.

Membrane technology can be employed in a similar way in the purification of other wastes and by-products of food refining. The extraction of protein from soybean meal, for example, yields liquids that contain both the desired protein and waste carbohydrates and salts; these could be separated and purified by a combination of ultrafiltration and reverse osmosis. The wet milling of corn produces a waste stream called high-solubles water; a typical large plant must dispose of 20 million gallons a day, which is now evaporated, at a cost of $200,000 a day, to 50 percent concentration and sprayed on silage as animal feed. Ultrafiltration could remove all the suspended matter and dissolved protein and recover 99 percent of the water, which could then be purified to drinking-water quality by reverse osmosis. Only the residue of this second membrane treatment, containing 13 percent solids, need be further concentrated by evaporation for use as animal feed. An economic evaluation of the membrane process suggests it would require only 20 percent as much energy as the direct evaporation of the waste stream.

One promising method of capturing solar energy lies in the cultivation of efficient crops that can be converted into a liquid fuel, such as methanol. Among the plants being considered as fuel crops are algae, which are efficient converters of solar energy and which can be grown in a variety of environments. There are two major impediments to such an energy economy, however: the algal cells, which range in size from one micrometer to 10 micrometers, must be harvested quickly and cheaply from the ponds in which they would be grown, and the enormous supply of nutrients required by the growing plants must be conserved. Membrane technology could alleviate both of these problems.

The algal cells are present in water



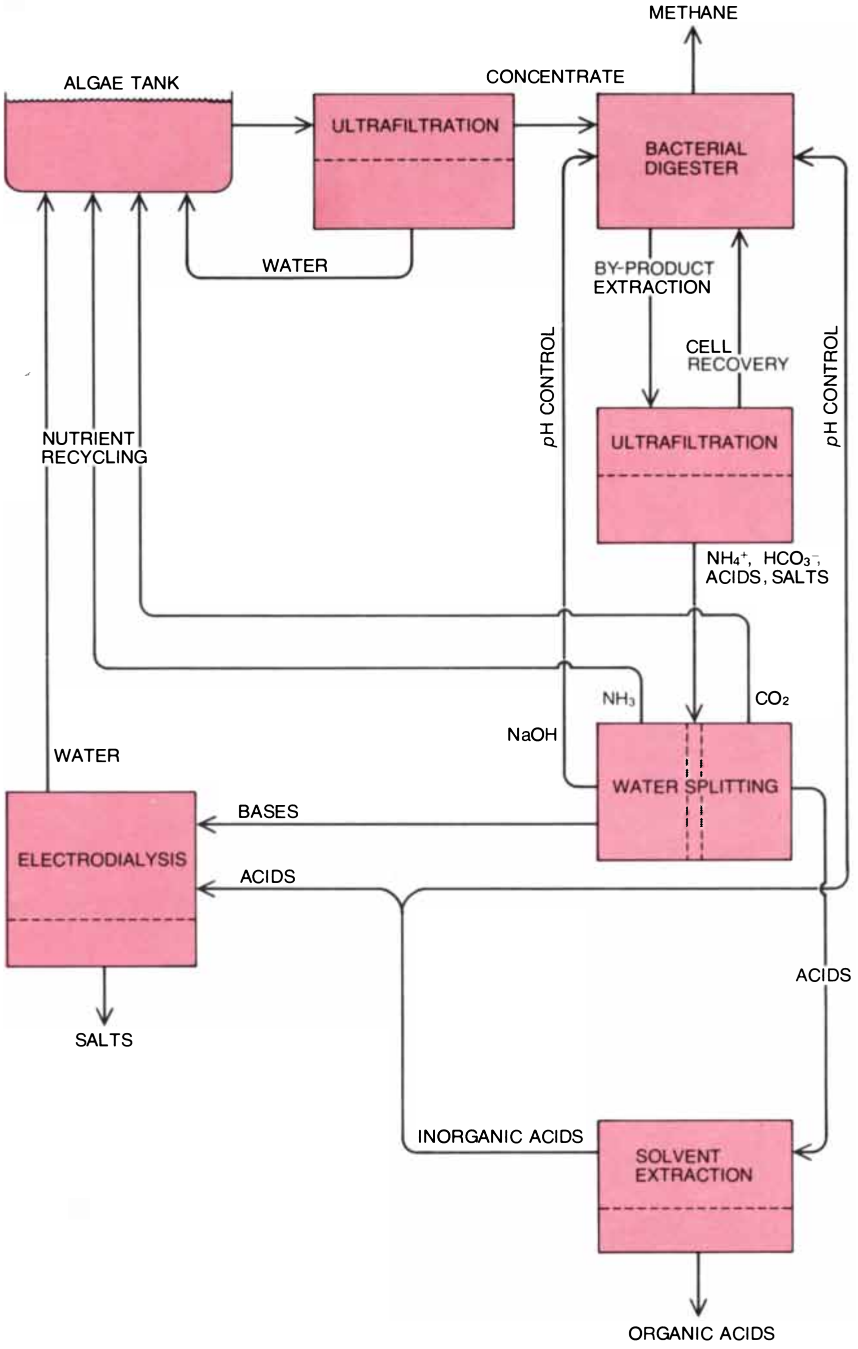
EXTRACTION OF PROTEIN from soybeans might be simplified by synthetic-membrane technology. In the conventional extraction (*top*) the protein is first dissolved in alkali, along with other substances, including certain undesired carbohydrates. The protein is then coagulated in acid, washed in water to remove most of the impurities, dispersed again in alkali and finally spray-dried. At each stage suspended wastes are removed by centrifugation. The membrane process also begins with an alkali extraction, but unwanted substances are removed by ultrafiltration. The permeate can then be treated by means of reverse osmosis to concentrate the residue, which consists mainly of carbohydrates, and to recover pure water. The membrane process requires less energy and leaves a smaller volume of wastes.

at a concentration of only .1 to .2 gram (dry weight) per liter, and they must be concentrated to at least 10 grams per liter before being fed into a bacterial digester where the fuels would be generated. Algae can be concentrated by chemical coagulation or by centrifugation, but the cost of these methods appears to be prohibitive for a crop that essentially is to be burned and therefore must be very cheap. Ultrafiltration offers a promising alternative, and since the cells are rather large on the scale of the pore diameters a loose membrane with a high flux can be employed. In experiments conducted by Joaquin Errazuriz and one of us (Harry P. Gregor) the concentration of the alga *Selenestrum* was raised from .2 gram per liter to 17 grams per liter by ultrafiltration. The cells are not broken, and the permeate liquid is entirely clear. It has been estimated that in large-scale aquaculture or mariculture the cost of ultrafiltration would be less than three cents per 1,000 gallons, well within the bounds of economic feasibility.

The supply of nutrients for the algae is a more difficult problem. John H. Ryther and Joel Goldman of the Woods Hole Oceanographic Institution have estimated that if the entire output of the U.S. fertilizer industry were dedicated to such fuel crops it would supply only 1 to 3 percent of the need. The requirement can be reduced only by conserving the nutrients, a process that could draw on several membrane technologies.

A prototype algal system, constructed by Daniel Omstead, Thomas Jeffries and one of us (Harry P. Gregor) is now in operation. Material in the bacterial digester is ultrafiltered to remove by-products, which has the benefit of substantially accelerating digestion. The digestion is designed to yield methane, and the by-products are ammonium bicarbonate, the salts of organic acids and common inorganic salts. The permeate containing these substances is passed through a water-splitting membrane cell, yielding ammonia and carbon dioxide. The carbon dioxide is dissolved in a basic stream and is returned to the digester to enhance methane production; the ammonia can be recycled or it can be sold as fertilizer. The organic acids, which have commercial value, are removed by solvent extraction. Finally, reverse osmosis and electrodialysis are employed to purify the remaining nutrients for recycling, to concentrate wastes for disposal and to recover pure water. Even in this elaborate system of closed cycles, of course, some nutrients are lost. One possible way of making up the deficit would be to employ sewage sludge as a fertilizer. When the sludge has been concentrated by ultrafiltration, it could be transported economically.

It is worth noting that all the known membrane technologies have models in the living cell. Indeed, a single organ, the kidney, carries on activities comparable to all six of the membrane technologies reviewed here: ultrafiltration, reverse osmosis, electrodialysis, water splitting, solvent extraction and the promotion of chemical reactions by membrane-bound enzymes. The synthetic membranes do not yet approach the biological ones in speed, compactness, selectivity or efficiency. The principles of operation of the synthetic membranes are now well understood, however, and their performance is ample for commercial exploitation.



CULTIVATION OF ALGAE FOR FUEL is a proposed technology that could employ several membrane processes. The algae could be harvested by an ultrafiltration membrane with comparatively large pores, the concentrated suspension of algal cells being fed into a bacterial digester where the fuel (such as methane) would actually be generated. By-products of the digestion would be removed by another ultrafiltration membrane and would be broken down by a water-splitting cell. Ammonia (NH_3) and carbon dioxide (CO_2) represent valuable nutrients that could be returned to the algal pond or tank. Organic acids, which have commercial value, could be removed by means of membrane solvent extraction. Inorganic acids and bases could be concentrated as salts by electrodialysis, with pure water being returned to the system.

Innovation. Imitation.

Olympus innovated the concept of the compact SLR. The introduction of the OM-1 startled the world of photography by putting so many big ideas into such a small body.

And after Olympus did the thinking, others did the following.

Today, the OM cameras still stand ahead of the pack with exclusive features.

OM-1: The #1 Innovation.

Enter the OM-1. Suddenly, the SLR camera is 33% smaller and lighter, yet incredibly rugged to meet the demands of professional wear and tear. Miraculously, the viewfinder is 70% brighter and 30% larger for faster, easier composing and focusing.

And suddenly the OM-1 became the #1 selling compact SLR. Its metering system is designed to give complete control to professionals and photojournalists. No distractions, blinking lights, or obscured images in the viewfinder.

A Quiet Innovation.

Olympus created a unique shock absorber/air damper system to eliminate noise and vibration, for sharper, unobtrusive photography. Especially vital for long tele shots and macro/micro photos.

The Motor Drive Innovation.

OM-1 is still unsurpassed in its continuous-view motor drive capability: 5 pictures per second. And a rapid winder that fires as fast as 3 shots a second! With no mirror "lock-up," regardless of lens used.

The Biggest System Innovation.

OM-1 is part of the world's most complete compact system. More than 280 components, all compact design, including 13 interchangeable screens so you can meet any photographic challenge. Ingeniously designed to change in seconds through the lens mount. And more compact lenses than any other system, each a marvel of optical design and performance.

The Automatic Innovation: The OM-2.

It's the fully automatic OM, with major differences from all other automatics! The only SLR with "off-the-film" light measurement for those photographers demanding the ultimate innovation in automatic exposure control. Which means each frame is individually exposure-controlled even in motor drive or rapid winder sequences. And it makes possible the unique Olympus 310 Flash whose flash duration is controlled by the camera's metering system.

And of course, the OM-2 shares every other innovation and system component with the OM-1.

We Wrote The Book On Compacts.

The OM System story is detailed in our full color in-depth brochure, yours free for writing OLYMPUS, Woodbury, New York 11797. Read it all. Discuss the advantages of an Olympus with your photographer friends.

Visit your camera store. Compare. Wouldn't you rather buy the innovator instead of the imitator?

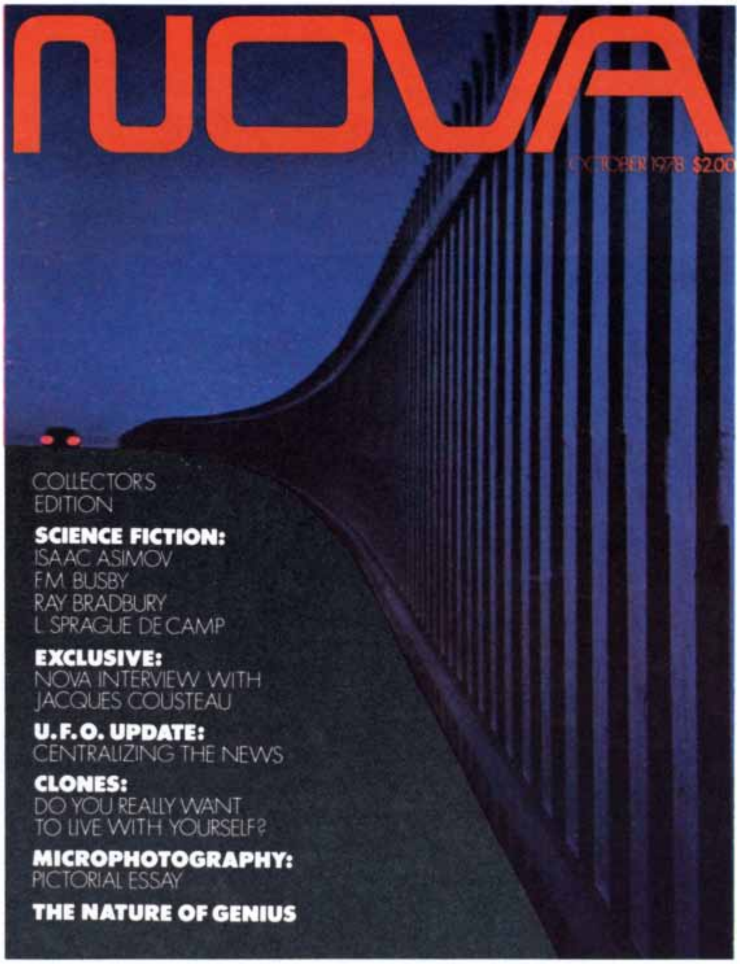
OLYMPUS OM SYSTEM

ANNOUNCING

NOVA

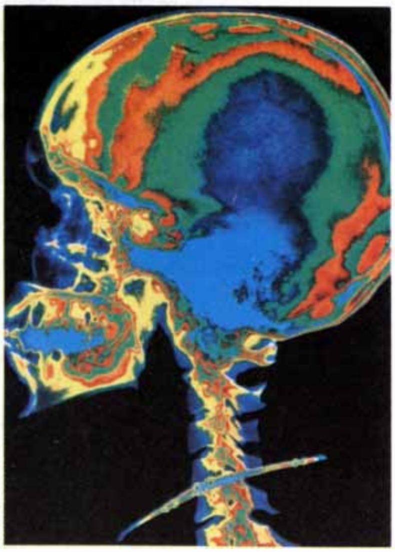
THE MOST SPECTACULARLY BEAUTIFUL NEW STAR
IN THE COMMUNICATIONS GALAXY

THE SCIENCE FICTION OF YOUR CHILDHOOD
IS THE SCIENCE FACT OF TODAY.



YOU ARE INVITED TO RESERVE THE PREMIER ISSUE NOW
UNDER OUR SPECIAL INTRODUCTORY CHARTER OFFER!

Ernest Fuchs



AGFA-GEVAERT

Michael Freeman

The Kiwi

New Zealand, where this flightless bird lives today, had no mammals for 80 million years. In filling ecological niches that would have been occupied by the mammals the bird evolved mammalian characteristics

by William A. Calder III

The animal kingdom is so rich in its diversity that evolution can appear to have been a random process. Consider the famous flightless bird of New Zealand, the kiwi. Natural selection is partly a matter of chance, affected by random mutations and historical accidents of parentage and location. At the same time it operates under certain constraints. For example, physical laws determine how large a bird can be and still fly. In their short life span human beings are not likely to witness the interaction between environment and heredity that gives rise to an organism adapted to an available environmental niche. The natural history of biological isolation, however, presents some suggestive case histories. The kiwi's is one of the most fascinating.

One often hears that birds exhibit marvelous evolutionary adaptations. The examples given may be the arctic tern's long migration, with its remarkable feats of navigation; the wingspan of those master soarers, the great albatrosses; the coevolution of the tiny hummingbirds and plants that bear red tubular flowers, and the elaborate mating displays of the bower birds. Can the same possibly be said for a bird that has lost its ability to fly, that probes for worms, that burrows in the dirt to lay a ridiculously oversized egg and that is known to nonbiologists mainly as the label of a brand of shoe polish? My answer, given the context of isolation and ancestry, is a clear yes.

There are three species of kiwi. The common or brown kiwi (*Apteryx australis*) is found on New Zealand's North Island and South Island and also on Stewart Island, off the southernmost tip of South Island. The little spotted kiwi (*A. oweni*) was until recently present on both North Island and South Island but now is found only on South Island. The great spotted kiwi (*A. haasti*) is also found only on South Island. These birds are known nowhere else in the world.

Kiwis are the smallest members of the seven families of ratite birds: nonflying birds that lack the keeled breastbone to which are attached the massive pectoral muscles that flap the wings of flying birds. ("Ratite" is from the Latin *ratis,* raft, an unkeeled vessel.) Among the other ratites are the ostrich, the emu and the extinct giant moa of New Zealand. In addition to lacking a keeled breastbone the kiwi shares with the other ratites an adult plumage that is the same in structure as the juvenile down of other birds. The kiwi's abundant feathers are long and flexible; they lack the usual interlocking mechanism, giving the plumage a shaggy, furlike appearance. The feathers do not ruffle as the kiwi digs the burrows where it spends the daylight hours, and they provide insulation both in the cool, damp underground habitat and in environments that range from semitropical in the far north of North Island to virtually subpolar in the Southern Alps of South Island.

The kiwi lacks an external tail, and its wings are reduced to vestiges that are hidden under its plumage. Its sense of touch is augmented by well-developed rictal bristles at the base of the bill and a sensitive region at the tip. Its legs are very strong; they serve the bird not only for locomotion but also for burrowing and fighting. The kiwi's posture when it is digging resembles the attitude of a swimming duck.

The kiwi feeds on insects, berries, seeds and other plant materials, but its principal food is the earthworm. New Zealand has a particularly rich earthworm fauna, and the kiwi has 178 native and 14 introduced species of worms at its disposal. The bird's foraging technique consists of probing the soil with its long bill and smelling out the worm. Betsy G. Bang of Johns Hopkins University finds that relative to the size of the forebrain the olfactory bulb of the kiwi is the second-largest among all the birds she has studied. In the kiwi the ratio of the diameter of the olfactory bulb to the diameter of the cerebral hemisphere is 34 percent. (The snowy petrel shows the highest ratio: 37 percent.) Such evidence with regard to function is only circumstantial, but it finds support in experiments conducted by Bernice M. Wenzel of the University of California at Los Angeles. She gave kiwis access to mesh-covered pots filled with earth, at the bottom of which were vessels containing either food or plain earth. If the vessel contained only earth, the kiwis ignored the pot and did not attempt to probe it.

The biology of the kiwi is best appreciated in a context of seemingly diverse matters: the natural environment of New Zealand, the overall evolution of ratite birds and the function of bird eggs and their shells. The topography of New Zealand is spectacularly steep, raised high by faulting and volcanism and sculptured by glacial ice. Rainfall is generally abundant, and before the European colonists began to alter the landscape it was predominantly forest. At the lower elevations and latitudes the forests were a mixture of broadleaf trees and podocarp conifers, a group of trees peculiar to the southwest Pacific and Chile. Above this belt of mixed forest and on up to the timber line, at 900 to 1,500 meters, the forest was predominantly southern beech. Patches of natural grassland were also present, mainly on the drier leeward side of the major mountain ranges such as the area east of the Southern Alps.

Except for two species of bat there were no mammals in New Zealand until man arrived in the first millennium A.D. Hence the evolution of the two ratite families—the kiwis and the moas—could proceed free from the predation and competition that were the rule where terrestrial mammals flourished. It was here that the ratites had their greatest success in diversification. W. R. B. Oliver of the Canterbury Museum in Christchurch, a student of the recently extinct moas, assigned their semifossilized remains to 22 different species. From a recent reanalysis Joel Cracraft of the University of Illinois Medical Center concludes that the fossils represent only 13 species. Nevertheless, combined with the three living kiwi species this is a total of 16 ratite species that are distributed over an area only the size of Colorado, although spanning a range of latitude equal to that between Daytona Beach, Fla., and Goose Bay, Newfoundland.

Some of the moas that evolved in island isolation were only a half meter tall at maturity; others grew to a height of three meters. Dean Amadon of the American Museum of Natural History estimates that the two races of the largest species, *Dinornis giganteus,* weighed between 230 and 240 kilograms, more than twice the weight of an ostrich. In contrast an adult brown kiwi stands no taller than a domestic chicken and weighs on the average 2.2 kilograms.

The first mammals (other than bats) to invade New Zealand were men: a group of seafaring Polynesians, known today only as the moa hunters, who reached the islands in the ninth century A.D. Whether they brought along dogs and rats is uncertain. If they did not, later Polynesian arrivals did.

