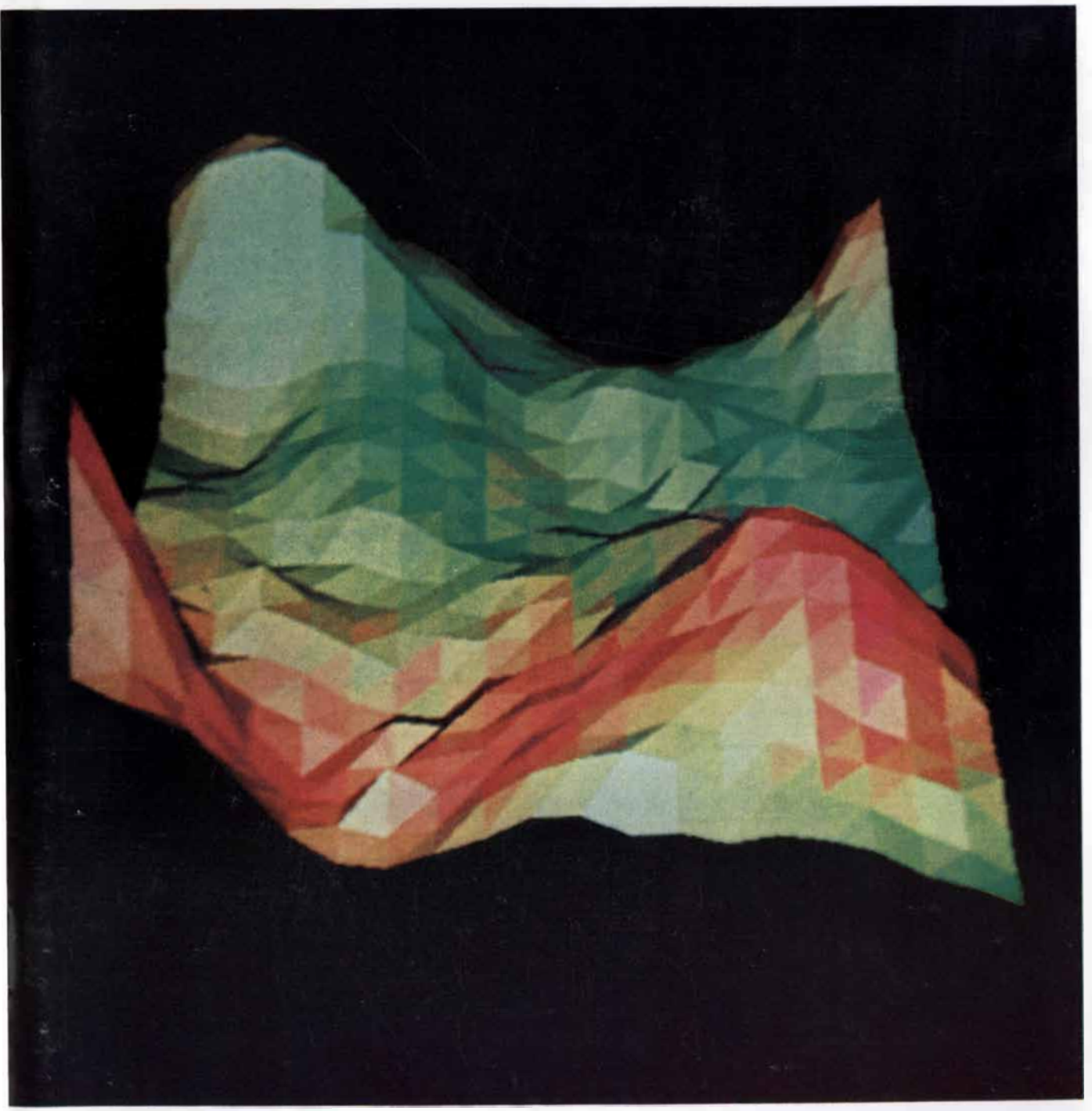
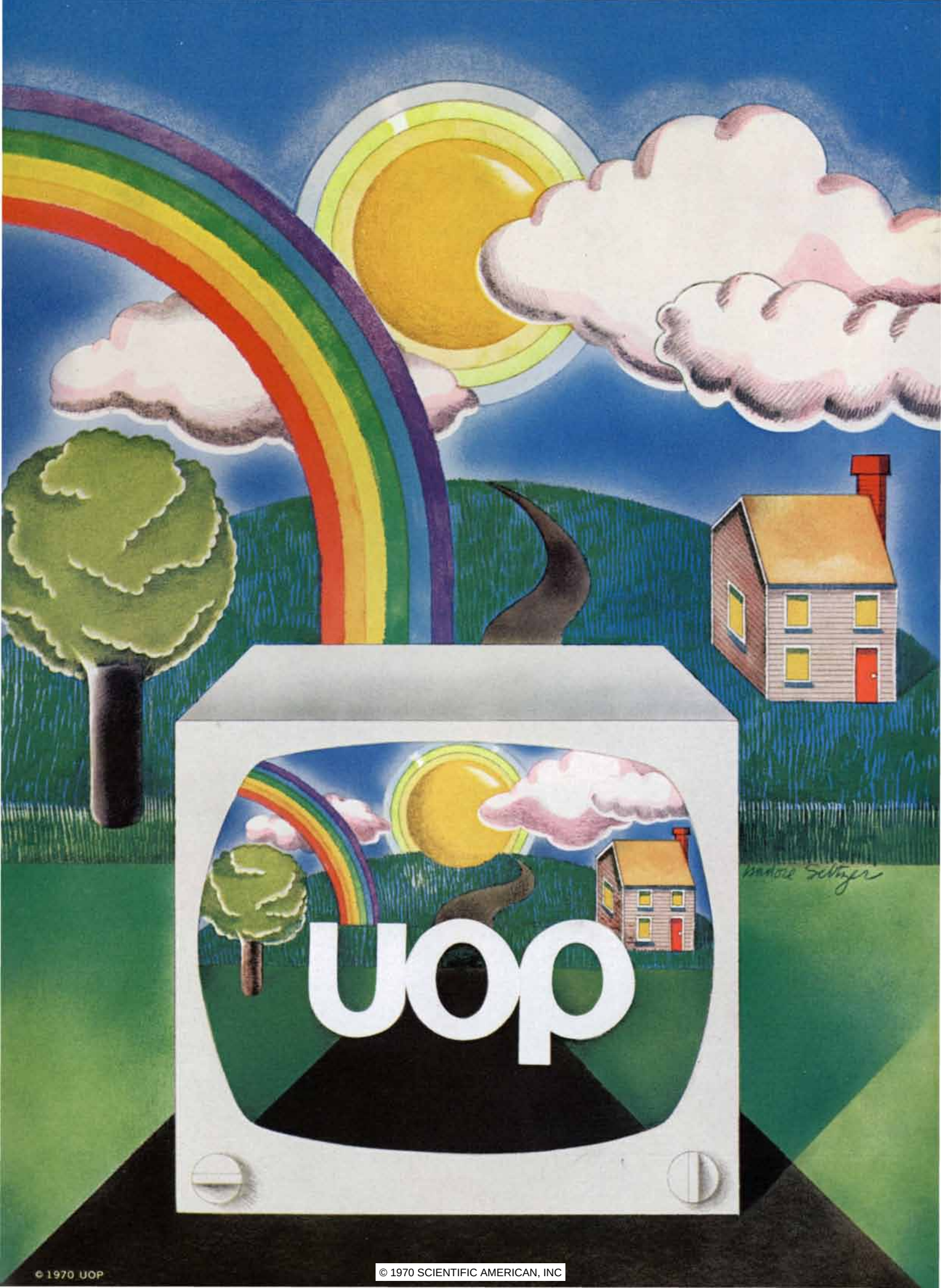
SCIENTIFIC AMERICAN



COMPUTER DISPLAYS

ONE DOLLAR

June 1970



UOP HAS A PIECE OF TV'S BIG, BRIGHT PICTURE.

Television's picture holds steady for the next year. And the next, and the next. Contributing importantly to the big picture is Universal Oil Products Company.

UOP is one of the prime suppliers to the industry with Norplex Brand copper-clad laminates that form the base of printed circuitry. If you're in the communications industry—TV, radio, radar, stereo, landing systems, whatever, you can profit, as many of the majors have, from UOP's broad experience in laminates.

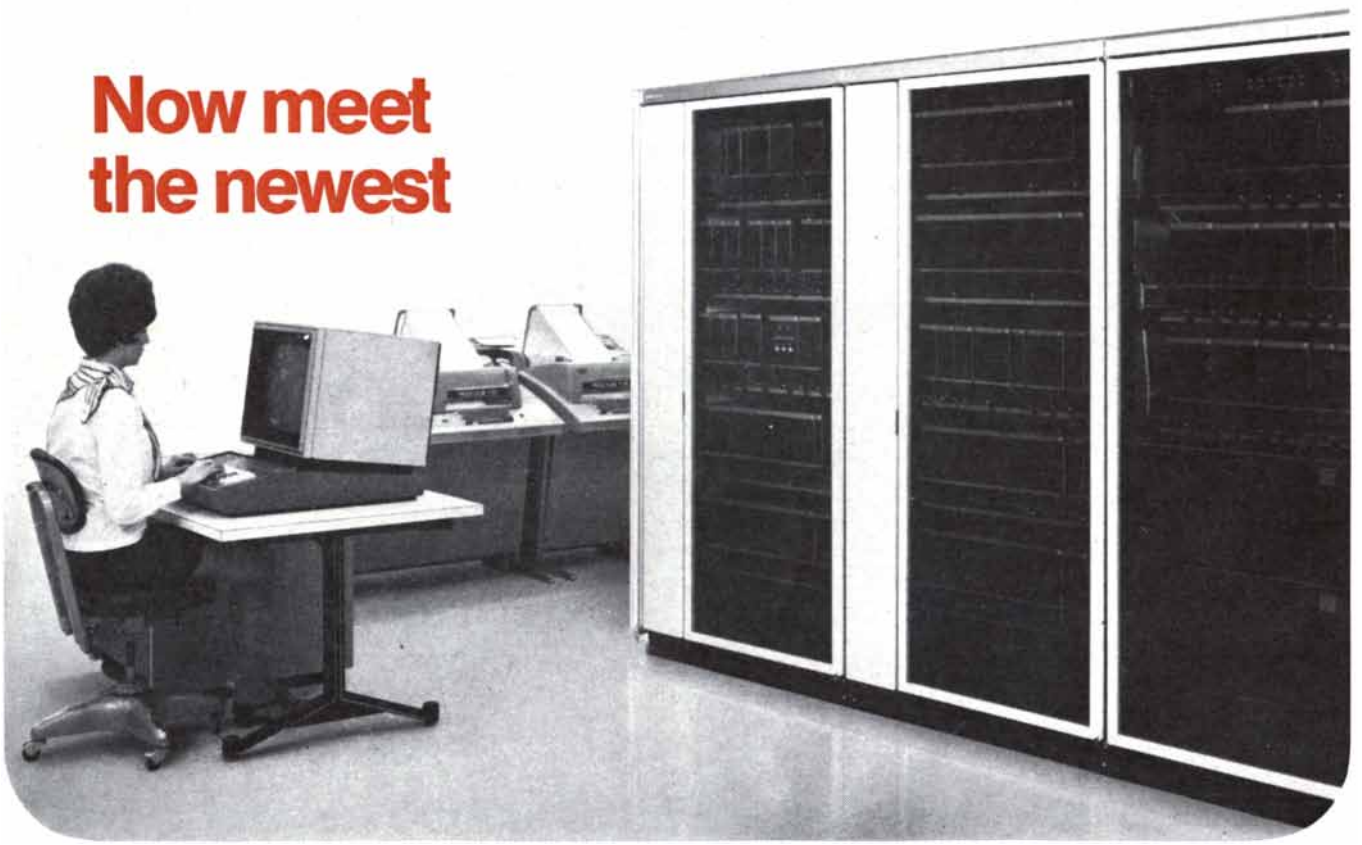
UOP serves these markets:
Environmental Control
Air and Ground Transportation
Petroleum and Petrochemical
Metal Fabrication
Fragrances, Flavors and Foods.

uop

30 ALGONQUIN ROAD, DES PLAINES, ILLINOIS 60016

Billions of messages ago
Collins introduced computer switching systems

Now meet the newest



Billions of messages ago Collins introduced computer-controlled message switching systems. Acceptance of these systems by airlines, railroads, and government agencies has given Collins the most extensive message switching experience in the world, both in hardware and software.

Collins now offers a new high-speed, high-volume, computer-controlled communication system to handle both message switching and data communication requirements. The system employs Collins C-System computer concepts.

This latest advance in communication control systems economically combines reliability and expandability through employment of advanced design, production, and testing techniques.

RELIABILITY. Continuous system up-time is achieved with advanced circuit technology, modular packaging, automatic switchover, and recovery without manual intervention or loss of messages.

EXPANDABILITY. System design permits easy expansion as requirements change. And the communication system hardware can serve as the initial configuration for a total management information and control system. A multiplexed input/output system provides high speed interconnection facility between all processors, storage devices and peripheral equipment, permitting the addition of processors, storage units, and terminal adapters to accommodate changes in system configuration. Low- and medium-speed circuit terminations interface the processor(s) via a lower speed, multiplexed channel.

ECONOMY. The C-System message switch with standard programming for various size systems is available now at competitive prices. New manufacturing techniques along with advanced computer technology have been combined to produce smaller system size and increased throughput.

Whether your application requires store-and-forward message switching, communication network control for a distributed data network, or on-line data retrieval, investigate the Collins C-System.

For more information on how Collins can solve your communication problems, write Dept. 300, Collins Radio Company, Dallas, Texas 75207. Phone: (214) 235-9511.

COLLINS

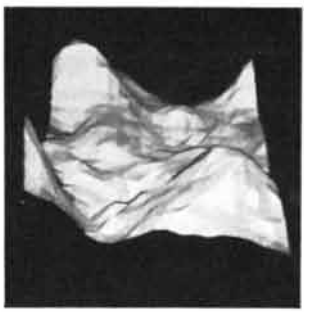
COMMUNICATION / COMPUTATION / CONTROL

Established 1845 **SCIENTIFIC AMERICAN** June 1970 Volume 222 Number 6

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: $10 PER YEAR.



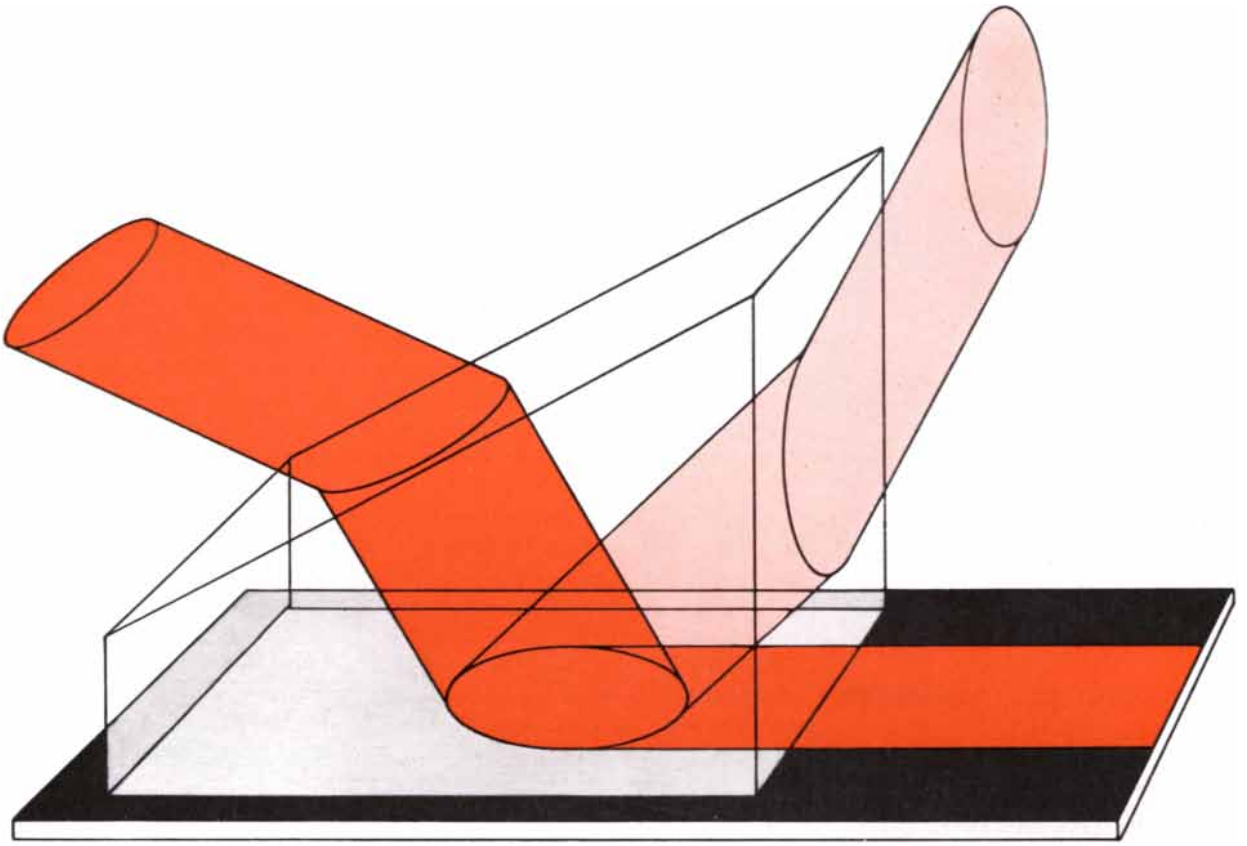
THE COVER

The picture on the cover, which resembles a colorful flying carpet, is actually a perspective view of a three-dimensional surface drawn by computer on the face of a cathode ray tube. Because the tube produces only black-and-white images, the color was achieved by having the computer display three separate pictures, one for each of the primary colors, and recording them through appropriate filters on a single sheet of color film. The height of the surface corresponds to the strength of the earth's magnetic field in a 25-square-mile area in southwestern Montana. Red denotes regions where the magnetic and gravimetric anomalies are both positive; green denotes regions where the anomalies are both negative. A new copper mine is located in the lower right corner of the map. The illustration was prepared by Robert B. Smith of the geophysics department of the University of Utah with the assistance of John Warnock. The making of such pictures is described this month by Ivan E. Sutherland in his article "Computer Displays" (*page 56*).

THE ILLUSTRATIONS

Cover photograph by Robert B. Smith and John Warnock, University of Utah

Page	Source
18	Survey Research Center, University of Michigan
19–25	Graphic Presentation Services, Inc.
27	Hale Observatories
28–29	Bunji Tagawa
30–31	Hale Observatories
32–34	Bunji Tagawa
35	Halton C. Arp, Hale Observatories (*left*); Kitt Peak National Observatory (*right*)
37	Jerome Kuhl
38	L. Kemp and A. F. Howatson (*top*), Jack Griffiths (*bottom*)
39–44	Jerome Kuhl
56	Gary Watkins
58–59	Paul Weller (*top*); Steve Allen, Fairchild Semiconductor (*bottom*)
60–65	Dan Todd
66–67	Bell Telephone Laboratories (*top*), Dan Todd (*bottom*)
68–69	Massachusetts Institute of Technology
70	Paul Weller (*top*), Nelson Wadsworth (*bottom*)
71–73	Paul Weller
74–79	Dan Todd
81	John Warnock (*top*), Bell Telephone Laboratories (*bottom*)
83	Carl Gans
84–86	Tom Prentiss
88–93	Carl Gans
94–95	Tom Prentiss (*top*), Carl Gans (*bottom*)
96	Carl Gans
100	Austin S. Post
102–103	Jim Egleson
104	O. Süssli-Jenny, Thalwil-Zurich (*top*); Swiss National Tourist Office (*bottom*)
105	W. O. Field
106–107	Jim Egleson
108	Bradford Washburn
110	W. O. Field
112–115	Newfoundland Memorial Museum
116–117	Eric Mose
118	Newfoundland Memorial Museum (*top*), Lee Boltin (*bottom*)
119	Newfoundland Memorial Museum
120	Eric Mose
121	Newfoundland Memorial Museum
123	Frank Hawking
124–127	Lorelle A. Raboni
128	Lorelle A. Raboni (*top*), Frank Hawking (*bottom*)
129–131	Lorelle A. Raboni
132–136	Alan D. Iselin
141–145	Roger Hayward



Breaking through an interface

When we need the great bandwidth of optical communications, we'll also need devices and circuits in which light can flow much like electricity in wires of conventional circuits. Laser beams may be guided into circuits made of thin films of semiconductor. In these, the transmitted signals will be modulated, switched, converted in frequency, etc.

But one difficulty is getting light into—and out of—films which may be less than a micron thick. It isn't practical to direct a light beam through the edge of such a thin film. Also it's difficult to make the edge optically perfect.

Now, scientist P. K. Tien of Bell Laboratories has found a way to feed light into the film through its surface.

The method is illustrated above. A prism is held close (within a wavelength) to the semiconductor. A laser beam through the prism refracts downward and strikes the base of the prism at some angle. The phase velocity of the light now has two components: one perpendicular to the film, one parallel to it. At the proper angle, the parallel component equals the phase velocity of light in the semiconductor and light energy is coupled into the film. (If a similar prism, facing the opposite way, is placed against the film in the path of the light beam, the light will be coupled back out of the film.)

In an experiment with red helium-neon laser light (6328-A wavelength), Tien has coupled 50% of the light's energy, in a single mode, into sputtered zinc-oxide or zinc-sulfide films on a glass substrate. From theoretical calculations, 80% should be possible.

By choice of angle at which the light beam strikes the "coupling face" of the prism, Tien can excite any of eight discrete propagation modes in the semiconductor film. Each can act as a separate communications channel. The number of possible propagation modes increases with film thickness.

This discovery may someday be a key to highly miniaturized electro-optic devices for processing information.

From the Research and Development Unit of the Bell System:

Bell Labs

LETTERS

Sirs:

Charles Darwin, I was pleased to find, is mentioned in regard to his barnacle *Balanus nubilis* ["How Is Muscle Turned On and Off?" by Graham Hoyle; SCIENTIFIC AMERICAN, April]. Dr. Hoyle remarks: "I venture to suggest that Darwin intended to do justice to this magnificent invertebrate by naming it *B. nobilis*, and that the typesetter, working from Darwin's scrambled handwriting, misread the name."

Darwin's son Francis states (in *The Autobiography of Charles Darwin and Selected Letters*) that Darwin wrote a rough copy hastily. "[This] was then reconsidered and a fair copy was made.—The fair copy was then corrected, and recopied before being sent to the printers. The copying was done by Mr. E. Norman, who began this work many years ago when village schoolmaster at Down." He adds that Charles Darwin's wife looked over the printer's proofs of the *Origin*, as did a daughter and Francis himself. In other words, the typesetter saw Mr. Norman's handwriting, and with the family that makes four people at least to misread a word.

Francis Darwin continues about a description of a larval cirripede "with six pairs of beautifully constructed natatory legs, a pair of magnificent compound eyes and extremely complex antennae [*Origin*, page 440]. We used to laugh at him for this sentence, which we compared to an advertisement." So Darwin's terms "beautifully," "magnificent" and "extremely complex" lead up to the spirit of Hoyle's phrase, "magnificent invertebrate," and confirm his suggestion that Darwin probably meant *B. nobilis*.

FLORENCE ELY DAY

Cambridge, Mass.

Sirs:

In Arnold C. Wahl's excellent article "Chemistry by Computer" [SCIENTIFIC AMERICAN, April] it is implied that the Hartree "self-consistent field" method of determining approximately the electronic structure of atoms was the joint work of William and Douglas R. Hartree. I think it is important to point out that the method of the self-consistent field, first published in 1928, and the associated numerical techniques for the solution of Schrödinger's equation (assuming radial symmetry) were in fact the work of Douglas R. Hartree alone. Douglas Hartree's father, William Hartree, undertook a considerable amount of the numerical work involved in the early electronic structure determinations (for the lighter atoms), but this was carried out under the direction of his son.

It may be of interest to note that probably the first application of an analogue computer to a self-consistent field determination was in 1933, when the late Douglas Hartree and I (his first Ph.D. student), using a differential analyzer of the Vannevar Bush type, which we built largely from Meccano components (at a cost of some 30 pounds), obtained a good approximation to the self-consistent field of the normal atom of hydrogen. This was the initial step (essentially to test out whether or not such a crude computer was adequate for such computations) in my M.Sc. thesis project, which was a determination of an approximation to the self-consistent field of the normal atom of chromium....

ARTHUR PORTER

University of Western Ontario
London, Ontario

Scientific American, June, 1970; Vol. 222, No. 6. Published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Gerard Piel, president; Dennis Flanagan, vice-president; Donald H. Miller, Jr., vice-president and 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 Allan Wittman, Advertising Manager, 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 20 cents.

Change of address (or other subscription correspondence) should be addressed to Circulation Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Please notify us four weeks in advance of change. If possible, kindly furnish an address imprint from a recent issue. Send both old and new addresses or complete and mail Post Office Form 357B.

Please enter change of address ☐
Please enter subscription ☐
☐ 1 year, $10 ☐ 2 years, $18 ☐ 3 years, $25

NAME

NEW ADDRESS

OLD ADDRESS

LAND OF THE BLABBERMOUTHS.

Blab blab blab. The communications flow in the U.S.
is turning into a gusher. Can technology keep up? Martin Marietta
communications people are studying unusual new digital
techniques to expand the volume of conventional circuits. We're also
conducting transmission experiments in the millimeter
wave region. It could give some much-needed tongue room in the
crowded electromagnetic blab spectrum.
Martin Marietta Aerospace Group. Headquarters:
Friendship International Airport, Maryland.

MARTIN MARIETTA

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

JUNE, 1920: "Perhaps the very experiment which started Einstein at work on his now famous 'theory of relativity' will be the means of demonstrating to the scientific world the correctness or incorrectness of his conclusions. The classical experiment performed by Michelson and Morley in 1887 and later repeated with better apparatus by Morley and Miller to determine 'ethereal drift,' upon the assumption that the ether of space is a stationary sea through which our earth moves, is the experiment referred to. New interest is aroused by the statement made recently by Prof. Dayton C. Miller, one of the original experimenters, that the work has never been carried far enough to be positive in its results. He has proposed moving the apparatus, which is still intact, to the top of Mount Wilson, near the famous Lick Observatory, and repeating the experiment there with the cooperation of the Lick scientists. Such an event would undoubtedly attract the attention of the entire scientific world, for although many scientists now recognize the probability that Einstein's theory will displace Newton's law of gravitation and indeed revolutionize our whole fundamental conception of the structure of matter and space, there are still many doubters."

"Many are the readjustments which the coming of prohibition has brought to the industrial and commercial world. Perhaps the most serious problem, as well as the one which has received most attention in the press, has been that of the ultimate fate of the breweries, with their buildings and machine equipment representing an investment of $700,000,000 or more. But the manufacturers were not the only ones who had to find a new field of operations. The plight of a middle-aged person with flat feet and soft hands who knows how to do nothing except serve drinks and who is barred by legislative enactment from this occupation furnishes ample material for humor or tragedy, according to the resourcefulness of the individual. Even more hopelessly involved than the bartender, however, seem to be those who have been engaged in the growing of the various crops from which have been manufactured the beverages that are now taboo. The raw materials employed in the making of distilled, malt and vinous liquors in 1919 were valued respectively at 37, 96½ and 6½ millions of dollars. Obviously these raw materials are almost entirely agricultural products such as hops, barley, corn, rye, grapes, etc. It might well seem at first glance hopeless to seek new markets for this bulk of croppage."

"The recent heavy fighting between the Polish forces and the Russian Bolsheviki has apparently involved the Kosciusko Squadron, which forms a unit of the Polish air service. The Kosciusko Squadron is composed of volunteer American fliers, commanded by Major Cedric E. Fauntleroy. The planes are of the latest single-seater fighter type, equipped with machine guns firing through the propellor sweep. The Kosciusko Squadron has been cited in the recent fighting, and at least one of its members has been shot down behind the Bolsheviki lines."

"In a recent popular article on the still undiscovered trans-Neptunian planet Dr. Eric Doolittle states that according to one of the latest calculations, based on disturbances in the movement of Uranus, this hypothetical planet lies at a distance of nearly 4,000,000,000 miles from the sun and occupies 283 years in completing its circuit around that luminary. Its mass, according to this calculation, is about six times that of the earth, making it much smaller than any of the other planets that are remote from the earth, and correspondingly difficult to observe. It is believed that, when discovered, its brightness will not exceed the 12th or 13th magnitude."

SCIENTIFIC AMERICAN

JUNE, 1870: "We are in receipt of the following letter from a firm of engineers and architects in Moscow, Russia: 'As we are considered by the postal authorities in Moscow the oldest subscribers here to your very valuable paper, we request your kind assistance in the following matter, and trust that some of your correspondents may favor us with their advice. In the cities and towns of Russia the street paving is entirely composed of pebbles, varying in size from three inches to nine inches in diameter; it naturally follows that at the breaking up of the winter season the streets are entirely out of repair, so that for a few weeks it is utterly impossible to carry more than half loads, or to travel with any degree of safety. Russian law requires that every householder shall patch his own frontage; therefore the cheaper that is done the better it is! At present repair consists of a stone in one hand for the hammer, the other stones laid being tapped affectionately on the head by the child, boy or man; the traffic has then to consolidate the road, and that is the system. Are there any cheap inventions, patented or otherwise, that would with one or two man power give the householder a chance of hammering the roads properly? If so, we should be most happy to entertain any proposal if the invention be of a real and practical service.' "

"As we go to press the telegraph brings us news of the death of the great novelist Charles Dickens, than whom no writer in his time has become more widely known and admired."

"The latest book of Alfred Russel Wallace, a series of essays entitled *Contributions to the Theory of Natural Selection*, merits an elaborate review, and we have kept it upon our table for some weeks, hoping to gain leisure to give it a thorough reading and to place in the form of a review its salient points before our readers. We have not yet found time to do this, but we can assure all lovers of natural history, and all who delight in original investigation of any kind, that no time can be better improved than by the perusal of this work. The author candidly and modestly sets forth his claims to the origination of the evolutionary theory of the origin of species, while acknowledging that few men but Mr. Darwin—and perhaps none now living—could have accomplished the great work from which that great philosopher derives his fame. The book is therefore written in no spirit of rivalry or detraction but with an earnest desire to advance true science."

"Mr. F. G. du Pont of Wilmington, Del., has sent us the result of his recent analysis of the celebrated Lancaster beer of Pennsylvania. In his covering letter he states: 'I believe this is the first analysis ever made of this beer. I think it may prove interesting to some of your many readers. I have been much interested in the Notes on Science in your valuable paper.' "

CHEVY'S NEW LITTLE CAR: IT WON'T STAND ALONE.

Our forthcoming mini will come forth this summer with more going for it than most cars twice its size.

It will have General Motors going for it.

It will have Chevrolet going for it.

It will have over 6,300 Chevrolet dealers going for it.

It will have the experience, resources, talent and track record of the world's largest, most successful car-maker going for it.

Even when we think small, we think big.

We couldn't build just one.

Unlike the other little cars that have been coming out lately, ours is going to be a complete line. And they'll all come out at once:

(1) A sporty little hatchback coupe.

(2) A sturdy little 2-door sedan.

(3) A slick little station wagon.

(4) A solid little panel truck.

All four will be exceptionally strong little vehicles with extraordinary stability, braking, acceleration, handling and comfort.

And they'll be right there with the best of them when it comes to gasoline mileage.

They'll stand apart, though.

Rest assured, the car won't be what you could call radical or far out. But it's going to have some features quite uncommon in cars this size.

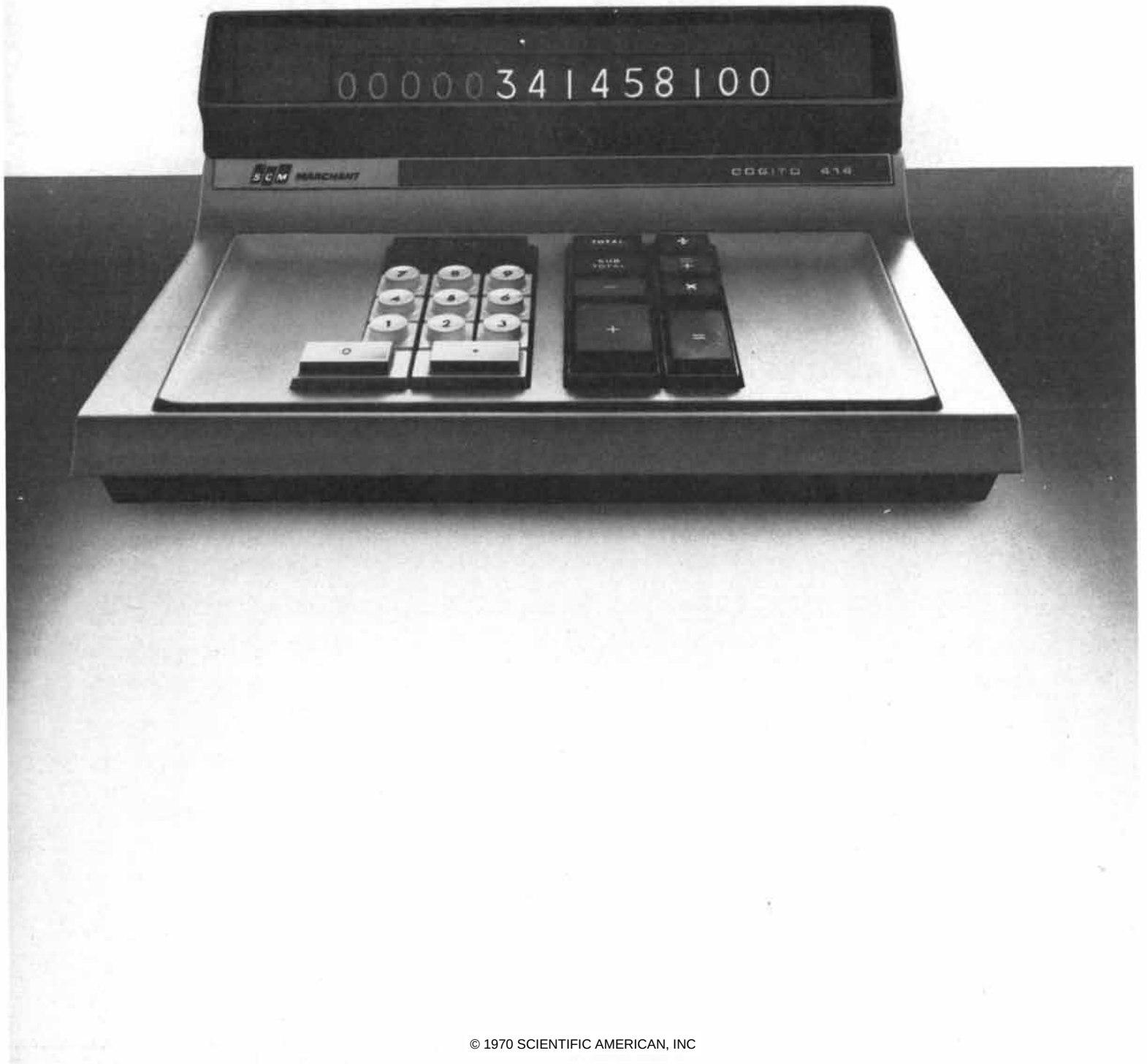
A new kind of little-car engine, for instance, that doesn't race and make a lot of noise when you want to make some time, or pass a truck, or merge onto the freeway.

Also: a hood that opens backwards, guard rails inside the doors, disc front brakes, and a finish you won't have to wax.

To be continued.

You'll see.
CHEVROLET

GM
MARK OF EXCELLENCE



There is no electronic calculator in the world to match the new Marchant electronic.

It's the only one to use a remarkable new technology: Super MOS.

When you buy an electronic calculator, there are dozens to choose from. How do you know which one's best? We think the answer is, simply pick the most advanced one you can find.

And that happens to be the Marchant® Model 414. It may look like other electronic calculators, but the resemblance ends right there. Because the 414 uses a totally new kind of technology in both its memory and logic functions.

It has a space age brain.

Our technology utilizes sub-miniaturization akin to the space age. Hundreds of transistors (which by themselves are marvels of electronic wizardry) are microscopically incorporated on a single chip so small 8 of them can sit comfortably on the face of a dime.

This tiny speck of material is the very brain of our calculator. It is a custom designed Metallic Oxide Semi-Conductor (we call it Super MOS for short). And, just to put it in perspective for you, it helps to know that each Super MOS chip contains 40 square feet of complex electrical diagrams.

Think of it! A blueprint 5 by 8 feet is reduced to something you can hardly see. And it works!

8 Super MOS chips.

Our 414 has 8 Super MOS chips. Each chip handles the work load previously assigned to 20 integrated circuits. We are the first manufacturer of calculators to use Super MOS chips. And to date, no other maker has achieved this extraordinary sophistication in design.

With Super MOS to do the work, the 414 can add, subtract, multiply, divide, place decimals automatically, and recall the last number displayed just as if you had entered it. And with fewer components and electrical connections, our 414 has fewer parts to wear out, fewer things to go wrong.

What does it all mean?

When all is said and done, it really comes down to just what we said up at the top. If you're going to buy a brand-new electronic calculator, you really ought to get the most singularly advanced one you can find.

The Marchant Model 414.

It may look like other calculators, but its brain thinks just a little bit ahead of all the rest.

SCM SMITH-CORONA MARCHANT
DIVISION OF SCM CORPORATION

Marchant. The First.

THE AUTHORS

PHILIP E. CONVERSE and HOWARD SCHUMAN (" 'Silent Majorities' and the Vietnam War") are professors of sociology at the University of Michigan and are associated with the Survey Research Center there. Converse was graduated from Denison University in 1949 and, after study at the University of Iowa and the Sorbonne, received his Ph.D. in social psychology from the University of Michigan in 1958. He has been closely involved with the continuing sequence of biennial-election studies at the Survey Research Center and has also done considerable comparative work on opinion-formation and electoral processes in France. Schuman received his undergraduate degree at Antioch College and, in 1961, his Ph.D. from Harvard University. He then spent several years studying social-psychological aspects of economic development; during much of the time he was in Pakistan and India. His current interests are race relations, survey research methods and public opinion.

MARTIN J. REES and JOSEPH SILK ("The Origin of Galaxies") are respectively at the Institute of Theoretical Astronomy of the University of Cambridge and the Princeton University Observatory. Rees obtained his master's degree in mathematics and his Ph.D. in astrophysics from Cambridge. His interests include cosmology, diffuse matter in space and theoretical radio astronomy. Silk was a Cambridge undergraduate and obtained his Ph.D. at Harvard University. This summer he will take up an appointment in the department of astronomy of the University of California at Berkeley.

MARK PTASHNE and WALTER GILBERT ("Genetic Repressors") are members of the department of biochemistry and molecular biology at Harvard University; Ptashne is a lecturer in biochemistry and Gilbert is professor of biophysics. Ptashne was graduated from Reed College in 1961 and received his Ph.D. in molecular biology from Harvard in 1965. He was a junior fellow of the Society of Fellows at Harvard from 1965 to 1968. His interests include work in a program that sends teachers to Cuba for summertime teaching at the University of Havana; he taught there himself last summer. Gilbert was graduated from Harvard College in 1953. He received his doctorate, which was in mathematics, from the University of Cambridge in 1957. Gilbert was originally a theoretical physicist; he became interested in biology a decade ago.

IVAN E. SUTHERLAND ("Computer Displays") is president of the Evans & Sutherland Computer Corporation in Salt Lake City; the firm manufactures computer-display equipment. Before joining the firm Sutherland was associate professor of electrical engineering at Harvard University. He is continuing his research on computer graphics on a part-time basis at the University of Utah, where he is associate professor of electrical engineering. Sutherland writes: "I have found the transition from university professor to corporate executive a very educational one. I had not accurately predicted how much effort goes into management and how little into engineering. I look forward to the day when I can concentrate again, as I did at Harvard, on technical activities and leave the management to someone else." Sutherland received a bachelor's degree in electrical engineering at the Carnegie Institute of Technology in 1959, a master's degree at the California Institute of Technology in 1960 and a Ph.D. from the Massachusetts Institute of Technology in 1963.

CARL GANS ("How Snakes Move") is professor of biology at the State University of New York at Buffalo. He began his career as a mechanical engineer, having received a bachelor's degree in mechanical engineering at New York University in 1944 and a master's degree at Columbia University in 1950. From 1947 to 1955, except for a year when he was studying Brazilian snakes, he was involved in the design and service of combustion equipment for the Babcock and Wilcox Company. In 1957 he received his Ph.D. in biology from Harvard University. Gans writes: "I consider myself a functional anatomist and have spent much of the past 12 years working on the amphisbaenians, an odd group of subterranean reptiles whose burrowing mechanics and prey-locating mechanisms involve some fascinating biomechanical problems. Fortunately the animals are found in interesting places, thus providing an excuse for much travel, the latest being a trip to Africa, where we spent four months digging up the landscape between looks at nonburrowing animals."

GEORGE H. DENTON and STEPHEN C. PORTER ("Neoglaciation") are respectively associate professor of geological sciences at the University of Maine and associate professor of geological sciences and associate director of the Quaternary Research Center at the University of Washington. Denton received degrees in geology at Tufts University in 1961 and from Yale University in 1964 and 1965. As a member of the staff of the American Geographical Society he carried out research on glacial history in northern Scandinavia, Yukon Territory, southern Alaska and Antarctica. Porter received degrees in geology from Yale in 1955, 1958 and 1962. His research has centered mainly on alpine glaciation and geomorphology in the Brooks and Cascade ranges of Alaska, the Himalayas and the Hindu Kush of India and Pakistan.

JAMES A. TUCK ("An Archaic Indian Cemetery in Newfoundland") is assistant professor of anthropology at the Memorial University of Newfoundland. "I have worked throughout the northeast," he writes, "including New York, New England and the Atlantic Provinces. My first researches investigated the development of the Onondaga Iroquois since the 12th century. Subsequent researches have led progressively northward to Newfoundland, where we have investigated numerous Eskimo and Indian sites. Currently I am directing a project in northern Labrador, investigating Indian and Eskimo cultural development on the Labrador peninsula." Tuck was graduated from Syracuse University in 1962 with a degree in botany; in 1969 he obtained his Ph.D. in anthropology there.

FRANK HAWKING ("The Clock of the Malaria Parasite") is a member of the staff of the National Institute for Medical Research in London. "My first interest was classics," he writes, "but as these did not seem to offer a livelihood, I turned to science, particularly medicine." Born on a farm in Yorkshire and educated at the University of Oxford and St. Bartholomew's Hospital, he had no special interest in tropical countries until 1930, when an advertisement drew him to the Liverpool School of Tropical Medicine. His first personal experience of the rhythms of malaria began in 1930, when he contracted the disease in Nigeria. During World War II he was secretary of the British Malaria Committee. He has been at the National Institute for Medical Research since 1940, but at different times he has studied in or visited most of the tropical countries of the world.

One day we count chickens. The next day we make hay.

We're synergistic. How do you keep track of 3½ million active chickens? One of our customers uses our UNIVAC 9200 Computer. This management system not only helps him run his business but even figures the cost of raising a bird to within four decimal points.

And farmers all the world over use our New Holland bale wagons which can handle up to 2,500 bales a day. That's a lot of hay.

We do a lot of things at Sperry Rand. And we do each one better because we do all the rest.

SPERRY RAND™

Wait till you see what we do tomorrow.

NEW HOLLAND • REMINGTON • REMINGTON RAND • SPERRY • UNIVAC • VICKERS

Registered trademarks of Sperry Rand.

What is U.S. Steel doing to make

tomorrow's cars better?

This is Innovari II.

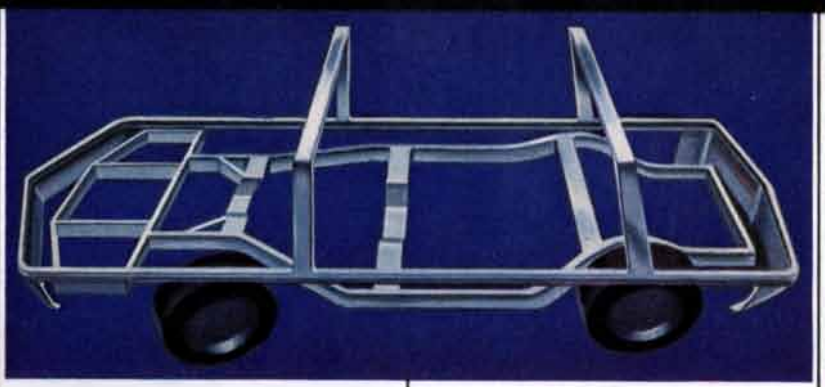
It's an idea car designed and built by U.S. Steel. It's one of many things we're doing to help put more durability, safety and convenience in your new car.

The "vinyl" roof is a good example.

It's embossed steel. Steel with an alligator pattern rolled in at the mill. After it's painted, this embossed steel looks just like vinyl. But with the durability of steel.

Then there's safety. Your safety.

Innovari's sun visor core is made of thin sheets of steel, curled into resilient fingers. On impact, they absorb the shock, and protect your head.



We also took a fresh look at ways to protect you in collisions, even in roll-overs.

Perimeter protection, including light but rugged steel roll bars front and rear, was our effective answer.

Innovari's new door hinge is for your convenience.

It's double-acting. This means the door opens several inches wider than it could with a conventional hinge. A little thing? Of course. But it's one of those tremendous trifles that could make your new car a lot more fun.

All in all, we've put 23 ideas to work in Innovari II. And not one of them is for sale.

But they're available for any auto manufacturer who wants to use them.

When tomorrow's cars roll off the line, we're not certain what they'll look like.

But we're doing a lot of things to help make them better.

USS is a registered trademark.

USS

We're involved.

This pesticide will start to self-destruct in seven days.

After destroying the bugs, it destroys itself.

In a week it begins breaking down into harmless parts.

Developed at Union Carbide, Sevin carbaryl insecticide gets rid of bugs without harming anything else. Like birds, fish, or people. (Sevin doesn't build up in living tissue.)

Which is why it's one of the most widely used bug killers in the world.

In America, it's the one used to combat the gypsy moth. A creature of the Northeast who thrives on oak and shade trees. But will take on an entire forest if he's in the mood.

The first year he feeds on a tree, he strips it bare. If he goes for seconds, the tree will likely die.

Because of Union Carbide, trees that are properly treated will continue to have leafy arms and be an inspiration to poets.

As for the gypsy moth.

His first meal is his last.

UNION CARBIDE

THE DISCOVERY COMPANY
270 Park Ave., New York, N.Y. 10017

Established 1845 SCIENTIFIC AMERICAN June 1970 Volume 222 Number 6

"Silent Majorities" and the Vietnam War

Is there a majority opinion on American involvement in the war in Vietnam? A study of public-opinion polls indicates that most Americans are against the war but that motivations are mixed

by Philip E. Converse and Howard Schuman

President Nixon has asserted that a "silent majority" supports his stand on the war in Vietnam. His critics respond that a majority of Americans want a quicker end to the war than the President offers. Each side seems able to cite opinion polls that appear to uphold its point of view. One is left wondering whether or not a national majority has in fact been identified and, if it has, what it favors.

At the Survey Research Center of the University of Michigan we have recently brought together a decade of the published results of polls on the Vietnam war. We have also subjected data from our own national surveys to a level of detailed analysis that goes well beyond the usual newspaper reports of public opinion. Although it would be reckless to claim that any simple or unequivocal picture emerges from this work, we can say that our findings do not present the kind of patchwork that defies coherent interpretation. At the very least they help to clarify what is accurate, as opposed to what is probably overdrawn, in either extreme of interpretation of the public mood on the war in Vietnam.

One fact is important to recognize at the outset. Although the standard national survey of some 1,500 interviews may seem far too small to represent a nation with an adult population exceeding 100 million, the procedures of the major reputable survey organizations are geared to mathematical sampling theory and can be expected to provide estimates of overall proportions that are accurate to within a few percentage points. It is true that there are practical problems in applying sampling theory to surveys of human populations, but the problems have to do not with size of sample but with failure to adequately represent all parts of the population. The main loss is unavoidable: some of the people who are selected refuse to be interviewed and others cannot be found during the interviewing period. The proportion missed in this way is often as large as 20 percent and could have serious effects on the results of a poll if such people differed consistently from the rest of the population. Fortunately studies of the characteristics of such "nonrespondents" indicate that omitting them has little effect on the conclusions one would draw about opinion on the Vietnam war. In any event these losses are similar in all surveys of opinion and so cannot serve to explain any seeming difference in the results.

It follows, then, that two national surveys of standard size taken at the same time by competent agencies should show results that differ from each other by no more than a small sampling error. We find this expectation is typically borne out, given one crucial proviso: the questions and the interviewing procedures must be identical or sufficiently similar. The difficulty is that this condition is met far less frequently than the casual observer might expect.

Where voting behavior is concerned there are not many different ways to assess voter intentions, and election polls can be compared readily. Most matters of public policy are not so easily approached. For the war in Vietnam the controversial aspects are almost innumerable. Each aspect requires a somewhat different question, which has its own unique meaning. Moreover, since polling agencies prefer to word their own questions, it is possible to obtain different results even when two agencies are examining essentially the same aspect of an issue. In the case of the war in Vietnam, for example, it is unusual to find instances where two national surveys were made at the same time with questions worded the same way.

The difficulties that can arise from differing questions are easily demonstrated. In the summer of 1966, for example, 70 percent of the people interviewed by the Gallup poll said they approved the bombing by the U.S. of oil storage dumps in Haiphong and Hanoi, 11 percent disapproved and the remainder had no opinion. Two months later the same organization sampled opinion about the demand of many "doves" that the U.S. submit the Vietnam problem to the United Nations "and agree to accept the decision, whatever it may be." In this case 51 percent of the sample thought the proposal was "a good idea" and only

32 percent disliked it. The reader who considered this result to be three to two in favor of the doves and compared it with the apparent majority of seven to one for the "hawks" a few weeks earlier might have had cause to wonder about the solidity of public opinion or the soundness of polling procedures. In actuality the confusion lies in too simplistic a view of the results, since the questions were quite different and could not be expected to give commensurate responses.

A more subtle problem is presented when questions identical in purpose differ markedly in wording or format. In June, 1969, the Gallup poll told respondents that the President had "ordered the withdrawal of 25,000 troops from Vietnam in the next three months" and asked for opinions on whether "troops should be withdrawn at a faster or a

38. How much attention have you been paying to what is going on in Vietnam:

/1.A good deal/ , /3.Some/, or /5.Not much/ ?

38a. Do you think we did the right thing in getting into the fighting in Vietnam or should we have stayed out?

/1.Yes, did right thing/ /5.No, should have stayed out/ /8.Don't know/

38b. Which of the following do you think we should do now in Vietnam? (HAND R CARD #1)

/1.Pull out of Vietnam entirely/

/2.Keep our soldiers in Vietnam but try to end the fighting/

/3.Take a stronger stand even if it means invading North Vietnam/ /8.D.K./

OTHER; DEPENDS: _____

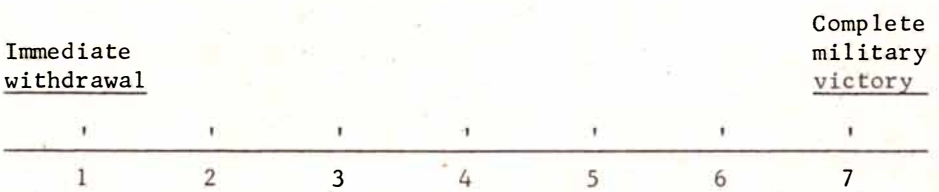
_____

38c. Which party do you think is more likely to take a stronger stand even if it means invading North Vietnam?

/1.Democrats/ /5.Republicans/ /3.No difference/ /8.Don't know/

POLLING FORM used by interviewers for the Survey Research Center of the University of Michigan included questions on the war in Vietnam. Under question 38b the interviewer is instructed to hand the respondent a card; the card enabled the respondent to see the alternatives as the interviewer stated them. The interviewer marked the response in appropriate box below the question.

67. There is much talk about "hawks" and "doves" in connection with Vietnam, and considerable disagreement as to what action the United States should take in Vietnam. Some people think we should do everything necessary to win a complete military victory, no matter what results. Some people think we should withdraw completely from Vietnam right now, no matter what results. And, of course, other people have opinions somewhere between these two extreme positions. Suppose the people who support an immediate withdrawal are at one end of this scale (SHOW CARD #4 TO R) at point number 1. And suppose the people who support a complete military victory are at the other end of the scale at point number 7.



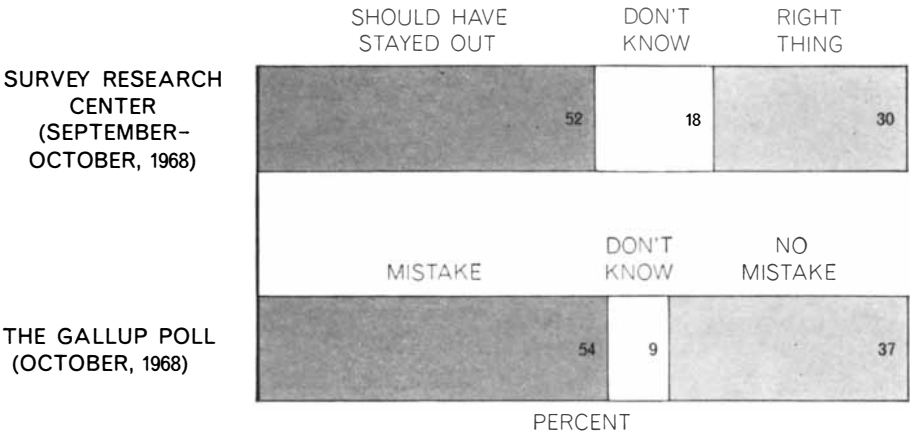
SCALE OF OPINION was presented to respondents by the interviewers for the Survey Research Center and the answer was marked by the interviewer on the form shown here. The respondent was also asked to place several public officials and himself on the scale.

slower rate." "Faster" won over "Slower" by 42 percent to 16 percent, with 29 percent refusing the alternatives presented and instead spontaneously declaring agreement with the President. Scarcely three months later the Harris poll asked a similar question but presented three choices: "In general, do you feel the pace at which the President is withdrawing troops is too fast, too slow, or about right?" Again the equivalent of "faster" won over "slower," by 29 percent to 6 percent, but this time 49 percent approved the current rate. The key to the difference in the two results is that the Gallup format made it easier for respondents to disagree with the existing rate of withdrawal.

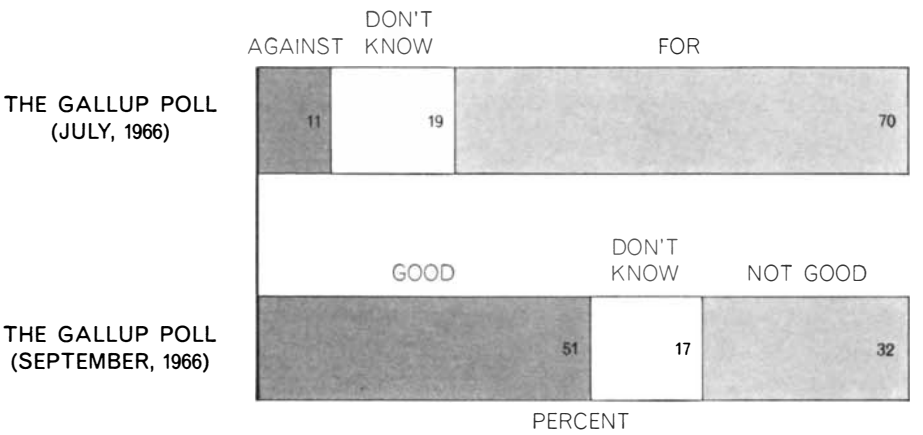
Several approaches are open to the observer who wants to avoid pitfalls of interpretation and make sense of majority opinion on Vietnam. One is to trace changes over a period of time in responses to a question, provided that the fundamental meaning of the question has not been altered by changes in the situation. A second approach is to pay strict attention to nuances in the wording of a number of questions presented at about the same time; in this way one can establish the broad boundaries of public opinion on the war in Vietnam and can isolate the ways that the wording of questions colors responses. A third approach is to explore differing reactions to the same question in various segments of the population.

The broadest trends in American opinion on the Vietnam war over the past decade appear clearly in all surveys. At the beginning of the period the attitude was one of inattentive tolerance toward the U.S. Government's actions in South Vietnam. For example, when respondents in a Survey Research Center poll in 1960 were asked to name the most important problem facing the country, only a tiny number mentioned the Vietnam war even indirectly. The war received 8 percent of first mentions in 1964, 45 percent in 1966 and 41 percent in 1968. This increased awareness was accompanied, particularly in its later phases, by a progressive disillusionment with the correctness of the U.S. commitment. Such disillusionment is best reflected by data published over a period of time by the Gallup poll [*see illustration on next page*].

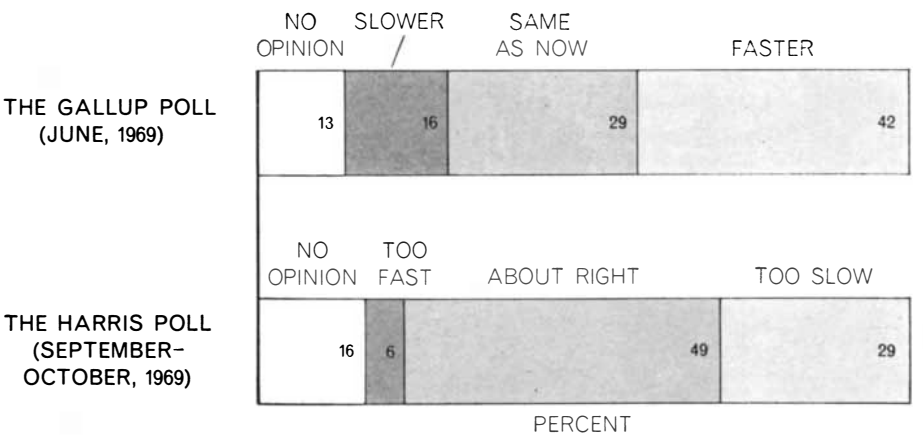
Although it is clear that the public mood about the Vietnam war is now rather firmly negative, it does not follow that the public reaction to alternatives of policy can be as succinctly character-



PARALLEL QUESTIONS asked at about the same time, even by separate polling agencies, should produce comparable results. In this case the Survey Research Center asked: "Do you think we did the right thing in getting into the fighting in Vietnam, or should we have stayed out?" The Gallup question was: "In view of the developments since we entered the fighting do you think the U.S. made a mistake in sending troops to fight in Vietnam?"



DIFFERENT QUESTIONS on the same broad issue can make it seem that many hawks have turned into doves in a short time. Gallup's question in July sought opinion on the U.S. bombing of oil storage dumps in Haiphong and Hanoi. September's question, dealing with a proposal that the U.S. submit problem of what to do about Vietnam to the United Nations and agree to accept the decision, was: "Do you think this is a good idea or not?"



METHOD OF PRESENTATION of a poll question can affect the results. In this case both surveys sought opinion on the rate of troop withdrawals from Vietnam. Gallup poll offered as alternatives only "Faster" and "Slower" but interviewers accepted "Same as now" if it was volunteered. Harris poll offered all three alternatives, thereby making it easier for respondents to go along with rate of withdrawal that President Nixon had announced.

ized. In particular, the simplistic grouping of respondents as hawks and doves breaks down rapidly. Hawks are supposed to believe the U.S. did the right thing in attempting to stem the tide of Communism in Southeast Asia and that the nation should now escalate its military efforts to achieve victory. Doves are supposed to rue the fact that American troops ever became involved in Vietnam and to demand their immediate withdrawal.

Equipped with no more than these definitions, an observer would derive the clear impression from the broad trends reflected by polls that the doves now constitute a majority. Indeed, the Gallup poll showed clearly a shift of doves from a minority to a majority between 1967 and 1969 [*see bottom illustration on opposite page*]. From time to time the poll asked: "People are called 'hawks' if they want to step up our military effort in Vietnam. They are called 'doves' if they want to reduce our military effort in Vietnam. How would you describe yourself—as a hawk or a dove?"

Other data show that the impression given by these results is quite wide of the mark. A Survey Research Center finding in 1968 is a case in point. Respondents were asked not only if the commitment in Vietnam was a mistake but also which of three courses should be pursued: pulling out, preserving the *status quo* or escalating. As one would expect, there was a tendency for people who thought the intervention was right to favor a stronger stand in the future and for people who deplored the intervention to favor withdrawal. Nonetheless, among those who viewed the war as a mistake almost as many favored escalation as were for withdrawal! All told, then, a five-to-three majority regretted the original intervention, but at the same time those calling for "a stronger stand even if it means invading North Vietnam" outnumbered those advocating complete withdrawal by about as large a margin.

How can the Gallup "dove" majority and the Survey Research Center results be reconciled? Some attention to the wording of the questions is rewarding. The description of "hawk" in the Gallup poll compares reasonably well with the "stronger stand" offered as an alternative by the Survey Research Center, and the proportions endorsing both positions at about the same time show no sharp discrepancy. The description of "dove," however, is more elusive. Whereas it was hardly Administration opinion in 1967, by 1969 even President Nixon was advocating a reduction of the U.S. military effort in Vietnam, thereby qualifying as a "dove" by the Gallup definitions. Nonetheless, his position was plainly different from that of the complete-withdrawal "doves" of the Survey Research Center question.

It is also illuminating to consider the trend of opinion from 1965 through 1969 on the President's handling of the Vietnam situation [*see illustration on page 23*]. The main trend is the sharp decline of confidence in President Johnson's management of the war. One further detail worth attention is the time lag before the decline set in. The Johnson electoral landslide in 1964 was seen in many quarters as a vote against escalating the U.S. involvement in Vietnam. The bombing of North Vietnam began only three months later, touching off the first highly vocal bitterness on college campuses because it was seen as exactly the policy of escalation just repudiated. Yet the confidence of the broad public in the President's handling of the situation seemed at first to be heightened. It did not begin to turn downward until a full year later. We shall return to this point.

President Nixon appears to have gained in public approval both over his predecessor and during the period since he took office. By last July he had acquired nearly as much support as President Johnson enjoyed in 1965. The President's announcement last November 3 of his plan for a progressive reduction of the nation's military activity in Vietnam appears to have increased his support still more, although this surge now seems to have passed its peak.

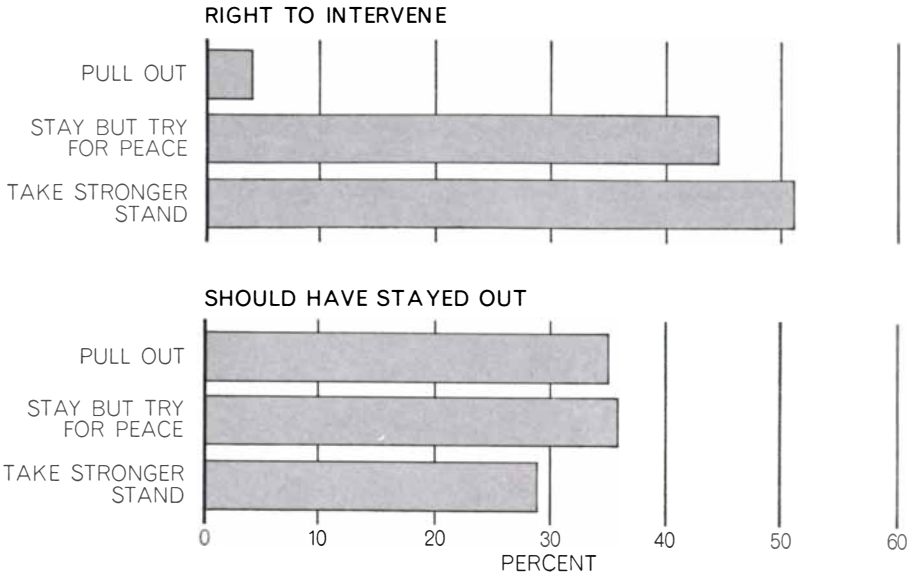
Thus at the turn of the year the President stood squarely aligned with several majorities: those regretting American involvement in Vietnam, those wanting to reduce the nation's commitment there and the large majority rejecting complete withdrawal as an alternative. How stable are these majorities? In what direction might they move under various plausible conditions? Any prognoses must rest to an uncomfortable extent on surmise and are subject to reversal by unforeseen events. It is possible, however, to probe somewhat more deeply into the structure of opinion on the Vietnam war in ways that provide further insights.

Among other things, one must consider the plasticity of opinion. It is widely recognized among workers on public opinion that some measured opinions are much softer, or less crystallized, than others. Soft opinions are likely to change dramatically in response to events, even minor ones, and to the pleas of esteemed leaders.

GROWING DISENCHANTMENT with the U.S. involvement in Vietnam is reflected in five years of responses to a Gallup poll question: "In view of the developments since we entered the fighting do you think the United States made a mistake in sending troops to fight in Vietnam?" Percentage calling it a mistake rose steadily (*black*) and percentage viewing it as not a mistake declined (*color*). The white area represents respondents with no opinion.

The evolution of opinion on Vietnam provides instances of the difference between soft and hard opinion. In March, 1968, for example, the Gallup poll noted that "the North Vietnamese have said that if we agree to stop the bombing of North Vietnam, they will agree to peace negotiations" and asked: "Should we stop the bombing or not?" The result was 40 percent for stopping and 51 percent against. Shortly afterward President Johnson announced his decision to stop the bombing, and in April the Gallup poll asked respondents if they approved; 64 percent did and 26 percent did not. It is safe to assume that the shift of some 25 percent in one month was largely due to the President's speech.

On the other hand, one can consider President Johnson's repeated attempts to keep the public believing that the commitment of troops to Vietnam had been necessary. The people accepted this judgment for a while, but events gradually convinced them that the Vietnam war was a mistake. It is likely that the President's later pleas to the contrary damaged his credibility more than it swayed what had become hard opinion.

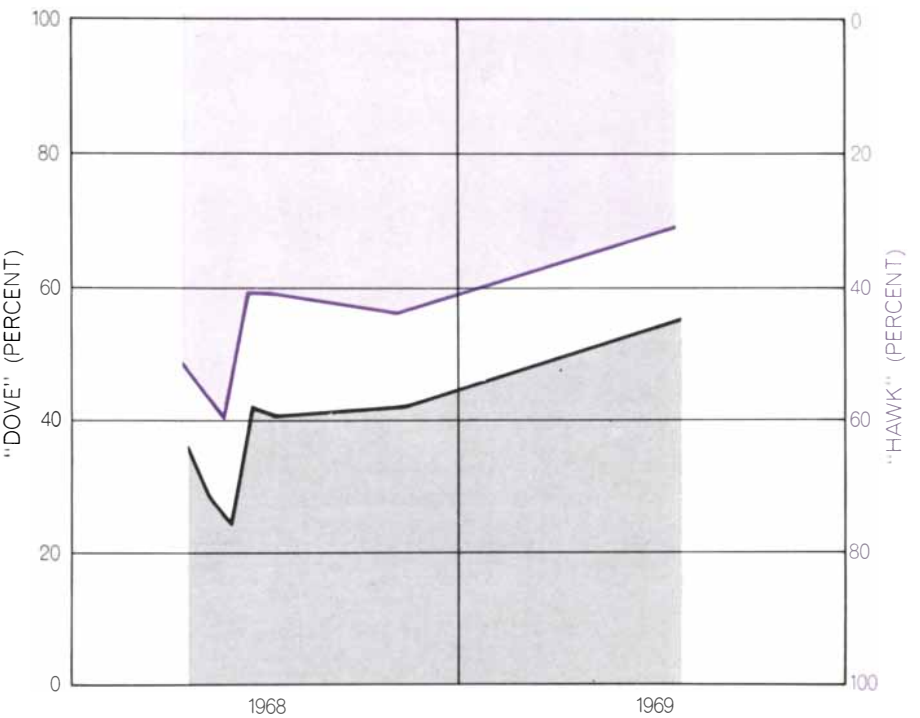


SEEMING PARADOX appeared when Survey Research Center sought opinion in 1968 on what to do in Vietnam. Respondent was asked whether U.S. was right in entering the fighting or should have stayed out; he then judged alternatives of pulling out, staying or taking a stronger stand. Supporters of original intervention tended to prefer stronger stand, but even among those regretting intervention a stronger stand was nearly as popular as pulling out. Both for regretting intervention and for taking stronger stand, margin was five to three.

In general, it is likely that crystallized opinion differs from soft opinion in being anchored more deeply in individual experience, information and motivation. Although there was a time when most Americans lacked enough information about Vietnam to form solid judgments on the nation's involvement there, experience has mounted steeply in terms of deaths, taxes and soldiers not home for Christmas. The end state, a satisfactory disengagement from Vietnam, has come rather clearly into the public's mind. As for the means of reaching that end, opinion is much more tentative. Moreover, there is no reason to believe it will ever become sharply crystallized. To a large degree the public entrusts officials to make the detailed policy decisions, although it may ultimately punish them at the polls if the selection of means turns out to have been unfortunate.

None of this is to say that all opinions favoring withdrawal or escalation are casual. Many of them are hard, and if the President succeeded in disengagement, there would surely be bitter recriminations from those whose preferences were slighted. Nor are even the deep regrets about the nation's involvement in the Vietnam war immutable. One can imagine circumstances, for instance a southward sweep by Communist China or the collapse of the Vietcong, when such regrets would be swept away as irrelevant. The important point is that if the basic elements of the situation remain as they are, it is likely that a successful disengagement by the President would probably win the strong approval of a vast majority, regardless of the means that might have seemed more sensible along the way.

This interpretation is supported by other observations about polls. People following polls have noted how the support of the President seems to rise after any new initiative, whether it is in the direction of escalation or a reduction of commitment. It is in this light that we interpret the rise in support for President Johnson after his decision to bomb



SELF-CLASSIFICATION of respondents as "hawks" (*color*) and "doves" (*black*) on the Vietnam war was made for Gallup poll over two years. Sharp change in 1968 marks Tet offensive. Although question wording remained the same, significance of the terms shifted.

North Vietnam. The support lasted until it became clear that the projected results had been illusory. Similarly, the reaction to President Nixon's speech of last November may chiefly reflect satisfaction that some kind of effort was being made to move the situation off dead center. The entry into Cambodia seemed to be received at first with a parallel surge of public support, in spite of sharp attacks on the policy by leading spokesmen.

What if there is no successful disengagement within a reasonable period of time? If our analysis is correct, we would expect strong public disapproval to set in. It would not be surprising if such a reaction gathered speed more swiftly than the comparable moods during the Korean war and the latter days of the Johnson Administration. In this sense the provisional timetable laid out in the speech of November 3 is a short fuse on the powder keg, much like the Johnson Administration's statements about having the troops home for Christmas of 1965. There is some evidence, for example, of considerable support for a faster withdrawal of troops than present policy envisions. Moreover, although support for immediate withdrawal has not been large in the past, by early this year there were signs that it might be growing significantly. A definite timetable of a year or so for withdrawal might well be a political platform that could attract widespread support.

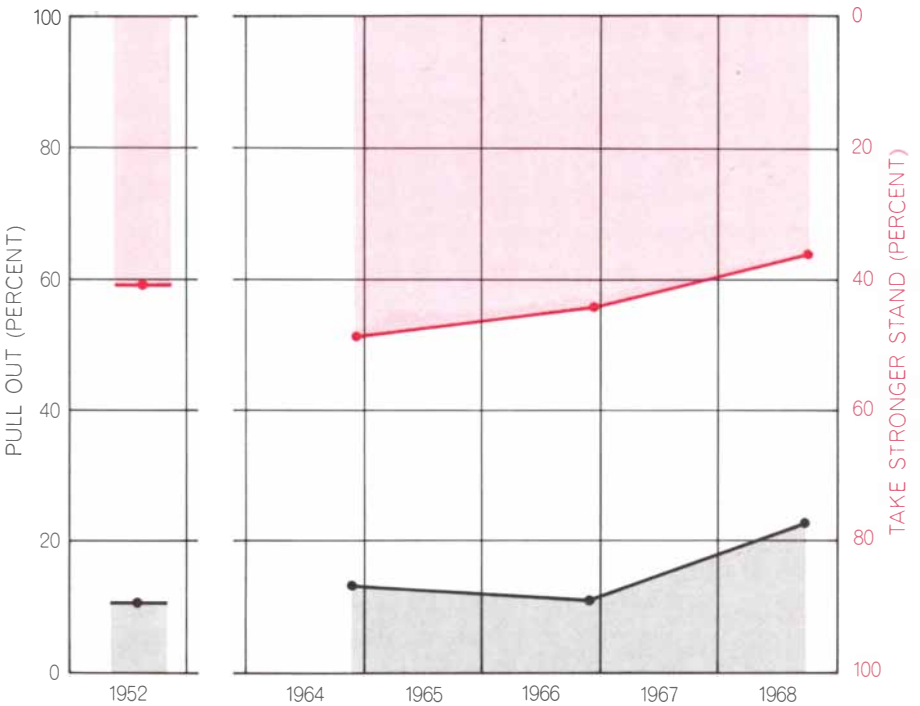
It is instructive, in examining the structure of public opinion on Vietnam, to look closely at the segments of the population where support for or resistance to the war is particularly strong. In 1968 the Survey Research Center asked: "How much attention have you been paying to what is going on in Vietnam? A great deal, some, or not much?" When views on the Vietnam war are divided in this way, one finds that the most attentive persons seem to show the firmest support for the war [*see illustration on page 24*]. This relation is most notable in the responses to an additional question on whether or not the original intervention in Vietnam was right. We are aware that self-gauged interest is not necessarily a reliable indicator of how well informed the respondent may be. Nonetheless, the results seem surprising in the light of the association between the concern and the negative feeling about the Vietnam war that has been evident on the campuses of leading universities.

Although attitudes toward the Vietnam war have consistently shown less-distinct differences among various standard groups of the population than is true of many other political attitudes, two exceptions stand out. They are race and sex. Blacks and women have shown more disenchantment with the involvement in Vietnam than white males over the entire period when relevant samplings have been made. The patterns are not surprising. The war in Vietnam has been a notable focus of alienation in the development of black consciousness during the 1960's. Women have traditionally been unenthusiastic about war involvement, and the largest sex differences in responses emerge when policies involving strong military initiatives are at stake. Moreover, blacks and women, for different but obvious reasons, typically show less interest in political affairs than white males do. Therefore some of the association between interest and relatively hard attitudes on Vietnam probably arises from these background differences.

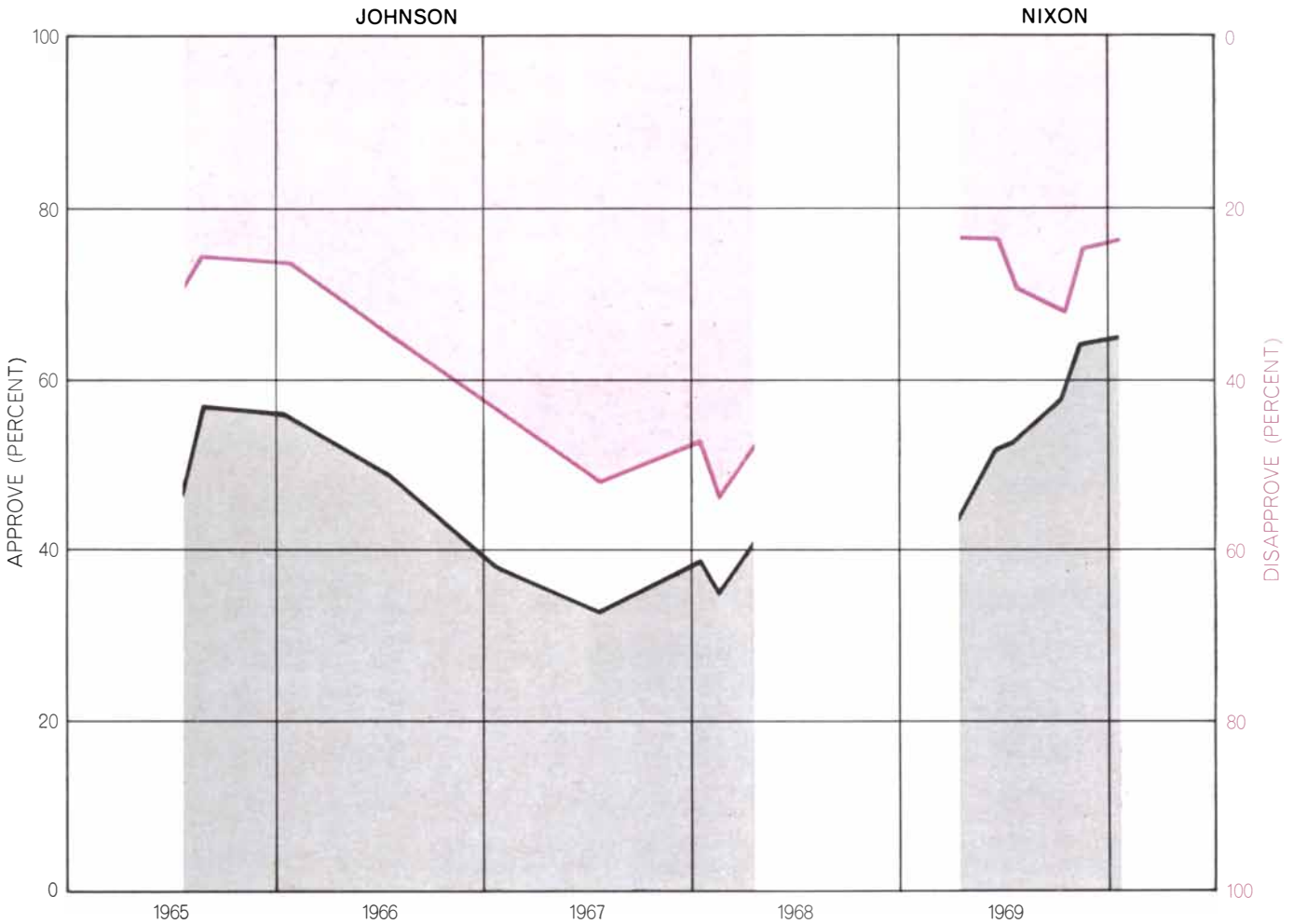
Another pattern, not quite as distinct, is a positive association between education and support for the nation's involvement in Vietnam. Respondents classed as "college-educated" register disproportionately among those who are most "hard-line," or least negative, about the Vietnam war. Here again, since there is typically a fairly strong relation between education and interest in public affairs, one sees a pattern that contributes to the association between attentiveness to the war and support for the nation's involvement.