As the largest land animals in New Zealand until the arrival of man, the moas were obvious prey for the invaders. Although some moa species may have disappeared even before the first human beings arrived, those that survived were probably wiped out before the time of the first European contact: Captain Cook's visit in 1769. The main mammalian invasion began early in the 19th century, when European colonists introduced 22 different kinds of mammals, including pigs, sheep and cattle. Many exotic immigrants were intentionally introduced as game animals or escaped from domestication, and several of them soon began to wreak ecological havoc.

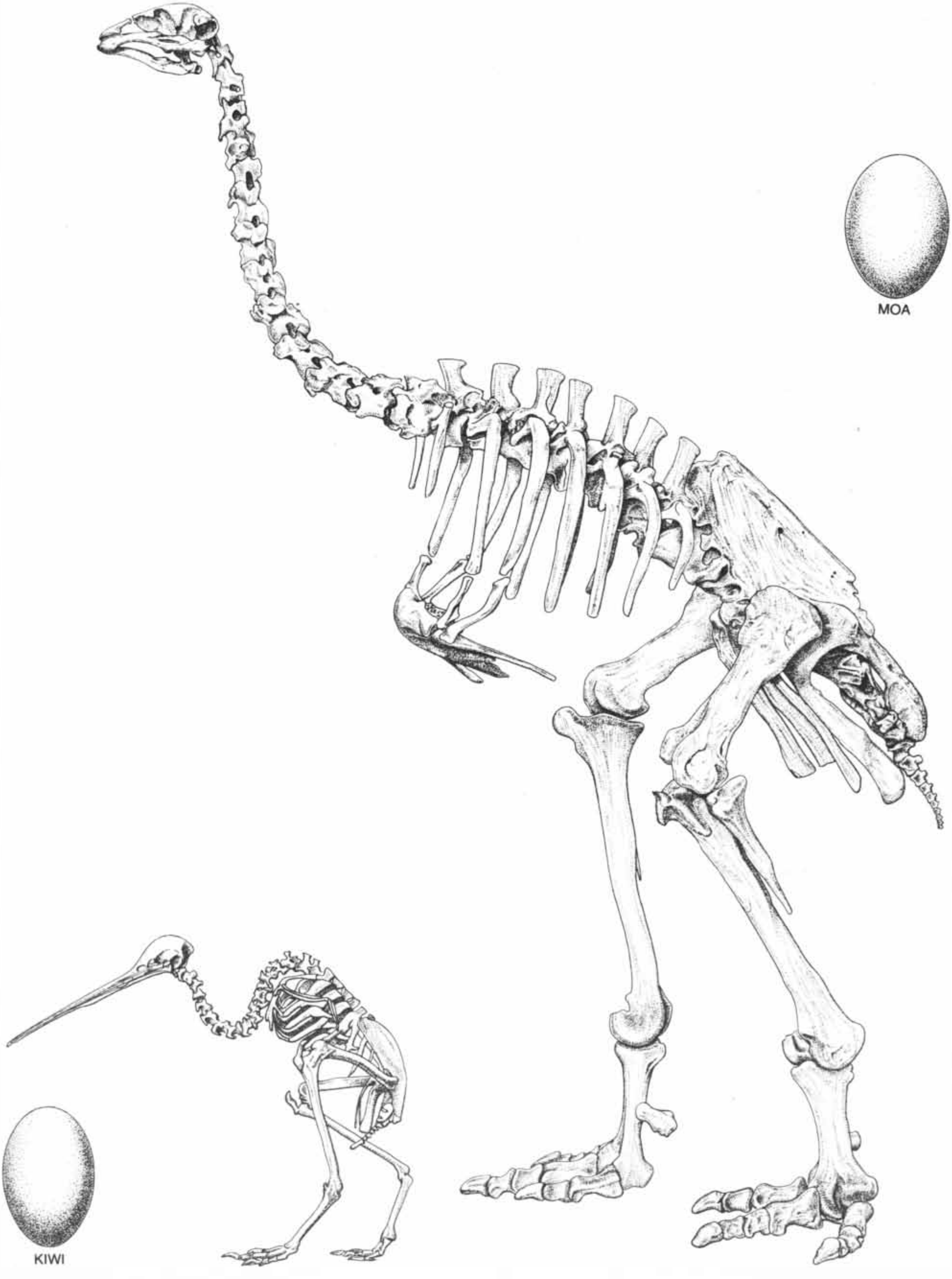
From the time of their arrival the Europeans have put pressure on the islands' long-isolated flora and fauna in a number of ways. The forest habitat of the kiwi shrank under a dual assault: the logging of native trees and the clearing of land for sheep pasture. In recent years much forest land has been cleared and replanted with the imported Monterey pine, because these trees grow faster than the native species. Traps and poisons have been deployed in an attempt to control the imported mammals that became pests. From what is now known about the persistence of poisons in food chains elsewhere in the world it seems likely that the poisons have adversely affected native animals, the kiwi included. Indeed, F. C. Kinsky of the Dominion Museum in Wellington has reported instances of kiwis killed by poisoning. The kiwi is also threatened by trapping, automobiles, predation by dogs and land clearance by burning.

The kiwis are now protected, but no one is sure why the little spotted kiwi is disappearing. Brian Reid and Gordon Williams of the New Zealand Wildlife Service believe that the bird had vanished from North Island before the European settlers arrived. It does not seem likely that it was hunted to extinction by the Polynesians, because the brown kiwi, which is equally vulnerable to hunting, is still present on North Island. To make the puzzle deeper, the process appears to be a continuing one. Reid reports that even though the little spotted kiwi has not been eliminated on South Island, it is rarely seen there.

The closely related moas and kiwis have more distant kin all around the world. In addition to the ostrich of Africa and the emu of Australia, the ratites include the cassowaries of Australia and New Guinea, the rheas of South America and the extinct elephant birds of Malagasy and the Canary Islands. How did these flightless birds evolve?

The anatomy of the ratites makes it clear that their ancestors were flying birds. For example, the ratites' wrist bones are fused for strength at the point where primary flight feathers would be attached, even though their wings are reduced in size and function. Flight entails a high oxygen demand, strong mechanical support and precise coordination, and the physiology of the ratites still shows that their ancestors met these requirements. The birds have retained the air-sac system for effective lung ventilation, a fusion of the lower vertebrae

THE COMMON OR BROWN KIWI (*Apteryx australis*) is the principal species of three still present in New Zealand. It has a shaggy appearance because its plumage resembles the juvenile down of other birds. It feeds largely by probing for worms with its long bill.



SKELETONS OF THE KIWI AND THE MOA, a giant extinct relative of the kiwi, are seen in profile. The specimens are respectively a foot and three feet high at the shoulder. The moa species is actually one of the smaller ones. Members of the flightless ratite, or unkeeled, suborder of birds, the kiwi and the moa are similar in anatomy, with a flat breastbone, vestigial wings and strong legs. Their eggs are about the same length; the moa egg is slightly greater in diameter and thus would have weighed about 35 percent more than the kiwi egg.

with the pelvis and a shortening of the tailbone for strength, and a large cerebellum for coordination.

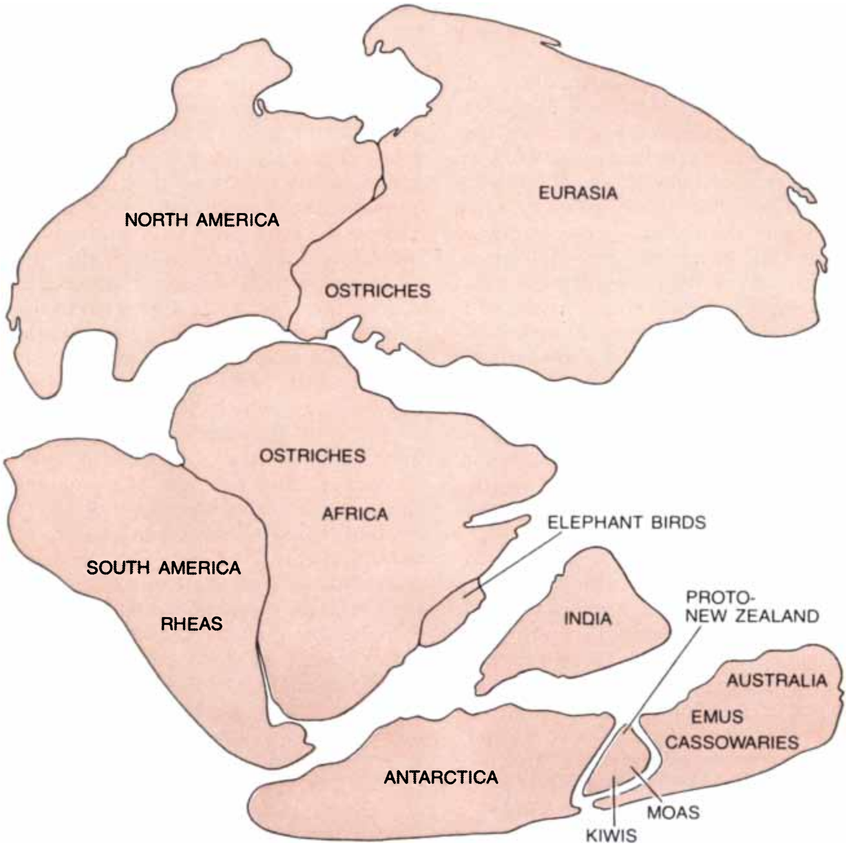
All birds are thought to have had a common pigeon-sized ancestor that arose some time in the Jurassic, 150 million years ago, and so it is reasonable to wonder just what factors favored the natural selection of birds as large as the ratites. The advantages of a large body include increased running speed, a larger home range and the potential for dominance over other animals. Another advantage is an improved endurance at times when food is in short supply; this is because the amount of food energy an animal can store increases in linear proportion to its mass whereas the metabolic rate at which the stored energy is burned does not increase.

At the same time a larger body means a greater total food requirement, which is potentially disadvantageous. Perhaps an even more serious disadvantage is that a large bird can weigh too much to fly. C. J. Pennycuick of University College in Nairobi has pointed out that the power required for a bird to become airborne increases out of proportion to the power available from the pectoral muscles. A bird that weighs more than about 12 kilograms is grounded. Thus the giant birds have made an evolutionary trade-off.

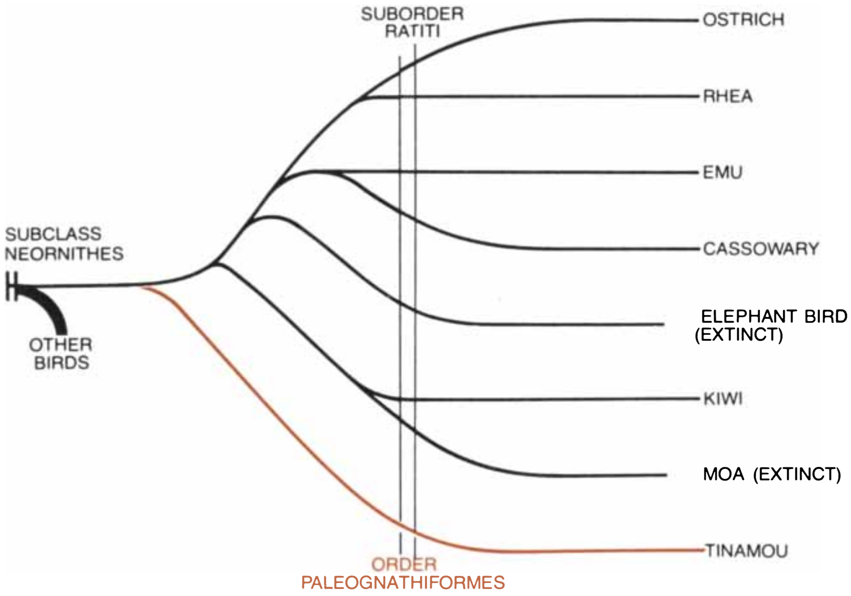
Another evolutionary puzzle was a good deal more tangled until recently. This is the problem of how the flightless ratites became so widespread geographically. The ratites were once regarded as primitive birds both because they are flightless and because they are separated from one another by oceans they could not cross. It seemed necessary in those days to construct hypotheses that explained the independent evolution of each member of the group. The biologist T. H. Huxley was almost alone in opposing such hypotheses of polyphyletic origin; he based his opposition on studies of the birds' bony palate.

Today anatomical studies by Cracraft and by Walter J. Bock of Columbia University (relating respectively to the birds' postcranial skeleton and their skull) indicate that the ratites belong among the advanced birds rather than the primitive ones. Chief among their distinguishing features is a palate similar in morphology to that of the tinamou, a partridgelike bird native to South Africa. It may therefore be assumed that the tinamous and the ratites share a common ancestry and that the ratites' departure from the family tree of the flying birds did not precede the rise of the tinamous.

Certain properties of the proteins in the egg white of ratite birds and their immunological relations have been analyzed by David Osuga, Robert E. Feeney, Ellen Prager and Allan C. Wilson at the University of California at Davis and the University of California at



PRESENT DISTRIBUTION OF THE RATITE BIRDS can be explained as the result of a common ancestor having lost the power of flight at a time before continental drift divided the ancient continent of Gondwanaland into the southern continents of today. According to Joel Cracraft of the University of Illinois Medical Center, the kiwis and the moas are anatomically more primitive than the other ratites. Their isolation can be attributed to the separation of New Zealand, as is shown here, from western Antarctica 80 million years ago. The elephant birds of Malagasy arose from the ancestral stock left behind, as did cassowaries and emus of Australia and New Guinea, ostriches of Africa and Eurasia and the rheas of South America.



FAMILY TREE OF THE RATITES, based on anatomical studies conducted by Cracraft, places the seven families on a common stem with the tinamous: South American birds resembling partridges that have retained the power of flight (*color*). Cracraft considers ostriches and rheas to be the most specialized ratites and closely related. Immunological studies suggest, however, that rheas are more closely related to the kiwis than they appear to be in this diagram.

Berkeley. Their findings make it possible to construct a ratite family tree very much like one that has been proposed by Cracraft. The only significant difference between the two has to do with the fact that Cracraft relates the ostriches closely to the rheas on the basis of anatomical characteristics. Others have pointed out that the two birds share lice of the same genus and that whereas the ratites of the Pacific have only vestigial wings, the ostrich and the rhea still have wings large enough to serve in courtship display. The immunological studies, however, suggest that the rheas are actually closer to the kiwi than to the ostrich. These findings, regardless of fine points still to be settled, seem to have scuttled the hypotheses of polyphyletic origin, assuming, of course, that the curious geographical distribution of the ratites can be otherwise explained.

Cracraft has noted that a simple explanation is in fact provided by continental drift. The pattern corresponds neatly with the breakup of the great southern continent of Gondwanaland beginning late in the Jurassic period. By the end of the Cretaceous, about 90 million years ago, Africa was completely severed from South America, thus providing the possibility of an early isolation of the rheas and the ostriches that fits the immunological distance between the two. New Zealand began to drift away from western Antarctica about 80 million years ago, leading to the isolation of whatever ancestral ratite gave rise to the moas and the kiwis.

When did the extremes in size among the ratites of New Zealand arise? Is the kiwi perhaps a shrunken moa? The only known moa remains are relatively recent, so that nothing is known of the original ratite stock. If one accepts the monophyletic hypothesis of ratite origin and the likelihood that the birds' ability to fly was lost in the course of their increase in size, one is led to the conclusion that the kiwis are considerably smaller than their ancestors.

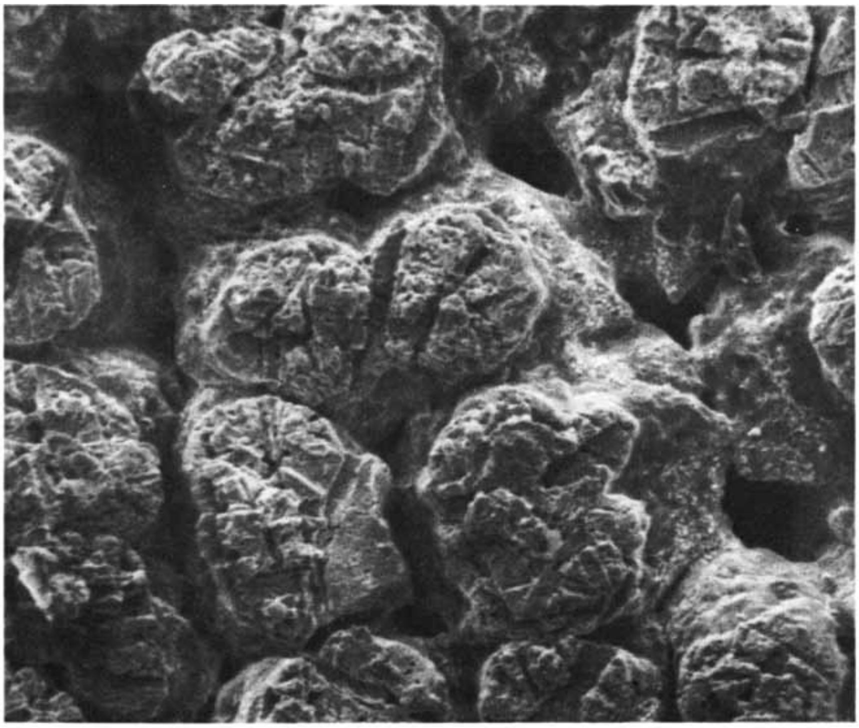
Such matters come under the heading of allometry, which Stephen J. Gould of Harvard University has neatly defined as "the study of size and its consequences." For example, if an organism's size is scaled upward, not all of its anatomical features can change in the same ratio if the same function is to be maintained. A skeleton or an eggshell must be disproportionately heavier in the larger organism in order to provide a margin of safety under the stress of greater mass. On the other hand, if the mass of an organism is doubled, its metabolic rate is not doubled but is only increased by 68 percent, and frequencies such as heart rate or breathing rate actually decrease by about 20 percent.

Allometric analysis is gratifyingly simple. This is because within any particular class of organisms not only structural dimensions but also the rates and frequencies of organic processes are exponential functions of body weight. When the functions are plotted on logarithmic scales, they usually show straight-line trends.

The first allometric study of the relation between the mass of bird eggs and the mass of the birds that lay them was published by T. H. Huxley's grandson Julian. For example, a hummingbird weighing 3.5 grams lays an egg weighing .5 gram, and an ostrich weighing 100 kilograms lays an egg weighing between 1.4 and 1.7 kilograms. In other words, although the ostrich egg is far heavier than the hummingbird egg, it is very light in proportion to the weight of the ostrich. Interpolated allometrically, a kiwi the size of a chicken should lay an egg the size of a chicken's or perhaps slightly larger, weighing between 55 and 100 grams. In actuality the egg of the brown kiwi usually weighs between 400 and 435 grams. What should a bird weigh in order to lay a 400-gram egg? The allometric equation answers 12.7 kilograms, about six times the weight of a kiwi hen.

The leading American workers in allometric analysis of bird eggs are Hermann Rahn and his colleagues at the State University of New York at Buffalo: Amos Ar, Charles V. Paganelli, Robert B. Reeves, Douglas Wangensteen and Donald Wilson. Their work has yielded equations that predict from the weight of an egg an amazing range of related properties: the thickness of the shell, the permeability of the shell to gases and water vapor, the probable incubation period, the maximum rate of oxygen uptake and the total water loss during incubation. For example, the diffusion of water vapor from the albumen, or egg white, through the porous shell is measurable as a weight loss over the period of incubation. The rate of loss can characterize the water-vapor conductance (permeability) in the same way that electrical conductivity or resistance is determined when the current and voltage of an electric circuit are known. Oxygen diffuses inward through the same pores, so that the maximum oxygen uptake is directly proportional to the water-vapor conductance. Therefore once one knows the rate of water loss, the metabolism of the embryo can be predicted up to the moment of hatching.

The development of the avian egg begins when the ovum, consisting of a germ cell and a ration of nutrient yolk, is formed in the ovary. The associated albumen, membranes and eggshell are added after ovulation, as the ovum passes down the oviduct. The proportion of yolk to the total mass of the egg is directly related to the extent of embryonic development before hatching. Altricial birds, those that emerge from the egg blind and featherless, come from eggs that are perhaps 20 percent yolk. Preco-



INTERIOR OF A KIWI EGGSHELL is revealed in this scanning electron micrograph, which shows the structure at a magnification of 325 diameters. The crystals of calcite that are secreted by the eggshell gland form a series of roughly circular deposits: the mammillary cones. Continued secretion causes each cone to grow; gaps that remain between cones are respiratory pores of the eggshell. Micrograph was made by Cynthia Carey of the University of Colorado.