The surprising feature of the relation between education and attitudes toward Vietnam, however, is that the pattern does not seem to be of a piece with the campus ferment against the war. Moreover, to compound the confusion, national surveys did not yield until recently any distinct relation between age and attitude toward Vietnam. The "generation gap" that one would have expected, wherein the young oppose the war and the old support it, simply failed to appear, and even now it is not very large. Indeed, putting these two findings together, it can be shown with Survey Research Center data as of 1968 that (among whites) college-educated people in their twenties were more likely than older people of grade school education both to justify the war and to favor an intensification of it. The differences are substantial, running to 20 percent or more.

We have subjected the relation between college training and support for the war to closer scrutiny, and the results are highly interesting. An obvious explanation of the seeming paradox is that most of the people who comprise the category of "college-educated" persons in poll results have been out of college for many years. Moreover, the "college-educated" group is a good deal more



ATTITUDES ON KOREA AND VIETNAM were parallel in findings made by the Survey Research Center. A small group (*black*) favored pulling out, a larger group (*color*) favored taking a stronger stand and a third group (*white*) favored staying but trying for peace.



SUPPORT FOR PRESIDENT fluctuates considerably in data obtained by the Gallup poll. Over a period of three years in the Johnson Administration and several months of the Nixon Administration poll asked respondents whether they approved or disapproved of the way the President "is handling the situation in Vietnam." Black curve shows approval; color, disapproval; white, uncertainty.

heterogeneous than one tends to think. It includes, for example, many people who attended a junior college or failed to complete the requirements for a degree at a four-year college.

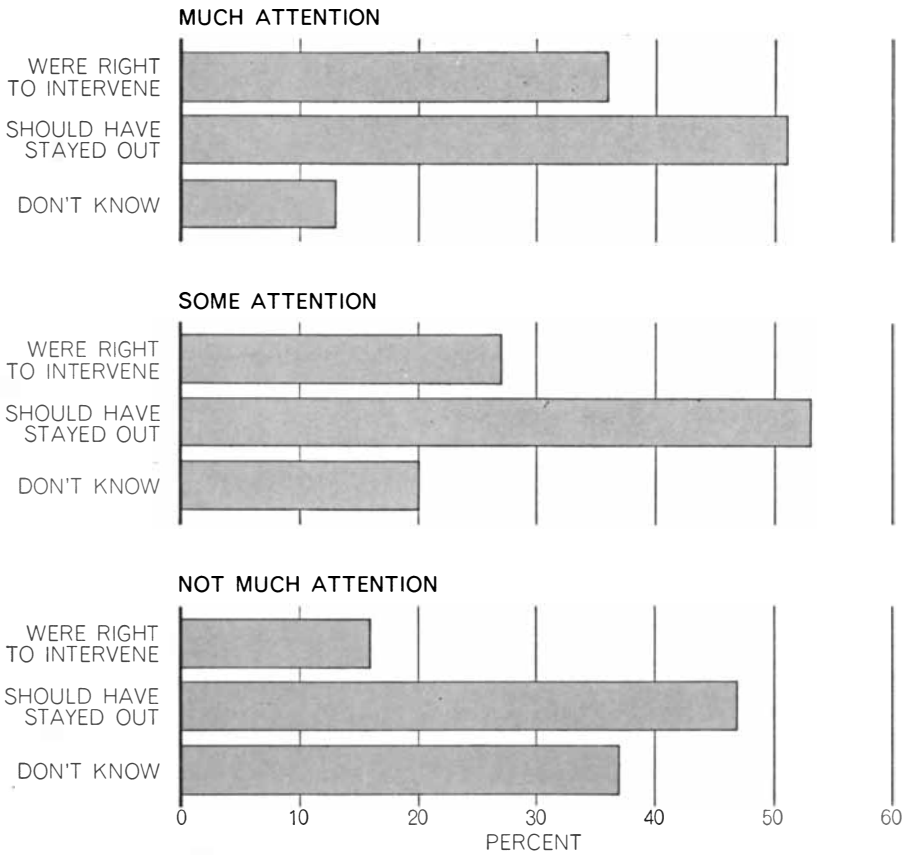
It is plain that before the entry into Cambodia last month most of the vehement antiwar sentiment had appeared at a fairly small number of campuses usually classed as leading universities. Although these institutions are large and each of them awards many degrees, they are vastly outnumbered by the hundreds of smaller colleges in the nation. As a result, fewer than 25 percent of the college-educated members of the population have ever had contact with a university where the antiwar sentiment has been strong for some time, and for many of those who have had such an affiliation the connection ended long ago.

With these considerations in mind, we have grouped college-educated respondents in Survey Research Center samples according to a general quality rating of the universities they attended. The rating is based on data concerning faculty salaries and on other information reflecting the academic quality of student bodies. We have also set apart as a separate group the college graduates who did postgraduate work.

Within this framework we have examined the responses to a number of questions, such as the one about whether the respondent favors a stronger stand, preservation of the *status quo* or withdrawal from Vietnam. The reader will recall that responses to this question by the total population inched in a dovish direction between 1964 and 1966 and moved more rapidly in that direction between 1966 and 1968. In each of those years the college-educated group, taken as a whole, was slightly more hawkish than the noncollege population [*see illustration on page 25*]. Yet even by 1964 people with graduate training were more dovish than the noncollege group had become by 1968.

More detailed analyses show that women from the "quality" institutions were already quite negative about the war in 1964, although the males from those institutions were at the time thoroughly hawkish. By 1966, however, males with a background in the prominent universities (but without graduate training) had become more resistant than the noncollege population to a hard line in Vietnam.

Throughout the entire period from 1964 to 1968 alumni of the smaller colleges, although they came to see the war as a mistake, clung to a harder line than even the noncollege population. It is this constituency from smaller colleges more than any other that has served as the backbone of popular support for the war. Since this group is numerically the largest in the college-educated population, its views explain why national survey data show people of college background giving relatively strong support to the war. The spread of strong antiwar protest to smaller campuses in recent months, particularly in reaction to the



ATTENTION TO WAR and attitudes on intervention were compared by the Survey Research Center. Respondents were asked whether they paid much, some or little attention to developments in Vietnam and then were asked their view of original intervention by U.S.

Cambodian campaign, may be a forerunner of broader disillusionment with the war even within the more supportive segments of the public.

In any event these findings suggest rather clearly that feeling against the war has consisted up to now of two currents that are widely separated from each other. One current is made up of a tiny fraction of the population, but one that is highly educated, articulate and visible. The other group tends to be less educated than the national average and is much less politically visible, although it is far larger than the set of vocal critics—perhaps by a factor of 10 or more.

It is likely that the grounds for antiwar sentiment are quite disparate between the two currents. There is of course plain evidence of moral outrage in the more highly educated and articulate group. The polls have made little effort to illuminate the bases of negativism in the broader public, but signs of any moral overtones to this larger discontent are few. General reaction to the reports of the massacre at Mylai, as recorded in a Harris poll in January, can perhaps best be described as bland. Most disenchantment with the war seems pragmatic and can be summed up in the attitude that "we have not won and have little prospect of doing so."

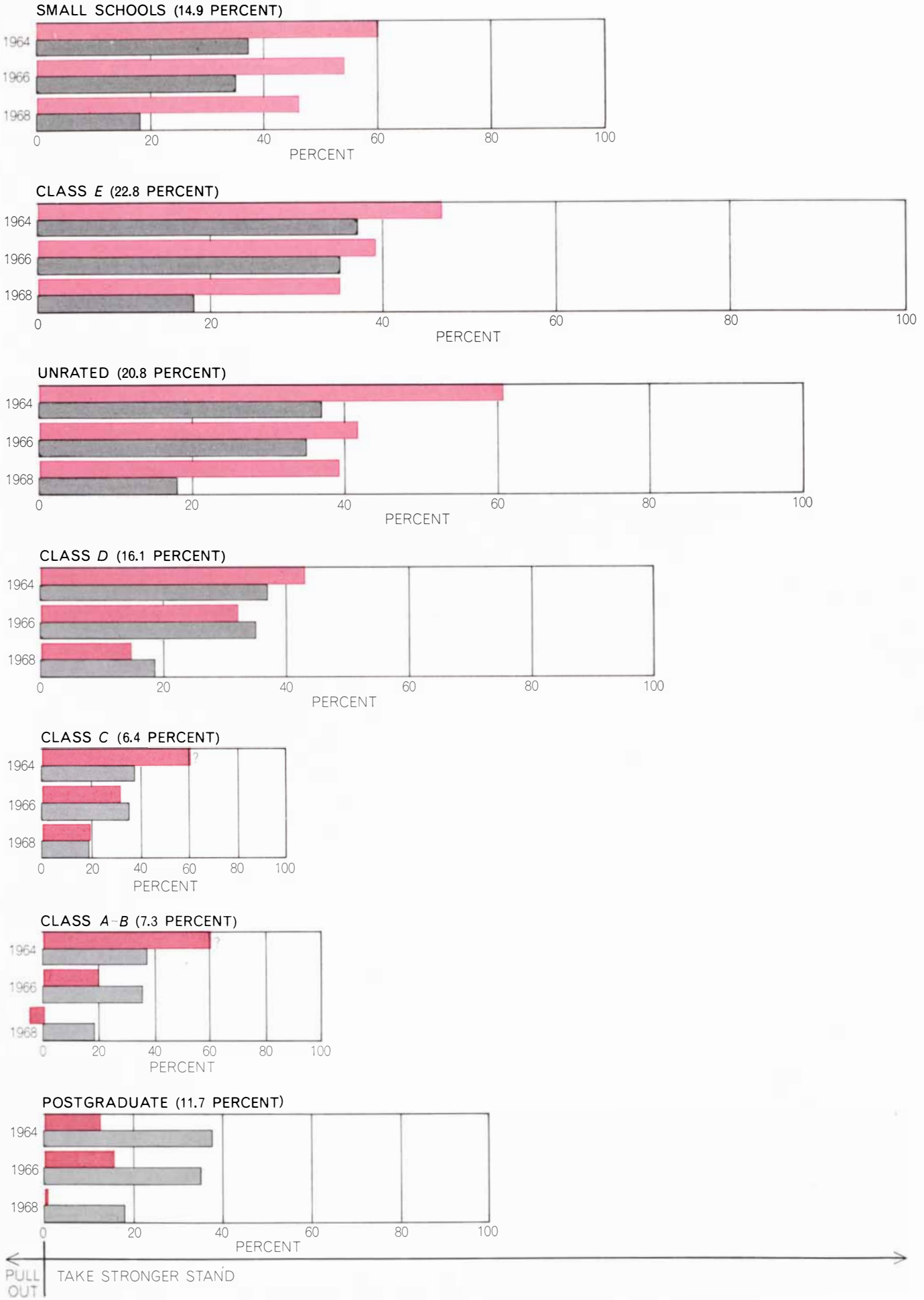
This simple war-weariness has colored poll results for some time. One of the most negative responses to any policy proposal for the Vietnam war appeared when the Gallup poll asked in 1967 about a "suggestion" that "income taxes be raised to help pay for the war in Vietnam": 70 percent of the respondents were opposed. Almost any proposal to share or somehow unload the burden of the war wins strong approval. We have mentioned the positive response to the idea of submitting the issue to the UN. In view of the President's policy of "Vietnamization" it is interesting to note that as early as 1966 the Gallup poll showed 62 percent favoring the transfer of "more responsibility for the fighting of the war" to South Vietnam, with 22 percent opposed and the remainder uncertain. A year later the edge had risen to 77 to 11.

The failure of most commentators to recognize the profound differences between the two currents of opposition to the war may have led to much of the perplexity over the behavior of public opinion. People who oppose the war for moral reasons are little affected by new events that suggest victory or defeat, whereas those who oppose the war for pragmatic reasons are greatly influenced by just such events.

There is much reason to believe the two groups have little communication with each other. The smaller group is in the vanguard of protest against the war, and many of its members are among the militant protesters. The larger group, although it also regrets the war, shares at the same time the antipathy that most Americans have to any political dissent that goes beyond the confines of conventional debate. It is therefore not surprising that when the Survey Research Center asked respondents in 1968 to evaluate a wide range of political leaders and groups on a "feeling" scale, ranging from extremely negative to highly positive, reactions toward "Vietnam war protesters" were by a substantial amount the most negative in the set. Nearly 75 percent of the respondents rated protesters in the negative half of the scale and more than 33 percent placed them at the extreme negative point, which was not otherwise much employed. Indeed, 63 percent of those believing the war was a mistake viewed protesters negatively, and even of the group favoring complete withdrawal from Vietnam, 53 percent put the protesters on the negative side of the scale. Plainly opposition to the war and opposition to active protest against it go together for a significant part of the population.

These findings help to fill in a profile of public opinion on the war in Vietnam. They also lend credence to the proposition that the net effect of vigorous protest in the streets has been to shift mass opinion toward renewed support of the President. On the other hand, the role of highly visible dissent in keeping a wider range of options in the public eye and in encouraging dovish spokesmen in Congress or skeptical commentators in the mass media could be quite real.

COLLEGE EDUCATION and attitudes toward Vietnam are compared in the chart on the opposite page, which uses a rating of colleges made by the Survey Research Center on a sliding scale that begins at *A* with the universities that pay the highest faculty salaries and have student bodies of high academic quality. Large bar encompassing each group of smaller bars reflects percentage of the U.S. adult population that attended each type of college. Gray bars show opinion of the noncollege population.



THE ORIGIN OF GALAXIES

The size, shape and other properties of the observed galaxies are traced to slight density enhancements in the expanding primordial fireball. Enhancements of certain mass were favored over others

by Martin J. Rees and Joseph Silk

Perhaps the most startling discovery made in astronomy this century is that the universe is populated by billions of galaxies and that they are systematically receding from one another, like raisins in an expanding pudding. If galaxies had always moved with their present velocities, they would have been crowded on top of one another about 10 billion years ago. This simple calculation has led to the cosmological hypothesis that the world began with the explosion of a primordial atom containing all the matter in the universe. A quite different line of speculation argues that the universe has always looked as it does now, that new matter is continuously being created and that new galaxies are formed to replace those that disappear over the "horizon."

On either hypothesis it is still necessary to account for the formation of galaxies. Why does matter tend to aggregate in bundles of this particular size? Why do galaxies comprise a limited hierarchy of shapes? Why do spiral galaxies rotate like giant pinwheels? Astrophysicists are trying to answer these and similar questions from first principles. The goal is to explain as many aspects of the universe as one can without invoking special conditions at the time of origin. In most of what follows we shall assume a cosmological model in which the universe starts with a "big bang." When we have finished, the reader will see, however, that some form of continuous creation of matter may not be ruled out.

Before the invention of the telescope the unaided human eye could see between 5,000 and 10,000 stars, counting all those visible in different seasons. Even modest telescopes revealed millions of stars and in addition disclosed the existence of many diffuse patches of light, not at all like stars. These extragalactic "nebulas," many of them beautiful spirals, are seen in all directions and in great profusion. As early as the 18th century Sir William Herschel and Immanuel Kant suggested that these nebulas were actually "island universes," huge aggregations of stars lying far beyond the limits of the Milky Way.

The validity of this hypothesis was not confirmed until 1924, when the American astronomer Edwin P. Hubble succeeded in measuring the distances to a number of spiral nebulas. Several years earlier Henrietta S. Leavitt had shown that Cepheid variables, named for the prototype Delta Cephei, a variable star discovered in 1784, had light curves that could be correlated with their magnitude. The distances of a number of Cepheids were later determined by independent means, so that it became possible to use more distant Cepheids as "standard candles" to establish a distance-magnitude relation. Hubble looked for Cepheid variables in some of the nearer external galaxies and found them. From their period he was able to deduce their absolute luminosity, and from this he was able to estimate their distance. Hubble soon established that the nearest spiral nebulas (or galaxies) were vast systems of stars situated a million or more light-years outside our own galaxy.

Subsequently Hubble developed a scheme for classifying galaxies according to their morphology, ranging from systems that are amorphous, reddish and elliptical to systems that are highly flattened disks with a complex spiral structure containing many blue stars and lanes of gas and dust [*see illustration on page 28*]. The spiral galaxies themselves vary in appearance. At one extreme are those with large, bright nuclei and inconspicuous, tightly coiled spiral arms. At the other extreme are galaxies in which the nuclei are less dominant and the spiral arms are loosely wound and prominent. The elliptical galaxies also form a sequence, ranging from almost spherical systems to flattened ellipsoids. In addition there are highly irregular systems showing very little structure of any kind.

In all these sequences there is a parallel progression in certain characteristics of the galaxies. In general spirals are rich in gas and dust, contain many blue supergiant stars, are highly flattened and rotate appreciably. Ellipticals, by contrast, seem to possess little gas or dust, usually contain late-type dwarf stars and exhibit scant rotation.

The masses of galaxies are found by several methods. Galaxies are often gravitationally bound together in pairs. If the distance between them and their relative velocities are known, Kepler's law can be used to find their total mass.

CLUSTER OF GALAXIES in the constellation of Hercules demonstrates the inhomogeneity of the distribution of galaxies in the sky. About 350 million light-years away, this cluster contains about 100 members and is some five million light-years across. It was photographed with the 200-inch Hale telescope on Palomar Mountain. Some very rich clusters contain 1,000 members or more and vary from one million to 10 million light-years across. There is some evidence that such clusters are in turn grouped together into superclusters of perhaps 100 members, spread over 100 million light-years. On scales larger than this the universe appears to be uniform. The bright circular spots with the spikes radiating from them are nearby stars; the spikes are produced by reflections within the telescope.

Another method, used mostly for spirals that are viewed edge on or obliquely, is to determine the velocity of rotation by measuring the Doppler shift of spectral lines emitted by ionized gas in various parts of the disk. (The spectral lines of approaching gas will be shifted toward the blue end of the spectrum, those of retreating gas toward the red end.) One can plot a rotation curve showing how the velocity of rotation varies with the distance from the center of the galaxy. The mass can then be estimated from the requirement that the centrifugal and gravitational (centripetal) forces must be in balance. It turns out that the masses of galaxies are typically about 10^{11} (100 billion) times the mass of the sun. The range, however, is fairly broad: from about 10^8 solar masses for some nearby dwarf galaxies to 10^{12} solar masses for giant ellipticals in more remote regions of the universe. The diameter of the larger spirals, such as our own galaxy, is about 100,000 light-years.

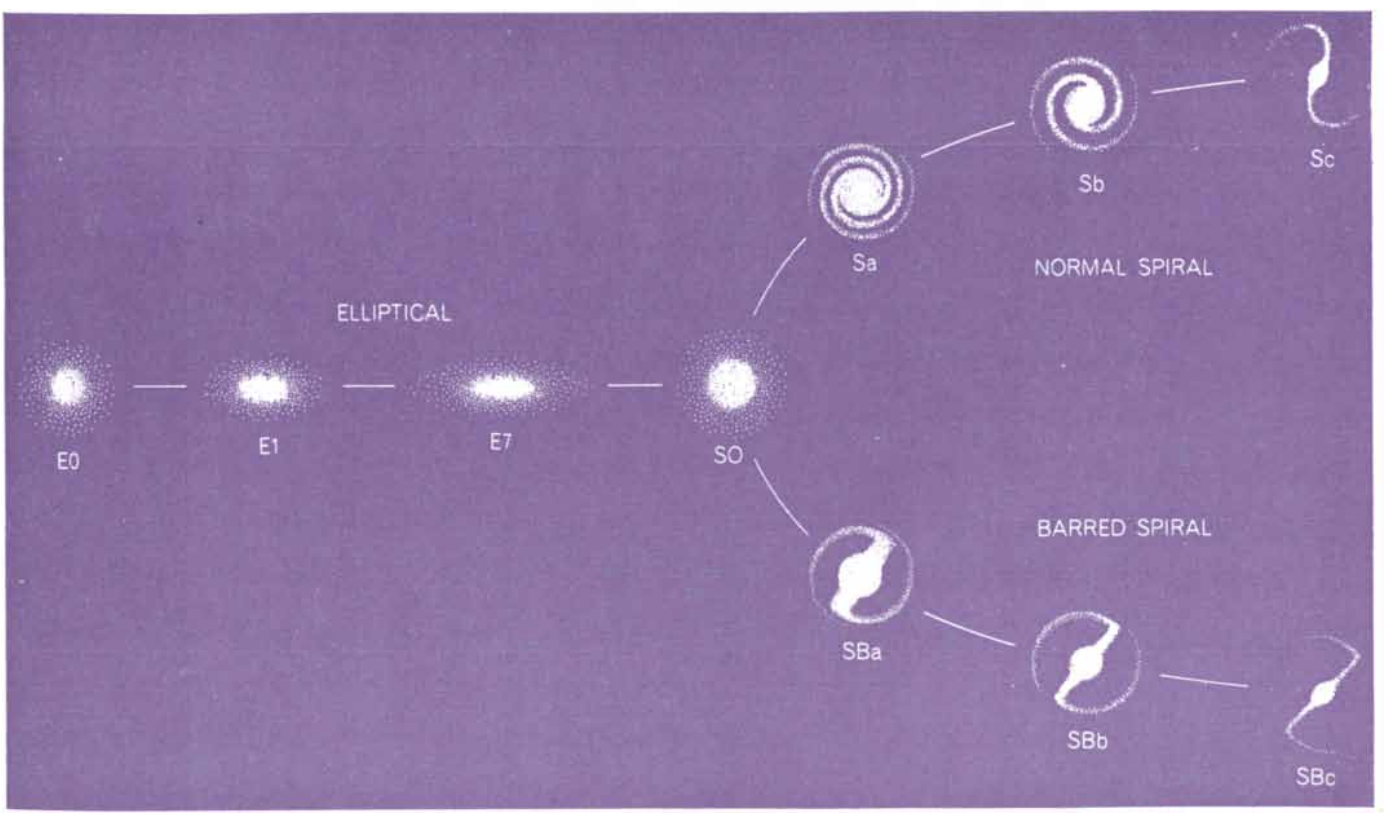
Galaxies also differ widely in the ratio of mass to luminosity. Taking the mass-to-luminosity ratio of the sun as unity, one finds that for large spirals, such as our own galaxy, the ratio varies from one up to 10. In other words, some spirals emit only a tenth as much light per unit of mass as the sun does. Ellipticals commonly emit even less: only about a fiftieth as much light per unit of mass. (Thus their mass-luminosity ratio is 50.)

The distribution of galaxies in the sky is quite inhomogeneous. There are many small groups of galaxies, and here and there some rich clusters containing up to 1,000 members or more. Such systems vary from one million light-years across to 10 million. Our own galaxy is a member of the "local group," an association of about 20 galaxies, only one of which, the Andromeda galaxy, has a mass comparable to that of ours. The local group is about three million light-years in diameter. The Andromeda galaxy is some two million light-years away; the nearest large cluster of galaxies is in Virgo, about 30 million light-years distant.

Even such clusters do not seem to be randomly distributed in space. Some astronomers have argued that there is evidence that clusters are grouped into superclusters of perhaps 100 members, spread over 100 million light-years. The universe appears to be uniform on scales larger than this.

Establishing the distance of galaxies was only part of Hubble's achievement. Working with the 100-inch telescope on Mount Wilson, he showed from red-shift measurements that the galaxies are in recession. Hubble found, moreover, that the red shift of a galaxy is directly proportional to its distance, as judged by its apparent luminosity. The most distant galaxies known are in a faint cluster in the constellation Boötes; Rudolph Minkowski discovered that the wavelength of light coming from this cluster is stretched by 45 percent. The corresponding velocity of recession is nearly half the speed of light. Light originating from some of the brilliant starlike objects known as quasars is red-shifted more than 200 percent, but astronomers disagree whether or not this red shift is due to the cosmological expansion of the universe.

The light from Minkowski's cluster of galaxies set out toward us about five bil-



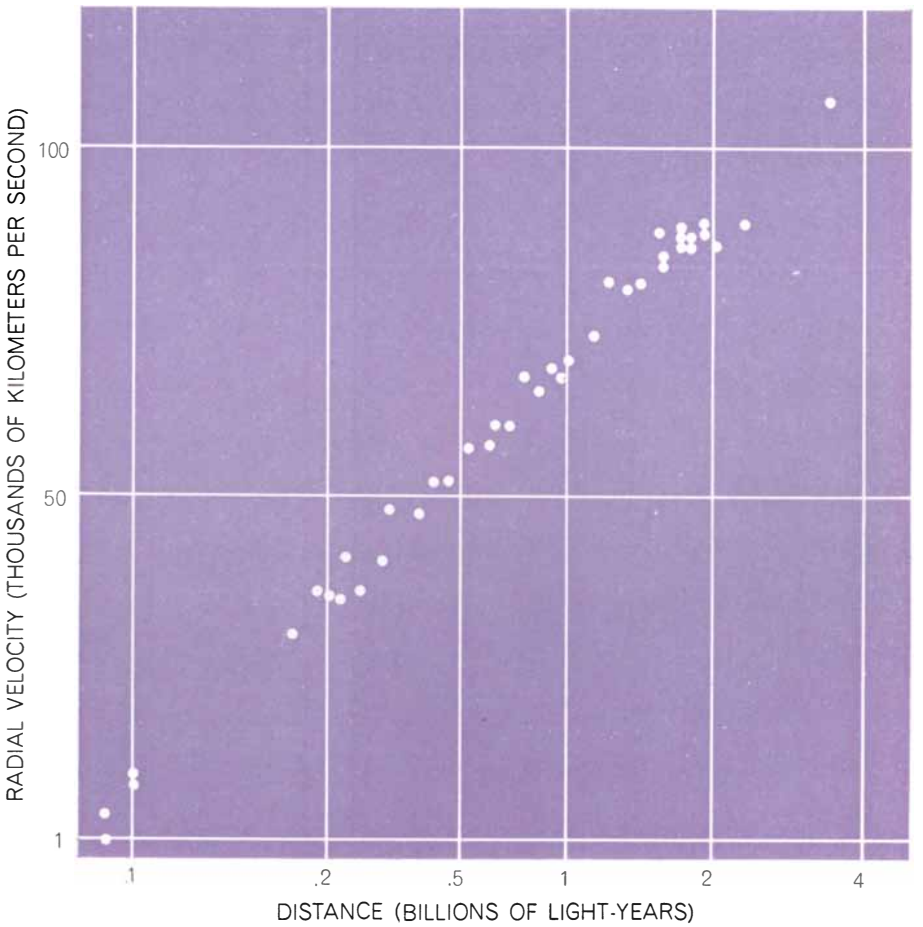
CLASSIFICATION SCHEME developed by Edwin P. Hubble in the early 1930's organizes galaxies according to their shape, ranging from amorphous elliptical systems containing many red stars and little gas and dust (*left*) to highly flattened spiral disks containing many blue stars and lanes of gas and dust (*right*). The elliptical galaxies range from almost spherical systems (designated E0) to highly flattened ellipsoids (E7). The spiral galaxies themselves form two sequences: normal spirals (*top right*) and barred spirals (*bottom right*). At one extreme in both sequences are galaxies with large bright nuclei and inconspicuous, tightly coiled spiral arms (Sa, SBa); at the other extreme are galaxies in which the nuclei are less dominant and the spiral arms are loosely wound and prominent (Sc, SBc). At the branching point of the diagram is a disklike form that resembles the spirals but lacks spiral arms (SO).

lion years ago, and so we can be sure that some galaxies are even older than that. On the other hand, as we have mentioned, all the galaxies must have been tightly packed together no more than 10 billion years ago, based on their present recession velocity. Estimates of the ages of stars suggest that our galaxy, and others like it, are unlikely to be much less than 10 billion years old. Hence we are presented with a remarkable coincidence: most galaxies appear to be about as old as the universe. This implies that galaxies must have formed when conditions in the universe were much different from those now prevailing.

It seems clear, then, that the formation of galaxies cannot be treated apart from cosmological considerations. The dynamics and structure of the universe in the large are beyond the scope of Newtonian physics; it is necessary to use Einstein's general theory of relativity. Because of the complexity of the theory, it is practicable to solve the equations only for cases having special symmetry. Until quite recently the only solutions for an expanding universe were those found in 1922 by the Russian mathematician Alexander A. Friedmann. In his idealized models matter is treated as a strictly uniform and homogeneous medium. The universe expands from a singular state of infinite density, with the rate of expansion decelerating as a consequence of the mutual gravitational attraction of its different parts. The universe may have enough energy to keep expanding indefinitely or the expansion may eventually cease and be followed by a general collapse back to a compressed state. Observations of the actual rate of expansion of the universe at different epochs, as determined by the red shift–luminosity relation of the most distant galaxies, fail to tell us unambiguously whether the expansion will finally stop and be reversed or whether it will continue indefinitely.

The clumping of matter into stars, galaxies and clusters of galaxies in the real universe might seem to make Friedmann's models, based on perfect homogeneity, empty exercises. In actuality the "graininess" we observe in the universe is on such a small scale that Friedmann's solutions remain valid. The reason is that the gravitational influence of local irregularities is swamped by that of more distant matter.

Perhaps the most convincing evidence in support of Friedmann's simple description of the universe was supplied



RECESSION VELOCITY OF A GALAXY is obtained by measuring the amount by which the radiation it emits is shifted to the red end of the spectrum. The velocity is directly proportional to the galaxy's distance, as judged by its apparent luminosity. In this diagram, adapted from a recent study by Allan R. Sandage of the Hale Observatories, the ratio of recession velocity to distance is shown for brightest galaxy in each of 41 clusters of galaxies.

in 1965 by the discovery that space is pervaded by a background radiation that peaks at the microwave wavelength of about two millimeters, corresponding to the radiation emitted by a black body at an absolute temperature of three degrees (three degrees Kelvin). This radiation could be the remnant "whisper" from the big bang of creation. The remarkable isotropy, or nondirectionality, of this radiation is impressive evidence for the isotropy of the universe.

The radiation was discovered independently and almost simultaneously at the Bell Telephone Laboratories and at Princeton University [see "The Primeval Fireball," by P. J. E. Peebles and David T. Wilkinson; SCIENTIFIC AMERICAN, June, 1967]. The radiation has the spectrum characteristic of radiation that has attained thermal equilibrium with its surroundings as a result of repeated absorption and reemission, and it is generally interpreted as being a relic of a time when the entire universe was hot, dense and opaque. The radiation would have cooled and shifted toward longer wavelengths in the course of the universal expansion but would have retained a thermal spectrum. It thus constitutes remarkably direct evidence for the hot-big-bang model of the universe first examined in detail by George Gamow in 1940.

Assuming the general validity of the Friedmann model for the early stages of the universe, it seems clear that the material destined to condense into galaxies cannot always have been in discrete lumps but may have existed merely as slight enhancements above the mean density. There will be a tendency for the larger irregularities to be amplified simply because, on sufficiently large scales, gravitational forces predominate over pressure forces that tend to oppose collapse. This phenomenon, known as gravitational instability, was recognized by Newton, who, in a letter to Richard Bentley, the Master of Trinity College, wrote:

"It seems to me, that if the matter of our sun and planets, and all the matter of the Universe, were evenly scattered

SPHERICAL GALAXY, classified as type E0 in Hubble's scheme, is a member of the Virgo cluster of galaxies. A representative of the most massive type of galaxy, this system, designated M 87, contains about 30 times as many stars as a spiral system such as our own does.

ELLIPTICAL GALAXY in the constellation of Cassiopeia is a member of the local group of galaxies. Designated NGC 147, it is an E4, or intermediate type of elliptical galaxy. Because of its comparative proximity to our system it can be resolved into individual stars.

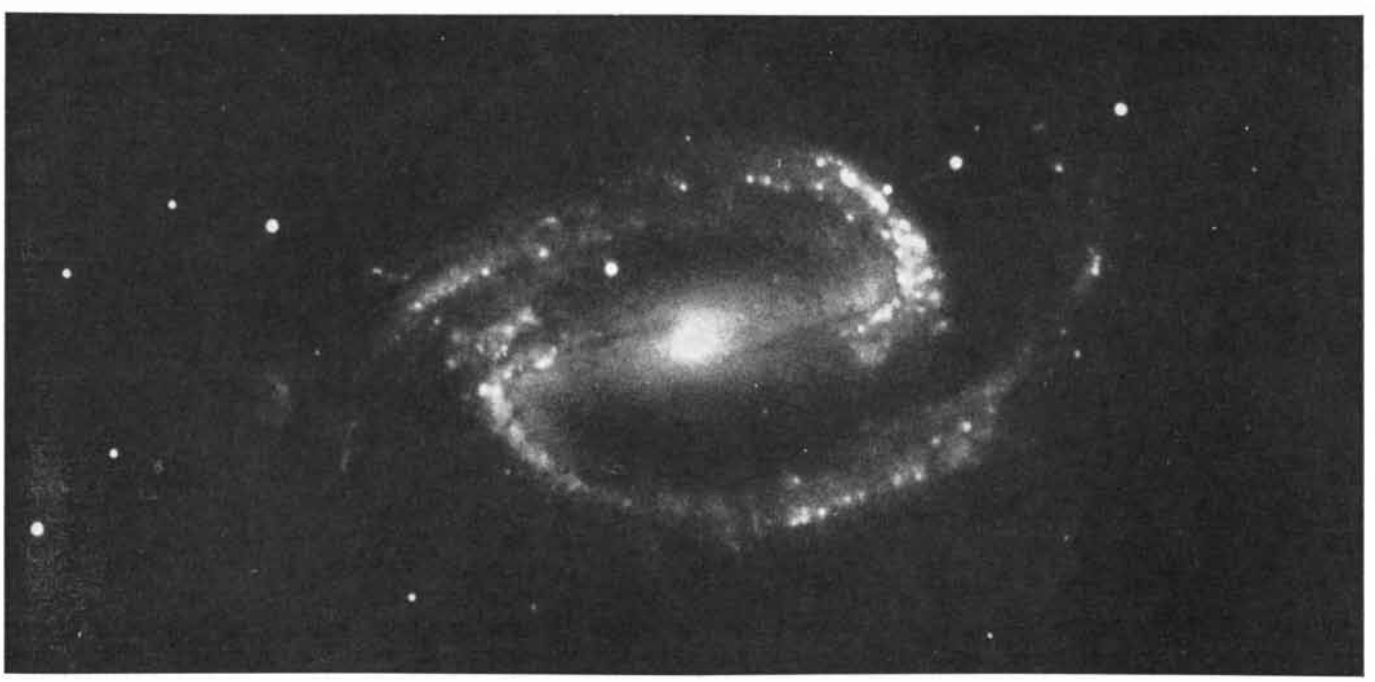
through all the heavens, and every particle had an innate gravity towards all the rest, and the whole space throughout which this matter was scattered, was finite, the matter on the outside of this space would by its gravity tend towards all the matter on the inside, and by consequence fall down into the middle of the whole space, and there compose one great spherical mass. But if the matter were evenly disposed throughout an infinite space, it could never convene into one mass; but some of it would convene into one mass and some into another, so as to make an infinite number of great masses, scattered great distances from one to another throughout all that infinite space. And thus might the sun and fixed stars be formed, supposing the matter were of a lucid nature."

Newton envisaged a static universe, but the same qualitative picture occurs in an expanding Friedmann universe, as was shown by the Russian physicist Eugene Lifshitz in 1946.

Because of the atomic nature of matter the early universe could never have been completely smooth. It would obviously be gratifying if the inevitable random irregularities in the initial distribution of atoms sufficed ultimately to produce the bound systems of stars we see throughout the universe today. Unfortunately this type of statistical fluctuation fails by many orders of magnitude to account for the observed degree of structure in the universe. Moreover, it remained a puzzle why agglomerations of a certain mass, notably galaxies, should be so plentiful. It appeared necessary to postulate initial fluctuations in a seemingly *ad hoc* manner, and nothing had really been explained; "things are as they are because they were as they were."

Only in the past two or three years has it been realized that the background radiation acts as a gigantic homogenizer on certain preferred scales. To understand just how this works we must look more closely at Gamow's model of the universe. In the early stages, when the universe consisted of a primordial fireball, no structures such as galaxies or stars could have existed in anything like their present form. All space would have been filled with radiation (photons) and hot gas, consisting of the nuclei of hydrogen and helium and the accompanying electrons. The photons would be repeatedly scattered from the electrons. For at least the first 100,000 years of its history (beginning roughly 10 seconds after its emergence

SPIRAL GALAXY M 101 in Ursa Major is representative of the Sc type, which is characterized by a relatively inconspicuous nucleus and prominent, loosely wound spiral arms. Our own Milky Way galaxy is either of this type or of the slightly less open Sb type.



BARRED SPIRAL GALAXY NGC 1300 in Eridanus is classified SBb, which means that it is an intermediate type on the barred-spiral branch of the Hubble sequence. All the photographs shown on these two pages were made with the 200-inch Hale telescope.

from the initial singularity) the universe can be pictured as a composite gas in which some of the "atoms" are particles and the rest are photons. For the universe as a whole there are now at least 10 million times more photons than particles. From thermodynamic considerations one can conclude that photons must also have greatly outnumbered particles in the fireball. For a gas in equilibrium each species of particle contributes to the total pressure in proportion to its number. This still holds (very nearly) for photons, so that the radiation would make an overwhelmingly dominant contribution to the pressure. (During the first 10 seconds, when the temperature exceeds a few billion degrees, the situation is less simple because pairs of photons can interact to form an electron and a positron.)

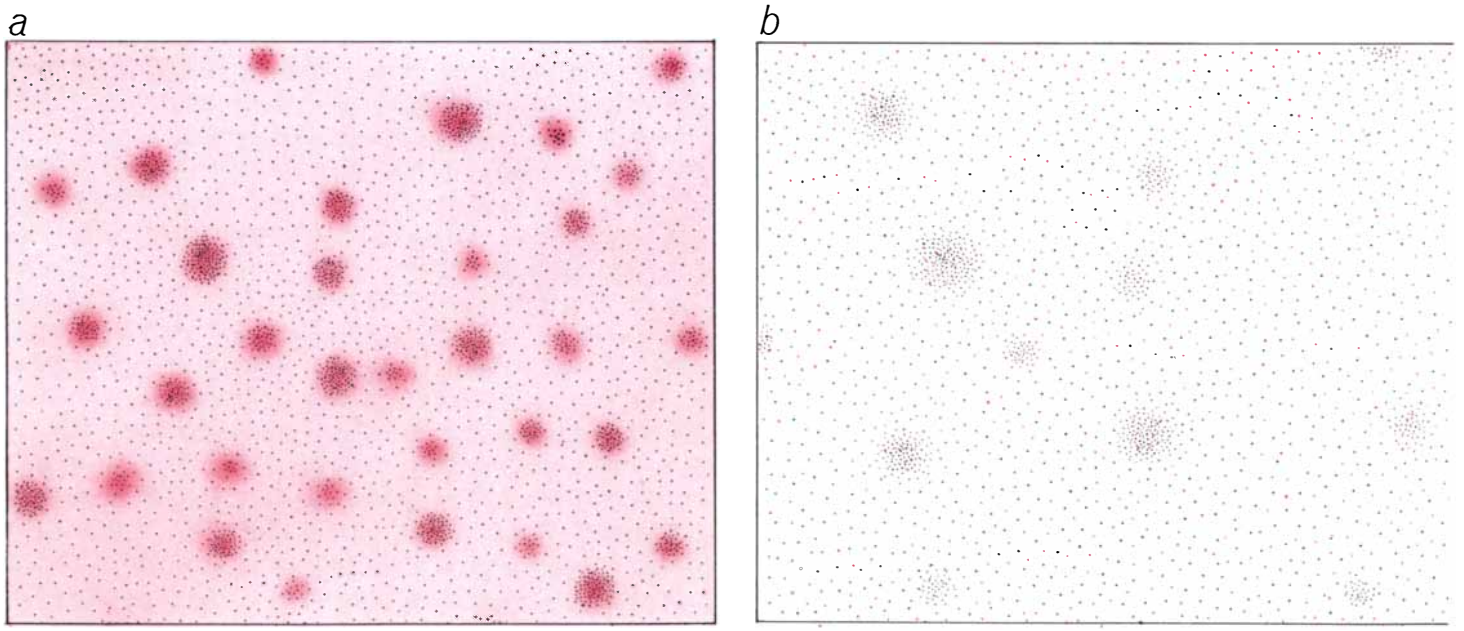
As the expansion proceeds and the density decreases, the photons lose energy, the temperature drops and the particles move less rapidly. A key stage is reached after about 10^5 years, when the fireball has cooled to 3,000 degrees. The electrons are then moving so slowly that virtually all are captured by nuclei and retained in bound orbits. In this condition they can no longer scatter photons and the universe becomes transparent. Inasmuch as the background temperature today is only about three degrees absolute, one can conclude that the universe has expanded by a red-shift factor of 1,000 since the scattering stopped. (Wavelength is inversely proportional to temperature.)

The microwave background photons have probably propagated freely since the universe became transparent and therefore they should carry information about a "surface of last scattering" at a red shift of more than 1,000. Compare this with the red-shift factor of about one-half for the most distant galaxy known! Because these photons have been traveling unimpeded since long before galaxies existed, they should provide us with remarkably direct evidence of physical conditions in the early universe.

Let us return now to the epoch of the primordial fireball and ask: How were inhomogeneities in the fireball affected by the presence of the intense radiation field? Radiation would inhibit the process of gravitational collapse. Under radiation pressure nonuniformities in the fireball would take the form of oscillations, pressure waves or turbulence. These disturbances, in turn, will be dissipated by viscosity and the development of shock waves. Some wavelengths will be attenuated more severely than others, so that inhomogeneities of favored size will be preserved whereas those less favored will tend to be destroyed. The aim of recent work has been to determine what scales of perturbation are most likely to survive the various damping processes until the scattering of photons comes to an end. Any perturbation whose survival and growth is specially favored should eventually dominate, almost irrespective of how nonuniformities were initially distributed in the primordial fireball. An encouraging result that has already emerged from these studies is that 10^{12} solar masses, roughly the mass of a large galaxy, is one such preferred scale [*see illustration on page 34*].

After the electrons in the initial plasma have been bound into atoms, radiation no longer affects the distribution of mass. At this point the surviving perturbations are free to amplify gravitationally. (It should be noted, however, that on small scales—less than 10^6 solar masses—the kinetic energy of atoms exerts a pressure of its own that inhibits gravitational collapse.) The first generation of bound systems will therefore condense from whatever scale of fluctuations had the largest amplitude at the time of decoupling, that is, when the fireball ceased to be a plasma of electrons and other particles.

At what stage did protogalaxies stop expanding and separate out from the rest of the universe? We might guess that this happened when the mean density was comparable to the present density in the outlying parts of galaxies. In 1962 Olin J. Eggen, Donald Lynden-Bell and Allan R. Sandage of the Hale Observatories investigated the likely early history of our own galaxy by studying



FORMATION OF GALAXIES is represented in this sequence of drawings in terms of the "big bang" cosmological model first examined in detail by George Gamow in 1940. For roughly the first 100,000 years after the explosion of the primordial atom the temperature of the expanding fireball was so high that all matter (*black stippling*) was ionized, that is, dissociated into electrically charged particles (*a*). In this situation photons of radiation could not travel very far without being scattered by the free electrons; as a result the universe during this period was effectively opaque (*light shade of color*). Nevertheless, slight random enhancements in the density of matter above the mean density presumably took place, usually accompanied by corresponding enhancements in the

very old stars in the galactic halo. These stars probably formed while the galaxy was collapsing to its present disklike shape (and before the birth of the stars in the Milky Way), and their orbits indicate that our galaxy attained a maximum radius of about 100,000 light-years. One can then tentatively estimate that galaxies such as our own formed when the universe was 1,000 times denser than it is now, about half a billion years after the expansion began.

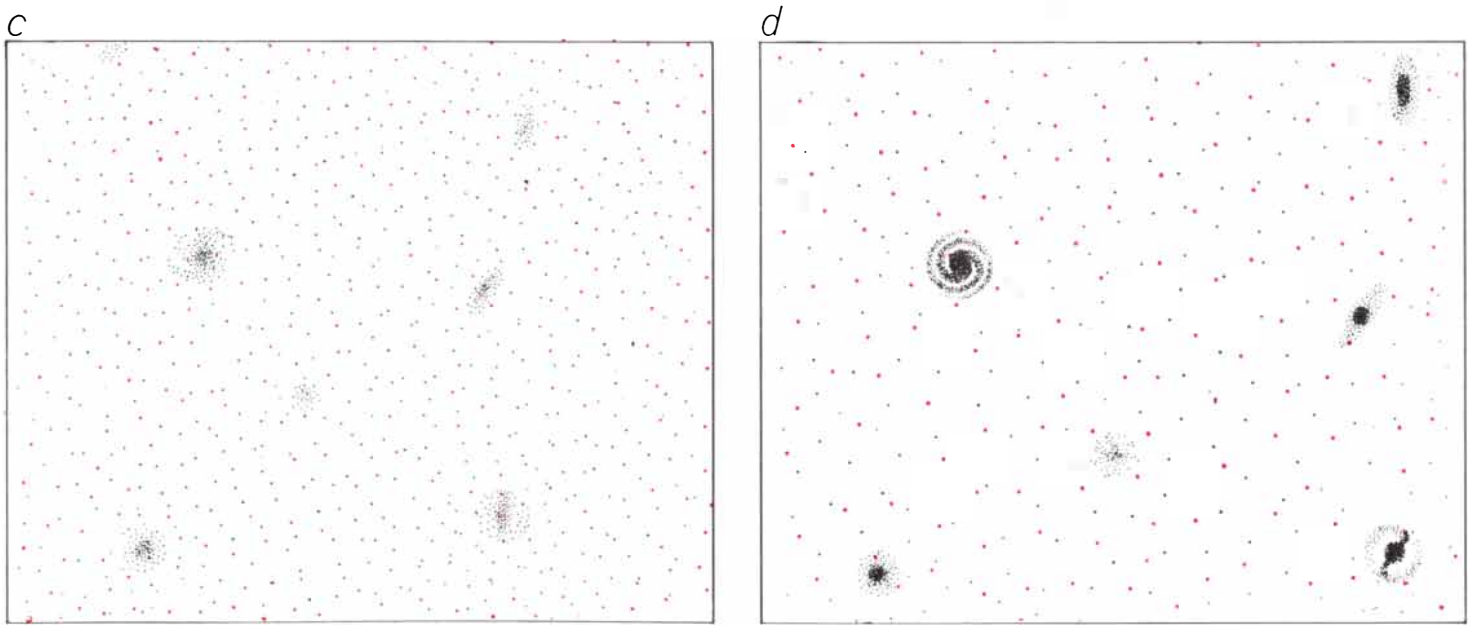
Extrapolating backward in time, we find that the protogalaxies would have taken the form of nonuniformities roughly 1 percent denser than the average density of the universe at the decoupling epoch. It is an attractive possibility that these are the dominant surviving irregularities, all smaller scales having been smoothed out during the fireball phase. There are, however, some types of fluctuation that are not eradicated in the fireball, so that smaller gas clouds may have formed first and later collided and agglomerated into galaxies. Robert H. Dicke and P. J. E. Peebles of Princeton have suggested that globular clusters—compact groups of about 10^5 or 10^6 stars that orbit around galaxies—may represent that small fraction of clouds which managed to avoid collisions, fragmented into stars and survived. Clusters of galaxies would have evolved from initial irregularities of smaller amplitude but larger scale than those destined to form single galaxies.

The only contribution of cosmologists to date toward explaining galaxy formation has been to calculate what scales of perturbation are most likely to survive or amplify in the fireball, thereby reducing the need to build these preferred scales into the initial conditions. This removes one element of arbitrariness in the initial conditions prescribed for the universe. There still remains, however, the task of explaining both the origin of the nonuniformity of the universe on all scales except the very largest, and the apparent uniformity encountered on the largest scales.

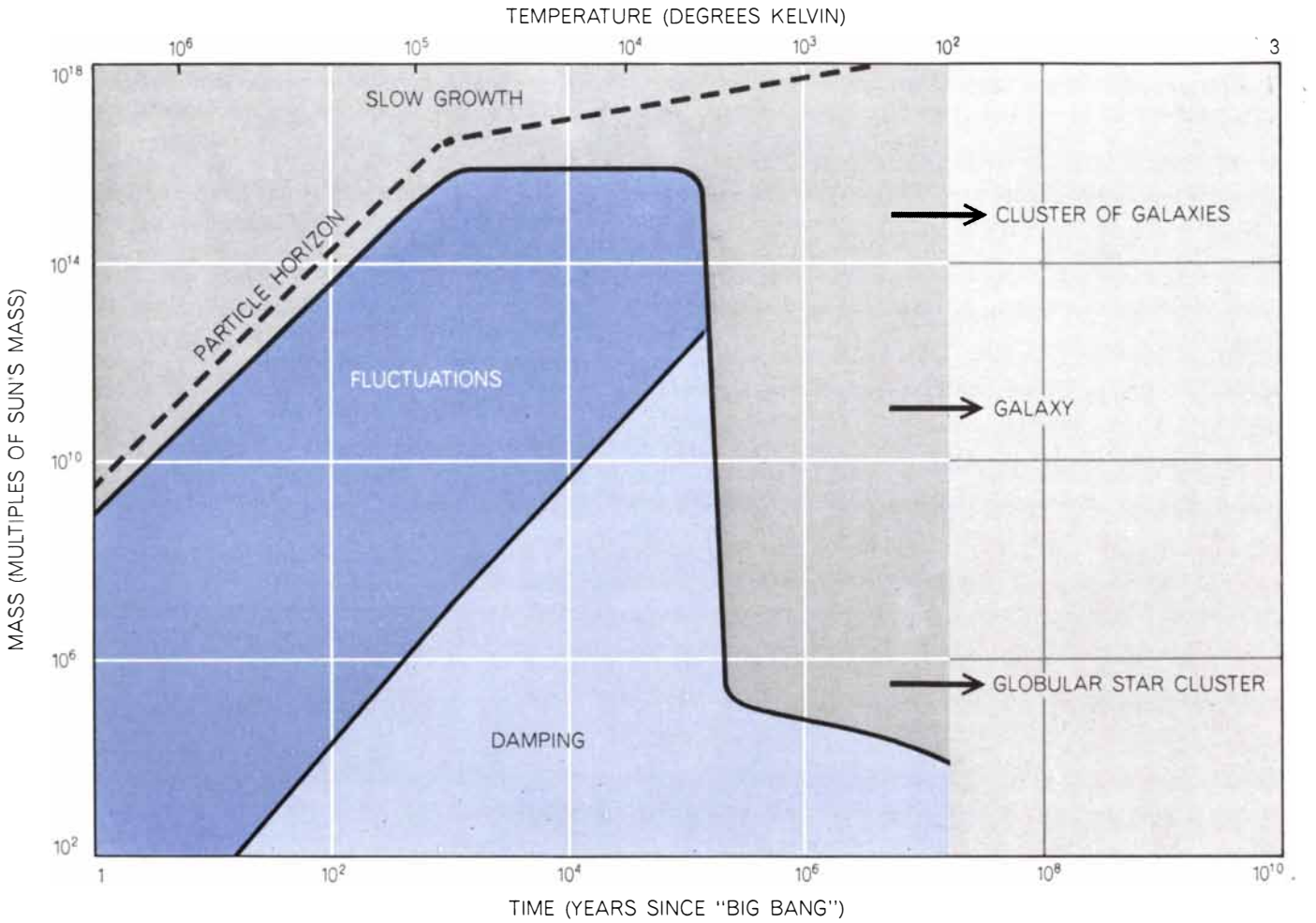
In fact, the Friedmann models may not provide an adequate description of the fireball when large inhomogeneities are present. It would be conceptually attractive if there were processes that could transform an initially chaotic universe into one that displayed the large-scale uniformity of a Friedmann model. An encouraging step toward this goal has been taken by Charles W. Misner of the University of Maryland, who has considered a "mix master" universe, which expands anisotropically in such a way that all parts of the universe are causally related very early in its history. At the outset matter would be so densely packed that even neutrinos would interact with other particles at a significant rate. Acting like a blender, the neutrinos would destroy the original anisotropy of the fireball by the time it had cooled to about 20 billion degrees. Thereafter the expansion would mimic a homogeneous Friedmann model.

Several types of observation may help to test this general picture of galaxy formation. The fluctuations that develop into galaxies and clusters would give rise to random motions on the surface of last scattering. As a result the microwave background photons would not all have been red-shifted by exactly the same amount; in some directions they might have been scattered off material with a random velocity toward us, whereas in other directions the last-scattering surface may have been receding from us. As a consequence the microwave temperature would be slightly nonuniform over the sky. Edward R. Conklin and R. N. Bracewell of Stanford University, Arno A. Penzias, Johann B. Schraml and Robert W. Wilson at the Kitt Peak National Observatory and Yuri N. Parijsky of the Pulkovo Observatory can detect temperature fluctuations as small as a tenth of a percent on angular scales of a few minutes of arc, but so far they have found no positive effect. This technique, however, has the potentiality of detecting embryonic galaxies or clusters of galaxies when they were merely small enhancements above the mean gas density.

There are reasons to expect galaxies that have just condensed to be brighter than typical galaxies at the present epoch. The energy released by the collapse of the protogalaxy would probably have been radiated away by hot gas be-



radiation density (adiabatic fluctuations). In such regions (*dark shade of color*) the radiation tended to damp fluctuations that would lead to further enhancements of matter if they were below a certain critical size (about 10^{11} solar masses). After about 100,000 years, when expanding fireball had cooled to about 3,000 degrees Kelvin, the negatively charged electrons were moving slowly enough to be captured by protons and retained in bound orbits, forming hydrogen atoms. In this condition electrons are much less effective in scattering photons and universe thus became transparent (*b*). Expansion and cooling of fireball continued and matter was progressively concentrated by gravitational forces, first into protogalaxies (*c*) and eventually into galactic types seen today (*d*).



ANOTHER REPRESENTATION of the formation of galaxies is a graph relating the mass of a system to the age and ambient temperature of the universe. Any density enhancement that reaches a minimum value of some 10^{15} solar masses when the universe is about 1,000 years old (the epoch at which the density of matter first equals the mass density of the fireball radiation) has enough gravitational force to overwhelm the effects of radiation pressure. Such an enhancement thereupon enters on a lifetime of slow growth, culminating in a large cluster of galaxies. In an intermediate range (between 10^{11} and 10^{15} solar masses at decoupling) fluctuations in density persist until the decoupling stage is reached and radiation pressure ceases to interact effectively with matter; surviving density enhancements in this range become individual galaxies. Below a certain threshold (10^{11} solar masses at decoupling) radiation pressure damps out most density enhancements. Within this range, however, some density enhancements not accompanied by increases in radiation pressure (isothermal fluctuations) may survive to form globular star clusters, ranging from 10^5 to 10^6 solar masses. "Particle horizon" is boundary of the observable universe, where objects would be receding at speed of light.

fore most of the stars formed. Moreover, the first generation of stars would tend to be heavier and more luminous in relation to their mass than the stellar populations in present-day galaxies. Although most of this energy would be radiated in the ultraviolet, it would be received in the near infrared, owing to the red shift. Robert Bruce Partridge and Peebles at Princeton have suggested that it might be feasible to detect such young galaxies even though these may now have red shifts of about 10.

We are plainly still far from understanding even the broad outlines of the processes whereby the observed aggregations of matter in the universe came into being. We are even further from understanding the detailed morphology of the bewildering variety of different types of galaxies. For example, we have not yet discussed the possible origins of the angular momentum or magnetic fields of galaxies. Peebles has argued that the rotation of galaxies may be induced by tidal interactions soon after formation. Other authors, notably Leonid Ozernoi of the P. N. Lebedev Physical Institute in Moscow, have considered galactic rotation to be of primordial origin. One remarkable feature of the primordial fireball is that it can store rotation in the form of "photon whirls"; subsequently this stored rotation could be transferred to matter whirls.

Galactic magnetic fields may be produced after the formation of the galaxy by a mechanism of the dynamo type. Alternatively, magnetic fields may very well be of primordial origin. Edward R. Harrison of the University of Massachusetts has pointed out that the shear between the photon gas and the matter gas in the fireball could have generated a small magnetic field; if primordial photon whirls are assumed to be present, this mechanism leads to the production of a "seed" magnetic field many orders of magnitude below the value of the magnetic field observed in the spiral arms of our own galaxy. Harrison argues that rapid rotation of the protogalaxy may have subsequently produced sufficient winding of the primordial magnetic field to enhance it by the dynamo mechanism to the field currently observed. Primordial fields alone, he feels, would be insufficient to account for the observed galactic fields. The amount of rotation and the strength of the magnetic field in the protogalaxy probably help to determine whether it will evolve into an elliptical galaxy or into a spiral.

Galaxies are observed to possess random velocities with respect to the cos-

mic expansion. It is a curious coincidence that the rotational velocity of galaxies is of just the same order of magnitude—hundreds of kilometers per second—as these random motions. Perhaps this is simply a consequence of the primordial turbulence, which may have been the source of all structure in the universe.

Present data on the sizes of clusters of galaxies, and on possible "superclusters," are too sparse to enable us to assess the validity of theories that predict the mass spectrum of condensations. Moreover, our knowledge of the masses of galaxies is bedeviled by selection effects. Large and bright galaxies can be seen out to great distances, but small and intrinsically faint ones would only be noticed if they were comparatively close to us. Such objects may therefore occur much more frequently than is believed. A more drastic possibility is that most of the material in the universe may be in some nonluminous form. Evidence for the existence of such material comes from studies of the stability of clusters of galaxies.

This basic problem was first discussed in 1933 by Fritz Zwicky of the California Institute of Technology. For example, if one estimates the mass required to make the Virgo cluster a gravitationally bound system, one finds that the total observed mass in the member galaxies falls short by a factor of 50 or more. One possible way around this paradox is to assume that the Virgo system may be exploding, as the Soviet astrophysicist V. A. Ambartsumian has suggested. Perhaps even more puzzling is the apparent deficiency in mass of the Coma cluster. This system is so spherically symmetric and centrally condensed that astronomers believe it must be a stable system. Yet the observed mass, predominantly in elliptical galaxies, falls short of the mass required for stability by perhaps a factor of five, even if one assumes that the mass-luminosity ratio for ellipticals is around 50.

Similar results have been found for other clusters. Some astronomers have attempted to explain this problem by arguing that nonluminous matter is present in sufficient quantity to stabilize these systems. This material probably cannot all be in gaseous form; neutral hydrogen or ionized hydrogen, whether uniformly distributed or in clouds, ought to be observable either by radio or by X-ray observations.

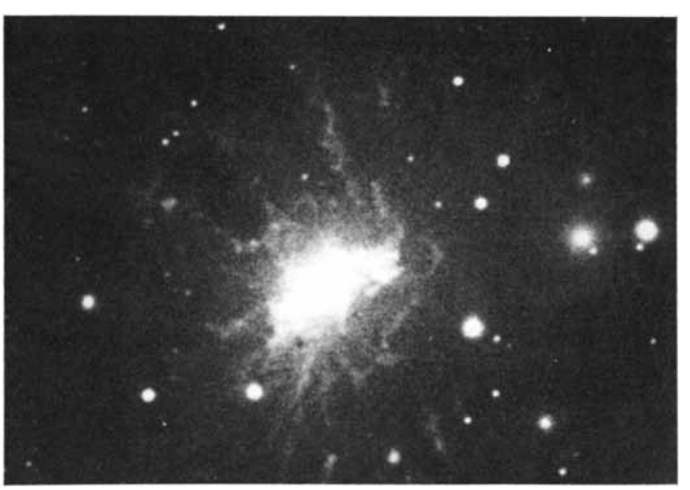
Alternatively, the "missing mass" may be in the form of "dead," or burned-out, galaxies. An even more intriguing possibility is that concealed within the clusters are many objects that have undergone catastrophic gravitational collapse, as predicted by the general theory of relativity. The gravitational field around such objects would be so strong that no radiation could escape from them; only their gravitational influence could be detected by a distant observer.

Other arguments that indicate the apparent youthfulness of some galaxies stem from observations of clusters of galaxies. To be stable, one such chain of galaxies would require a mass-luminosity ratio of more than 5,000, or 100 times as much mass as the cluster seems to possess. One seems forced to the conclusion that here are newly formed galaxies, born within the past 100 million years. Zwicky has discovered an entire class of compact galaxies whose surface brightness resembles that found only in the nuclei of ordinary galaxies. Even more baffling is the discovery that some quasars emit as much radiation as 1,000 galaxies, the energy apparently coming from a colossal explosive event in a region less than 1 percent the size of the solar system. Seyfert galaxies display the same energetic phenomenon on a somewhat reduced scale.

Ambartsumian has long maintained that galactic nuclei are sources of matter and that indeed the galaxies themselves emerge out of dense primordial nuclei. In recent years Halton C. Arp of the Hale Observatories and Erik B. Holmberg of the University of Uppsala have found evidence that small galaxies may even have been ejected from larger galaxies. These phenomena certainly suggest that violent events, involving perhaps the birth of galaxies, are continually taking place in the nuclei of existing galaxies. One is reminded of Sir James Jeans's prescient conjecture, written in 1929, that "the centers of the nebulae are of the nature of 'singular points,' at which matter is poured into our Universe from some other, and entirely extraneous, spatial dimension, so that, to a denizen of our Universe, they appear as points at which matter is being continually created."

Further progress in this field must await fuller information on the distribution, masses and velocities of galaxies. Moreover, satellite observations in infrared, ultraviolet and X-ray wavelengths may soon reveal completely new and unsuspected types of objects, and should in any case give us confidence that we have a fairly complete inventory of the contents of the universe. We shall then be better able to relate theoretical abstractions to the universe in which we dwell.

CHAIN OF GALAXIES VV 172 was photographed by Halton C. Arp with the 200-inch Hale telescope. Four of the galaxies are 600 million light-years away; the fifth appears to be twice as distant. Conceivably it has been ejected from the cluster at high velocity.



EXPLODING GALAXY NGC 1275 was recently photographed in red light by C. Roger Lynds with the 84-inch telescope at Kitt Peak. The radiating filaments of gas, reminiscent of the Crab nebula, were not visible in earlier photographs taken in white light.

Genetic Repressors

Genes do not operate continuously but are switched on and off. One control mechanism is repression. Now the first specific repressors have been isolated, confirming hypotheses put forward a decade ago

by Mark Ptashne and Walter Gilbert

How are genes controlled? All cells must be able to turn their genes on and off. For example, a bacterial cell may need different enzymes in order to digest a new food offered by a new environment. As a simple virus goes through its life cycle its genes function sequentially, directing a series of timed events. As more complex organisms develop from the egg, their cells switch thousands of different genes on and off, and the switching continues throughout the organism's life cycle. This switching requires the action of many specific controls. During the past 10 years one mechanism of such control has been elucidated in molecular terms: the control of specific genes by molecules called repressors. Detailed understanding of control by repressors has come primarily through genetic and biochemical experiments with the bacterium *Escherichia coli* and certain viruses that infect it. In this article we shall first outline the current view of the action of repressors and then report some of the experimental evidence, involving the isolation of specific repressors and the description of how they work, that supports this picture.

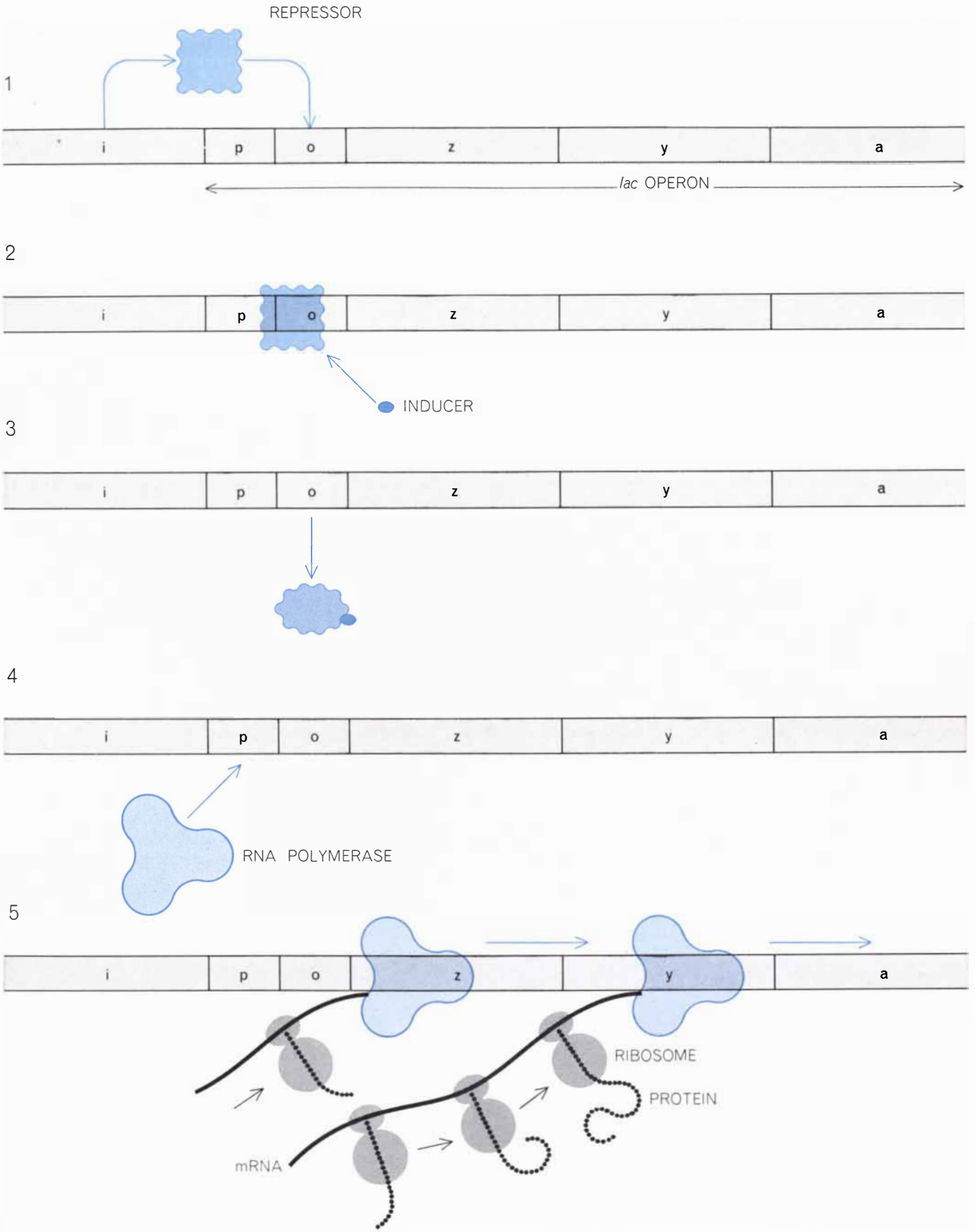
A gene is a particular sequence of subunits called bases, ranged along a molecule of deoxyribonucleic acid (DNA), that is ultimately translated into a corresponding sequence of amino acids constituting a molecule of protein [*see illustration on opposite page*]. An enzyme, RNA polymerase, first transcribes the DNA sequence into a corresponding sequence of bases of ribonucleic acid (RNA), beginning at a specific site termed the promoter. This messenger RNA molecule then becomes attached to the particles called ribosomes, where its genetic information is translated into protein. For certain genes this process may be controlled by the intervention of a protein, a repressor, the product of a separate control gene. The repressor binds, or attaches, directly to the DNA molecule at the beginning of the set of genes it controls, at a site called the operator, preventing the RNA polymerase from transcribing the gene into RNA and thus turning off the gene. Each set of independently regulated genes is controlled by a different repressor made by a different repressor gene.

The repressor determines when the gene turns on and off by functioning as an intermediate between the gene and an appropriate signal. Such a signal is often a small molecule that sticks to the repressor and alters or slightly distorts its shape. In some cases this change in shape renders the repressor inactive, that is, no longer able to bind to the operator, and so the gene is no longer repressed; the gene turns on when the small molecule, which here is called an inducer, is present. In other cases the complex of the repressor and the small molecule is the active form; the repressor is only able to bind to the operator when the small molecule (here called a corepressor) is present.

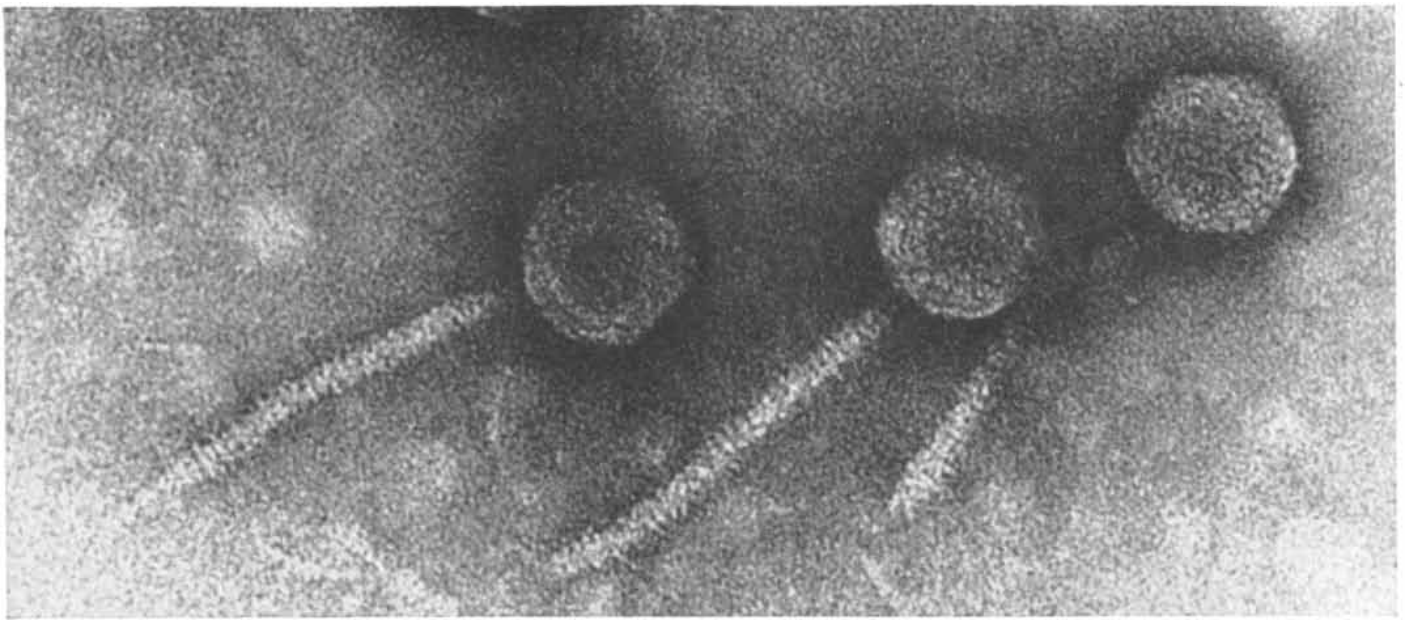
For example, consider the *E. coli* gene that codes for the enzyme beta-galactosidase. If the bacterium is to grow with lactose (milk sugar) as its source of carbon, this enzyme must be produced in order to split the sugar into its two components, galactose and glucose. In the absence of lactose the beta-galactosidase gene is not needed, and it is switched off by a specific repressor, the *lac* repressor. In the presence of lactose the required enzyme is synthesized because a breakdown product of the sugar acts as an inducer: it attaches directly to the repressor, causing the repressor to release the DNA and allow the synthesis of the messenger RNA corresponding to the beta-galactosidase gene. When the lactose supply is exhausted, the repressor is freed of the sugar inducer and once again switches off the beta-galactosidase gene.

According to this picture the lactose repressor is a protein molecule that interacts (at two different sites) with two entirely different kinds of molecules. One site recognizes the unique sequence of DNA bases that constitutes the operator; the other binds the inducer (the lactose derivative). The binding of the inducer to the repressor induces a conformational change in the repressor that prevents its sticking to the operator [see "The Control of Biochemical Reactions," by Jean-Pierre Changeux; SCIENTIFIC AMERICAN, April, 1965].

Much of the genetic basis of the picture we have just described was presented in 1961 by François Jacob and Jacques Monod of the Pasteur Institute in Paris. One of their most illuminating examples was an analysis of the complete set of genes for the metabolism of lactose by *E. coli*, the *lac* system. Genetic and biochemical experiments showed that two kinds of genes are involved in the utilization of this sugar. There are three genes of one kind, one coding for beta-galactosidase (the *z* gene), one for a permease protein that concentrates lactose in the cell (the *y* gene) and one for a transacetylase enzyme with an unknown function (the *a* gene). Mutations in any one of these three genes change the structure of the corresponding enzyme but have no effect on the regulation of their synthesis. Mutations in the other kind of gene fail to change the structure of these enzymes but do affect their regulation. These mutations affecting regulation are located in



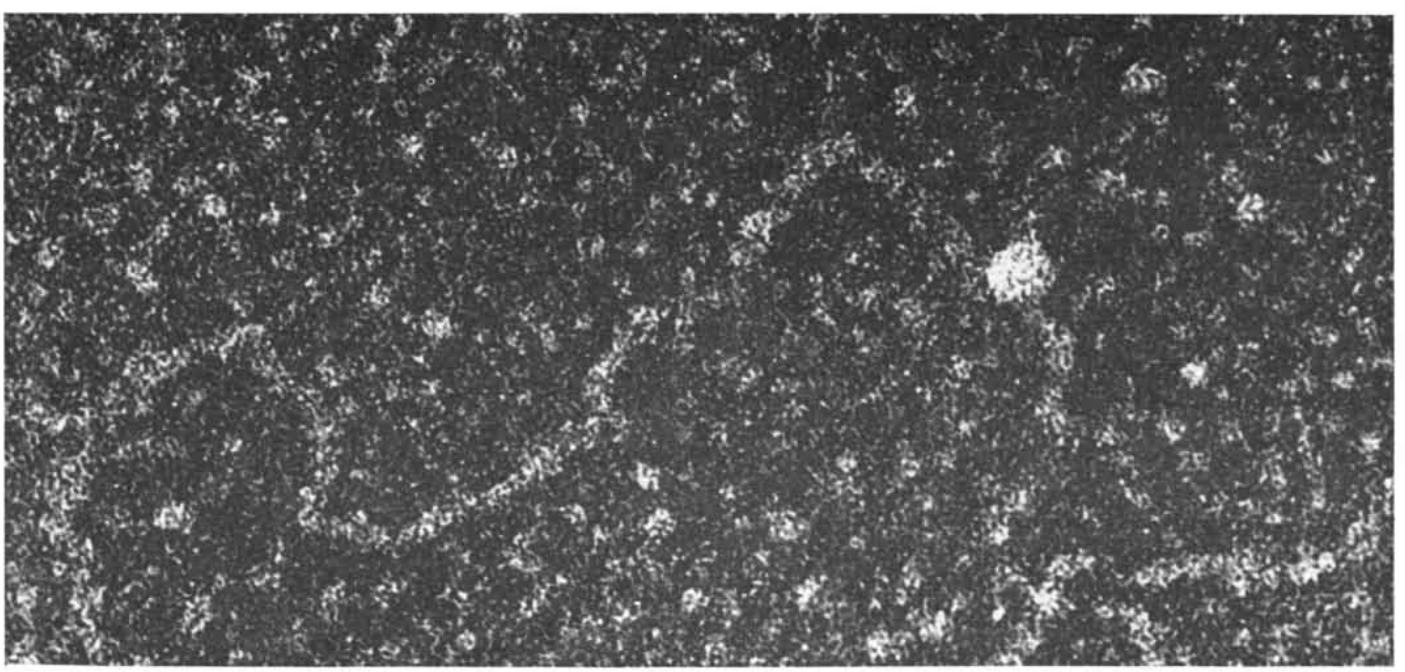
REPRESSOR FUNCTION is illustrated for the *lac* repressor, which controls genes that provide for the metabolism of lactose in the bacterium *Escherichia coli*. The repressor, produced by the *i* gene, binds to the DNA at the operator site (*o*), thus preventing the transcription of the *z*, *y* and *a* genes into RNA (*1*, *2*). In the presence of an inducer (a lactose derivative), which attaches to the repressor and changes its shape, the repressor releases the DNA (*3*), switching the gene on. Beginning at the promoter (*p*), RNA polymerase (*4*) transcribes the DNA into messenger RNA, which is translated into protein on particles termed ribosomes (*5*).



PHAGE LAMBDA, a virus that infects *E. coli,* is another organism whose genes are controlled by a repressor. The phage consists of a DNA molecule wrapped in a protein head to which a long tail is attached. In the electron micrograph, made by L. Kemp and A. F. Howatson of the University of Toronto, the lambda particles are enlarged 350,000 diameters and stained with uranyl acetate.

a single gene, the *i* gene. Typical *i* gene mutations studied by Jacob and Monod caused continual synthesis of *lac* enzymes, even in the absence of lactose.

What was the nature of this control? It might be positive, that is, the product of the *i* gene might keep the *lac* genes turned on. Or it might be negative: the *i* gene product might somehow turn the *lac* genes off. The actual nature of the control can be shown by selectively eliminating the *i* gene product from the cell by introducing a mutation that deletes the *i* gene. In such a mutant the *lac* enzymes are synthesized at the full rate whether or not the inducer is present. If the absence of the *i* gene results in continual synthesis, the *i* gene's control must be negative in character. This is confirmed if several different pieces of DNA carrying the *lac* genes are introduced into a cell, some with and some without the *i* gene. No *lac* enzymes are synthesized, indicating that the presence of any *i* gene is enough to turn off the *lac* genes on all the pieces of DNA, and therefore showing again that *i* gene control is negative. Jacob and Monod proposed the simplest possible mechanism based on this evidence: that the *i* gene product acts as a repressor—a substance that diffuses through the cell and shuts



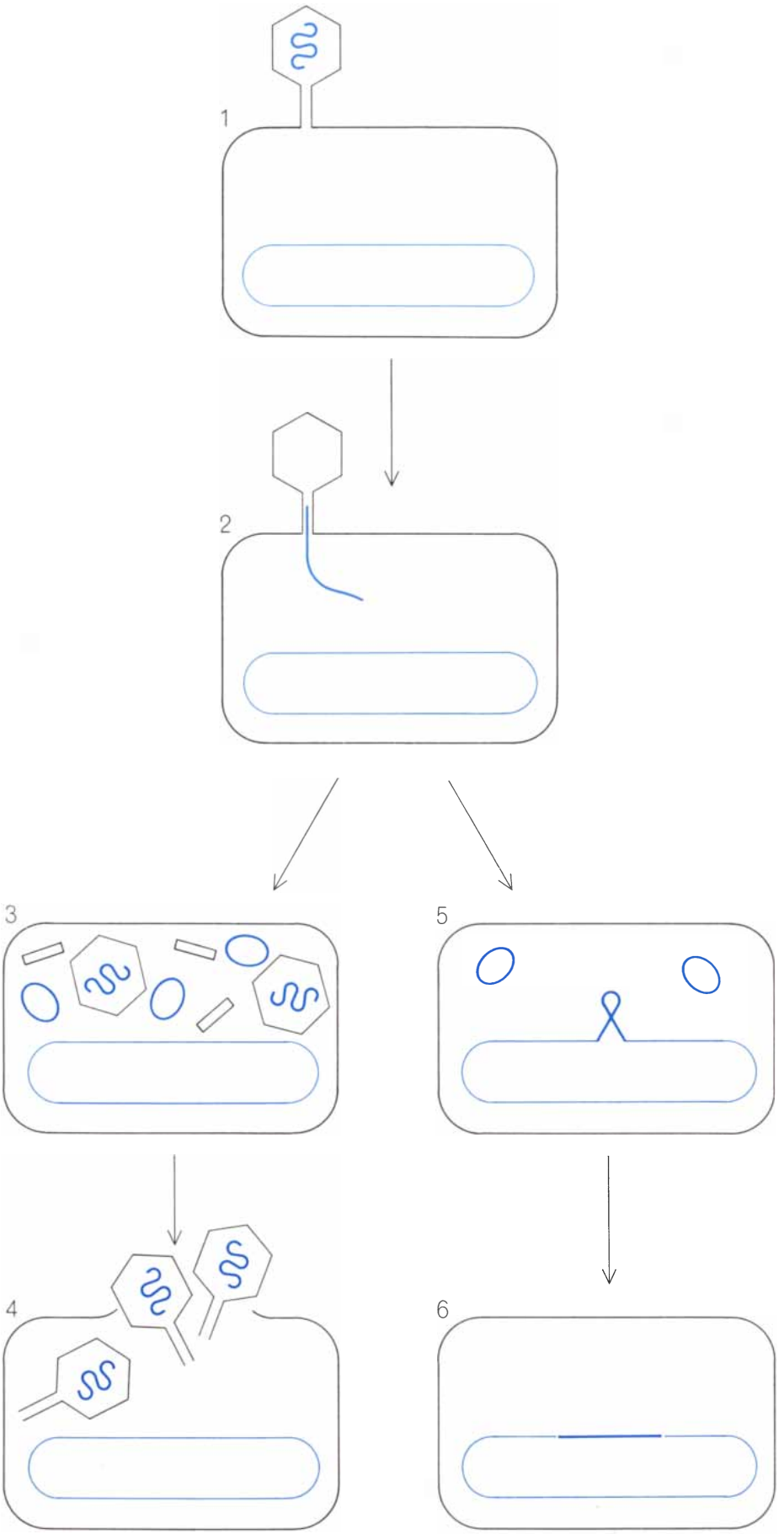
LAC REPRESSOR MOLECULE (*large white object*) sits on DNA in an electron micrograph made by Jack Griffiths of Cornell University. Purified *lac* repressor was put into a solution with DNA bearing the *lac* operon. When DNA molecules were caught on a carbon film and were shadowed by rotation in tungsten vapor so as to build contrast evenly on all sides, repressor molecules were seen bound to the DNA. The molecules have been enlarged 500,000 diameters. The granular background texture is due to the shadowing.

off any *lac* genes that are present, not just the *lac* genes on the same piece of DNA.

Jacob and Monod went on to predict and find mutations that render the *lac* genes insensitive to the *lac* repressor—mutations that cause continual enzyme synthesis even in cells known to contain active repressor. These mutations, which are located on the bacterial DNA just before the beginning of the *z* gene, define the genetic determinant of the repressor's site of action. Note that one must distinguish between a genetically defined region of DNA, which can harbor mutations affecting a certain function, and the ultimate product of that region, which actually carries out the function. The genetic experiments could show only that mutations in an operator gene locus (*o*) render inactive some structure (the operator) that otherwise would interact with the repressor. The promoter, the operator locus and the three genes controlled by the *lac* repressor constitute what is called the *lac* operon.

Several other sets of genes controlled by specific repressors have been studied in *E. coli*. One that has played an important role in illuminating the mechanism of repression is not properly a group of bacterial genes at all but rather the genes of a bacterial virus, phage lambda, that infects *E. coli*. The phage particle consists of a single DNA molecule enclosed in a protein coat with a tail through which the DNA is injected into a bacterial cell. Once inside the cell the DNA can follow either of two courses [*see illustration at right*]. It can complete the ordinary course of infection, utilizing the cell's machinery to replicate its own virus DNA and wrap it in a protein coat, producing large numbers of progeny phage that lyse, or burst, the bacterium and go on to infect more cells. Alternatively, instead of completing this lytic cycle, the phage genes are switched off soon after they enter the cell. Then the dormant phage DNA becomes integrated into the host cell's chromosome and remains there inertly, replicating with the bacterial DNA and giving rise to a population of cells each of which contains a dormant phage chromosome, or prophage.

What shunts an ordinarily infectious phage into the dormant, or lysogenic, state? Genetic investigations conducted by Jacob, by A. Dale Kaiser of Stanford University and by others showed that the phage's ability to turn its genes off is conferred by a single phage gene designated C_I. Experiments and arguments



INFECTION can take either of two courses after the phage becomes attached to the bacterium (*1*) and injects its DNA (*2*). In lysis (*left*) the viral DNA (*thick colored line*) replicates and forms new particles (*3*), which burst the cell (*4*). In lysogeny (*right*) most of the viral genes are switched off. Instead of replicating, the viral DNA integrates into the bacterial DNA (*thin colored line*) by crossing over (*5*) and remains there, dormant (*6*).

strictly analogous to the ones that elucidated the *lac* system showed that the C_I gene makes a repressor, the lambda phage repressor, that switches off certain genes of the phage. Treatments (such as irradiation with ultraviolet) that turn the genes on again and restore the phage to the lytic state are believed to act by leading to the production of a small molecule—an inducer—that inactivates the phage repressor. Recent experiments have shown that the lambda repressor acts at two separate operator sites on the phage chromosome, controlling two different sets of genes that are read in opposite directions on the DNA [*see illustration below*]. The "early" genes directly controlled by the repressor include genes whose products are necessary to turn on the "late" phage genes. The mechanism of this positive control is not yet known.

The presentation of the repressor theory in 1961 caused great excitement among biologists because it proposed to explain gene control, which had been regarded as an extremely complicated matter, in terms of just three elements: the repressor, the operator and the inducer. The most important aspect of the model was the proposition that the *i* gene product or the C_I gene product functions directly to turn off its target genes without the intervention of other large molecules. A counterhypothesis would be, for example, that the *i* gene product is actually an enzyme that is involved in the formation of the ultimate repressor. Although the genetic evidence was entirely consistent with the simpler model, it could not exclude this more complicated possibility. Furthermore, the genetic analysis did not reveal the molecular nature of either the repressor or the operator. It was suggested that the repressor might consist at least partly of RNA, and in fact experimental evidence, now known to be incorrect, was presented to support this notion.

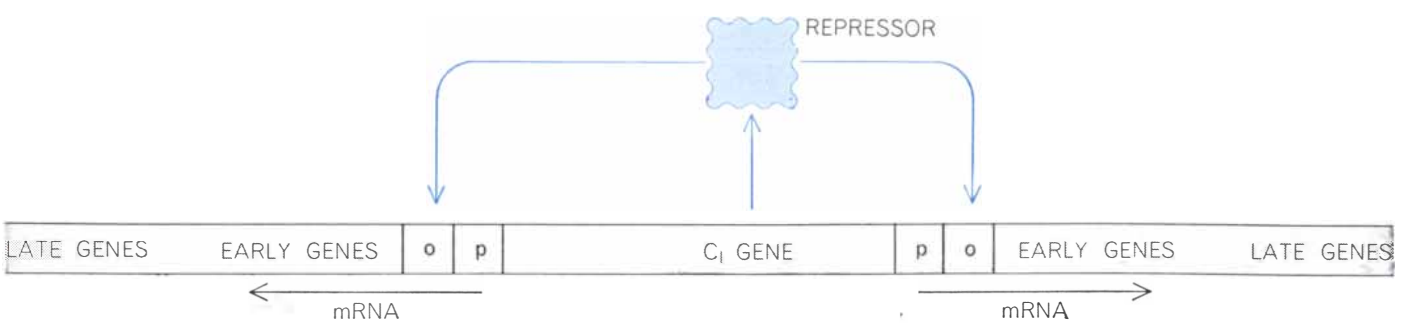
The question of the molecular nature of the operator was even more intriguing. One model had the repressor binding to the operator sequence on the DNA and directly blocking the transcription from DNA to RNA. It was possible, however, that the repressor had its effect at some other stage of the process. For example, if the operator were transcribed into RNA at the head of the messenger RNA molecule, the repressor might function by binding to this operator sequence on the messenger RNA and blocking the translation into protein. The results of experiments performed in many parts of the world in the five years following the introduction of the repressor theory made it clear that the only way to answer these questions was to isolate a genetically defined repressor and to show how it worked in the test tube.

In the summer of 1965 we began experiments at Harvard University to try to isolate specific repressors. One of us (Gilbert), in collaboration with Benno Müller-Hill, a postdoctoral fellow from Germany, attacked the *lac* repressor. The other (Ptashne) began a search for the lambda phage repressor aided by Nancy Hopkins, now a graduate student at Harvard. The actual experimental approaches used to isolate these two repressors were entirely different, although the same major problems were faced by both groups.

The problems were both scientific and psychological. We had no biochemical assay for repressors and could not devise a functional assay because we did not know how they would function. We suspected (correctly, as it turned out) that repressors are present only in extremely small quantities in the cell. (The average *E. coli* cell contains only about 10 or 20 molecules of the *lac* repressor, constituting only .002 percent of the cell's protein.) We did not even know whether to look for repressors among proteins, nucleic acids or other substances. Since there was no way to assay function, we had to surmise some other property that would give us a foothold, but not until the experiments were completely successful could we know that our working hypothesis was relevant. In fact, we could not be sure that the experiments would have any outcome at all. It was possible that we would search but not find anything; such a failure would not prove the negative (that repressors did not exist) but would simply mean that the question was still open. Even if we could isolate the product of the repressor gene, we might still fail to understand how that product functioned.

Gilbert and Müller-Hill based their experiments on the simplest interpretation of the induction phenomenon of the *lac* operon: that a lactose-like inducer inactivates the repressor by binding directly to it. The inducer they used was isopropyl-thio-galactoside (IPTG), a substance that resembles the sugar inducer in molecular form and that is the most active experimental inducer known for the *lac* genes. They reasoned that they might detect the presence of repressor molecules in solution by the ability of these molecules to bind radioactively labeled IPTG molecules. Hoping to measure the extent of this binding (and thus establish a method of assay for the repressor), they resorted to the technique of equilibrium dialysis.

A concentrated extract of *E. coli* cells was placed in a bag made from a cellulose membrane. The membrane has pores that are small enough to prevent the passage of typical protein molecules but large enough to allow water and IPTG molecules to pass through freely. The bag was floated in water containing radioactively labeled IPTG molecules. Since the IPTG molecules could pass through the pores in the bag, the concentration of these molecules free in solution would soon become equal inside and outside the bag. If, however, there were large molecules inside the bag that would bind IPTG, then the total concentration (free plus bound) inside the bag



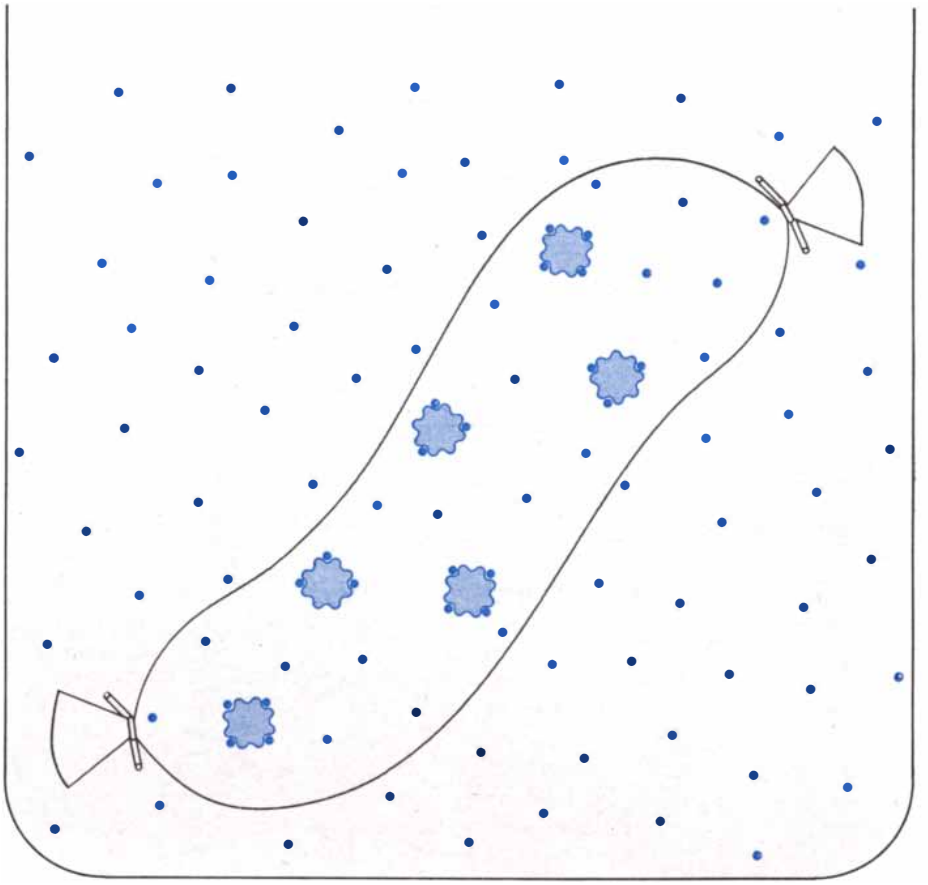
LYSOGENY of the lambda phage is brought about by a repressor, a product of the C_I gene on the viral chromosome mapped here. The repressors bind to two different operators, blocking the transcription of two sets of "early" genes read in opposite directions.

would come to be greater than the concentration outside. Since the IPTG was radioactive, even a small difference in concentration would be revealed by comparing the radioactivity of samples taken from inside the bag and from the surrounding solution.

After nine months of false starts and failures, a way was finally found to prepare a highly concentrated extract of bacterial cells that would draw a tiny excess of radioactivity into the bag. That excess became the basis of the needed assay. After that it was a straightforward matter to fractionate the bacterial proteins and, by testing each fraction separately for the ability to bind IPTG, to purify the substance that could bind the inducer.

Was this stuff indeed the repressor? One critical test involved strains of cells that in genetic experiments made no repressor. These cells turned out not to provide protein that would bind IPTG in the dialysis experiment. An even more convincing proof was based on the existence of cells with special mutations in the *i* gene that increase the affinity of the repressor for the inducer in the living cell. Extracts from these cells were shown to contain molecules that bound radioactive IPTG more tightly in the dialysis assay.

Meanwhile the search for the lambda repressor was proceeding along different lines. Ptashne's strategy was to attempt to create a situation in which the only protein synthesized by the cell would be the lambda repressor. Then radioactive amino acids, the protein precursors, could be incorporated in the repressor only, making the properties of the repressor molecule easy to study. To approach this ideal situation he tried first to stop all protein synthesis directed by the cellular DNA and second to stop the functioning of all the phage genes except the C_I gene itself. The first requirement was met by exposing *E. coli* cells to massive doses of ultraviolet radiation, which so damages the DNA that it cannot be copied into RNA, halting the further synthesis of bacterial proteins. The cells' protein-synthesizing machinery remains functional, however, so that when phages inject undamaged DNA into irradiated cells, phage proteins can be synthesized. In order to meet the second requirement the experiment was done with *E. coli* cells that contained a special lambda prophage, one whose repressor is not inactivated by ultraviolet. After irradiation this repressor shuts off all the genes on the incoming phage except for the repressor gene itself.



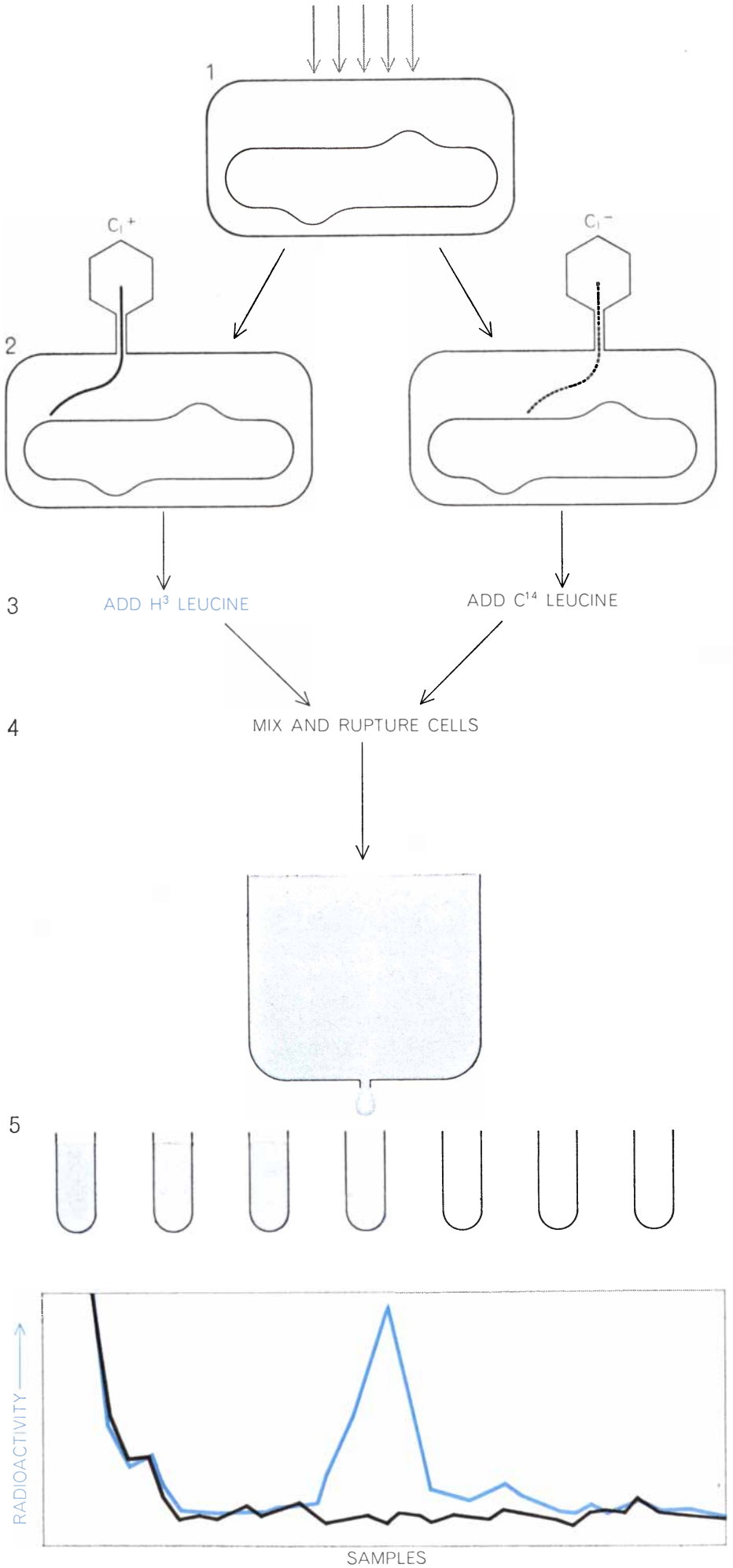
***LAC* REPRESSOR was identified by its ability to bind molecules of IPTG, an effective inducer. An *E. coli* extract was put inside a bag made of a semipermeable membrane. The bag was floated in a solution containing radioactive IPTG (*colored dots*). The IPTG diffused through the membrane, making the concentration of free IPTG the same inside and outside the bag. Repressor molecules (*light color*) in the bag bound some IPTG molecules, however. The excess of radioactivity inside the bag measured the extent of the binding and provided an assay for purifying the repressor: the protein that was binding IPTG.**

If these stratagems worked perfectly, only the repressor would be synthesized in the irradiated and infected cells. In reality the procedures are quite imperfect, and so repressor synthesis accounts for only 5 to 10 percent of the total residual protein synthesis. An experiment using two different radioactive labels was therefore necessary in order to pick out the repressor from among the other labeled proteins [*see illustration on next page*].

The irradiated cells were divided into two portions. One was infected with lambda phages carrying an intact repressor gene and the other with phages bearing mutations that blocked the synthesis of the repressor. Both cultures were fed radioactive amino acids. The cells with an intact repressor gene got amino acids containing radioactive hydrogen, or tritium (H^3), whereas the cells that could not make repressor got amino acids containing radioactive carbon (C^{14}).

The beta particle emitted when the tritium nucleus decays has only a tenth the energy of the beta particle emitted by the carbon. This energy difference is easily detectable, so that proteins made from amino acids labeled with H^3 can be distinguished from those synthesized from amino acids labeled with C^{14}. Ptashne mixed the labeled cells, extracted their proteins and spread them out on a chromatographic column, which separates proteins from one another. In the pattern of radioactivity corresponding to the proteins synthesized in both kinds of cells, and thus labeled with both isotopes, one peak appeared that was labeled only with H^3—not with C^{14}. The H^3-labeled protein, made in one culture but not in the other, was presumably the repressor. This was confirmed in a second experiment when another phage, bearing a mutation in the C_I gene that modifies but does not eliminate the repressor, caused the synthesis of a protein with properties slightly different from those of the unmutated repressor.

Many biologists had expected that repressors would turn out to be similar to histone proteins, which are found



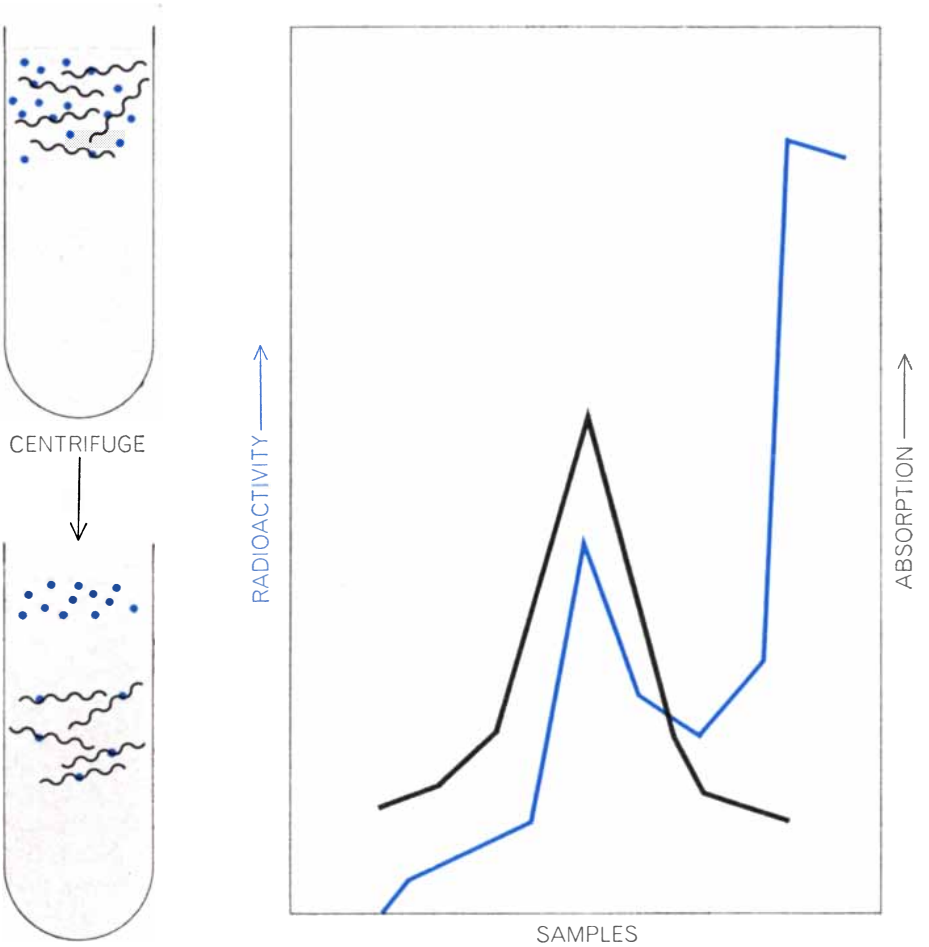
in conjunction with DNA in the chromosomes of higher organisms. The histones are very small proteins characterized by a large excess of positive charges carried on their molecules at a neutral *p*H. Both the lambda and the *lac* repressor are weakly acidic, however, meaning that at neutral *p*H they have an overall negative charge. Furthermore, they are large proteins. The *lac* repressor has four identical subunits, each with a molecular weight of 38,000. Although the lambda repressor was originally isolated as a single polypeptide chain with a molecular weight of 28,000, we now know that the functional structure is a complex of four of these subunits. (Vincenzo Pirrotta, working in Ptashne's laboratory, has recently isolated another phage repressor that behaves like the lambda repressor, although it bears a slight overall positive charge at neutral *p*H.) Some biologists had thought repressors would contain RNA, which would allow them to recognize other nucleic acids by base pairing, but both the *lac* and the lambda repressor turned out to be quite ordinary proteins, with no striking properties other than their action.

Having isolated and purified these two products of control genes, we were still faced with fundamental questions. Were these substances in fact repressors in the sense that they would interact directly with their respective operators? And what, in fact, was the operator? One simple version of the Jacob-Monod theory was that repressors might bind directly to the operator region on a DNA molecule and so block the transcription into RNA. We verified this notion by examining directly the interaction between isolated repressor and purified DNA in the test tube. The first experiment involved mixing radioactive lambda repressor with DNA purified from lambda phage particles to see if the radioactive repressor would stick to the DNA and sediment with it in a centrifuge [*see il-*

LAMBDA REPRESSOR was identified as shown at left. *E. coli* cells were irradiated with ultraviolet to halt bacterial protein synthesis (*1*). The cells were divided (*2*) and half were infected with phages producing repressor (*left*), half with phages that could not produce repressor (*right*). The former were fed amino acids (*3*) labeled with radioactive hydrogen (H^3), the latter amino acids labeled with radioactive carbon (C^{14}). The cells were mixed, their proteins were extracted and separated by chromatography (*4*) and the protein fractions (*5*) were tested for radioactivity. One protein (*peak in colored curve*) appeared that was labeled only with H^3. This was the repressor.

lustration at right]. A sample of the mixture was layered on top of a solution containing a gradient of sucrose and then spun at high speed in an ultracentrifuge. (The sucrose gradient, a solution that is denser at the bottom than at the top, simply prevents mixing and allows large molecules to move individually down the tube in the centrifugal field, which is several hundred thousand times normal gravity.) The DNA was much more massive than the repressor and would sediment faster. If the repressor was bound sufficiently tightly to the DNA, it would move at the high rate characteristic of the DNA; otherwise the repressor would remain near the top of the tube. When samples from different levels of the centrifuge tube were examined, it was clear that the radioactive lambda repressor did bind to lambda DNA. On the other hand, it did not bind to DNA isolated from a different phage.

Similarly, the *lac* repressor binds to DNA that contains the *lac* genes but not to DNA without *lac* genes. For these experiments the necessary radioactive repressor was made by purifying the *lac* repressor from bacterial cells that were grown on radioactive nutrients so that their proteins were labeled. The *lac* DNA was obtained not from *E. coli* but from a special phage that had been bred (from a phage very similar to phage lambda) by Jonathan R. Beckwith and Ethan Signer, who were then working at the Pasteur Institute. This phage carried the *lac* genes in place of some of the phage genes; thus it provided us with a conveniently small DNA molecule carrying the *lac* genes and in effect with a concentrated source of *lac* genes.



INTERACTION BETWEEN repressor and DNA was demonstrated by centrifugation. Lambda DNA and radioactive lambda repressor were layered on a sucrose gradient and spun in an ultracentrifuge. The DNA, being heavier, sedimented more rapidly than the free repressor. Much of the repressor, however, sedimented with the DNA, indicating that it had bound to the DNA. This was shown by the fact that a peak of radioactivity (*color*) coincided with the DNA peak (*black*). The remaining free repressor formed a second peak.

How specific is the repressor-DNA interaction? To verify the hypothesis we must know that the presumed repressor molecule interacts specifically with a unique site along the DNA. To show this we turned to mutant DNA's: molecules isolated from phages with mutations in their operators. The lambda repressor fails to stick to such mutant DNA's if both phage operators are damaged. If the *lac* operator is damaged (that is, if the DNA carries a mutation in the operator region), the *lac* repressor no longer binds. Such experiments show that these repressors do indeed interact with unique sites on the DNA molecules. If mutation alters even one base in the binding site, the protein molecule can no longer bind.

The action of the inducer of the *lac* operon can also be studied in the test tube. After radioactive *lac* repressor has been bound to *lac* DNA, the inducer IPTG is added. The repressor then releases the DNA and, in the type of experiment we have described, no longer sediments with the DNA.

An ingenious alternative method of observing the interaction of the repressor with DNA was developed by Suzanne Bourgeois and Arthur D. Riggs of the Salk Institute for Biological Studies. They observed that when a solution of DNA molecules mixed with repressor was filtered through a cellulose nitrate filter, the repressor molecules would stick to the filter substance. If the molecule had previously been bound to DNA, the repressor molecule, in sticking to the filter, held the DNA on the filter. (The holes in the filter are gigantic compared with the repressor molecule, and they are too large to retain even the DNA alone.) This sticking property provides a simple assay of extraordinary sensitivity for the repressor-DNA interaction: one puts a radioactive label into the DNA and asks whether or not the label is bound to the filter. The *lac* repressor must stick to the filter in an active form because, when the filter is washed with a solution of IPTG, the inducer interacts with the repressor and releases the DNA, leaving the repressor bound to the filter. Miss Bourgeois and Riggs have used this assay to study many of the quantitative aspects of the binding of the *lac* repressor to DNA.

The DNA binding experiments show directly that the operator, defined functionally as the physical target for the repressor, is a stretch of bases along a DNA molecule. We infer that when the repressor is bound to the operator, it prevents the RNA polymerase from initiating transcription. This has been confirmed recently for phage lambda: purified phage repressor directly blocks the copying of the two operons into RNA. Such a final demonstration has not yet been achieved for the *lac* repressor.

The mechanism by which the repressor recognizes the appropriate DNA sequence is not yet understood. It is clear

that the interaction is with the native double strand of DNA, because if the strands are separated, or denatured, by heating, the repressor will not bind. Part of the interaction is presumably an attraction for the phosphates of the DNA backbone; although the repressors themselves have no strong overall charge, they have regions containing positively charged groups that would attract them to the negatively charged phosphates. In order to tell one sequence from another along the DNA molecule so as to bind at a specific site, however, the protein molecule must interact in some way with the bases in order to read the sequence information. It may be, for example, that the repressor forms hydrogen bonds with the outside of the bases, in the big groove of the DNA double helix, and in this way "sees" enough differences in the shapes of the bases to recognize the specific region.

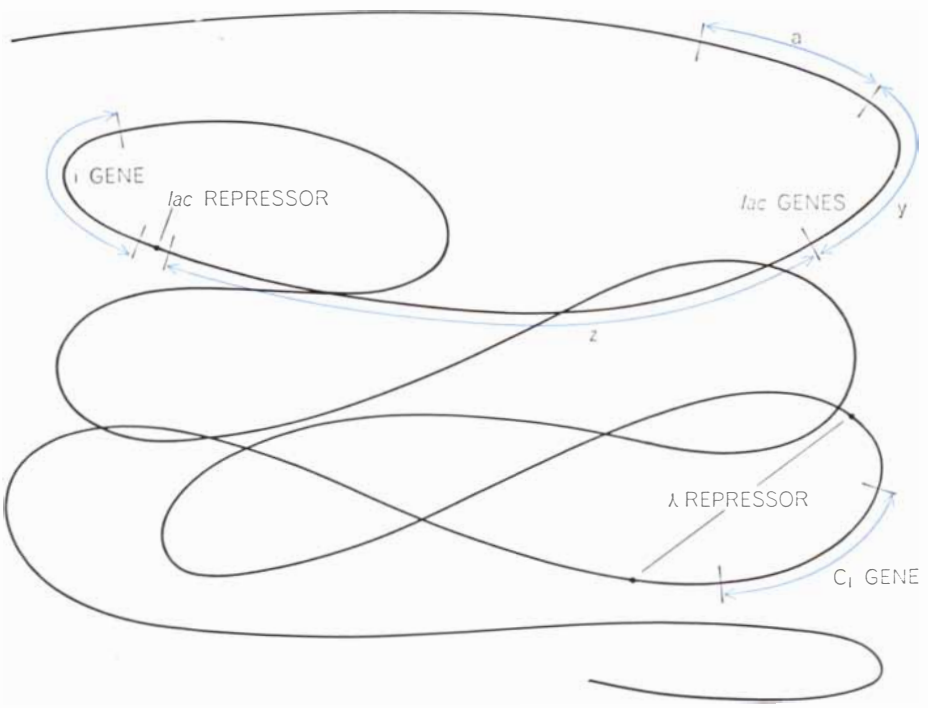
How large is the operator? There is no direct evidence as yet, but a likely guess is that it is on the order of 12 bases long. (This is the shortest sequence that could be unique if the three million base pairs of the *E. coli* chromosome are considered as though arranged at random.) Such a sequence would be 40 angstroms long, a little more than one turn of the DNA helix and a short enough region to fit easily against a molecule of the repressor's size.

The gene for the *lac* repressor functions at a very low rate, producing only 10 molecules of repressor during the time between cell divisions. On the other hand, a very active gene in *E. coli,* such as the gene for beta-galactosidase, can make 10,000 molecules in each generation. Could one alter the rate at which the repressor gene functions? This was achieved by Müller-Hill, who found mutant forms of the *lac* repressor gene that make tenfold more repressor. Further mutations, found by Jeffrey Miller of the Harvard Medical School, raised this another five times, so that the final form of the gene works 50 times as rapidly as the unmutated form. We believe these mutations are changes in the promoter for the repressor gene that somehow make the promoter more active. In order to obtain still more repressor these superactive forms of the gene were introduced into the special *lac*-containing phage DNA we mentioned above. When this phage DNA multiplies inside a cell, many copies of the repressor gene are synthesized and there is a further twenty-fold increase in repressor synthesis. The overall thousandfold increase means that the *lac* repressor is now one of the more easily obtainable proteins.

Control by repressors will probably turn out to be only one of many different mechanisms for gene control. For example, specific positive control elements are known to exist. The enzymes required to utilize the sugar arabinose, which have been studied by Ellis Englesberg of the University of California at Santa Barbara, have a single control gene, as lambda and the *lac* operon do, but the product of this gene seems to be required to turn the operon on rather than to turn it off. That is, if the control gene is deleted, the structural genes do not function. Such control could work through the attachment of arabinose to the product of the regulatory gene, with the resulting complex in turn binding to DNA and modifying the DNA structure in such a way as to permit the operon to be read. Confirmation of this notion awaits the isolation of the product of the control gene.

One specific molecular process of positive control is now known. Richard Burgess and Andrew Travers of Harvard and Ekkehard Bautz and John J. Dunn of Rutgers University have shown that RNA polymerase, which initiates the synthesis of RNA chains at the promoters, contains an easily dissociated subunit that is required for proper initiation. This subunit, the sigma factor, endows the enzyme to which it is complexed with the ability to read the correct promoters. Travers has shown that the *E. coli* phage T4 produces a new sigma factor that binds to the bacterial polymerase and enables it to read phage genes that the original enzyme-sigma complex cannot read. This change explains part of the timing of events after infection with T4. The first proteins made are synthesized under the direction of the bacterial sigma factor; among these proteins is a new sigma factor that directs the enzyme to read new promoters and make a new set of proteins. This control by changing sigma factors can regulate large blocks of genes. We imagine that in *E. coli* there are many classes of promoters and that each class is recognized by a different sigma factor, perhaps in conjunction with other large and small molecules.

Both the turning on and the turning off of specific genes depend ultimately on the same basic elements we have discussed here: the ability to recognize a specific sequence along the DNA molecule and to respond to molecular signals from the environment. The biochemical experiments with repressors demonstrate the first clear mechanism of gene control in molecular terms. Our detailed knowledge in this area has provided some tools with which to explore other mechanisms.



SPECIAL PHAGE CHROMOSOME was utilized in some experiments on the repressor-DNA interaction. The chromosome carries *lac* genes in addition to lambda genes. The diagram shows approximate relationships among some of the elements. The repressors are about twice as wide as the DNA molecule, which is some 8,000 times as long as it is wide.

We want to be useful ...and even interesting

Kodak

What Harry does

• **New tissue solubilizer for liquid scintillation counting (LSC).** We have stimulated Kodak chemists to develop a strongly basic quaternary ammonium hydroxide and solvent combination that is high in acceptance of water at 0-5C and low in quenching. Order from lab suppliers as EASTMAN 13108.

• **Neutron detection by LSC.** We have arranged for Cadmium 2-Ethylhexanoate, known from the literature, to be reliably prepared in Kodak laboratories and distributed as EASTMAN 10756.

• **Numerous LSC fluors and solvents in quality adequate for Kodak's own research use.** We have had repeated sessions with a committee representing various Kodak laboratories. Object: conformance at lowest cost to performance standards that safeguard against waste of work. Now added: Butyl PBD (EASTMAN 13003), POPOP (13005), Dimethyl POPOP (13009), p-Terphenyl (13057), p-Xylene (13071), 2-Methoxyethanol (13072), 2-Butoxyethanol (13073), 2-Aminoethanol (13074), Anthracene (13076), α-NPO (13077), BBOT (13109), PBBO (13111). (Previously announced in this same Scintillation Grade have been p-Dioxane (EASTMAN 13011), Toluene (13016), Naphthalene (13007), PPO (13000), and Phenethylamine (13021).)

• **For extreme scintillation yields in radiochromatography with tritiated material, use EASTMAN CHROMAGRAM Sheet 6061, apply benzene, and drop to −190C.** We are happy to note *(Anal. Chem. 42:*156 (1970)) this additional interesting connection between our earlier contribution to thin-layer chromatography (TLC) and our present activity on behalf of LSC.

• **"Inorganic Thin Layer Chromatography of Thirty Common Cations."** Title of paper in a British journal *(Metallurgia 80:*209 (1969)). We are also happy to note that the newest alternative to industrial emission spectrography with KODAK Spectrum Analysis Film and Plates is TLC with EASTMAN CHROMAGRAM Sheet.

"We" as used above refers mostly to a Mr. Harry Anonymous (not his real name), whose personal scintillations are proving quite effective. He didn't even have a B.S. when he came in 1958, low man in a team constituting the U.S. end of a collaboration with our French affiliate Kodak-Pathé toward reproducible sheet for TLC. Ph.D.'s supplied the brain power, Harry the hands. Don't admit this to your children, but sometimes the advantage lies with the hands. Progress is then made either by becoming entranced with the fine details, or conversely, by prying the subject loose from the grip of scholars. Since Harry Anonymous had personally manipulated so much of the product that became EASTMAN CHROMAGRAM Sheet, it became clear he could be equally useful to outside scientists interested in the stuff. Harry put a higher interpretation on his being paroled to the outside world. Right off he saw himself as a communication channel between the people outside who want us to make something and the people inside who know how to make it. He won acceptance as a matchmaker. "Product planner" sounds more dignified. Now he does it for all EASTMAN Organic Chemicals, not just CHROMAGRAM Sheet.

What Tom does

Tom Deever spent three-quarters of his career at Otterbein College preparing for the ministry. Upon graduation in June, 1968 he set to work so that Kodak cameras for ordinary family use might yield pictures of a sharpness that has long been considered to call for much costlier equipment.

Tom grew up in a family environment of emphasis on social responsibility. He liked the "hard" sciences and math, but merely for fun. In his junior year, a teacher who had been a clergyman advised, "Never go into the ministry unless you couldn't be happy anywhere else." Tom knew physics made him happy. A physics professor sent him to us.

With some hesitation, we took a chance on him, told him how lens designers assume the centers of curvature of their refracting surfaces to be collinear even though in an imperfect world this cannot be, told John McLeod that a new young man would be available to fetch and carry for him as he explored the potential of axicon-laser combinations.

It is hard to believe that a totally new kind of optical element—neither a lens nor a prism—could have waited to be invented as late as 1954. Invented the axicon was by this same McLeod. Then it was interesting for imaging a point source everywhere along an axial line; now, with lasers as light sources, it is more interesting in the opposite direction.

Tutelage by a man of his grandfather's generation has not kept young Tom Deever from contributing 1) a way of placing axicon and laser off-axis from experimental lenses and 2) guidance to the computer programmers for a program we call SOMEX (Simulated Optical Manufacturing Experiment). Together, his contributions enable us to deal with manufacturing perturbations as parameters right along with the idealistic parameters —curvature, index, dispersion, thickness, spacing.

Membrane by the acre—NOW!

Just over a year ago, EASTMAN Membrane for reverse osmosis was announced here. Readers were warned that the membrane was supplied wet because drying killed it. They were also bluntly warned they were on their own to devise mechanical support for it against the applied pressure.

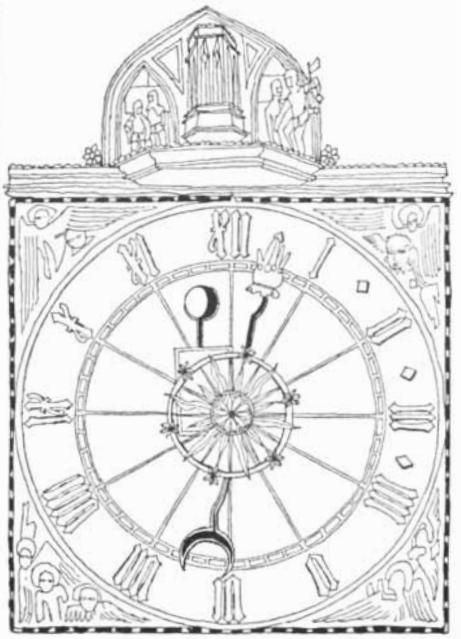
A year's progress has softened these demands. EASTMAN Membrane that stays dry until the customer's equipment is ready to operate can now be produced in high volume. It can be supplied as half-inch tubing. Much more information is now available on supports and on high-temperature operation. During the same year many communities have imposed industrial sewer charges of a magnitude to provide plenty of incentive for reclaiming water and/or solutes.

Now the sales curve for EASTMAN Membrane can better show whether faith is justified that what's good for the environment is good for business.

Let's go!

Please direct commercial inquiries to "Membranes," Eastman Chemical Products, Inc., Kingsport, Tenn. 37662.

SCIENCE AND



Cruelest Month

April was a month of fateful developments in the nuclear-arms race. The substantive phase of the U.S.-U.S.S.R. strategic-arms-limitation talks (SALT) opened in Vienna. In the U.S. steps were taken to increase the level of offensive and defensive nuclear armament in the face of urgent proposals for at least a temporary halt in the deployment of strategic weapons. China entered the space age.

The SALT meeting comes at a critical time. Both superpowers appear anxious to slow the arms race, if only to limit its enormous cost. Each, on the other hand, is developing new offensive and defensive weapons that tend by their nature to destabilize the balance of terror and accelerate the arms race. The juncture of negotiation and escalation has prompted urgent proposals in the U.S. for a mutual halt in the deployment of multiple independently targeted reentry vehicles (MIRV) and anti-ballistic-missile systems (see "The Limitation of Strategic Arms," by G. W. Rathjens and G. B. Kistiakowsky; SCIENTIFIC AMERICAN, January). In April the general advisory committee of the U.S. Arms Control and Disarmament Agency recommended that the U.S. seek a comprehensive agreement limiting strategic weapons; the Senate approved, by 72 to six, a resolution urging President Nixon to propose an immediate moratorium on the deployment of such weapons. The President said the resolution was "irrelevant," since it merely restated the Administration's own general objective for the talks. The President did decide, however, to aim for a comprehensive limitation on strategic weapons in Vienna rather than taking the more cautious weapon-by-weapon approach recommended by some of his advisers.

Meanwhile the testing and production of the MIRV missiles Poseidon and Minuteman III went forward. The Administration pressed for Congressional approval of a third Safeguard anti-ballistic-missile site and preliminary work on five more sites. (Construction of the first two sites was authorized last fall.) The Department of Defense announced that a Russian missile-launching submarine was on station in the Atlantic Ocean. The Institute for Strategic Studies, a British research group, reported that the U.S. now has about 4,200 nuclear warheads ready for delivery and would have about 11,000 by 1975 on the basis of current plans; it said the U.S.S.R. has about 1,900 warheads and could triple that number by 1975.

China announced the launching on April 24 of its first earth satellite, a 381-pound spacecraft that orbits the earth every 114 minutes broadcasting a popular song, "The East Is Red." China is the fifth nation to boost a satellite into orbit (after the U.S.S.R., the U.S., France and Japan). The launching was said to have been anticipated by U.S. defense and space officials. Nevertheless, it signaled a major advance in China's capability for long-range nuclear warfare. China had tested nuclear devices, including a hydrogen bomb, and was known to be producing nuclear weapons; now it has shown that it may soon be able to deliver a nuclear weapon to a target. The spacecraft was considered a heavy payload for a country's first successful launching attempt; Japan's, orbited several months ago, weighed 50 pounds. The Chinese satellite could have been launched by the first stage of the medium-range ballistic missile China is known to have developed or by the first stage of an intercontinental missile.

Nuclear Navy

U.S. warships with nuclear-powered propulsion plants now in operation number 86 missile-launching or attack submarines, a submarine research vessel designed for deep dives, and four sur-

face vessels: the aircraft carrier *Enterprise,* the missile cruiser *Long Beach* and two missile frigates. Planned for the future, the Atomic Energy Commission notes in its annual report to Congress for 1969, are further attack submarines of two new kinds: three of a "high speed" design and an unstated number of "quiet" electric-drive vessels. The total number of nuclear submarines authorized by Congress through fiscal year 1970 is 110, including 41 of the Polaris class; the AEC report does not comment on the 19 or so submarines authorized but excluded from its report.

Additional surface units planned or under construction include two more of the present class of missile frigates, four missile frigates of a new class and three *Nimitz*-class supercarriers. Two of the supercarriers have been authorized by Congress; the keel of the first was laid in June, 1968. The *Nimitz* and its sister ships will be able to operate for 13 years without refueling.

Easier Abortion (Cont.)

The trend toward making a legal abortion easier to obtain in the U.S. is accelerating. In recent weeks Alaska, Hawaii and New York have enacted laws that in essence make an abortion in the first four to six months of pregnancy a matter to be decided by the woman and her physician. The Maryland legislature has passed a similar law; it would replace one, enacted only two years ago, that was based on a model put forward by the American Law Institute. The model provides that a physician may lawfully terminate a pregnancy when there is a substantial risk to the woman's life or health, when there is a threat that the child may be born with serious defects or when the pregnancy is the result of rape or incest. Ten states besides Maryland have adopted essentially this reform since Colorado led off in 1967.

Reform of one kind or another has thus been enacted by 14 states since 1967. All the other states still have laws of long standing that severely restrict the conditions for legal abortion. Most of the laws permit an abortion only when it is necessary to save the life of the mother. In California, Wisconsin and the District of Columbia such laws have been over-

turned by lower courts; the expectation is that one or more of the decisions will be reviewed by the U.S. Supreme Court. (California is one of the states that have adopted a reform along the lines of the American Law Institute model; the court decision was made in response to a challenge of the old law.)

A considerable amount of dissatisfaction appears to exist in the states that have adopted reforms patterned on the American Law Institute model. (None of the more liberal laws has been in effect long enough to evoke comment.) A survey by *The New York Times* found that in eight states with laws based on the model, criticisms were made that the laws had little effect on the poor, who could not afford abortions or the medical consultations that are often necessary preliminaries, and that many doctors still refuse to perform abortions either on moral grounds or because they are uncertain of their liabilities under the new laws. Advocates of more liberal abortion laws also argue that, because of the restrictions in the American Law Institute model, only about one woman in 10 seeking an abortion can get one under statutes based on the model.

Quartet in Amino Acids

The three-dimensional structure of lactate dehydrogenase has been worked out by a group of investigators headed by Michael G. Rossmann of Purdue University. The accomplishment is unique in two ways. It is the first such structural determination of an enzyme whose primary structure—the sequence of amino acids in its molecule—is still unknown. It is also the first structural determination of an enzyme made up of more than one subunit. (A protein subunit is defined as a polypeptide chain produced by the chemical linkage of several dozen amino acids.) Lactate dehydrogenase is made up of four identical subunits, each consisting of somewhere between 310 and 326 amino acids; the exact number is still uncertain. In having four subunits lactate dehydrogenase resembles the blood protein hemoglobin, whose complete structure was worked out six years ago. (Hemoglobin is not ordinarily regarded as an enzyme.)

Lactate dehydrogenase plays an important role in muscle cells when sudden stress has depleted the muscle's oxygen supply. At such times lactate dehydrogenase cooperates with other enzymes to meet the cell's energy needs by breaking down glucose. There is evidence that the several subunits in an enzyme shift slightly but precisely as they carry out their chemical function. This rearrangement of subunits may provide the "jaws" for gripping and then modifying a specific substrate. Lactate dehydrogenase turns out to have a large cavity for holding the substrate (lactate) that it helps to alter chemically. When the substrate is in position, part of the molecule closes like a trapdoor.

Rossmann and his colleagues have spent more than five years making X-ray-diffraction measurements of lactate dehydrogenase and derivatives in which certain atoms were replaced by the atoms of heavy metals. Knowing the location of these heavy atoms provides important clues for interpreting the X-ray patterns. So far Rossmann's group has established the position of 310 of the amino acids in a subunit of the enzyme. Unraveling the complete amino acid sequence of a subunit is a separate chemical problem still to be solved.

Upward and Hahnward

A new element, number 105 in the periodic table, has been synthesized by workers at the Lawrence Radiation Laboratory of the University of California. The element has been named hahnium, after the late German physicist Otto Hahn. The announcement of the discovery was made in Washington at the spring meeting of the American Physical Society by Albert Ghiorso, head of the Berkeley group and a codiscoverer of elements 95 through 104. His colleagues in the latest experiment were Matti Nurmia, James Andrew Harris, Kari A. Y. Eskola and Pirkko L. Eskola.

Element 105 was produced by bombarding a minute target of californium 249 (element 98) with a beam of high-energy nitrogen-15 nuclei from the laboratory's heavy-ion linear accelerator (HILAC). When a nitrogen-15 nucleus was absorbed by a californium-249 nucleus, four neutrons were emitted and a new atom, hahnium 260, was formed. The half-life of the new element turns out to be unexpectedly long: 1.6 seconds as opposed to the few milliseconds predicted on the basis of the half-lives characteristic of the other known transuranium elements. Moreover, the atoms of 105 were produced in comparative abundance: about six atoms per hour for as long as the experiment was continued. Since the first atoms of 105 were detected conclusively in early March, many hundreds have been observed.

The identification of the new element was based both on its observed half-life

and on the energy of the alpha particles emitted when its nuclei decayed. The validity of the identification was further confirmed by observations of the decay of a by-product of the 105 decay; the decay properties of this by-product, lawrencium 256, are well known.

Russian physicists had reported observing atoms of element 105 as much as two years ago, but the Berkeley group believes these reports are in error. A similar dispute followed the announcement last year by the Russian group at Dubna that they had succeeded in synthesizing element 104, which they named kurchatovium in honor of the Russian physicist I. V. Kurchatov. The Berkeley group then synthesized element 104 by a different method and challenged the Russian finding, naming their discovery rutherfordium in honor of the British physicist Lord Rutherford. The official names of elements 104 and 105 have yet to be authorized by the International Union for Pure and Applied Chemistry.

The Berkeley group next plans to bombard einsteinium 253 with carbon-13 nuclei in an effort to produce the more stable isotope of element 105, hahnium 262. The half-life expected for this isotope, a minute or two, should make it possible to perform tests of the chemical properties of the new element. Various searches for elements even heavier than element 105—including possibly long-lived "superelements"—are under way at Berkeley and elsewhere.

More Interstellar Chemistry

One of the most active areas of research in radio astronomy for the past year or so has been the investigation of clouds of various gases in regions of interstellar space where new stars are believed to be in the process of formation. The molecular species detected so far in these regions include the hydroxyl radical (OH), water (H_2O), ammonia (NH_3) and formaldehyde (H_2CO). Recent improvements in observational technique have led to two more important developments. One is the discovery of carbon monoxide (CO) in a number of regions within our galaxy. The other is the resolution of water-vapor clouds with extremely small angular sizes in regions where radio waves characteristic of the presence of both water molecules and hydroxyl radicals had been detected before.

The discovery of the interstellar carbon monoxide clouds was reported recently by Keith B. Jefferts, Arno A. Penzias and Robert W. Wilson of the Bell

Some things are changing for the better.

Math comes alive for grade school students.

And they're learning it faster than ever before, according to Mr. William Rybensky, program director for computer-assisted instruction (CAI) at Ravenswood City School District, East Palo Alto, California. In his district, students are averaging 15 months of math progress in just 7 months. And at Willow School (where 400 children now regularly use the HP time-share system) some have finished 1½ years of math in less than three months!

And students really like working with our computer, because they receive immediate recognition for correct answers and avoid the embarrassment of making a mistake in front of the whole class. Teachers also appreciate the system. It gives them up-to-the-minute progress reports on every student and eliminates the need for checking answers. So the teacher has more time to coach slow learners and deal with the "why" questions.

For about 25 cents per student per day, any school district can enjoy the benefits of a Hewlett-Packard system, complete with software based on the Stanford Math Drill and Practice curriculum for grades 1 through 6. This $150,000 system is ready for delivery now, completely field-proven and fully supported.

If you'd like to find out how CAI from Hewlett-Packard can help your school district, just drop us a line.

New ways to bypass the big computer.

Enthusiasm for our 9100 Computing Calculator's capabilities has led to a most rewarding exchange of ideas with its users. A typical letter from a California university professor states in part:

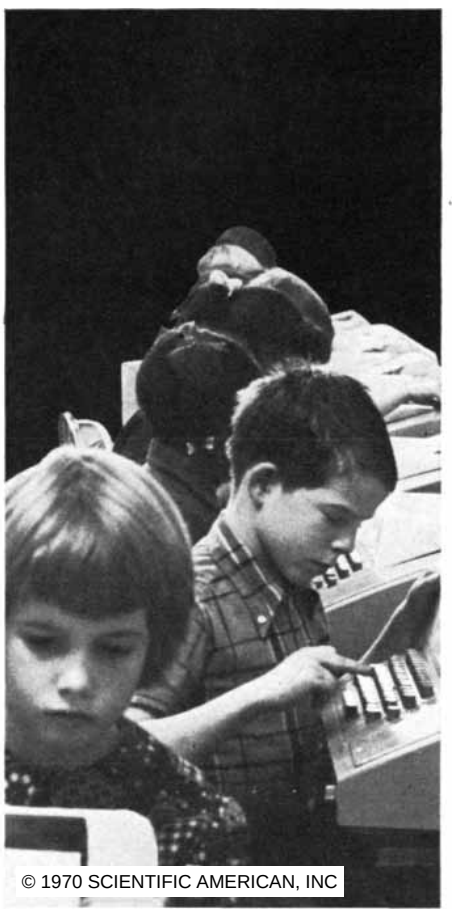
"I do not normally write letters of appreciation to manufacturers of products I use, but my joy in ownership of one of your new 9100A calculators impels me to write you my congratulations and praise for a unique product. The design of this machine seems perfectly matched to the needs of a scientist who is continually confronted with minor problems requiring either the use of tables or the fairly laborious construction of a computer program. All too many pertinent problems are ignored because of the work involved.

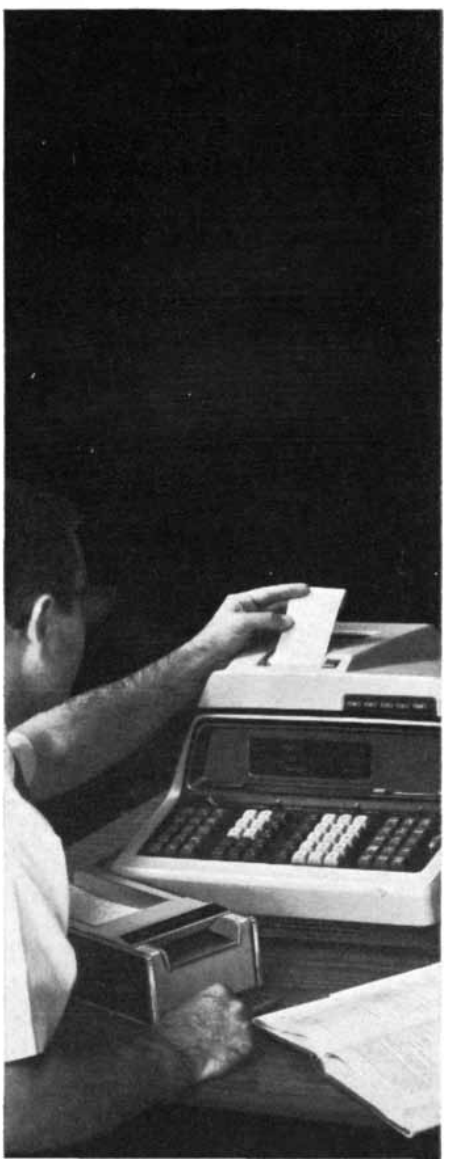
"We have our own CCI terminal to our (big computer) which is a blessing for big problems, but a headache for little ones. We have repeatedly found mistakes in our Fortran and PL-1 programs by spot-checking the routines with our 9100A.

"The ease of programming and program editing immediately attracts the novice as well as the expert and the ability to step through a routine is helpful in debugging.

"Your choice of parameters seems particularly felicitous—with sufficient speed to provide accuracy without noticeable delay and all the hard-wired subroutines one needs to do anything."

Users like this have helped us compile libraries of programs for our calculator system. The newest of these relates to applied statistics, consists of over 100 programs, and covers the following: General Statistics; Distribution Functions; Test Statistics; Curve Fitting; Harmonic Analysis; Sampling Theory; Analysis of Variance Operations;





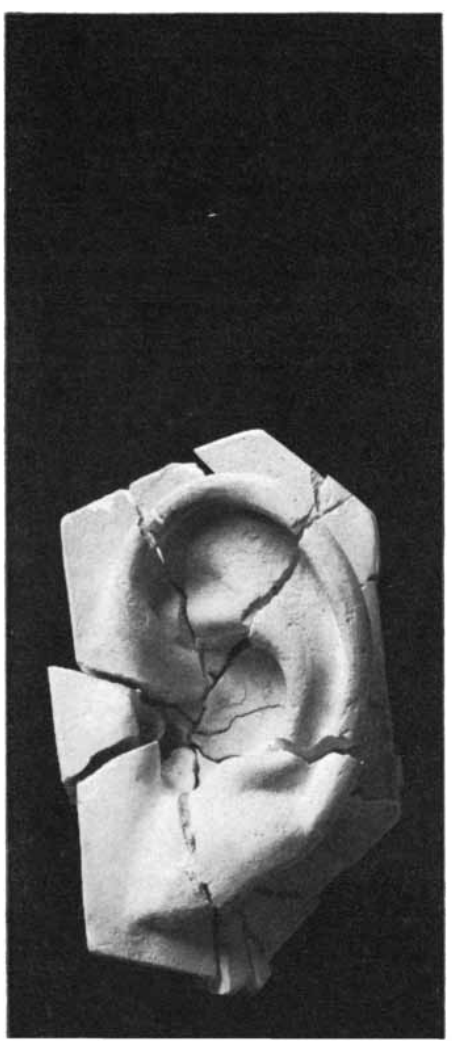
Reliability and Quality Control; and Plotter Programs.

This series of programs on magnetic cards gives you instant access to the calculators so you can spend all your time on creative data interpretation rather than computation. We call it the HP STAT-PAC. Used with the HP 9100 Calculator System, it can help free you from the laborious reduction of statistics and endless computations. Write us if you'd like more information.

Help for companies with noise problems.

Occupational noise exposure and its potential solutions are currently undergoing extensive investigation by government, industrial, legal and medical experts. Last year's enactment of the Walsh-Healey Act really put teeth in the demands for better industrial noise environment, especially with any firm having Federal contracts of $10,000 or more per year.

If you're one of the people looking for an inexpensive way to measure industrial noise, a new Sound Level Recorder we developed can help you.



It's a simple push-button instrument that accurately measures and records all the noise exposure levels and peak sounds covered by the Walsh-Healey Act. If you need certified documents of performance, we have a Sound Level Calibrator that goes with it.

The Recorder costs $2,200, the Calibrator $285. We'll be happy to send you details on both instruments, plus information on how you can use them to meet Walsh-Healey requirements.

A better way of learning to cope with heart failures.

When caring for cardiac patients, how can a nurse be sure she will recognize a life-threatening arrhythmia? Or effectively defibrillate in the few vital minutes available? Or simply make good electrode connections to assure continuous, accurate monitoring information?

The key, obviously, is the right training. A highly effective training tool is our Arrhythmia Trainer and Torso. With it, nurses and interns quickly learn to recognize individual arrhythmias and the sequences in which they occur. And it duplicates human response to correct and incorrect electrical treatment, so they can practice countershock techniques. It's the kind of training no hospital should be without. The price is just $1750.

This is another example of how Hewlett-Packard technology is helping to change things for the better. More detailed information on any of the above, or any other HP instrument is available by writing to: Hewlett-Packard, 1504 Page Mill Road, Palo Alto, California 94304; Europe: 1217 Meyrin-Geneva, Switzerland.

HEWLETT hp PACKARD

Measurement, Analysis, Computation

00010

Telephone Laboratories. In order to detect radio waves characteristic of emission by carbon monoxide molecules they equipped a radio telescope at the Kitt Peak National Radio Astronomy Observatory with a special high-frequency receiver developed by C. A. Burrus of Bell Laboratories. The new equipment picked up signals emitted by carbon monoxide molecules in at least five different regions of the Milky Way, including the Orion nebula and the direction of the galactic center in the constellation of Sagittarius. On the basis of the unexpectedly large amounts of carbon monoxide detected in these regions the investigators now expect to find an abundance of this gas in many other regions of space. Carbon monoxide is considered a likely end product of processes associated with the birth and death of stars.

The improved resolving power of the instruments used to measure the extent of interstellar gas clouds was reported by a group of radio astronomers from the Massachusetts Institute of Technology and the Naval Research Laboratory. Using the technique of very-long-base-line interferometry, in which pairs of radio telescopes separated by up to hundreds of miles are focused on the same celestial object at the same time, they measured four known sources of water-vapor radiation. In one target region (the radio source known as W49) the group measured five different water-vapor clouds, each no more than .003 second in angular diameter. Since W49 is some 45,000 light-years away, this works out to a linear diameter no greater than 40 times the distance from the earth to the sun. A second target in the constellation of Orion was found to have two water-vapor clouds with the same angular diameter. Since these sources are only about 1,500 light-years away, they must extend over no more than one and a half times the earth-sun distance. Single water-vapor clouds were discovered in two other regions.

Who's Who at Swartkrans

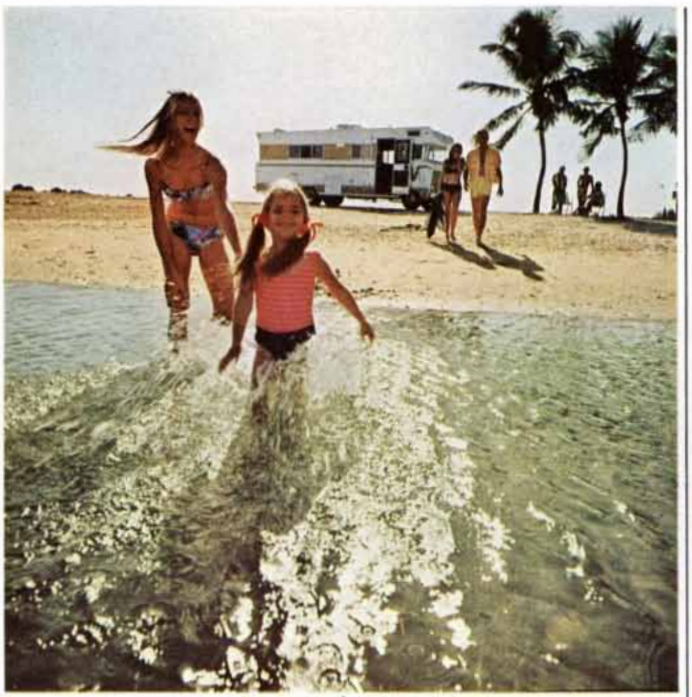
The limestone quarry at Swartkrans, a fossil-rich site first investigated in 1948 by the late Robert Broom and his colleague John T. Robinson, is one of five sites in South Africa that have yielded remains of the subhuman hominid *Australopithecus.* Renewed work there since 1965 by the staff of the Transvaal Museum and other scholars has yielded a number of fresh glimpses into the nature of the region's ancient population and the question of how the fossil deposit came into being over several thousand years during Pleistocene times. First in importance is the confirmation of Robinson's long-standing belief that a more advanced hominid was present at Swartkrans along with *Australopithecus.*

Until last summer this second hominid, named *Telanthropus capensis* by Robinson on its discovery in 1949, was represented by only five fragments: a lower jaw, a smaller piece of jaw, a single tooth, one end of an arm bone and a piece of palate. Examining the collection of *Australopithecus* fossils from Swartkrans at the Transvaal Museum last July, R. J. Clarke of the Kenya Center for Prehistory and Palaeontology noticed among them some skull fragments, including a cheek bone and the orbit of an eye, that appeared to differ substantially from the rest. Clarke, together with F. Clark Howell of the University of Chicago and Charles K. Brain of the museum staff, found not only that these fossils could be fitted together but also that the fragment of palate found in 1949 matched the assembly perfectly.

The combination represents the left side of a face and most of the upper jaw; when they are joined to the *Telanthropus* lower jaw, the result is the partial image of a hominid with a heavy brow ridge and a face substantially shorter and smaller than the faces of the robust Swartkrans australopithecines. Reporting their conclusions in *Nature,* the three investigators state that the assembled fossil should be regarded as a species of *Homo* and that it resembles *Homo erectus,* an identification Robinson has long suggested for *Telanthropus.* Until further fossil material is recovered, however, they prefer to write it down simply as *Homo* (species indeterminate).

Meanwhile Brain has concluded from his study of the 14,000 fragments of bone thus far obtained from Swartkrans that the deposit represents a typical carnivore's accumulation, in particular the kind that results today from the arboreal feeding and storage habits of leopards. The prey at Swartkrans included antelope, hyrax, baboon and the australopithecines themselves. Brain's identification of leopards as the probable predators is neatly supported by puncture wounds in the skull of one immature *Australopithecus*—wounds that are frequently cited as evidence of man's long record as a killer of his own kind. The punctures are spaced 33 millimeters apart; this is roughly the distance between the tips of a leopard's lower canine teeth. Brain suggests in *Nature* that the holes were made as a leopard dragged the dead australopithecine by the head to its feeding retreat.

Winnebago. The fifty-two week funhouse on wheels.



Exciting Winnebago "motor homes" are not to be confused with those other vehicles called "mobile homes." Not that there's anything wrong with a mobile home. It's just that we don't make them. We make a self-propelled, self-contained, fifty-two week funhouse on wheels. In fact, we make ten different models of them. And we make them so you can go places, do things, meet people, and have a downright good time while you're at it.

Like touring wherever and whenever you wish. Vacationing with your family. Taking the gang to the ball game, or your cronies out duck hunting. You can even supply the ski lodge on your next Aspen outing. Or use it as a warming house for any winter sport. Winnebago's your cottage by the sea. Or it's anything you want it to be. Anytime you want it to be.

It's your lounge, your kitchen, your dining room. Your shower and bedroom. Plus your transportation. And that eliminates the hassel of motel reservations and restaurants and schedules and layovers and packing and unpacking and packing again.

You see, a Winnebago motor home offers new vistas, an enjoyable change in the way you do things with your family and friends. You'll meet new people. People who enjoy life, people already part of the "fun set," whatever your brand of fun.

So, when you get a Winnebago, you're getting more than a two-week vacation in Acapulco. A whole lot more. You're getting all the things you ever imagined being a part of, fifty-two weeks a year, for years to come. That's a happy investment, and a wise one. Because you can get a Winnebago motor home for about the same price as one of America's finer made automobiles.

They come in 10 different models, 5 different lengths, 7 different floor plans and from standard to luxurious. So, sure as a bear has hair, you'll find the one Winnebago motor home that's your cup of tea.

For instance, the industry's most popular selling motor home is Winnebago's model D-22. Twenty-two feet of fun on wheels, with pecan colored, vinyl clad plywood cabinetry. High-fashion colors of gold, avocado and blue, coordinated among quilted upholstery, window curtains, kitchen appliances and wall-to-wall carpeting. A dinette with 5" thick foam cushions covered with vinyl one side and fabric the other. It makes into one of three beds sleeping six (optional eight). And in the kitchen, the four burner deluxe range with automatic

oven sits between the built-in refrigerator-freezer and the stainless steel double sink. There are overhead cabinets, drawers and work space with drop leaf extension. The bathroom has the usual facilities, lavatory, marine toilet, and telephone shower head. A bed, accommodating six-footers, is made up from the rear lounge and there's a divider curtain for privacy. Another divider curtain up front is for use during nighttime driving. Otherwise the driver's compartment is a part of the entire motor home complex.

Rich black vinyl bucket seats for the driver and copilot match the padded dash. Panorama windows all around. Power steering, power brakes and three-speed transmission let even grandmothers drive it with ease and assurance. Ducted forced air gas heat keeps it comfortable in the winter time. And optional air conditioning is available to help you keep your cool in the simmering summer.

Camp America's Fly In - Camp Out rental program uses Winnebagos exclusively.

Call this toll-free number for the name of your nearest Winnebago dealer. We'll send you a colorful brochure, too.
800/553-9550

(In Iowa call collect 319/242-1867)

WINNEBAGO®
Designed from the inside out

Both printing and visual display models available with or without programming.

The idea behind a remarkable new product from The Calculator Company. It will revolutionize the working habits of the scientists, mathematicians, engineers and statisticians of the world.

THE IDEA: An application of the latest MOS/LSI micro-miniaturization. (That little chip on our finger contains several thousand transistors.)

THE BREAKTHROUGH: A desk-top calculator that makes decisions like a computer but weighs-in at a mere twelve pounds.

THE RESULT: A whole new generation of desk-top calculators for engineering and science. So if you're a scientist your calculator speaks to you in the scientific language you understand. Ditto if you're an engineer. Or mathematician. Or statistician.

THE ADVANTAGE: A unique portable programmable calculator. In both printing and display models. Also available without programming if you don't need it. Fast. Silent. Efficient. Compact. Easy to use.

OUR EDGE: Soon to be distributed worldwide, our new calculators are available now coast to coast in over 350 cities. Supported by over 3,300 sales and service technicians. Monroe men who know everything there is to know about calculators.

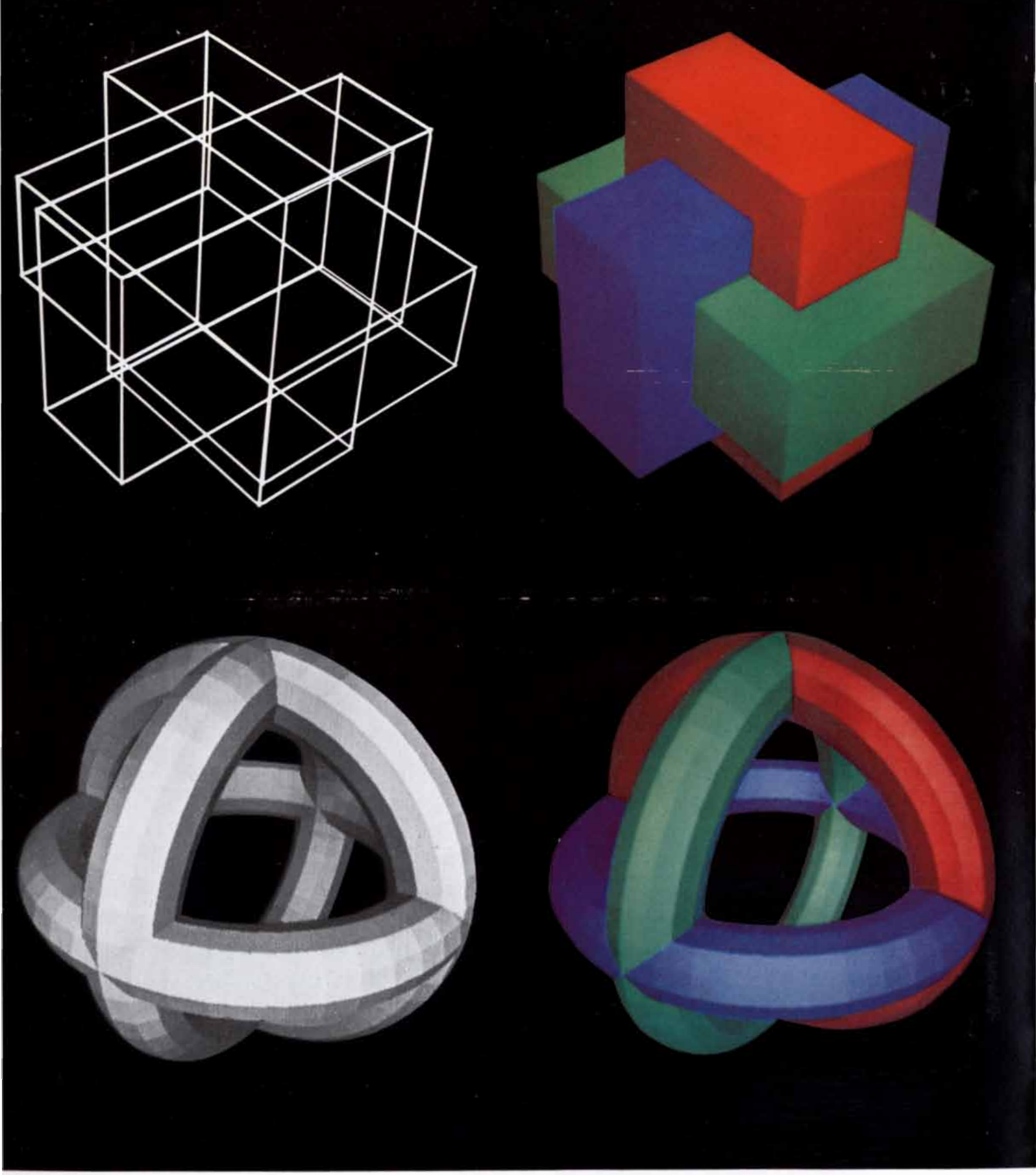
You see, calculators are our only business. Not just a sideline.

That's why year after year, more people buy Monroe Calculators than any other kind.

For more information, write to Monroe. The Calculator Company. 550 Central Avenue, Orange, N. J. 07050. In Canada: Monroe. 81 Advance Road, Toronto, Ontario

Monroe. The Calculator Company.

A Division of Litton Industries



THREE-DIMENSIONAL STRUCTURES were created by a computer using programs developed by Gary Watkins, a graduate student at the University of Utah. The simple "wire frame" version of the interlocking blocks (*top left*) is readily produced because no effort is made to remove hidden lines. By using special programs, however, the hidden lines can be identified and the structure can be displayed as a solid (*top right*). In this case the electron beam in the display tube generates surfaces by a raster technique, point by point and line by line; the entire process takes about a minute per image. To make pictures in color three different black-and-white images are created on the screen, one for each primary color, and then photographed through filters on color film. As the two pictures at the bottom demonstrate, the computer program can also create highlights, as if the pictures were taken with a flashbulb.

COMPUTER DISPLAYS

Pictures "painted" rapidly and accurately by an electron beam under computer control are helping designers and engineers at many tasks. Computer displays can also help to clarify complex physical phenomena

by Ivan E. Sutherland

Whereas a microscope enables us to examine the structure of a subminiature world and a telescope reveals the structure of the universe at large, a computer display enables us to examine the structure of a man-made mathematical world simulated entirely within an electronic mechanism. I think of a computer display as a window on Alice's Wonderland in which a programmer can depict either objects that obey well-known natural laws or purely imaginary objects that follow laws he has written into his program. Through computer displays I have landed an airplane on the deck of a moving carrier, observed a nuclear particle hit a potential well, flown in a rocket at nearly the speed of light and watched a computer reveal its innermost workings.