Discover

the <u>most</u> organized (and inexpensive) way to keep up with what's new in computer techniques and management—

The Library of Computer and Information Sciences

Take any 3 books for only $3.95

(values to $60.00)

if you will join now for a trial period and accept only 3 more books—at handsome discounts—over the next 12 months

(Publishers' Prices shown)

62811. MODELING AND ANALYSIS USING Q-GERT NETWORKS. *A. Alan B. Pritsker.* Shows how to write program inserts, retrieve sub-programs, edit and link similar subnetworks, and much more. **$17.50**

78350. SIMULATION WITH GPSS AND GPSS V. *Bobillier, Kahan, and Probst.* A detailed approach to the effective planning and execution of simulation experiments and projects. With concrete examples. **$18.50**

82560. TECHNIQUES OF PROGRAM STRUCTURE AND DESIGN. *Edward Yourdon.* Over 300 illustrative problems, 90 valuable references, and four demonstration projects. **$17.50**

54450. THE IBM 5100 PORTABLE COMPUTER: A Compreh nsive Guide for Users and Programmers. *Harry Katzan, Jr.* Includes BASIC and APL programs for payroll/inventory control applications. **$19.95**

70115. PRINCIPLES OF COMPILER DESIGN. *Alfred V. Aho and Jeffrey D. Ullman.* Step-by-step approach to developing the "software tools" necessary to design the major component of a compiler. **$18.95**

80845. STRUCTURED SYSTEMS DEVELOPMENT. *Kenneth T. Orr.* Basic steps for analyzing, developing, and installing systems with the emphasis on a structured "problem-solving" approach to systems building. **$17.00**

70110. PRINCIPLES OF DATA-BASE MANAGEMENT. *James Martin.* Covers data structures, security, file storage, real-time vs. non-real time, more. **$19.95**

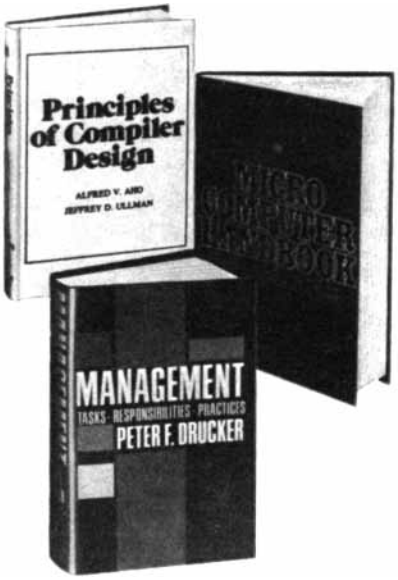
32485. ALGORITHMS + DATA STRUCTURES = PROGRAMS. *Niklaus Wirth.* A classic work on structured programming with practical examples. **$19.00**

39980. COMPUTER-COMMUNICATION NETWORK DESIGN AND ANALYSIS. *Mischa Schwartz.* An organized, quantitative approach to the design of communications networks. **$18.95**

37580. CHESS AND COMPUTERS. *David Levy.* With illustrative games that show how to apply computer skills to chess problems. *Softbound.* **$7.95**

39970. COMPUTER ARCHITECTURE. Second Edition. *Caxton C. Foster.* Covers microprogramming, switching circuitry, memory speed increases in mainframes. **$16.95**

42010. THE DESIGN AND ANALYSIS OF COMPUTER ALGORITHMS. *Aho, Hopcroft, and Ullman.* Practical all-inclusive survey of principles and techniques. Covers such topics as use of lists, queues, trees, and graphs; sorting. **$18.95**



62651-2. MICROPROCESSORS AND MICROCOMPUTER SYSTEMS. *G.V. Rao.* Analysis of hardware and software developments, built-in and associated LSI solid state memories, and incidental external bulk-memory input/output devices. *Counts as 2 of your 3 books.* **$24.50**

42217. A DICTIONARY OF MICROCOMPUTING. *Philip E. Burton.* Over 900 definitions, complete with supporting examples, sketches, diagrams, and tables. **$12.50**

83400. THEORY AND DESIGN OF SWITCHING CIRCUITS. *Friedman and Menon.* Practical solutions to optimization problems for large-scale systems, plus flip-flop memory elements, mathematical models. **$19.95**

84385. TOP-DOWN STRUCTURED PROGRAMMING TECHNIQUES. *Clement McGowan and John R. Kelly.* Practical solutions in actual code serve to introduce this revolutionary concept. **$14.95**

62620. MICROCOMPUTER HANDBOOK. *Charles J. Sippl.* Detailed references to the technology and applications of microprocessors. **$19.95**

41670-2. DATA BASE DESIGN. *Gio Wiederhold.* Excellent sourcebook for the methods, principles, and concepts relevant to the design of data base software. *Counts as 2 of your 3 books.* **$21.50**

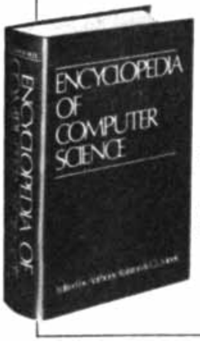
49355. FUNDAMENTALS OF DATA STRUCTURE. *Ellis Horowitz and Sartaj Sahni.* Includes in-depth coverage of file organizations, indexing techniques, internal and external sorting and tree structures. With detailed illustrations. **$17.95**

58805. LOGICAL DESIGN OF DIGITAL SYSTEMS. *Arthur D. Friedman.* Basic reference to the components of the digital system: gate level logical design, combinational circuits, sequential circuits, register transfer languages, more. **$16.95**

79165. SOFTWARE METRICS. *Tom Gilb.* Practical techniques for estimating program reliability and efficiency. Details the experiences of leaders in the field like TRW-Systems and Lockheed. **$15.95**

59920. MANAGEMENT: Tasks, Responsibilities, Practices. *Peter F. Drucker.* A brilliant task-focused handbook of the skills, techniques and tools a good executive needs. **$17.50**

EXTRAORDINARY VALUE!



44900-3. ENCYCLOPEDIA OF COMPUTER SCIENCE. *Edited by Anthony Ralston and Chester L. Meek.* More than 1550 outsize pages of information on every aspect of computer science—from algebra to automata theory, from basic terminology to string-processing languages. 470 articles by 208 experts. Over 700 charts, tables, graphs, and diagrams. *Counts as 3 of your 3 books.* **$60.00**

If the reply card has been removed, please write to
The Library of Computer and Information Sciences
Dept. 7-A2F, Riverside, N.J. 08075
to obtain membership information and an application.

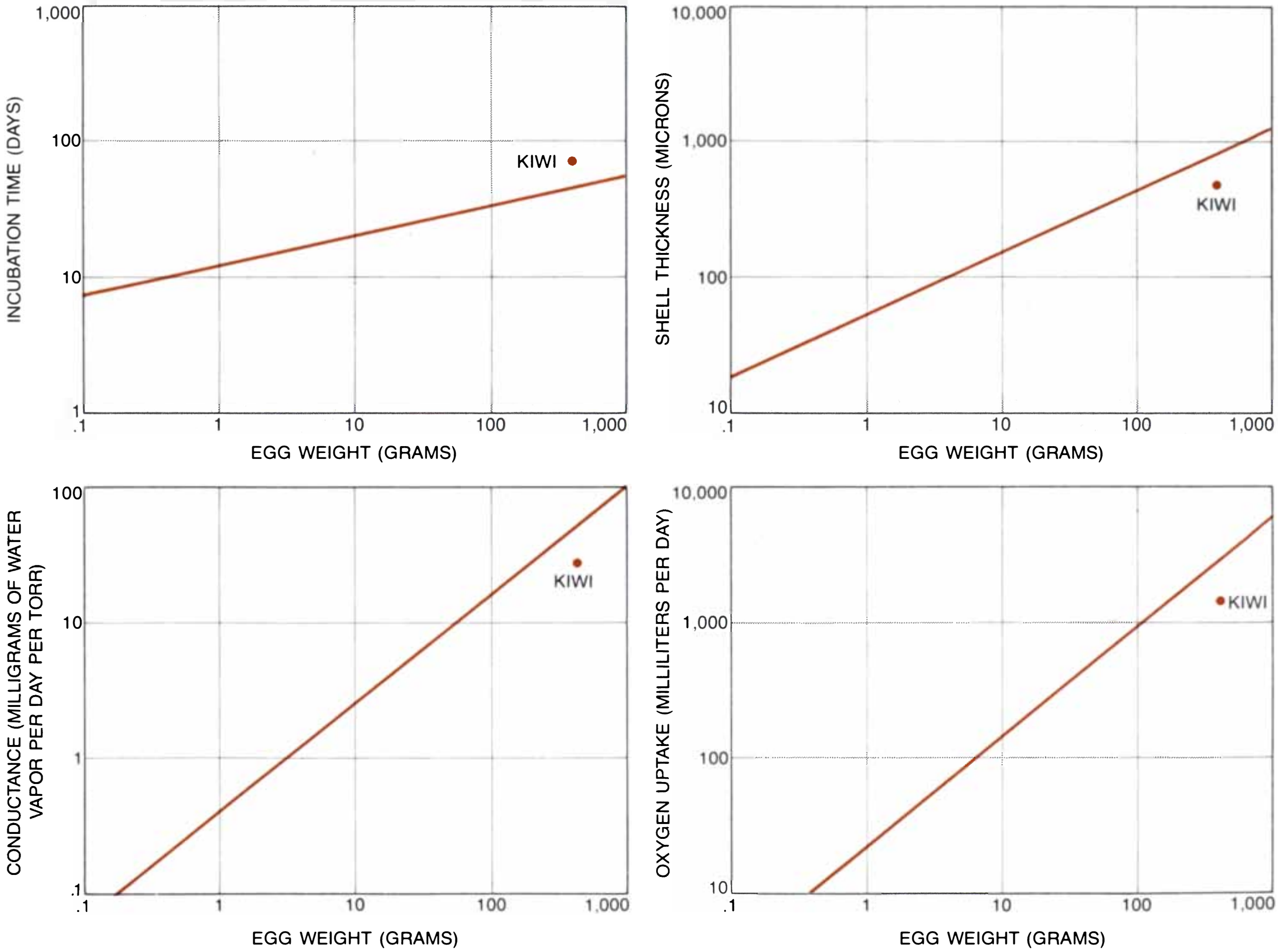
cial birds, which hatch already covered with down and able to feed themselves, come from eggs that are about 35 percent yolk. Given the large size of the kiwi egg and its prolonged incubation (71 to 75 days instead of the 44 days that would be normal for a 400-gram egg) it is no great surprise that the young kiwi hatches fully feathered and receives no food from its parents. What is more noteworthy is the extent of the preparations that lie behind the kiwi's precocity. Reid has found that the contents of the kiwi egg are 61 percent yolk, a proportion half again as large as that found in the eggs of typical precocial birds.

Once the shell of the egg has been deposited no more food or water can be added to the egg. Thus by that time the contents of the egg must include enough yolk to nourish the embryo. They must also include enough water, and the structural integrity of the eggshell must be such that the water is adequately conserved. Rudolph Drent, who is now at the University of Groningen in the Netherlands, has measured the daily rate of water loss from large and small eggs in the course of natural incubation. From Drent's findings, taking into account the relation between incubation time and egg size, Rahn and Ar have established that the percentage of water loss is much the same for all eggs: 18 percent for the smallest eggs and 13 percent for the largest.

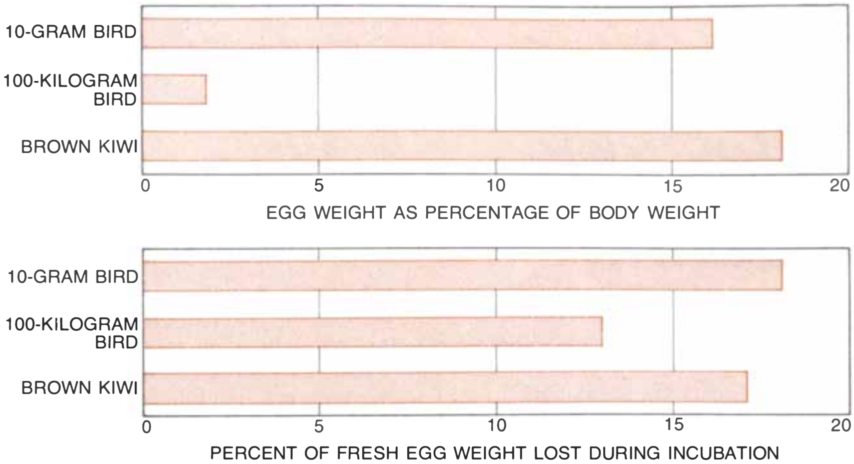
It is exciting to see such generalities emerge. Correlations as reliable as these strongly suggest that natural selection operates within quite strict physical limitations that among birds impose a certain order on egg size, incubation time and such structural details as the porosity of the shell. If the evolution of the avian egg has indeed proceeded within, so to speak, close engineering tolerances, what about the kiwi? Are the rules of evolutionary scaling actually understood? Can the rules be real if some birds are exempt from them? It was with just such questions in mind that I recently took a sabbatical leave to visit New Zealand and study kiwis.

One of the questions turning over in my mind during the 14-hour flight from Los Angeles to Auckland concerned the kiwi's extremely long incubation period. Could the extended incubation be the result of poor physical contact for heat transfer between a relatively small bird and a relatively large egg? I looked forward to monitoring egg temperatures.

I had settled down to my first meal in New Zealand, tea at a restaurant overlooking Lake Taupo, when my attention was drawn to a television report in progress. Barry Rowe of the Otorohanga Zoological Society was being interviewed about his work with kiwis. I stared in disbelief as Rowe opened the door of an artificial kiwi burrow and removed an egg that he opened into two halves, revealing a small transmitter that broadcast a continuous record of the egg's temperature. After I had recovered from the shock of learning that another investigator had beaten me to the draw I telephoned Rowe. He invited me to share the facilities and hospitality of the zoological society. I promptly joined Rowe in his work and was soon busy



RELATIONS AMONG AVIAN EGGS are evident in the straight-line plots that appear in these logarithmically scaled graphs plotting increases in egg weight (the horizontal scale) against four factors determined by Hermann Rahn and his co-workers at the State University of New York at Buffalo. These are (*top left*) increased incubation time, (*top right*) increased thickness of the eggshell, (*bottom left*) increased rate of evaporation from the egg and (*bottom right*) increased oxygen uptake by the egg. In each instance the kiwi egg fails to conform. Its incubation time is some 60 percent longer than the norm for an egg of its size, its shell is nearly 40 percent thinner and its rate of water-vapor diffusion and oxygen uptake are below the norm even though shorter pores of a thinner shell should accelerate diffusion.



WEIGHT OF THE AVIAN EGG is not scaled in linear proportion to the weight of the female. The bars in the top graph show egg weight as a percentage of parental weight for a hypothetical small bird weighing only 10 grams, for a hypothetical large bird weighing 100 kilograms and for a brown kiwi weighing 2.2 kilograms. The weight of the large bird's egg is less than 2 percent of the weight of the parent. The bars in the bottom graph show the weight lost during incubation by the same three eggs, expressed as a percentage of the weight of the fresh-laid egg. The small egg would require incubation for 13.4 days, the large egg for 62.3 days and the kiwi egg for 71 to 74 days. The percentage differences in weight loss are not statistically significant. The calculations follow formulas developed by Rahn and his co-workers.

measuring the temperatures of reinforced, water-filled eggs as they were incubated.

Rowe recorded the weight changes in a female kiwi during the interval between successive layings. The accumulation of material to be deposited in the energy-rich yolk occupies the first seven and a half days of the 24-day egg-laying cycle. In terms of energy investment the female during that week must exceed her basal metabolic rate by between 174 and 203 percent. The bird's feeding activity in the wild probably increases dramatically in the same period, but there is no quantitative information about kiwi behavior in the wild to confirm this assumption.

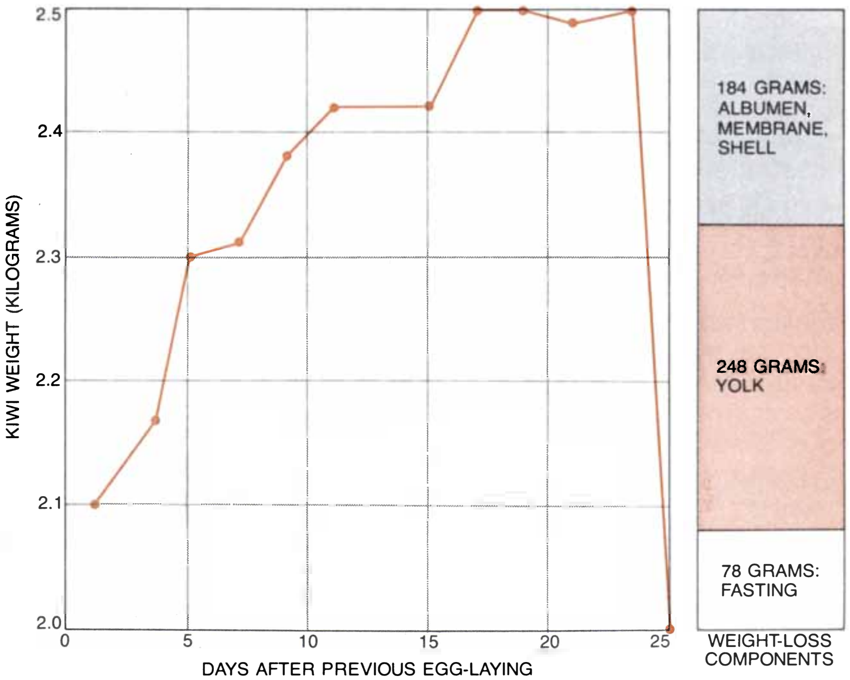
The female kiwi, having invested a great deal of energy in egg synthesis over a 24-day period, stops feeding in the last two days of the cycle. As a result when she leaves the burrow after laying her egg she has lost more weight than the weight of the egg alone. The energy investment in the chick-to-be is now completed. The female is free to feed and may have already entered a second cycle of egg synthesis.

The male now takes over the task of incubating the egg for nearly two and a half months. In captivity he leaves the burrow to feed once or twice a night. It has been suggested by some observers that in the wild the female also occasionally brings food to the male during this period; there are reports of males remaining in the burrow several days without feeding.

At embryo depth within the kiwi egg the temperature of incubation averages 35.4 degrees Celsius. When we placed an instrumented egg under a bantam hen, a bird smaller than a kiwi, the temperature of incubation rose to 37.7 degrees. The difference reflects the difference in the body temperature of the two birds. Donald S. Farner of the University of Washington and his colleagues, Norman Chivers and Thane Riney, measured the body temperatures of various kiwi species. They found that the average basal temperature of the adult brown kiwi is 38 degrees C. This is two degrees lower than the body temperature of other birds and more like the temperature of a mammal. In any event if the smaller bantam hen could maintain the kiwi egg at 37.7 degrees, the lengthy period of incubation could not simply be the result of the egg's being too large to be kept warm.

Knowing the natural kiwi incubation temperature, Rowe was able to achieve the first successful artificial incubation of a kiwi egg. Over the 71-day incubation period I measured the oxygen uptake of the egg at intervals of two to three days. Meanwhile, kiwi eggs being difficult to come by, we salvaged infertile eggs from the society's breeding program for other comparative measurements. The kiwi had proved to be a nonconformist not only in egg size but also in incubation temperature and duration. Did the kiwi egg conform in any way to the predictions of Rahn and his associates?

I determined the shells' water-vapor conductance, measured their thickness and separated their yolk and albumen. The separated components were taken to the Ruakura Agricultural Research Centre, where C. R. Parr, D. P. Karl and R. Whaanga determined the water content of both and prepared samples for measurement of the eggs' energy content by means of the bomb calorimeter.



SYNTHESIS OF THE KIWI EGG is accompanied by a 400-gram increase in the weight of the female. Some 200 grams are gained during the first seven days of the 24-day cycle. The female's food requirement must be highest during that week. Some 180 grams of albumen, membrane and eggshell are synthesized as the egg travels along the oviduct. Minerals required for eggshell are provided in the diet or are drawn from the bones. For the last two days of the cycle the female fasts. Barry Rowe of Otorohanga Zoological Society monitored the changes.