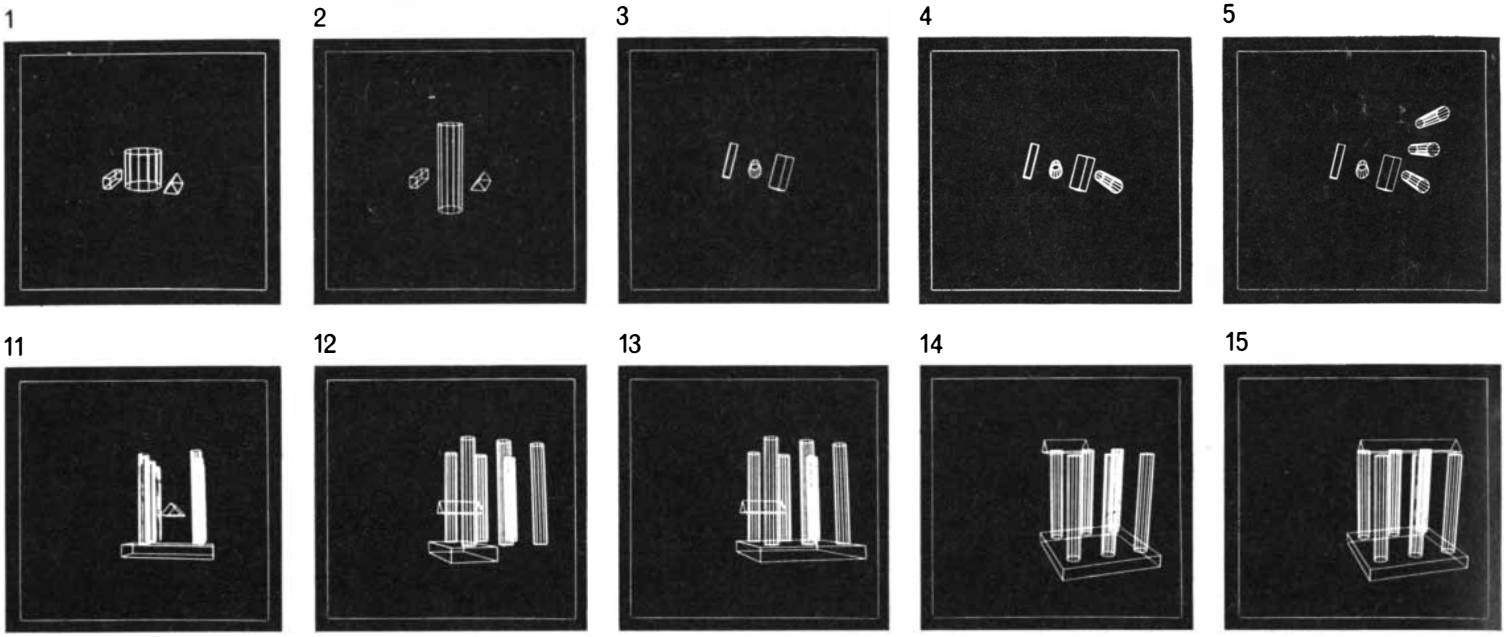
My interest for some years has been the programming of computers to draw pictures on the face of a cathode ray tube. Obviously computers can produce pictures in other ways, such as by means of mechanical plotters or by printing arrays of symbols, but because the essential problems in creating pictures by computer can be understood by considering images produced by electron beams I shall ignore these other forms of output. I shall also ignore the kinds of computer-driven picture tubes that present modest amounts of text at high speed in such places as airline terminals and brokerage offices; their operation is straightforward.

Computer displays have become of major importance to two groups of people. One group has a pictorial problem in the workaday world for which it would like computer help. These users, for example, may want to shape a metal part on a computer-controlled machine tool; they begin by describing the part to a general-purpose computer, which draws a picture of the part and verifies that the description is accurate. Other users employ computers to produce the intricate high-resolution photographic masks required for making integrated electronic circuits [*see bottom illustration on next two pages*]. Similar pictorial problems in which computers can help arise in highway planning, automobile and aircraft design, topographical mapping, architecture, the layout of publications and the production of clothing patterns. In these and many more areas, written language is far from adequate.

The other group using computer displays is interested in gaining insight into complex natural or mathematical phenomena. These users simulate physical situations of various kinds in the computer and use display devices to present the results of the simulation. For example, an organic chemist may want to synthesize a particular molecule; he creates a picture of the molecule on a display screen and then initiates a program by which the computer presents a selection of simpler molecules from which the desired substance can be synthesized [*see illustrations on pages 72 and 73*]. An engineer designing a communication circuit asks the computer for a graph showing how circuit response varies with frequency. A physician studying how blood flows through the arteries obtains a plot that reveals high vorticity at exactly the locations where the lesions of atherosclerosis are most common. A physicist programs a computer to illustrate how elementary particles interact with their own electric fields to give his students some feeling for quantum-mechanical behavior. A circuit designer draws a circuit and asks a computer to simulate its operation and to plot its performance in a graph of voltage and current. A feedback theorist describes the location of poles and zeros on a complex plane and watches as the computer plots the root locus. A mathematician enters the equations for conformal mappings and observes the maps produced by each equation. A pilot practices takeoffs and landings on a simulated airfield that can assume any orientation on the display screen as he operates "controls" for engine power and aircraft attitude. All these people, interested in educating themselves or others, use computer displays as one of many tools for gaining deeper understanding of a problem.

Two Kinds of Display

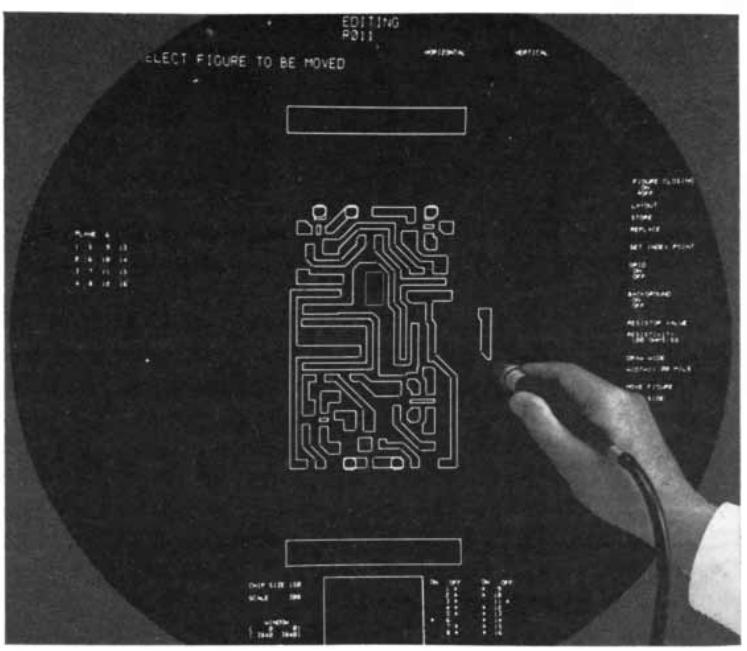
Two broad classes of computer-display systems are now in common use: calligraphic displays and raster displays. Calligraphic displays "paint" the parts of a picture on the cathode ray tube in any sequence given by the computer. The electron beam in a calligraphic display is moved from place to place in a pattern that traces out the individual lines and characters that make up the picture. Raster displays make pictures in the same way that television sets do: the image is painted in a fixed sequence, usually from left to right and from top to bottom. The calligraphic display has the advantage that information to be displayed can be stored in computer memory in any order, whereas information for a raster display must first be sorted from top to bottom and from left to right so that it can be put on the screen in the correct sequence. On the other hand, the deflection amplifiers, deflection yokes and other electronic components needed for a raster display are much less costly than those required for a calligraphic display, not only because they need not be so carefully made but also because suitable components are now mass-produced.

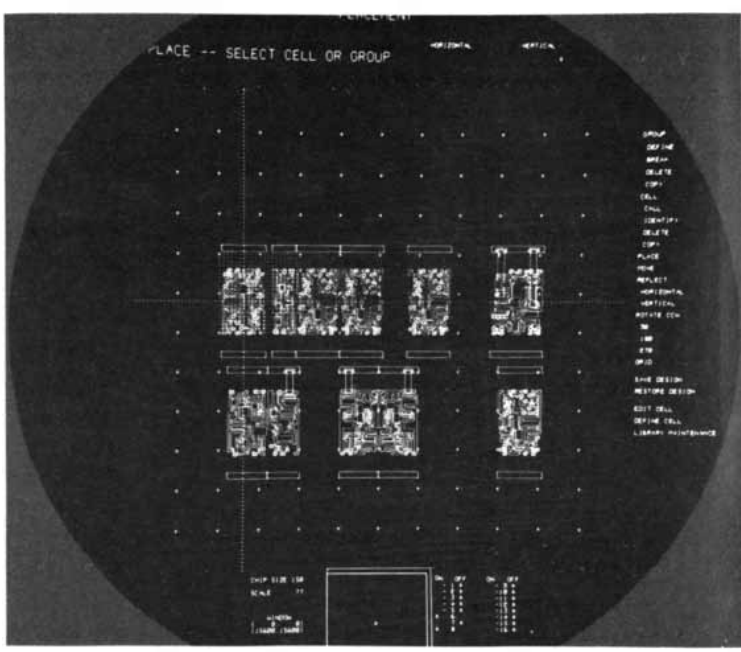


COMPUTER DESIGN OF "GREEK TEMPLE" demonstrates the versatility and speed of present-day graphic displays and computer programs. The sequence was photographed at the International Business Machines Research Laboratory; the computer was an IBM 1130 with a 2250 display unit. The "draftsman" begins by asking the computer to display three geometric solids selected from a library of shapes on file in a disk memory: a hexahedron, a cylinder and a triangular prism (*1*). The remaining operations are carried out by pressing keys on an "interactive" console that can copy, move, rotate or distort anything placed on the screen. The drafts-

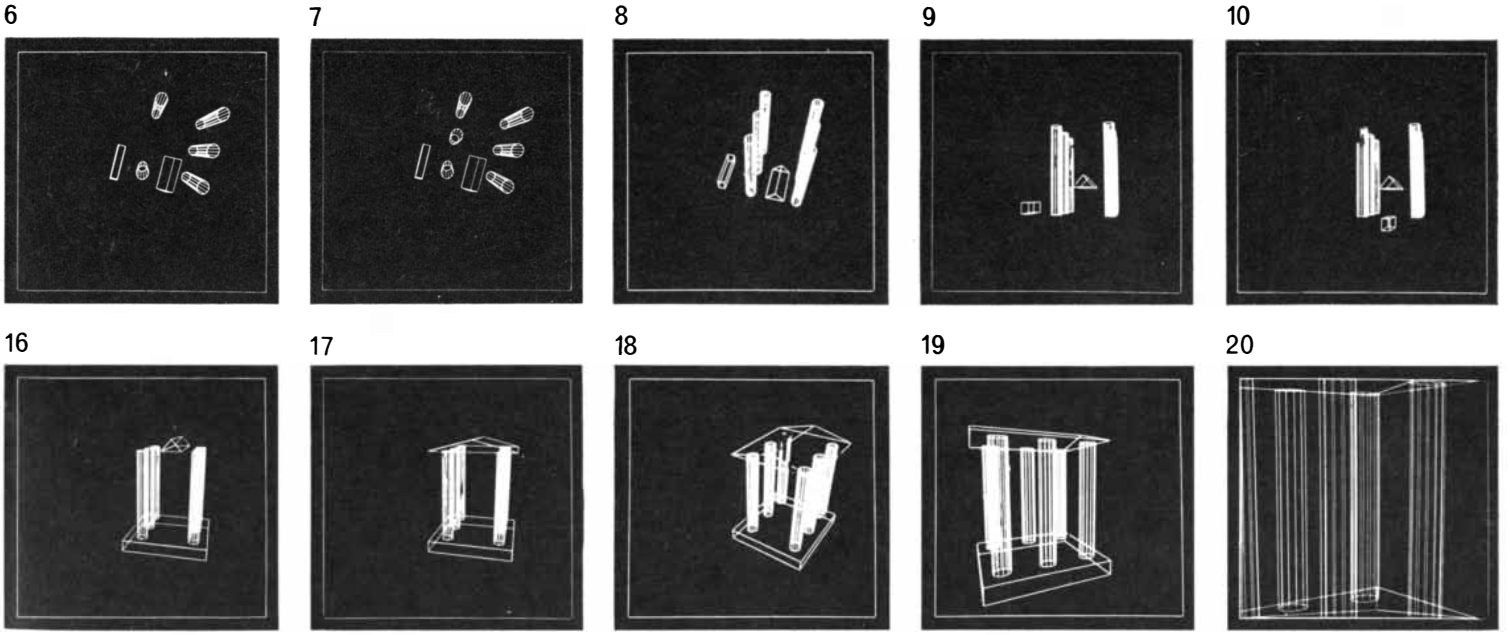
The task of sorting information from top to bottom and from left to right for presentation on raster displays has largely precluded their use for anything but presentations of text. In principle, however, a raster display has the potential of producing pictures with a range of light and dark tones, in color if desired, that provide a realism unequaled by the line drawings of a calligraphic display. The potential of this type of presentation can be seen in the pictures on page 56, which were made in our laboratory at the University of Utah by photographing a low-speed simulation of a raster display. I shall describe later some recent algorithms, or mathematical routines, for sorting picture information to eliminate hidden surfaces in making raster displays. I believe such displays will have become a common form of computer output within a very few years.

The information to be projected on the face of a computer display tube is speci-



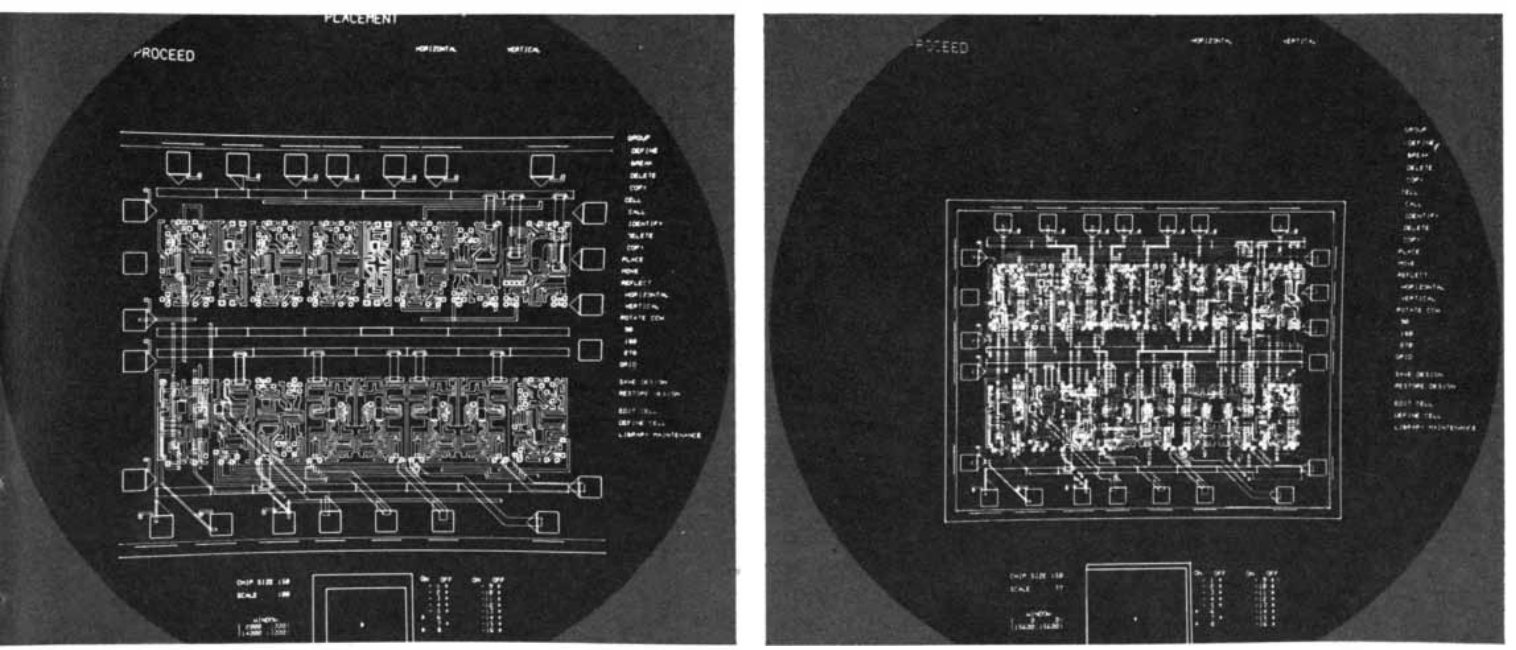


DESIGN OF INTEGRATED CIRCUITS exploits the flexibility and precision now provided by interactive computer displays. This sequence, photographed at Fairchild Semiconductor, was made with a Sanders display unit coupled to an IBM 1130 computer. The goal in this application is to end up with photolithographic masks in which line patterns are located with an accuracy of 50 millionths of an inch. The resolution of the display is defined by a coordinate grid of 1,024 lines and 1,024 columns. The coordinates of the lines drawn on the face of the cathode ray tube are recorded magnetically and then used to control digitally driven machines that cut greatly enlarged versions of the masks. Subsequently the masks are optically reduced in size to the tiny dimensions characteristic of



man lengthens the cylinder (*2*) and rotates the three elements to a plan view (*3*). He then adds five more columns (*4, 5, 6, 7*). The display is rotated (*8*) until the columns are viewed at eye level (*9*). The hexahedron is then moved to a position below the columns (*10*) and widened (*11*). The emerging temple is rotated to a side view (*12*) and the hexahedron is lengthened until it becomes a floor slab (*13*). The prism is raised (*14*) and lengthened (*15*). Again the temple is rotated (*16*) and the prism is widened to form a roof (*17*). The finished temple is displayed from above (*18*). Back at ground level the viewer walks toward it (*19*) and enters (*20*).

fied in terms of a coordinate system that covers the tube face. This coordinate system can be thought of as a grid of points that covers the face of the tube even though the grid is not actually etched on the glass. The resolution of the display system can be specified by saying how many rows and columns of points there are in the coordinate system. Most modern display systems have a minimum resolution of 1,024 rows and 1,024 columns and a few systems have four times as many rows and columns. To specify a point on a 1,024-by-1,024 array requires that each x and y value contain 10 binary digits, or "bits" ($2^{10} = 1{,}024$); to specify a point in a 4,096-by-4,096 array requires 12 bits for each coordinate. Higher resolution is not generally sought because it is difficult to maintain accuracy at higher resolutions and also because the diameter of the electron beam is so large that having higher deflection resolution improves the appearance of



integrated circuits. This circuit will be 150 mils on a side and contain a total of 16 "cells" of seven different types. The pattern in frame 1 is a mask that specifies the location of metal interconnections in a logic cell containing 13 diodes and transistors. In designing the pattern the operator uses a light pen to indicate the two end points of a line segment and the display program automatically completes the line. The pen is also used to move elements around. In this frame the ultimate cell is magnified 308 diameters. In frame 2, at lower magnification (77 diameters), a variety of cells are being summoned up from memory and are being moved into position with the help of grid points. In frame 3 the routing for one level of metal interconnections is being added. In frame 4 the job is done.

the picture only slightly. Recently, however, workers at the Bell Telephone Laboratories have built a device that uses a very fine laser beam to produce computer-controlled film images with a resolution of 32,000 lines and 26,000 columns. The device is being used to make high-resolution circuit masks [*see top illustration on pages 66 and 67*].

Although modern calligraphic displays have the ability to post dots, straight lines and text, the earliest display systems could post only individual dots on the screen. The computer specified the x and y coordinate for each dot to be posted. The coordinate specifications were converted from digital form to analogue voltages, which were then amplified to drive the deflection system of the cathode ray tube [*see top illustration on page 65*]. After the transients in the conversion process and the deflection system had subsided, the electron beam was turned on for a microsecond or so to plot the indicated dot. Characters and lines could be built out of individual dots, but of course not very quickly.

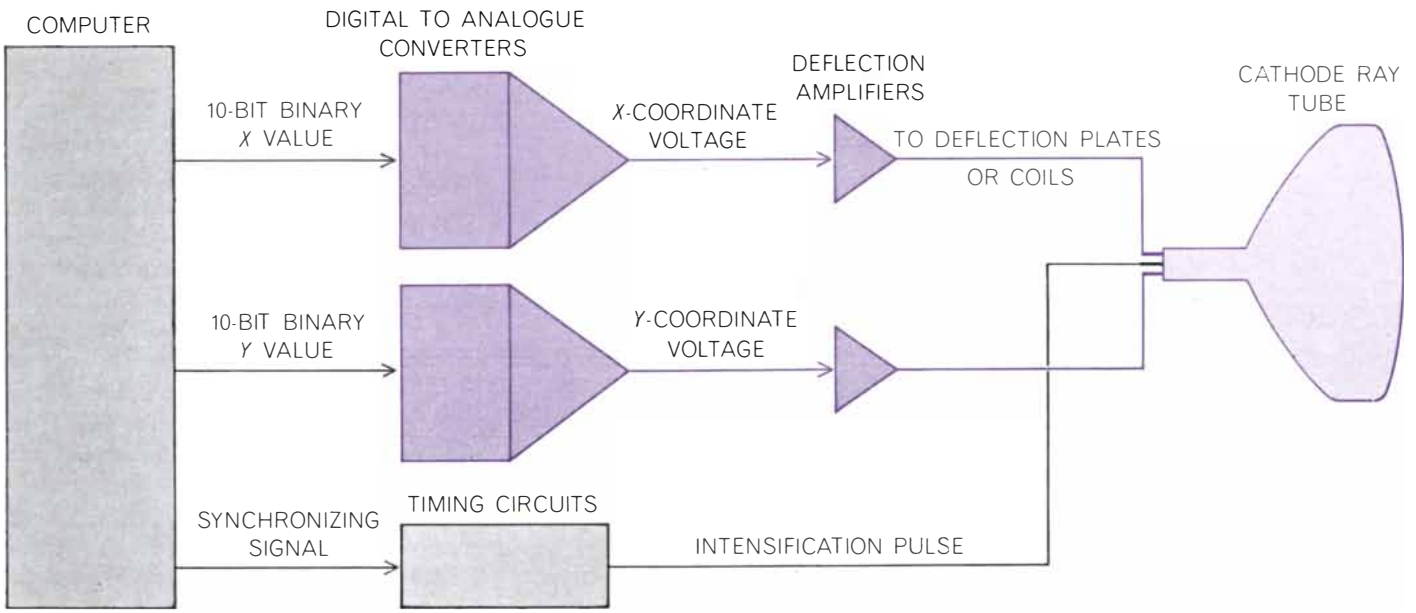
To speed up the process of drawing lines and characters, special analogue circuits have been developed that automatically trace outlines on the screen in a continuous process. To post a character the character-generator circuit produces a wave form for the x and y deflection systems and an associated pattern for turning the beam on and off; the two combined cause the character to appear [*see bottom illustration on page 65*]. To post a straight line on the screen, the analogue line generator will produce coordinated "ramp" wave forms for the x and y deflection systems so that the electron beam moves smoothly and continuously from one end of the straight line to the other. Modern display systems can produce characters in about five microseconds each and straight lines in three to 40 microseconds, depending on length. Some display systems include analogue equipment for producing simple curves, such as circles, automatically, but such facilities are still rare.

To produce a straight line all the way across a display screen in 40 microseconds is a prodigious feat of electrical engineering. The human eye has a remarkable ability to detect even slight deviations from straightness. Moreover, if two lines are plotted end to end, any error in their length such as might be produced by overshooting or a failure of the electron beam to be turned on promptly will show up either as an overlap or as a gap in the lines. The lines must be traced not only straight but also at a constant rate because any change in the beam's rate of motion will show up as a change in the brightness of the line. To be sure, one can compensate for changes in the writing rate of the line by changing the electron-beam current, but such a compensation circuit must be accurate or discrepancies will be visible. No display system that I know of is entirely satisfactory in performance, and even the best need periodic tuning to keep operating well. Considering the difficulty of the task, it is amazing that the systems work at all.

The Logic of Displays

In addition to the analogue circuits for generating lines and characters, a computer-display system must incorporate digital logic circuits to connect the display to the computer. Because the logic needed for this purpose often has characteristics in common with the processing unit of a computer, I have come to call these digital circuits the display processor. The display processor fetches information from a memory where the picture to be posted is defined and converts it to the elementary driving signals for the analogue circuits that generate lines and characters.

The display processor interprets different formats of information in the memory to give them meaning in terms of the picture produced. For example, it may interpret sequences of characters to produce rows of text on the screen, or it may distinguish sequences of connected line segments from isolated line segments, in each case producing the individual commands required by the analogue line generator to yield the desired picture. The display processor accounts for more than half of the cost of many display systems and strongly af-



COMPUTER DISPLAY EQUIPMENT consists basically of a digital computer, which specifies the point-by-point coordinates of the desired display, together with circuits that convert the digital output into analogue form for deflecting the electron beam in a cathode ray tube. If the values of x and y are each specified by 10 binary digits (0's and 1's), the electron beam can be directed to any point in an array consisting of 1,024 lines (the y axis) and 1,024 points per line (the x axis). With only the circuits shown here, the image on the screen would have to be built up from discrete points. More modern displays include analogue function generators to sweep the electron beam in a continuous fashion to generate smooth lines, numbers, letters of the alphabet and special symbols.

The sun breaks through a winter overcast to light the undercast of the New York metropolis. RYSZARD HOROWITZ

The September issue of **SCIENTIFIC AMERICAN** will be devoted to

the single topic of

TECHNOLOGY AND THE ENVIRONMENT

under the title of

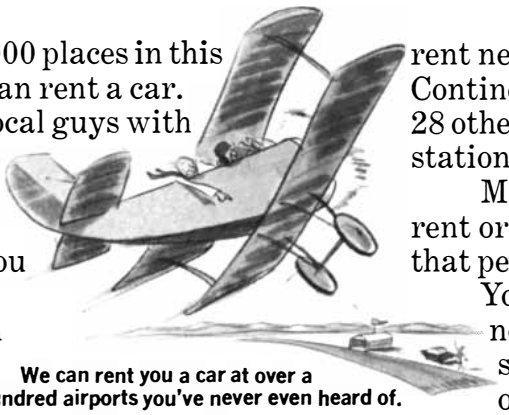
THE BIOSPHERE

We can help you whether you drill for fuel or use it.

Every rent a car company will rent you a car. And that's where the comparison ends.

There are about 15,000 places in this country where you can rent a car.

They go from local guys with a handful of cars to national companies with thousands. Some merely hand you the keys to a car. Some, in comparison to the ones who do nothing, do a lot. But nobody else does everything Hertz does. First, we make it easier to rent a car.



We can rent you a car at over a hundred airports you've never even heard of.

We can have a car waiting for you in any of over 100 countries, in practically every major airport, and in many that aren't so major.

We even can tell you about places that take children and dogs.

If you want to rent a car in one city and drop it off in any of over 1500 others, we'll arrange it for a small drop-off charge. (In 50 U.S. cities and many overseas, there's no drop-off charge. Ask us for details.)

In a number of cities, we have the only office where you can rent a car or return it at any hour of the day or night.

If you want to drive something other than a regular Ford Galaxie, ask for it when you make your reservation. You can rent new Ford Torinos, Mustangs, Lincoln Continentals, Mercury Cougars, and about 28 other kinds of cars, including convertibles, station wagons, sports cars, and even trucks.

Most of these things help you when you rent or return the car. But we've learned that people can use a lot of help in between.

You can know where you want to go but not how to get there. So we give you state, city, and even airport maps on which we'll trace the best route.

We know places where you can ward off an attack of the midnight hungries.

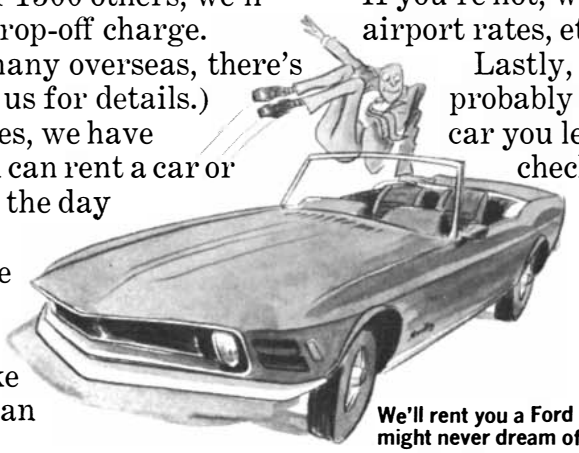
If you don't have a place to stay or if you want a good meal at an odd hour, we can suggest a few places.

We even wrote a book for each of many U.S. cities that'll tell you everything you need to know about them.

If you're on vacation, we have special rates for the United States, Canada, Hawaii, Europe, and practically anyplace else. If you're not, we have commuter rates, airport rates, etc., etc.

Lastly, we rent you a Ford that's probably better maintained than the car you left at home, a car that's checked out, washed, and serviced every time it's rented.

As you can tell, we're not just a place to rent a car. We consider it our business to help people travel as painlessly as possible.



We'll rent you a Ford your wife might never dream of letting you own.

Hertz

You don't just rent a car. You rent a company.

fects the programming techniques that must be used with the system. Moreover, not much progress has been made in standardizing the design of display processors. For these reasons prospective users must carefully weigh the characteristics of competitive systems.

For the past six or seven years my own professional interest has been in improving the design of display processors. My colleagues and I have designed or influenced the design of a series of display processors of increasing complexity. Although we know that the evolution of display-processor design is far from complete, we have learned much about how the design of a display processor interacts with the programming requirements for the display system. Like computer systems themselves, displays require a great deal of programming to make them useful; enlightened display-processor design simplifies the programming task.

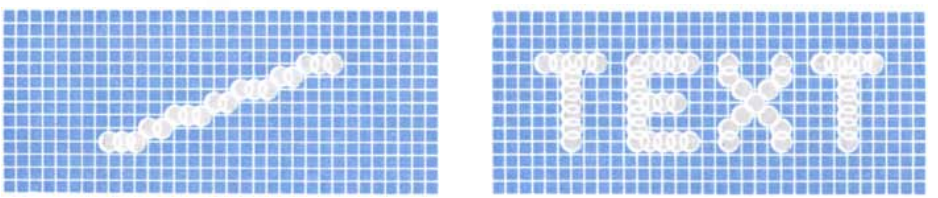
The objective of most computer-graphics programs is easily stated: to represent objects of some kind and to provide a means for manipulating them. For example, if one is designing the mask for an integrated electronic circuit, the objects to be represented are the transistors, resistors, gates, wiring and other elementary components from which the circuit is to be built. In depicting an organic molecule the objects to be represented are the individual atoms of the molecule and the bonds between them. In each application the objects are represented in the memory of the computer in a format suitable to the application. For instance, it may be convenient to list the atoms one after another in a table, giving the type and the three-dimensional coordinates for each atom. It may be convenient to represent the bonds by listing each bond in a table with the index numbers of the two atoms it joins. Such a storage format is convenient not only for display but also for the scientific programs that will manipulate this "data base."

The data base for a computer-display application is utilized in many ways. If the user wants to change the program he is working on, he may give commands that cause the computer to change the data base by replacing, adding or deleting information. Thus the user might add a new atom to his molecule. If he is exploring some line of thought, he may ask that computations be done on the data base. For example, the computer might compute the total weight of the molecule represented, or it might compute new coordinates for the atoms corresponding to some quantum-mechanical law the user is trying to understand. Finally, the user will want to see some pictorial representation of the data base on his computer-display system. It is important to note that the pictorial representation of the information in the data base is only one of many ways the information is used.

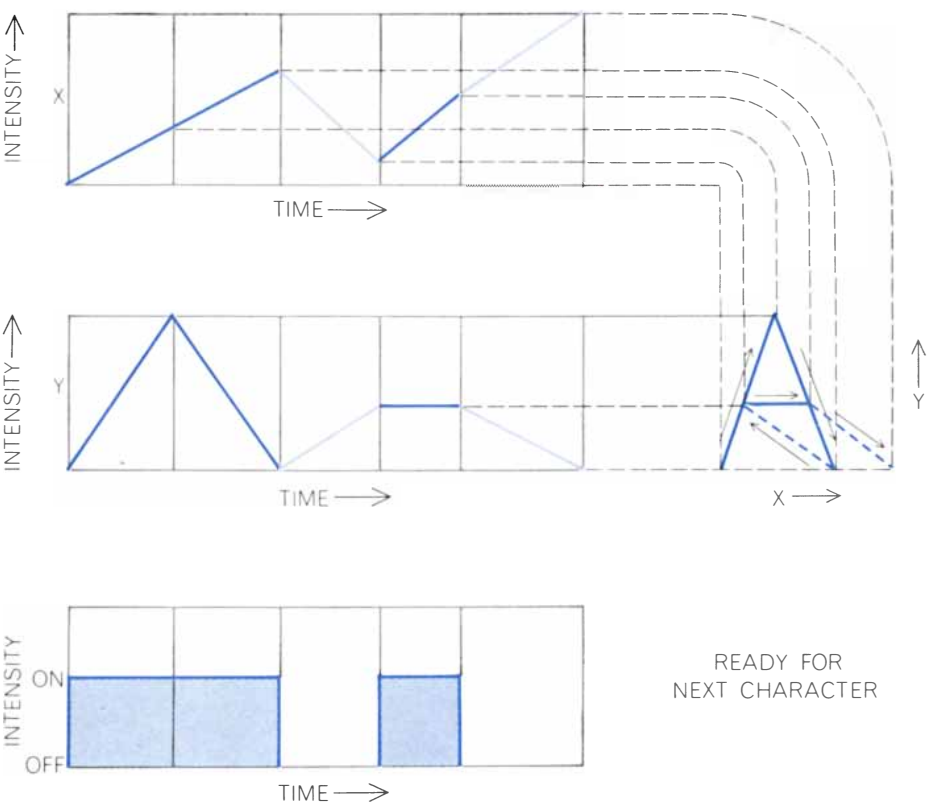
The Data behind the Display

The format of the information in the data base is not usually suitable for direct use by the display processor. The data base represents all the relevant facts about the objects being considered, whereas the display processor needs information about the particular appearance of the objects. For instance, the data base may distinguish among different types of chemical bond, all of which look the same on the display. Or the data base may identify different kinds of atoms according to their atomic number, whereas in the display they are to be presented with a different code letter, such as *C* or *H* or *N*. The appearance of the objects can be derived from the data base by a suitable program and stored for use by the display processor. The results of this computation form a "display file" [*see bottom illustration on next page*]. The display file is a secondary representation of the objects stored in the data base.

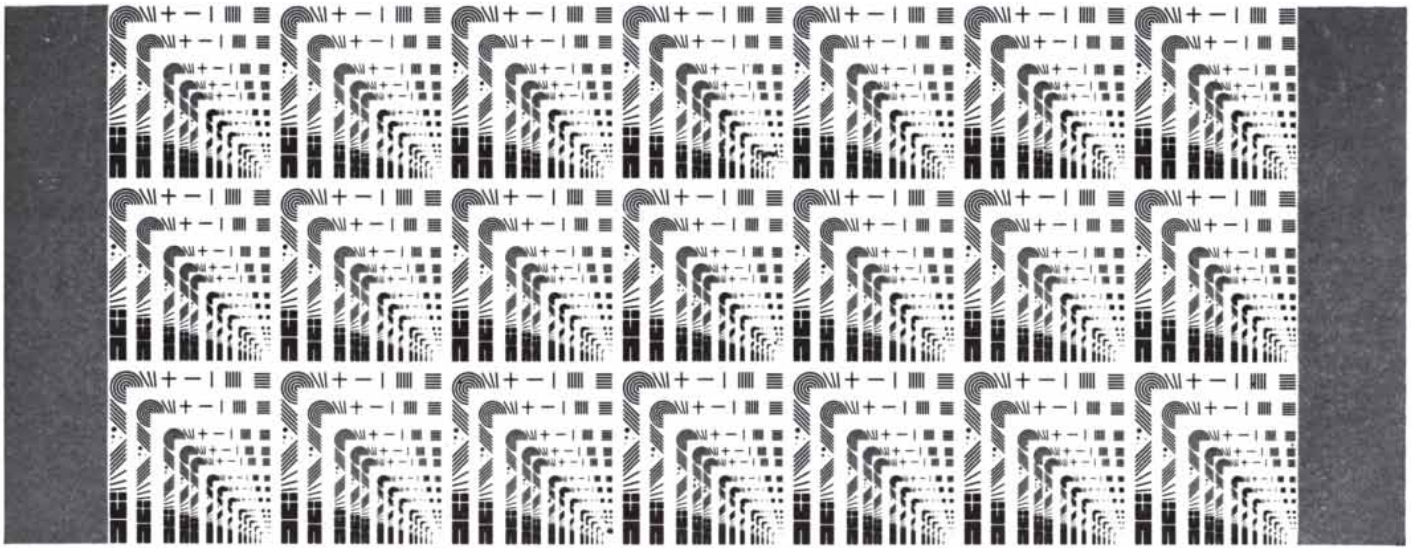
Conversion of the information from the data base to the display file involves geometric computation, selection of the relevant information and expansion of compact notations into individual lines and characters for display. Thus the two-dimensional coordinates required to make a perspective presentation of the molecule from a selected viewpoint might be computed from the three-dimensional coordinates known for each atom. If a magnified view of some portion of the object is desired, it can be



DOT-BY-DOT method of generating lines and characters was originally used in computer graphics. Newer analogue techniques are much faster and create cleaner images (*see below*).



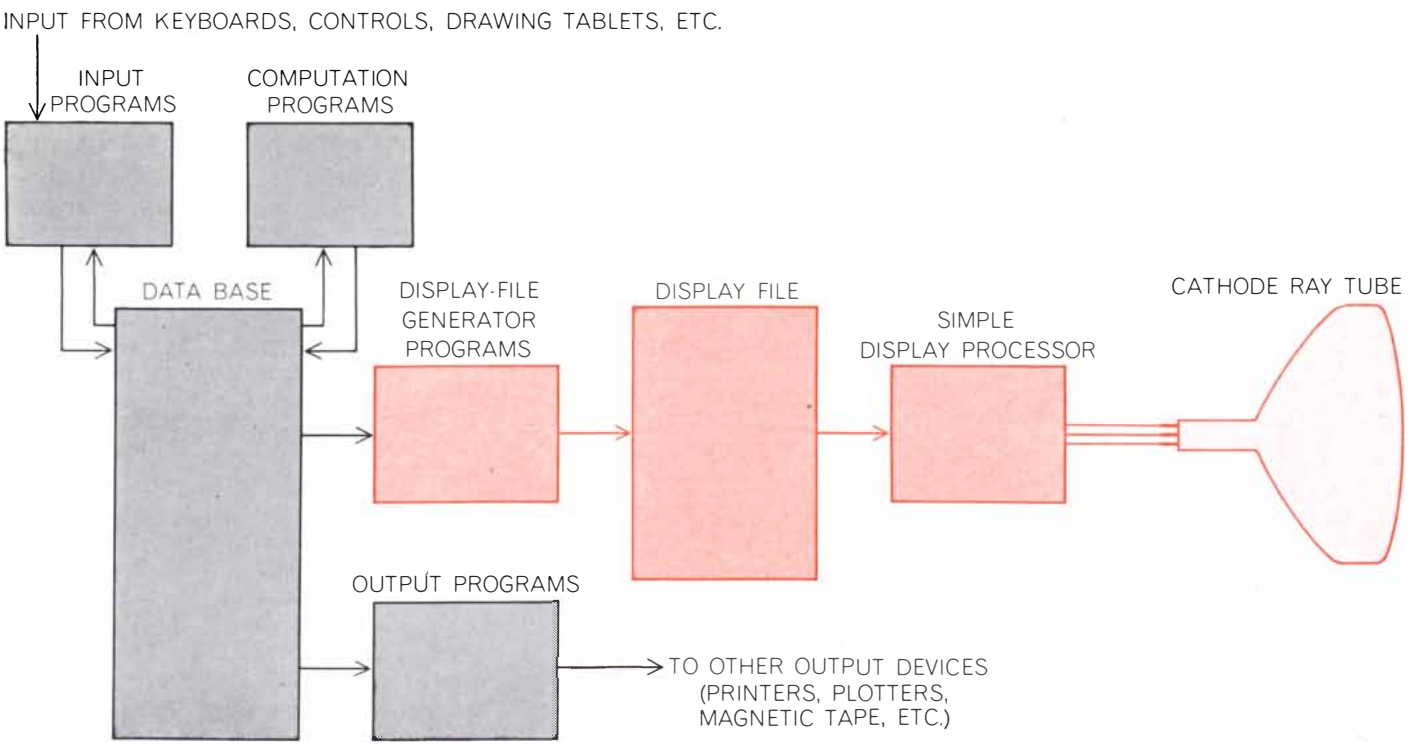
ANALOGUE CHARACTER GENERATOR produces separate wave forms for the *x* and *y* deflections of the electron beam, along with an on-off intensity pattern. Such display systems can write a character in a few microseconds and a page of text in several milliseconds.



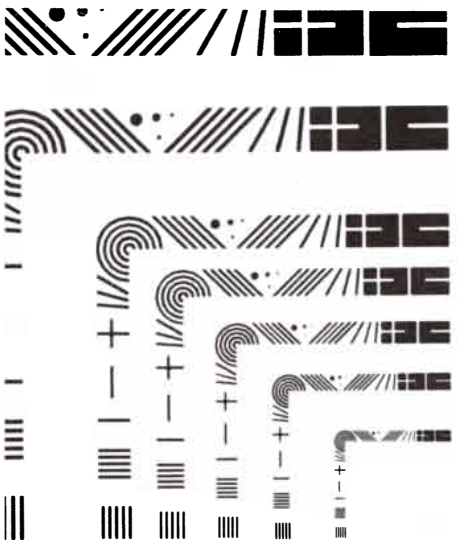
HIGH-RESOLUTION LASER BEAM produced this test pattern using a new computer-display technique developed by the Bell Telephone Laboratories. One-third of the 15-by-21-centimeter test pattern is reproduced here actual size. Normally the device is used to make masks for integrated circuits. Modulators and lenses control the fine beam of an argon laser, which is deflected to scan a photographic plate by a 10-sided mirror rotating on air bearings. The laser beam produces an image consisting of 32,000 scan lines with 26,000 image points per line. The accuracy of the beam can be controlled to better than one arc second, equivalent to a deviation of less than a third of an inch in a line a mile long. Not only is the device extraordinarily accurate but also it can produce

provided by geometric scaling as the display file is prepared. Particular items of information in the data base can be selected for display or omitted entirely. For example, it may be desirable to omit from the picture all hydrogen atoms and all bonds connected to them so that the heavier part of the molecule can be seen more clearly. Finally, the compact notations of the data base can be expanded into appropriate symbols on the screen. Each atom might be converted into a circle with a letter in it; each bond might be converted into a pair of lines (in perspective, if that is desired).

It has long been recognized that the dual representation commonly used in display applications—the data base and the display file—leads to a complex program. In many display systems the user can point to an object of interest on the screen by using a light pen, a device that identifies a specific point on the screen by producing a brief pulse when the electron beam sweeps past that point. This tells the computer that some object in the display file is to be changed. The



TYPICAL PROGRAM STRUCTURE for computer displays has the form shown here. Information about objects is stored in the data base. Information about the appearance of the objects is stored in the display file. The actual appearance of the objects is specified by programs in the display-file generator. These programs perform geometric operations and expand the definition of objects.

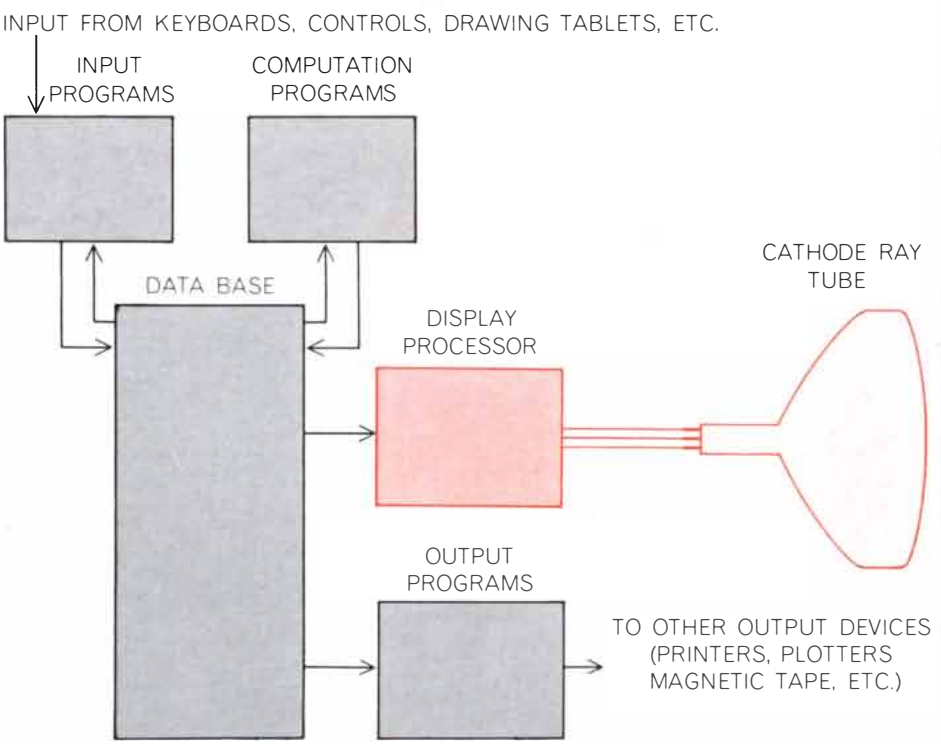


in 12 minutes a circuit mask that formerly took 12 hours of machine time. A portion of the test pattern is enlarged 10 diameters at the right. The effective resolution exceeds 40 lines per millimeter, which is remarkable over a field measuring 15 by 21 centimeters.

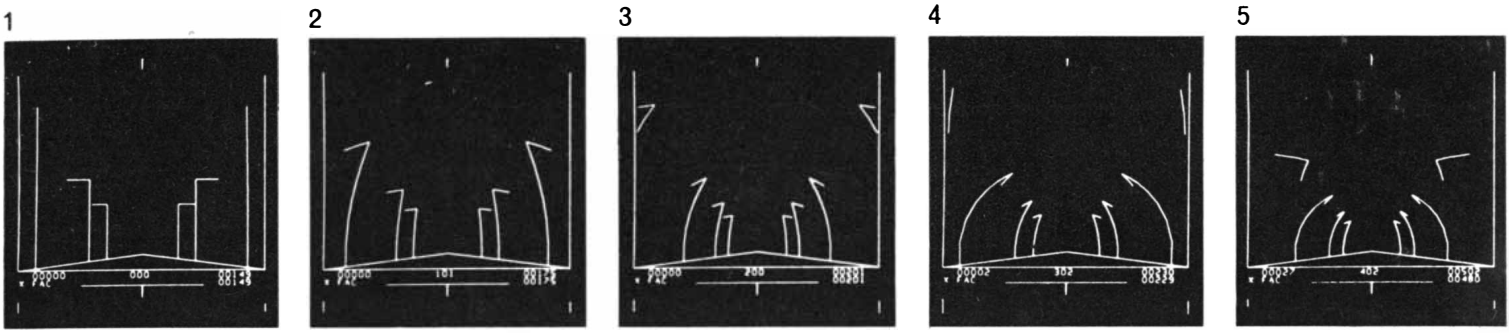
computer must somehow be able to relate each item in the display file to the item in the data base that caused it to appear.

If, having identified the item in the data base, the computer is instructed to change it, the computer must then also change the relevant parts of the display file. If the display picture is to be kept up to date so that changes appear promptly after they are initiated, computation of the display file must be done quickly. Unfortunately general-purpose digital computers are not well suited to this kind of geometric computation. For all these reasons designers of display processors have begun to provide more computing capability in the display processor to enable it to interpret the data base directly, thus eliminating the need for a display file altogether [*see illustration below*]. If the design is successful, changes in the data base will instantly be reflected in the picture and an identification of a part of the picture will instantly identify the corresponding item in the data base. Specialized display processors designed to handle such graphical computations are now beginning to reach the market.

An early innovation in display processors, now almost universal, provided them with the ability to display many symbols from a single symbol definition in memory. With this capability the display file can contain a single definition of a transistor symbol and several references to that symbol's definition. For each reference the display processor fetches the individual lines of the symbol from the single definition and traces the transistor symbol again in a different position on the display screen [*see "a" in illustration on page 79*]. Obviously all coordinate information in the symbol is



ADVANCED DISPLAY PROCESSOR eliminates the display file and display-file generator by providing more powerful computation and selection functions. This simplifies programming by eliminating elaborate cross-references between the data base and the display file.



RIDE IN "ROCKET CAR" down a "country road" was programmed by a group under Judah L. Schwartz at the Massachusetts Institute of Technology, which has found computer displays valuable in demonstrating abstruse physical phenomena. The road is lined with telephone poles shaped like inverted *L*'s. The middle number under the picture shows the rocket's velocity as a decimal fraction of the velocity of light (the decimal point before the first digit is omitted). The number at bottom left of the screen is the number of telephone poles passed. The upper number at bottom right is the reading of clocks by the side of the road; the lower number shows the time aboard the rocket. The rocket-borne clock shows a lower cumulative reading than the roadside clocks do,

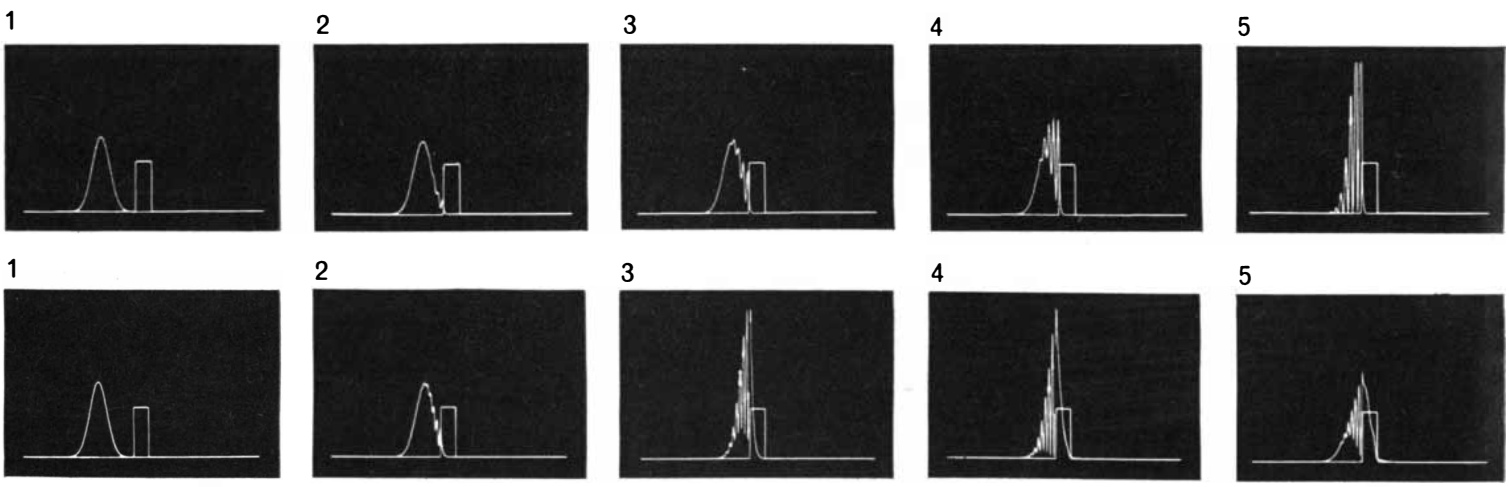
stored with respect to the location of the symbol, and the location for the symbol must be established before the symbol definition is called on. This symbol capability in display processors is identical with the subroutine capability found in all general-purpose computers. Like the subroutine capability, it makes possible more compact coding because many symbols can be created from a single definition.

Unfortunately including subroutine capability in the display processor tends to introduce new complexities in the display programming. A display processor capable of handling subroutines must remember during the execution of each subroutine the location of its call; if the display file is changed, such information may be lost. Moreover, the number of parameters that can be provided to display subroutines is generally limited to a simple indication of position. Symbols of different sizes must have different definitions in the display file even though the size of an object may be represented in generalized form in the data base. The display file for a processor equipped for subroutining takes on a structure that may correspond in part, but not entirely, to the structure of the data base, and the display-programming system has to decide when to call on this structure. The subroutining capabilities in the display processor prove most useful when the structure that the display processor can interpret is sufficiently rich to represent everything wanted in the data base. If the symbol parameters available in the subroutining capability cannot handle the variability of the symbols represented in the data base, it is often not worth the effort of using the subroutining capability at all.

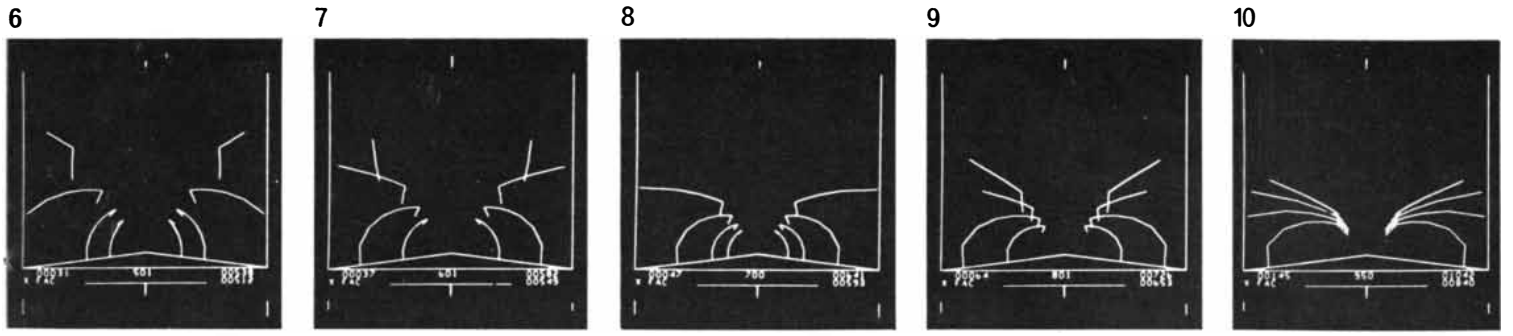
The fact that information in display subroutines can be specified with respect to a symbol origin makes it possible to specify a position that is outside the boundary of the display screen. Unfortunately a capacity of only 10 or 12 "bits" of information is usually provided in the memory registers that hold the screen coordinates. Thus if the displacement of an image produces values that lie outside the screen, those values will be truncated to the required 10 or 12 bits by the removal of the most significant parts of the sum.

The "Windowing" Problem

The effect of such truncation is that the coordinate positions immediately beyond the right edge of the screen appear at the left edge of the screen, and the coordinate positions just above the top of the screen appear at the bottom of the screen. Symbols that overlap an edge of the screen will reappear at the opposite edge. Consequently if screen coordinates are merely truncated to 10 or 12 bits, the screen will appear to have



QUANTUM-MECHANICAL SCATTERING, representing what happens when a particle of given energy meets a barrier, is depicted in these M.I.T. computer displays. The particle is represented by a Gaussian curve, or wave packet, which is the probability distribution for finding a particle, such as a proton, in a given region of space. In the upper sequence the particle has only half the energy needed to pass through a square barrier. Nevertheless, at the peak of the collision (*frames 5, 6*) the wave packet penetrates the barrier, thereby indicating a finite probability for finding the particle in a region that should be inaccessible to it according to



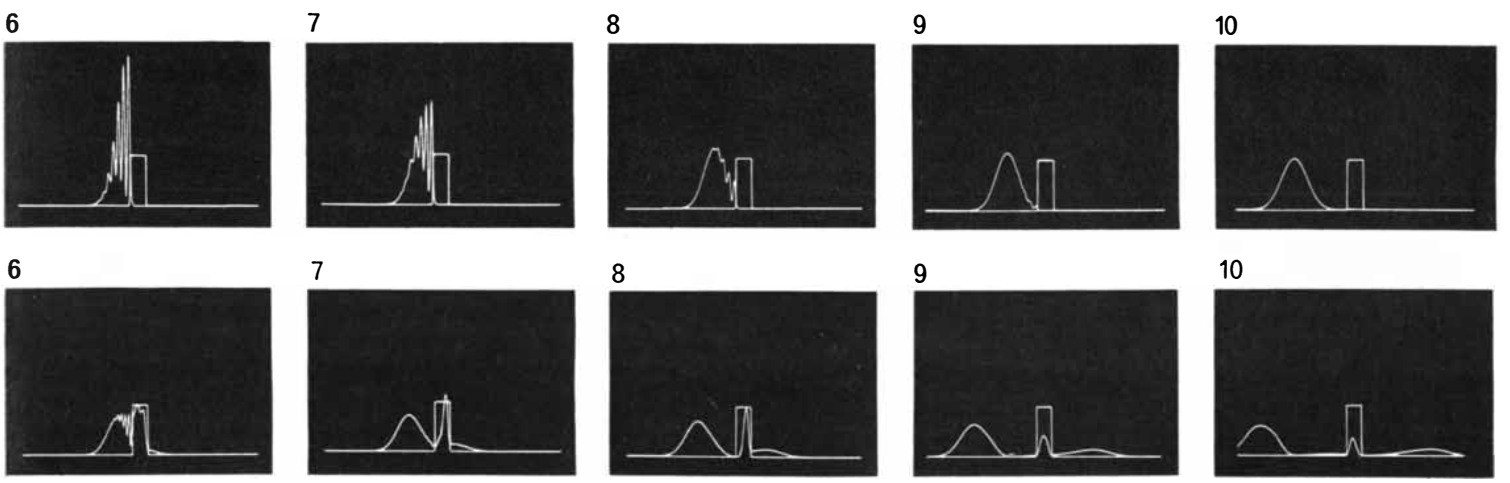
demonstrating the dilation of time experienced by objects (and observers) traveling at a large fraction of the velocity of light. As the rocket car accelerates, the telephone poles appear to bend over the road, as viewed by an observer in the car. Actually they are leaning away from the observer because the tops of the poles are farther away from him than the bottoms of the poles and thus light from the tops takes slightly longer to reach his eye. Therefore he sees the top of each pole at an earlier point in time, hence farther down the road, than the bottom of the pole. The horizontal members atop the poles appear to rotate away from the observer and finally to bend down. This is a more complicated result of the special theory of relativity and was not put into computer program.

the topology of a torus. The phenomenon is particularly annoying if one end of a short line moves off the edge of the screen only to reappear at the other edge. The line between such end points will cross the entire screen going in the wrong direction.

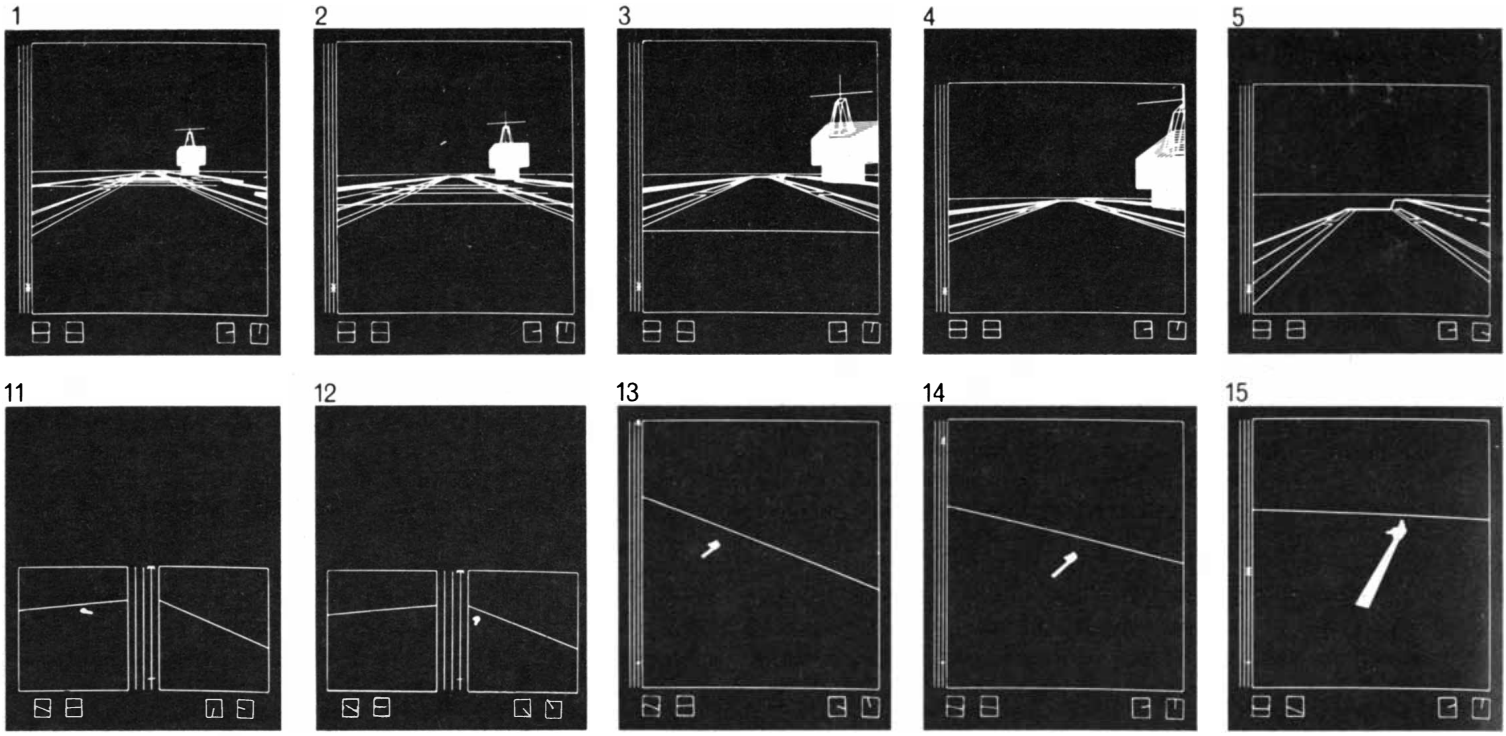
The cures for these well-known difficulties all provide the display user with a drawing area much larger than the actual display screen. By one means or another the display tube is made to behave as if it were a window through which this much larger drawing can be seen. The computations involved in providing this illusion are called windowing. When properly provided, windowing seems to be a natural part of the way the display operates. Objects that move off the screen disappear smoothly at a well-defined edge. Only when windowing is missing from a computer-display system is the absence of this feature immediately evident.

There are two basic methods of windowing. One method is to provide the display system with larger analogue deflection voltages than are needed for covering the screen itself and simply to blank out the beam anytime it is deflected beyond the edge of the screen. This procedure, although it is widely used, has the shortcoming that it places an additional burden of accuracy on the analogue deflection system and that it must of necessity operate slowly, since there is no way to speed up the beam when it is drawing parts of the picture outside the viewing area. The second method of windowing is to provide a mechanism to precompute digitally the actual coordinates of the part of the picture that should appear in the window. In this method the display tube is provided only with valid data and so does not waste time drawing invisible parts of pictures. This alternative, often called clipping, is widely programmed in software for sophisticated computer-graphics systems.

Windowing is the most important of several kinds of geometric computation needed in preparing information for presentation on a display screen. Unfortunately the programs required for windowing in an ordinary general-purpose digital computer, even a big and fast one, can process only about 1,000 line segments per second. If additional computations are needed, for instance to present a two-dimensional perspective view of a three-dimensional object, the operation takes substantially longer. Recently my colleagues and I have spent much time developing special display processors that overcome these delays by transferring the necessary computations to high-speed, special-purpose digital circuits. These processors make it easy to obtain rapid, dynamically changing views of three-dimensional objects [*see top illustration on next two pages*].

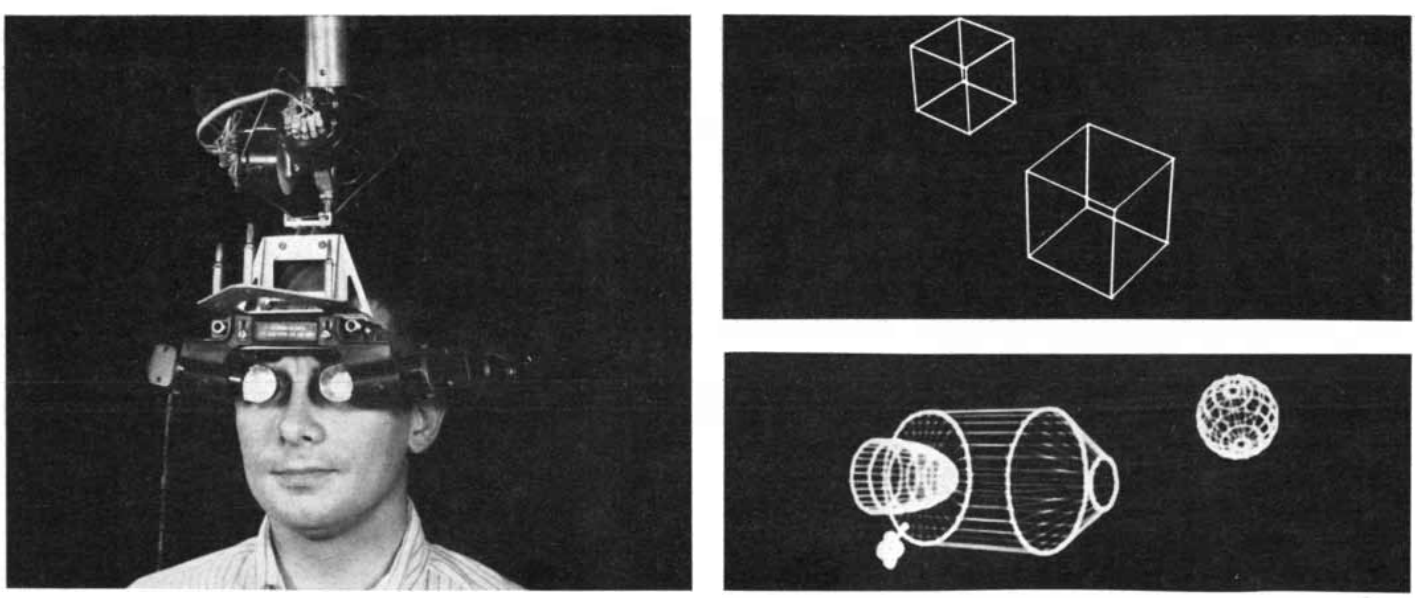


the pre-quantum-mechanical description of nature. Subsequently the packet is reflected and the probability distribution is flattened. In the lower sequence the energy of the wave packet is exactly equal to the barrier height. In classical physics the particle should lose all its energy on entering the barrier and remain there forever. In quantum physics a small probability packet gets temporarily trapped inside the barrier (*frames 7, 8, 9, 10*) while most of the packet is reflected. In addition a small, flat packet is transmitted. This indicates the probability of a particle's "tunneling" through the barrier when the experiment is repeated many times.

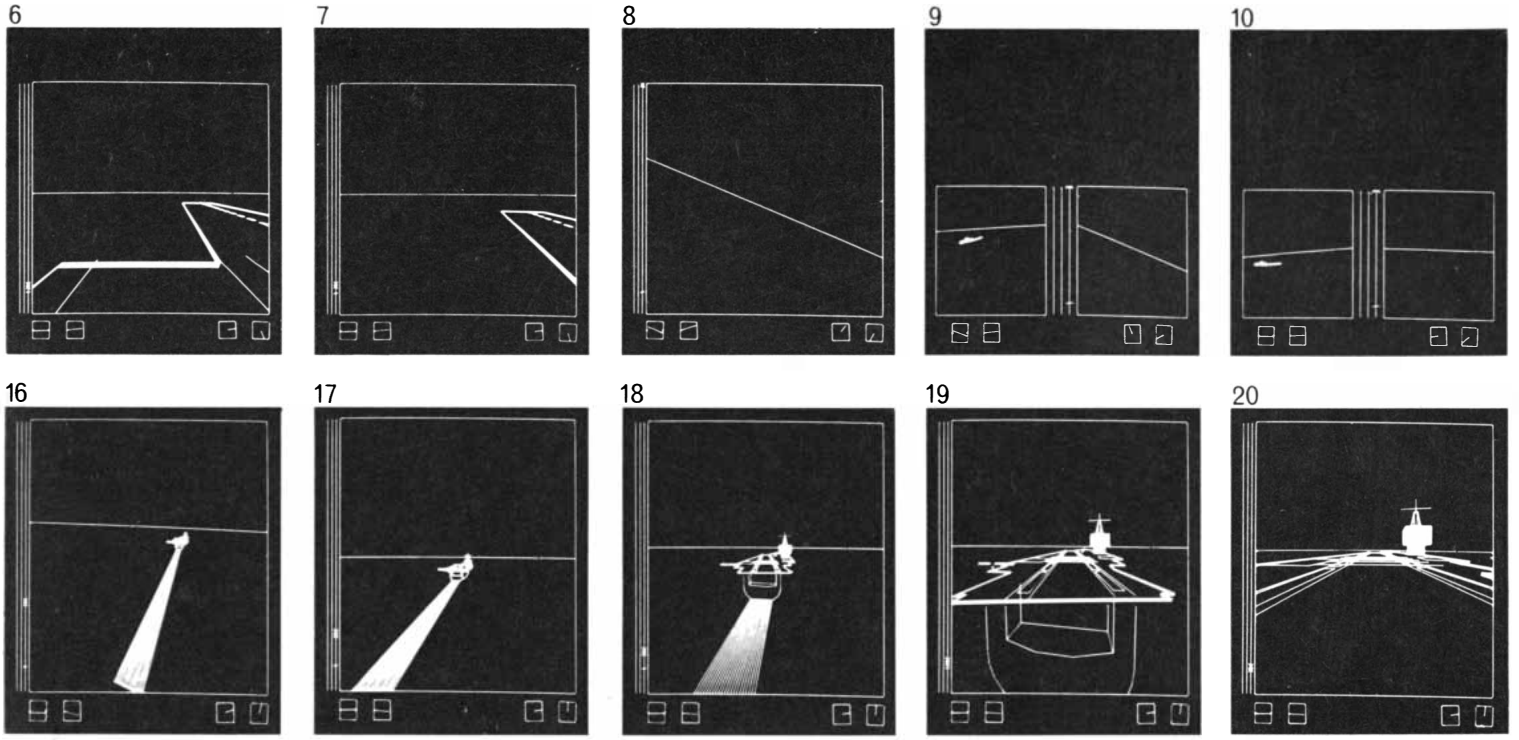


TAKEOFF AND LANDING ON AIRCRAFT CARRIER have been simulated by the Evans & Sutherland Computer Corporation for the Naval Training Device Center. The display system used for this demonstration is installed at Bolt, Beranek and Newman Inc. It follows the design illustrated at the bottom of page 67, in which circuit hardware in the display processor replaces a "software" program for handling information stored in the data base. Only such a system provides the computing speed needed for this kind of real-time simulation in which two complex objects are moving rapidly in three-dimensional space. Frame 1 shows what the "pilot" sees looking down the deck before takeoff. Through a keyboard on the computer he can adjust the engine throttle and the attitude of the aircraft. At the left of the screen is a vertical line with a cross that indicates the aircraft's altitude. The two boxes at the bottom right

I became involved in the development of such high-speed displays through a desire to present an observer with a simulated three-dimensional environment. My idea was very simple: Mount miniature cathode ray tubes on the user's head, one tube in front of each eye, so that the computer can control exactly what he sees. Measure the position of the user's head and compute a perspective picture appropriate to that viewing position. As the user turns his head, the perspective picture should change just as if the object portrayed were really in the room with him. There is no need to measure his eye position because the picture presented on his glasses is not changed by the part of it on which he chooses to focus. Because the three-dimensional objects he sees will appear to remain stationary in space as he moves around



HEAD-MOUNTED DISPLAY at the University of Utah employs a display processor similar to that used for the aircraft carrier sequence. Two miniature cathode ray tubes are built into the goggles. A mechanical linkage tells the computer where the viewer is looking at each instant. The display processor instantaneously supplies the correct image for each head position. The viewer is free to look anywhere in a 360-degree circle and can look up and down through an angle of about 45 degrees. Two samples of what he sees are shown at the right. The objects grow larger or smaller and move with relation to one another as the observer moves around.



of the screen indicate the aircraft's compass heading (east at takeoff) and the direction to the carrier's superstructure. At bottom left are an artificial horizon and a rate-of-climb indicator. Frames 2 through 7 show the aircraft moving down the deck and becoming airborne. The pilot sees the horizon tilt (*8*) as he banks to the left. In the next four displays (*9, 10, 11, 12*) the view from the left window of the aircraft is shown on the left half of the screen while the view ahead appears on the right half. Ordinarily the two windows are on separate screens in their normal relation to the pilot. As the pilot continues to bank, the carrier disappears from the left window and reappears in the front window (*12, 13*). Still flying counterclockwise in a large circle, assisted by being able to see the carrier's wake, the pilot lines up the carrier for his approach (*14, 15, 16, 17*), and then finally levels off for a perfect landing (*18, 19, 20*).

them, an illusion that the objects are present in the room with the observer should be created. The equipment illustrated at the bottom of the opposite page was designed to create this illusion.

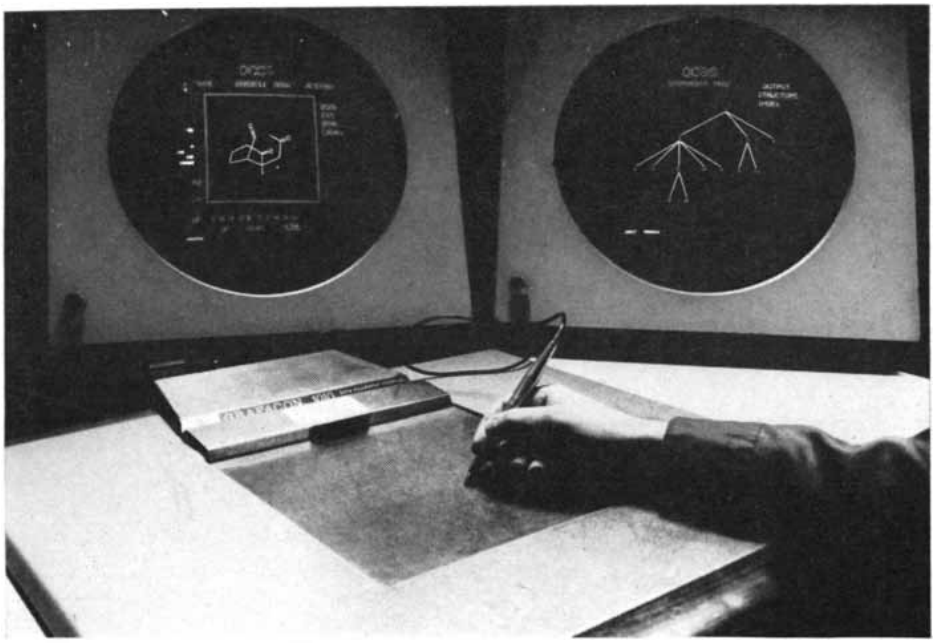
The head-mounted display does indeed provide the illusion of seeing three-dimensional objects at identifiable locations in the user's environment. Because the objects presented are transparent and made of glowing lines essentially free of texture, no one is fooled into thinking they are real. On the other hand, even naïve observers are able to identify the size of the objects portrayed and are able to move themselves into an appropriate viewing position to examine a particular feature of the objects portrayed.

Logic of High-Speed Displays

The significant thing about the head-mounted display project is not the operational results but rather that it forced us to think carefully about the problems of presenting a real-time perspective picture. We needed a solution to the windowing problem in three dimensions because as the user turns his head away from objects they move out of the field of view provided by the optical system. If the objects are to move smoothly out of view, the correct end points for the parts still in view must be very rapidly computed, and prompt decisions must be made in order to omit parts that are entirely out of view. Since no general-purpose digital computer is capable of doing these computations fast enough to present dynamic views of any but the most trivial objects, we were forced to design and build special-purpose equipment to do rotation, windowing and perspective computation. Having a good solution to the windowing problem in three dimensions quickly led us to a solution in two dimensions for our latest design for a display processor.

Computing a perspective picture of a "wire frame" three-dimensional object is really rather easy. The position of each part of the object on the screen is computed by projecting the object onto the screen in straight lines [*see top illustration on page 74*]. Fortunately straight lines in three-dimensional space project into straight lines on the screen. Thus the rotation and projection computations for each line segment need be done only for the ends of the line; the center parts of the line can be filled in by the two-dimensional analogue line generator. If windowing or clipping in three dimensions is done before projection, the problems of dividing by a zero or a negative depth coordinate are avoided [*see bottom illustration on page 74*]. The digital processing equipment we developed for the head-mounted display project is able to accept information in three dimensions, convert it to an appropriate viewpoint and viewing direction and present valid display-screen coordinates to an ordinary line generator for drawing a perspective picture on the screen.

We have found it particularly useful to have a display system in which a three-dimensional computing engine is built into the display processor. In the first place, coordinate information for display can be in three dimensions instead of in two. Symbols can be scaled on the drawing as well as merely being positioned, and three-dimensional symbols can be rotated to any desired angle with respect to other objects in the drawing. Since clipping is done in the display processor, the processor can accept the stored data to a resolution much finer than the actual resolution of the display itself. Such data can be regarded as a drawing on a huge piece of paper (a piece covering a square mile, say), any part of which, or all of which, can be presented on the display screen. For many applications such capabilities can eliminate the need for a display file altogether and so greatly simplify the programming of the display. Again our hope is to operate the display

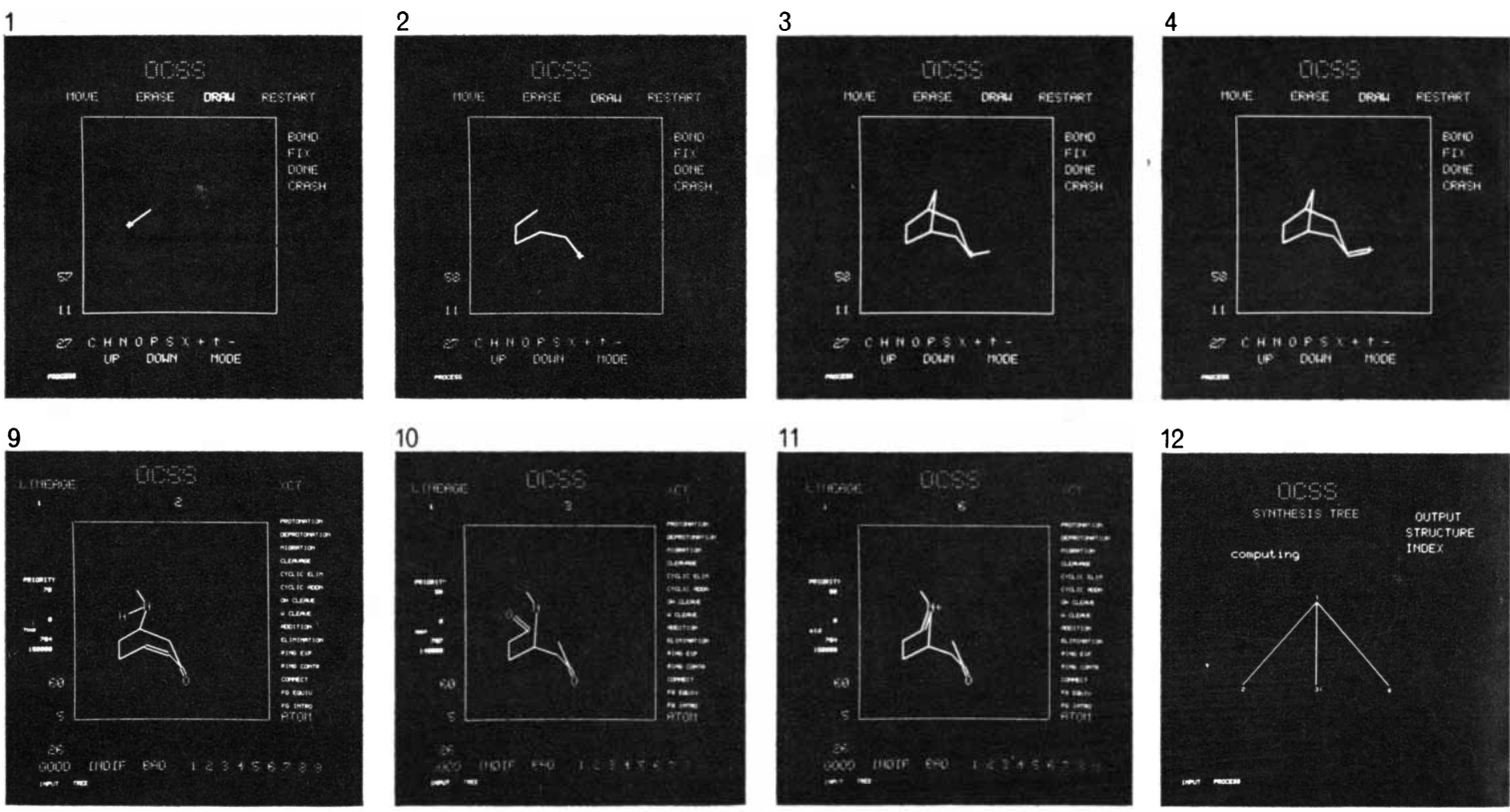


COMPUTER-AIDED CHEMICAL SYNTHESIS program was developed at Harvard University by W. Todd Wipke under the direction of Elias J. Corey. The computer is supplied with a library of chemical compounds and chemical reactions, together with a program for manipulating this information. By interacting with the computer the chemist can select the most efficient way to synthesize a desired compound. The chemist starts his "conversation" with the computer by drawing the structure of the target molecule on a writing tablet. The tablet contains a coordinate grid that relays the position of the pen to the display screen. Here a molecule related to the target molecule appears on the screen at the left, while the screen at right displays a variety of chemical pathways to the target in the form of a "tree." The development of these displays is depicted in the pictures at bottom of these two pages.

processor directly on the data base so that changes in the data base are instantly reflected in the display.

To provide for flexibility in the format of the data bases, we have chosen four typical forms in which information commonly appears in applications [*see illustration on page 79*]. In the first form a sequence of coordinate values is simply connected by a sequence of straight lines. This form is the common one for connecting data points on a graph. In the second form disjoint lines are represented by pairs of pointers that indicate the location of the coordinate data to be used for the end points. This format would be useful for representing the bonds in a molecule. The coordinates can be either two- or three-dimensional. The disjoint-line format is useful in a variety of pictorial representations.

The third form is the simple subroutine, where standard format parameters for position, size and rotation are provided. For all fixed-geometry symbols, such as those used in making circuit diagrams, this form is adequate. We have also designed the display processor so that it does not execute any subroutine that is entirely outside the viewing area. This provision greatly



CHEMICAL SYNTHESIS PROBLEM is presented to a venerable PDP-1 computer at Harvard by a chemist who uses a writing tablet to draw a picture of the molecule he wants to make. By frame 3 he has drawn the basic structure of a molecule with eight carbon atoms. He adds a double bond at the extreme right of the molecule (*frame 4*), then bends the bond into the proper stereochemical configuration (*5, 6*). Addition of an oxygen atom (*O*), a nitrogen atom (*N*) and another carbon atom (*indicated by the line projecting from the nitrogen*) completes the skeleton structure of the target molecule: longifolene (*7*). The computer is now asked to search for somewhat simpler compounds that can be converted to longifolene in a single chemical step. The result of the search is shown by the tree in frame 8. Longifolene is designated by the 1 at the top of the tree; three related compounds, designated 2, 3 and 6, are

speeds up the display of magnified views of complex objects. Finally, complete subroutines are available for displaying objects that are similar in a nongeometric way. For example, cross-hatched boxes can be defined by position, size and type of cross-hatching. Such boxes are not geometrically similar because the number of cross-hatch lines must be computed from the size of the box; moreover, the angle of the cross-hatching must be constant regardless of the size or orientation of the box.

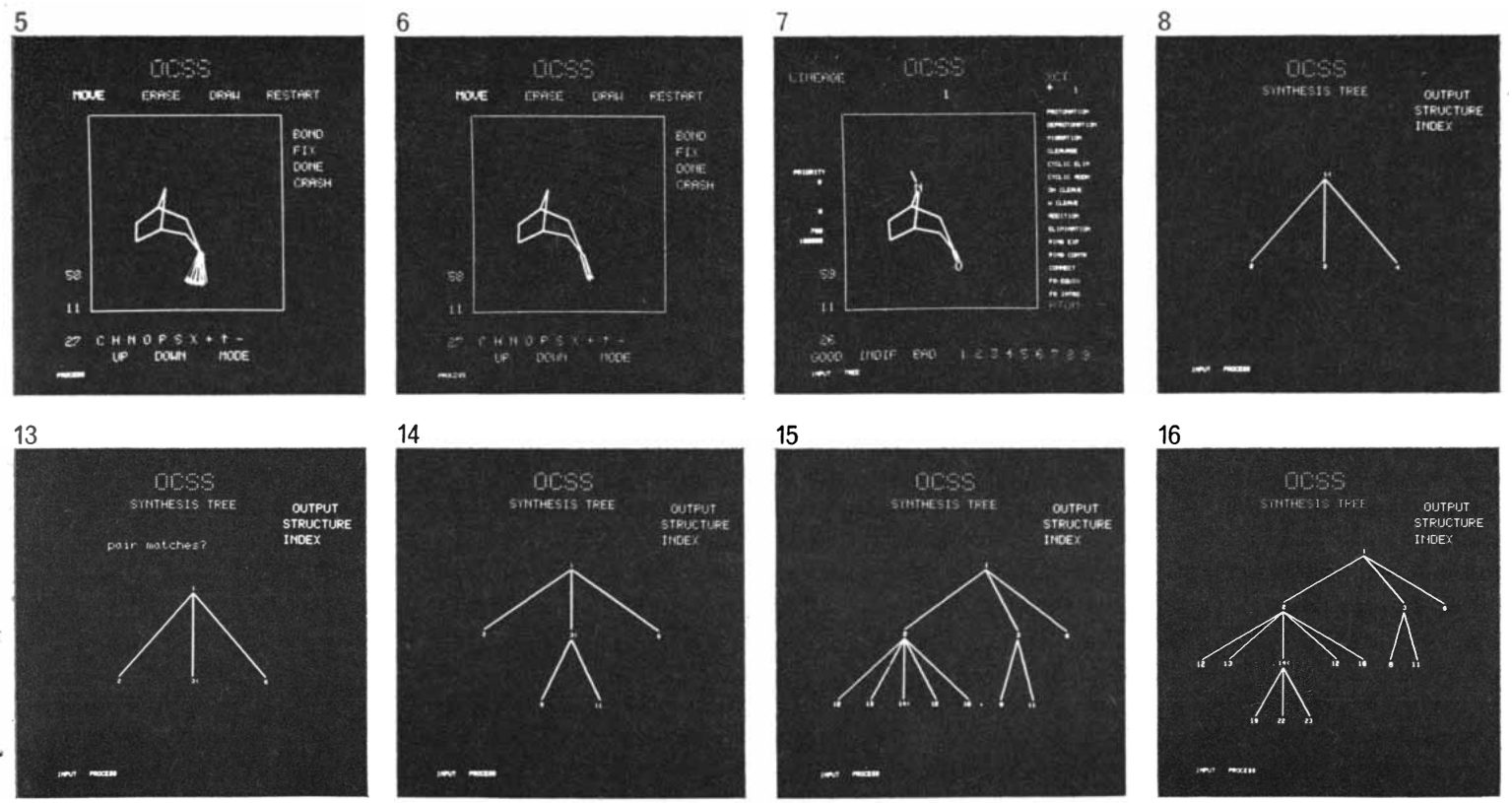
The Hidden-Surface Problem

In spite of the advances that have been made in computer graphics in the past decade, most of the objects displayed are still in the form of line drawings. It is difficult to present solid objects with today's computer displays; indeed, only half a dozen research installations even try. The preliminary efforts are nonetheless rewarding. The difficulty is in computing what parts of an object are visible from a given viewing position and what parts are not. If an object represented in three dimensions is to be rendered as being solid, some parts of the object must be able to obscure other parts hidden behind it. Many workers have tackled this "hidden-surface problem," and in the past few years several interesting solutions have been developed.

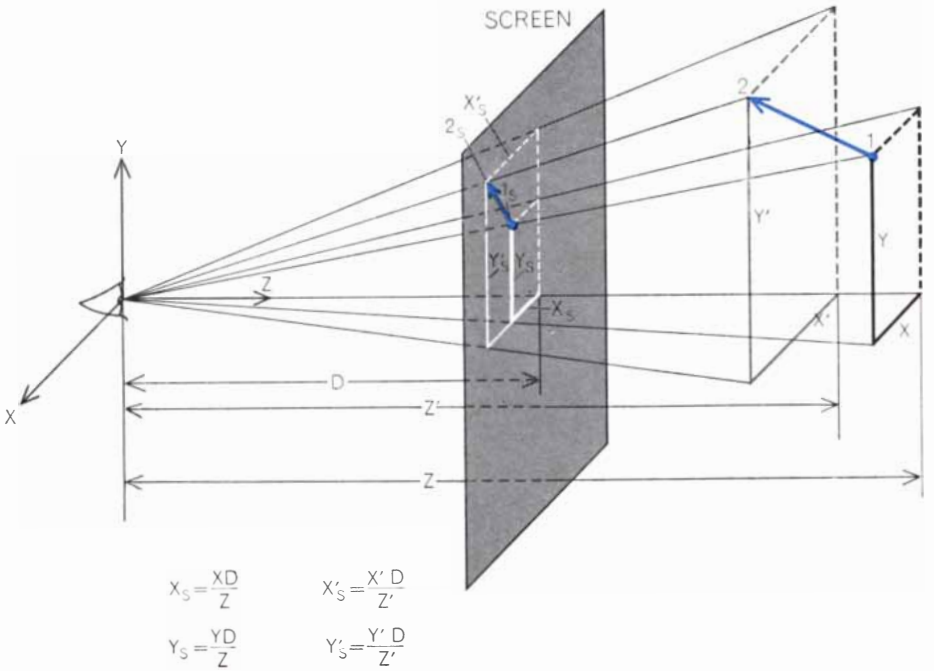
The effectiveness of a solution to the hidden-surface problem can best be described in terms of the growth of the solution time with the complexity of the situation. Almost any method will work adequately with very simple situations, but because the computation time required by some methods grows with the square of the situation complexity, much computation time may be required for complex situations. Four years ago I showed in *Scientific American* a computer-drawn picture of a solid object produced by Lawrence G. Roberts, which was typical of the best then available [see "Computer Inputs and Outputs," by Ivan E. Sutherland; SCIENTIFIC AMERICAN, September, 1966]. His program took 1.5 minutes to draw a picture containing 40 separate blocks from which hidden lines had been removed; the computing time was roughly proportional to the square of the number of objects examined. Today general-purpose computers of not much greater power can produce pictures such as those shown on page 56 in about a third of the time, and the growth law of these programs appears to be about $N \log N$, where N is the number of objects.

The traditional approach to the hidden-surface problem was to compare each object in the picture with every other object to see which was obscured. This procedure requires $N - 1$ comparisons for the first object, $N - 2$ for the second object, $N - 3$ for the third object, and so on, for a total of $(N^2 - N) \div 2$ comparisons. Moreover, the computations must be executed in three dimensions with high precision because they are done before projection onto the screen. The new methods consider only the task of computing the picture rather than the task of comparing objects. If a picture of the television type is to be produced, there are, after all, only some 250,000 points on the television screen at which intensities can be displayed. Why bother to compute to any better precision than that?

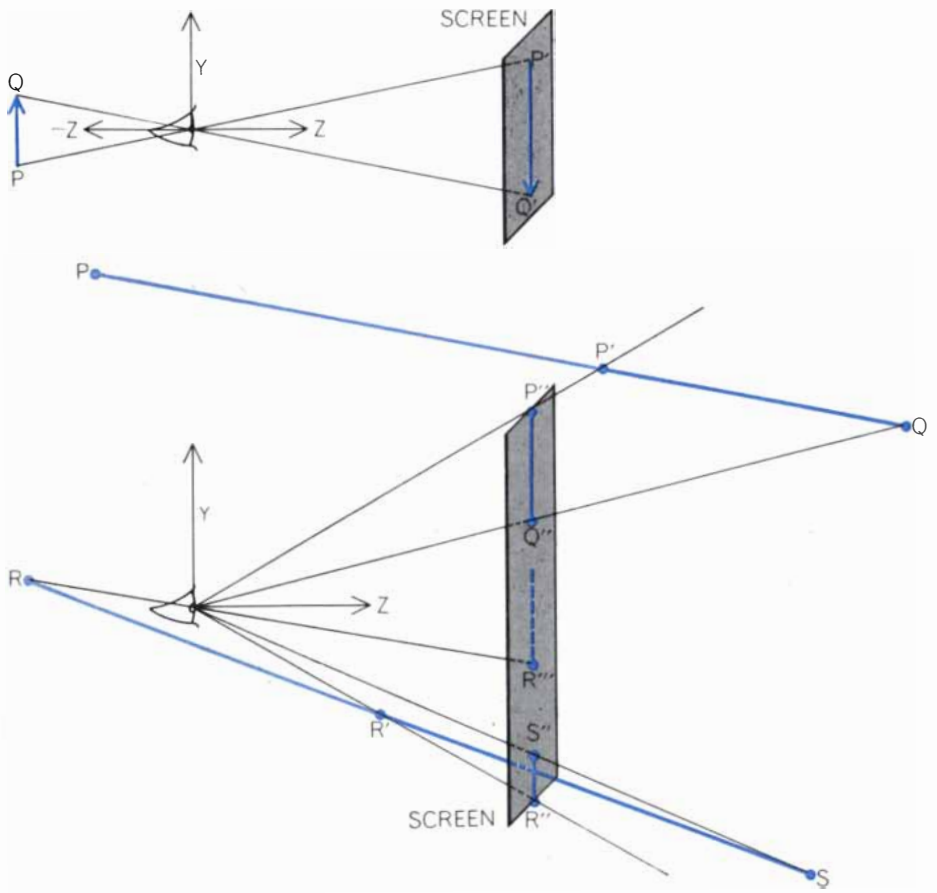
The most elegant, although perhaps not the fastest, algorithm to date is one invented by John Warnock of the University of Utah and somewhat embellished by others (including myself). Warnock's algorithm concentrates on computing shades for parts of the



listed at the ends of the branches of the tree. The chemist asks to see the structures of these three molecules and they are presented to him in frames 9, 10 and 11, together with coded information at the left of the screen that describes the chemical reaction that will convert each to longifolene. The computer also evaluates each reaction and assigns a priority, or preference. The computer is then asked to find related compounds that can be converted into compound No. 3. While the search proceeds the computer can be asked what it is "thinking." The words "Computing" (*frame 12*) and "Pair matches?" (*frame 13*) appear on the screen. The last remark indicates the computer is puzzled whether a particular molecular structure is identical with another one. Two routes to compound No. 3 appear in frame 14 and five routes to compound No. 2 appear in frame 15. In the last frame the search is carried still deeper.



PERSPECTIVE PROJECTION for a computer display is obtained by placing the origin of the coordinate system at the observer's eye. The projection of all points on the screen is then readily computed from the geometry of similar triangles. The diagram shows how the two end points, *1* and *2*, of an arrow are projected to the points 1_s and 2_s. The first equation states that x_s is to x as the distance to the screen, D, is to z. Other dimensions are obtained similarly. Objects extending beyond the screen must be "clipped," as shown below.



"CLIPPING," OR "WINDOWING," PROBLEM arises when objects in real space extend beyond the limits of the screen or lie behind the observer. In the latter case (*top*) z has a negative value. Although the equations are correct for such points, the result would be to project the object upside down. The problem can be avoided (*bottom*) if all lines are first clipped off at planes defined by the edges of the screen. If one applied the usual equations to the point P, a numerical overflow would result. Point R in the line RS would project to point R''', and points close to R would project upward on the screen. By proper windowing, or clipping, the line $P'Q$ appears on the screen correctly as $P''Q''$ and $R'S$ appears as $R''S''$.

screen, and also on computing shades for as large an area as is possible at one time. I shall describe Warnock's algorithm in some detail not only because it shows much ingenuity in the solution of a difficult problem in computer graphics but also because it utilizes a particularly appealing technique of computation.

Data for Warnock's algorithm are defined in terms of a collection of surfaces that form the objects to be shown. The simplest computations result if these surfaces are planes bounded by straight lines, but the algorithm could be applied to more complex surfaces. Each surface is assigned a color and a shade, and each surface is considered to be opaque. Rotation and translation are used to get the surfaces into the observer's coordinate system, and a perspective division is performed to position the objects on the screen. Depth information is preserved through the perspective projection.

Warnock's program now breaks the screen up into a few subregions and applies a standard procedure to each one. The standard procedure is to sort all surfaces by x and y position to determine whether they are entirely outside the subregion, surround it or partially intersect it. This sorting is done by first considering gross measures of the surface's position (such as: Is its leftmost point to the right of the subregion?) and then considering, if necessary, the individual edges of the surface. Obviously surfaces that are outside the subregion need not be considered further. The algorithm then considers the surfaces that surround or intersect the region and attempts to prove that certain of them are hidden by others. For example, if a surrounding surface can be shown to be in front of all other surfaces throughout the region, all the other surfaces are eliminated and the program can shade the entire region according to the shade of that surface. If there is only a single edge in the subregion, perhaps the program can fill in the appropriate pair of shades.

The important thing about Warnock's algorithm is its response to a subregion in which the situation is too complicated to handle. In such cases the algorithm simply subdivides the subregion into smaller regions. The sorting task for a subdivided region is simpler than it was for its predecessor because all surfaces that surrounded the predecessor will surround any of its subdivisions, and all surfaces that were found to lie off to the side of the predecessor will also lie to the side of its subdivisions. Only those surfaces that partially intersected the prede-

Are you sure you're at the right address?

We believe career advancement may depend as much on being in the right place at the right time as it does on ability. If you agree, then now (while there is lots of summertime for making a move) you should look into what is happening at NCR.

NCR is now the second largest company in the world in the computer systems business — supplying hardware and software answers for a whole world full of expanding markets. But the more answers we work out, the more questions we stir up.

That's why we are still growing rapidly. Why we need more engineers, of all degrees. More physicists, too. And software people. That's why NCR could be the right place — and *now* the right time — for you to advance your career.

In many positions you would be handling significantly more responsibility than you do now, because our employees advance as fast as they are able to assume additional duties.

Unless you are absolutely sure you're at the right address, tell us about your background and career objectives.

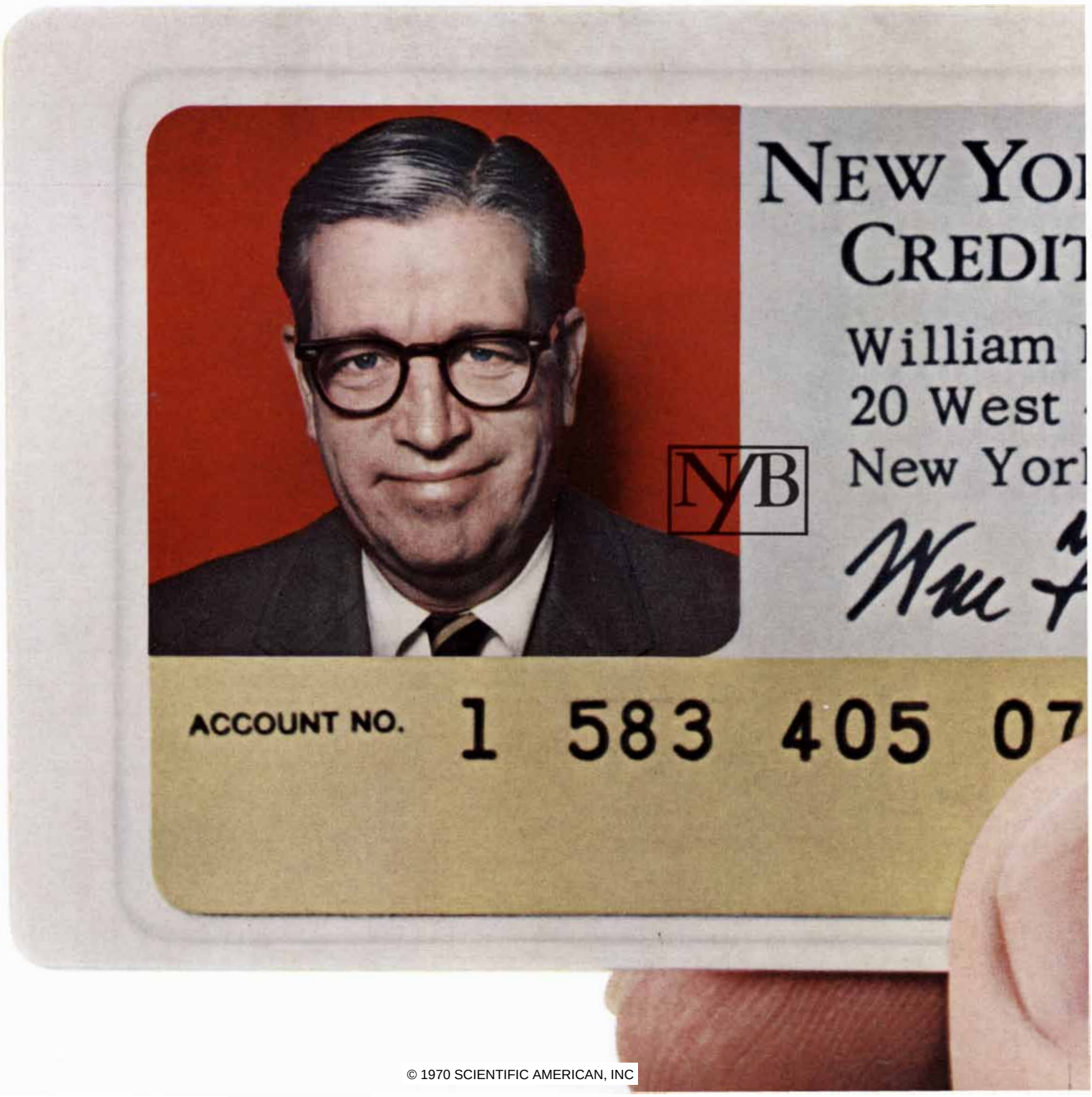
Write:
T. F. Wade, Executive and Professional Placement
The National Cash Register Company, Dept. SC-6
Dayton, Ohio 45409

We are an Equal Opportunity Employer M/F

NCR

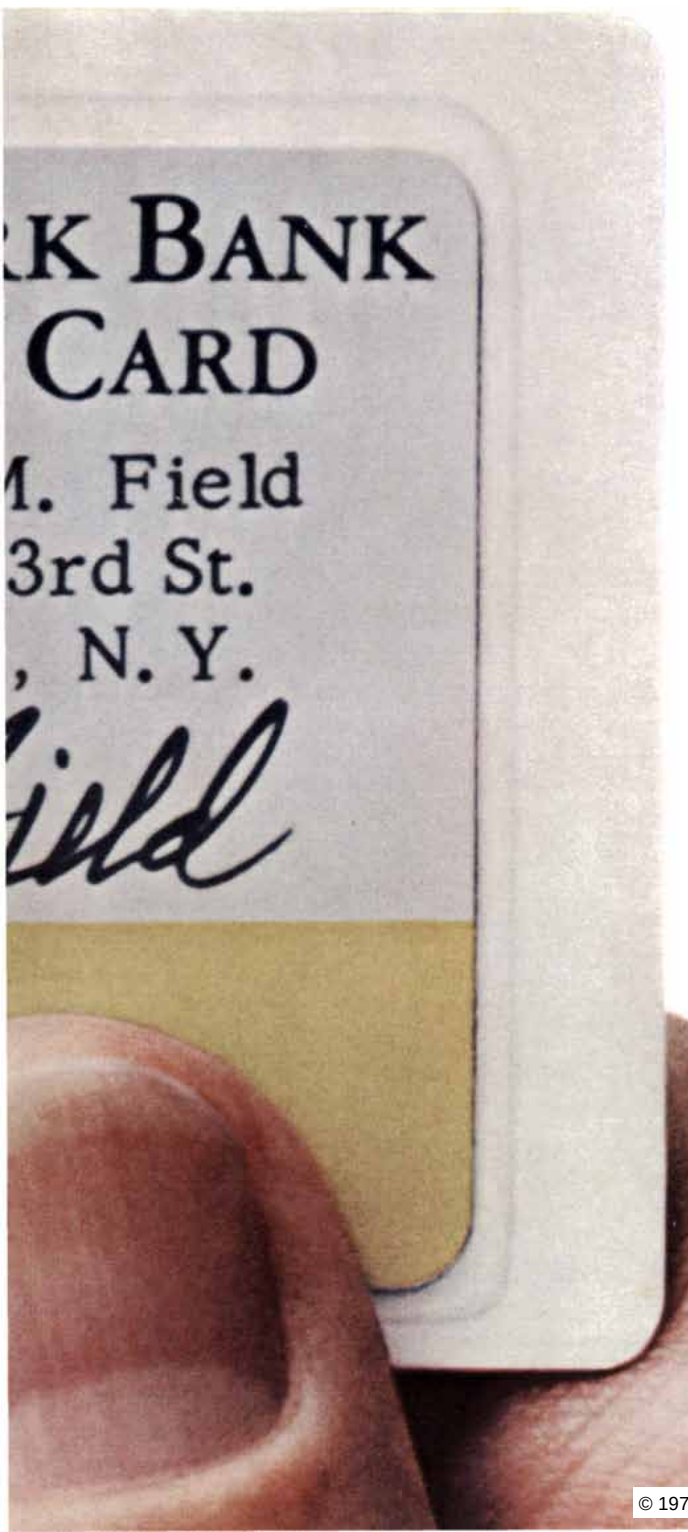
Credit card fraud is now costing Ar

But there
staring us ri



ıericans a record $200,000,000 a year.

s a remedy
ght in the face.



The number of credit cards in Americans' wallets has zoomed to 300,000,000. And the upsurge has been a bonanza for the swindler.

Credit cards are stolen from wallets, handbags, bureau drawers, and mailboxes. Some are thrown carelessly in trash baskets, and are found by scavengers uninhibited by scruples.

Even organized crime has gotten into the act. They have developed sophisticated techniques of forgery, and organized a thriving black market, where the going rate for a credit card is over $50.

Unless it provides complete and positive identification, no credit card is immune, no credit-card holder free of the danger.

Worst of all, the efforts by law enforcement and private agencies to stop the thievery have had little success. The problem has been attacked by state and federal legislatures. But the losses keep mounting.

However, there is an answer.

Of all the methods devised to keep credit cards out of the hands of criminals, none has proven as effective as the Polaroid Portrait ID System.

It puts the owner's picture (in full color) and pertinent data on a single piece of film, and seals it with a chemically bonded lamination.

It's almost impossible to alter it without detection.

As a result, few criminals would be foolish enough to try to use a Polaroid ID card.

They can forge a signature, but not a face.

Recently, there has been dramatic evidence of the degree of its security. One of the many banking institutions that have installed the Polaroid Portrait ID System gave their cardholders a guarantee of complete and unconditional freedom from financial loss due to lost or stolen cards. No deductible; no limit based on date of notification; simply no loss at all.

The Polaroid portrait ID card is made in a number of formats. So there's a style that suits just about any situation that calls for fast, positive identification.

If you'd like more detailed information, write to: Polaroid Corporation, Dept. 57-190, Cambridge, Mass. 02139.

We'll show you how to solve the problem in the snap of a shutter. Polaroid®

The Polaroid Portrait ID System.

She'd hug us if she could.

More precisely, she'd hug our John Welsh.

One of John's customers manufactures concrete statues. He casts them in 900 pound steel molds which must be cleaned of residual concrete with a hot mineral acid after each casting.

Hot acid cleans away concrete easily. So easily, that if left unchecked, it can also attack and pit the steel molds. That's why John's customer was testing MAIN™ 33S acid inhibitor—for the safety, economy and prolonged bath life it promised. But he didn't expect to test it under such severe conditions.

Using his electrically powered crane, he lowered one of his molds into the hot mineral acid. And then it happened. The power failed.

With no power to lift the mold out of the acid, John's customer was sure it would be corroded beyond use by the time power was restored.

And it would have been too if it weren't for MAIN 33S. This unscheduled acid exposure test proved that MAIN 33S establishes a barrier between the steel mold and the acid, preventing corrosion.

John knew MAIN 33S could effectively inhibit acid corrosion under normal acid cleaning conditions. But it took his customer's particular acid cleaning operation and a power failure to substantiate our claims that MAIN 33S is effective even under severe acid exposure conditions.

This MAIN 33S success is only one example of the way Abbott Chemicals works to help you find solutions to your chemical problems.

One of our jobs as the chemical marketing arm of Abbott Laboratories is to find industrial applications for many of the 1000 chemicals Abbott synthesizes each year.

To do this, we rely heavily on our people. The John Welshes who help their customers find solutions to their chemical problems. Who find new uses for our existing chemicals. Who uncover new applications for never-before-used chemicals.

In the process, we've developed a series of economical polymer catalysts used in manufacturing urethane foams.

A better antioxidant/stabilizer for middle distillate fuels.

And a special polymer catalyst for overprint varnishes and other specialty coatings.

Helping solve your problems is what our product development is all about.

North Chicago, Ill. 60064
Montréal, P. Q., Canada
Zurich, Switzerland
Cronulla, Australia
Buenos Aires, Argentina

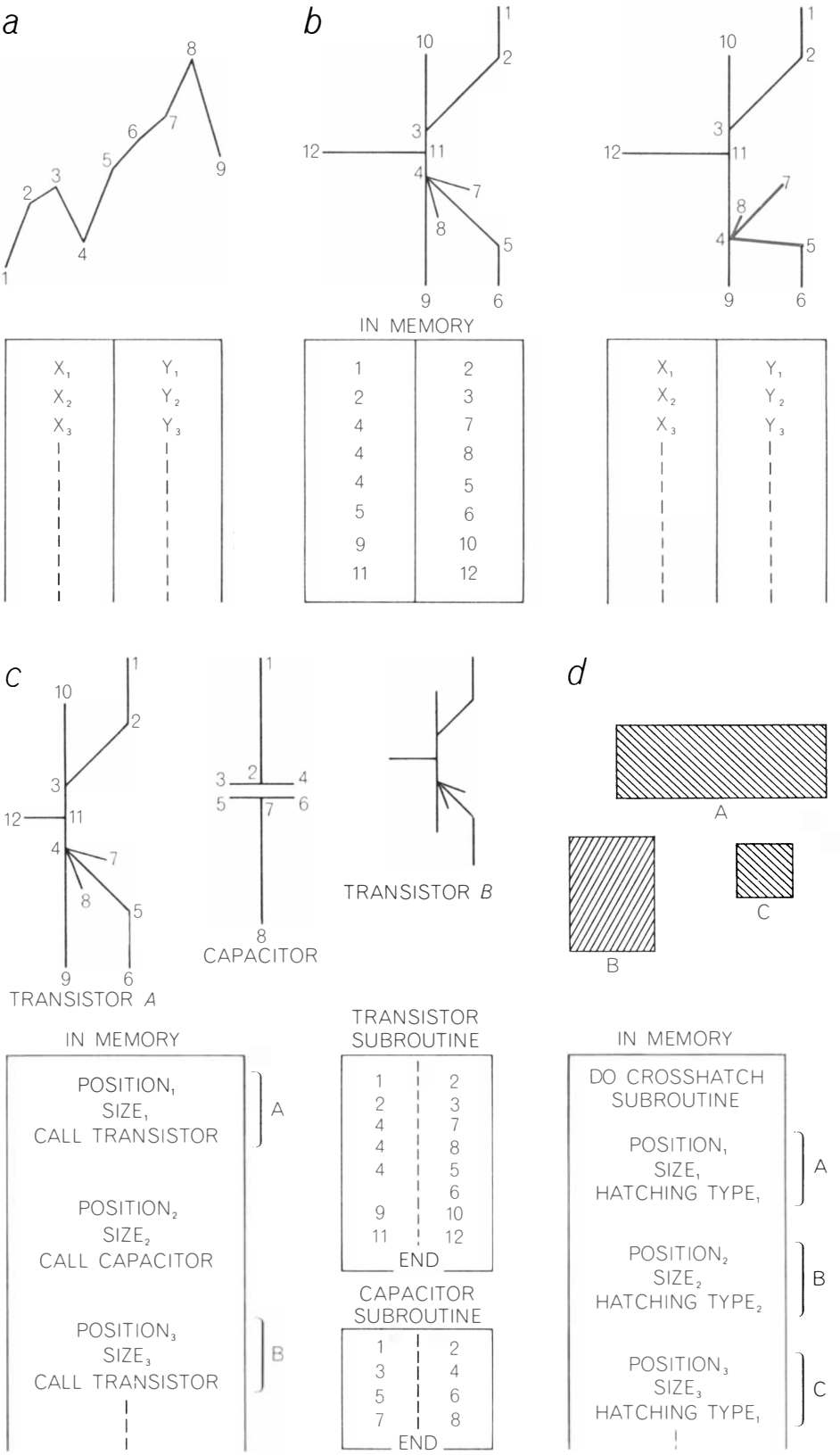
Abbott Chemical Products Division

cessor region have to be considered in dealing with its successor regions. Thus the number of surfaces to be considered in each subregion becomes quite small as the area of the subregion decreases. If the situation in a subregion of the smallest possible size is too complex, the algorithm simply picks a shading rule from the nearest adjacent surface, because such a choice affects only a single point on the display screen.

Many varieties of the Warnock algorithm are possible, depending on how complex a situation the algorithm can handle without having to subdivide, and also on the method of subdivision. Warnock's own programs always divided regions into four equal parts and yielded an output only when the entire square was the same shade. The subdivision of such an algorithm is shown in the illustration at top left on page 81. A smarter display program, which avoids subdivision when there is only a single edge in view, will produce the subdivision pattern shown in the adjacent illustration on the same page. This program, of course, requires slightly more time for the computation of each square. Thus there is a trade-off between the complexity of the display program and the number of subareas that need to be considered. We are currently investigating this trade-off statistically in an attempt to discover what kind of algorithm may be the best.

The work of Gary Watkins and Gordon Romney of the University of Utah has developed some interesting statistics about the pictures we are drawing. The most significant and surprising fact is that the sorting task in the hidden-surface problem is not complicated in depth but is quite complicated laterally. The pictures shown in this article, although complicated, are not very deep. That is, if you penetrate the picture with a ray at any particular position, you will hit fewer than 10 surfaces in all but specially contrived cases, and the average number of surfaces struck is more nearly three. We have concluded that the problem in producing hidden-surface pictures of the type shown here is largely the problem of sorting the pictures in the x and y directions to discover those few surfaces for which a depth computation is required.

Once the hidden-surface problem is solved, shading and color are relatively easy to introduce. The hidden-surface computation develops information to tell which surface is visible at each point on the screen. Since each surface in the computer includes a color, as well as other properties, the program can compute precisely what hue, saturation and



REPRESENTATION OF OBJECTS IN DATA BASE takes four principle forms. Line segments can be represented either as strings of connected coordinates (*a*) or as sets of disjoint lines, each connecting two coordinates (*b*). Once such linear displays have been created they can be stored in the memory for recall at any future time. The representation at the upper left in *b* is the symbol for a transistor. If point No. 4 were shifted downward by changing y_4, all the lines terminating at No. 4 would be shifted simultaneously without, however, destroying the topology. Another common representation (*c*) stores symbols of fixed geometry, such as elements for designing circuits. These can be called to the screen, placed in any desired location and scaled to any size. The record in the memory will then list "Position," "Size" and a code for calling the subroutine that draws the device. If the device is a transistor, the subroutine might read like the sequence of numbers in *b*. A capacitor will be called by its own subroutine. Another representation scheme (*d*) defines a class of objects that are similar in some simple but nongeometric way, such as hatched boxes.

A high-speed digital MODEM transceiver -- one of the smallest, lightest, most versatile in the defense communications field -- was delivered to the U.S. Air Force recently by Hughes. The 70-lb. MODEM (modulator-demodulator) replaces a current 250-lb. device. It transmits or receives 1200 or 2400 digital bits per second in a narrow frequency range over normal telephone circuits, and will be used in conjunction with military computers, teletypes, radars, and other sensors.

Switzerland's computerized air defense system -- built in California by Hughes but code-named "Florida" by the Swiss -- has been accepted and put in operation. It consists of several radar stations that provide simultaneous range, bearing, and height data to the high-speed Hughes computer, nerve center of the system. Should a target be identified as an immediate threat, the commander can request all-weather aircraft or surface-to-air missiles to intercept and destroy.

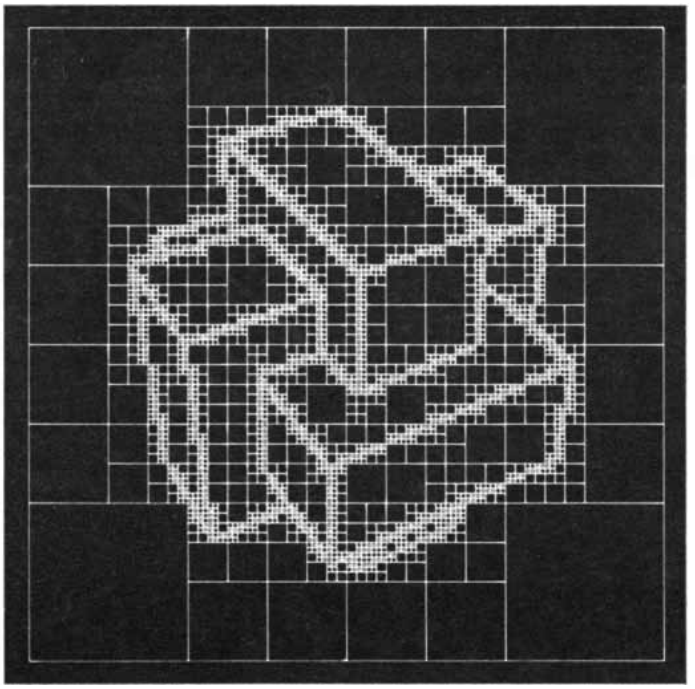
New uses for lasers developed by Hughes include cutting clothing patterns for Genesco, Inc., Nashville garment manufacturer; prototype of the computer-controlled system can cut different patterns singly from a variety of materials much faster than a blade....and trimming thick or thin film resistors with a laser resistor trimmer that can cut nichrome, gold, chromium, and tantalum films as well as standard inks; its single Q-switched laser is split and distributed to as many as four individually controlled trimming stations.

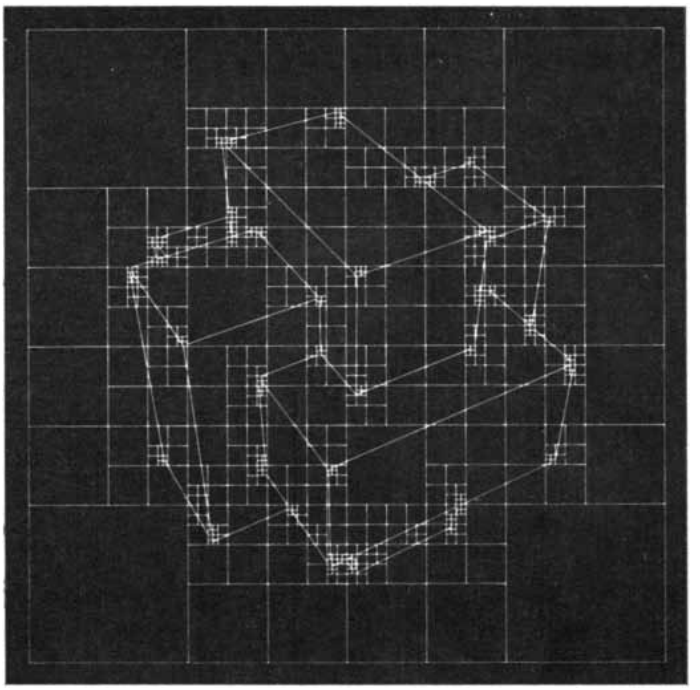
Three custom-designed calibration vans, first of their kind, have been delivered to the U.S. Navy by Hughes. The mobile vans conform to interstate highway requirements and will be driven to the dock to calibrate the Pacific fleet's electronic and ordnance equipment -- eliminating the present expense of removing ships' equipment to land-based laboratories and substantially reducing down-time.

Current programs at Hughes Research Laboratories include: 1) Study of the physics and applications of low-voltage gas discharges. 2) Study of "double stream" plasma instabilities, noise generation in plasmas, and behavior of plasmas in crossed electric and magnetic fields. Physicists or electrical engineers with experience in solid-state microwave are needed. Accredited degree and U.S. citizenship required. Please write: Mr. Al J. Simone, Hughes Research Laboratories, Malibu, California 90265. Hughes is an equal opportunity employer.

Millimeter wave frequencies -- which offer great promise for future space communications and data links -- will be studied in an experiment Hughes is building for NASA's ATS-F satellite. Advantages include: wide bandwidth, plasma penetration, reduced spectrum crowding, reduced size and weight of components. The Hughes-built equipment for the experiment includes the millimeter-wave tube.

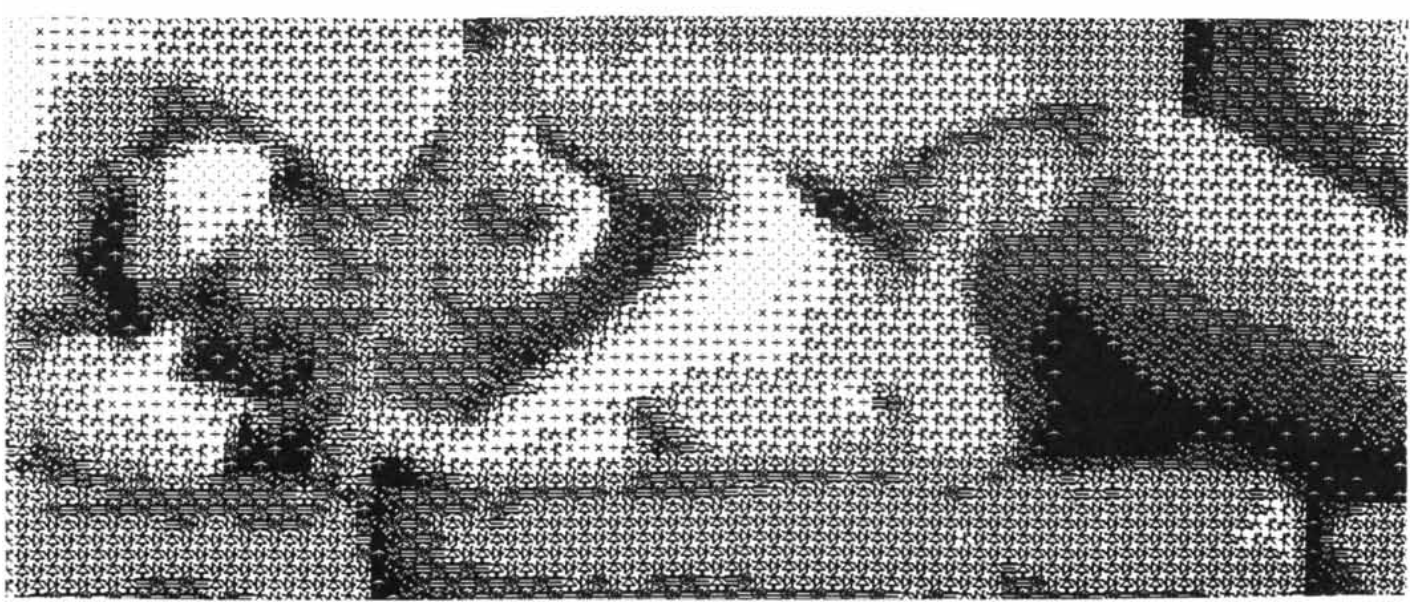
A new, all-electronic mass memory device developed by Hughes promises to make computer information processing faster, simpler, and more efficient. Trademarked DYNABIT, it stores vast quantities of information on magnetic wire thinner than a human hair. It has no moving parts, can withstand severe environmental conditions, and can be adapted to new or existing computer applications.





HIDDEN-SURFACE PROBLEM has been solved by special algorithms that answer the question: Which of several surfaces lies in front of all the others? It is no trick, for example, to have a computer draw transparent structures such as that shown at top left on page 56. The computer plots above represent algorithms for analyzing such a structure and removing hidden lines. The approach at the left, invented by John Warnock of the University of Utah, employs an algorithm that systematically divides the screen into subregions (in this case units of four) and applies a standard procedure to each one. All the surfaces that occur in a given subregion are sorted out and the subregion is divided, if necessary, to end up with only a single surface that lies in front of all others. The method is described more fully in the text. A similar algorithm, which often proves faster, is represented at the right. It avoids subdivision when only a single edge appears in any subregion. Improvements on these algorithms led to the "solid" figures on page 56.

brightness to display. In the simplest pictures the surface is simply displayed in its appropriate color. It is only slightly more difficult to compute the shade, or brightness level, based on a two-part reflectance property: one part for diffuse reflectance and the other for specular reflectance. The photographic output routines can also compensate for the color sensitivity of the film used in photographing the display, so that the resulting prints come out exactly as desired. As a surface of a given color becomes more nearly perpendicular to the observer's line of sight, its specular reflectance dominates its diffuse reflectance and it becomes whiter as well as brighter. This effect can give objects a shiny appearance if much specular reflection is included, or a dull appearance if only the diffuse reflectance is used.



BELL LABORATORIES EXPERIMENT illustrates how a photograph can be analyzed by computer and then redrawn with arbitrary symbols. In this case the Greek letter "alpha" was used as a primary symbol to build up more complex patterns that range from very light to very dense in terms of average density per unit "cell." These cells were then assigned by a computer to match the point-by-point reflectivity of the photograph. Leon D. Harmon and Kenneth C. Knowlton did the experiment as a study in perception.

How Snakes Move

They have four modes of progression, termed lateral, rectilinear, concertina and sidewinding. In lateral progression, the commonest mode, the snake uses its loops to push not downward but sideways

by Carl Gans

"The way of a serpent upon a rock" is one of four phenomena described in the Book of Proverbs as being beyond understanding. A snake undulates in a smooth and asymmetrical fashion, so that the size and curvature of the loops formed by its body change as the animal glides along. Because the parts of the snake's body that exert motive forces are never as clearly defined as those in species with limbs it is not obvious to a casual observer how such movement propels the animal forward. It would seem that the snake's undulations would be just as likely to move it backward or leave it thrashing in one place. Such considerations make the movement of snakes a challenging subject for investigation.

Slithering and related forms of motion employed by snakes are the consequence of the snakes' evolution. Snakes presumably evolved from animals that lost their appendages when they took up a burrowing existence. Since they lacked limbs, the entire body had to become an organ of locomotion. This function favored elongation of the trunk and necessitated a repackaging of the viscera. Both of these characteristics were further developed by most snakes when they emerged to radiate into a broad spectrum of terrestrial, arboreal and even aquatic niches. As the trunk lengthened it remained flexible because the number of bony segments in the vertebral column increased. Thus snakes have between 100 and 400 trunk vertebrae, a number that contrasts strikingly with the 16 or so found in a typical mammal. Ophidian trunk vertebrae are also joined by at least one extra set of articulations, and each vertebra bears a pair of ribs that reach down to form a flexible yet strong support for the sides of the elongated visceral cavity.

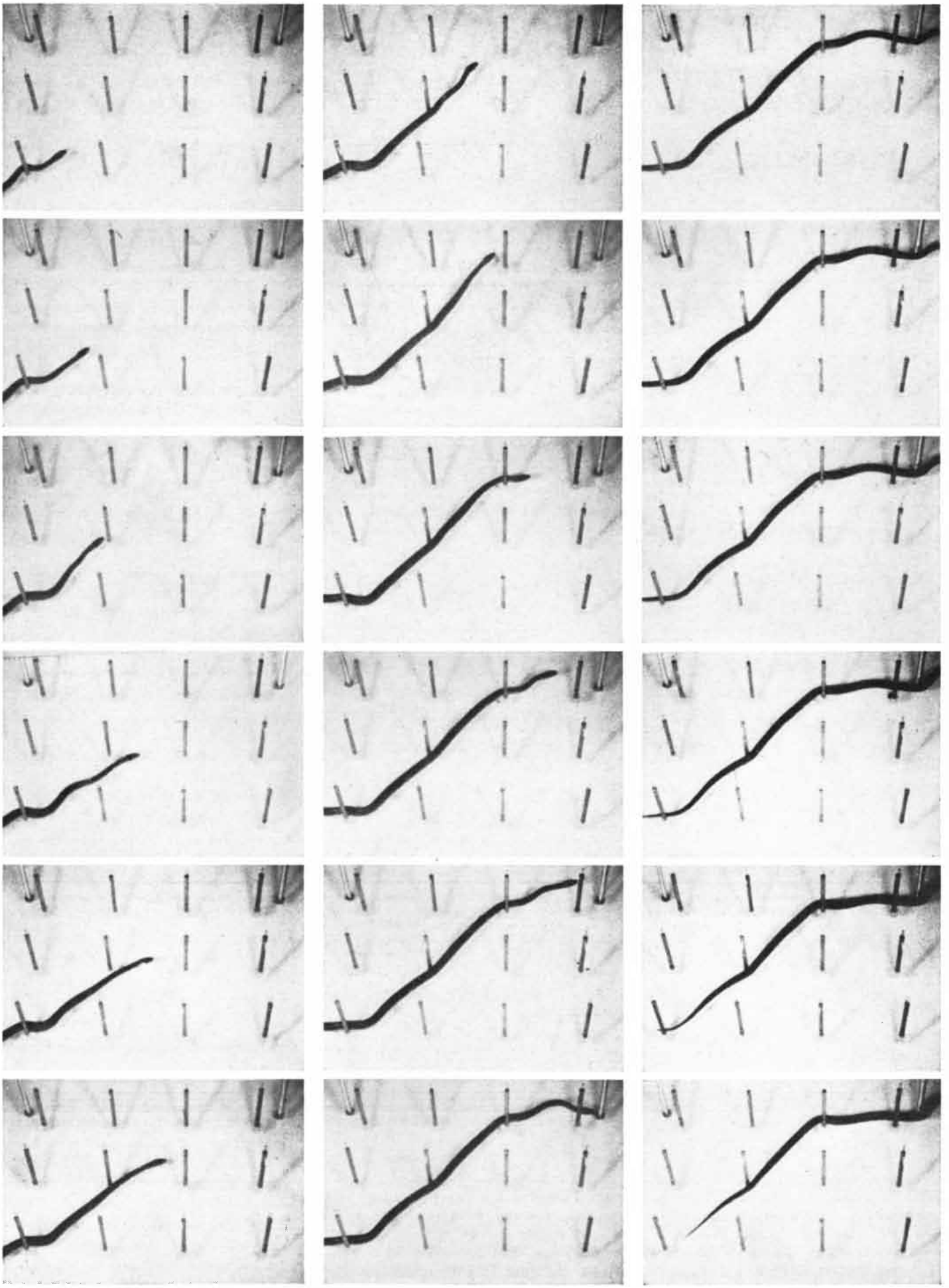
The increase in the number of the segments gives the snake its flexibility; it also complicates the snake's control problems. Some 30 years ago the late Walter Mosauer of the University of California at Los Angeles showed that a snake's vertebrae are joined in such a way that one vertebra can bend up and down in relation to another through approximately 28 degrees and swing from side to side through about 50 degrees. Although the bulk of tissue probably reduces a living snake's flexibility, the number of degrees through which a snake can bend is still remarkable. Even twisting movements can be approximated by a combination of up-and-down and side-to-side bending, in spite of the fact that the joints between any pair of vertebrae do not allow one segment actually to twist in relation to the other. The posture assumed by a snake at any one moment is established by the muscles that together with the skeleton form the axis of the body. The most important of these muscles are arrayed in 12 or more bundles that connect one vertebra to another and to the ribs on each side of the trunk, and in two to four bands that attach the skin to each side of the snake's main mass.

How does a snake utilize all these elements in order to move? An exact description of slithering requires that the position of each element at any given instant be specified. It is easily seen that investigators who seek such a detailed description must solve a truly staggering number of simultaneous equations. The task is complicated by the fact that as the snake travels its parts move and accelerate at different rates. The solution gains several magnitudes of complexity when one wants to know which of the snake's thousands of muscles are actually exerting the forces that induce and maintain motion.

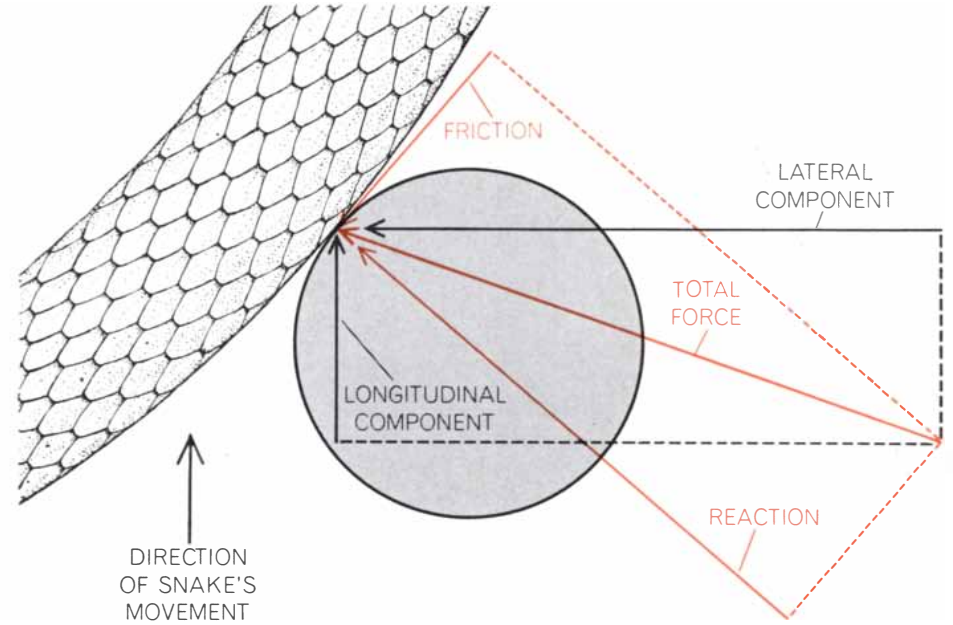
One method of achieving a simpler solution is to work inward from the outside. The snake is treated as a special kind of "black box" capable of exerting forces on areas within an envelope of space considerably greater than its own volume. One can then deal with the snake's external responses without, for the moment, concerning oneself with the way they are generated internally. One first asks where and how these external forces are exerted; then it can be asked what internal forces and mechanisms are responsible for the externally observed efforts. Ultimately this approach allows investigation of the function of the specific muscle groups, and of the motor units and their control.

Since the immediate area of contact between the snake and the environment is the skin it is important to consider the texture of the skin's surface. The sides and back of most snakes are covered with longitudinal rows of lozenge-shaped and more or less rigid scales. Such scales are generally mounted, so that their trailing edges are free, on a flexible and accordion-pleated intermediate layer of skin that is often lined by a layer of muscle. The skin of a snake has a remarkable capacity for extension, as is evidenced when the animal swallows prey more than twice the diameter of its body. The snake's bottom surface is covered by a series of scutes: wide scales whose free trailing edges overlap like plates of armor. Each scute tends to be associated with a single row of scales on the snake's sides and back, and to a single vertebra and a single set of muscles; the scutes thus reflect the snake's internal organization. The side and back scales often have sensory pits of unknown function and may be ridged; the scutes do not have pits and are normally smooth.

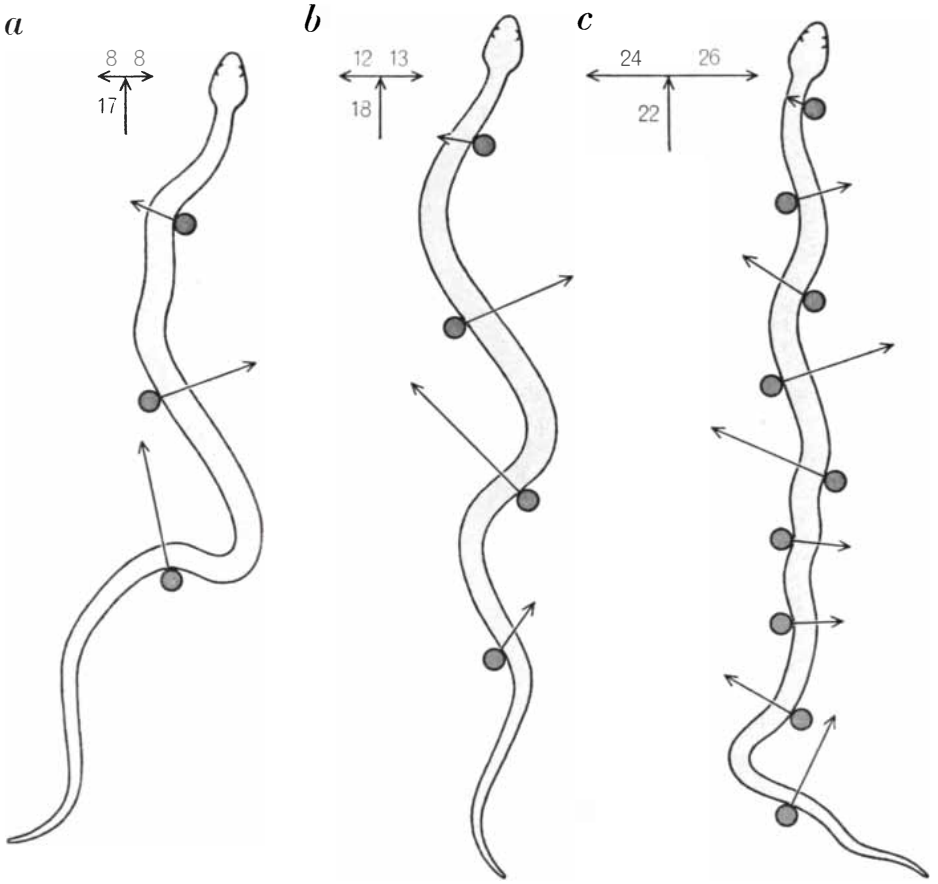
It is generally accepted that the mis-



GARTER SNAKE slithering from left to right across an experimental surface demonstrates lateral undulation. These motion-picture frames show that the snake advances by forming curves with its body that it braces against the brass pegs projecting from surface.



FORCES EXERTED by a snake's body against an object propel the animal. When the snake flexes its body, a reaction force is exerted that pushes the snake away from the peg, as indicated in diagram of forces (*color*). At the same time the snake slides past the peg because of pressure exerted elsewhere. This movement of its body against the peg produces a frictional force. Frictional force together with the reaction force equals the total force. This force consists of a propulsive longitudinal component and a lateral component.



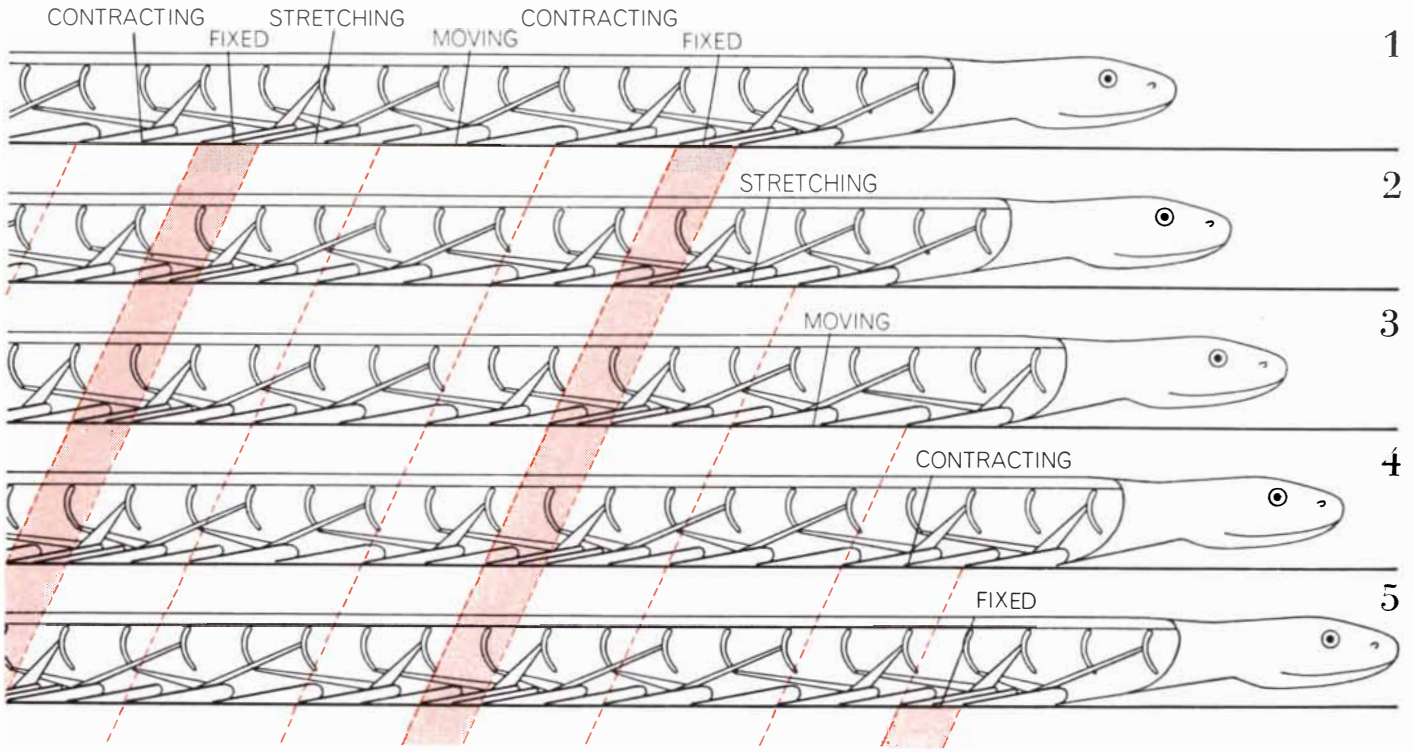
EFFICIENCY of lateral undulation, that is, the ratio of longitudinal forces propelling snake forward and lateral forces expended to the sides, declines as number of contact points increases. In *a* the snake propels itself forward by exerting force on two points at right and using one at left to balance the other two. Vector sum shows that forward-acting force is 17 grams and that 16 grams of force are expended to the sides. In *b* the snake uses four contact points and exerts 18 grams of force in a forward direction but expends about 25 grams of force to the sides. In *c* number of contact points has doubled and sum indicates that more force is expended to each side than is exerted in a forward direction.

cellaneous ways by which snakes propel themselves can be subdivided into four distinct patterns: lateral undulation, concertina progression, rectilinear locomotion and sidewinding. These categories are somewhat artificial. Most species of snakes can move in more than one of these patterns; indeed, a snake moves in more than one pattern at the same time. My colleagues and I at the State University of New York at Buffalo have investigated these patterns of movement. We are currently looking into their control mechanisms.

Lateral undulation is the kind of snake movement most frequently seen. It is the preferred movement not only of snakes but also of limbless lizards. Under certain conditions this kind of movement is even employed by limbed species, but their limbs are first folded against the body. The pattern can be shown to be the most primitive one in vertebrates. Lateral undulation involves the movement of the body along an S-shaped path. In an idealized situation the snake's track can be seen to be a single wavy line scarcely wider than the animal's body; each portion of the snake's trunk has traced this path. The snake begins to undulate laterally by bending the forward part of its body. This movement establishes a wavelike muscular contraction that travels down the snake's trunk.

How does this wave of muscular contraction cause the snake to move? Because human beings walk by exerting force directly on the surface under their feet, it is generally assumed that a snake too propels itself by somehow exerting a force directly on the ground. In actuality undulatory progression operates on a completely different principle. The wavelike contractions flowing down a snake's body do not propel it by acting on the ground under the snake's belly. Instead they cause the snake's body to exert force laterally on irregularities in the snake's path: small elevations and depressions, pebbles, tufts of grass and so on. The snake in effect pushes itself off from such points the way a man sitting in a chair with casters can push himself away from a desk. Since a snake can push only against a limited number of fixed objects, it sets its body in a corresponding pattern of loops, with the outside of each loop forming a contact point. The speed at which a loop progresses down the snake's body toward its tail equals the forward speed of the snake; the moving loops thus seem to be stationary with respect to the ground.

This analysis of lateral undulation has



RECTILINEAR LOCOMOTION enables a snake to advance in a straight line as it stalks prey or crosses a flat surface. In *1* the snake pulls together two sets of scutes, its large abdominal scales, and fixes them against the ground. By pulling against these fixed zones that have established frictional contact with the ground the snake propels itself forward. For this task it tenses the bands of muscle running from the scutes back toward the ribs. In *2* the fixed zones have moved toward the snake's tail (*colored bars*). In order to move the fixed zones the advancing snake picks up the forward scutes in each zone, using muscles that run from scutes forward to ribs, and pulls them clear of the ground (*stretching areas*). Simultaneously it pulls additional scutes forward from behind the fixed zones (*contracting areas*). In *3* snake has begun to contract scutes near its head. In *4* these scutes are being gathered into a new fixed zone, while the hindmost fixed zone has disappeared. In *5* two zones are again fixed at more forward sites and cycle is complete.

been confirmed by Hans W. Lissmann and Sir James Gray in a series of experiments they conducted to characterize the forces exerted against the ground by certain snakes. Their snakes crawled across a smooth board over which hung a regular array of heavy, pendulum-like pegs. Any lateral force exerted against a peg would cause it to swing, indicating both the magnitude of the applied force and its direction. Calculations of these forces showed them to be consistent with the force and speed of a snake's movement. When the pegs were taken away, the snakes were totally unable to move by lateral undulation; they slipped about on the smooth board because they had no objects to exert lateral force against.

Although a snake propels itself by exerting force laterally, it does apply other forces to the ground. One of them is a vertical force induced by the weight of the animal. Another is the horizontal force due to friction. Only the frictional force is exerted by muscle activity; it is the result of the movement produced when a snake curves its trunk in order to apply propelling forces against surrounding objects.

What effect do these forces have on a snake's movement? Friction is produced by lateral undulation in two ways: by the underside of the body sliding along the ground and by the flanks of the body sliding along objects to the side. The magnitude of such sliding friction equals the force pressing the two surfaces together multiplied by a coefficient determined by their nature. In the first instance this amounts to the weight of the snake times the coefficient of friction between the snake's belly skin and the ground. In the second instance it is a function of the force exerted against the object to the side times the coefficient between the flank skin and the object.

In both instances the frictional forces act in opposition to the force inducing motion; thus in lateral undulation friction has to be overcome by muscular effort. This being so, there is a selective evolutionary pressure to reduce the frictional coefficient of the snake's skin, and various kinds of snake scales have been shown to have the low frictional coefficient of about .35. That friction is unnecessary to the locomotion of snakes can be demonstrated by placing a snake in a virtually frictionless experimental environment: a smooth surface lubricated with powder and studded with vertical pegs that are fixed in position but able to rotate. The snake will traverse such a surface even more rapidly than it can cross one studded with nonrotating pegs.

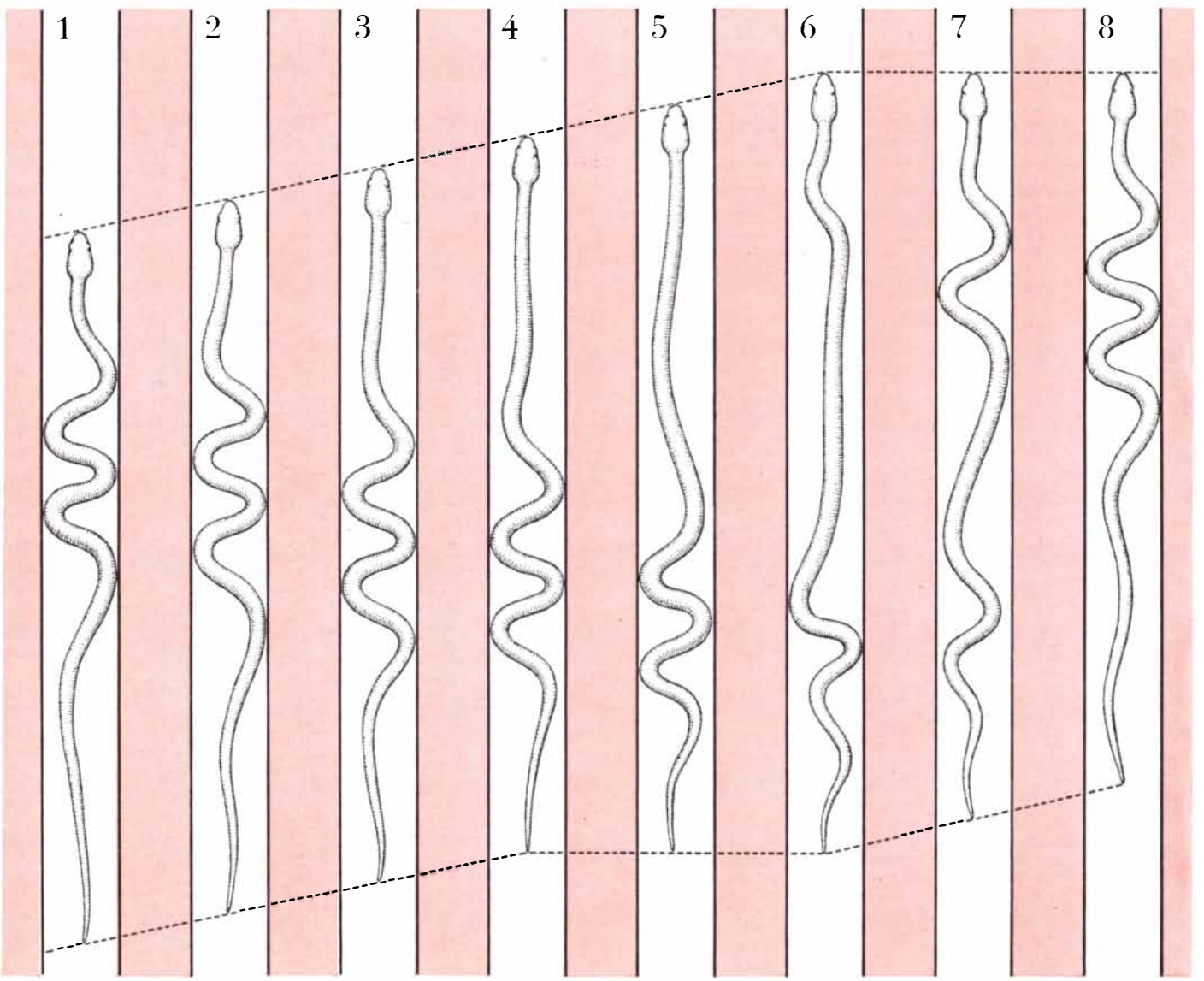
Continuous forward progression by a snake requires a minimum of three contact sites at all times. Moreover, the contact sites must have a particular spatial arrangement. The snake pushes against two of the sites to generate force and uses the third to balance the forces produced at the other two so that its body can move in a particular direction. If the snake is to move, the contact points need to be on both sides of its body and the forces on the contact points must be exerted at an acute angle to the rear. Such considerations suggest that as the snake moves from one contact point to another it shifts its position slightly, and this is confirmed in films we have made. The transition from one set of three contact points to the next must be made quickly; otherwise the snake must utilize four or more contact points at a time and will lose efficiency.

The minimum number of three contact points represents a compromise between two opposing factors. One factor

is the snake's flexibility. Since a snake is notably flexible it must expend a certain amount of energy to make its body rigid in order to transmit force from one contact point to another. The amount of energy applied for this purpose can be reduced if the number of contact points is increased; less energy is needed to maintain rigidity over several short distances than is needed over a few long ones. Yet the greater the number of contact points, the shorter the amplitude of the snake's curves, the wider the angle of force application and the greater the waste of force in the snake's direction of motion. The snake balances the requirement for internal force transmission against the requirement for efficient force application by establishing a number of contact points that is consistent with both. The number of points is also affected by the "desired" speed, the relative elongation (diameter divided by length) of the body and the spacing of objects to the side.

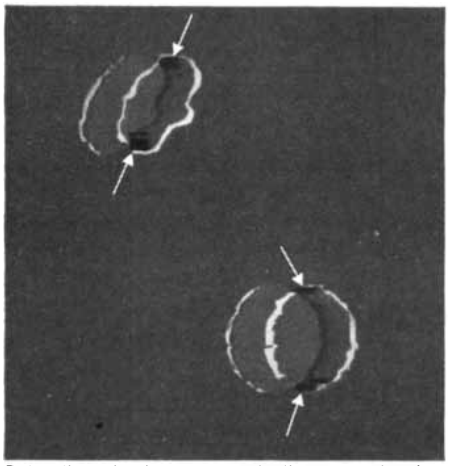
Experiments in which snakes of varying sizes are chased through a fixed array of pegs suggest that the speed of a snake is not proportional to its size. In fact, the limitations on the speed of snakes remain obscure. There may well be an optimum peg-spacing for a particular size of snake, where optimum means the facilitation of a maximum velocity rather than a minimum energy expenditure per traverse. No one has yet done experiments with different peg-spacings to see if such spacings would actually affect the snakes' velocities. Nor have there been studies relating speed to work output or speed to the time or distance a snake is capable of traveling.

Such questions about peg-spacings and efficient movement lead one to ask further questions about control mechanisms. As we have seen, the body of a snake can twist around its long axis to some extent. Thus a snake is capable of lifting a loop over one contact point and making use of another, an ability that enables the snake to select advantageous contact points in its environment. What environmental factors affect such loop placement? Experiments indicate that the key factor may well be the relative resistance of contact points. A snake moving by lateral undulation through grass or sand will increase the size of its loops in order to displace loose surface material until it can exert the maximum force possible under the circumstances. Yet how does the snake monitor information about such resistance, and what is the feedback sequence that shifts the information from one segment of the snake's body to the next as the snake



CONCERTINA MOVEMENT enables a snake to move in a narrow channel. In *1* the snake begins to move by bracing *S*-shaped loops against sides of channel. Snake then begins to extend its head and forebody (*2*) while pushing against passage walls until forward extension is complete (*6*). In order to bring tail section up snake forms a new *S*-curve behind its head (*7*) and braces against the sides of channel. Snake can now bring tail section along by pulling it forward and forming more loops against the passage walls.

How Western Electric spots dots with TV.



Delay line circuitry creates duplicate, overlapping "ghost" images for pencil mark (upper left) and regular .025" dot (lower right). The points where the images cross can be made to produce black areas (marked by arrows). The ghosts are moved right or left, and the positions of the black areas tested

How would you use automation for drilling holes when every hole has to be in a precise spot, but every spot isn't exactly where you expect it to be?

That became a critical problem for Western Electric when we started using a polyethylene with very superior electrical properties as the base laminate in printed circuit boards. Unfortunately this material showed a tendency to shrink when the circuit pattern was etched in the copper on the board—not enough to affect its electrical properties, but enough to dislocate the dots which indicated where holes were to be drilled for placing components. The shrinkage was unpredictable but could move a dot by as much as 1/32"—more than its own diameter.

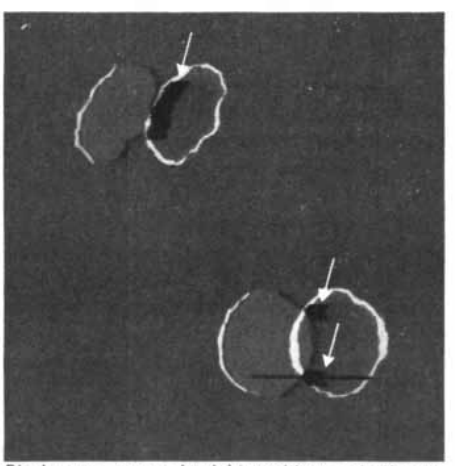
What that meant was that we could not use a conventional tape-controlled drilling machine, which put the drill in precisely the same spot every time. We needed a machine that could, in effect, "see" a dot, no matter where it was, tell it apart from an accidental marking of about the same size and shape, and put a drill right through it.

Engineers at our Greensboro, N.C., Works made such a machine. It consists of a TV camera hooked up to circuitry which removes the grays from the picture, turning it into a series of true on-or-off digital pulses; and logic circuitry which can respond to such pulses and activate a mechanism which moves the board around.