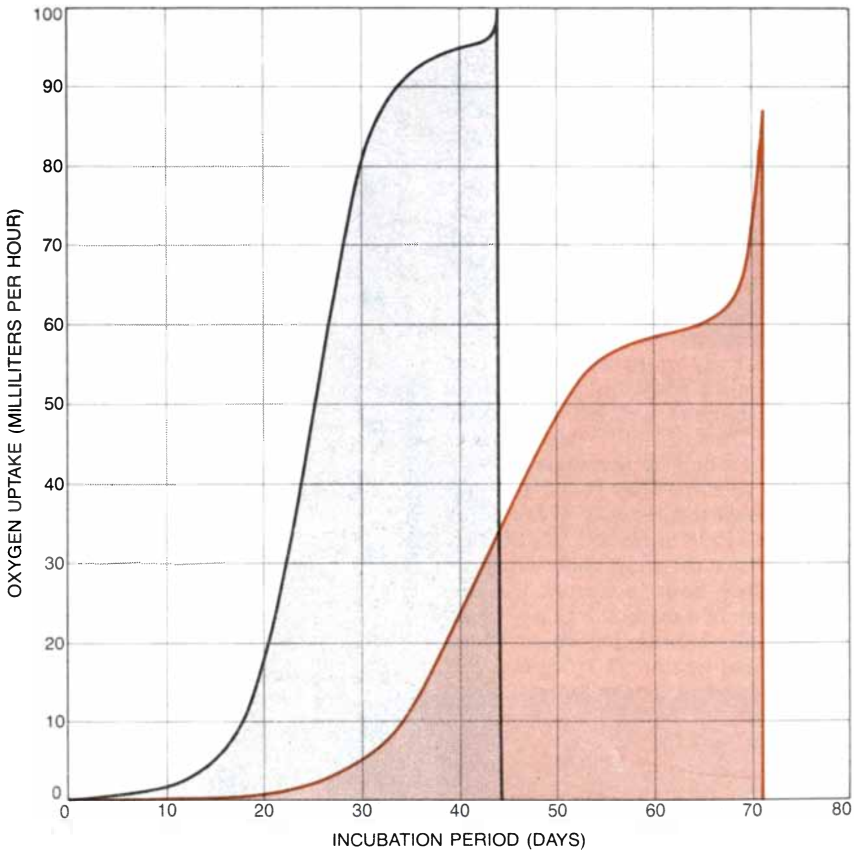
The mean energy content of five eggs was 4,014 kilojoules (959 kilocalories). The figure verified an earlier estimate by Reid, who had assumed a similarity between the kiwi egg and the eggs of domesticated fowl.

This finding means that in proportion to its size and metabolic rate the kiwi leads all birds in the amount of energy invested in the egg. This enormous energy reserve sustains the kiwi embryo during its lengthy development. Moreover, by the time the kiwi chick emerges from the egg only a little more than half of the yolk has been consumed. What remains in the yolk sac is retracted within the chick's abdomen and provides nourishment for the next two weeks or so.

What about water loss during the 71-to-74-day period of incubation? This period is 61 percent longer than the allometric prediction of the period required to hatch a normal egg of the same weight. With evaporation continuing through the pores for this long a period, one might expect to record a water loss that was 61 percent greater, other things being equal. Other things, we found, were not equal.

The kiwi eggshell is much reduced in porosity compared to the avian average. The reduction results from a decrease in the number of pores and in the size of the individual pores. Consequently the water-vapor permeability of the shell is only 60 percent of the predicted normal rate. This slower-than-normal departure of water vapor from the egg perfectly compensates for the longer-than-normal incubation period. The inward diffusion of oxygen would of course be similarly slowed, since the same pores are involved. That, in fact, was what my oxygen-uptake measurements indicated: the uptake rate was only 63 percent of normal.

The question arises: Does the embryo's slower oxygen uptake represent a partial oxygen deprivation? Measurements were made of the oxygen consumption of a two-week-old chick that had finally consumed the content of its yolk sac, and also of the oxygen consumption of two older chicks and five adult brown kiwis. For the chicks the mean value for oxygen consumption at rest was 65 percent of the predicted rate for birds of their weight. The adult kiwis averaged 61 percent of the predicted



UPTAKE OF OXYGEN by the embryo in the kiwi egg over the 71-day incubation period (*color*) is compared with the predicted oxygen uptake for a normal 400-gram egg (*black*). The incubation period for the normal egg is 45 days. As the kiwi embryo passes the first month of incubation the oxygen uptake rises sharply until after the second month it reaches a plateau at about 60 milliliters per hour (1.4 liters per day). Areas under the two curves are the same, showing that the total metabolic cost of development, whether slow or fast, is about the same.

metabolic rate. When individuals of the other two species of kiwis were similarly tested, their metabolic rates also proved to be low. Hence the reduced porosity of the kiwi eggshell limits the oxygen uptake during incubation to roughly the same level that is normal for kiwis after hatching.

The basis of our comparisons was the allometric equation for the metabolism of all birds except the Passeriformes (the order that includes all songbirds). Could it be that the other ratite birds also have a lower-than-average metabolic rate? In collaboration with Terry Dawson of the University of New South Wales we measured the metabolic rate of emus. They too were found to have a lower rate of oxygen consumption than would be predicted on the basis of their body weight. Perhaps the basal metabolism of all the ratites should be reexamined. It may prove to be below the avian average, as the songbirds are above the average (and the marsupials are below the average for placental mammals).

Just as laying the egg marks the termination of the female's obligations, the male eventually fulfills his obligation at hatching—unless the female has laid another egg. The kiwi chick now takes nourishment from its yolk sac until it is able to feed itself. This life-support system works very well. For example, a kiwi chick that weighed 325 grams immediately after hatching dropped to 225 grams before starting to gain weight again as a result of foraging for itself. The 100-gram difference in weight was the lunch mama had packed almost three months ago.

Does the ratio of egg weight to adult weight in the kiwi represent an evolutionary decrease in body size or an evolutionary increase in egg size? As one way of finding an answer, let us suppose that the kiwis' ancestors were flying birds that arrived in New Zealand independently of moas or other ratite birds. Then the large egg of the living kiwis, rather than being the legacy of some moa-related ancestor, would have developed by natural selection after the arrival of these hypothetical flying pre-kiwis.

What are the selective advantages of a large egg? It would favor a more extended embryonic development than is typical of precocial birds, which might serve to schedule energy demands in accordance with the natural abundance of food. The food supply in New Zealand shows little annual fluctuation, however, and so there is no apparent necessity to store up any great reserve of nutrients in the egg.

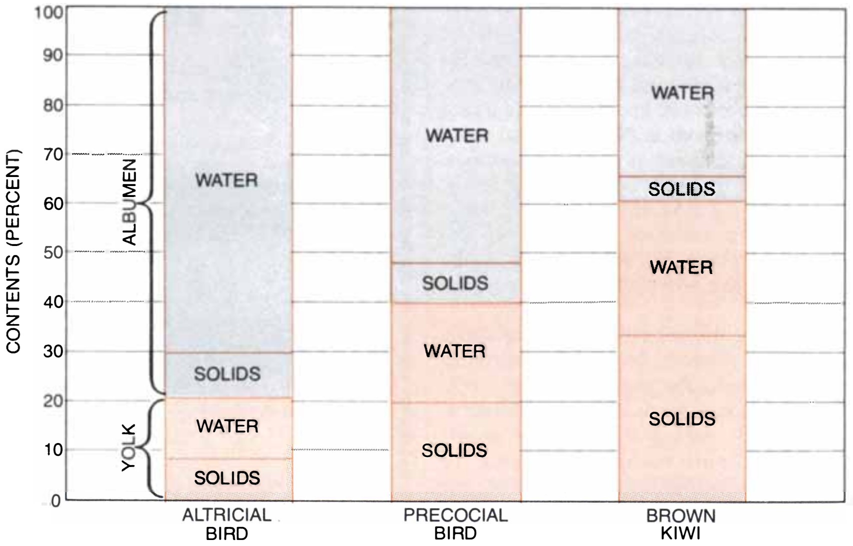
Several species of birds did reach New Zealand on the wing and have since lost or almost lost the power of flight. Among these are the wekas, two species of now-flightless rails (*Gallirallus australis* and *G. hectori*); the takahe, a flightless water bird of the gallinule group (*Notornis hochstetteri*); the Auckland Island flightless duck (*Anas aucklandica*), and the kakapo, a parrot that can no longer fly but can still glide (*Strigops habroptilus*). Each of these birds, although having evolved to a flightless state, has maintained the same proportion of egg size to body size as that of their flying relatives. This suggests that there is no environmental imperative in New Zealand that equates flightlessness with enlarged egg size. For that matter, moa eggs, although large, were not disproportionately so.

For egg size to increase there must be a selective advantage that overrides whatever forces preserve the general allometric norm. If, however, some selective pressure for a smaller body size once existed, say in the evolution of a kiwi from an ancestral moa, the retention of a disproportionately large egg could merely represent the absence of a strong selective pressure for economy in egg size, perhaps abetted by the advantages of hatching a more developed chick.

The pieces of the kiwi story can be put together in more than one way. I prefer to look on this curious bird as a classic example of convergent evolution. In this view an avian organism has acquired a remarkable set of characteristics that we generally associate not with birds but with mammals. That the temperate, forested New Zealand archipelago provides good habitats for mammals is indicated by the success of the exotic mammals introduced there. When there were no mammals present to lay claim to the niches in this hospitable environment, birds were free to do so.

The kiwi must still lay eggs; after all, it is a bird. It is nonetheless mammal-like in a number of ways. For example, Kinsky has reported that kiwis are unique among birds in retaining both ovaries fully functional, so that the female alternates between ovaries during successive ovulations, as mammals do. Also as with mammals the prolonged development of the kiwi embryo proceeds at a temperature below the avian norm. The 70-to-74-day incubation period of the kiwi is much closer to the 80-day pregnancy of a mammal of the same weight than it is to the 44-day period that should be enough to hatch a kiwi-sized egg.

When one adds to this list the kiwi's burrow habitat, its furlike body feathers and its nocturnal foraging, highly dependent on its sense of smell, the evidence for convergence seems overpowering. Only half jokingly I would add to the list the kiwi's aggressive behavior. In the course of my research at the Otorohanga Zoological Society I often had to enter a large pen that was the territory of a breeding male kiwi. When I intruded on his domain at night, he would run up to me snarling like a fighting cat, seize my sock in his bill and drive his claws repeatedly into my ankles until I went away. For this behavior and for the many other reasons I have cited I award this remarkable bird the status of an honorary mammal.



CONTENTS OF THREE EGGS are compared. For both the egg and the albumen the solid content is distinguished from the water content. A chick's state of development at hatching is correlated with the percentage of yolk in the egg. The chick of what is called an altricial bird hatches naked, blind and helpless from an egg that is perhaps 20 percent yolk. The chick of a precocial bird, incubated in an egg that is 40 percent yolk, can run after the hen an hour after it has been hatched. The brown kiwi chick receives no parental care. It is sustained in its long incubation by a yolk that represents 61 percent of the egg's contents; the yolk solids are about a third protein and two-thirds fat. The albumen solids are largely protein. Data for the contents of the kiwi egg were provided by Brian Reid of the New Zealand Wildlife Service.

The new Colt. Isn't a Datsun. Isn't a Toyota. It's a lot of little Dodge.

With a low price, good looks, economy, handling, and performance . . . the 1978 imported Dodge Colt is loaded with the kind of value you buy an import for. And to ensure that it holds its value, Colt is serviced by a nationwide network of Dodge Colt Dealers.

NEW MCA-JET ENGINE SYSTEM.
The economy comes from this brilliant new engineering achievement that manages to squeeze a lot of pep and power out of 1.6 litres and still be rated by the EPA at 46 MPG on the highway and 35 MPG in the city.*

LOTS OF GOODIES.
The value comes from a list of equipment all standard even in our lowest priced two-door coupe. Including whitewall tires, two reclining bucket seats, tinted glass all around, carpeting, adjustable steering column, simulated wood-grained instrument panel, four-speed manual transmission, trip odometer, locking fuel filler door, and electric rear window defroster.

MAKE NO MISTAKE.
It's not a Datsun. Not a Toyota. For $3634,** it's a lot of little Dodge. And you can buy one—or lease one—at your Dodge Colt Dealer's.

*EPA mileage estimates for Colt two-door coupe equipped with standard 1.6 litre engine, four-speed manual transmission, and 3.31 rear axle ratio. Your mileage may vary according to your car's condition, equipment, and your driving habits. California mileage lower.
**Base sticker price, not including taxes and destination charges.

Computer Poker

The familiar card game has interested mathematicians, economists and psychologists as a model of decision-making in the real world. It is now serving as a vehicle for investigations in computer science

by Nicholas V. Findler

In a recent poker session one player bluffed repeatedly, over-representing his hand to his opponents. He won several games, although most of the time another player had better cards. After a number of games the second player caught on, adjusted his playing style to accommodate the bluffing and began to win. The second player, the one whose play improved with time, was a computer program. Computers are now being programmed to play poker and to learn, like human players, from experience.

Games are of interest to economists, psychologists, politicians, military strategists and many others because they provide microcosms of many real decision-making situations. The branch of mathematics known as game theory has been developed to determine the optimum strategy to pursue in games or abstract gamelike situations, involving elements such as competition, conflict of interest, choices of action and incomplete information. The game of poker is particularly interesting because it provides a realistic environment for decision-making involving risk and uncertainty.

Poker presents, among other problems, the question of how to manage money on a long-term basis. (Although a poker session usually lasts no longer than a few hours, 100 or more games, each involving a dozen or more critical confrontations, can easily be played in that time.) Hence the game is relevant to many different areas (such as economic planning, political campaigning and war) where some quantifiable element (in these instances say money, votes and military objectives) must be maximized over a considerable period.

Poker shares another feature with these areas of human activity. In poker, decisions must be made on the basis of imperfect information. Compare poker to chess. Chess is a game of perfect information, that is, all the elements of past and present situations are known to each player, who must in a limited period of time carry out the complex calculations needed to determine a satisfactory move. In poker each player has incomplete information about past and present situations. Some of the player's information is certain (for example the cards in his hand or the betting history of the game), but some of it is risky (for example the probabilities of drawing certain cards) and some of it is uncertain (for example the future behavior of the opponents). Players must form assumptions about their opponents' hands, a task made more difficult by the major role bluffing plays in the game. The players must integrate a variety of different types of information and refer to a continually changing psychological model of their opponents to make an estimation of their status in the game.

This type of decision-making is difficult to reproduce in a computer program because it relies heavily on human judgment. It is this difficulty, however, that makes the programming of poker an attractive problem to computer scientists. Solutions to the problems inherent in poker-playing programs would shed light on other problems also known to depend on human insight and intuition. In particular, an understanding of poker-playing itself can be made use of in many other areas of decision-making.

My students and I at the State University of New York at Buffalo are now at work on a project aimed at exploring the cognitive processes involved in poker-playing. Experiments are being conducted with human subjects in order to identify the decision-making processes utilized in playing the game, and at the same time a major programming effort is being implemented. In this article I shall describe the focus of our research: the development of machine strategies, or computer programs that play poker according to a variety of different principles. To do this I should first outline the form of poker used in our research: draw poker.

At the beginning of a game of draw poker each player pays an ante, or entrance fee, and is dealt a hand of five cards face down. When all the players have been dealt their cards, the player to the left of the dealer is given the option of opening the first betting period of the game. (In most forms of draw poker the opening bettor must have in his hand a pair of jacks or better; otherwise the option of opening the betting passes clockwise around the poker table.) After the opening bet has been placed play continues to pass clockwise around the table. Each player in turn must decide among three options: folding (dropping out of the game and therefore losing whatever money he has contributed to the pot of ante and bet money), calling (putting enough money in to bring his total contribution to the pot up to that of the player who was last active in the game) or raising (making a new bet by putting more money into the pot than is required for calling).

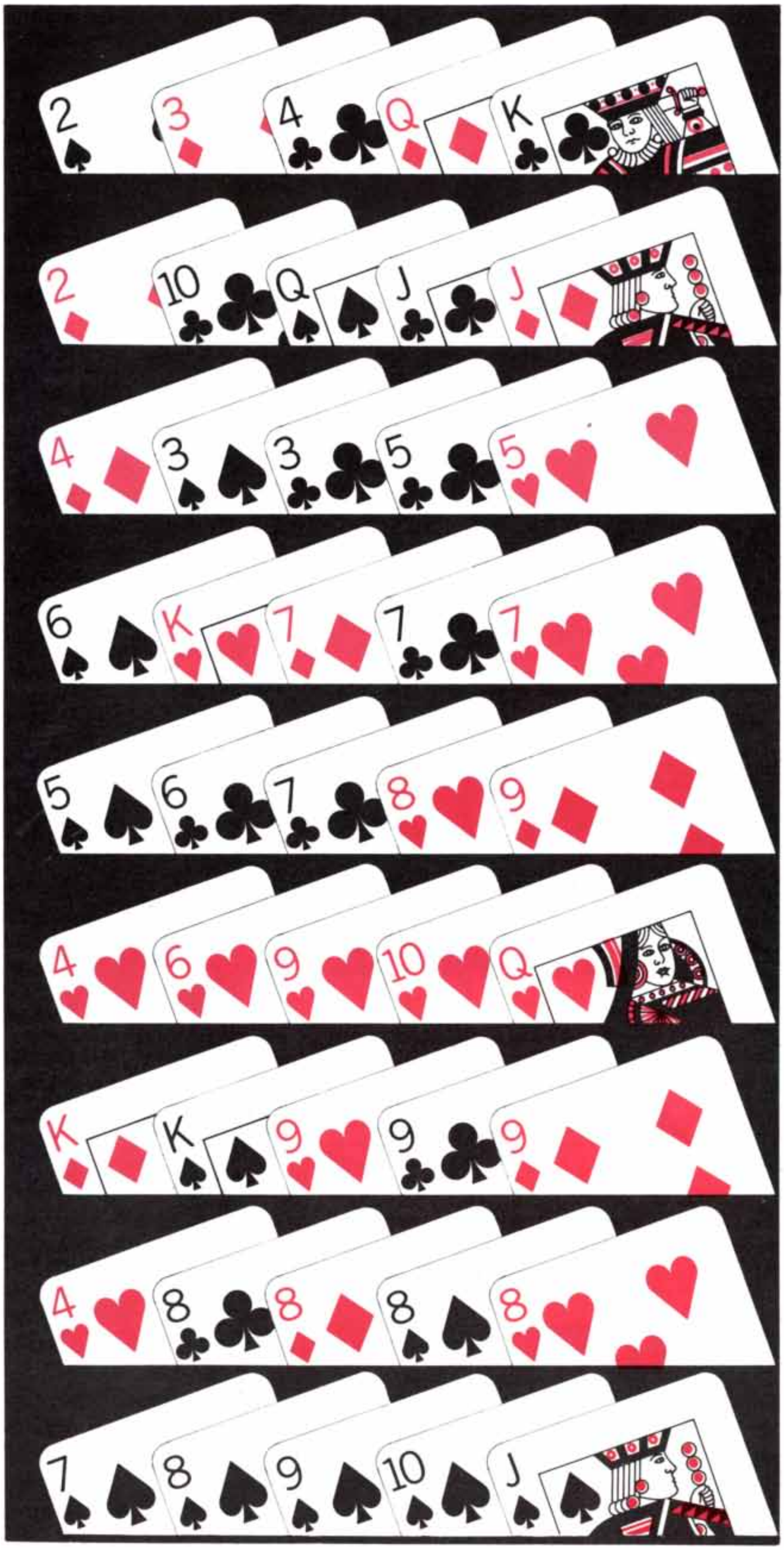
The first betting period is terminated when each active player has put an equal amount of money into the pot. (The number of betting rounds in a period is limited by a rule decided on at the beginning of the poker session.) The next period of play is the draw phase, in which the remaining active players can exchange one or more of the cards in their hand for cards from the undealt portion of the deck. A second betting period follows, ending in one of two ways: If all but one of the players fold, then the remaining player wins the game. Or if several players are still active in the game when the betting stops, there is a showdown: each of the remaining players shows his hand, and the highest-ranking hand wins.

The decisions that a poker player makes throughout the game are based on a number of criteria, the most important of which is usually the strength of his hand. In many instances, however, a player will act in a way that does not truly reflect, say, his position in the game or the strength of his hand. This behavior is called bluffing, and it can be employed for several different objectives. By over-representing a weak hand and thus convincing his opponents to fold, a player may win a game. Or by under-representing a strong hand (sandbagging) and thus keeping his opponents from folding a player may increase the pot he expects to win.