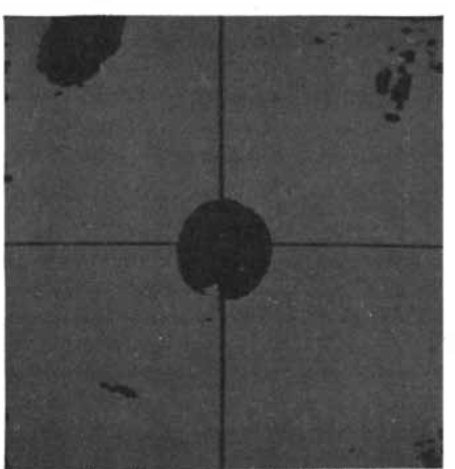
Once the dot has been roughly centered beneath the camera, delay line circuitry creates a second image which is first superimposed on the original, then moved by successive stages to the right. The points where the two images meet—i.e., where the two circles cross—will be in certain positions if and only if the original dot is a perfect circle of the proper size. Logic circuitry tests these positions, and if they are not exactly right, it won't drill and the circuitry sets itself to search for another dot. If they are, it lines the dot up, moves the camera aside, moves the drill into position, and there you are.

A simple, fast, economical, and very, very precise solution to a vexing problem. The kind of solution that helps us keep the quality of the equipment we make for the Bell System high, and the costs low.

Western Electric



Black areas appear in right positions only if marking is a circle of the proper size.



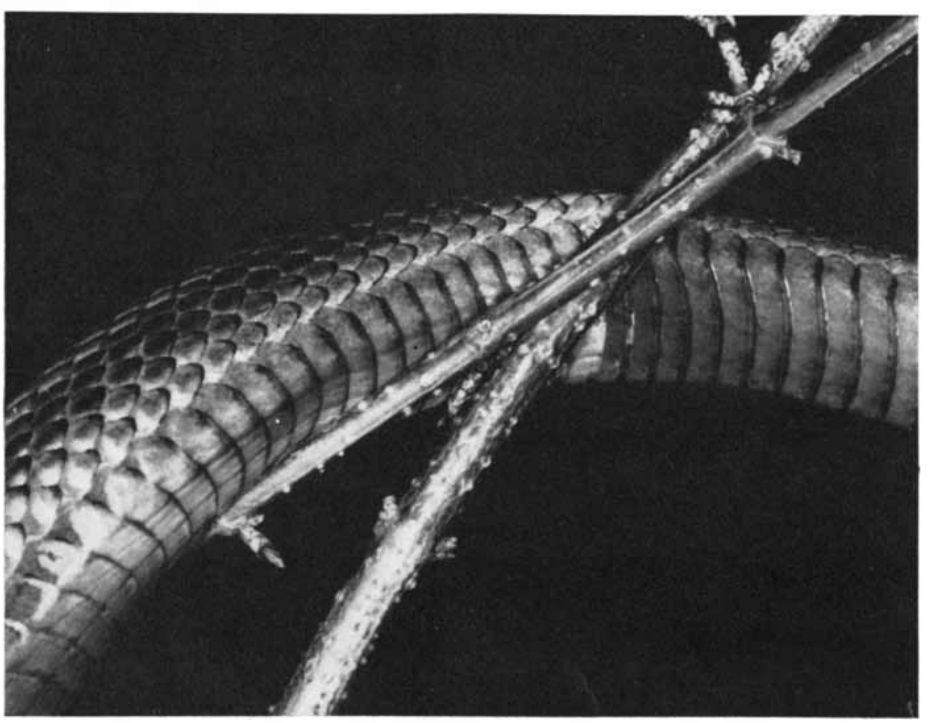
The pencil mark having been rejected, circuitry lines up acceptable dot for drilling.

moves past a particular contact point?

The lateral undulatory progression of a snake has the inherent advantage that the capacity for establishing staggered, backward-traveling waves of muscular contraction is phylogenetically much older than snakes. Snakes have only had to develop a feedback sequence for the control of such waves. This kind of progression is also effective because it permits the snake's body to travel at a relatively constant speed with only limited changes in its velocity. Yet lateral undulation does not work under all circumstances. It is useless for traveling down a tunnel with parallel sides: no backward-directed forces can be exerted against the sides of the tunnel. It is also ineffective on a flat surface that lacks elevations and depressions, and on a rounded surface such as a branch or a fallen tree trunk. Indeed, some snakes find it difficult to use lateral undulation at all. This mode of locomotion is well suited to relatively slender snakes, but it is less effective for short, stout ones. When a snake species has the option of invading an environmental niche that requires a relatively stout body, there is immediate selective pressure for a substitute method of locomotion.

An alternative is rectilinear movement. This mode of locomotion differs from lateral undulation in two respects: it involves the application of force somewhat downward instead of laterally, and it is effective only if friction is established between the snake's skin and the ground. Rectilinear locomotion is made possible by the very loose skin found in many snakes (perhaps as a result of such a specialization as prey constriction). In order to move in this mode the snake fixes several series of scutes and starts to move the skin between them. For example, it pulls together a series of scutes near its head, fixes them against the ground and then moves the rest of its body with respect to this fixed zone and several similar zones behind it. As the body of the snake moves forward the skin is stretched, pulling the forwardmost scutes of each series out of contact with the ground, while additional scutes are continuously pulled up to the rear edge of the series. In this way a constant length of belly surface remains fixed to the ground. Normally a snake fixes two or three of these zones to the ground at once, and they can be seen to move continuously to the rear as the snake progresses [*see illustration on page 85*].

In rectilinear locomotion the two sides of the snake must move symmetrically rather than in alternation. X-ray motion pictures have confirmed an earlier observation that in this mode of movement the ribs and the vertebrae do not move with respect to one another. The scutes are shifted forward by slender muscles that stretch forward from the sides of each scute to points high on the side of each rib. A second set of muscles runs to the rear at a shallower angle to attach the skin to the bottom ends of the ribs. Contraction of the muscles in the first group pulls the skin forward and up, out of sliding contact with the ground; contraction of the stouter and mechanically better-placed muscles of the second group accelerates the snake's mass and maintains its constant forward movement.



SCUTES on the belly of this snake, a Japanese species called the *awo daishu* (blue general), correspond to rows of back and side scales. This pattern is found in all species of snakes.

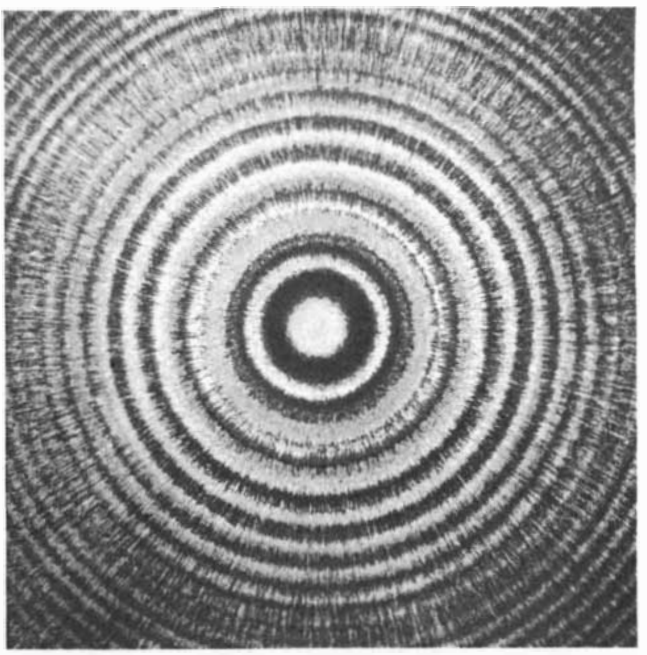
The fact that rectilinear locomotion relies on friction places a certain limitation on this form of movement. The particular form of friction involved is called static friction. Static friction is defined as the force that must be applied parallel to the contact surface between two objects in order to start them sliding past each other. Then static friction becomes sliding friction. It follows that a snake can only exert a force parallel to the ground that is less than the force of static friction; otherwise the animal will slip. A snake in rectilinear movement must accelerate smoothly, because even a temporary imbalance may shift a group of scutes from static to sliding contact and make the effort ineffective.

Even those snakes that have achieved the necessary body form to move by rectilinear progression find this mode of locomotion limiting because of the need to keep some parts in static contact. Rectilinear progression enables the snake to cross flat surfaces and to advance in a straight line when it is stalking prey. Such progression, however, is slow because each scale must establish stationary contact with the ground before it can transmit horizontal forces. It seems that motor control is again of critical importance in keeping the contact zones from sliding.

Most snakes that do not have the muscle and bone structures necessary for rectilinear progression can still use static friction in locomotion by employing concertina progression. In order to move in this way the snake draws itself into an S-shaped curve similar to the posture assumed in lateral undulation, and sets the curved portion of its body in static contact with the ground. Motion begins when the head, the neck and the forward part of the body are extended by forces transmitted to the ground in the zone that remains in stationary contact. These forces produce movement by acting against the force of friction generated by the weight of the snake's body. This reserve of friction is called the static-friction reservoir. It should be noted that the area of contact between the snake's body and the ground does not affect the "capacity" of the reservoir. Friction is a function solely of the force pressing two surfaces together multiplied by a coefficient that represents

"LASER LIGHT"

A SCIENTIFIC AMERICAN FILM



Almost everyone has heard of lasers, but relatively few people have seen them in action. The Editors of SCIENTIFIC AMERICAN now present "LASER LIGHT," a 16-millimeter sound film about lasers: what they are, how they work, the marvelously pure and curiously scintillating light they produce, how they are being used and how they may be used in the near future. The film is in color and lasts 37½ minutes. It is now available for sale or rent.

A few highlights of the film are:

- **Computer-generated animation explaining stimulated emission and resonant optical cavities.**
- **Ripple-tank and oscilloscope demonstrations explaining the wave principles underlying laser action and holography.**
- **Holograms, their three-dimensionality dramatically evoked by the moving camera.**
- **A 600-foot, 8.8-kilowatt laser in action.**
- **Tunable lasers.**
- **A television picture transmitted by laser beam.**
- **The laser chalkline for the San Francisco Bay tunnel.**
- **Laser interferometry.**
- **Gas, solid and organic-liquid lasers.**
- **An experiment on the use of holography in a computer memory.**
- **Original musical score.**

"LASER LIGHT" is recommended for general audiences with an interest in science and technology, and for use in conjunction with the teaching of physics and optics. The film is accompanied by a selection of five SCIENTIFIC AMERICAN articles on lasers and holography, written by leading authorities in these fields.

The sale price per print is $375, the rental price $37.50 for a booking of three days. If the film is purchased after rental, the rental price will be deducted. If rental booking is desired, kindly specify date. Write Motion Picture Department, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017

Coexistence on

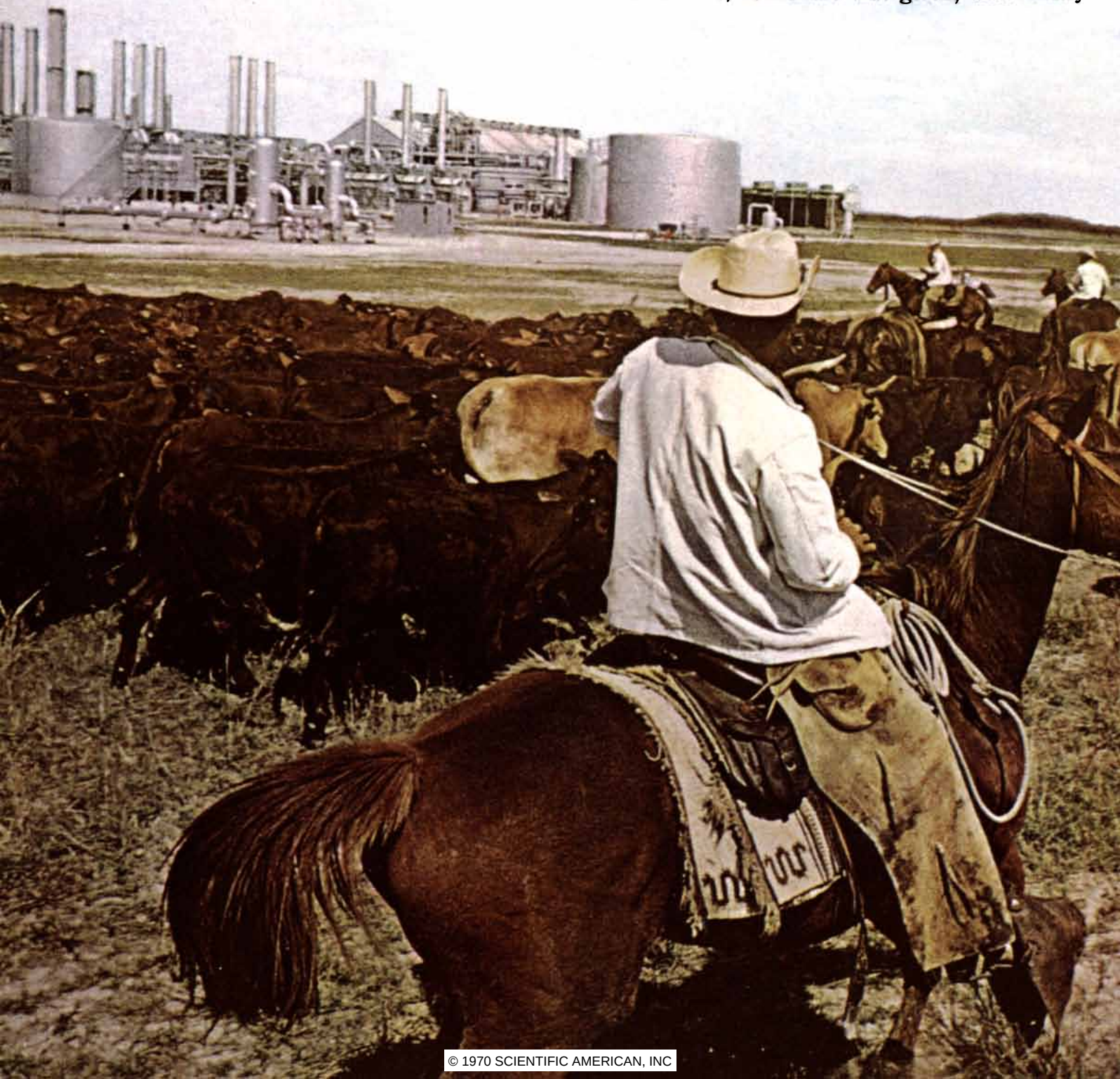
King Ranch is a hunk of pure Americana under the huge Texas sky. It is a lot of things in a large place. And they all get on remarkably well together.

There are thirty thousand cattle and twenty-eight oil and gas fields; two thousand miles of fencing and one thousand miles of buried pipeline; over three hundred windmills and the world's largest natural gas plant. You can see it in our picture.

There are schools and stores and fire stations. Five hundred cowboys. Two thousand cow ponies. A training track for race horses. And five hundred miles of private road. Yet there is more wildlife on the ranch today than ever before in its history.

We, at Jersey, believe this last fact proves something about our affiliate, Humble Oil & Refining Company. Good housekeeping.

They have been drilling for oil and gas on the ranch for forty-five years. If running an oil field were the noisy, messy business some people think it is, how come wild geese, wild turkey



the King Ranch.

and bobwhite quail elect to stay and multiply by the thousand?

If pipelines and gas plants pollute the water, how come deer, nilgai and javelinas guzzle it? And thrive.

One more point about this precious water. The King Ranch gas plant uses thousands of gallons an hour for cooling purposes. When this water is returned to the ponds and creeks, the cattle drink it.

These include the famous Santa Gertrudis, the first breed of cattle developed in the Western Hemisphere. They are now bred in forty-eight foreign countries. Quite an export.

Would the owners of King Ranch expose these rich, red beauties to anything less than good water? Not on your life.

Standard Oil Company (New Jersey)

ESSO

Picture this.

Your four-door family sedan with a body designed by Pininfarina. With the same flair, the same sleek treatment of contours that mark the Ferrari and Lancia.

Result? The kind of car that gathers a curbside crowd wherever you park it. The unrivaled Italian sporty look you've so often stopped to admire yourself.

The interior, the "soul" of the car, is elegantly French. The fully reclining bucket seats, the thick wall-to-wall carpeting, the built-in adjustable headrests and foldaway armrests—each feature, each detail reflecting that centuries-old French love of luxury.

And the bright warmth of sunlight—how the French love that. A solid steel sliding sunroof lets you enjoy it, too!

All these opulent interior components are standard equipment.

Standard equipment, too, is every engineering advance that gives this car its "international muscle". Each of the four wheels has its own independent suspension, its own power disc brake. A newly designed engine gives you as much as 25 miles to the gallon; lets you cruise at 90 mph all day. The entire body is built of .9mm sheet steel—heavier, safer, longer-lasting than you've had before. It comes with Michelin radial tires. And four-speed synchromesh transmission. This is the car that's built to take any road, anywhere in the world.

The car international auto writers voted "Car of the Year".

The Peugeot 504. Italian style, French soul, International muscle. All new, all yours, the complete car. Air conditioning and fully automatic transmission available, if you want them.

If you want a new car, you'll want to test drive the 504.

It can be your "Car of the Year" for years to come.

For more information see your local Peugeot dealer or write Peugeot, Inc., 107-40 Queens Boulevard, Forest Hills, New York 11375.

$3195, p.o.e. East Coast ($3295, p.o.e. West Coast)*

*suggested retail price. Local taxes and other dealer delivery charges, if any, additional.

Italian style, French soul, International muscle.

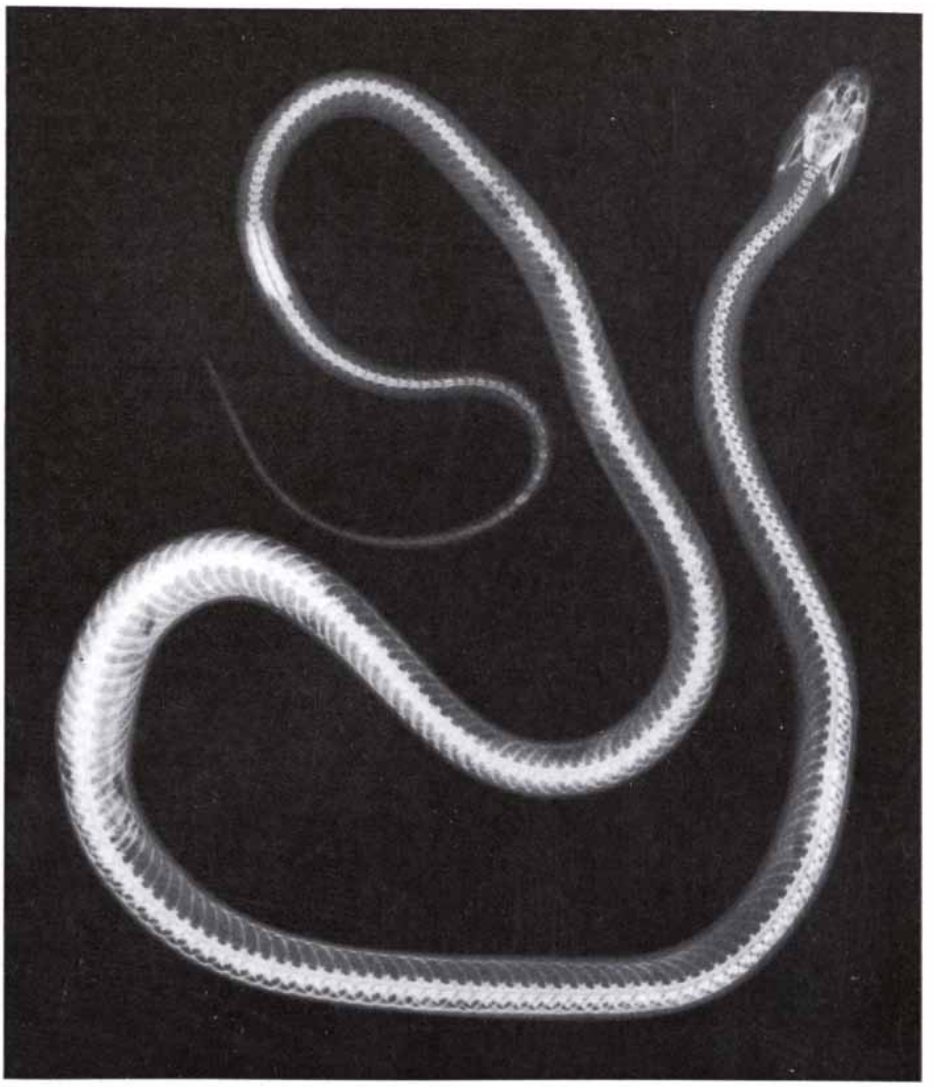
The New Peugeot 504

PEUGEOT SALES & SERVICE NATIONWIDE

their roughness. The force available for acceleration in the static-friction reservoir remains constant as long as the snake continues to support its weight on the stationary zone.

After the front end of such a snake has moved forward a short distance it stops. This establishes a new zone of stationary contact in which horizontal forces are exerted against the ground, and the rear end of the snake's body is pulled forward. The concertina progression of a snake is thus rather like the locomotion of an inchworm [*see illustration on page 86*].

As long as the moving portion of the snake's body can be lifted out of contact with the ground, the ratio of the moving portion to the stationary one depends only on the snake's stability. Yet if the moving parts slide forward or are dragged along they will induce sliding friction that must be overcome. Moreover, since each portion of the snake's weight that induces sliding friction is unavailable for maintaining the static-friction reservoir, the snake will attempt to keep such areas to a minimum. It is therefore possible to predict the approximate fraction of a particular snake's body that can be kept in sliding motion over a particular kind of surface.



SKELETON of *awo daishu* in this X-ray photograph has more than 300 vertebrae and 400 ribs. A large number of vertebrae enables a snake to bend and twist through many degrees.

Snakes sometimes supplement weight-induced friction by muscular force, thereby enlarging the static-friction reservoir. Such modified concertina progression is often used to traverse a straight-sided channel or to climb tree branches. In the first instance the snake widens the amplitude of its S-shaped loops, actively pushing them into contact with the channel's walls. In the second instance the snake forces its stationary surfaces into contact by constricting around a branch. Such enhancement of concertina progression probably arose among species that were already adapted for it through the accident of having an overdeveloped axial musculature for the constriction of prey.

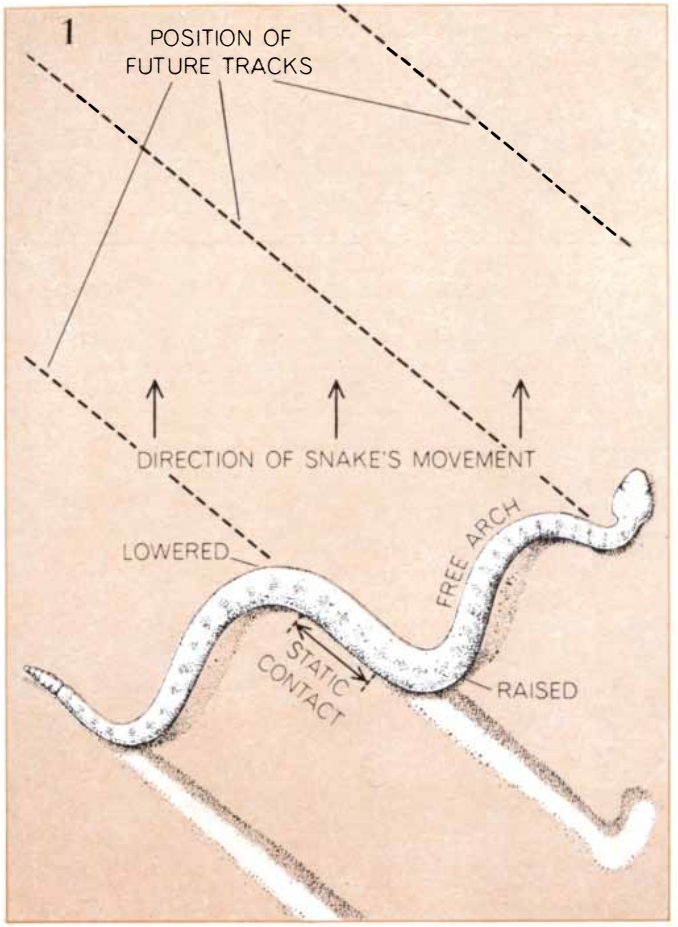
This observation pays an unexpected literary dividend. In "The Adventure of the Speckled Band" Sherlock Holmes solves a murder mystery by showing that the victim has been killed by a Russell's viper that has climbed up a bell rope. What Holmes did not realize was that Russell's viper is not a constrictor. The snake is therefore incapable of concertina movement and could not have climbed the rope. Either the snake reached its victim some other way or the case remains open.

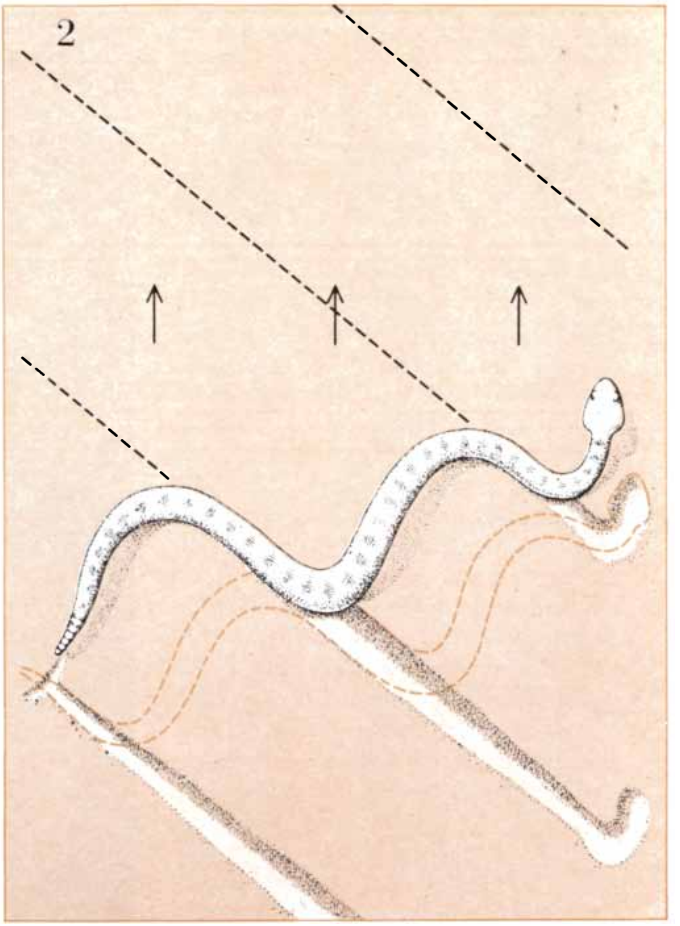
Concertina progression is almost as common as lateral undulation, and one often sees a long snake combining the two kinds of movement. Various species of rat snake use this combination of movements to climb trees; they use the irregularities in the bark as channels. Many such snakes have a double keel on their belly. The structure enables the snake to push laterally against irregularities in the surface it is climbing. Certain burrowing snakes increase their propulsive force by digging in their tail and then straightening their trunk.

Since concertina progression is effected by a bending of the body, it proceeds by asymmetrical contractions. Observation confirms that these intermittent contraction waves move along the body toward the tail. It thus appears that concertina progression is essentially a low-speed pattern of movement, because the animal must pause while it brings its rear end forward so that it can proceed again. Although concertina and rectilinear progression are suited to surfaces where lateral undulation would be ineffective, the price is lack of speed.

For pursuit—or escape—snakes need a mechanism that enables them to employ static friction without sacrificing speed. What is needed is a means of locomotion that keeps the zones that are in static contact from slipping, allows contraction waves to pass continuously and regularly down the snake's body and prevents minor irregularities in the surface from affecting the pattern. Such a form of locomotion is sidewinding, which was first properly described by Walter Mosauer in 1932. A sidewinding snake achieves firm static contact by moving so that its body lies almost at right angles to the direction of its travel. In this orientation the snake is more likely to encounter rocks and other slight irregularities against which it can brace itself than it would if it were moving straight ahead, for the reason that it is sweeping out an area many times wider than the width of its loops.

The track of a sidewinding snake that has traversed a smooth flat surface appears as a series of straight parallel lines,



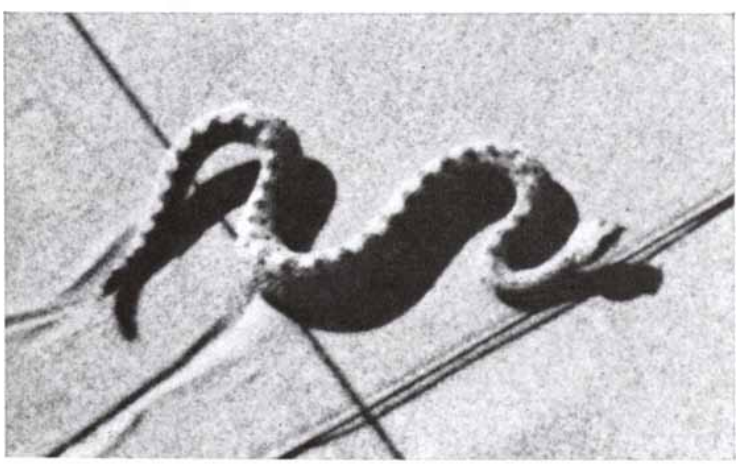


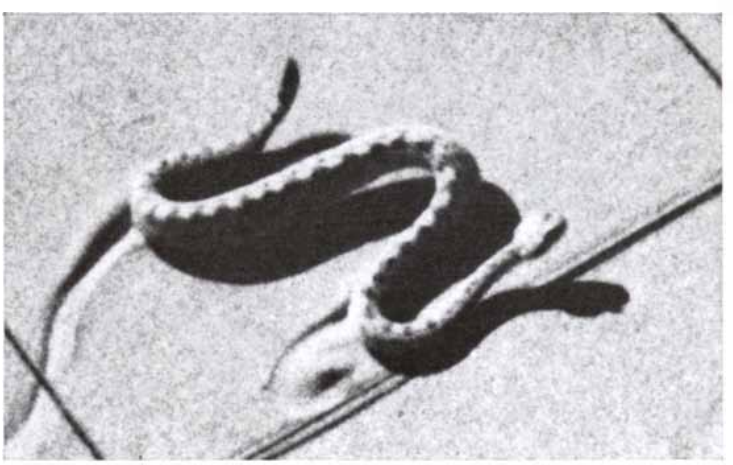
SIDEWINDING enables desert vipers to move rapidly. In *1* a snake with tail in first track (*lower left*) has lifted head and forebody from hooked second track (*lower right*) and arched forward to begin a third track. In *2* head has risen from this new track at

each inclined some 60 degrees to the snake's direction of motion and each about as long as the snake itself. The rear end of each line is bent into a short, forward-pointing hook. The track of the scutes is well defined; slippage, if any, can therefore only be sideways.

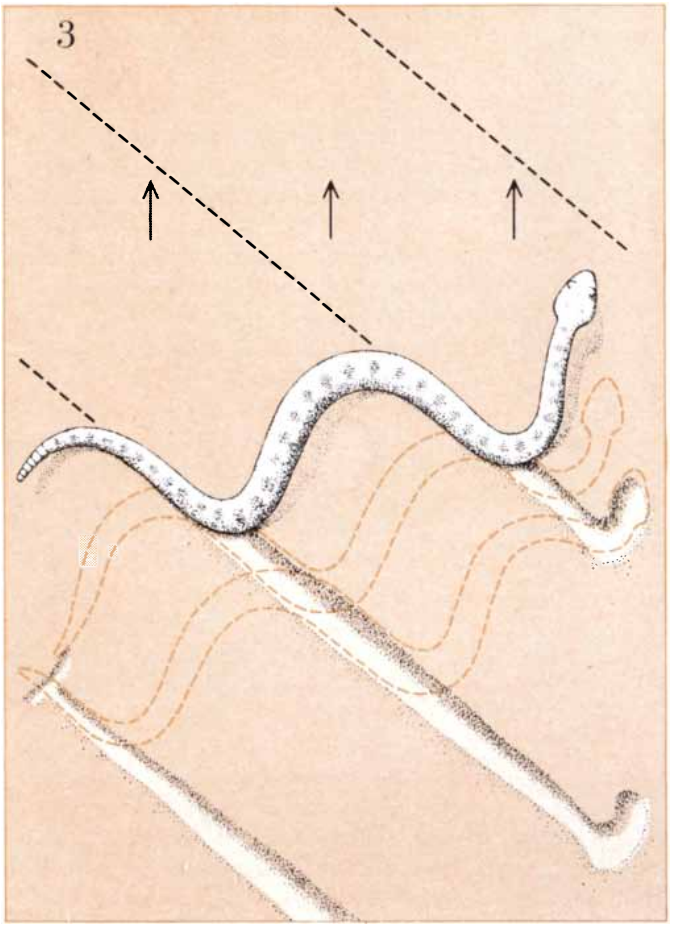
The series of actions producing these tracks can be visualized by assuming that the snake is lying so that its tail points about 60 degrees away from the direction in which the animal is going to move. The snake's head is raised and turned through an obtuse angle, so that it faces in the direction of travel. Only the bend of the snake's neck is in contact with the ground; this produces the hook at the rear end of each track.

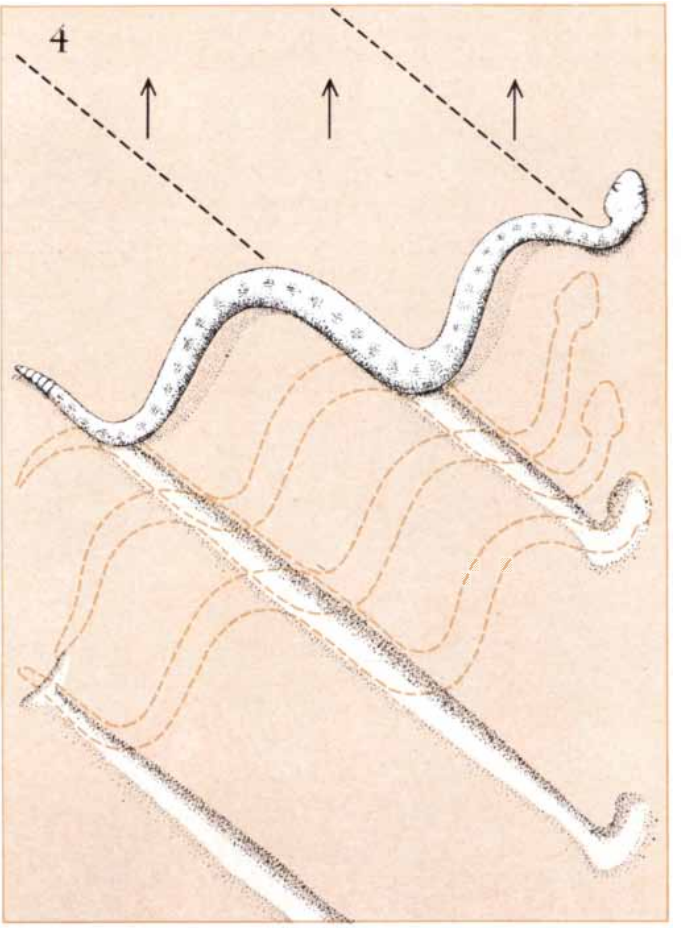
As the snake starts to move forward it lifts its head and neck off the ground. In order to reach the next track the forward part of the body has to curve in a smooth loop. The extension of the snake's front end continues until approximately a quarter of the trunk is cantilevered out of contact with the ground. The head then arches downward as the cantilevered trunk remains off the ground; in making the first contact the neck bends at the next track of the sequence, so that another hook is produced. Successive sections of the body and then the tail follow along the new track, which parallels the preceding one. Considerably before the tail has been pulled into the second track the snake's front end starts





SIDEWINDER (*Crotalus cerastes*) undulates at an angle of about 60 degrees to its direction of travel. This angling helps the snake, a western U.S. species, to avoid slipping. First the sidewinder sweeps out such a large area that it is likely to encounter objects to brace



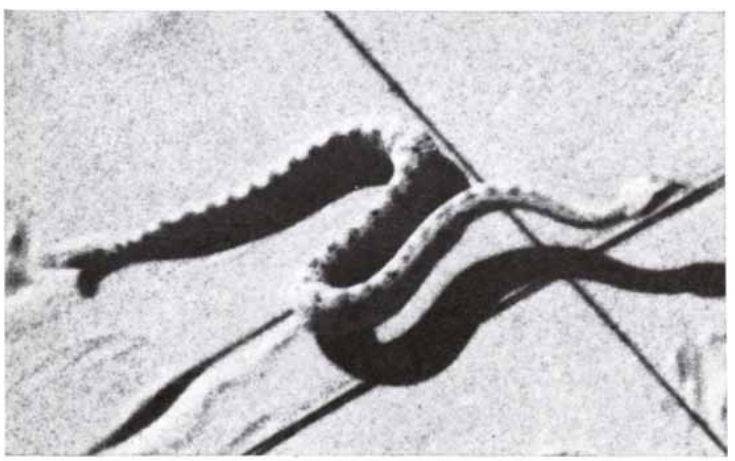


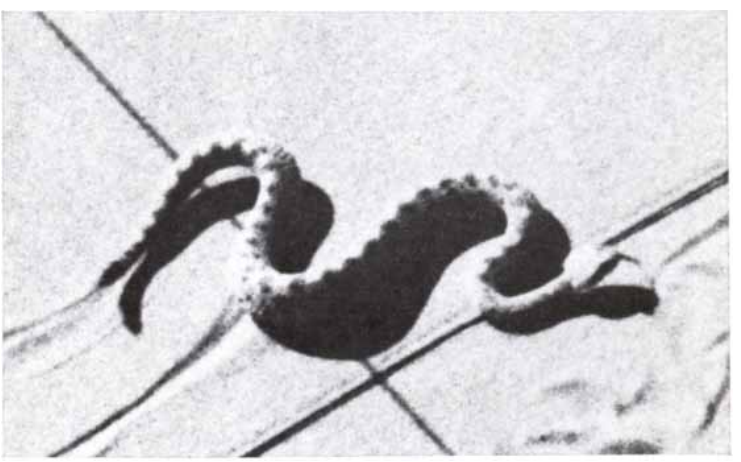
right. Meanwhile the snake's body has lifted clear of the first track and begun to move into the second one (*broken outline of snake's body indicates earlier position*). In *3* snake's body lies in both the second and the third track. In *4* a fourth track has been started.

into the third track. The body of a sidewinding snake thus lies on two or three separate tracks, with the body parts between tracks held off the ground.

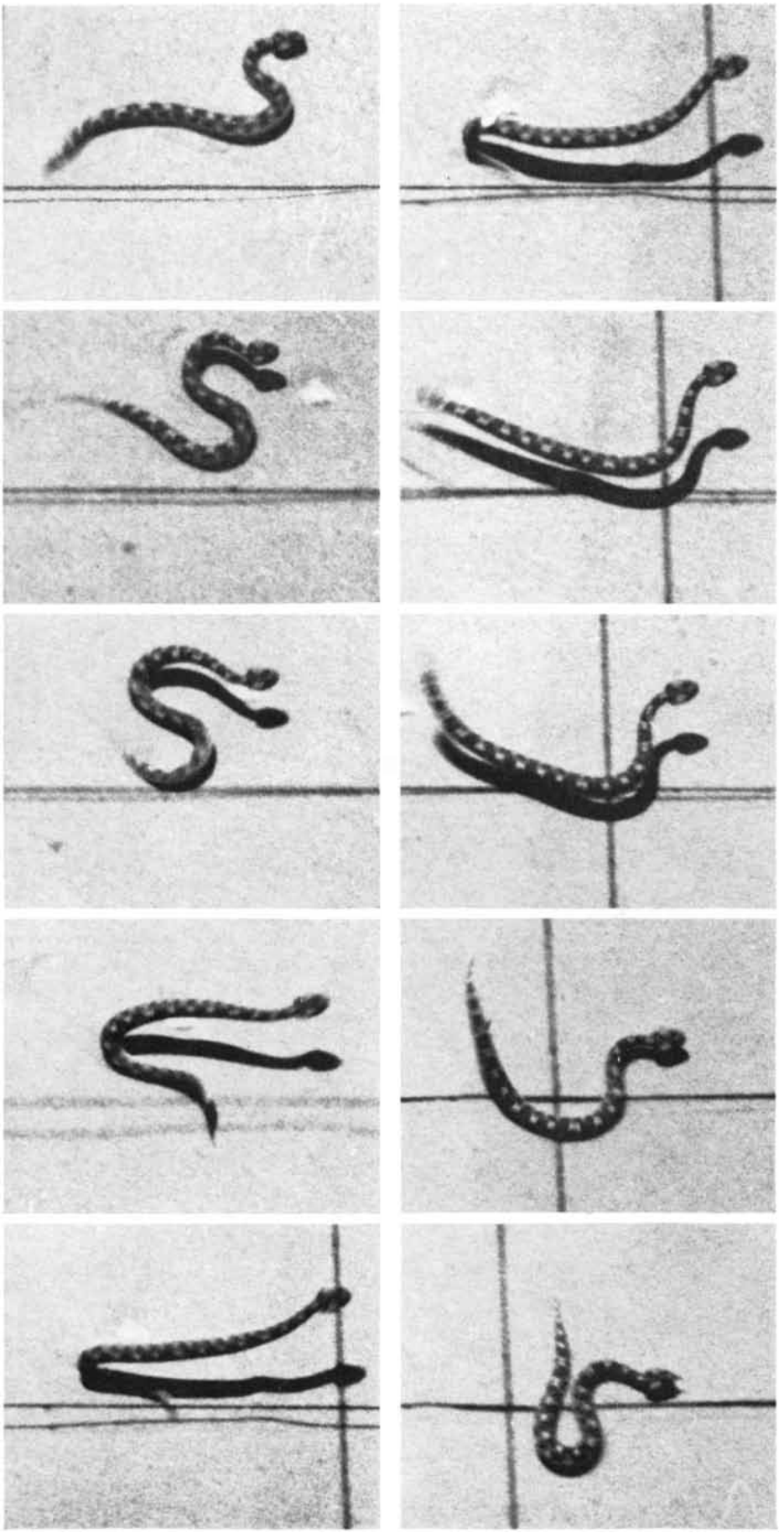
Although the snake's body swings through loops that are reminiscent of lateral undulations, one can see that the force-transmission pattern is more like that of the concertina sequence. The force for the initial acceleration of the front part of the snake, and for maintaining the velocity of the moving parts, must be transmitted by static friction of the snake's belly. It is this pattern that gives sidewinding its peculiar advantages. When a nonsidewinding snake travels over loose sand, its lateral force shifts the sand and its track becomes a trough. Moreover, a snake that exerts a lateral force exceeding the resistance capacity of the force points tends to slip sideways instead of moving forward. Similarly, a snake in rectilinear locomotion will just dig in, slipping within its track, if the force exerted by the contact zones exceeds the limit of the static-friction reservoir. For the sidewinder such slippage proceeds sideways and causes the snake's body to dig in, piling up sand along its length and increasing the forces that produce forward movement. In other words, a slipping sidewinder quickly stabilizes itself by converting sliding friction into static friction.

H. Mendelssohn of the University of





against. As the snake advances it also pushes sand back with its body until the sand piles up (*ridges in tracks*) and resists further pressure. In this way the sidewinder is able to arrest the backward slippage of its loops and move its entire body forward.



JUMPING SNAKE, the African desert viper *Bitis caudalis*, has developed this variation of the sidewinding pattern in order to escape intense heat. In the first of a series of frames from a motion-picture film this snake begins its leap by sending muscular contractions down its body toward its tail. In fourth frame down the impulses have lifted half of the snake from the ground. In the next two frames all of the snake except the tail is in the air. In the last frame the snake lands on a black line that is part of a measuring grid.

Tel Aviv and I are currently analyzing the feedback mechanisms that control sidewinding. The energy required for sidewinding is apparently reflected in the height to which the traveling loops are cantilevered off the ground. The loop height is lowest for those surfaces that are smoothest and these therefore require the least amount of energy. Indeed, when the snakes are allowed to travel over a smooth, low-friction surface such as a polished terrazzo floor, they will slide rather than lift the moving parts of their body. If their tracks were visible, they would tend to be connected. When the snakes traverse a smooth but high-friction surface such as sandpaper, the tracks are quite separate. A rough surface such as a layer of crushed, sharp-edged aggregate causes a sidewinding snake to lift its loops quite high. Frame-by-frame analysis of films suggests that the loops continuously change their height.

Sidewinding has the further advantage that the contraction waves can start at the neck and pass down the snake's lateral musculature. Each motor sequence is thus continuous even though parts of the body stop and start. The snake is accordingly able to cross relatively flat areas, liberated from the need to search for lateral irregularities to propel itself.

Sidewinding seems to be a method of locomotion available to any kind of snake. Various conditions elicit it, and no special structural modification is required. Yet although many snakes sidewind under stress, only a few species do it effectively, and the beautiful control and minimal energy utilization described here were observed only in desert vipers. Such snakes seem to have found this method advantageous for traveling quickly over flat, sandy and rocky surfaces. Sidewinding would seem to be particularly useful when the crossing has to be made at midday, when speed and minimum contact may prevent overheating.

The various sidewinders differ in their track angles and in the number of tracks they occupy at one time. One of the most spectacular of the movement patterns that have been observed so far is seen in juvenile specimens of the Southwest African desert viper, a species that has evolved a special escape sequence when it is faced with heat stress. The muscle-contraction waves travel down the snake's body at very high speed, and some specimens jump completely off the ground from track to track [*see illustration at left*].

If Avis rents you a Plymouth with over 19,000 miles on the odometer, you can keep it.

This pledge is good any time through June 30, 1970. At any Avis office. In any of the 48 contiguous states.

But we don't expect to have to fulfill it, because we try harder to rent you the newest Plymouths in the business.

If we should slip, and you rent a Plymouth with over 19,000 miles on it, it's yours.

Lock, stock, and barrel.

Avis.
We try harder.

Freeze-dried frogs' legs, an old Inca favorite.

Hundreds of years before Columbus, the Incas of Peru were freeze-drying much of their food. They'd lug meat, fish, frogs' legs, potatoes and such up into the Andes and expose them to the bitter night cold. The food would freeze, partially thaw when the sun came up, then freeze again at night. Before long all the food's moisture sublimated, leaving a dry, nearly imperishable material that sprang back to life when boiled in water.

Today, there are hundreds of freeze-dried foods, snack foods, soup mixes, salad mixes, processed meats and dairy items on

The Dow Chemical Company, Packaging Department, Midland, Michigan 48640.

grocers' shelves. And chances are a fair portion of these will be protected by some form of package or packaging material from Dow.

It could be flexible film. Or a laminate of film and foil. It could be sheet. Or extruded foam. Or printed cups with lids.

It could be most anything, in fact, of a plastic nature. And the same materials could just as easily be holding six-penny nails, hand cream or quail eggs.

If you're involved with packaging—whether of food or any other product—think of Dow on your next assignment. We'd like to lend a hand.

DOW

NEOGLACIATION

Mountain glaciers all over the world have advanced several times since the end of the last major "ice age." Their fluctuations are a generalized record of global climatic changes over 6,000 years

by George H. Denton and Stephen C. Porter

Glaciologists and geologists have known for some time that glaciers respond to small variations in climate by growing or shrinking in size. The response of a glacier is of course not instantaneous, and many factors combine to determine the time lag between a change of climate and a detectable change in the position of the glacier's terminus. The response time for neighboring glaciers of differing size and flow characteristics may be quite varied, particularly with regard to short-term changes of climate. For long-term changes, on the other hand, a dominant regional pattern of glacier activity may become manifest, with most glaciers responding in a similar manner, although perhaps at quite different rates. On the time scale of a major "ice age" glaciation, which may encompass several tens of thousands of years, global cooling leads to a worldwide growth of existing glaciers as well as to the generation of new glaciers, including extensive ice sheets, in previously nonglaciated regions.

Evidence of early glacier fluctuations, found widely throughout middle and high latitudes, records the growth and retreat of large ice-age glaciers during the Pleistocene epoch. The last major glaciation, called the Wisconsin in North America and the Würm or Weischel in Europe, was marked by the growth of several large ice sheets in the Northern Hemisphere and of smaller glaciers in alpine and polar areas. At the time of maximum Wisconsin glaciation, some 20,000 to 14,000 years ago, sea level was about 130 meters lower than it is today, and glacier ice covered about 27 percent of the earth's land surface, compared with about 10 percent today. A widespread glacier retreat, accompanied by a rise in sea level, followed the Wisconsin ice maximum. This period of recession was marked by many glacier fluctuations, some caused by changes of climate and others possibly by factors unrelated to climate.

By about 7,000 years ago the Scandinavian ice sheet had almost completely disintegrated, and by 5,000 years ago the last remnants of the once huge Laurentide ice sheet had melted away in areas peripheral to the Hudson Bay region. Glaciers in alpine areas throughout the world also dwindled in size, and many disappeared entirely. Throughout the world the altitude of the treeline rose. Carbon-14 dates of remnants of ancient trees, as well as interpretation of pollen records from bogs and lakes in many areas, indicate that the time of maximum warmth was reached close to 6,000 years ago. At that time the mean world temperature probably was about two or three degrees Celsius higher than it is now. This period of relative warmth, which followed the Wisconsin glaciation and preceded today's cooler climate, is commonly referred to as the hypsithermal interval.

More than two decades ago François E. Matthes of the U.S. Geological Survey recognized the complexity of recent glacier fluctuations and reviewed the evidence pointing to a post-Wisconsin resurgence of glacier activity in western North America following the mild hypsithermal interval. Up to that time it had been widely assumed that modern alpine glaciers throughout the world were shrunken remnants of former Wisconsin-age glacier systems, and that the fresh glacial sediments bordering them were merely the youngest of a series of such deposits left as the glaciers dwindled in size.

Matthes believed the major trunk glaciers of coastal Alaska and Canada, together with the long valley glaciers on volcanoes of the Cascade Range in the Pacific Northwest, probably persisted through the post-Wisconsin interval of mild climate. He maintained, however, that the same assumption could not be extended to all the numerous small "cirque" glaciers that now exist in the Cascades, the Sierra Nevada and the principal ranges of the Rocky Mountain system. Small ice bodies in the Sierra Nevada, he suggested, were regenerated when warm, dry conditions during the height of the hypsithermal interval gave way to a cooler and moister climate that favored glacier growth. By inference he extended this concept to include the majority of cirque glaciers in other mountain ranges of western North America and designated the period of renewed glacier activity the "little ice age." Matthes' hypothesis of the post-Wiscon-

NEOGLACIAL MORAINES fringing the terminus of the Donjek Glacier in southwestern Yukon Territory in Canada are shown in the vertical aerial photograph on the opposite page. Such ridgelike accumulations of sediment deposited along the margin of a glacier are among the most common glacial landforms encountered in alpine valleys. The outermost moraines of the Donjek Glacier, which are dark in color owing to a cover of vegetation, are about 200 to 300 years old and represent the maximum position attained by the glacier in the past 10,000 years. The numerous small ridges superimposed on the massive embankments are end moraines that have probably resulted from minor advances of the terminus during periodic surges of the glacier. Meltwater streams have cut through moraine system at several points to flow onto the outwash plain at bottom. Photograph was made in 1961 from an altitude of 13,400 feet by Austin S. Post of the U.S. Geological Survey.

sin rebirth and growth of glaciers in western North America has never been seriously questioned.

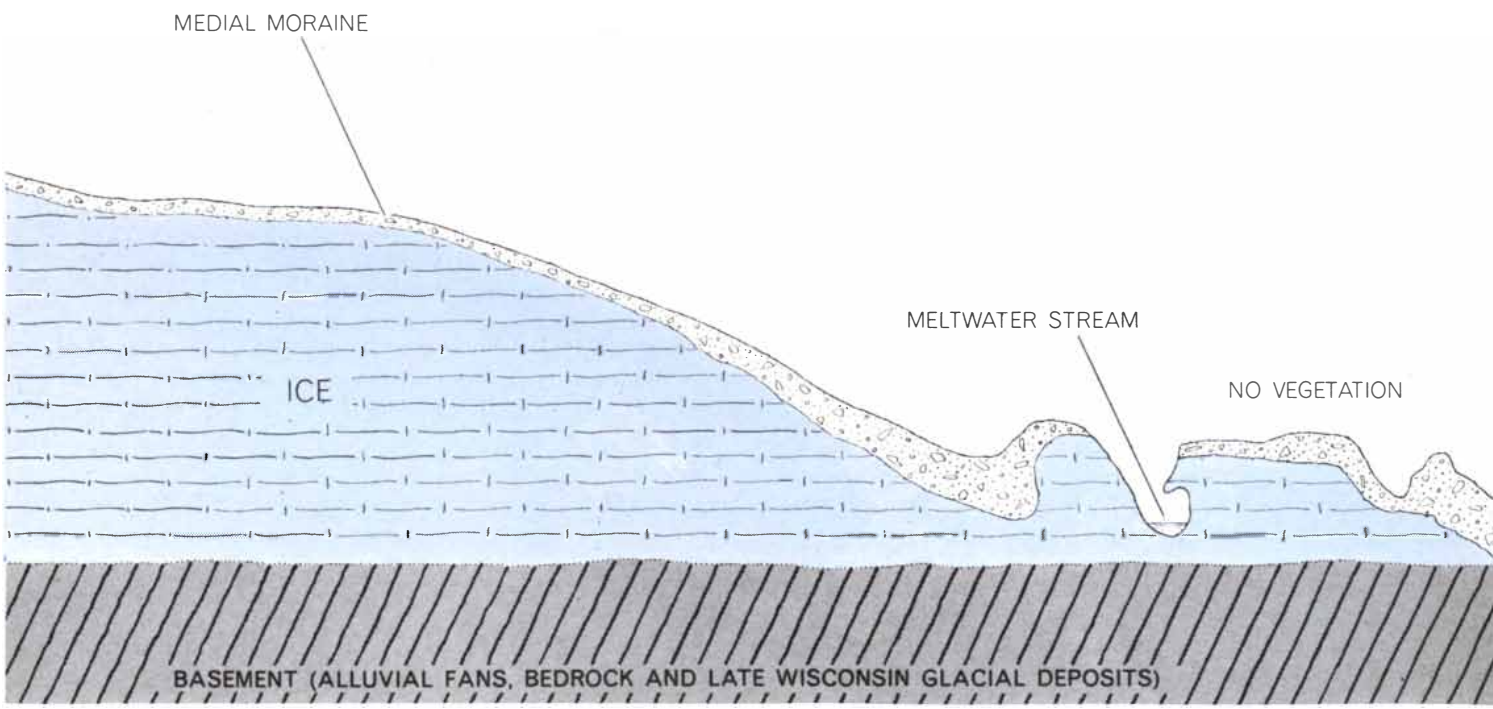
Extensive field research in recent years on post-Wisconsin glacier fluctuations has made it possible to extend Matthes' idea and to reconstruct a reasonably detailed chronology of glacier activity during the past few millenniums. The evidence is varied, and its interpretation has required the combined efforts of specialists in geology, botany, soil science and carbon-14 dating. The post-Wisconsin behavior of the earth's two present-day ice sheets, which mantle Antarctica and Greenland, is still largely unknown. Relatively minor glacier fluctuations, however, have undoubtedly characterized most polar and alpine glaciers throughout post-Wisconsin time. These glacier fluctuations occurred both during and after the hypsithermal interval.

The term neoglaciation," replacing Matthes' "little ice age," is used to encompass the interval of rebirth or renewed growth, and all subsequent fluctuations, of glaciers after the time of maximum hypsithermal glacier shrinkage. Very likely maximum shrinkage coincided with the period of maximum warmth, which from various lines of evidence is thought to have occurred sometime between 8,000 and 5,000 years ago, probably close to 6,000 years ago. The hypsithermal interval originally was conceived of as a period of rather uniformly mild climate and was strictly defined on the basis of dated pollen-zone boundaries from northwest Europe as extending from about 9,000 to 2,500 years ago. It is now known that rather complex low-order changes of climate characterized the hypsithermal interval, resulting in several early neoglacial episodes of glacier expansion. Therefore, in some regions at least, neoglaciation and the hypsithermal interval, as they are currently understood, partly overlap in time.

Historical records of the latest glacier fluctuations during neoglaciation are available from many alpine regions. In the European Alps the oldest records involve the Great Aletsch Glacier in Switzerland, which during the 13th century advanced over part of an aqueduct used to transport meltwater to a local village. More recent records from throughout the Alps chronicle glacier advances of the past few centuries, a period during which many glaciers reached their post-Würm (Wisconsin) maximum. Villages built in what had been considered safe places were overwhelmed by glaciers in the early 17th century. Several of these villages are still ice-covered today. Many distinct advances characterized this period of glacier expansion, major ones culminating in the early 1600's, the 1820's and the 1850's. A general recession, interrupted by a minor readvance between 1912 and 1920, characterized the period from 1850 to about 1960. Since 1960 about a third of the glaciers in Switzerland have reversed the trend of recession and have begun to readvance.

The Rhone Glacier in Switzerland, the source of one of Europe's major rivers, has been studied intensively during the past several centuries. Although it is not in a settled region, the Rhone Glacier has a long history of observation because of its location near two key mountain passes, and because the glacier was a major tourist attraction while it occupied the Gletsch Valley prior to its recent drastic retreat. For more than 250 years, from about 1600 to 1860, the terminus of the Rhone Glacier remained near its neoglacial maximum position, which was reached in 1620. Slightly less extensive advances culminated in 1818 and 1856. A rapid retreat followed, interrupted by a small readvance about 1921, but during the past several years the glacier has once again begun to advance.

An early widespread advance of gla-



NEOGLACIAL HISTORY of the Kaskawulsh Glacier in southwestern Yukon Territory is represented by the cross-sectional diagram on these two pages. During the final phases of the Late Wisconsin glaciation (the last major "ice age" glaciation) the Kaskawulsh Glacier retreated from north to south (*right to left in the diagram*). By about 10,000 years ago the glacier had withdrawn behind the position of its later neoglacial maximum extent. The layer of Kluane, or Late Wisconsin, loess (an airborne sediment) originated during this recession from windblown silt derived from the active outwash plains. During the subsequent warmer period, known as the hypsithermal interval, the Kaskawulsh Glacier receded considerably, the outwash plains became inactive, the dep-

ciers in western Norway, which is well chronicled by tax records, appears to have begun between 1660 and 1700. The first half of the 18th century was marked by a general advance (amounting to as much as several kilometers for some glaciers), which culminated between 1740 and 1750. This advance was associated with climatic cooling that reached a peak in about 1741. As a result crops failed to ripen, famine set in, and the death rate in Norway far exceeded the birthrate. Subsequent regional glacier recession has been interrupted by a number of minor readvances, particularly during the intervals 1807–1812, 1835–1855, 1904–1905 and 1921–1925.

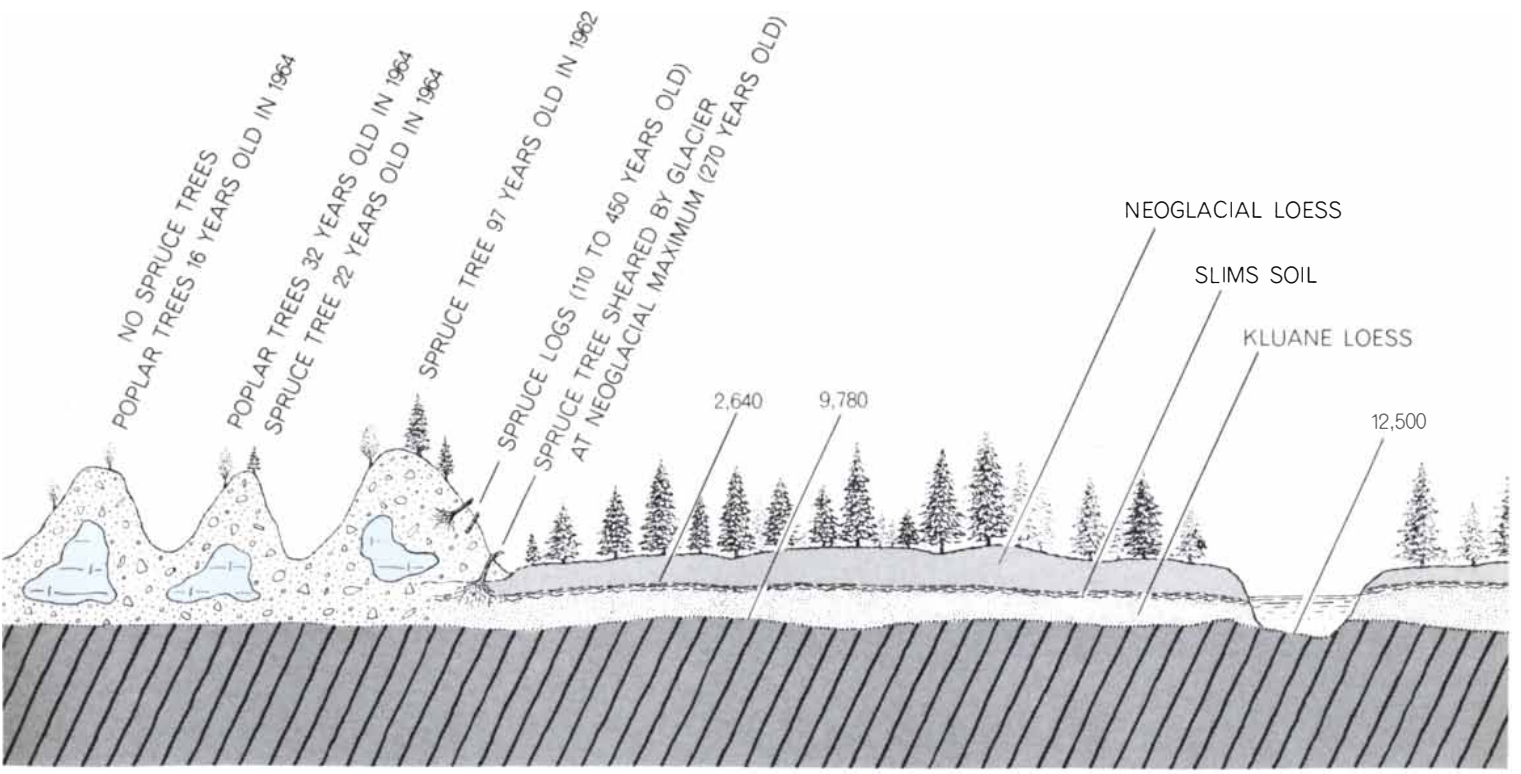
Historical records from Iceland indicate that from the beginning of colonization in A.D. 870 until at least 1200 glaciers were more restricted than they are now. During that time farms were built near glaciers at places that were overwhelmed or nearly overrun by advancing ice early in the 18th century. In Iceland, as in Norway and the European Alps, the past three centuries have been generally characterized by widespread glacier expansion that has included several major and minor advances. Recent glacier behavior has been marked by a general recession, and at present Iceland's glaciers are smaller than at any time since the late 17th century.

Paralleling the records of glacier activity are other historical references to the climatic changes that accompanied glacier fluctuations. These include records of Viking voyages, the amounts of arctic sea ice near Iceland and southern Greenland, changes in habitats of animals and plants, agricultural data from throughout Europe and direct meteorological measurements. Pollen records also indicate that the Roman Iron Age in northern Europe (which lasted from about A.D. 50 to 400 in Scandinavia) was characterized by a mild climate. Between 400 and 1200 the climate was also relatively mild, with the warmest interval apparently occurring between 800 and 1000, when summer temperatures were a degree or two higher than they are now. At that time arctic sea ice virtually disappeared near Iceland and southern Greenland, and Viking voyages of discovery, colonization and trade in the North Atlantic became numerous.

Climatic cooling characterized the subsequent period from 1200 to 1400. Arctic sea ice expanded, clogging sailing routes between Iceland and Greenland. This climatic change, together with social changes in Norway, greatly hampered and finally prevented trade between Greenland and the outside world. The combination of climatic deterioration and severance of trade links, possibly coupled with attacks by Eskimos, led ultimately to the extinction of the Norse settlements in Greenland. The final reference to the colonies was made in 1410 by an Icelander who caught what may have been the last ship from Greenland.

The period from 1400 to 1550 was characterized by a slight warming, which was followed between 1550 and 1880 by further cooling, leading to a vast increase in the amount of sea ice near Iceland and southern Greenland. On at least one occasion the edge of the sea ice is believed to have reached the Faroe Islands, only 250 miles north of Great Britain. It was during this interval that several Eskimos reached Scotland by kayak, presumably traveling along the margin of the arctic ice pack.

The earliest recorded observations of glaciers in North America were made along the southern coast of Alaska. Count de La Pérouse visited Lituya Bay on a voyage of exploration in 1787, and shortly thereafter, in 1794, George Vancouver observed the position of ice in Glacier Bay. At that time glaciers in southern Alaska had begun to retreat from a maximum neoglacial position at-



osition of loess nearly ceased, and the Slims soil formed on the surface of the Kluane loess. An early neoglacial readvance of the glacier led to the deposition of a layer of neoglacial loess, beginning about 2,640 years ago near the present terminus of the glacier. Sometime between 130 and 300 years ago the Kaskawulsh Glacier deposited its outermost neoglacial moraine. The general recession that followed this event has been interrupted by several minor readvances of the glacier, which have resulted in the construction of additional moraines. The dates of the living trees were obtained by making borings in the trees and counting their annual growth rings. The dates of the logs and other organic debris at various depths in the soil were obtained by means of the carbon-14 dating technique.

RECENT DRASTIC RETREAT of the Rhone Glacier in Switzerland is evident from a comparison of these two pictures, made about a century apart. The engraving at top shows the terminus of the glacier as it appeared in 1850, when it occupied much of the Gletsch Valley (*foreground*) and was a major tourist attraction. The photograph at bottom shows the glacier's terminus as it appears today, after having retreated far up the mountainside out of the valley. In the past few years glacier has begun to readvance.

tained earlier in the 18th century. Detailed regional observations of glaciers in southern Alaska, begun in the late 19th century, have documented a general retreat from the neoglacial maximum of the 18th century. For Glacier Bay the recession locally has exceeded 100 kilometers, owing in part to the fact that the glaciers terminate in the sea. Scattered observations of glacier fluctuations in other North American localities provide evidence of general glacier retreat through the first half of the 20th century. Within the past two decades, however, several glaciers in the Cascade Range and in the Canadian Rocky Mountains have readvanced, perhaps in response to a worldwide cooling trend first detected in the 1940's.

The history of glacier fluctuations during the long span of neoglacial time before written records must be inferred largely from glacial deposits. End moraines, which are ridgelike accumulations of sediment deposited along the margin of a glacier, are among the most common glacial land forms encountered in alpine valleys. Well-preserved systems of neoglacial moraines fringe the termini of most modern alpine glaciers. Often associated with the moraines are well-defined forest or lichen "trimlines" recording the former position of the glacier's margin.

In many areas the outermost moraines and trimlines date from the past few centuries, when a large number of glaciers attained their neoglacial maximum. In such areas evidence of possible hypsithermal ice retreats or of earlier neoglacial advances has been obscured and can be recognized only by stratigraphic studies of glacial sediments and interbedded soils. In cases where a small glacier has experienced numerous fluctuations over a relatively short distance, a large moraine composed of a complex of layered deposits may have resulted. Careful sectioning of such moraines has revealed a record of multiple neoglacial advances not recognizable from an examination of the surface morphology alone.

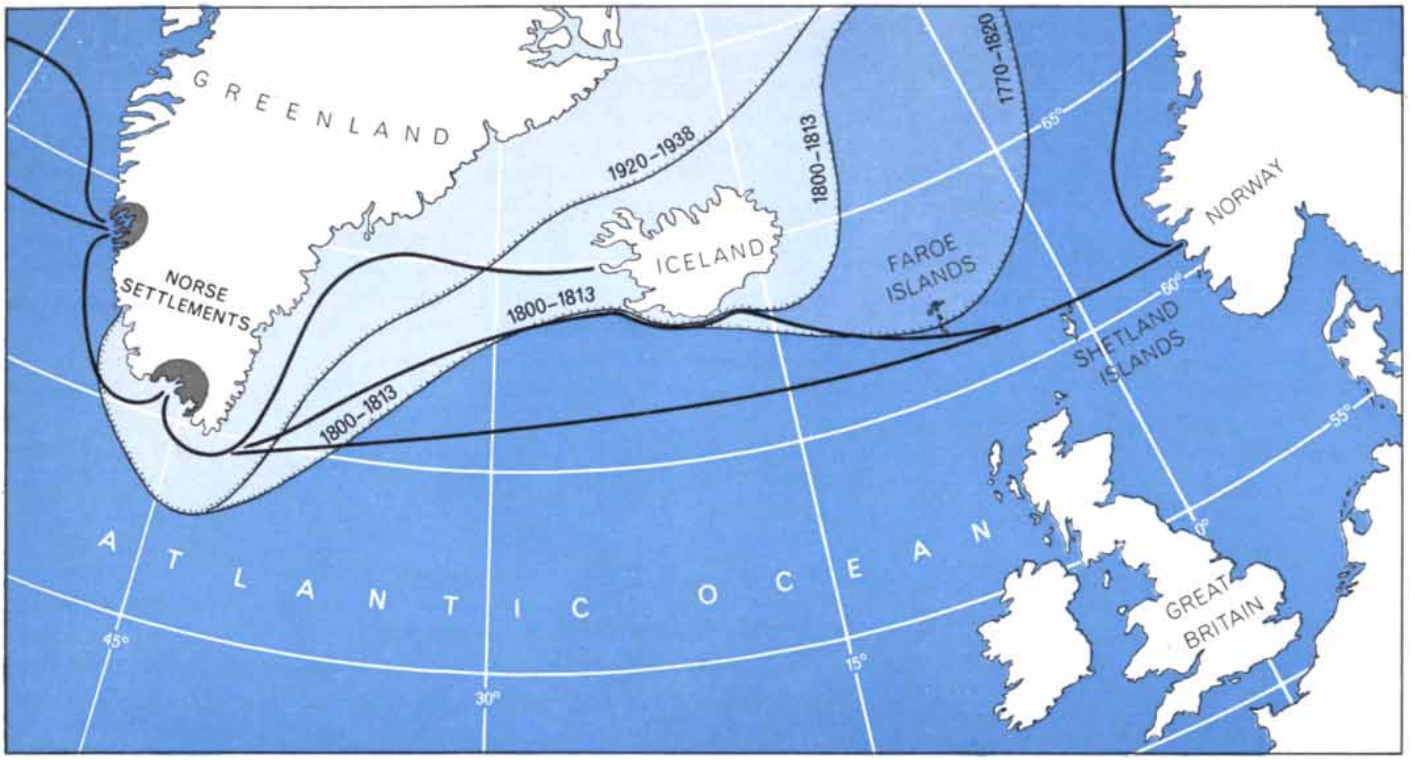
Carbon-14 dates of vegetation, peat or soil overrun by glacier ice provide direct ages for neoglacial advances, and dating of organic matter in recessional or advance deposits associated with neoglacial moraines may provide important limiting ages. Although the carbon-14 method is applicable to all neoglacial time, the resulting ages normally are not true calendar ages and must be corrected before they can be compared directly with ages obtained by other methods. The discrepancy between radiocarbon years and calendar years is a result of known fluctuations in the carbon-14 content of the atmosphere, which possibly are related to solar variation, changes in climate or variations in the intensity of the earth's magnetic field.

Because datable organic material cannot always be found in glacial deposits, several other methods have been employed to obtain the age of neoglacial moraines. In regions where recurring volcanism has produced widespread layers of pumice and ash these deposits have provided a unique tool for correlation and dating. Frequently pumice and ash falls have buried vegetation or soils that can be dated by the carbon-14 technique, thereby providing close limiting ages for the time of eruption. The presence or absence of airborne volcanic deposits of known age on moraines provides a basis for assigning limiting ages to the moraines. This has played an important role in dating neoglacial moraines in Iceland, western Canada, southern Alaska, the Cascade Range and the Sierra Nevada, and is potentially important in several other glaciated regions that have experienced recent volcanic activity.

Throughout much of northwestern North America it has been possible to date recent moraines by counting the rings of living trees. The age of the oldest tree growing between the present glacier terminus and the outermost position of a given advance, frequently marked by an end moraine, provides a minimum estimate of the elapsed time since the ice began to recede. The age of the tree is usually obtained by making a boring near the base of the trunk and counting the number of annual rings exposed on the recovered core. It may be necessary to sample a large number of trees in order to find the oldest, which will provide the closest estimate of the age of the moraine.

The limiting ages are necessarily approximate; an unknown amount of time elapses between the retreat of the glacier and the germination of the first tree that ultimately grows to maturity. Nevertheless, estimates of the age of moraines commonly can be made to the nearest decade with considerable confidence. In rare instances living trees that

DISTINCTIVE FOREST "TRIMLINE" marks the neoglacial maximum of Robson Glacier in the Canadian Rockies. The glacier receded drastically in the first half of this century. The photograph was made in 1953 by W. O. Field of the American Geographical Society.



SOUTHERN LIMIT OF SEA ICE in the North Atlantic Ocean has fluctuated widely since the time of the Viking voyages. Estimations of the Norse sailing routes between Scandinavia, Iceland, Greenland and North America (*broken black lines*) indicate that between the 10th and the 14th century the sea ice was about as restricted as it is today. By the 18th century, and possibly earlier, these former sailing routes were partly blocked by floating ice, which reached as far south as Iceland and the Faroe Islands. During the past century and a half there has been a general retreat of the sea-ice limit toward the north and west. The maximum advance of the sea-ice limit was essentially contemporaneous with the most recent worldwide advance of glaciers during the period of neoglaciation.

were disturbed or tilted by advancing ice can also be found. Examination of their growth rings may indicate the time of tilting, thereby providing an even closer date for an advance. Although this method has found its greatest use in dating glacier advances of the past several centuries, in some areas the method can be extended over at least the past millennium.

Another botanical method that has been successfully used for both relative and absolute dating of neoglacial features involves the measurement of the diameter of the largest lichens growing on a moraine or within a trimline in order to arrive at an estimate of the time elapsed since deglaciation. Lichens normally begin to grow on moraines once their surfaces become stabilized. Following an initial period of rapid growth (commonly several decades) their growth rate becomes essentially linear. For old lichens the diameter is a direct indication of age. Because the rate of growth varies among species and from region to region, independent growth-rate curves must be established for each lichen species in each region. Where reliable growth-rate curves cannot be determined, lichens can still prove useful in correlating moraines within restricted areas. Although not as precise or reliable as other methods of dating neoglacial moraines, lichenometry has proved to be particularly useful in certain nonforested arctic and alpine regions and, under ideal circumstances, is applicable over intervals as great as 4,000 years.

A combination of historical, geological, botanical and carbon-14 data has served to delineate several widespread and long-term episodes of glacier expansion during the period of neoglaciation. Because the neoglacial behavior of the Antarctic and Greenland ice sheets is still largely unknown, most of the pertinent data are from comparatively small mountain glaciers in middle and high latitudes. The magnitude of the advances of noncoastal mountain glaciers ranges from more than 25 kilometers for Kaskawulsh Glacier in northwestern Canada to less than a kilometer for many small Scandinavian and North American glaciers. In areas marginal for glacier support, numerous glaciers were reborn during periods of general expansion. Some advances have occurred in geographically restricted regions; although these may eventually prove to be of general significance, for the present they are considered only local phenomena.

Such advances may have been caused either by local climatic changes or by anomalous rapid ice movements resulting from instabilities in glaciers. Recent studies in Iceland, Spitsbergen, Pakistan and the coastal mountains of northwest Canada and southern Alaska indicate that certain alpine glaciers in these regions are inherently unstable and from time to time advance with exceptional rapidity, an occurrence called a surge. Surges appear to be unrelated to regional climatic trends, but their origin remains obscure. The moraines they produce can easily be confused with those built by a glacier during a climatically induced advance. Thus climatic implications cannot unequivocally be assigned to all local advances until more is learned about the cause of glacier surges.

The earliest indication that there may have been a worldwide post-Wisconsin glacier growth comes from several widely separated geographic areas. In Washington, British Columbia, Baffin Island, Patagonia, Norway and the Swiss Alps carbon-14 dates associated with glacial deposits point to an increase in the size of certain glaciers 5,500 to 4,500 years ago. In addition pollen studies indicate that the climate was both cooler and wetter than it is today in southeastern

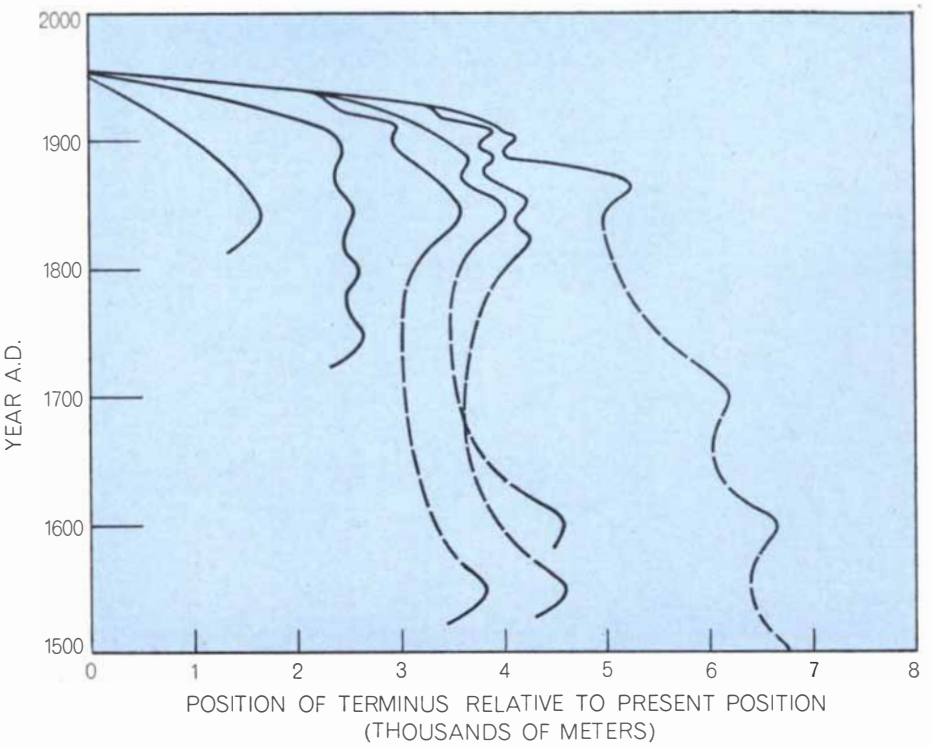
Alaska between 5,500 and 4,500 years ago and in southern Chile between 6,500 and 4,500 years ago, a condition that is compatible with glacier expansion at that time.

Studies of fossil pollen from lake sediments in Keewatin and Manitoba in Canada also point to a possible cooling trend beginning about 5,100 years ago and lasting for several hundred years. Similar data from northern Europe suggest a possible climatic cooling at about the same time, but the evidence is ambiguous in that it can also be interpreted as representing vegetational changes brought about by advancing neolithic agricultural practices. The increasing number of carbon-14 dates that are turning up for this interval suggests that the advances occurred in response to a worldwide change of climate that may have marked the first significant shift to cooler global temperatures following the culmination of the post-Wisconsin warm interval.

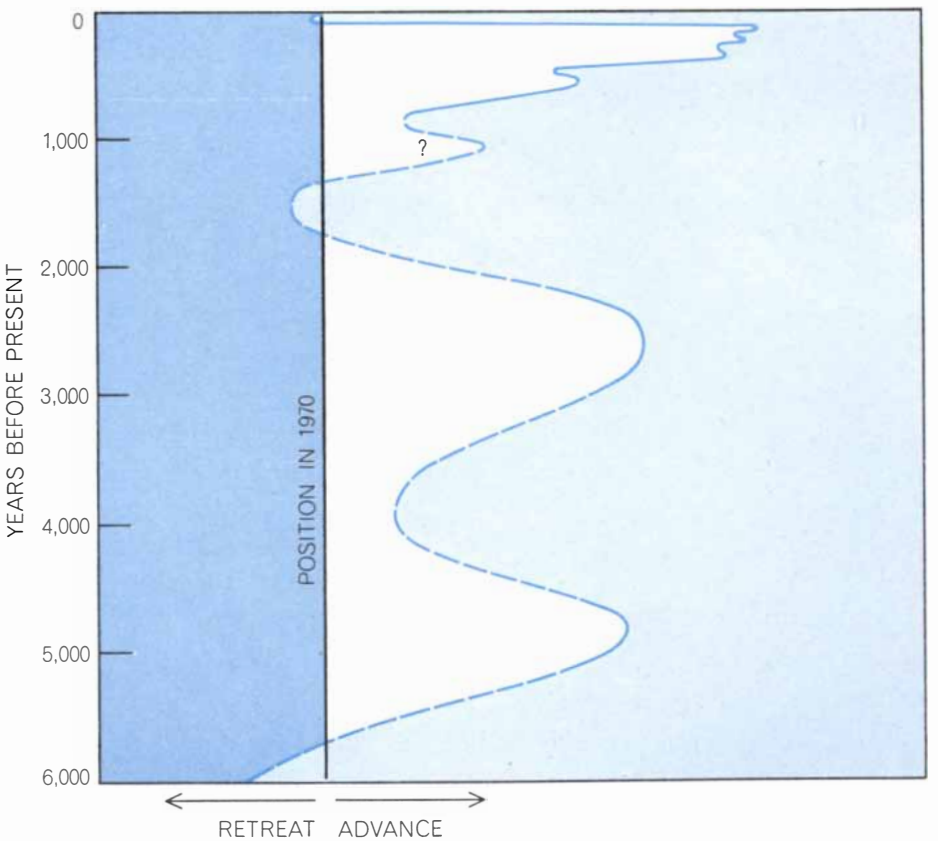
A cluster of carbon-14 ages between about 2,800 and 2,600 years marks another interval of significant glacier growth in many parts of the world. Lichen studies indicate that certain neoglacial moraines in arctic Canada and in the Colorado Rockies were built about 2,400 years ago; comparable studies of moraines fronting glaciers in West Greenland suggest glacier advances in the third or fourth millennium before the present. Several dates from the Alps provide maximum limiting ages for neoglacial advance in central Europe that were essentially contemporaneous with those from other continents.

Dates from the Southern Hemisphere are scarce, but recent studies indicate that an interval of major glacier expansion in Patagonia culminated between 2,300 and 2,000 years ago. In Australia the downslope movement of hillside debris associated with an interval of cool climate began about 3,000 years ago. Many of the available dates within this interval define the time when glaciers were at their maximum neoglacial positions and suggest that a widespread growth of glaciers began sometime before about 2,800 years ago.

Varied evidence from the lands bordering the North Atlantic Ocean suggests that a relatively mild climate characterized most of the long interval between the neoglacial advances of 2,800 to 2,600 years ago and the equally widespread advances of the past few centuries. Historical records from Iceland indicate a mild climate between the

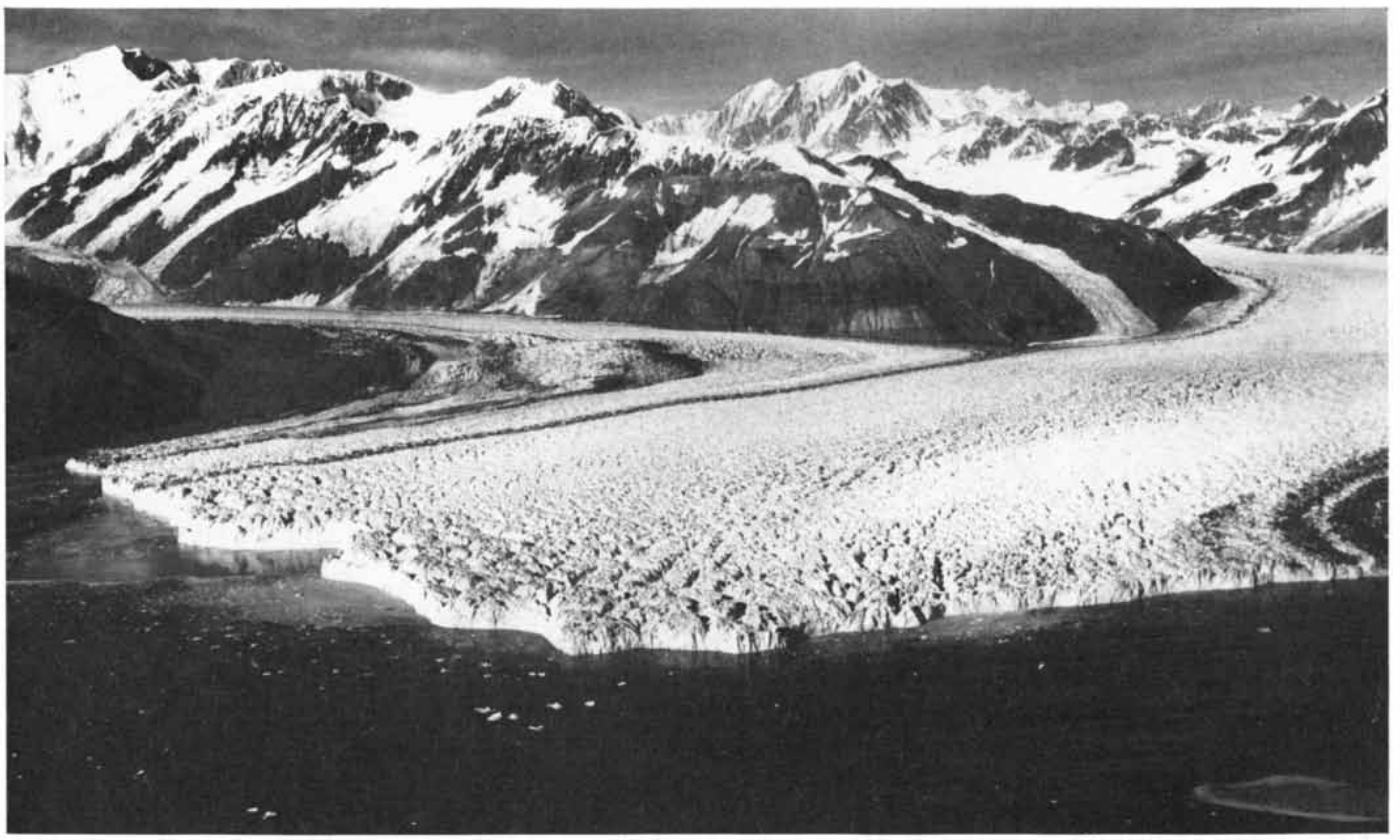


MEASURED FLUCTUATIONS in the terminal positions of six valley glaciers in the Cascade Range of Washington in the past 500 years are indicated in this graph in relation to the present position of the terminus, taken as zero. Each glacier was at or near its maximum neoglacial position during the middle of the 19th century but receded markedly beginning toward the end of the century. The recession of the glaciers paralleled a steady increase in mean global temperature in the same period that amounted to nearly one degree Celsius.



ESTIMATED AVERAGE FLUCTUATIONS of mountain glaciers throughout the world are shown by this curve representing the entire period of neoglaciation. Although most glaciers appear to have reached their greatest neoglacial extent sometime during the past three centuries, the maximum advance for some occurred during the third or fifth millennium before the present. The curve is based primarily on carbon-14 measurements. The vertical straight line in the chart indicates the position of the glacier margins in 1970.

CIRQUE GLACIER drains the south face of Aiguille Verte, a peak in the French Alps. Meltwater seeping into cracks in the rocks may freeze, fracturing and breaking away material that will eventually be deposited at the terminus of the glacier to form a moraine. In this process of erosion the glacier carves out a theater-shaped depression, known as a cirque, in the side of the mountain.



COASTAL GLACIER terminates at Disenchantment Bay in Alaska. Fluctuations in the position of the terminus of such a glacier are often much greater than they are in the case of glaciers that terminate on land. Large blocks of ice "calve" off the front face of the glacier to form icebergs. Both photographs on this page were made by Bradford Washburn of the Boston Museum of Science.

ninth and the 13th century, when glaciers were smaller and sea ice more restricted than at present. An even milder interval of climate apparently coincided with the beginning of the Christian Era.

We have mentioned the evidence from bogs in western Norway suggesting that from about A.D. 50 to 400 the climate of Scandinavia was warmer than it was earlier or later. In addition, during the 1930's several arrows, with wood shafts intact, were found in the mountains of Norway at sites from which snowbanks had recently disappeared. The oldest arrows date from 400 to 600. Because the shafts could have escaped decay only by being covered by a continuous mantle of snow and ice, the preservation of the arrows suggests that snowbanks had a similar volume in the periods 400 to 600 and 1930 to 1940, and that during the intervening time they were larger. If these inferred changes in snowbank volume paralleled similar trends in glacier volume during this time interval, glaciers probably were at least as restricted during the early centuries of the Christian Era as they are now.

The evidence from the North Atlantic area thus indicates that most of the interval between the beginning of the Christian Era and the 13th century was relatively warm. However, a growing number of carbon-14 and lichenometric dates for moraines from southern Alaska and the western U.S. leaves open the possibility that one or more episodes of glacier expansion, as yet unrecognized in most areas, may have occurred during this interval.

During the past few centuries glaciers in mountain regions throughout the world have advanced markedly, and many have attained their maximum neoglacial position. The time of this interval has been rather accurately fixed by the historical records mentioned above, as well as by carbon-14 and botanical dates. Although some glaciers appear to have advanced as early as the 12th or 13th century, in most areas recent glacier maxima were reached at various times from the middle of the 14th century to the middle of the 19th. No consistent worldwide pattern is discernible among the many moraines deposited during these centuries, owing either to insufficient data or to the varied behavior of the different glaciers. In some regions, such as the Alps, consistent regional behavior seems to be recorded. In other areas this is not so.

The culminating advances of the mid-

dle and late 19th century were followed by a drastic retreat of alpine glaciers on a worldwide scale. This marked reversal closely coincides with a distinct global warming trend that led to an increase in the mean world temperature by nearly one degree C. Many large glaciers receded significantly and became markedly thinner, and many small glaciers became stagnant and began to disappear. This trend toward milder temperatures continued into the 1940's, when there was a reversal leading to cooler and moister climate in several glaciated regions. The response of glaciers to this reversal was first noted in the late 1940's and early 1950's. Since then rates of glacier recession have generally decreased and some glaciers in the maritime coastal mountains of northwestern North America and in Switzerland have advanced. If the trend toward cooler and moister climate continues, glacier growth may well become a worldwide phenomena leading to a significant increase of ice in alpine regions.

To what degree can we now predict future glacier fluctuations? Although significant insights into the relation of climate and glaciers have been gained in the past decade, knowledge of glaciers has not yet reached the point where confident predictions of their future behavior can be made for wide geographic regions. Since climate apparently does closely control glacier behavior, an even more important question can be asked. What effect will man have in altering worldwide climate trends and ultimately the extent of glaciers?

Experimentation clearly indicates that man can directly control glacier melting and glacier growth on a limited scale by increasing or decreasing the albedo, or light-reflecting power, of glacier surfaces. Of even greater potential importance, however, is man's progressive alteration of the atmosphere. The injection of increasing amounts of carbon dioxide into the atmosphere by burning industrial fuels can lead to global warming and glacier retreat. On the other hand, the accumulation of solid pollutants in the upper atmosphere could have the opposite effect, decreasing the world temperature by reducing the amount of incoming radiant energy. This might well lead to glacier expansion and perhaps even generate a new ice age. It is becoming increasingly obvious that the balance between glaciers and climate is extremely delicate and that this balance might easily be disturbed.

RECENT ADVANCE of the Meares Glacier in Alaska is demonstrated by the fact that it is overrunning the trees in a forest (*left*). The photograph was made recently by W. O. Field.

Basic Research at Honeywell
Research Center
Hopkins, Minnesota

The Effect of Insecticides on Enzymes

The study of the effect of pollutants on the olfactory system has led to a possible explanation of how common chlorinated hydrocarbon insecticides kill.

Although developed to control unwanted insects, chlorinated hydrocarbon insecticides (DDT, Chlordane, Lindane and others) are now regarded under certain conditions as pollutants which can endanger other desirable forms of life. Like all pollutants, the first step in controlling usage is to set permissible standards of concentration of the insecticide in the environment. However, unlike most pollutants, these insecticides become stored in the body tissues of the humans and other animals who acquire them from water and food. For example, fish kills have been noted where the concentration of DDT in the fish was far greater than the concentration in the water. To further complicate the problem, the mechanism by which these insecticides operate is unknown so that toxic limits cannot be set.

Honeywell scientists have been working on means of detecting, identifying and measuring pollutants in the water and atmosphere for a long time. One promising avenue of research in detecting pollutants has been the study of the olfactory mechanism.

Previous studies in olfaction suggested that the absorbance spectra of water soluble extracts from olfactory mucosa were altered after they were exposed to certain odorants. Because the first odor-sensing action is believed to take place on the plasma membrane where enzymes could play a major role in olfaction, the odorant effects on membrane-bound enzymes are under investigation.

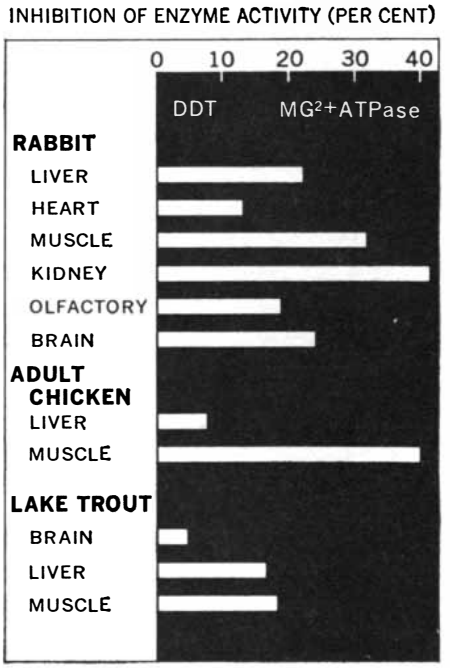
The ATP (adenosine triphosphatase) enzymes, are extremely important in metabolic processes such as cation transport through plasma membranes and oxidative phosphorylation, a process by which food energy is converted and stored as usable chemical energy in the form of ATP. Inhibition of these enzymes could have a profound effect on metabolic processes.

Sodium-potassium ATPase is found exclusively in plasma membranes and plays an important role in maintaining the electrical potential across nerve-cell membranes by cationic transport. This enzyme is able to transport sodium out of the cell and potassium into the cell, both against high concentration gradients of the ions. To carry on this process, ATP is required as an energy source and is produced within the cell itself in a subcellular organelle called the mitochondrion.

The inhibition of the ATPases by chlorinated hydrocarbon insecticides may therefore seriously change cellular energy production and nerve action.

Honeywell scientists succeeded in isolating the plasma membranes from olfactory tissue by differential centrifugation and were able to concentrate the enzymes. Experimentation confirmed the toxicity of chlorinated hydrocarbon insecticides. DDT, chlordane, lindane, aldrin, dieldrin and dicofol were all shown to have an inhibitory effect on the ATPase enzymes.

The lethal hyperactivity commonly observed in cockroaches exposed to DDT concentrates may be associated with the inhibition of the magnesium ATPase. The response was similar to asphyxiation, but it appeared that the agent was interfering not with oxygen intake but with the primary energy production (ATP). In chlordane exposure, the opposite effect has been observed; there was a strong inhibitory effect on the sodium-potassium ATPase and the cockroach seemed to be paralyzed, as if a nerve had been shorted.



The inhibitory effects of the chlorinated hydrocarbon insecticides on the ATPase systems could account for the observed toxic effects. However, results of fairly extensive studies indicate that although these insecticides inhibit the ATPase system in general, there may be particular tissues which are more sensitive to specific insecticides (see figure).

Scientists at Honeywell are now working with entomologists at the University of Minnesota to find out the specific reactions of various animals exposed to insecticides, since a wide variety of effects have been seen. Studies have shown that animals concentrate DDT and other insecticides in their tissues. The amount of DDT present in some lake trout has been reported to be more than 13,000 times greater than the solubility of DDT in water. This concentration will, of course, affect any limits which must be set to protect the fish. Also detoxification of insecticides occurs slowly and knowledge of the extent and speed of this process is essential for determining threshold limits. In addition, the site that the insecticide affects may have a bearing on its toxicity. For instance, Honeywell and University scientists have found that in bluegills DDT affects muscle, primarily, and brain ATPase activities to a lesser extent.

After extensive study of these and other variables, they hope to be able to determine whether or not the ATPase enzymes can be used as a measure of the toxicity of chlorinated hydrocarbon insecticides. If they can, studies on animals will be extended to determine which tissues are most affected by which insecticides.

The goal is to set threshold limits for the use of these insecticides but the problem is extremely complex because of the variety of mechanisms which must be understood before the toxicity can be measured on the one hand and the obvious beneficial effects of these insecticides on the other.

If you are working in the area of olfactory sensing or the effects of chlorinated hydrocarbons on ATPase inhibition and wish to know more of Honeywell's work, please contact Dr. Robert Koch, Honeywell Corporate Research Center, 500 Washington Ave. S., Hopkins, Mn. 55343. If you have an advanced degree and are interested in a career in research at Honeywell write to Dr. John Dempsey, Vice President, Science and Engineering, Honeywell, 2701 Fourth Ave. S., Mpls. Mn. 55408.

Honeywell
AUTOMATION

DOUBLE BURIAL at the Port au Choix cemetery area designated Locus II includes the well-preserved skeleton of a young woman who was interred resting on her left side and cradling a small child in her arms. Grave goods accompanying the burial include a number of quartz pebbles (*small white spheres*) and a barbed point made of caribou antler (*below lower jaw*), perhaps used as a spearhead.

An Archaic Indian Cemetery in Newfoundland

Little has been known about the hunters and gatherers who roamed northeastern North America 4,000 years and more ago. The richness of their culture is now disclosed by work at a beach burial ground

by James A. Tuck

A remarkable accident of preservation has helped to resolve a problem that first began to trouble New World archaeologists early in this century. At that time the first in a series of enigmatic archaeological sites, best described as great boneless cemeteries, was discovered in northeastern North America. Plainly locales where many persons had been ceremonially buried, the sites—which now number in the dozens—showed only traces of human bone and of artifacts made of perishable materials.

Because red ocher was almost always present in the graves the cemetery-makers were given the name (among others) "Red Paint People." Various objects made of stone, found along with the ocher, provided the small amount of information about the Red Paint People that was available. The cemeteries were found from Maine to Newfoundland, but few dwelling sites were unearthed. Indeed, until 1967 archaeologists had little beyond the evidence of mortuary ceremonialism provided by the cemeteries themselves to help reconstruct how the Red Paint People lived. The lack of organic remains in the graves made the burials almost impossible to date with precision. It was safe to say only that the stone artifacts were typical of the New World Archaic (or intensive hunting-gathering) period. The material resources of the Archaic Indians, although they were more sophisticated than those of the earlier, big-game-hunting Paleo-Indians, did not include a knowledge of agriculture or the art of pottery-making. In spite of this handicap and the pressures of a sometimes precarious existence based on hunting and gathering, the Archaic Indians managed to live a full life.

Just how full this life was, both technologically and ceremonially, is now revealed by the contents of a unique Archaic Indian cemetery discovered three years ago on the island of Newfoundland. Unlike the boneless cemeteries unearthed earlier, this site proved to contain the well-preserved skeletons of more than 100 individuals, together with a vast wealth of objects made of perishable materials that, like the skeletons, had been safely preserved for thousands of years. The Newfoundland cemetery was a preferred Archaic Indian burial ground over a span of nearly 1,000 years, from sometime late in the third millennium B.C. down to the final centuries of the second millennium.

The site is located at Port au Choix, a small fishing village on a point of land on the west coast of Newfoundland's Great Northern Peninsula [*see illustration on page 117*]. The village overlooks a stretch of water known as the Back Arm, a sheltered embayment that opens into the Gulf of St. Lawrence. In the past the water level was higher and the Back Arm extended farther inland; several raised beaches remain as evidence of the earlier encroachment.

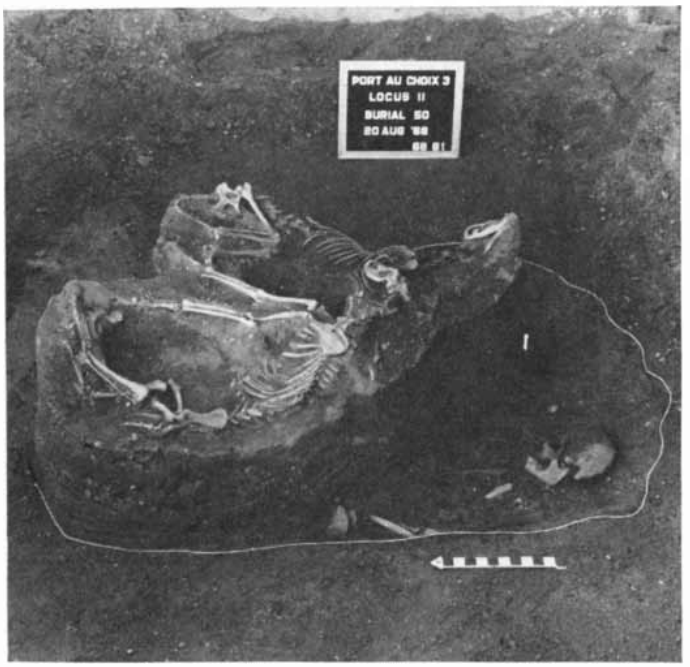
One of the raised beaches, which lies about 19 feet above high water today, is almost a mile long and varies from 30 to 70 feet in width; a good part of Port au Choix is built on this natural terrace. The fine sand of the ancient beach is easily dug with the simplest of tools, which may partly account for its selection as a burial ground by the Archaic Indians. The beach has been kind to archaeologists in another way: it is extremely alkaline, with a *p*H that averages close to 8.0. This condition and the excellent natural drainage of the beach account for the burials' remarkable state of preservation.

Soon after the initial find was made the Memorial University of Newfoundland sent a group under my direction to Port au Choix to excavate the site with the sponsorship of the Department of Provincial Affairs and the National Museum of Canada. We discovered that there were three distinct sets of burials. The first, which we called Locus I, had been exposed in the fall of 1967 as the foundations for a new building were being prepared. The second and largest set of burials, which we called Locus II, was still mainly undisturbed. We also found two infants' graves that were evidently not associated with the other burials; this set we called Locus IV. (A settlement site that belonged to a later Dorset Eskimo group had already been named Locus III.)

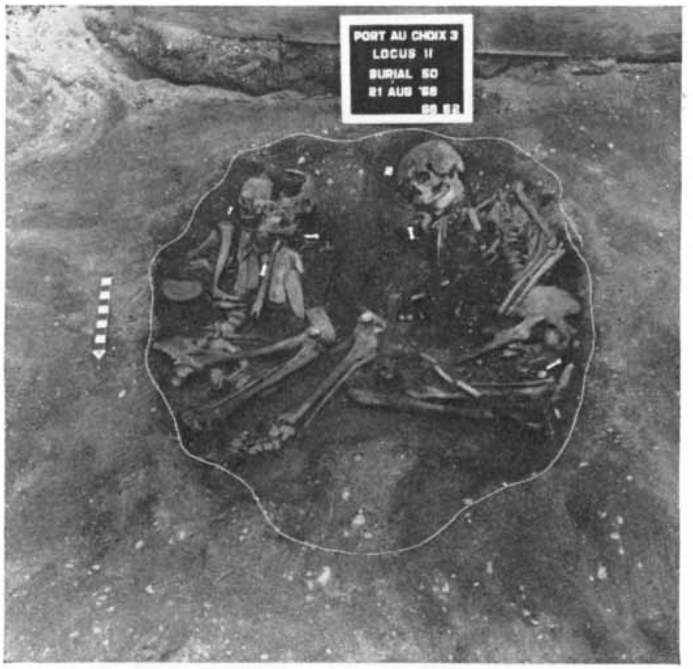
Most of the material from Locus I had been removed at the time of discovery. The foundation builder and other observers reported that the eight skeletons then unearthed had their legs drawn up in a "flexed" position rather than stretched out in an "extended" one. The graves may have been arranged in a rough circle; at least one of the eight skeletons, later identified as male, had a bundle of grave goods cradled in its arms. Red ocher was present in most if not all of the graves.

Our group was able to locate one additional burial at Locus I. It contained the skeletons of two females in a grave pit lined with red ocher. One skeleton rested on its back, its lower legs doubled backward. The other lay in a tightly flexed position, probably because the body was partially decomposed at the time of burial and had been bound up in a shroud or a blanket. Our discovery of the additional grave at Locus I brings the total of individuals buried there to 10. Only two were males.

The second cemetery at Port au Choix is the most extensive Archaic Indian burial site known in Newfoundland and



MOST ELABORATE BURIAL at Locus II contained the remains of two large dogs that had been placed above a grave containing an adult man and woman. No cause of death was evident for the dog in the foreground; the skull of the other had been crushed by a blow.



LOWER LEVEL OF SAME BURIAL contains a number of grave offerings in addition to the woman (*left*) and the man (*right*). The white arrows, placed in the grave by the excavators, point to some of the less visible objects. The posture of both skeletons is flexed.

DECORATION OF CLOTHING is evidenced by the double row of beads that forms an inverted "*V*" with its apex at the skull of a young woman whose burial was unearthed at Locus I. The beads, made from the shells of small marine snails, were probably sewn to the edge of a hooded skin garment that was interred with the body. No garments survived millenniums of burial at the site.

TIGHTLY FLEXED BURIAL of a woman, unearthed at Locus II, was accompanied by three large stone axes. The degree of flexing suggests that the body was quite decomposed. Those Archaic Indians who died during the winter months evidently did not receive burial until after the ground thawed in spring or early summer. Sometimes only the skull and a bundle of bones were finally buried.

one of the largest in all of northeastern North America. It consists of 53 separate burials, arranged in three relatively compact groups. The burials contain a total of nearly 100 complete or incomplete skeletons and literally thousands of hunting implements, tools and other grave goods. Boulders or slabs of rock were placed over many of the burials, either to act as markers or to prevent disturbance by wild animals. Many of the graves nonetheless showed signs of intrusion. Usually the intrusions were the result of natural events, such as tree-root growth or the shifting of soil by the fall of a tree. Some of the intrusions, however, were evidently the result of visits to the site by later Archaic Indians.

After the last of these visits the site was covered with forest for perhaps 3,000 years, as we can tell by the accumulation of a stratum of dense, peaty humus as much as a foot thick on top of the beach sand. Where the sand and the humus meet, a thick calcareous crust was formed by the interaction of the alkaline sand and the acid humates percolating downward. The crust was undisturbed above all the burials found in Locus II.

The methods of burial at Port au Choix varied in both the posture of the dead and the condition of the individual. Infants and young children were generally buried in an extended position. Older children and adults who had been buried soon after death were usually interred in a flexed position. A practice known as secondary burial, which involves preliminary decomposition of the body and the later interment of the skull and a bundle of bones, was also known at Port au Choix. It is clear, however, that no rigid line separated the two practices. Skeletons were found reflecting all stages of decomposition, which suggests that people who had died during the winter were not buried until the ground thawed in spring or early summer.

Most of the articulated skeletons we unearthed at Locus II had been buried resting on their left side, but we could detect no particular orientation of the dead with respect either to the water of the Back Arm or to the points of the compass. The entire cemetery area, however, occupies an east-facing slope that overlooks the water; Archaic Indian burial grounds in similar locations are found elsewhere along the northeastern seaboard.

James E. Anderson of McMaster University is analyzing the skeletal material from Port au Choix. Some preliminary findings, particularly concerning the sex and age of the sample, are available. For example, it is apparent that the dead at Port au Choix were interred without regard to sex or age. The sex of each skeleton cannot always be determined, but of the burials at Locus II, 23 can be established as male and 18 as female. Adding to these the two males and eight females at Locus I, it appears that the ratio of the sexes buried at Port au Choix is almost equal.

As for the ages of the individuals at Port au Choix, the large sample from Locus II provides a good average distribution. Almost exactly half of the skeletons—a total of 44—are those of adults. Only seven of these individuals were more than 50 years of age; one was

IMPORTANT YOUNG MAN, perhaps a shaman, was buried with grave goods of both ceremonial and utilitarian type. The latter include a barbed harpoon head near the skull (*uppermost arrow*) and two harpoon foreshafts that lie across the man's chest. The chief ceremonial object is the large killer-whale effigy near lower jaw (*see illustration on next page*).

between 18 and 21. The remaining adults can be called "middle-aged." Fifteen of the individuals at Locus II were between six and 18 years old and two between two and six years old. Infant mortality was high among these Archaic Indians: 15 of the dead were less than two years old and 12 were newborn.

The causes of death among the people buried at Port au Choix are obscure. The condition of the skeletons suggests that most individuals were healthy and robust. Tooth wear—perhaps partly due to softening hides by chewing—frequently exposed the adults' pulp cavities and even caused the loss of some teeth, but tooth decay was virtually absent. Some adults suffered from a mild arthritis of the spine, and an unusual manifestation of arthritis also affected the finger joints of several others. A perforating bone abscess, possibly the result of the bone-marrow malignancy known as multiple myeloma, appears on one skull. The principal evidence of trauma consists of broken ribs and several well-knit fractures of the arm and the leg.

The Port au Choix people are clearly unrelated to the Eskimos who have occasionally inhabited Newfoundland in both ancient and modern times. Comparison with other prehistoric New World skeletons is now in progress. So far the most significant finding is a close relationship between the Port au Choix population and a small but important series of Archaic Indian skulls discovered in the early 1960's on Morrison Island in the Ottawa River valley of Ontario. Carbon-14 dating shows that the Morrison Island skulls are about 4,400 years old: the dates fall between 2500 and 2300 B.C. This is in good agreement with the carbon-14 determinations recently obtained for the Port au Choix material. Locus II has proved to be the oldest of the three sites; its earliest date is 2340 ± 110 B.C. Other material from Locus II graves is dated from 1880 to 1740 B.C. Locus I is younger than Locus II. The single carbon-14 date so far available from Locus I is 1460 B.C. The site of the two infant burials, Locus IV, is the most recent of all, with a carbon-14 date of 1280 ± 220 B.C.

When one considers that the grave goods at Port au Choix were produced over a span of 1,000 years, it is not surprising that the artifacts show some differences in style. Nonetheless, they clearly belong to a single tradition. In the following reconstruction of this Archaic Indian tradition I have concentrated on the material from Locus II, which is the most abundant. The reconstruction nonetheless applies equally well to the Archaic Indians who left their dead at the other two loci in later years. It seems hardly necessary to warn the reader that, as with all such extrapolations, my reconstruction incorporates many speculative elements.

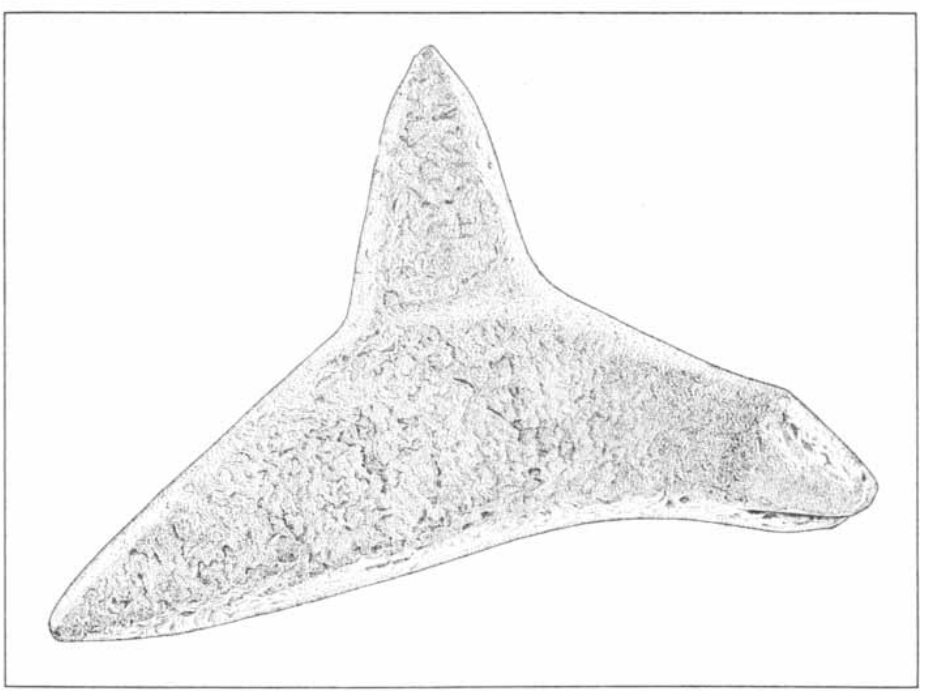
How did the Archaic Indians wrest a livelihood from this forested northern coast? We can begin to answer this question by making an inventory of the hunting gear found among the grave goods at Port au Choix. We uncovered a number of "bayonet" points (probably the heads of lances) fashioned by grinding pieces of slate to the desired shape. Several spear points with stemmed bases were also made of ground slate [*see top illustration on page 119*]. Stone was not the only material used to make objects of this kind; we found four stemmed spear points, three of them at Locus II and one at Locus I, that had been made by grinding pieces of bone instead of slate. There were also several "bayonet" points fashioned from bone and marked with deep barblike slashes near the tip [*see illustration on page 121*].

Presumably both lance points and spear points had once been mounted on wooden shafts and had been used as thrusting weapons to kill various kinds of game, particularly migrating caribou. A ceremonial aspect to the inclusion of these points in the burials at Port au Choix is evident in the fact that some of them were "killed," that is, purposely broken, before interment. It is noteworthy that, of all the stone points found at Port au Choix, only one was fashioned by chipping rather than grinding; it is a side-notched point made from gray quartzite.

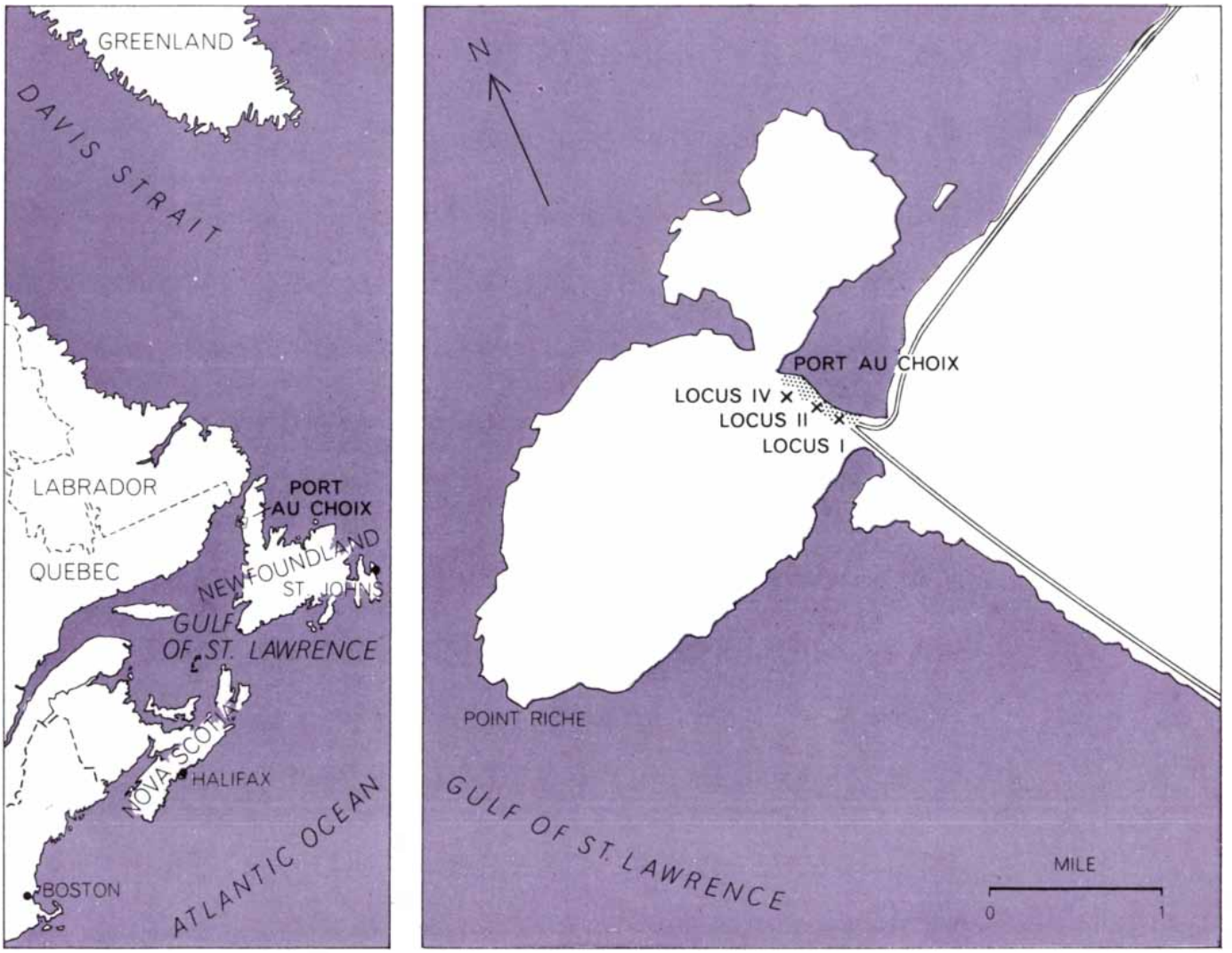
In addition to thrusting weapons, the Archaic Indians were equipped with daggers, which were presumably useful for dispatching wounded game. We found one such dagger made from walrus ivory, another made from antler and several made from the leg bones of caribou. The walrus-ivory dagger may once have been carried in a sheath that combined pieces of antler with some perishable substance; only the antler portion remained intact.

Both the thrusting lances and the spears may have been used for hunting at sea as well as on land. The grave goods also include artifacts specifically designed for hunting sea mammals. These are harpoon heads of both the barbed and the toggling type, made from bone and antler and perforated for attaching a line. The toggling harpoon heads had an open socket to receive a foreshaft [*see illustration on page 120*]. Only two such foreshafts, however, were found in the graves. The barbed harpoon heads are equipped with from one to three barbs, and the base of the head is tapered, presumably for socketing in a wooden handle.

The graves also contained a number of unusual foreshafts made of whalebone. All of them have a slotted base, apparently for "tongue and groove" at-



KILLER-WHALE EFFIGY represents the animal in motion; its head, mouth closed, is at right and its distinctive high dorsal fin dominates the sculpture. Two killer-whale effigies were found at Port au Choix; one is shown *in situ* in the illustration on the preceding page.



CEMETERY SITE lies on a promontory of the west coast of Newfoundland's Great Northern Peninsula (*left*), now occupied by a fishing village. The embayment where the village stands opens northeastward and the three sets of burials (*right*) face the sunrise.

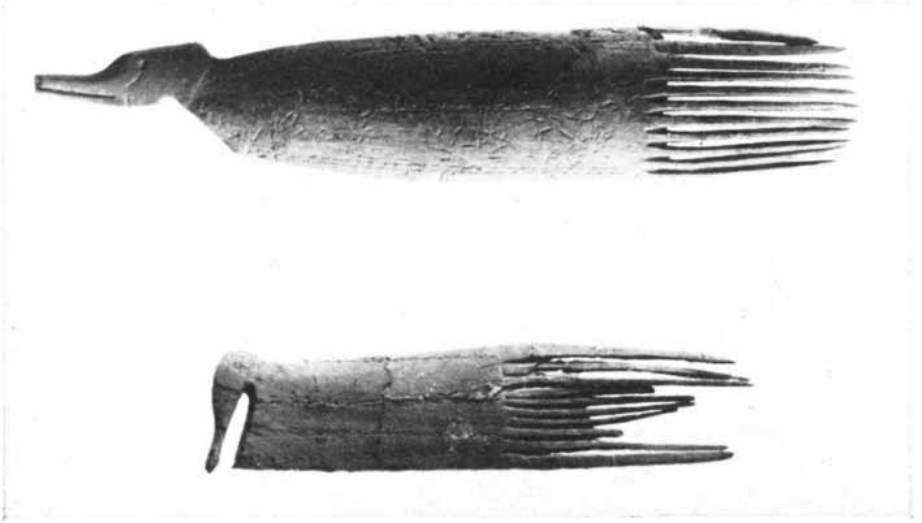
tachment to a wooden handle. The tips of the foreshafts exhibit a variety of shapes. Some are slotted and have lugs or small holes that would aid in lashing a harpoon head into place. Others merely have a broad "*V*" or are bluntly tapered. Curiously, none of the harpoon heads from the graves nor any of the lance points we unearthed would fit the whalebone foreshafts. Perhaps the foreshafts had originally been fitted with heads of chipped stone that for some reason were not included in the burials. The one chipped-stone point unearthed at Port au Choix was not found in association with the whalebone foreshafts.

A slender foreshaft helps a hunter to drive the tip of his weapon deep into his prey, and a harpoon enables him to hold the animal at the end of a line. It thus seems logical to suppose that the Archaic Indians used both tipped foreshafts and harpoons in the pursuit of free-swimming animals whose vital organs are not easy to penetrate: seals, walruses and perhaps even small whales. This, in turn, supposes that the hunters were also sailors. The key evidence that the Archaic Indians were familiar with overwater travel is the fact that they reached the island of Newfoundland in the first place. The nearest point on the mainland is the coast of Labrador to the north, a minimum of 10 miles away across the often treacherous Strait of Belle Isle; to reach Newfoundland from Nova Scotia to the south calls for a 60-mile voyage.

The evidence for marine hunting from the graves at Port au Choix suggests that the Archaic Indians of Newfoundland knew more than just how to cross water; they must have been skillful boatbuilders and competent seamen. No boats or boat fragments have been found in the graves. Tools suitable for the construction of dugouts or skin boats were available, however, as were the raw materials for both kinds of craft.

Unless lance points and spear points also served the Archaic Indians as knives, there is a curious absence of implements for the dressing of game and the preparation of other food at Port au Choix. The easiest explanation would be that such work was done with knives made of chipped stone, and that these were not considered a necessary accompaniment for the dead. In contrast, the various tools used to prepare animal skins are well represented among the grave goods. They include scrapers made from the shoulder blade of caribou and two-handed "beaming" tools made from the split long bones of various mammals. Bone awls and a number of extremely fine bone needles attest to the use of sewn garments. Eight of the needles were found in a decorated needle-case that had been made from a caribou limb bone.

With the exception of a few whetstones, tools for working stone also appear not to have been considered necessary grave furnishings. We know, however, that the Port au Choix people used hammer stones for chipping and peck-



TWO COMBS made from caribou antler have been carved to represent birds' heads. The top comb decoration represents a merganser, the bottom a swan, goose or broad-billed duck.

ing. Unfinished stone tools in the graves clearly bear the marks of these implements.

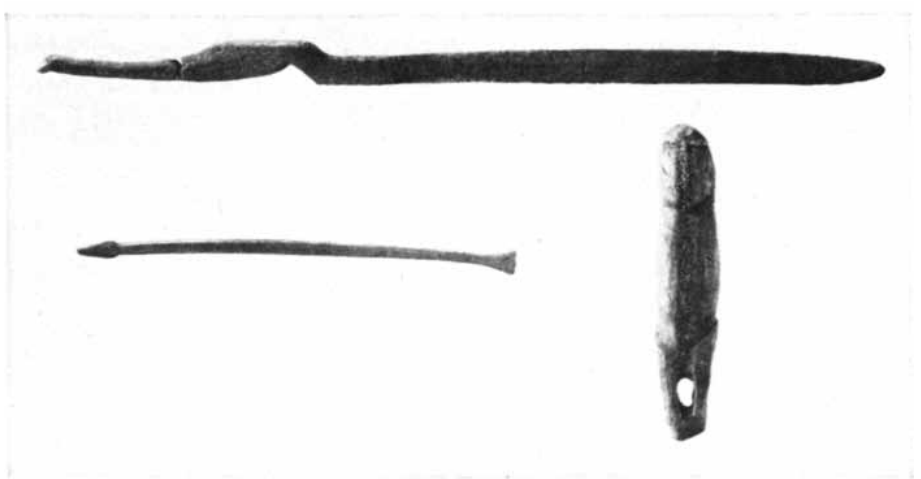
The accident of preservation that enabled us to recover so many artifacts made of bone, ivory and antler from the graves of Port au Choix unfortunately did not protect objects made of wood and bark. Of the numerous artifacts the Archaic Indians must have fashioned from these materials only a few scraps of birch bark survive. Woodworking tools, however, are found in abundance. These include axes, adzes and gouges of several sizes, styles and materials. Most of the axes and adzes were made of slate, but some adzes were fashioned from ivory [*see illustration on page 120*]. The incisor teeth of beavers were also among the grave goods. They had been modified into woodworking tools by beveling and pointing their working edge or by shaping them into miniature gouges. A handle made of antler, with a beaver incisor hafted at each end, might even be regarded as foreshadowing the modern multipurpose implement of the subarctic: the "northern crooked knife." Some of the stouter bone awls among the grave goods may have been used in the manufacture of bark vessels for storage and cooking as well as for working with leather. The number and variety of such implements suggest a major woodworking industry among the Archaic Indians of Newfoundland.

A large proportion of the material unearthed at Port au Choix is nonutilitarian and may be classified as being either decorative or—with more caution—magico-religious. In the first category are thousands of beads, made from the shells of small marine snails, that were presumably used to decorate clothing and skin pouches and perhaps were sewn onto strips of hide. Strips decorated in this way were described by the first European visitors to Newfoundland. Another animal product used for beads by the Archaic Indians was the teeth of a species of skate common in Newfoundland waters; these distinctive objects may also have been used as hair ornaments.

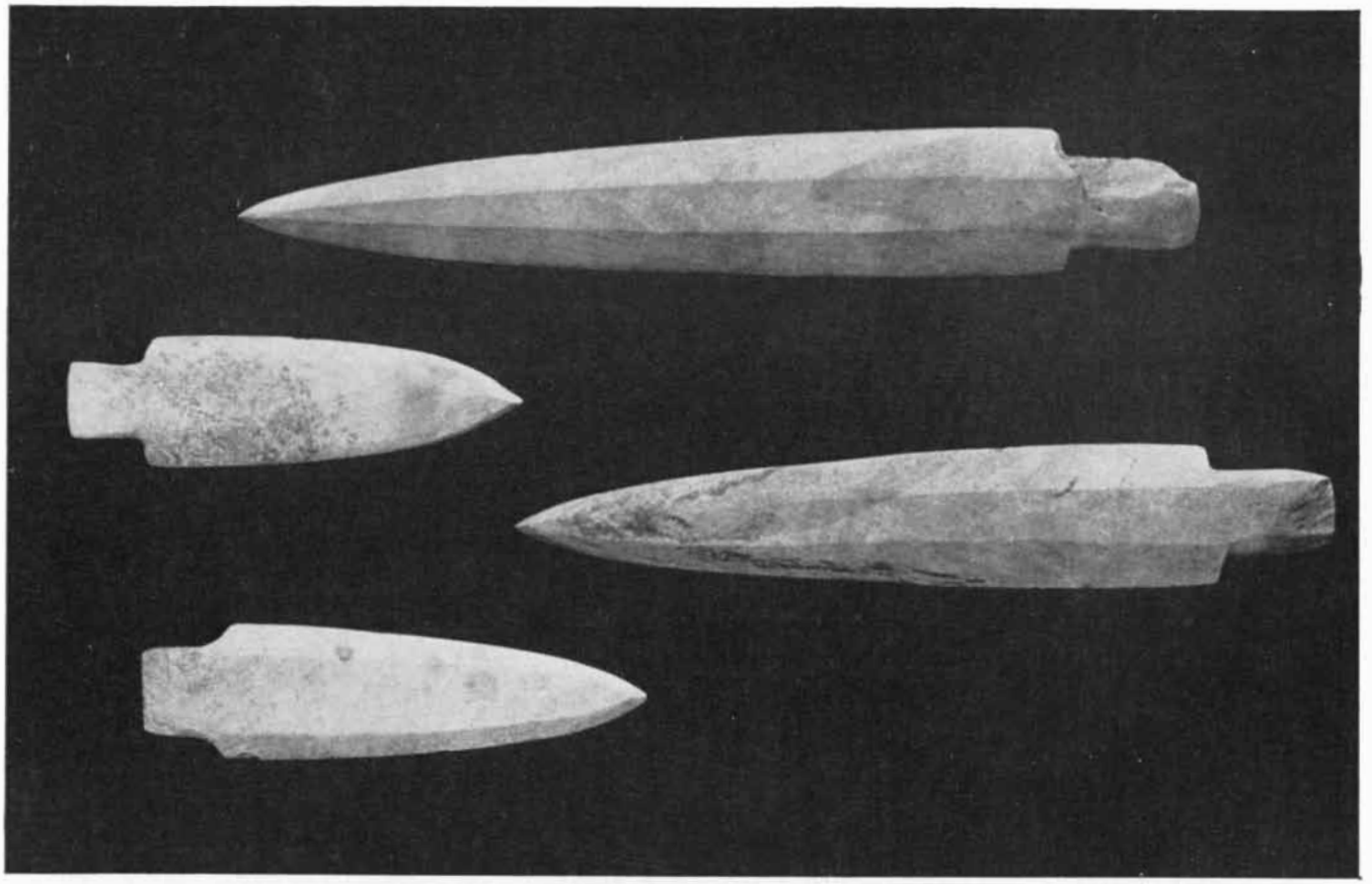
A number of bone and antler pendants and pins, together with three bone combs, are objects that, although they are certainly decorative, must also reflect the religious beliefs of their makers. Some of the pendants depict birds such as ducks and loons; others are human figures, geometric forms and objects that have spearlike and swordlike shapes. Some are of slate and soapstone and one is formed from native copper. The pins display a variety of decorations; perhaps the most striking is one that ends with a bird's head, probably meant to represent the now extinct great auk [*see bottom illustration on this page*]. Of the three combs, two are decorated with a bird's head. One of them evidently represents a merganser; the other could be a swan, a goose or a broad-billed duck.

The grave goods also include collections of "exotic" stones, such as crystals of quartz, amethyst and calcite, and pebbles of quartz and iron pyrite. The pyrite pebbles may have belonged to fire-making kits, although the striking stones needed to complete the kits are rare and in only one case was a "striker" found in direct association with a pyrite. Other stones, either left unmodified or shaped to a greater or lesser extent, apparently represent different animal species and could have served as amulets. This seems indisputable in the case of two large, three-lobed sculptures that apparently represent killer whales [*see illustration on page 116*]. They invite speculation about parallels with the killer-whale cults found among the hunters of sea mammals in many parts of the world.

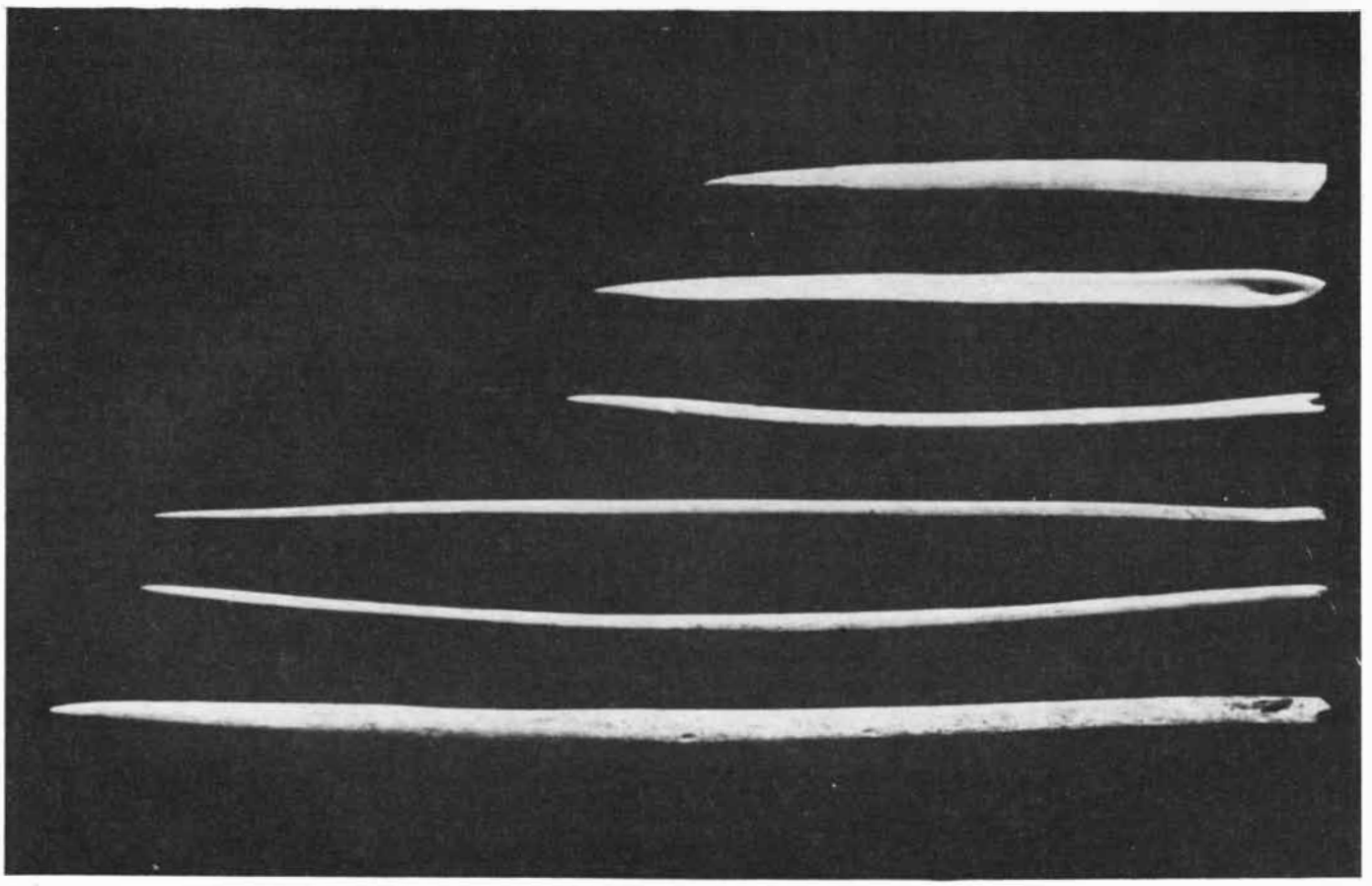
In addition to the snail shells and skate teeth used for decoration, other remains of animals are present in the burials. Some of them, such as seal claws and the teeth and jaws of caribou, bear, dog, wolf, fox, marten, seal and otter, may once have been sewn on clothing. The heads, bills and feet of gulls, mergansers and great auks may also have been sewn on clothing, but many of these objects seem instead to be the components of medicine bundles. Some of the objects still have specific significance as charms and amulets among the living aboriginal inhabitants of the subarctic and Arctic. The likelihood that



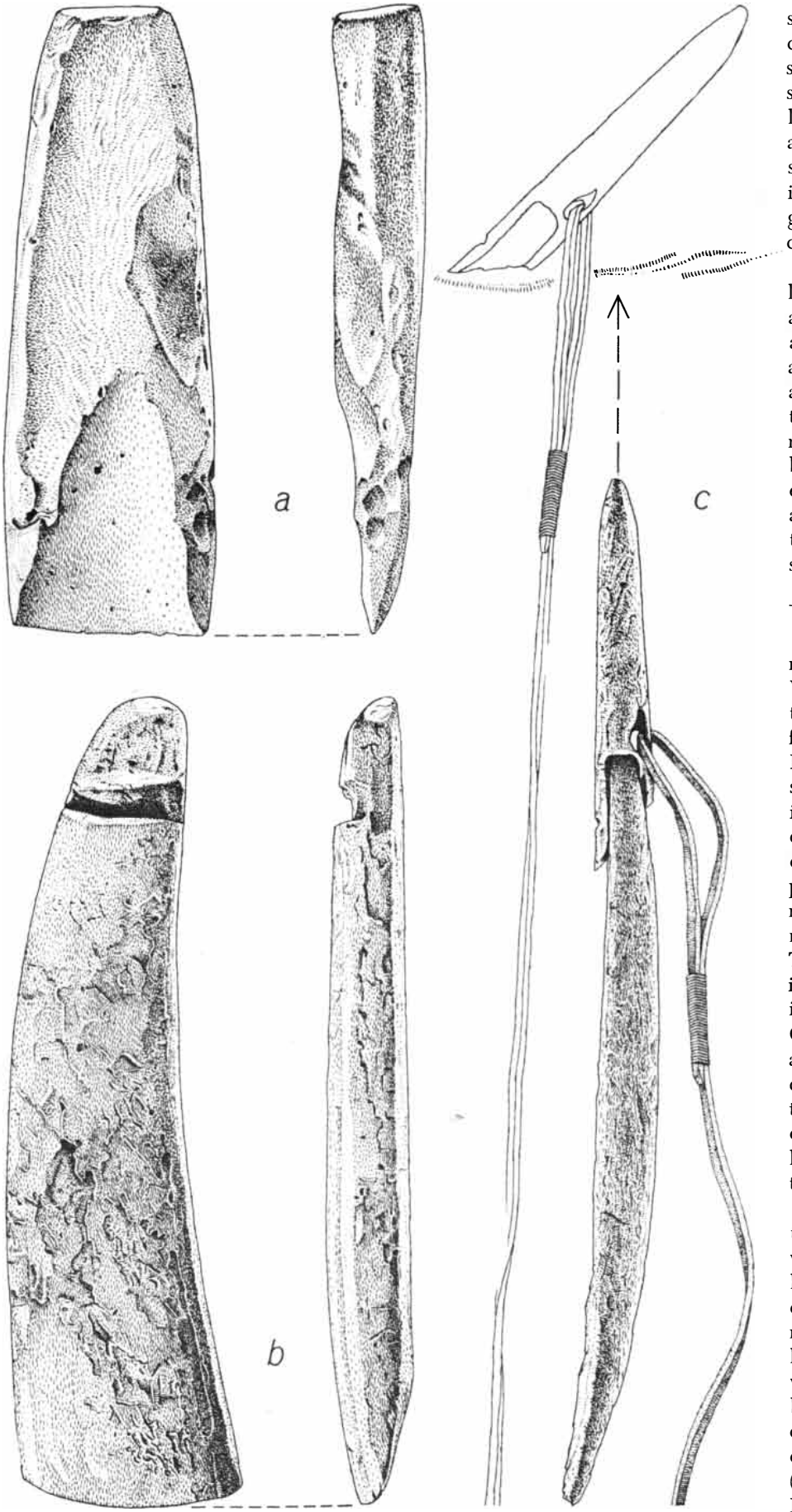
THREE DECORATIVE OBJECTS found among the grave goods at Port au Choix are a human figurine, made from antler, and two bone pins. The carving on the longer pin represents a great auk; the carving on the second may have been meant to represent a spear.



POINTS FOR WEAPONS were skillfully formed from slate slabs that were shaped by grinding. Points of both sizes were probably hafted to wooden shafts to make spears and lances and both kinds of weapon used in hunting caribou ashore and sea mammals afloat.



FINE NEEDLES, some made from the bones of birds, were formed by carving and grinding. The abundance of needles in the graves at Port au Choix, along with tools for dressing animal skins, suggests that the Archaic Indians dressed in tailored skin garments.



GOUGES FOR WOODWORKING were made from stone (*a*). Adzes were occasionally made from ivory (*b*). The number and variety of woodworking tools at Port au Choix suggest that the Archaic Indians were skilled in the use of wood. The principle of the toggling harpoon head is illustrated at right (*c*): the head, fitted over its foreshaft, is driven deep into the quarry. Tension on the harpoon line then twists the head, setting it sideways in the wound.

such beliefs also existed among the Archaic Indians is strengthened by our observation that items of possible ritual significance are not distributed randomly among the burials. For example, when a comb, a pin or a pendant that represents a particular species of bird is found in a grave, the heads, bills or feet in the grave frequently belong to the same species of bird.

So far we have completed only preliminary correlations between the age and sex of an individual and the kinds and quantity of goods in the grave. It appears, however, that adult males were accorded the most grave goods. Although some children's graves were richly accoutered, others held nothing but a few beads and ornaments. Women's graves frequently contained charms and amulets of the kind that modern tradition associates with male activities such as hunting.

What conclusions about Archaic Indian life can be drawn from the material unearthed at Port au Choix? We know that much of the durable contents of the graves resembles material found from Maine to as far north as Labrador (where Archaic Indian living sites have recently been discovered). It is thus apparent that a specialized type of culture existed in this part of northeastern North America during the latter part of the third millennium B.C. and most of the second millennium. I have named this culture the Maritime Archaic Tradition. The area where it flourished includes three separate biotic provinces—the Eskimoan, Hudsonian and Canadian—each with a distinctive flora and fauna. All three have enough in common ecologically, particularly in terms of the animal species most exploited by man, to have allowed a relatively uniform type of culture to flourish throughout the region.

At the time of the first European contacts the sea mammals of the region that were important to man included the harp seal, the harbor seal, other seal species and the walrus. It was probably much the same 4,000 years earlier. The land animals of greatest importance were the woodland caribou, the black bear and the beaver, and the birds included the great auk, the murre, the eider duck and the Eskimo curlew (whose flights once darkened the sky). Every summer spawning salmon filled the coastal rivers of the northeast. In the more southerly parts of the region there were the elk, the moose and perhaps the Virginia deer in addition to the caribou. This bountiful animal life, supplemented

with certain wild vegetable foodstuffs, provided the economic basis of the Maritime Archaic Tradition.

The Archaic Indians, presumably organized in small bands, would have followed a more or less fixed seasonal round in exploiting the region's food resources. In the late fall each year, when the caribou congregate after the first snow to move from summer to winter grazing, the migrating herds were a prime target. It is not difficult to kill caribou as they migrate: the animals can be diverted into dead-end runways or speared as they swim across or ford rivers and lakes. The caribou meat can then be dried or, if the temperature allows, preserved by freezing. A few days of caribou-hunting might furnish enough meat for the winter months, and this basic winter ration could be supplemented with fresh kills of beaver, hare or larger mammals such as bear.

During the winter months the Indians of the Maritime Archaic Tradition probably stayed close to the inland sites where the caribou were killed. As spring approached and the caribou meat dwindled, the bands probably moved to the coast to hunt the harp seals and their young as they drifted southward with the pack ice. After the seal harvest there might be another inland caribou hunt to take advantage of the migrating herds' return to summer pasture. The caribou meat from these kills that was not eaten on the spot could be dried for later consumption.

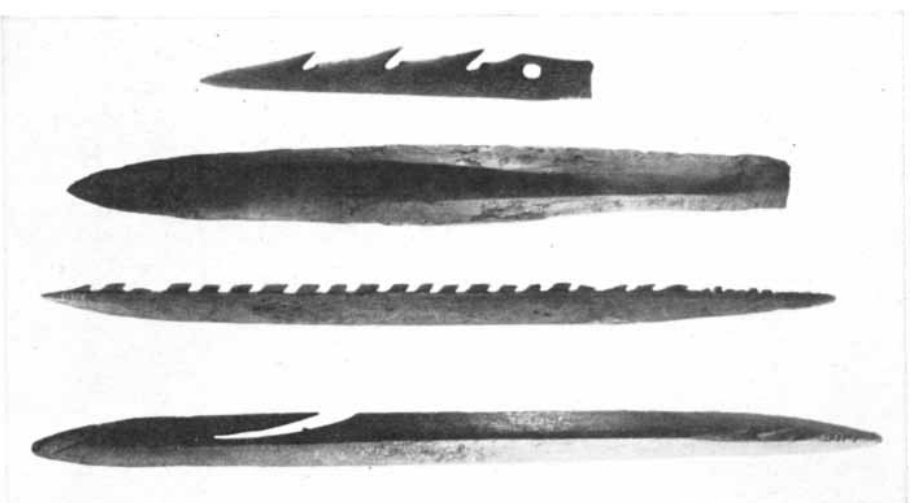
The Indians of the Maritime Archaic Tradition would then have returned to the coast to spend the summer fishing for the migrating salmon, preying on molting aquatic fowl and their nestlings and venturing offshore in pursuit of sea mammals. In the same season a harvest of such wild vegetable foods as gooseberries, blueberries, squashberries and bakeapples was available to be eaten when picked or to be stored for future consumption. With the first snow the time came for the bands to travel inland again for the late-fall caribou hunt.

The detailed evidence of technological versatility provided by the grave goods at Port au Choix, including the artifacts of stone, bone, antler and ivory and the tools for working wood, bark and skin, is clear testimony to the material success of the Maritime Archaic Tradition in adapting to life in northeastern North America. The grave goods also offer us tantalizing glimpses of the tradition's magico-religious beliefs. In contemporary Indian and Eskimo practice, for example, a row or two of caribou incisors stitched to one's clothing ensures success in hunting these animals. Bear teeth bring strength and strike fear into the heart of an enemy. Dog teeth make the wearer a good fighter. Seal teeth bring success in sealing; seal claws give one strong arms. The heads or feet of gulls, mergansers and other diving birds bring luck in fishing for salmon.

One could cite a dozen more such beliefs of today, and it is tempting to see the charms and amulets of Port au Choix in a similar light. Moreover, if the presence in a single grave of both the image and the physical remains of a single animal species is more than a matter of chance, the use of amulets may have gone beyond hunting magic to embrace the concept of a personal guardian spirit or totem animal. Possibly the guardian was believed to be instrumental in conveying the dead individual and his buried goods to another world. Although the precise imagery must always remain obscure, the red ocher in the graves, and perhaps the crystals and pebbles too, may have been additional symbols of existence in an afterlife. The east-facing burial ground may be related to symbolism involving the rising sun. Lavish grave goods, including "killed" items, ocher, east-facing cemeteries—all this may be the first manifestation in the northeast of what William A. Ritchie of the New York State Museum calls the "basic core of religiosity" noted at many North American prehistoric sites of the following millenniums.

How did the Maritime Archaic Tradition evolve? Its artifacts resemble those of several Eskimo cultures that existed in Newfoundland and the Canadian Arctic at a somewhat later date. The similarities, however, are only the product of the tendency toward convergence in the face of the same environmental conditions; they are not the result of a racial relationship. Nor is there any direct relationship between the Maritime Archaic Tradition and the now extinct Beothuk ("Red Indians"), the aboriginal people of Newfoundland at the time the Europeans arrived. On the other hand, when comparisons are made with the Archaic Indian cultures in the forested belt of the Great Lakes region, significant similarities emerge.

These cultures to the west are the ones that Ritchie places in the Laurentian Archaic Tradition. Their relationship to the Maritime Archaic Tradition is unmistakable. The Morrison Island skulls, which closely resemble those from Port au Choix, were associated with artifacts of the Laurentian Tradition. A number of basic tools, including gouges and points made by grinding slate, are also common both to the Laurentian and to the Maritime Archaic. It appears likely that a forest-adapted population of hunters with a Laurentian-related material culture moved from the Great Lakes region along the St. Lawrence valley to the sea, where they quickly learned to exploit the animal resources of a new domain. Thereafter, with a culture economically, technologically and socially adapted to the maritime environment, the Archaic Indian hunters flourished along the coast from Maine to Labrador for at least 1,000 years.



LAND AND SEA WEAPONS used by the Indians of the Archaic Maritime Tradition include, in addition to toggling harpoon heads, simple barbed ones (*top*). The ornamented bone "dagger" (*second from top*) could have served on land or at sea and may have been attached to a wooden shaft rather than being held in the hand. The square-barbed leister point (*third from top*) is made of bone and evidently formed part of a fish spear. Why the bone "bayonet" point (*bottom*) was made with an exaggerated single barb is not known.

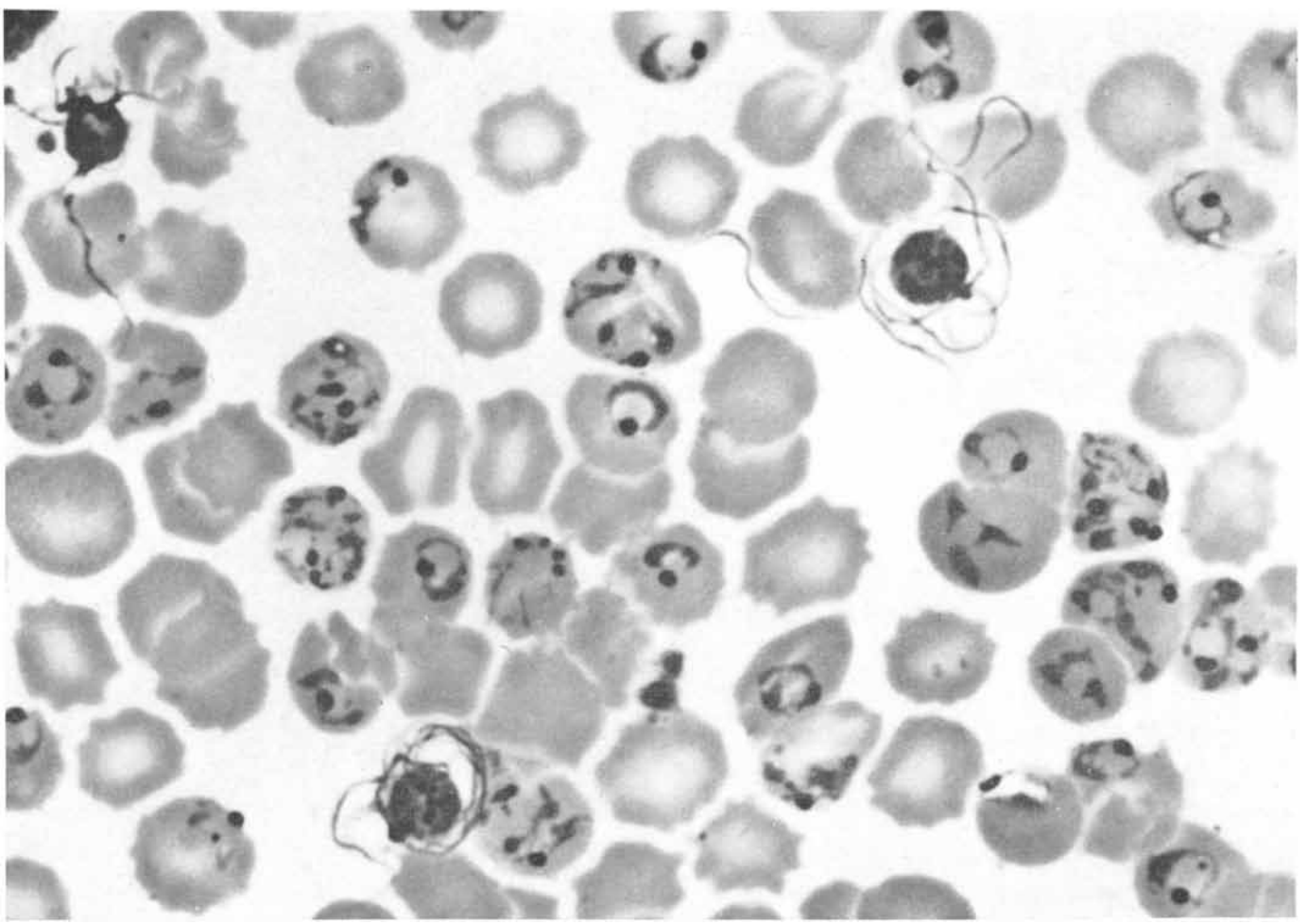
The Clock of the Malaria Parasite

The life cycle of the protozoans that cause malaria is carefully timed so that they are most ready for transfer from victim to mosquito during the hours of darkness, when the mosquitoes are most likely to be feeding

by Frank Hawking

A remarkable feature of malaria is the precise timing of its recurrent attacks, which are always at some multiple of 24 hours. Since the cause of malaria is a small protozoan (single-celled) parasite that grows in the red blood corpuscle, the attacks show that this rather simple organism has a highly accurate biological clock. As a result of recent investigations it is now possible to explain why the parasite needs a clock and (a question that was somewhat more difficult to answer) how the organism determines what the time is, so as to adjust its rhythm accordingly. The findings are of both medical and biological interest.

Even in the early days of medicine, as far back as the ancient Greeks, malaria was recognized as being different from all other fevers because of its regularly recurring attacks, which in the most



MALARIA PARASITES are seen in this photomicrograph at the stage in their sexual cycle that occurs after a mosquito has drawn blood from a diseased host. The organism is *Plasmodium knowlesi*, a malaria parasite of monkeys. Present within many of the monkey's red cells are male and female parasites at the gametocyte stage of development. Some males are exflagellating, an act that transforms one male into several flagella that, like spermatozoa, swim off to fertilize the female gametocytes. Two exflagellating males are visible in the micrograph at top left and bottom left; near a third, at top right, are two of the free-swimming flagella.

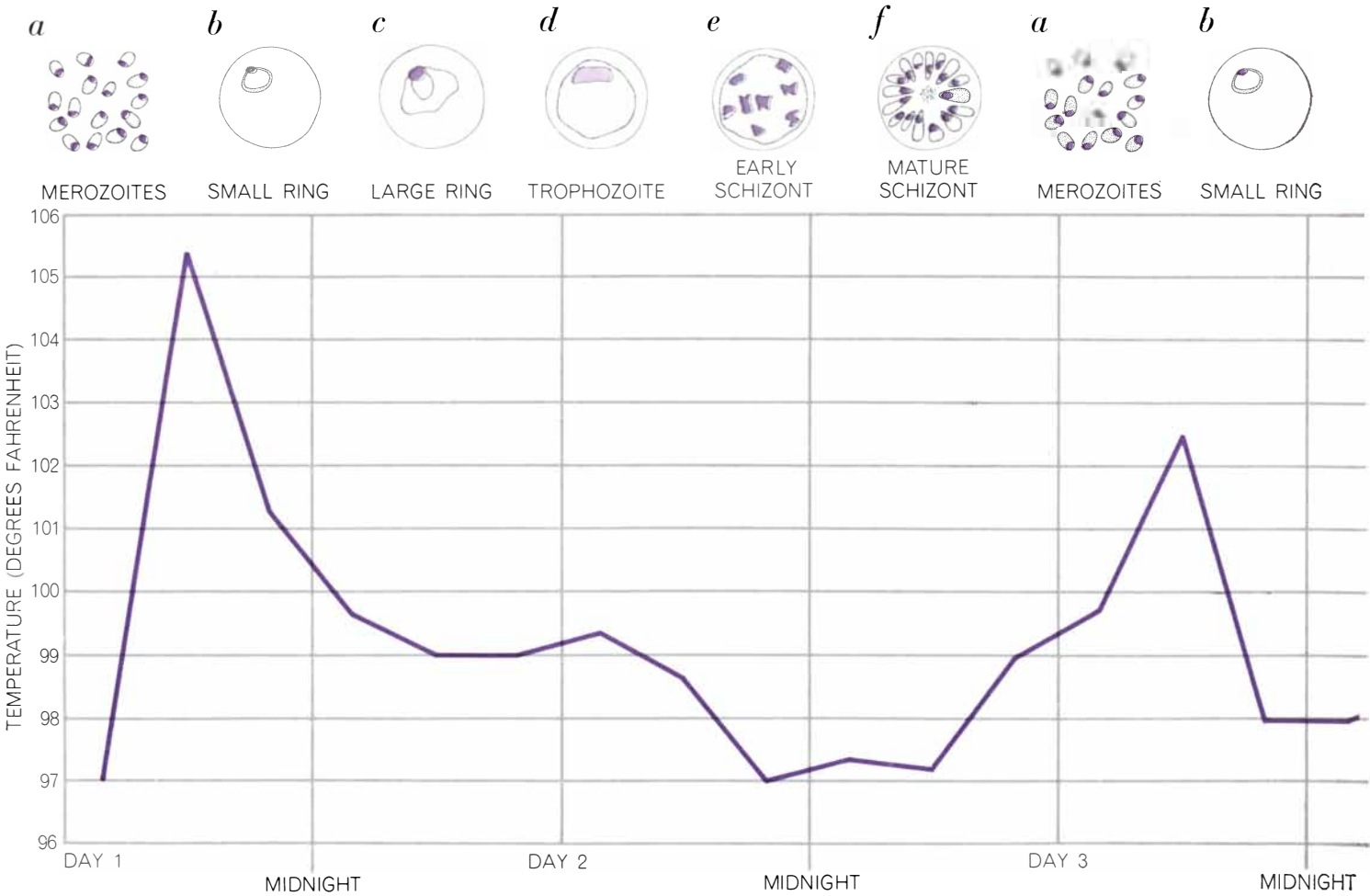
common form of the disease come every other day. On the day of an attack the patient feels all right until about noon. Then he begins to experience sensations of coldness and to shiver, although to observers his skin feels hot; he also suffers from headache and back pain. The symptoms grow worse until late evening, when the patient begins to feel extremely hot and to sweat profusely, although it is at this time that his measured body temperature starts going down and the other symptoms abate rapidly. After a few more hours the attack passes and, apart from exhaustion, the patient feels more or less normal. On the following day he is all right, but on the third day at about noon all the symptoms begin to reappear. This is the typical tertian fever or ague. (Sometimes the attacks come every fourth day and the fever is called quartan.)