In poker, as in many types of business transactions, deciding how much to bluff in a particular situation and with a given past history is an extremely important issue. Too much bluffing or too little (that is, poor choices regarding the frequency and the intensity of bluffing) can be financially disastrous. Consider the completely rational poker player who depends on what is called a mathematically fair strategy. This player does not bluff at all; he decides how much to bet at a particular point in a game by equating the expected values of his wins and losses, that is, he forms an equation with the expected value of the win on one side and the expected value of the loss on the other and computes the bet from the equation. (An expected value of an event is the product of the probability that the event will happen and the value associated with the event.) The problem with this no-bluffing strategy is that it will succeed only if the opponents are such poor players that they are not capable of identifying the strategy or if all the other players in a game are also mathematically fair. Otherwise the opponents of the mathematically fair player need only raise his bets, regardless of their own hand strength. Eventually he will fold and lose all of his contributions to the pot.

As the example of the mathematically fair player indicates, a successful poker strategy must include a judicious amount of bluffing and an ability to de-

POKER GAMES in which computer programs participate can be monitored on the cathode-ray screen of a computer terminal. The illustration shows the cathode-ray-screen format that enables a human player to interact (in this instance compete in draw poker) with player programs. A player program is a computer program designed to play poker according to a particular strategy. The seven small blue squares around the dark circle on the screen represent player programs; the pink square below the cards of a poker hand represents the human player, who would be seated in front of the screen. The poker hand is the human player's. The current game situation can also be read on the screen. The game is in the betting phase before the draw. There are 18 chips in the pot. Player 4 dealt, Player 6 opened with a bet of 10 chips and Player 9 folded. (The other eight chips in the pot are the antes of one chip from each of the eight players.) The human player, with a pair of kings in his hand, must act next. He may fold, call or raise by pointing a light pen at the appropriate labeled pink square at the lower right. He indicates the amount he wants to raise on the scale across the bottom of the screen. After the human player acts, play continues to pass clockwise around the dark circle. (The names of the players Exploit, Caller 2 and Asprate refer to the strategies they follow.) A similar screen format is used to examine situations in games in which only machine players are competing.



RANKING OF POKER HANDS, from the lowest (*top*) to the highest (*bottom*) is: high card (the card in the hand with the highest value), one pair, two pairs, three of a kind, straight (five sequential cards of two or more suits), flush (any five cards of the same suit), full house (a pair and three of a kind), four of a kind, straight flush (five sequential cards having the same suit).

tect bluffing. Bluffing is used not only to win single games but also to obtain information and to prevent opponents from obtaining information. For example, a player can bluff his way to a showdown—that is, stay in a game that he will probably lose—in order to see an opponent's hand. (In draw poker the winning player is not required to show his hand if all his opponents have folded.) The bluffing player is buying information, paying to see the relation between the strength of an opponent's hand and his betting behavior in the game. In other words, in return for the money the bluffing player loses he receives information about the playing style of one or more of his opponents. By bluffing in different ways a player can also "sell," or pass, misleading information about his own playing style to his opponents. (This type of bluffing might also be called false advertising.) The buying and selling of information that characterizes poker-playing is found in many other decision-making activities.

In order to program a computer to play poker well it is necessary to understand the cognitive processes employed when human beings play poker. (The mathematical theory of games can only treat simplified versions of the game.) In general, a person facing a decision constructs (consciously or unconsciously) a model of the decision problem. Such a model describes the objects and events involved in the problem, and gives their meaning and interrelations in a simple abstract manner. This model-building can be viewed as an ongoing learning process because when new information is received, the model is augmented or modified accordingly. Or if the information is in strong disagreement with the model, a new model is created. Models of this type enable the human mind, a processor of limited capacity, to interpret its environment correctly and to make valid decisions about it.

The model-building involved in game-playing can be loosely divided into three stages. In the first stage a player acquires the rules of the game. By establishing an internal representation of the game and the problem it presents the player also acquires some playing strategies that, although primitive, are superior to random moves. The learning process active at this stage should yield (1) a set of evaluation routines, or procedures for evaluating hand strength, opponents' playing style and so on, (2) the acquisition of information, including playing rules, routines to follow in frequently encountered situations and certain known probabilities pertaining to the game, (3) adequate structures for storing information about a game in progress and (4) decision-making routines for combining the preceding types of information to plan an action.

In the second stage of game-playing

the player learns to generate good moves by formulating possible moves and testing them against different criteria. In the third stage he learns to play a better game. By looking back at the history of play and comparing expected and actual results of various moves the player is able to correct and improve his strategies.

This description of game-playing leaves unanswered many questions about the way that human players develop and employ problem-solving rules. On what criteria are decisions in poker based and what is the relative importance of those criteria? How does a poker player evaluate his position and the position of his opponents? What events or game configurations induce a player to alter a strategy? When is a player willing to buy information and how much will he pay? How do a player's images of his opponents affect his decisions?

The series of experiments my students and I have been conducting with human players are designed to answer these questions and others. Our subjects are experienced poker players who, as in many psychology experiments, are paid a fee for participating in the project. The subjects play poker with their fee money, so that the experiments, conducted under controlled laboratory conditions, entail a real financial risk. (Permission to carry out the experiments was obtained from the Buffalo police department and the university's ethics committee.) Hence observers can make valid generalizations about risk-taking in different situations.

It is beyond the scope of this article to describe these ongoing experiments in detail. Many of the findings have already been incorporated into various machine strategies. It is to this part of our project, the programming of computers to play poker, that I should now like to turn.

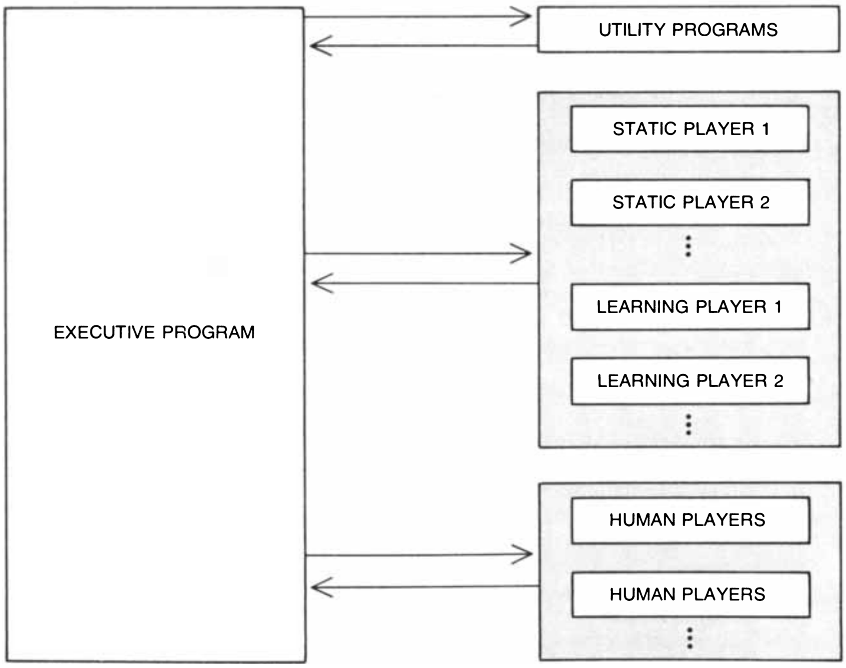
The programming system we have devised consists of three interacting parts. An executive program manages the flow of information and the transfer of control between various components. A large set of utility programs performs tasks such as checking errors, collecting and computing statistical data, calculating probability values, managing game sessions and so on. Finally, a growing collection of machine strategies, or player programs, complete the system.

These machine strategies are not collections of what are known as pure poker strategies, or strategies that prescribe a specific action for every conceivable game situation. A game situation, determined by variable factors ranging from the betting history of a game and the personality of the opponents to the placement of certain players around the poker table, is such a general concept that enumerating all the conditions for even a single pure strategy is a problem of enormous complexity. Moreover, for poker, as for many other games, there are a great many different pure strategies. (The difficulty of formulating pure strategies and the vast number of such strategies explain the practical impossibility of finding an optimum strategy for poker and for games such as bridge and chess.)

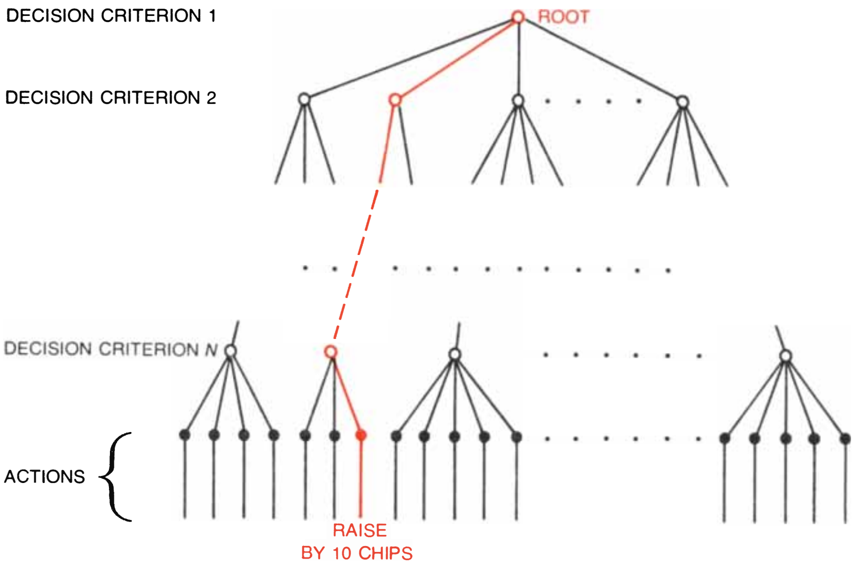
The machine strategies I shall outline prescribe specific moves in response to certain variables pertaining to a game situation, its history and the personalities of its opponents. Each strategy consists of an ordered set of decision components, each relevant to a particular stage of the game. There are several different ways of representing strategies in the computer. For example, some machine strategies are depicted as a set of production rules. Each rule consists of a pattern, or a string of symbols characterizing a configuration of elements that might occur in a game, associated with an action to be taken in the corresponding game situation. When such a representation is used, a pattern is generated to describe the current game and personality situation, and then that pattern is compared with the patterns given in the production rules. When a match (even an approximate one) is discovered between the current pattern and a production-rule pattern, the program transfers the control to the procedure associated with that particular rule.

Another way to represent decision-making is with a decision tree. A tree is a graph in which there are no isolated nodes, or points, and no closed circuits. The nodes of the decision tree are arranged in several levels, with a single node called the root at the top; branches connect each node to one node in the next higher level and one or more nodes in the next lower level. The nodes in a particular level correspond to a game variable pertaining to the strategy the tree represents (say the value of a player's hand or the number of players still active in the game). The branches that lead down from each node correspond to different subranges of values the variable associated with the node might assume. Single "leaves" hang down from the terminal nodes in the bottom level of the tree, corresponding to actions in the game. A game and personality situation is entered at the root of the decision tree, and the current values of the game variables determine a path from the root to a particular leaf, or action.

Strategies can also be represented by an associative network: a graph where the nodes are not arranged in hierarchical levels: any two nodes can be connected, and any number of arcs, or connecting lines, can emanate from a node. Nodes may be arranged on different "planes," if the information they contain is classified according to some criteria, such as the phases of the game with which they are associated. In the asso-



PROGRAMMING SYSTEM designed for computer poker has three parts. The executive program controls the system and manages the flow of information and the transfer of control (*arrows*). The utility programs perform various housekeeping duties including collecting and computing statistics, debugging, error checking and managing game sessions. Player programs play poker games. There are two types of player programs: static players, whose structure and content remain fixed, and learning players, whose structure and content change with experience. Human players can interact with machine players by means of the executive program.



A MACHINE STRATEGY prescribes different actions to be taken according to the values of a set of decision criteria, or variables, pertaining to the current game situation, the history of the game and the personalities of the players. One of the several ways to structure the decision-making process of a machine strategy is with a graph called a decision tree. In the decision tree shown here each level of nodes corresponds to a variable of the strategy the tree represents. The branches leading down from a node correspond to subranges of values that can be assumed by the variable associated with the node. The "leaves" that hang down from the bottom level of nodes correspond to actions in the game. Thus when a game situation is entered at the top node of the tree, current values of variables determine a path to an appropriate action (*color*).

ciative network each node corresponds to some game element (say the number of players still active or an action to be taken), and each arc bears a label characterizing the relation between the two nodes it connects. A game situation is entered at a special entry node of the network and makes its way from node to node through the network until a node corresponding to some action is reached. The labels act as guides for traversing the network, that is, they help in choosing the arc connected to the currently active node that best applies to the game situation under consideration. It can be shown that the three computer representations of strategies are equivalent but one or the other may be more efficient for a given strategy.

The various machine strategies differ not only in their structure but also in their approach to decision-making. Some of them make decisions strictly on the basis of machine intelligence, that is, by following mathematical and logical rules. Others simulate human decision-making, relying on recommendations from poker books and experimental findings about human players. In fact, most of the strategies include at least some elements of human decision-making, which is characterized by the use of ill-defined decision criteria and what is termed heuristic reasoning.

Decision criteria are ill-defined when they are imprecise or generally qualitative. For example, in order to sell an image of himself as a timid player a poker player might plan to fold early in several games where the stakes are high. Although no exact value can be assigned to the attribute "early," computer programs can be designed to find out how best to respond to opponents exhibiting given types of behavior.

Heuristic reasoning, based on what is often regarded as insight or intuition, enables human beings to obtain solutions without examining all of a (possibly enormous) mass of relevant information. In a heuristic program a problem is solved not by conducting an exhaustive search for a solution but by making use of certain rules of thumb and the various approximations and shortcuts that characterize human judgments. By analyzing the actions of expert poker players it is possible to formulate heuristic rules that can be incorporated into computer programs. These heuristic programs obtain solutions that are not necessarily the best but are usually fairly good and reasonably efficient.

The player programs my students and I have developed fall into two categories: static strategies and learning strategies. In static strategies neither the content nor the structure of the strategies changes with time, whereas in learning strategies one or both do change as experience is accumulated.

A static strategy does not necessarily respond to identical configurations of game elements with the same action; bluffing may be programmed to occur at random or in response to certain configurations of other elements. The only nonbluffing static strategy is named the Mathematically Fair Player. Following the style of the completely rational human player I described above, this program decides when and what to bet by equating expected values of wins and losses; in other words, it bets strictly according to the odds of winning with a particular hand. Of course, this strategy cannot defend itself against sustained bluffing, but it performs rather well in games against relatively unsophisticated static players. In fact, it is convenient to rank different static players according to how well they do against the Mathematically Fair Player.

Another static strategy, the RH Player, makes decisions on the basis of a wider range of information than the Mathematically Fair Player does. The RH Player considers not only the strength of its hand but also the values of what are called second-order factors in a poker game, factors that are less important than strength of hand but nonetheless play an important role in determining the outcome of a game. The combined value of these second-order factors can be expressed by a variable called the RH index. (The variable was designed by, and named for, two of my graduate students, Jean Rachlin and Gary Higgins.)

The RH index measures the desirability of staying in a game for a particular player at a particular moment. This task is accomplished by putting the variable expressing positive motivation in the numerator of a fraction and the product of the variables expressing negative motivation in the denominator. The positive motivation is the amount of money to be won, expressed by a variable designated TABLEPOT. The negative motivations all relate to the possibility of losing. Among these factors the variable LIVE is equal to the number of players still active in the game. The variable RAISECOUNT is equal to the number of times the amount of money in the pot has been raised during the current betting period and can indicate (among other things) the determination of the active players to stay in the game. The variable FOLLOWERS is equal to the number of players who can act after the RH Player takes some action. Finally, the variable RAISE is equal to the cost of staying in the game, or the amount of money the player must add to the pot. Thus the RH index is equal to:

$$\frac{\text{TABLEPOT}}{\text{LIVE} * (\text{RAISECOUNT} + 1) * \text{FOLLOWERS} * \text{RAISE}}$$

(RAISECOUNT + 1 is used so that zero never appears in the denominator.)

Rachlin and Higgins conducted a series of computer experiments to determine the characteristic distribution of the values of the RH index. They generated random game situations on the computer, computed the values of the

RH index for each situation and then checked books by poker experts in order to determine the recommended action in each situation. It was possible to correlate actions and index values. In fact, Rachlin and Higgins were able to partition the entire range of values of the RH index into segments for which, depending on the strength of a player's hand, there is a single recommended action. This partition has proved to be very satisfactory. The RH Player does very well in games against other static players, including the Mathematically Fair Player.

Static players are not the most interesting part of the research my students and I are conducting. The focus of our project is the development of learning players: programs that raise their level of performance and become better with experience. At present some of these learning players are little more than a learning component attached to a basic set of game rules. These players cannot yet be considered to be independent, competitive strategies like the static players.

Poker tournaments are being conducted with learning players in order to determine how the learning components function under different conditions, how they interact with each other and how the rate of improvement in their performance can be changed. Static players are also included in these experiments in order to reduce the complexity of the evaluation of the results when learning programs learn only from other learning programs.

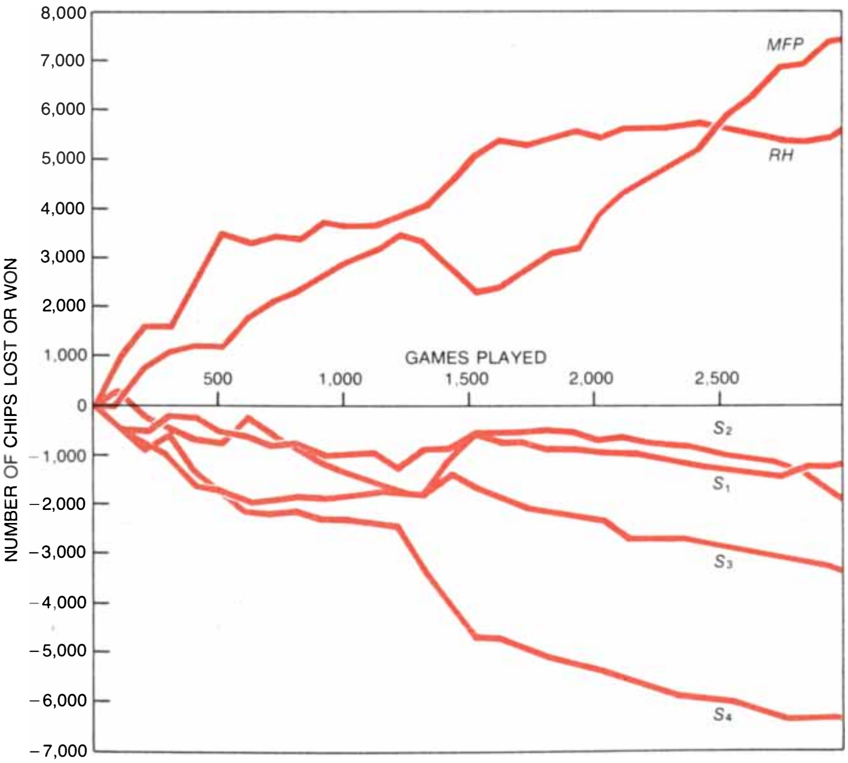
A number of learning components performing different functions have been devised and are now being tested. For example, the player called the Adaptive Evaluator of Opponents makes an initial rough estimation of each opponent's situation and then uses accumulated experience with the opponent's playing style to improve the estimation continually. This estimation takes the form of a list of possible strengths of the opponent's hand, ranked according to their plausibility. As the player gathers information about the style of the other players, the list of possibilities is progressively culled so that it becomes progressively more reliable. The learning process takes place along three dimensions in the program: in compiling statistical information about each opponent's activities, in choosing the most informative combination of events to observe for each opponent and in refining the logic of past judgments made about each opponent's hand.

The Adaptive Aspiration Level Player is based on the belief that in human cognition, decisions to undertake new courses of action are activated by an intermediate mechanism that might be called an aspiration level. A current aspiration level can be considered to be the result of balancing factors such as, on the one hand, the cost of a quest for better outcomes and, on the other, the value of satisfaction with the current rate of achievement. Human aspiration levels can rise and fall in response to changes in the configuration of perceived events and their estimated value. For example, a rather conservative poker player who suddenly loses a large sum of money might begin to play recklessly because he intensely desires to recover his losses. The learning component of the player program incorporates an aspiration-level mechanism, which is adjusted when, say, the financial status of the player changes much more than expected or the assumed probability of certain outcomes is significantly altered. This player is rather skillful in choosing a compromise concerning conflicts between offensive and defensive behavior, loss recovery and profit protection, success-seeking and failure-avoiding modes of operation, and so on.

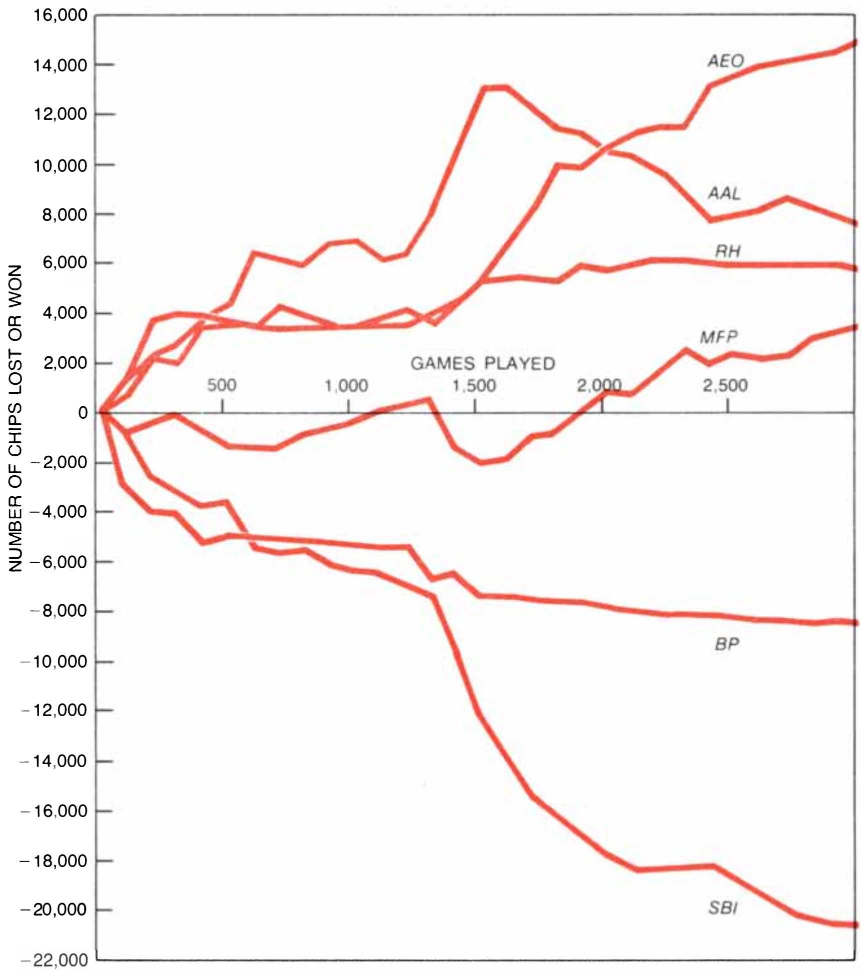
The next two learning players reproduce the kinds of bluffing found in human play. The player that Sells and Buys Images bluffs in order to impart misleading information about itself and to obtain information about the true personality of the opponents. Just like a human player, this program sells an image by adhering to a particular style of play over a long period of time and then acting in violation of that style in order to cash in on an opponent's disorientation. The program buys an image by establishing a certain relation between an opponent's actions and that opponent's situations in the game.

The learning component of this player is built around a metric that characterizes the "distance" of the image being sold or bought from the mathematically fair strategy. Elements of the opponents' personalities are also considered by this player. Broadly speaking, the farther away the image to be sold is from the Mathematically Fair Player, and the more cautious the opponents appear to be, the longer it takes to accomplish the transaction of selling the image. Similarly, the more extreme an opponent's playing style is, the more games this player goes through before buying it (believing in it).

The Statistically Fair Player bluffs in a different way. It corrects the shortcomings of the Mathematically Fair Player with a learning component that



LONG-TERM MONEY MANAGEMENT is called for in poker, as it is in many other human activities; the success of a poker strategy is measured not by the amount of money a player wins or loses in a single game but by the total amount won or lost over a long series of games. The illustration shows the cumulative wins and losses of six static machine players competing over 10 tournaments, or 3,000 games. The most successful players are the Mathematically Fair Player (*MFP*) and the RH Player (*RH*). The Mathematically Fair Player bets strictly according to the odds of winning with a particular hand. The RH Player considers, in addition to hand strength, several secondary decision criteria combined in a variable called the RH index. The other four players are static strategies based on decision trees of varying complexity.



LEARNING PLAYERS improve their own strategy as they accumulate experience. The graph shows the cumulative wins and losses over 3,000 games of four learning players and two static players. (Static players are included as bench marks for the evaluation of the rate of improvement of learning strategies and as reference points for the learning processes.) The learning player named the Adaptive Evaluator of Opponents (*AEO*) makes an initial judgment about the strength of each opponent's hand and then refines those judgments as play continues. The player that Sells and Buys Images (*SBI*) bluffs like a human player in order to project ("sell") a misleading image of itself and to identify ("buy") the image, or playing style of its opponents. The Bayesian Player (*BP*) is able to make inductive inferences and learns to improve its play by comparing the predicted consequences of actions with the actual consequences at various points in a series of games. The Adaptive Aspiration Level Player (*AAL*) reproduces a mechanism believed to exist in human cognition: a flexible aspiration level that decides whether risks are worth taking by considering the risk-taker's current status and past record. The aspiration level in player program is adjusted as a result of unexpected changes in the game situation.

identifies opponents' bluffing, judges the extent and frequency of the bluffing and adapts its play accordingly. It also identifies nonbluffing strategies, such as the Mathematically Fair Player, and bluffs to defeat them. The learning component of the Statistically Fair Player is very flexible, and it can adjust its playing style to deal with the entire range of player personalities, from the most extravagant to the most timid. Hence it does well in games against other player programs.

Four additional learning strategies have been developed that are able to make inductive inferences, that is, to draw general conclusions from specific events. All four strategies derive conclusions and make generalizations from them by considering game actions that proved to be successful in given situations. They differ in the amount and types of information they collect. Hence the rates of improvement of the strategies differ, as do their playing abilities.

These strategies are called Bayesian, after Bayes's theorem in probability; they continually adjust their own decision-making rules by comparing predicted outcomes of events and actual outcomes. There are basically three ways to adjust the rules so that the actual outcomes will be closer to the expected ones. If a number of parameters are included in the heuristic rules, a learning component can make their values converge to near-optimum values. Or an optimum hierarchical ordering for the heuristic rules can be found experimentally, by observing the frequency of their successful employment. It is also possible to generate new heuristic rules, test them and incorporate the successful ones into a strategy. This procedure, however, is the most difficult and time-consuming of the three.

Four more learning programs are still under development. I shall describe them briefly since each one has a wide range of applicability. For example, the Pattern-Recognition Strategy examines past games in order to discover possible causal relations between events in past games and opponents' actions. These events may be tendencies or cyclic or isolated phenomena. (The program also allows for a time lag between an event and a resulting action.) The postulated correlations, weak as they may often be, help the player to discern the reasons behind opponents' actions, a very important aspect of poker-playing.

The Advice Taker and Inquirer functions in a way that is different from all the player programs I have described so far. The program interacts with a human user in a language that is reasonably close to English. Fairly sophisticated strategies, in the form of game principles and specific examples, can be taught to this player, which checks all such material for consistency and completeness, asks prompting and clarifying questions and structures the information received so that it can be applied to actual game-playing.

The last two learning players under development differ from the other players in another respect. They aim at creating normative strategies: strategies that are optimum in the sense that they tell the player what to do in order to maximize the amount of money gained in the long run. (Up to this point I have discussed only descriptive, or non-normative, strategies.) Consider the EPAM-like Player. It was named for the similarity of its philosophy to that of the Elementary Perceiver and Memorizer, a project modeling verbal-learning behavior designed by Herbert A. Simon of Carnegie-Mellon University and Edward A. Feigenbaum of Stanford University. This player constructs a descriptive decision tree for each opponent, characterizing the opponent's method of discriminating between different game situations and of deciding on appropriate action. The EPAM-like player also constructs another tree with respect to the entire set of opponents: a tree that shows the optimum way to act for itself in all possible game situations. The two types of tree grow as the need arises. They are modified both in small continuous steps and in jumps, simulating the way human beings learn.

The Quasi-Optimizer seeks to generate a normative theory of playing poker by analyzing various descriptive machine strategies (descriptive theories of play) and combining different components from them into a "super" player. The Quasi-Optimizer operates in three stages. In the first stage the player accepts a number of machine strategies (as impenetrable "black box" programs) and builds a model of each in the form of, say, a decision tree. In the second stage credit is assigned: a series of games are played with the strategies to determine which of their components are responsible for good consequences and which are responsible for bad ones. In the third stage a new player program is synthesized by selecting the best element of the available machine strategies for each distinguishable segment of the new strategy. The new strategy is said to be quasi-optimum because it is optimum only in play with the particular set of opponent strategies used in the experimentation stage. A new optimum strategy must be generated for a different set of opponents.

There are obviously many applications outside of poker for learning strategies like the Adaptive Evaluator of Opponents, the Advice Taker and Inquirer, and so on. Learning programs in general are important in pointing to future directions in computer programming. Ultimately it should be possible for programmers to exert less effort in perfecting their programs and for people with less programming skill to communicate effectively with computers.

Computers can be used to solve problems for which there are no algorithmic, or step-by-step, solutions or for which the formulation of such solutions is cumbersome or prohibitively expensive. As the research I have described has demonstrated, computer programs can model the human approach to decision-making. The fact that the programs do reproduce human actions indicates that they make decisions for the same reasons human beings use in their cognitive processes. This statement is corroborated by a series of experiments currently being performed.

In these experiments a human player is seated at a computer terminal and is asked to play poker against a number of opponents. Some of the opponents are player programs and others are human players seated at terminals in different rooms. The subject must determine which of the, say, seven opponents are human players. (The player programs are adjusted so that their speed in computing moves does not give them away.) The results of these experiments are not yet conclusive, but it appears that unless the subject knows the decision-making mechanisms of player programs, and he and his human opponents are very experienced poker players, he will not be able to differentiate between machine and man in a statistically significant number of cases.

THE AMATEUR SCIENTIST

Chemical systems that oscillate between one color and another

by Jearl Walker

If you mixed several chemicals thoroughly, you would expect the mixture to have a single uniform color. Certain chemical reactions, however, yield a surprising oscillation of color. About once a minute for an hour or so they change from one color to another and back again.

One of the most famous oscillators of this kind was discovered in 1958 by the Russian chemist B. P. Belousov and was subsequently investigated by A. M. Zhabotinsky and others. The original reaction involved a mixture of potassium bromate, ceric sulfate and citric acid in dilute sulfuric acid. The oscillation was from colorless to yellow. Other mixtures that have been found since then oscillate between red and blue, violet and blue, red and green and even from colorless to gold to blue. Here I shall describe how you can make several of these oscillators and how you might look into the events that determine their periodicity.

In this work you will be dealing with questions that have not been fully answered. Oscillations might be expected in nonuniform mixtures, because they can be attributed to the diffusion of the ingredients. Oscillations in a closed system consisting of a homogeneous mixture, however, were thought to be impossible on the ground that they would violate a basic law of physics and chemistry to the effect that any spontaneous reaction must steadily lower the Gibbs free energy of the system. (I shall return to the matter of Gibbs free energy.)

In 1970 Zhabotinsky and A. N. Zaikin reported finding periodic structures (circular waves) running through a solution that was a slightly altered version of the original mixture. Later Arthur T. Winfree, who is now at Purdue University, improved the formula for the solution and extensively examined the periodic structures of the rotating spiral waves that appeared in it [see "Rotating Chemical Reactions," by Arthur T. Winfree; SCIENTIFIC AMERICAN, June, 1974]. Winfree has devised for me a do-it-yourself kit that I shall describe.

He suggests the following procedure for mixing the reagents for a red-to-blue oscillator. Add two milliliters of concentrated sulfuric acid and five grams of sodium bromate to 67 milliliters of pure water (distilled or deionized) to get 70 milliliters of solution. *As a rule of safety always add the acid slowly to the water, not the water to the acid, because of the danger of explosion.*

Pour six milliliters of this solution into a glass container and add .5 milliliter of a sodium bromide solution made by adding one gram of sodium bromide to 10 milliliters of water. Next add one milliliter of a malonic acid solution (malonate) made by putting one gram of malonic acid in 10 milliliters of water. After the bromine color disappears mix in one milliliter of .025-molar phenanthroline ferrous sulfate (a dye sometimes called ferroin). Add the tiniest trace (about one gram per liter) of either Triton X-100 surfactant or "photoflo" (a substance employed in photographic darkrooms) to reduce the surface tension and thereby aid the spreading of the fluid in a thin layer.

Pour out such a layer in a clean vessel such as a culture dish. Stir the solution well and then wait five minutes or so for the color oscillations to begin. From time to time thereafter gently swirl the solution around in the dish or gently stir it. The colors will pulse with a period of about one minute, although the blue will persist for only about five or 10 seconds at a time. Small bubbles of carbon dioxide will also form in the solution. They result from one of the reactions involved in the oscillation and can be removed by stirring the fluid occasionally.

The concentrations are not crucial. You must take care, however, to use only clean containers and to keep your fingers out of the solution and off of surfaces that the solution will touch. The chloride in the salt on your skin can prevent the oscillation reactions.

Before Winfree puts the sodium bromate into his mixture he usually recrystallizes it to remove impurities that can prevent oscillations. You can do this by dissolving as much sodium bromate as possible in a clean container of warm distilled water, taking care not to bring anything soaked with the sodium bromate near an open flame because of the danger of rapid oxidation. Wrap the closed container in an insulating material such as Styrofoam and leave it undisturbed for a week or two. Then pour off the remaining liquid, since you might want it for the next recrystallization, and scrape the sodium bromate off the walls with a clean instrument.

Before I go into variations on the basic mixture I should explain why color oscillations arise. The reactions that appear to be responsible for the oscillations are somewhat complicated. In the solution both the bromide and the bromate react with the malonate to form bromomalonate. If it were not for an inhibitory effect of the bromide, the bromate would also react with the phenanthroline dye by oxidizing the iron in it (adding one electron to the outer shell of electrons of each iron atom). The reduced (ferrous) state of the dye is red, whereas the oxidized (ferric) state is blue, so that if the bromide were to stop inhibiting the oxidation, the solution would switch from red to blue.

Eventually that is exactly what happens, because the bromide is almost totally consumed in the reaction with the malonate. Then the oxidation can proceed, the phenanthroline dye is oxidized and the solution turns from red to blue.

Why does it subsequently change back to red? The bromomalonate that has been and is still being produced reduces the ferric form back to the ferrous form. (Reduction is the opposite of oxidation.) The solution therefore becomes red again. Moreover, bromide is released from the bromomalonate as a product of the reaction, so that the inhibition of the oxidation of the dye by bromate is reinstated. Soon the bromide has been consumed again, and the cycle is renewed. The oscillations continue for an hour or so until the solution becomes permanently blue or red, depending on the initial concentrations.

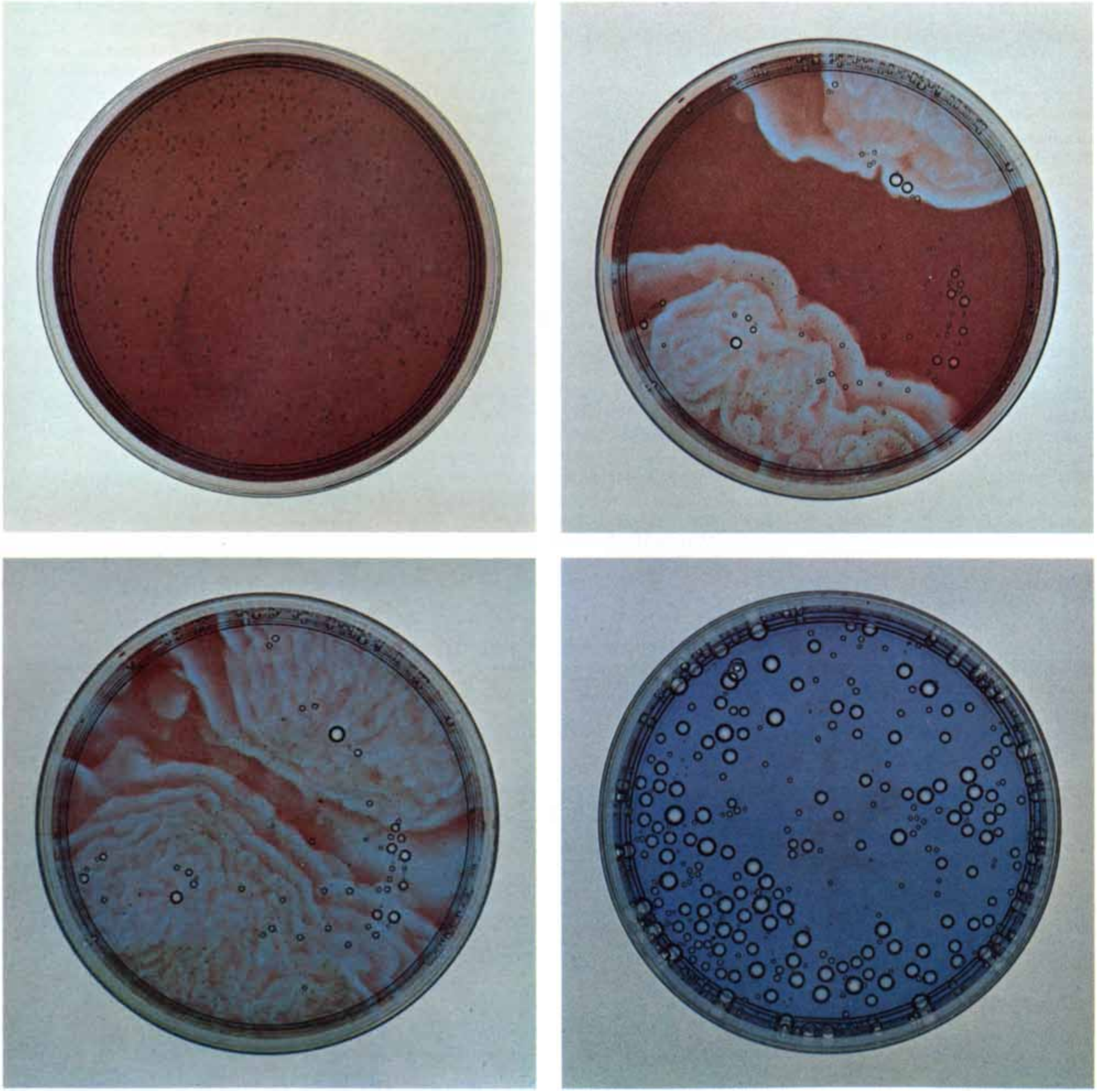
Winfree has described several variations that you can make on his basic mixture. In order to try them you will need to be familiar with the terms molar and normal as they apply to concentrations. The term molar is based on the atomic weight of the molecule or complex being put into a solution. In a one-molar solution the number of grams of the solute put into one liter of the solvent is equal in number to the atomic mass of the solute (expressed in atomic mass units). The number is usually printed on the container of the chemical. The phenanthroline as sold is already in the concentration required. A normal solution is similar to a molar solution but is also proportional to the charge number of one of the ionic species released in dissociation. For example, a one-normal solution of sulfuric acid is also a two-molar solution, two because once the acid is in solution the two hydrogen atoms from a dissociated molecule have a net charge of plus two.

Now for Winfree's variations. One is to leave out bromide (which will, of course, still be produced when the bromomalonate reduces the ferrous phenanthroline). The color oscillations are then more frequent, since less of the bromide must be consumed in each color cycle before the bromide's inhibiting effect on the oxidation of the iron is eliminated and the color of the solution can change to blue.

You can enhance the oscillations and obtain a deeper view into the solution by replacing some of the ferroin with ceric sulfate having an initial concentration of .1 molar. If you want to avoid acid stains that result from spilling the reagent in the experiment, substitute sodium bisulfate for the sulfuric acid. You might want to vary either the temperature of the solution or the concentration of the basic ingredients to see how the period of oscillation is altered. (Teachers interested in setting up a laboratory so that students can try these variations will find helpful an arrangement worked out by John F. Lefelhocz of Virginia Commonwealth University; his paper describing the setup is cited in the bibliography of this issue.)

Mixtures other than Winfree's can produce color oscillations. The original Belousov reactions involved ceric-cerium rather than feric-ferrous solutions; the transitions were from colorless to yellow and back. Richard J. Field of the University of Montana suggests adding ferroin to produce the more striking variation from red to blue. The concentrations he recommends for a lecture demonstration are shown in the top illustration on page 156. The colors you see in such mixtures depend on the relative amounts of the cerium-ceric ions (which oscillate between colorless and yellow) and the iron ions (which oscillate between red and blue). With appropriate amounts of yellow the colors may oscillate between blue and violet or green and red.

Oscillations between brownish-yel-



Arthur T. Winfree's bulk oscillator changing from red to blue low and pink can be obtained from solutions of manganese II and III, which replace the cerium III and IV and the iron II and III couples of the Belousov and Winfree reactions. (The Roman numeral describes the ionization state; iron II is the doubly ionized iron atom, or the ferric state.) The malonic acid can be replaced by other acids: citric, maleic, malic, bromomalonic and dibromomalonic. In many of the solutions the oscillations arise only if the solution is gently stirred. For some solutions one must wait as long as 40 minutes until the oscillations begin.

A system that changes from colorless to gold to blue and back again can be achieved at relatively low expense with an iodine clock that has been described by Thomas S. Briggs and Warren C. Rauscher of Galileo High School in San Francisco. I have tripled their published concentrations (in distilled water) in order to be clear about the amounts of chemicals to be put into each of three containers. The containers hold equal amounts, so that the final mixing (in which the contents of the three containers are put into one container) will yield the published concentrations.

Pour into the first container 3.6-molar hydrogen peroxide, made by adding 40 milliliters of 30 percent hydrogen peroxide to 60 milliliters of water. *Handle the hydrogen peroxide with extreme care.* In the second container mix .201-molar potassium iodate (4.3 grams per 100 milliliters of solution with water) and .159-molar perchloric acid, made by adding 2.3 milliliters of 70 percent perchloric acid to enough water to make 100 milliliters of solution. You will probably have to warm the potassium iodate solution to dissolve the chemical completely. In the third container mix .150-molar malonic acid (1.5 grams per 100 milliliters of solution with water), .0201-molar manganese (II) sulfate (.3 gram per 100 milliliters of solution with water) and .03 percent starch. (Percentage refers to the constituent's percentage of the weight of the solution. Here .3 gram of starch in one liter of distilled water is a .03 percent starch solution.)

When you are ready for the oscillations, mix together equal amounts of the three solutions. The blue comes from the blue starch complex that develops periodically when the iodide concentration is near its maximum value. Briggs and Rauscher obtained brief oscillations when they replaced malonic acid with 2,4-pentanedione. The oscillations are faster if you replace manganese with cerium.

In order to get bulk oscillations of color in many of these mixtures the solution must be continuously stirred (with a magnetic stirrer if you have access to one) or swirled. Otherwise another phenomenon occurs, although it is one of much interest in itself. Winfree has studied it in detail. Waves of color appear on and in the solution, propagate through it at a few millimeters per minute, rotate around points of origin, destroy each other as they meet and describe a variety of shapes: rings, ellipses and spirals.

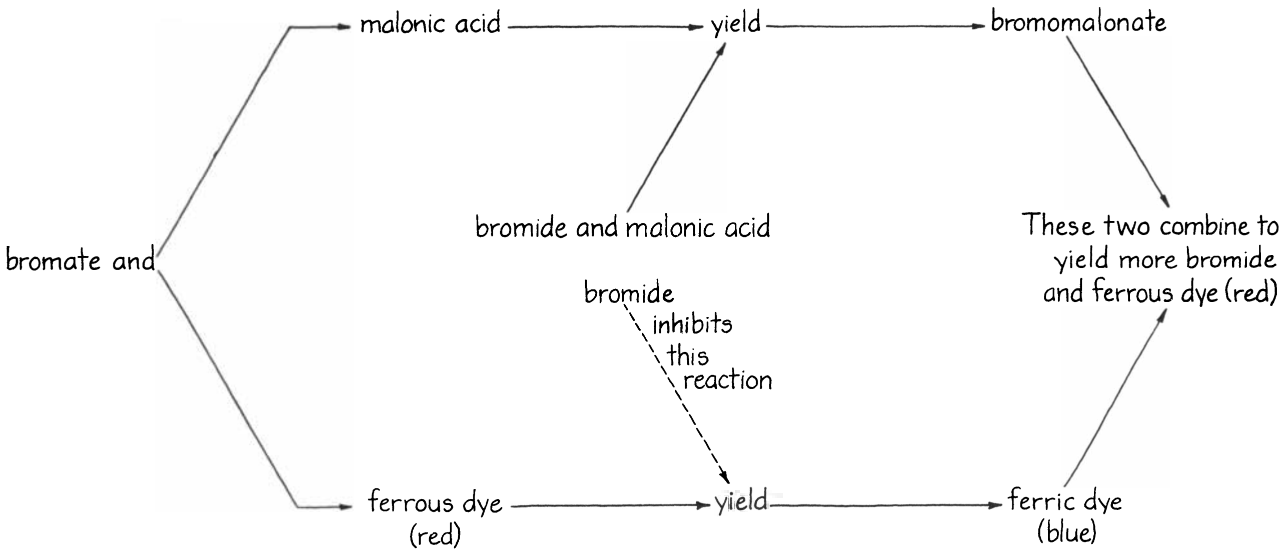
To enhance these color waves (by diminishing the bulk oscillations) Winfree suggests using more bromide and less sulfuric acid in the basic mixture. For example, in the sodium bromate solution use one milliliter of concentrated sulfuric acid and 67 milliliters of distilled water.

To enhance the visibility of the waves pour some of the final solution to a depth of about a millimeter in a container such as a plastic tissue-culture dish. The container must be very clean; any small grains of dirt or scratches will promote the formation of bubbles of carbon dioxide during the experiment and may generate too many simultaneous waves. Visual contrast is improved if you place the dish over another dish containing blue copper sulfate and a few drops of sulfuric acid and then place a light below the two dishes. The solution in the bottom dish will also protect the chemical waves from disruption by absorbing heat from the light. Cover the tissue-culture dish with its top to protect the waves from air currents and to eliminate evaporation. Leave the dish undisturbed. If the waves do not begin within a few minutes, touch the solution with a hot needle, which will set off waves. Usually, however, there will be enough bubbles, scratches or dust motes in the solution to give you a few waves.

Presumably the waves result from the same chemistry that gives rise to bulk oscillations. The solution stays in the red state with some equilibrium concentration of bromide. Once the mixture is disturbed by a bubble, a scratch, a dust mote or a hot needle a portion of the solution consumes its bromide. The bromate can then switch the color of solution to blue by oxidizing the iron in the phenanthroline. This conversion propagates outward as a wave because as diffusion brings bromide from the red area just outside the blue ring into the blue area the additional bromide is also consumed and the formerly red area turns blue. Once the blue ring has passed a particular area the solution regains its redness as the bromomalonate reduces the iron in the phenanthroline and releases more bromide.

The waves are truly chemical waves rather than hydrodynamic waves because the fluid is motionless. Moreover, they will not reflect from barriers placed in their way, and waves traveling in opposite directions will not pass through each other.

You will find that the waves, which have a length of a few millimeters, assume several shapes. Among them are concentric closed rings forming a bull's-



Reactions in the Winfree oscillator

eye, with new rings appearing in the center at regular intervals. These patterns do not rotate. Other waves form spirals, all of which have the same period, a shorter one than any of the variety of periods of the closed-ring patterns. These spiral patterns rotate. If the axis of rotation lies perpendicular to the surface of the solution, scroll waves (as Winfree has named them) propagate from the axis, producing a spiral structure around a central region surrounding the top of the axis as you look from above the solution. The structure is technically an involute spiral, which could be drawn mechanically by having a pencil tied to a string that is slowly unwound from a central cylinder instead of a single point. In the chemical involute spiral the central core is less than a millimeter in diameter.

If the solution is deep enough (deeper than the width of the spiral's core), the axis may lie tilted to one side. From above you then see elongated spirals surrounding the central area of the pattern. Also in an appropriately deep solution the scroll axis may bend over to form a *U* with its top on one surface of the solution and its bottom on the other. From one orientation you see a nest of rings surrounding a central point. If the *U* shape shrinks away from the surface having the rings, the rings disappear on that surface just before becoming complete circles. When the axis of rotation lies parallel to a surface and in a closed ring, the surface has concentric circles with the inner one propagating inward as the outer one propagates outward.

If you would like to place barriers in the way of the wave, a more viscous solution is desirable. It can be made by adding two milliliters of colloidal silicon dioxide to each milliliter of solution. Winfree describes the resulting solution as having a consistency similar to that of peanut butter.

You can preserve the patterns and lift them out of the solution by putting filters in the solution. The formation and propagation of waves are unaltered by the filters, which can be obtained from the Millipore Corporation (Ashby Road, Bedford, Mass. 01730) at $16.20 for a package of 10 type GSWP 142 00 filters. You must guard against contamination of the solution when you put the filters in it. Use nylon forceps and do not touch the filters with your fingers. If you first warm up the filters (even with body heat), the patterns are more lively.

Once the solution has been absorbed into a filter Winfree removes the filter and puts it between sheets of plastic (such as a food wrap), places it in oil to reduce evaporation and exclude oxygen, or sticks it to the inside of the lid of a plastic dish. A filter can be used again if you rinse it out well with water.

When Winfree puts a filter in oil, he can play a game with the chemical waves by touching the filter briefly

Reactant:	Concentration range (molar):	Convenient concentration (molar):
Cerium ammonium nitrate	.0001 — .01	.002
Malonic acid	.0125 — .5	.275
Sulfuric acid	.5 —— 2.5	1.5
Ferroin	.0006	.0006
Potassium bromate	.03 —— .0625	.0625

Richard J. Field's concentrations for the cerium oscillator

through the film of oil with a piece of iron such as the edge of a razor blade. The resulting red layer, which he calls the "iron curtain," blocks waves in that area for several seconds. The redness apparently results from the fact that the acid in the solution pulls ferrous-state iron out of the razor blade.

You may want to photograph the waves as they develop. You can also stop them in place on a filter by following a procedure that Winfree devised. (It will not work with a filter that has been put in oil.) At a time when waves are propagating through the filter remove the filter and put it in ice-cold saturated salt water to stop the reaction. Leave the filter there for about five minutes (until all the chemicals have diffused out of it except for the red ferrous phenanthroline) and then blot it dry, dip it in a solution of sodium iodide (one gram per 40 milliliters of solution with water) and blot it dry again. Then leave it for 10 minutes in the iodine vapor emanating from a bottle containing iodine crystals. This procedure permanently fixes the ferroin.

Rinse the filter in distilled water, dry it in air and clear it by floating it on paraffin oil. Blot it again. Finally, sandwich it between two sheets of clear sticky plastic such as Contact paper. Winfree says that such preserved specimens of chemical waves have lasted for years with full sharpness. Preservation for about a day can be achieved more easily by dipping the filters in an ice-cold solution of 3 percent perchloric acid, but the sharpness is diminished somewhat, and the pattern vanishes when the filter dries fully.

Spiral-wave patterns in an unstirred Winfree mixture

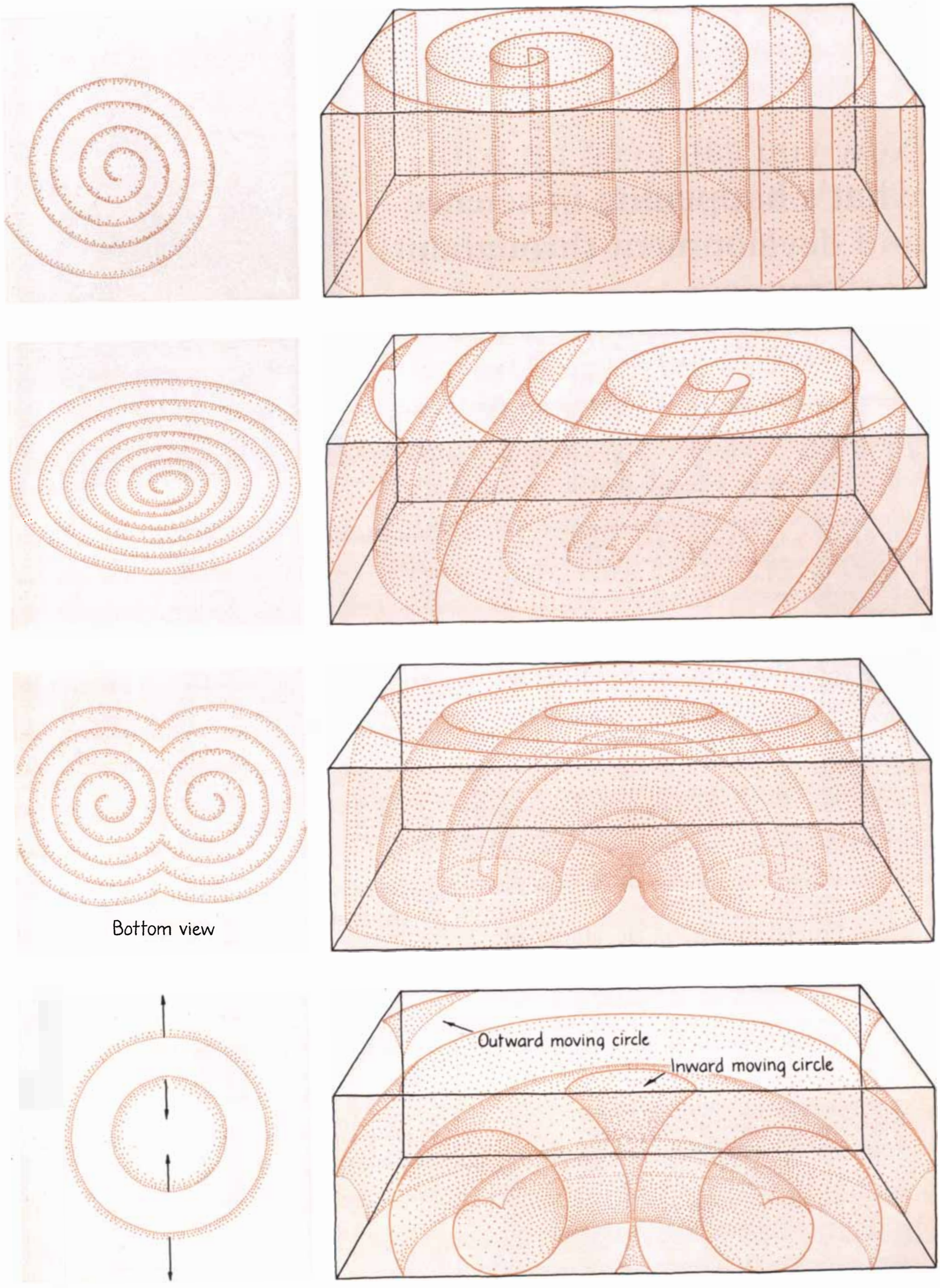
If you develop waves in a stack of filters (you need at least three per stack to get the elongated structures), you can peel them apart (as Winfree did) to see the vertical structures of the waves. In this way you can follow the scroll axis downward or unfold a *U*-shaped structure. Bubbles of carbon dioxide can be a bother if they separate the filters, which happens within about five minutes.

A stack of Millipore filters leads to a better understanding of the horizontal scroll ring. In such a ring the scroll axis is horizontal and closes back on itself to form a circle. Place a stack of five filters in the solution and touch the center of the top filter with a hot needle to initiate a cylindrical wave. After the wave has propagated outward through the top filter lower another stack of five filters onto the first stack, with the new stack in the initial red state of the solution. Soon a cylindrical wave will appear on the top of the new stack and the circle will split into two parts, one moving inward and one outward. When the new stack of filters is put into place, the scroll wave in the initial stack moves into the new stack and eventually emerges on the new top.

Winfree has another technique that he employs to prepare filters for a demonstration. Mix one-molar ammonium bromomalonate with an equal volume of .025-molar ferrous phenanthroline. Combine the mixture with an equal volume of four-molar ammonium bisulfate (46 grams per 100 milliliters of solution with water). Place a Millipore filter on the solution, lift it up to drain the surface liquid and then put it on a plastic surface to dry. Meanwhile, put a Whatman No. 1 paper filter in .33-molar sodium bromate (five grams per 100 milliliters of solution with water) and then dry it. When you want to demonstrate the scroll waves, wet the filter with pure water, put it on a clean surface (such as a plastic dish) and cover it with the Millipore filter. After waiting a few minutes you will find blue dots on the top filter; eventually you will see blue waves propagating through the filter at a rate of a few millimeters per minute.

A 12-square-inch sample of a Millipore filter already impregnated with chemicals for generating waves can be bought for $4 from Winfree (Arthur T. Winfree, Institute for Natural Philosophy, 51 Knoll Crest Court, West Lafayette, Ind. 47906). Send a self-addressed label with your order. The filter need only be wetted to give a 30-minute display. The filter can be warmed, cooled, cut with scissors, touched with an iron object to block the waves, stimulated with a hot needle and stacked to make three-dimensional waves.

How can a uniformly mixed solution at constant temperature and pressure



Types of spiral waves

oscillate in color, in concentration of molecular species, in conductivity or in anything else? The question has puzzled chemists since the 19th century. The oscillations are still debated, and a few chemists continue to argue that oscillations in such circumstances are impossible. In scientific terms the question is: Can oscillations occur in a homogeneous closed system at constant temperature and pressure? It is well accepted that oscillations can occur in mediums having inhomogeneous constituents and gradients of density in the molecular species undergoing diffusion. It is also well accepted that oscillations can arise if the system is open to the addition of more materials. In the bulk oscillators, however, oscillations are apparently turning up in a system that is homogeneous and closed.

In unraveling the puzzle some investigators concentrate on the Gibbs free energy of the system because any spontaneous reaction in a closed homogeneous system at constant temperature and pressure must decrease the Gibbs free energy of the system. The Gibbs free energy is an indicator of the spontaneity of a reaction (the tendency of the reaction to proceed on its own with no external activation). In any spontaneous reaction the system tends to reduce its energy (measured by enthalpy) and to increase its disorder (measured by entropy). It is not inevitable that both of these ends will be reached. For example, some reactions may be able to proceed spontaneously even though their energy (enthalpy) increases; the reason is that the disorder (entropy) increases more. Other reactions may be spontaneous for the converse reason. Hence neither enthalpy nor entropy alone is a good indicator of spontaneity. What one employs is a combination of the two called the Gibbs free energy. Any reaction that will reduce the Gibbs free energy is one that can proceed spontaneously.

At first the oscillating reactions I have described appear to violate the universally accepted rule that any spontaneous reaction must move steadily toward its final equilibrium state while steadily losing its Gibbs free energy. Actually oscillations seem to violate the second law of thermodynamics, which prohibits oscillations around the equilibrium point. The paradox is resolved rather easily, however, by noting that no thermodynamical rule prevents oscillations when the reactants are far from equilibrium and as the net reactions decrease the Gibbs free energy. For example, in Winfree's mixture the net chemical reactions consume very little of the bromate and malonic acid in each cycle of oscillation, and the cycles began with the reactants far from their final equilibrium values. In the steady drive of the net reactions toward equilibrium, the concentrations of the minor constituents bromide and phenanthroline can vary considerably, although they remain small with respect to the concentrations of bromate and malonic acid. Eventually the bromate and the malonic acid are sufficiently consumed in the net reactions so that the entire process stops and all the concentrations are in equilibrium. Thus the solution can oscillate in the concentration of its species (with a color change in the present case) as long as the principal reactants are plentiful and far from the final equilibrium concentrations. Once the equilibrium state is reached the oscillations cease. The second law of thermodynamics is not violated.

There must be other requirements for a solution to be able to oscillate or a chemistry class would be far more colorful. Indeed, the reactions taking place must be complex, and at least one reaction must catalyze itself (be autocatalytic) and be coupled to the other reactions in some way such as a dependence on the concentrations in the other reactions or by an inhibition or activation of the reactions. This coupling to other reactions is commonly called feedback, and most oscillating systems are characterized by it.

In the Winfree mixture the coupling is by way of the bromide ion and the phenanthroline. The bromide reacts with the malonic acid, inhibits the bromate reaction with the ferrous phenanthroline and is a product of the reaction of bromomalonate and ferric phenanthroline. One autocatalytic reaction is a bromate reaction with malonic acid. In other examples of bulk oscillators, such as the iodine clock, I am not certain that the complex mechanisms responsible for the oscillations have yet been sorted out and understood in thermodynamical terms.

Many of the chemicals you will need are available from Fisher Scientific Company (1600 West Glenlake Avenue, P.O. Box 171, Itasca, Ill. 60143), including the sodium bromate at $15.75 per pound, the sodium bromide at $5.95 per pound, the phenanthroline ferrous sulfate at $3.35 per fluid ounce, the sodium bisulfate at $5.75 per pound, the cerium ammonium nitrate at $10.95 per pound, the potassium bromate at $11.40 per pound, the sodium iodide at $7.30 per quarter pound, the iodine crystals at $7.50 per quarter pound and the ammonium bisulfate at $10.65 per pound. (The prices are from the 1977 chemical list.) Fisher also sells Whatman filters at about 55 cents for 100. The malonic acid is sold by the J. T. Baker Chemical Company (222 Red School Lane, Phillipsburg, N.J. 08865) at $21.65 per 100 grams; the same company sells Triton X-100 at $9.75 per pint and ceric sulfate at $35.60 for 500 grams. Colloidal silicon dioxide is available from the Apache Chemical Co. (P.O. Box 126, Seward, Ill. 61077) at about $20 for 25 grams.

"This is the most clever geometry book I have ever seen!"

—Thomas E. Lepperd, Evanston Township High School, Evanston, Illinois

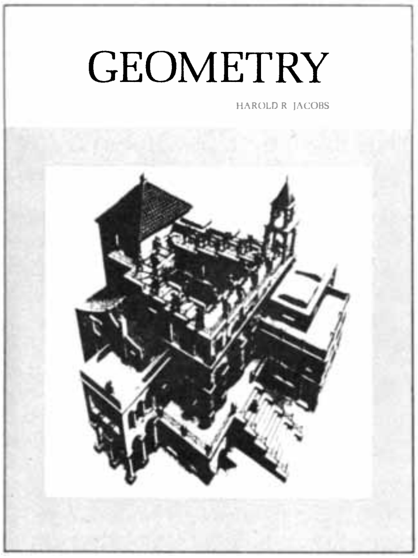
"Another typical Jacobs masterpiece!"

—Al Bigger, Auburn High School, Auburn, Alabama

"This book may single-handedly save geometry in the high schools."

—Patrick J. Boyle, Santa Rosa Junior College, California

Strong words about a strong book



GEOMETRY

1974, 701 pages, 1100 illus.,
ISBN 0-7167-0456-0, $9.00
Teacher's Guide $5.00
Transparency Masters $40.00
Test Masters $2.50

Also by Harold Jacobs
MATHEMATICS
A Human Endeavor
A Textbook for Those Who Think They Don't Like the Subject
Adopted by more than 2000 high schools
Foreword by Martin Gardner
1970, 529 pages, 634 illus.,
ISBN 0-7167-0439-0, $8.50
Teacher's Guide $5.00

"Einstein could have used Mr. Jacobs' book and still not have become so bogged down in the run-of-the-mill education that he could have missed the theory of relativity. Remember, Einstein said 'common sense is the collection of prejudices acquired by age 18'—no prejudices would be acquired by the geometry student and quite a few would be eradicated through this text."

—James J. Dolan, Chaminade Preparatory School, Canoga Park, California

"This is a book that belongs in every geometry class in the country. It has something to offer students at every level, combining wit and imagination with some very solid mathematical principles."

—Diana Brown, Wheeler High School, Marietta, Georgia

"This is a fascinating book with much more than the usual theorems. The cartoons and photographs make this a geometry book that is enjoyable reading. Most important, the problems are very good—they make one *want* to think and understand. . . . Congratulations, Freeman and Jacobs have done it again."

—Mark Saks, Community College of Philadelphia, Pennsylvania

The underlying theme of *Geometry* is a simple one: the principles of geometry may be more easily learned when they are shown to have immediate, practical applications in the student's world. *Geometry* is an exciting blend of humor, clear explanations, and an uncompromising attention to basic concepts—a blend that really does inspire students. We think that once you've examined this book, you'll have some strong words of praise too.

W. H. FREEMAN AND COMPANY
660 Market Street, San Francisco, California 94104
58 Kings Road, Reading, England RG1 3AA

BIBLIOGRAPHY

Readers interested in further explanation of the subjects covered by the articles in this issue may find the following lists of publications helpful.

MATHEMATICAL GAMES

THE CATEGORIES OF CHARLES PEIRCE. Eugene Freeman. Open Court, 1934.

THE PHENOMENOLOGY OF CHARLES S. PEIRCE: FROM THE DOCTRINE OF CATEGORIES TO PHANEROSCOPY. William L. Rosensohn. B. R. Grüner, 1974.

ALTERNATIVE AUTOMOBILE ENGINES

REPORT OF THE COMMITTEE ON MOTOR VEHICLE EMISSIONS. National Academy of Sciences, 1974.

AN EVALUATION OF ALTERNATIVE POWER SOURCES FOR LOW-EMISSION AUTOMOBILES. John W. Bjerklie, Elton J. Zairns, Charles W. Tobias and David Gordon Wilson. SAE No. 750929. Society of Automotive Engineers, 1975.

THE CHEMICAL DIFFERENTIATION OF NERVE CELLS

EVIDENCE FOR CHOLINERGIC SYNAPSES BETWEEN DISSOCIATED RAT SYMPATHETIC NEURONS IN CELL CULTURE. Paul H. O'Lague, Kunihiko Obata, Philippa Claude, Edward J. Furshpan and David D. Potter in *Proceedings of the National Academy of Sciences,* Vol. 71, No. 9, pages 3602–3606; September, 1974.

CHEMICAL TRANSMISSION BETWEEN RAT SYMPATHETIC NEURONS AND CARDIAC MYOCYTES DEVELOPING IN MICROCULTURES: EVIDENCE FOR CHOLINERGIC, ADRENERGIC, AND DUAL-FUNCTION NEURONS. Edward J. Furshpan, Peter R. MacLeish, Paul H. O'Lague and David D. Potter in *Proceedings of the National Academy of Sciences,* Vol. 73, No. 11, pages 4225–4229; November, 1976.

RAT SYMPATHETIC NEURONS AND CARDIAC MYOCYTES DEVELOPING IN MICROCULTURES: CORRELATION OF THE FINE STRUCTURE OF ENDINGS WITH NEUROTRANSMITTER FUNCTION IN SINGLE NEURONS. Story C. Landis in *Proceedings of the National Academy of Sciences,* Vol. 73, No. 11, pages 4220–4224; November, 1976.

NEUROTRANSMITTER SYNTHESIS AND UPTAKE BY ISOLATED SYMPATHETIC NEURONS IN MICROCULTURES. Louis F. Reichardt and Paul H. Patterson in *Nature,* Vol. 270, No. 5633, pages 147–151; November 10, 1977.

DETERMINATION OF TRANSMITTER FUNCTION BY NEURONAL ACTIVITY. Patricia A. Walicke, Robert B. Campenot and Paul H. Patterson in *Proceedings of the National Academy of Sciences,* Vol. 74, No. 12, pages 5767–5771; December, 1977.

NATURE AND NURTURE IN DEVELOPMENT OF THE AUTONOMIC NEURON. Richard P. Bunge, Mary Johnson and C. David Ross in *Science,* Vol. 199, No. 4336, pages 1409–1416; March 31, 1978.

THE HIGHEST-ENERGY COSMIC RAYS

COSMIC-RAY SHOWERS. Kenneth Greisen in *Annual Review of Nuclear Science: Vol. 10,* edited by Emilio Segrè, Gerhart Friedlander and Walter E. Meyerhof. Annual Reviews, Inc., 1960.

COSMIC RAYS. Bruno Rossi. McGraw-Hill Book Company, 1964.

COSMIC RAYS AT GROUND LEVEL. Edited by A. W. Wolfendale. The Institute of Physics, London, 1973.

NIGHTTIME IMAGES OF THE EARTH FROM SPACE

MISSION TO EARTH: LANDSAT VIEWS THE WORLD. Nicholas M. Short, Paul D. Lowman, Jr., Stanley C. Freden and William A. Finch. NASA SP-360. U.S. Government Printing Office, 1976.

LEARNING AND MEMORY IN BEES

THE DANCE LANGUAGE AND ORIENTATION OF BEES. Karl von Frisch. Harvard University Press, 1967.

DAS GEDÄCHTNIS DER HONIGBIENE FÜR SPEKTRALFARBEN, I: KURZZEITIGES UND LANGZEITIGES BEHALTEN. Randolf Menzel in *Zeitschrift für vergleichenbe Physiologie,* Vol. 63, pages 290–309; 1969.

THE DYNAMICS OF LEARNING IN THE HONEYBEE. Jochen Erber in *Journal of Comparative Physiology,* Vol. 99, No. 3, pages 231–255; 1975.

NEURAL MECHANISMS OF LEARNING AND MEMORY. Edited by Mark R. Rosenzweig and Edward L. Bennett. The MIT Press, 1976.

SYNTHETIC-MEMBRANE TECHNOLOGY

DESALINATION BY REVERSE OSMOSIS. Edited by Ulrich Merten. The MIT Press, 1966.

MEMBRANE PROCESSES IN INDUSTRY AND BIOMEDICINE. Edited by Milan Bier. Plenum Press, 1971.

POLYMERIC MEMBRANES. R. E. Kesting. McGraw-Hill Book Company, 1971.

INDUSTRIAL PROCESSING WITH MEMBRANES. Edited by Robert E. Lacey and Sidney Loeb. Wiley-Interscience, 1972.

REVERSE OSMOSIS MEMBRANE RESEARCH. Edited by H. K. Lonsdale and H. E. Podall. Plenum Press, 1972.

THE KIWI

PHYLOGENY AND EVOLUTION OF THE RATITE BIRDS. Joel Cracraft in *Ibis: A Quarterly Journal of Ornithology,* Vol. 116, pages 494–521; 1974.

THE AVIAN EGG: INCUBATION TIME AND WATER LOSS. Hermann Rahn and Amos Ar in *Condor,* Vol. 76, No. 2, pages 147–152; Summer, 1974.

THE KIWI. Brian Reid and Gordon Williams in *Biogeography and Ecology in New Zealand,* edited by G. Kuschel. Dr. W. Junk B. V. Publishers, The Hague, 1975.

INCUBATION TEMPERATURES OF THE NORTH ISLAND BROWN KIWI. Barry Rowe in *Notornis,* in press.

COMPUTER POKER

THEORY OF GAMES AND ECONOMIC BEHAVIOR. John von Neuman and Oskar Morgenstern. Princeton University Press, 1953.

GAMES AND DECISIONS. Robert Duncan Luce and Howard Raiffa. John Wiley & Sons, Inc., 1957.

GENERALIZATION LEARNING TECHNIQUES FOR AUTOMATING THE LEARNING OF HEURISTICS. D. A. Waterman in *Artificial Intelligence,* Vol. 1, pages 121–170; Spring, 1970.

STUDIES ON DECISION MAKING USING THE GAME OF POKER. Nicholas V. Findler, Heinz Klein, William Gould, Alexander Kowal and John Menig in *Information Processing 71. Proceedings of IFIP Congress 71, Vol. 2: Applications,* edited by C. V. Freiman. North-Holland Publishing Company, 1972.

STUDIES IN MACHINE COGNITION USING THE GAME OF POKER. Nicholas V. Findler in *Communications of the ACM,* Vol. 20, pages 230–245; 1977.

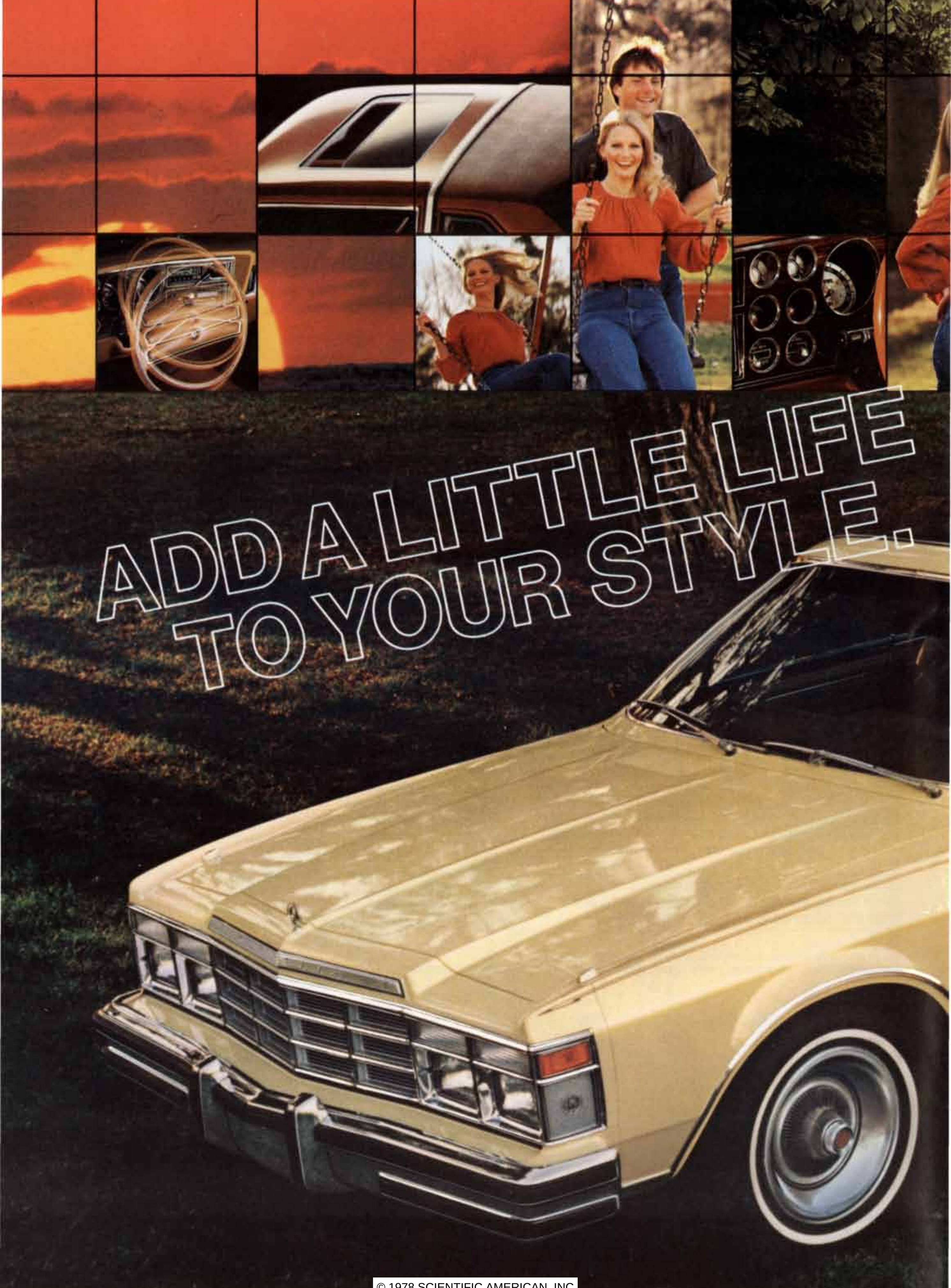
THE AMATEUR SCIENTIST

THE COLOR BLIND TRAFFIC LIGHT. John F. Lefelhocz in *Journal of Chemical Education,* Vol. 49, No. 5, pages 312–314; May, 1972.

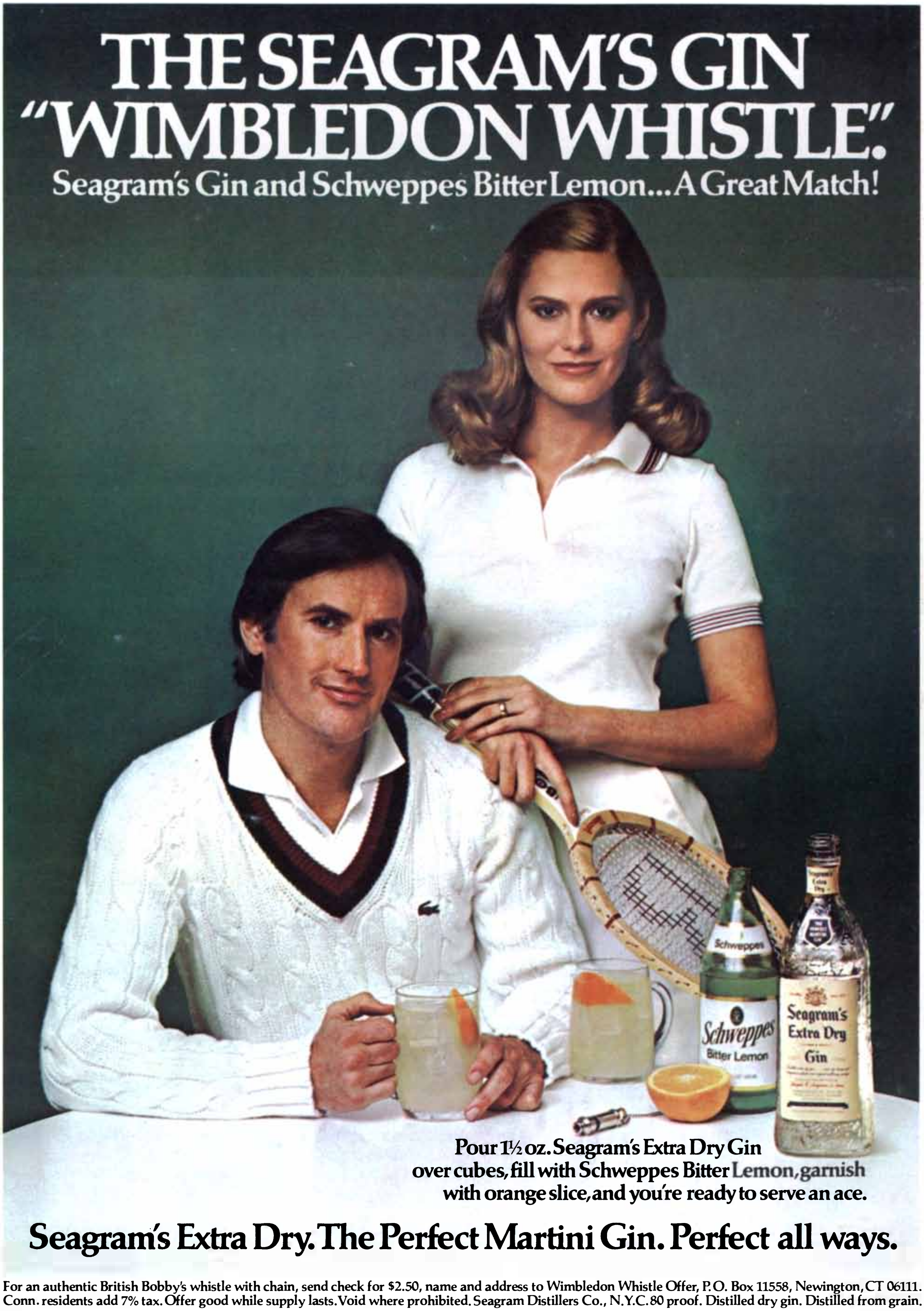
OSCILLATING CHEMICAL REACTIONS IN HOMOGENEOUS PHASE. Hans Degn in *Journal of Chemical Education,* Vol. 49, No. 5, pages 302–307; May, 1972.

A REACTION PERIODIC IN TIME AND SPACE. Richard J. Field in *Journal of Chemical Education,* Vol. 49, No. 5, pages 308–311; May, 1972.

AN OSCILLATING IODINE CLOCK. Thomas S. Briggs and Warren C. Rauscher in *Journal of Chemical Education,* Vol. 50, No. 7, pages 496; July, 1973.


ADD A LITTLE LIFE
TO YOUR STYLE.


Your style is like no other . . . and no other car brings more life to it than the new-size Chrysler LeBaron.
25 MPG HWY / 17* MPG CITY
It fits in the size and price class of cars like Cutlass Supreme, comparably equipped.
But we gave it a style all its own. A young aggressive style that fashionably avoids looking like all the others.
And there are plenty of ways to make it fit your style with interesting options like soft genuine leather seating, T-bar roof, AM/FM stereo. The kinds of extras that can make LeBaron *your* car.
It all adds up to this. Chrysler LeBaron is an exciting new way to add life to your style. And style to your life.
*EPA estimates for 6-cylinder engine with manual transmission. Your actual mileage may differ, depending on your driving habits, the condition of your car and its optional equipment. Mileage lower and optional automatic transmission mandatory in California.
CHRYSLER LeBARON.
CHRYSLER
A PRODUCT OF CHRYSLER CORPORATION


THE SEAGRAM'S GIN
"WIMBLEDON WHISTLE."
Seagram's Gin and Schweppes Bitter Lemon...A Great Match!
Schweppes
Bitter Lemon
Seagram's
Extra Dry
Gin
Pour 1½ oz. Seagram's Extra Dry Gin
over cubes, fill with Schweppes Bitter Lemon, garnish
with orange slice, and you're ready to serve an ace.
Seagram's Extra Dry. The Perfect Martini Gin. Perfect all ways.
For an authentic British Bobby's whistle with chain, send check for $2.50, name and address to Wimbledon Whistle Offer, P.O. Box 11558, Newington, CT 06111.
Conn. residents add 7% tax. Offer good while supply lasts. Void where prohibited. Seagram Distillers Co., N.Y.C. 80 proof. Distilled dry gin. Distilled from grain.