In 1880 the French army physician Charles Louis Alphonse Laveran, working in Algeria, discovered the causative parasite: a member of the genus *Plasmodium*, namely *P. vivax*. It took a number of years to work out the life cycle of the parasite, but eventually it was shown that the organism starts as a small ring-shaped structure in the red blood corpuscle [*see illustration below*]. The small ring grows into a large ring and then into a large solid parasite (called a trophozoite) that fills the corpuscle. Next the nucleus of the parasite divides successively into two, four, eight and 16 pieces, which collect enough cytoplasm around themselves to form 16 new small parasites called merozoites. (This process of nuclear division is called schizogony, and the forms with several nuclei are called schizonts.)

The merozoites now break out of the much damaged corpuscle and invade new blood corpuscles to form new small rings; thus the cycle begins again. The entire cycle takes 48 hours, and all the parasites keep approximately in step. The time when the schizonts break up and liberate new parasites into the blood is the time when the patient begins to have his attack. Presumably the fever and the other symptoms are due to the release of plasmodial waste products and other foreign substances into the patient's bloodstream.

The life cycle of *P. vivax*, which is the typical malaria agent of the more temperate zones, gives rise to tertian fevers every 48 hours. The quartan type of fever was found to be caused by a slightly different parasite, *P. malariae;* it has a cycle of 72 hours. There is also the parasite of the Tropics, *P. falciparum*, which has a 48-hour cycle like *P. vivax* but is more virulent and also differs from *P. vivax* in other ways. Other malaria parasites have been discovered in birds, monkeys and African rodents. Most of these parasites show similar regular cycles of 24, 48 or 72 hours.

The discovery of the cycle in the parasite seemed to explain the regular tertian (or quartan) course of the fever, which soon became so well established as a characteristic of malaria that it ceased to attract any further attention. People thought it quite commonplace that each individual parasite should take



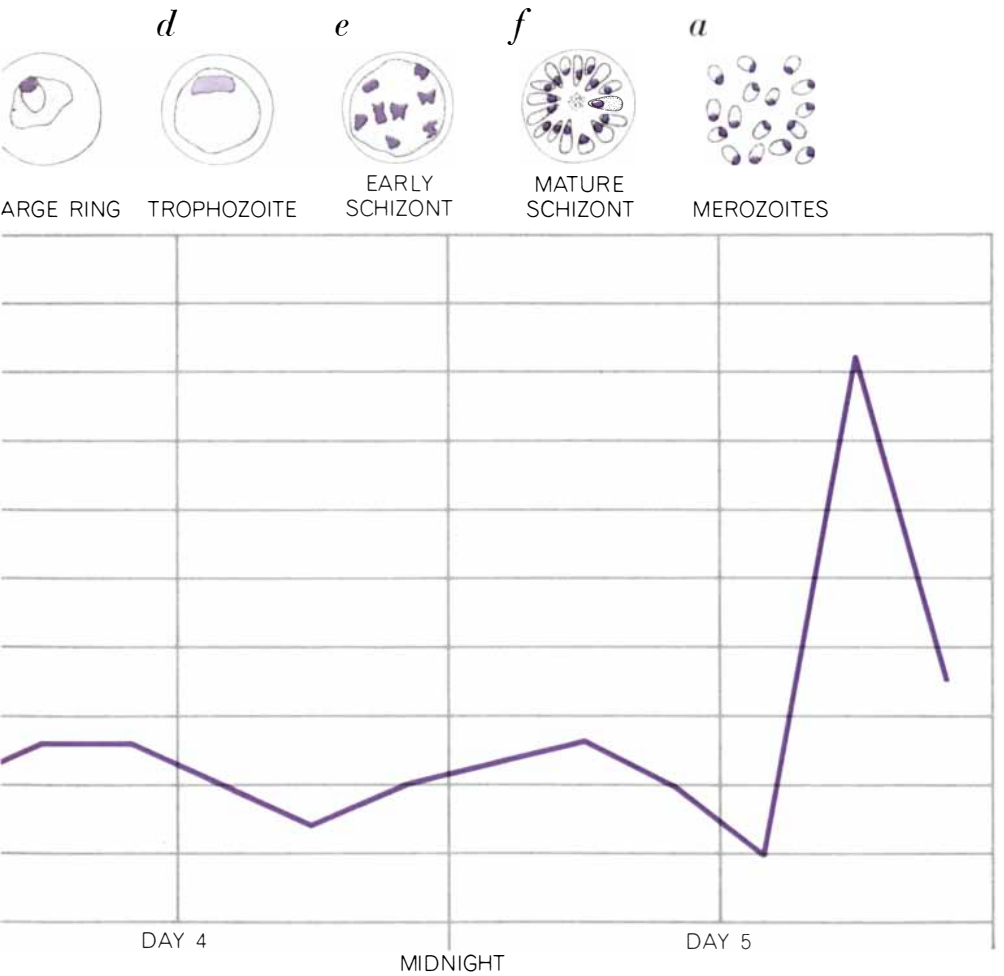
RECURRENT FEVER of the malaria victim is the result of the asexual reproductive cycle of the parasite that lives in the bloodstream of the victim. The variation in the temperature of a human who is host to *Plasmodium vivax* shows an onset of fever at about 4:00 P.M. every other day; this is the "tertian ague," an illness known since classical times. Above the fever chart are shown stages in the 48-hour life cycle of the parasite; the onset of fever coincides with the moment of reproduction. Many of the host's red cor-

exactly 48 hours for its development from one cell division to the next. As a result they overlooked three remarkable phenomena: (1) the periods of malaria parasites—24, 48 or 72 hours—were all simple multiples of 24 hours; (2) all the parasites behaved synchronously, keeping in step so that they all came to schizogony at the same time, and (3) cell division always took place at an hour of the day constant for each species of malaria parasite (usually about midday for malaria parasites of men and monkeys but in early morning or late evening for some malaria parasites of birds).

At the National Institute for Medical Research in London, Kenneth Gammage, Michael J. Worms and I became interested in the matter because we had been working for many years on the 24-hour cycle of microfilariae. These are small larval worms that circulate in the blood of many patients in the Tropics and give rise to the grotesque deformities known as elephantiasis. During the daylight hours the worms accumulate in the small arteries of the lungs, but at nightfall they swarm out into the peripheral blood.

It seemed to us that the synchronous behavior of malaria parasites, all coming to cell division at the same time every 24 hours or a multiple thereof, was remarkably like the behavior of microfilariae, all entering or leaving the peripheral blood at the same time every 24 hours. Microfilariae are carried from one patient to another by mosquitoes, which suck blood mostly at night, and the swarming of the microfilariae in the peripheral blood is arranged to coincide with this time of sucking blood. The biological purpose of the cycle of the microfilariae is clearly to help them encounter mosquitoes and so get transmitted to new patients.

Since malaria is also carried from one person to another by mosquitoes, it seemed to us that the periodic behavior of the malaria parasites in the blood was similarly designed somehow or other to facilitate transmission by mosquitoes. Unfortunately this hypothesis encountered a number of difficulties arising from the then existing orthodox conceptions of malaria parasites. The major source of the difficulty will become clear when I describe another peculiarity in the complex development of the malaria parasite.



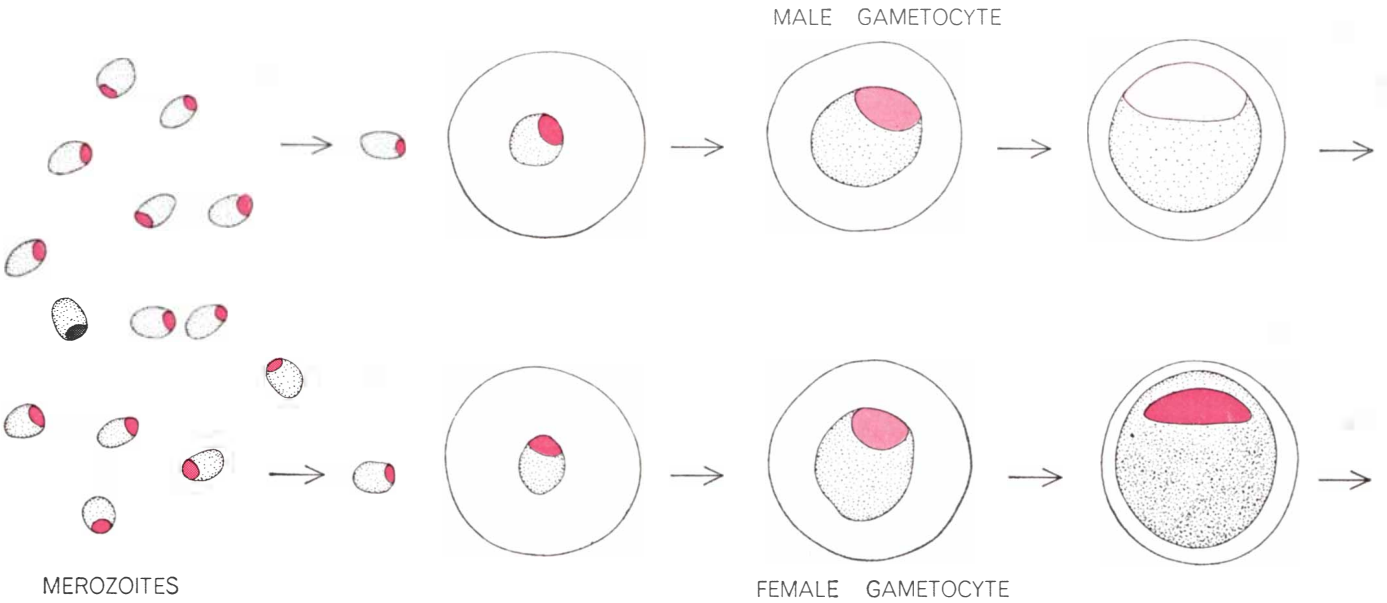
puscles harbor mature schizonts, each containing 16 merozoites. When the corpuscles rupture (*a*), releasing the merozoites, this also releases waste products that cause fever. Each merozoite then invades a fresh corpuscle and passes through small-ring (*b*), large-ring (*c*), trophozoite (*d*) and early schizont (*e*) stages. The cycle repeats at schizont maturity (*f*).

The regular synchronous cycle of the parasite, going round and round every 48 hours in the red blood corpuscles, is the asexual cycle. It serves only for the multiplication of the parasite inside the body. In order to provide forms that can develop in mosquitoes and so be transmitted to new patients, other forms called gametocytes (cells that marry) are budded off during cell division. Some gametocytes are male and some are female. When they are sucked up by a mosquito, they develop a bit further, and then the male and female combine to produce a series of forms that ultimately wind up as tiny parasites (called sporozoites) in the salivary glands of the mosquito. When the mosquito bites another man, the sporozoites are injected under his skin along with the saliva.

The difficulty in all of this for our hypothesis (that the timing of the synchronous asexual cycle had something to do with the fact that mosquitoes bite mainly by night) lay in the fact that the asexual forms, which show the cycle, do not develop in mosquitoes, whereas the gametocytes, which do develop in mosquitoes, were believed not to show any cycle. The conventional view was that gametocytes, once they had developed in the human body, circulated in the bloodstream indefinitely, waiting to be taken up by a mosquito. We, however, were so attracted by the hypothesis of a connection between the asexual cycle of the malaria parasite and transmission of the parasite by night-biting mosquitoes that we decided to investigate the matter further.

For these experiments we mostly used rhesus monkeys. Human malaria parasites cannot readily be transmitted to animals, but monkeys have many kinds of malaria parasites that are suitable for experiment. The one we used most is *P. knowlesi,* which has a 24-hour asexual cycle with cell division at about midday. We also used *P. cynomolgi,* which has a 48-hour cycle with cell division at midday and closely resembles *P. vivax* of man, and *P. cathemerium,* which infects canaries and has a 24-hour cycle with cell division at 10:00 P.M.

What we wanted to ascertain first was whether the gametocytes really persisted unchanged in the blood for many days or whether their mature stages were in



SEXUAL CYCLE of the malaria parasite has its inception when certain merozoites undergo a development within the red corpuscle different from that shown in the illustration on the preceding two pages. Instead of appearing as a small colored nucleus linked to a thin ring of cytoplasm enclosing a large vacuole (as in *b* and *c* in that illustration), they mature within the corpuscle to form a large, solid body that, like the trophozoite (*d* of the asexual cycle), has no vacuole. These are gametocytes; the males have large, pale nuclei and cytoplasm that is pale blue when stained. Females have small, colored nuclei and cytoplasm that stains a purplish blue. It was once believed that after parasites had reached the gametocyte stage they would circulate indefinitely in the host's bloodstream.

fact quite brief. To this end we allowed batches of mosquitoes (*Anopheles stephensi*) to feed on a lightly infected monkey. We used a new batch every four hours day and night for 48 hours, so that each group of mosquitoes took up gametocytes at a different time.

We then stored the mosquitoes for six days at a temperature of 28 degrees Celsius (82.4 degrees Fahrenheit). This was to give the gametocytes time to develop in the mosquito. Each pair of gametocytes that have combined grows into a small cyst, called an oöcyst, in the stomach of the mosquito. Oöcysts are easy to see under a low-power microscope, and by counting them we could measure how many ripe gametocytes had been in the monkey's blood at the hour when the mosquito sucked blood.

The results showed that there were few oöcysts in mosquitoes fed during the daytime (noon to 5:00 P.M.) and that the number increased rapidly during the evening, with a peak just after midnight. Thereafter the number fell quickly, being low again at 5:00 A.M. and 9:00 A.M. During the night of the second day the number of ripe gametocytes showed a similar big rise and fall. Similar results were obtained in our experiments with *P. cynomolgi* and *P. cathemerium*.

Testing the maturity of gametocytes by feeding batches of mosquitoes at four-hour intervals, keeping the mosquitoes for six days and then dissecting them to count the number of oöcysts on the stomach is a slow and cumbersome procedure. There is a quicker and simpler way of ascertaining whether or not gametocytes, at least the male ones, are mature. It depends on observing a process known as exflagellation, which means the putting out of the whiplike structures called flagella.

When blood containing ripe gametocytes is taken out of the body by a mosquito, its temperature falls from 37 degrees C. (body heat) to the temperature of the outside air. This change stimulates a further development in the male gametocyte. The nucleus goes into division twice and produces four small nuclei. They begin to dance about inside the parasite. Soon a long structure like a whip sprouts from each nucleus, and the parasite takes on the appearance of an octopus with four arms. The arms lash about vigorously for a while. Then each arm and its attached nucleus breaks off and swims away to look for a female gametocyte. (In effect the male gametocyte has transformed itself into the equivalent of four spermatozoa.) When one of these transformed gametocytes finds a female gametocyte, it penetrates it and fertilizes the nucleus, thus beginning a new generation of parasites. The term exflagellation applies to the part of the process up to the time when the transformed male gametocytes break off and swim away. The course of development up to that point of spermatozoa-like free swimming covers a period of about 20 minutes.

Exflagellation can easily be observed in a drop of blood under a microscope. It is most dramatic. It was seeing this active wriggling of the flagella that led Laveran to discover the malaria parasite. For our purpose we took blood every four hours, put it in a thin film on a slide and stored the slide in a moist chamber for 20 minutes. In this way the blood did not dry and the male gametocytes that were ripe had time to put out flagella. After 20 minutes we dried the slides, fixed them with alcohol and stained them with a Giemsa stain, which colors the nucleus red and the cytoplasm blue. Then at our leisure we could examine them under the microscope and count the number of male gametocytes that had been exflagellating per 100,000 red blood corpuscles.

In a typical experiment with a monkey heavily infected with *P. knowlesi* few gametocytes were exflagellating during the daytime, up to 6:00 P.M. Then the number rose rapidly until it reached a plateau at about 11:00 P.M. It stayed at the plateau level until about 5:00 A.M. and then fell rapidly. We obtained similar results with *P. cynomolgi* and *P. cathemerium*.

Clearly there are many more ripe gametocytes—ripe in the sense that they are ready to develop in mosquitoes—during the night than there are during the day. Judging by the duration of the plateau, the time a male gametocyte re-

mains ripe (as gauged by its readiness to exflagellate) is only about six hours. It seems unlikely that a gametocyte that has reached its ripeness for infecting mosquitoes and then has ceased to be infective will ever become infective again. It is much more probable that each gametocyte becomes mature and infective for a brief period (from six to 12 hours) and then degenerates, and that it is replaced the next night by a new gametocyte. In other words, the nightly increase in the number of infective gametocytes is due to a succession of short-lived individuals and not to a waxing and waning of the same gametocytes.

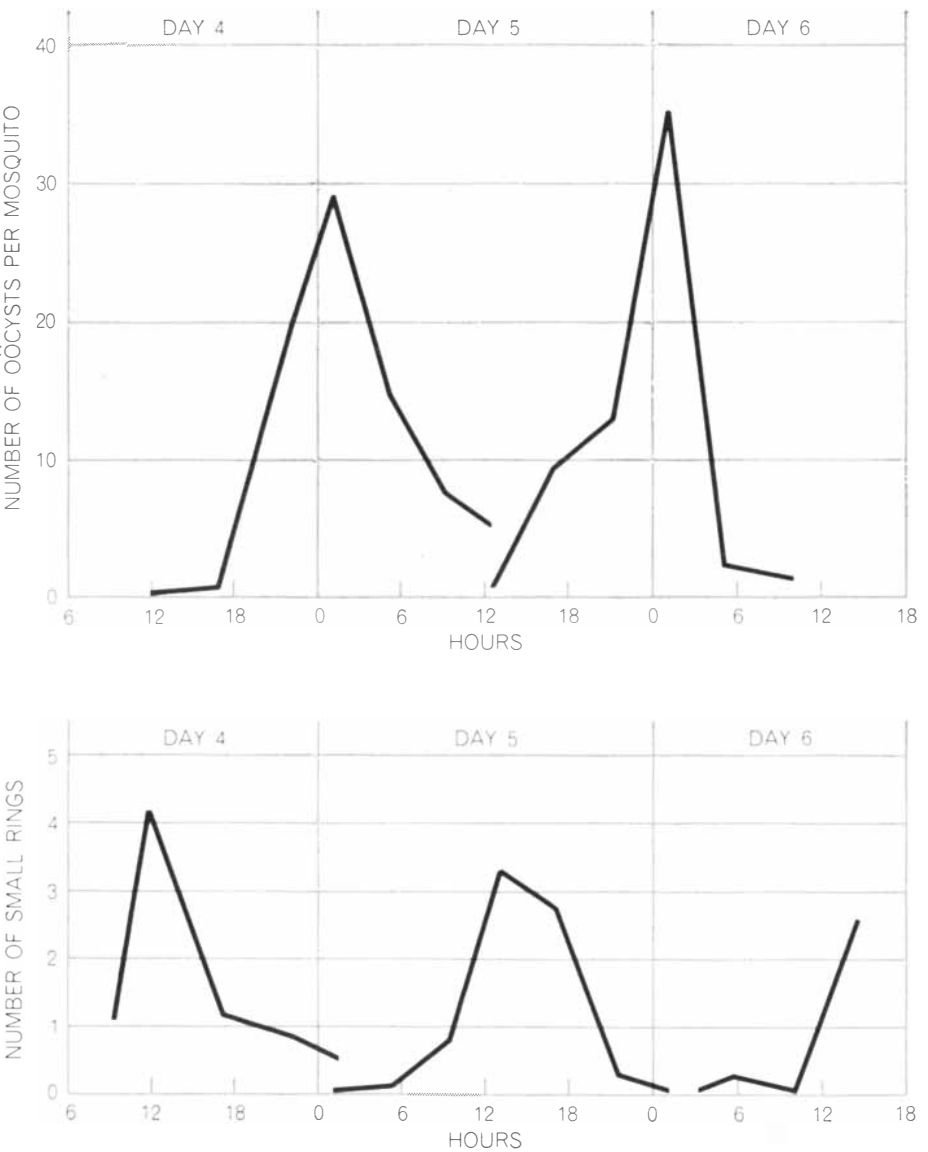
In view of these findings we thought it well to take a closer look at the appearance of gametocytes throughout a 24-hour period to see if they were really as unchanging in appearance as was commonly supposed. Not many investigators have studied the appearance of gametocytes around the clock. Most malariologists take a daily blood film at the same time each day; the few who have taken films every four hours have been studying the asexual cycle, not the gametocytes.

My colleague Worms took the blood films we had made every four hours during the study of exflagellation and reexamined the gametocytes in them carefully. He found that during the first 18 hours the early gametocytes are indistinguishable from the early asexual forms, but after that he was able to see several stages of gametocyte development. He divided them into five types [*see illustration on page 130*].

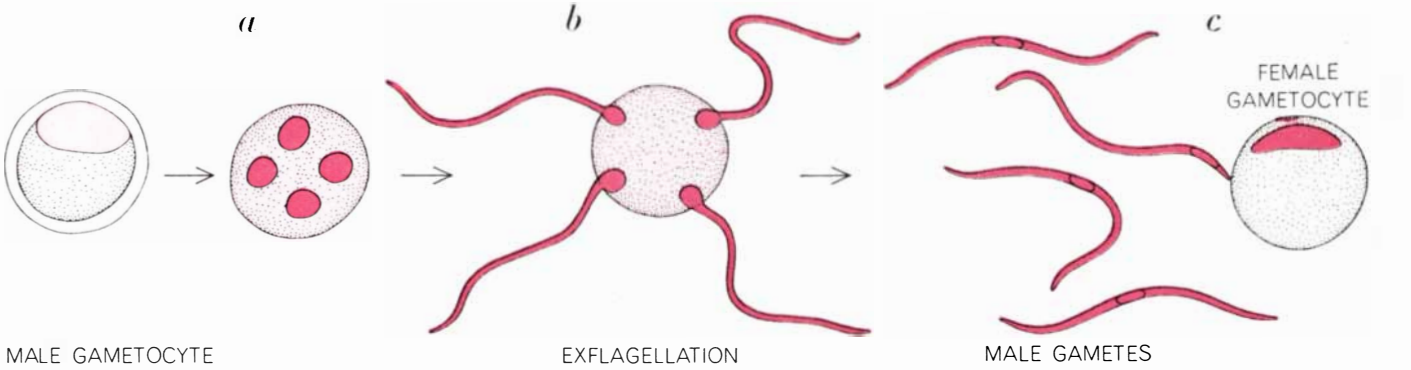
In order to determine when these respective types were most common, and to make sure that they really formed a sequence of development, Worms counted the number of the different types per 100,000 red blood corpuscles at the different times. The results can be summarized as follows: Type I is most numerous at 9:00 A.M., Type II at 1:00 P.M. and Type III at 3:00 P.M. Type IV, which is the mature type for both male and female gametocytes, reaches a high peak at midnight. Type V, which is the degenerate kind, is never very numerous (presumably these gametocytes disappear quickly from the blood), but it is more common from 5:00 A.M. to 9:00 A.M. than at other times. This morphological cycle of development quickly followed by degeneration was found in all the parasites we studied: *P. knowlesi, P. cynomolgi* and *P. cathemerium.*

Thus through these experiments we found that the gametocytes of the malaria parasites are ripe and able to infect mosquitoes for a period of about six hours. Therefore it is obviously important to the parasite, if it is going to get itself transmitted, that as many gametocytes as possible should manage to have their brief period of infectivity at the time during the 24 hours when mosquitoes usually bite and suck blood, which is nighttime. The problem confronting the malaria parasite has been how to time the gametocytes to match the mosquitoes. It has solved the problem by the device of the synchronous asexual cycle in which all the asexual parasites keep in step.

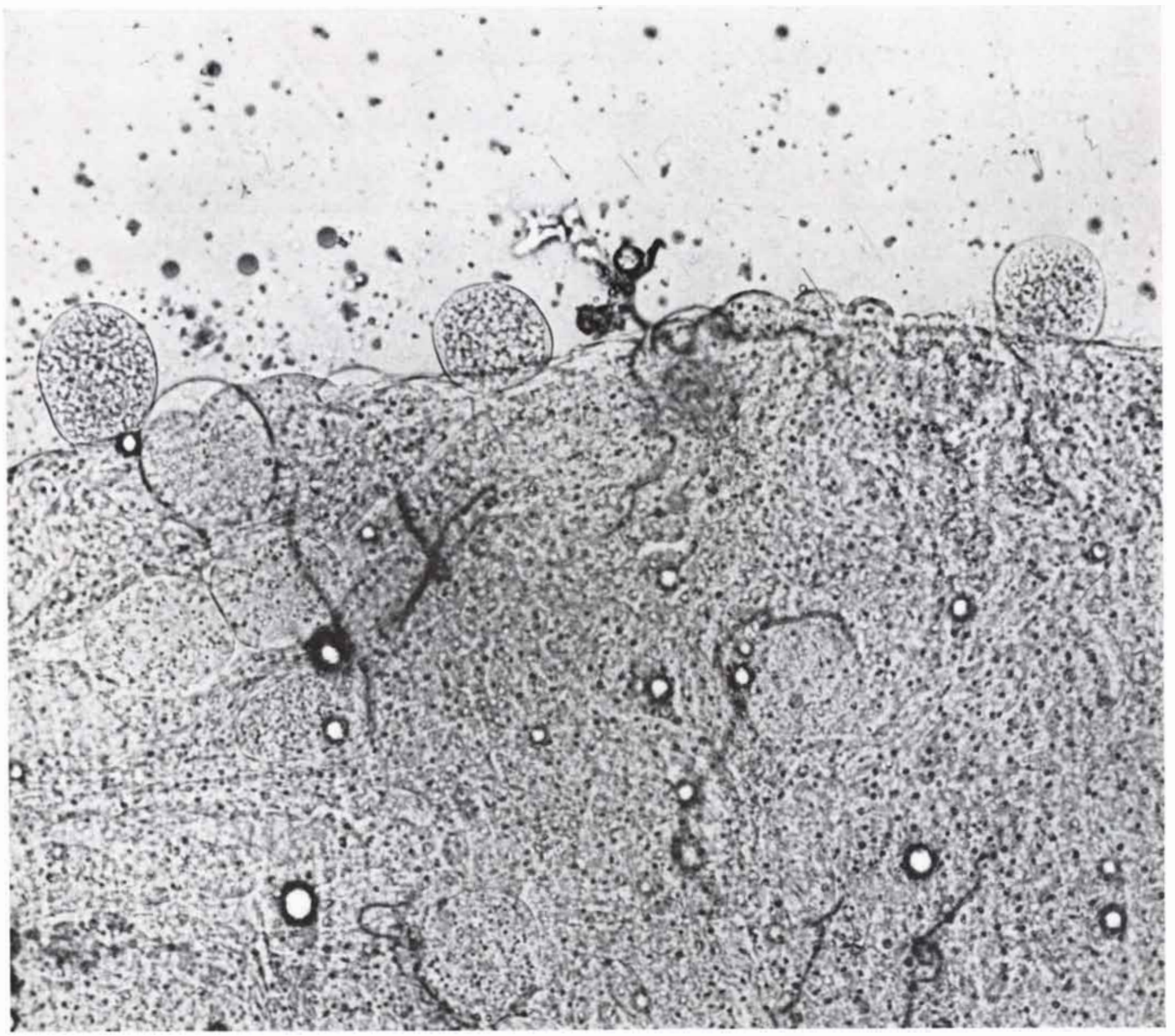
Apparently the system works as follows. The asexual forms all come to cell division at the same time. During this cell division some new small parasites (merozoites) are marked by their chromosomes to become gametocytes. In *P. knowlesi,* which has a 24-hour cycle, the gametocytes require between 30 and 35 hours to grow and become infective for mosquitoes. Accordingly the cell division



GAMETOCYTE LONGEVITY in a victim's bloodstream was tested in a monkey harboring *P. knowlesi*. Successive batches of mosquitoes were allowed to feed on the monkey at four-hour intervals, and the number of gametocytes in its blood at each feeding was later estimated from the number of oöcysts found in the mosquitoes (*top graph*). The greatest number were found in mosquitoes that had fed at night, the peak occurring near midnight. The number of oöcysts that developed as a result of daylight feedings was substantially lower, indicating a wide fluctuation in gametocyte numbers over a 24 hour period. This corresponded with the 24-hour asexual cycle of *P. knowlesi*, as shown by changes in the number of small-ring parasites in the monkey's red blood corpuscles (*bottom graph*), but was 12 hours out of phase. The total number of sexual and asexual parasites increases greatly with each successive cell division. To facilitate comparison here the vertical scales for the successive 24-hour periods have been reduced. In the count of small rings, for example, the numbers are given per 1,000 red cells in Day 4, per 100 cells in Day 5 and per 20 in Day 6.



EXFLAGELLATION, the key event in the sexual cycle of the malaria parasite, is triggered by the cooling of the blood as it is transferred from host to mosquito. The boundary between the nucleus and the cytoplasm in the male gametocyte disappears and four (or sometimes eight) solid bodies develop in the nucleus (*a*). The bodies soon put out long, whiplike flagella (*b*), which lash about wildly. Each flagellum, with chromatin attached, breaks free to seek and fertilize a female gametocyte (*c*); the result is an oöcyst.



OÖCYST STAGE in the sexual reproductive cycle of *P. knowlesi* is seen in this photomicrograph of the stomach (*dark area*) of a mosquito that has fed on an infected monkey. Three oöcysts are clearly outlined along the stomach margin; four others are less readily apparent. All are about .06 millimeter in diameter. Most of the parasite's sexual cycle takes place within the mosquito, beginning with union between gametocytes. Each union produces an oöcyst, from which a new generation of parasites arises. When the new generation reaches the sporozoite stage, living in the insect's salivary glands, the mosquito's bite injects them into a new host.

of the asexual forms of *P. knowlesi* is timed to occur approximately 32 hours before nightfall, that is, about noon of the preceding day. The baby gametocytes thus started off at midday are ripe and infective on the night of the next day, when the mosquitoes bite. The exact time intervals are different for the different species of parasites in order to achieve the same final result.

Therefore we see that, thanks to the timing of the asexual cycle, the gametocytes ripen in the blood just when the mosquitoes habitually bite. In this way transmission of the parasite is greatly facilitated. This matching of the gametocytes to the mosquitoes is the biological purpose of the synchronous asexual cycle.

With the biological purpose of the cycle established, the next question was how the cycle is managed. How are all the asexual parasites controlled so that they keep in step and all come to cell division at the right time? Although the rhythm is a 24-hour one, it cannot be directly related to the alternation of light and darkness, because the parasites are enclosed in the dark, opaque host. On the other hand, if a man, a monkey or a bird infected with malaria is made to sleep by day and wake by night, so that the normal 24-hour rhythm becomes inverted, the cycle of the malaria parasites also becomes inverted. Clearly the time-marker for the malaria parasites must be derived from some component of the 24-hour cycle of the host.

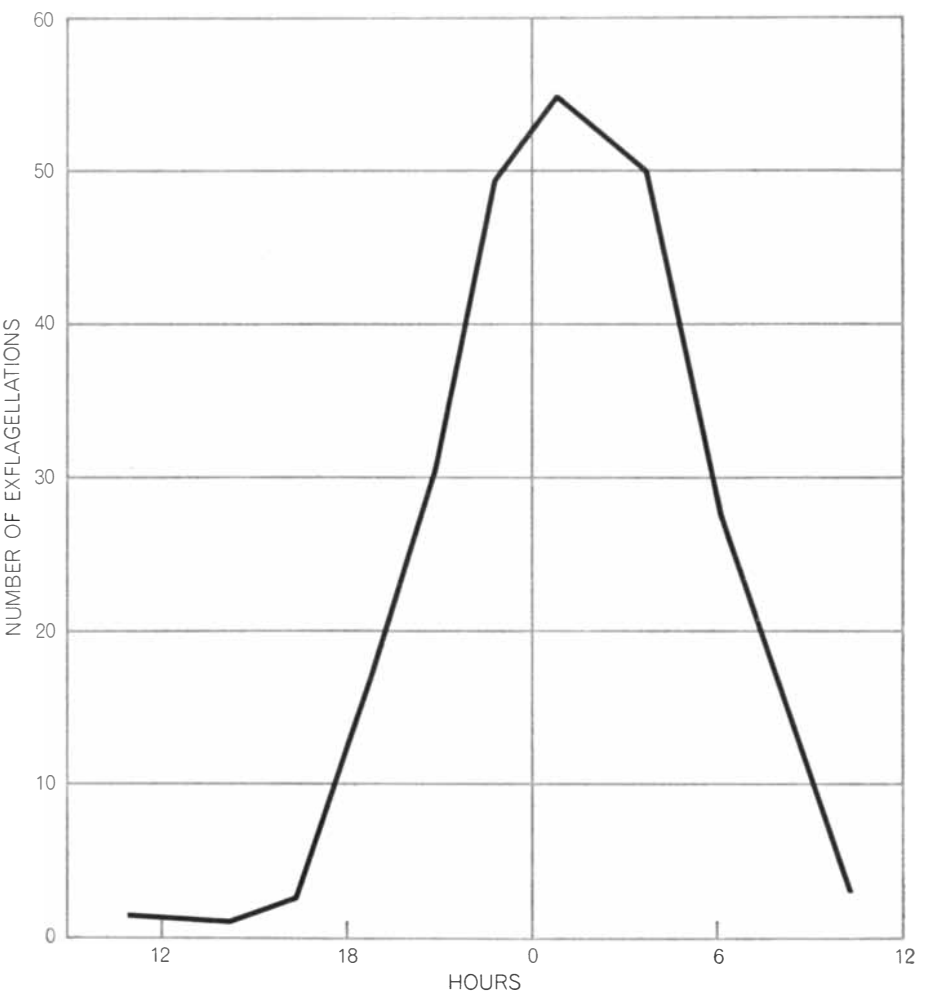
As is well known, almost all the systems of the body show 24-hour cycles. The temperature goes up by day and down by night; the blood becomes less acid by day and more acid by night; the adrenal glands secrete smaller amounts of corticosteroids by night than by day; the kidney excretes less urine during the night than during the day, and so on. The problem was: Which of these many cycles acts as a time-marker for malaria parasites?

From the start we were attracted by the hope that the temperature cycle would prove to be the time-marker. One of our reasons was that temperature change is one of the most fundamental physical conditions in any biological system. Moreover, we had just found temperature change to be the time-marker for the microfilaria of the parasite *Loa loa* in man and for various other microfilariae in certain monkeys. Most of all, we were appalled at the prospect of having to test hundreds of other possible host cycles in the laborious experimental system required for the study of malaria parasites.

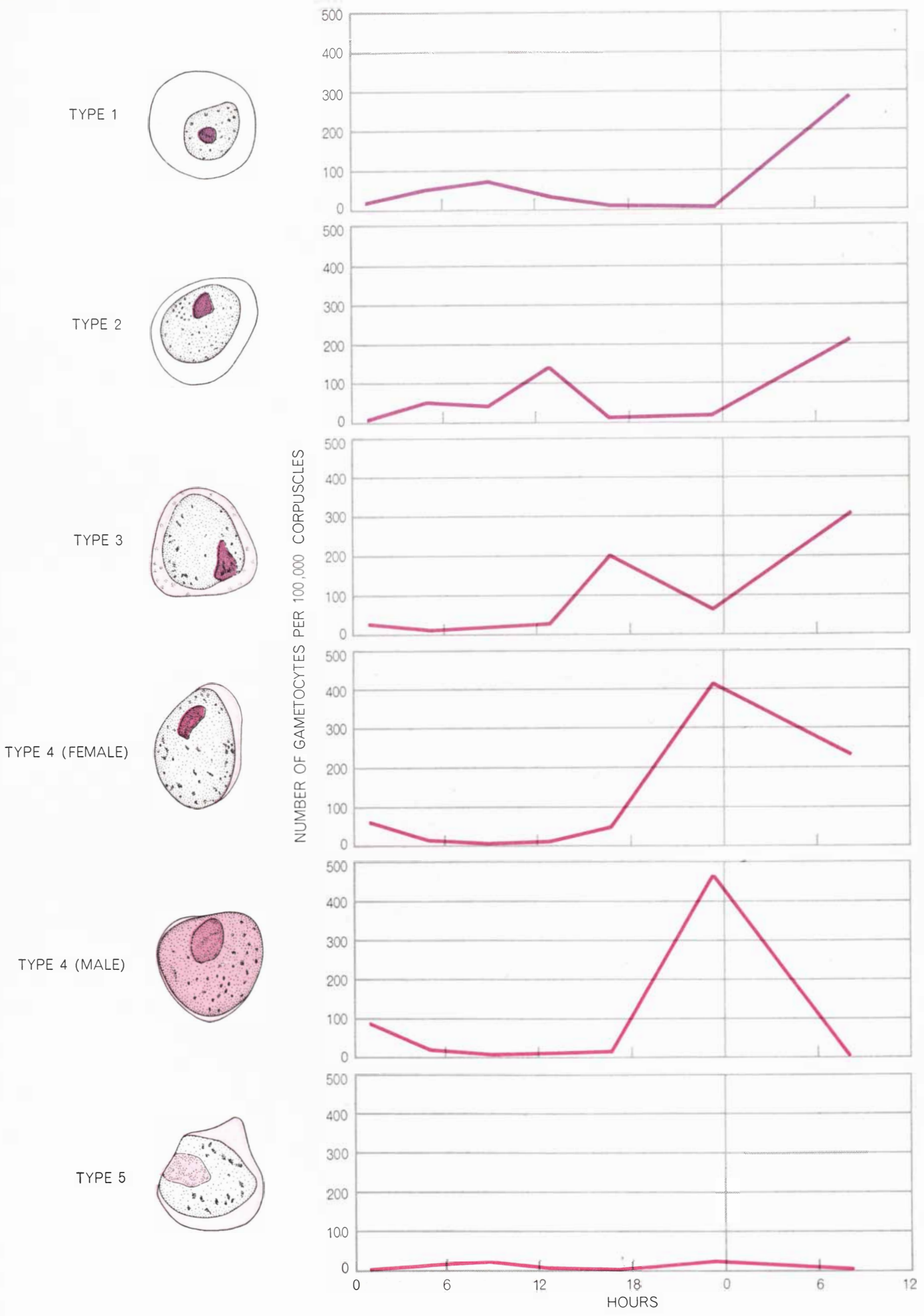
Our first efforts to alter the malaria cycle were made by changing the temperature cycle of a monkey host for a single day. They were unsuccessful (we later found that we had not kept at it long enough), and for two years we abandoned the subject. The temperature hypothesis continued to be attractive, however, and so we started again, this time using chick embryos because it is easy to alter the temperature of eggs merely by moving them from a hot incubator to a cool one and vice versa. The best malaria parasite to grow in incubated eggs is *P. lophurae,* and this is the one we used. It gave us some encouraging but ambiguous results. Later we found that even in partly grown chickens and ducks, which have a good 24-hour cycle, this particular parasite did not show a highly synchronized cycle. Evidently it was not very sensitive to host rhythms and so could not be expected to yield a reliable answer to our problem.

Lured on, however, by the success we had seemed at first to achieve with eggs, we went back to monkeys. Taking two monkeys infected with *P. knowlesi,* we kept them under light anesthetic between 10:00 A.M. and 5:00 P.M. so that we could alter their body temperature. One monkey, which served as a control to ensure that the anesthetic did not affect the malaria parasites, was kept at normal temperature. The other monkey's temperature was lowered from the normal 37.5 degrees C. to about 35 degrees, which we chose as being not too great an exaggeration of the normal nocturnal variation.

The experimental monkey was housed in a room where the lights came on at 7:00 A.M. and went off at 7:00 P.M. Accordingly the animal had three hours of light, during which it could run around and eat, and then it was anesthetized for



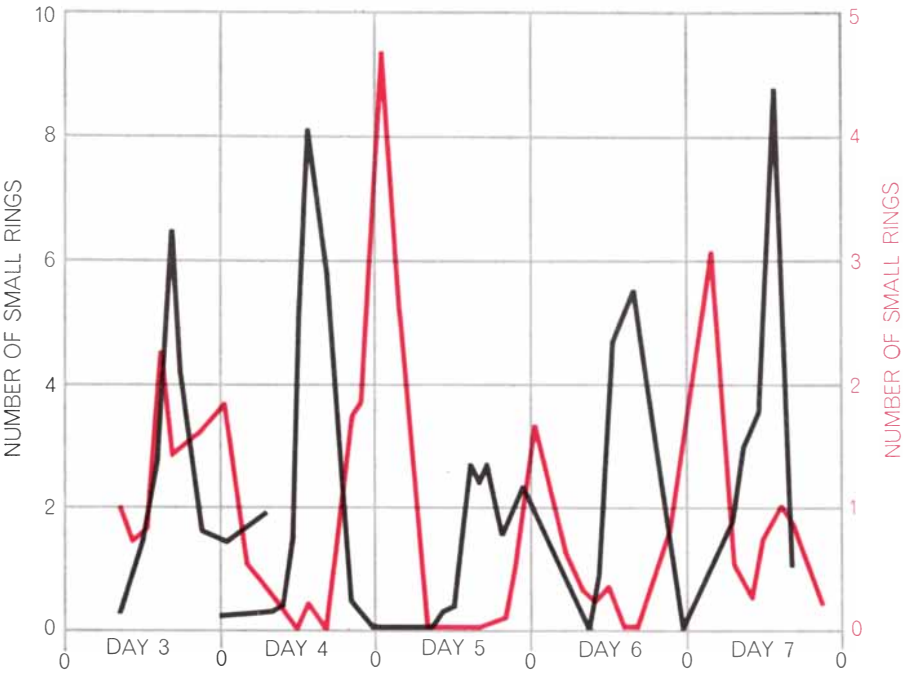
DAILY RISE in the number of ripe gametocytes, the parasites that transform mosquitoes into carriers of malaria, coincides with the daily period of maximum mosquito activity. Successive blood samples were taken from a monkey harboring *P. knowlesi* and the number of exflagellating gametocytes was noted; the maximum came during the hours of darkness.



seven hours. During that period its body temperature was brought below the normal nighttime temperature, thereby inverting the normal temperature cycle. In the evening the monkey had two more hours to run around in the light, and then it had 12 hours of darkness for sleeping. Thus the general 24-hour rhythms of the body were not altered except for the temperature cycle.

To follow what was happening to the cycle of the malaria parasites, we took blood films every four hours. We used them to count the number of early parasites (small rings) around the clock. Since 16 of the small rings are produced every time a mature parasite goes into cell division, the number of them in the blood rises sharply at the time of cell division and then declines as the parasites mature. Therefore by recording on graph paper the count of small rings at different times, it is easy to detect when cell division is occurring.

The results in a typical experiment were that the time of cell division in the control monkey was the same each day (soon after midday), whereas in the cooled monkey cell division beginning on the second day of cooling was delayed until almost midnight. Other experiments with other parasites produced similar results. In short, when the temperature cycle was inverted, the malaria cycle became inverted.



EFFECT OF HOST'S TEMPERATURE on the malaria parasite's asexual cycle was shown by artificially altering the body temperature of one monkey harboring the 24-hour parasite *P. knowlesi* while leaving a similarly infected monkey undisturbed. The body temperature of the control monkey was high during the day and low at night; the parasites it harbored underwent division (*black peaks*), as evidenced by increased numbers of small-ring parasites in its blood samples, at about noon each day. The other monkey was anesthetized and chilled for seven daylight hours, beginning on Day 3. By the next day and thereafter its parasites delayed division (*color peaks*) until after midnight, inverting the normal cycle.

The effects of inverting the temperature cycle can be seen more clearly with *P. cynomolgi*, which produces a milder and longer infection. This parasite has a 48-hour cycle, so that each cycle takes two days. The results in a typical experiment merit close attention. Monkey 380, the control, was kept at normal temperature by day for seven days (3½ cycles) and then observed for 17 more days. The time of cell division varied slightly and irregularly during all 24 days, but it was never earlier than 10:00 A.M. or later than 3:00 P.M.

In Monkey 379, which had its temperature cycle inverted by daytime cooling, the time of cell division was 1:00 P.M. on the first day (at the end of the first cycle) and became 6:00 P.M., 7:00 P.M. and 10:00 P.M. in the second, third and fourth cycles respectively. After 3½ cycles (seven days) the cooling was stopped and the monkey was allowed to revert to its normal temperature cycle. Interestingly enough, cell division thereafter continued to be later by about one hour every cycle until eventually, by the 12th cycle, it was back to its normal time of midday (but a day later).

In this case the malaria parasites had returned to their normal time of cell division—after the temperature cycle returned to normal—by lengthening the cycle from 48 to 49 hours. Sometimes in other monkeys the same effect was produced by shortening the cycle from 48 hours to 47 hours. By means of other experiments, using duck eggs that we placed in hotter or cooler incubators every 12 hours, the same effect could be shown with *P. cathemerium*, which normally grows in thrushes and canaries. Clearly when malaria parasites need to know the time of day, in order to adjust their asexual cycle and make their gametocytes match the biting time of the mosquitoes, they take the temperature cycle of their host and arrange their own cycle accordingly.

Since the rhythm of day and night is so deeply impressed on most animals and plants, it is not surprising that 24-hour rhythms are found in some of their parasites. Only a few of these rhythms have been investigated. Any such study must deal with the questions of what benefit the parasite derives from the cycle and how the parasite tells the time.

The answer to the first question will almost certainly be that somehow the rhythm helps the parasite to get itself transported to new hosts. Thus in the case of malaria the gametocytes are most ripe and able to infect mosquitoes at night, which is when the mosquitoes usually suck blood.

The answer to the second question seems to differ from parasite to parasite. Malaria parasites and some of the less common microfilariae, such as those of *Loa loa*, take their time from the temperature cycle of the host, but the more widespread human filaria, *Wuchereria bancrofti*, takes its time from a small day-to-night variation in oxygen tension in the lungs. Therefore it seems that to answer the second question each parasite—indeed, each species of parasite—must be investigated individually.

GAMETOCYTES change appearance as they mature and pass through five recognizable phases once they can be distinguished from asexual malaria parasites. The five types are present in different numbers at different times. The first (*top*) is most numerous at about 9:00 A.M., the second at about 1:00 P.M. and the third at about 5:00 P.M. The fourth is the mature gametocyte; male and female are illustrated. These are most numerous at about midnight. The fifth, a degenerate type, is never found in large numbers.

MATHEMATICAL GAMES

Elegant triangle theorems not to be found in Euclid

by Martin Gardner

One might suppose that the humble triangle was so thoroughly investigated by ancient Greek geometers that not much significant knowledge of the polygon with the fewest sides and angles could be added in later centuries. This is far from true. The number of theorems about triangles is infinite, of course, but beyond a certain point they become so complex and sterile that no one can call them elegant. George Polya once defined a geometric theorem's degree of elegance as "directly proportional to the number of ideas you see in it and inversely proportional to the effort it takes to see them." Many elegant triangle discoveries have been made in recent centuries that are both beautiful and important but that the reader is unlikely to have come across in elementary plane geometry courses. This month we shall consider only a minute sample of such theorems, emphasizing those that have suggested puzzle problems.

"Ferst," as James Joyce says in the mathematical section of *Finnegans Wake*, "construct ann aquilittoral dryankle Probe loom!" We begin with a triangle, *ABC*, of any shape [*see illustration below*]. On each side an equilateral triangle is constructed outward [*left*] or inward [*center*]. In both cases, when the centers (the intersections of two altitudes) of the three new triangles are joined by straight lines [*shown in color*], we find we have constructed a fourth equilateral triangle. (The theorem is sometimes given in terms of constructing three isosceles triangles with 30-degree base angles, then joining their apexes, but since these apexes coincide with the centers of equilateral triangles, the two theorems are identical.) If the initial triangle is itself equilateral, the inward triangles give a "degenerate" equilateral triangle, a point. It is a lovely theorem, one that holds even when the original triangle has degenerated into a straight line as shown at the right in the illustration. I do not know who first thought of it—it has been attributed to Napoleon—but many different proofs have been printed in recent decades. An unusual proof using only group theory and symmetry operations is given by the Russian mathematician Isaac Moisevitch Yaglom in *Geometric Transformations* (Random House, New Mathematics Library, 1962), page 93.

Another elegant theorem, in which a circle (like the fourth equilateral triangle of the preceding example) seems to emerge from nowhere, is the famous nine-point-circle theorem. It was discovered by two French mathematicians, who published it in 1821. On any given triangle we locate three triplets of points [*see top illustration on opposite page*]:

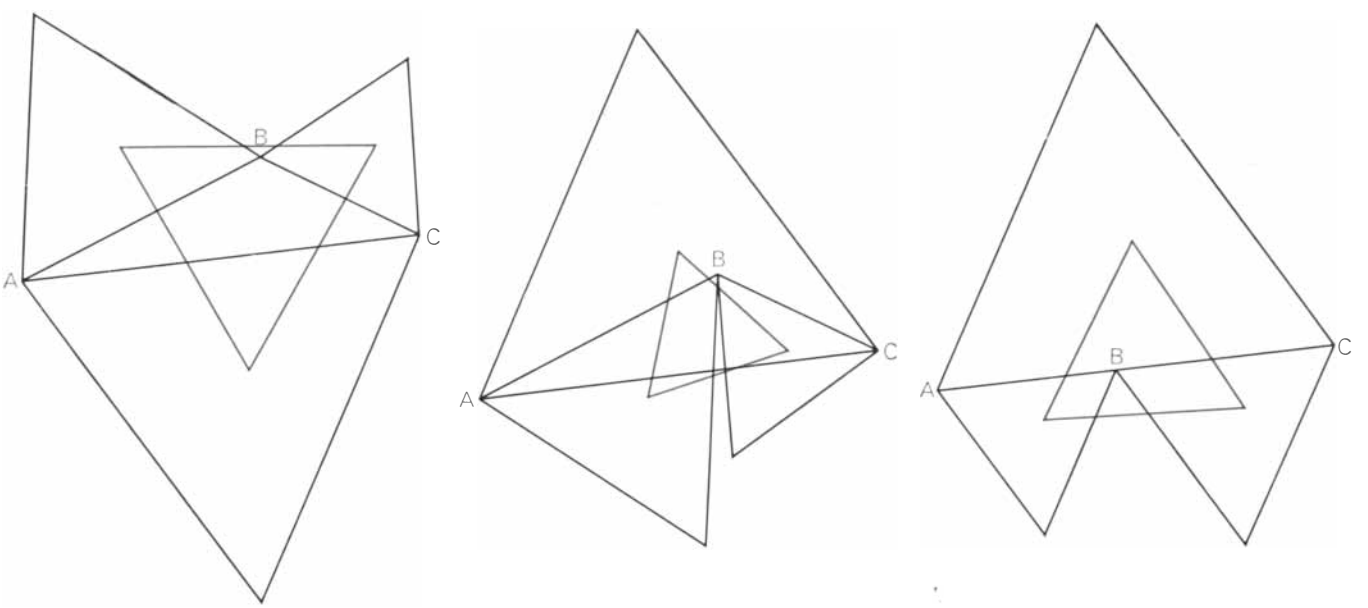
(1) The midpoints (*a*, *b*, *c*) of the three sides.

(2) The feet (*p*, *q*, *r*) of the three altitudes.

(3) The midpoints (*x*, *y*, *z*) of line segments joining each corner to the "orthocenter" (the spot where the three altitudes intersect).

As the illustration shows, those nine points lie on the same circle, a startling theorem that leads to a wealth of other theorems. It is not hard to show, for instance, that the radius of the nine-point circle is exactly half the radius of a circle that circumscribes the original triangle. The fact that the three altitudes of any triangle are concurrent (intersect at the same point) is interesting in itself. It is not in Euclid. Although Archimedes implies it, Proclus, a fifth-century philosopher and geometer, seems to have been the first to state it explicitly.

Three lines joining each midpoint of a side to the opposite vertex are called the triangle's medians [*see bottom illustration on opposite page*]. They too are al-

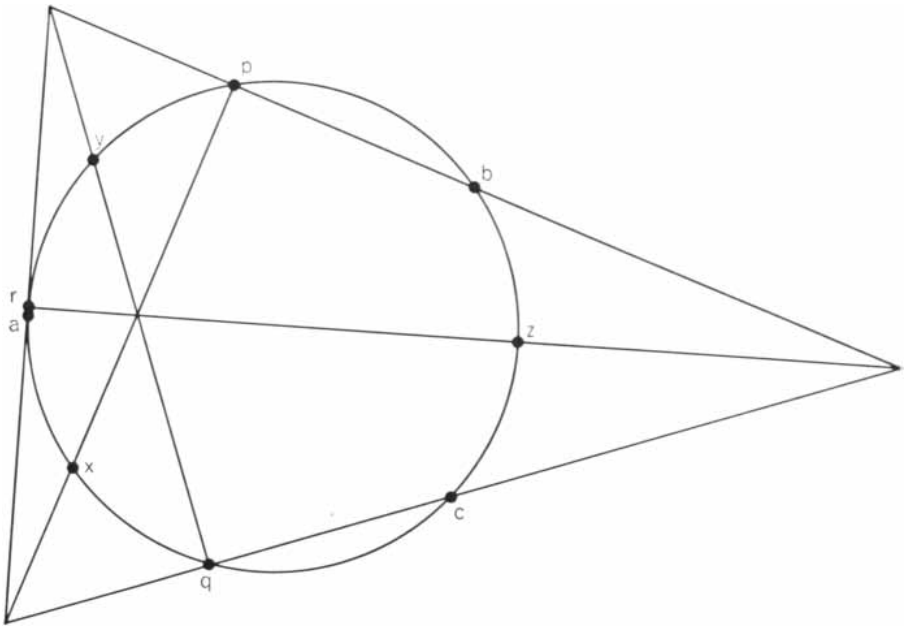


Joining the centers of three equilateral triangles creates a fourth one (color)

ways concurrent, intersecting at what is known as the triangle's centroid. The centroid trisects each median and the three medians carve the triangle into six smaller triangles of equal area. Moreover, the centroid is the triangle's center of gravity, another fact known to Archimedes. Your high school geometry teacher may have demonstrated this by cutting a scalene triangle from cardboard, drawing its medians to find the centroid, then balancing the triangle on a pencil by putting the centroid on the pencil's point.

The median is a special case of a more general line called a "civian" (after a 17th-century Italian mathematician, Giovanni Ceva). A civian is a line from a triangle's vertex to any point on the opposite side. If instead of midpoints we take trisection points, three civians drawn as shown in the top illustration on the next page will cut the triangle into seven regions, each a multiple of 1/21 of the original triangle's area. The central triangle, shown shaded, has an area of 3/21, or 1/7. There are many clever ways to prove this, as well as the results of a more general case where each side of the triangle is divided into n equal parts. If the civians are drawn as before, to the first point from each vertex in a clockwise (or counterclockwise) direction around the triangle, the central triangle (as Howard D. Grossman has shown) has an area of $(n-2)^2/(n^2-n+1)$. A still broader generalization, in which the sides of the original triangle may vary independently in their number of equal parts, is discussed by H. S. M. Coxeter in his *Introduction to Geometry* (Wiley, 1961), Section 13.55. A formula going back to 1896 is given and Coxeter shows how easily it can be obtained by embedding the triangle within a regular lattice of points.

Every triangle has three sides and three angles. Euclid proved three cases in which two triangles are congruent if only three of the six elements are equal (for example two sides and their included angle). Is it possible for two triangles to have five of the six elements identical and yet *not* be congruent? It seems impossible, but there is an infinite set of such "5-con" triangles, as they have been called by Richard G. Pawley, a California mathematics teacher. Two 5-con triangles are congruent if three sides are equal, and therefore the only situation that permits noncongruence is the one in which two sides and three angles are equal. The smallest example of such a pair with integral sides is shown in the bottom illustration on the next page.



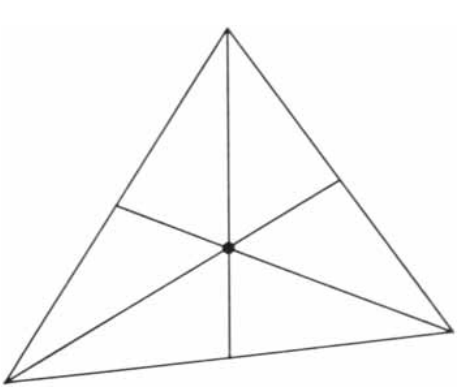
The nine-point circle

Note that the equal sides of 12 and 18 are not corresponding sides. The triangles are necessarily similar, because corresponding angles are equal, but they are not congruent. The problem of finding all such pairs is intimately connected with the golden ratio. A reference on this connection is Chapter 4 of Verner E. Hoggatt, Jr.'s *Fibonacci and Lucas Numbers* (1969), a booklet in the Houghton Mifflin "Mathematics Enrichment Series."

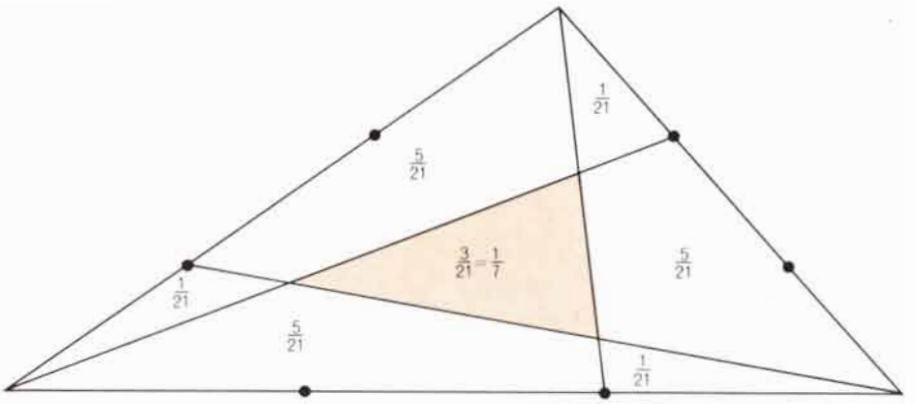
There are many ancient formulas for finding a triangle's sides, angles or area, given certain facts about its altitudes, medians and so on. The expression $\sqrt{s(s-a)(s-b)(s-c)}$, where a, b, c are the sides of any triangle and s is half of the sum of the three sides, gives the triangle's area. This amazingly simple formula was first proved in the *Metrica* of Heron of Alexandria, who is now known to have lived in the first or second century. The formula, Heron's chief claim to mathematical fame, is easily proved by trigonometry. Heron's geometric proof can be found in W. W. Rouse Ball's *A Short Account of the History of Mathematics* (Dover, 1960), Chapter 4. Heron, or Hero as he is sometimes called, is best known today for his delightful treatises on Greek automata and hydraulic toys, such as the perplexing "Hero's fountain," in which a stream of water seems to defy gravity by spouting higher than its source [see "The Amateur Scientist," December, 1966].

A classic puzzle of unknown origin, the solution to which involves similar triangles, has become rather notorious because, as correspondent Dudley F. Church so aptly put it, "its charm lies in the apparent simplicity (at first glance) of its solution, which quickly evolves into an algebraic mess." The problem concerns two crossed ladders of unequal length. (The problem is trivial if the ladders are equal.) They lean against two buildings as shown in the top illustration on page 136. Given the lengths of the ladders and the height of their crossing point, what is the width of the space between the buildings? The three given values vary widely in published versions of the puzzle. Here we take a typical instance from William R. Ransom's *One Hundred Mathematical Curiosities* (J. Weston Walch, 1955). The ladders, with lengths of 100 units (a) and 80 units (b), cross 10 units (c) above the ground. By considering similar triangles Ransom arrives at the formula, $k^4 - 2ck^3 + k^2(a^2 - b^2) - 2ck(a^2 - b^2) + c^2(a^2 - b^2) = 0$, which in this case becomes $k^4 - 20k^3 + 3{,}600k^2 - 72{,}000k + 360{,}000 = 0$.

This formidable equation is a quartic, best solved by Horner's method or some



Centroid trisects medians



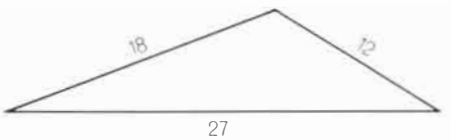
Trisecting civians

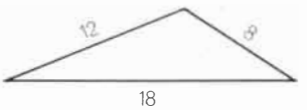
other method of successive approximations. The solution gives k a value of about 11.954, from which the width between buildings $(u + v)$ is found to be 79.10+. There are many other approaches to the problem. A good trigonometric solution is given in L. A. Graham's *The Surprise Attack in Mathematical Problems* (Dover, 1968), Problem 6. Other ways to solve the problem are in Graham's earlier paperback (Dover, 1959), *Ingenious Mathematical Problems and Methods*, Problem 25.

A difficult question arises at once. Are there forms of this problem (assuming unequal ladders) in which all four values are integers? As far as I know this was first answered by Albert A. Bennett of Brown University in *The American Mathematical Monthly*, April, 1941, solution to Problem E433. Bennett's equations have since been rediscovered many times. The smallest integral values, minimizing both the height of the crossing and the width between the buildings, result when the ladders are 119 units and 70 units long, the crossing is 30 units above the ground and the width 56 units. It turns out that if those four values are integral, the lengths of all other line segments in the diagram are also integral. I lack space to go into Bennett's formulas, but the interested reader will find them in the reference cited, in Ransom's book and in Alan Sutcliffe's article "Complete Solution of the Ladder Problem in Integers," *The Mathematical Gazette*, Vol. 47, May, 1963, pages 133–136. In addition there is an infinity of solutions in which the distance between the *tops* of the ladders is also an integer. (See Gerald J. Janusz' answer to Problem 5323, proposed by Sutcliffe, in *The American Mathematical Monthly*, Vol. 73, December, 1966, pages 1125–1127.)

When we are given no more than the distances from a point to the three vertexes of a triangle, there obviously is no way to construct the triangle or to determine the lengths of its sides. If the triangle is known to be equilateral, however, the side *can* be determined. The point may be inside or outside the triangle. A problem of this type, of unknown origin, is frequently sent to me by readers, usually in the following form: A point within an equilateral triangle is three, four and five units from the triangle's corners. How long is the triangle's side?

There are many ways to solve this problem, one of which, using a simple construction and similar triangles, will be given next month. I know of no version in which all four values are integral.





Smallest "5-con" triangle pair

February's set of short problems prompted—as usual—far too many letters to allow more than a cursory summary of their contents at this time. The problem of the three squares was proved by so many readers in so many different ways (Charles Trigg alone sent 42 proofs) that I cannot list names. Scores of correspondents avoided construction lines by making the diagonals equal to the square roots of 2, 5 and 10 and then using ratios to find two similar triangles from which the desired proof would follow. Others generalized the problem in

Many great ideas are scrapped for lack of timely financial support.

But yours needn't. Because now there's a book that can put you directly in touch with many willing sources of investment capital.

Not banks. Not credit companies. These firms are heavyweights in the field of high-risk investment. Some familiar names. Some surprises. But all on the lookout for a profitable venture.

Never before has this information been compiled in a single directory. For each company you'll find such useful facts as which investment areas the firm prefers; which ones it avoids; how much it invests; what other specialized help it can provide; data on the firm's last three investments; and, of course, whom to contact. You'll also find helpful advice on how best to present your ideas to potential investors.

The name of the book is *Venture Capital.* It's not for "get rich quick" dreamers. And it's not cheap.

But if you've got a great idea (or company) in need of an innovative source of money, this book could be the best investment you'll ever make.

Send today for your copy of Venture Capital.

TECHNIMETRICS, INC.
527 Madison Avenue
New York, N. Y. 10022

Please send me a copy of **Venture Capital**

☐ I wish to take advantage of the pre-publication 15% discount. I enclose a check for $17.00 ppd.

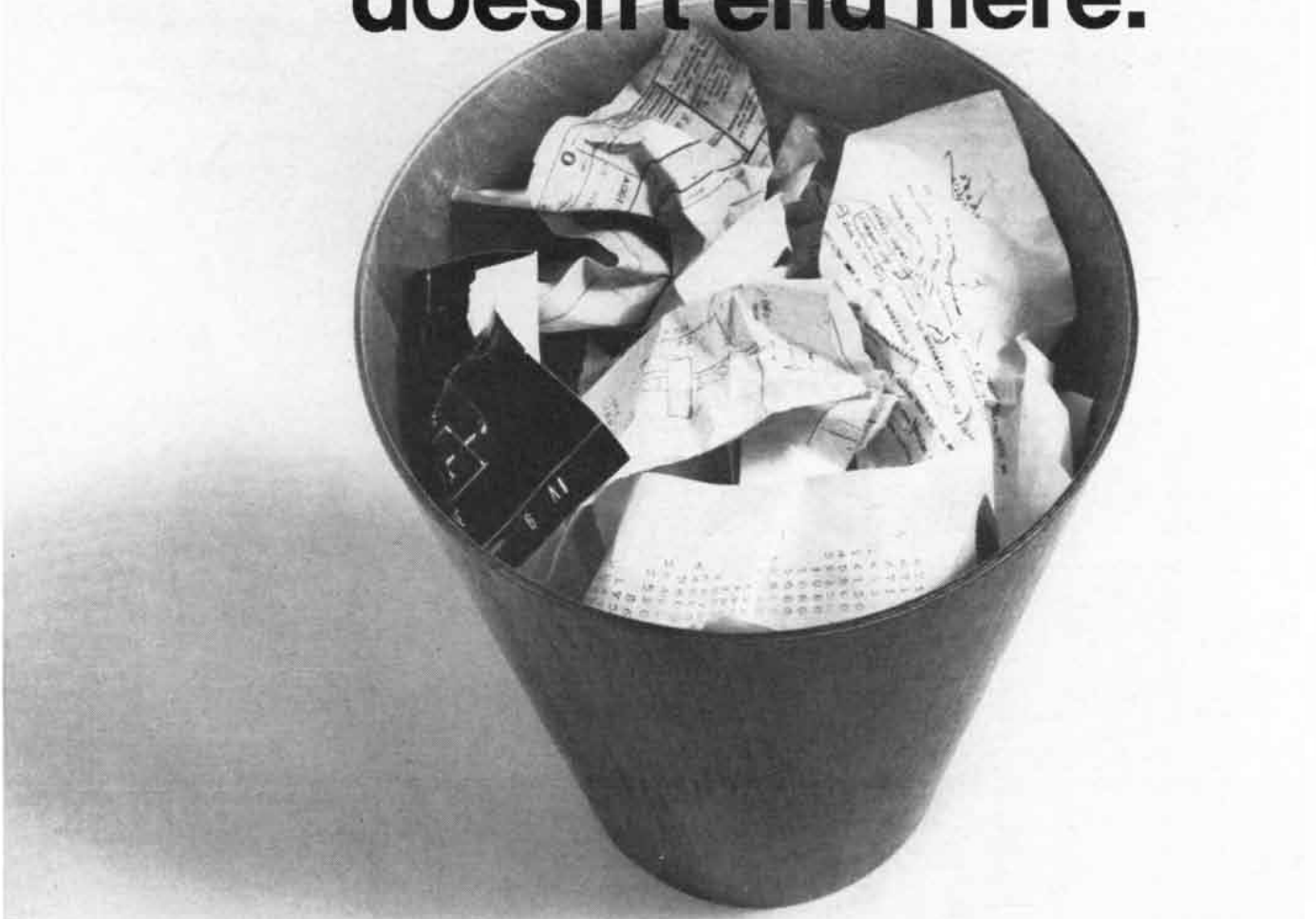
☐ Bill me at the regular price of $20.00 ppd.

Name_____

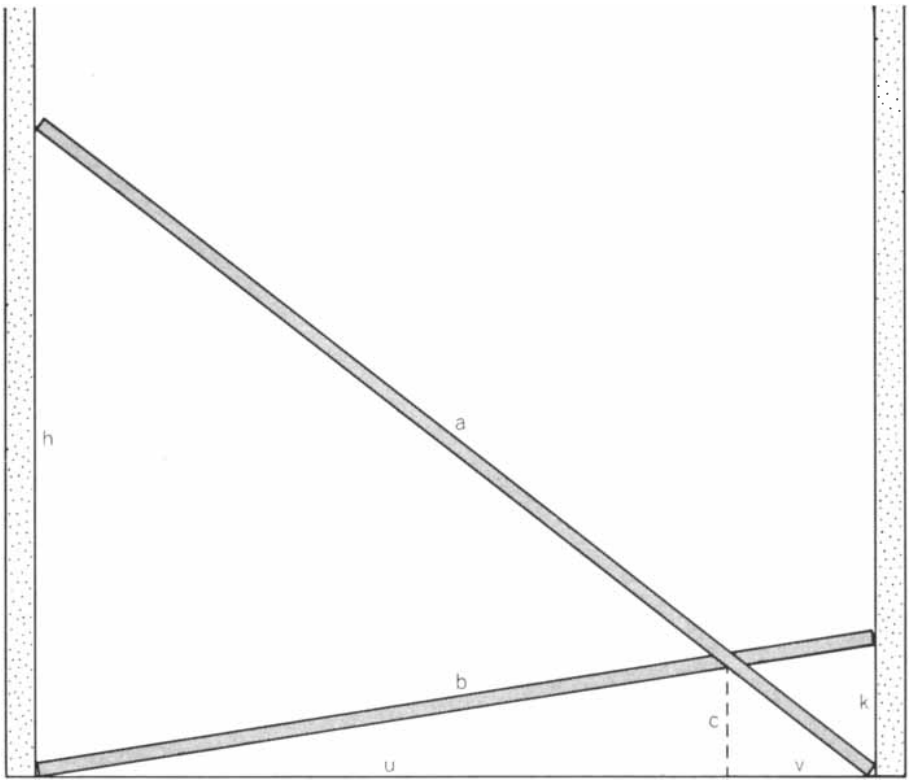
Address_____

City_____ State_____ Zip_____

With "Venture Capital" your great idea doesn't end here.



Venture Capital is a service of TECHNIMETRICS, INC. 527 Madison Avenue, New York, N.Y. 10022 Tel: (212) 688-6910



Crossed-ladders problem

MOVE	BLOCK	DIRECTION
1	6	U, R
2	1	D
3	5	L
4	6	L
5	4	D
6	5	R
7	2	D
8	3	L
9	5	U
10	2	R
11	6	U, L
12	4	L, U
13	7	U
14	10	R
15	9	R
16	8	D
17	1	D
18	7	L
19	2	D
20	4	R, D
21	5	D
22	3	R
23	6	U
24	5	L, U

MOVE	BLOCK	DIRECTION
25	4	U
26	2	U
27	10	U
28	9	R
29	7	U
30	1	U
31	8	U, R
32	1	D
33	7	L
34	10	L
35	2	L
36	4	D
37	2	R, D
38	3	D
39	6	R
40	5	R
41	7	U
42	10	L, U, L, U
43	8	U, L, U, L
44	4	L, U
45	2	D, L
46	9	U
47	2	R
48	1	R

John W. Wright's 48-move solution to sliding-block puzzle

unusual ways, but I must postpone their suggestions to a time when the column goes into a book collection.

Escott's sliding-block puzzle, for which I gave a 66-move solution, was solved in 48 moves (the record) by George R. Cook, David Elwell, Ruprecht Franzel, Louis G. Melio, Eva L. Milbouer, Ira B. Penniman, Maurice Povah, Wesley L. Walker and John W. Wright. No two methods were exactly alike. The table at the bottom of this page gives Wright's typical solution. The letters *U*, *D*, *L* and *R* stand for up, down, left and right; in every case the numbered piece moves as far as possible in the specified direction.

Since the initial pattern has twofold symmetry, every solution has its inverse. In this case the inverse starts with piece 5 moving down and to the left instead of piece 6 moving up and to the right and continues with symmetrically corresponding moves.

Dean Hickerson, Dean Hoffman and Gerson B. Robison generalized the puzzle of the 10 cells (Problem No. 7) to obtain a single formula that yielded all possible solutions in all number bases. Many readers sent what they mistakenly believed to be solutions different from the one given for the 10-pennies problem. Their mistake was a failure to note that the bowling-pin formation contains two tipped equilateral triangles with nonhorizontal bases.

David Silverman's geography game proved to have four first moves that win for the first player on his second move: Tennessee, Delaware, Rhode Island and Maryland. (It must be assumed that all players play rationally and without collusion to trap the first player.) Other states, such as Vermont, Texas and Connecticut, lead to wins on the first player's third move. Readers who found all four quick winners are J. W. F. Juritz, Friend H. Kierstead, Jr., Frederick S. Koehl, Jerry A. Menikoff and Carter V. Smith. The problem will appear in Silverman's collection of original puzzles *It's Your Move,* to be published by McGraw-Hill.

A large number of readers called my attention to ingenious shortcuts by which one can search for cyclic numbers, the topic for March, but again the subject is so complex that I must postpone it to a future time.

To climb in 10 steps to the top of the Penrose staircase shown last month, walk up four steps, turn right, continue up three more steps, go back around the level *U*-shaped path, down three steps, then up three to the top.

Announcing America's first (and only)

ALL-BOOK/ ALL-PUBLISHER DISCOUNT CARD

*Not available to dealers, wholesalers, retail book establishments or their representatives or sales agents.

- **COMPUTER REGISTERED** *in bearer's name for life (or until de-registration is requested).*
- **ENTITLES BEARER** *to purchase from us any book in print sold in any bookstore in the United States or its Territories (or listed in the* Master Catalog of American Books) *at guaranteed discounts up to 81% off original publishers' prices. (See Footnote 1 below)*
- **THIS DISCOUNT PRIVILEGE INCLUDES** *all best sellers such as* The Peter Principle, The Godfather, The Inheritors, The Graham Kerr Cookbook, Andromeda Strain, I Ching; *all fiction and non-fiction, hard cover and paper back, scientific and technical, political, historical, biographical, religious, philosophical, art, even reference works, unabridged, dictionaries, college texts, and professional books.*
- **THIS DISCOUNT PRIVILEGE ALSO INCLUDES** *books of all publishers. E.g., Random House, Knopf, Doubleday, Harper & Row, Simon & Schuster, Harvard University Press, Grove Press, etc. Even rare or "specialist" publishers.*
- **ACCEPTING THIS CARD** *involves absolutely no obligation or commitment to purchase anything, at any time, for any reason.*
- **COMPLETE COST** *of the All-Book/All-Publisher Discount Card, with full privileges, is $5.* There are no other charges. *This one-time Registration Fee is good for life. You need never renew or pay another fee for as long as you live. (See Footnote 2 below)*

HOW TO ORDER AT DISCOUNTS UP TO 81%

Simply look up selection(s) you want in the Master Catalog of American Books (provided free; see description below) and note the order number listed beside each book. Or, you may prefer to order a book currently offered by some other book club (that, too, is perfectly all right). Or, feel free to copy down the name of *any* book reviewed or advertised in *any* publication or sold in *any* bookstore. Then simply send in your order —accompanied, of course, by your membership number.

HOW MUCH CAN YOU SAVE IN DOLLARS AND CENTS?

If you order at the maximum discount of 81%, then the price you may have formerly paid for only *one* book will now bring you up to *five* books. Or, said another way: five $5.95 books, originally published at $29.75, might now be purchased (provided you have a card) for as little as $5.96. A total saving of $23.79. **In any event, no cardholder will ever pay full price for any item purchased with his card.**

Here are some orders processed by our IBM-S/360 computers during a typical week. (No comment, criticism or endorsement of any book, publisher or author is implied or intended. Prices subject to change.)

Title	List Price	Card Price
GLOSSARY OF THE CONSTRUCTION, DESIGN & USE OF ARMS & ARMOR, Stone. (Brussel)	$50.00	$ 9.95
AMERICAN HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE	8.95	6.93
BIRDS OF PREY OF THE WORLD, Grossman & Jamlet. (Bonanza)	25.00	9.93
JAMES BEARD OUTDOOR COOKBOOK, Beard. (Ridge Press)	12.50	5.89
EARLY AMERICAN HERB RECIPES	10.00	2.98
THE ANTI-SEX, Masters & Lea	8.50	3.89
PICTURE BOOK OF SEXUAL LOVE: A PHOTOGRAPHIC GUIDE	15.00	9.95
THE MEMOIRS OF CASANOVA, 6 Vols. (Putnam)	$30.00	$ 7.89
LAROUSSE ENCYCLOPEDIA OF THE EARTH (Prometheus)	17.50	9.89
COMPLETE PIANO MUSIC OF BEETHOVEN. Zubin Mehta, others. 21 stereo LP records. Boxed. With analysis and program notes	100.00	19.89
THE FINE ART OF CHINESE COOKING, Dr. Lee Su Jan	3.95	1.00
HANDWRITING ANALYSIS, Bunker. (Nelson-Hall)	10.00	2.49
DICTIONARY OF APHRODISIACS, Wedeck. (Philosophical Library)	10.00	5.88
COMPLETE STRING QUARTETS OF BEETHOVEN. Fine Arts Quartet. 9 stereo LP records. Boxed.	40.00	9.79
COMPLETE ILLUSTRATED BOOK OF YOGA, Vishnudavananda (Bell)	10.00	2.49
THE EROTIC REVOLUTION, L. Lipton	7.50	2.98
PICTORIAL HISTORY OF MAGIC & THE SUPERNATURAL, Bessey. (Sprint)	10.00	5.89

(cont'd top of next col.)

FOOTNOTE 1 Master Catalog of American books is sent free with your card. (You are also entitled to request, at a slight additional cost, the Master Catalog of British Books and—if you so desire—to order any and all British books in print, as well as obtain book quotations and bibliographic searches for research and scholarly purposes.

FOOTNOTE 2 We do not profit from your $5 Registration Fee (which, incidentally, cardholders generally "amortize" many times over on their first order alone). It goes at once, in part, to offset the considerable cost of programming your name and order-number into the battery of IBM-S/360 computers comprising our new Computer Control Network, which can store over 1½ BILLION characters of information, scan and process book orders at rate of TWO MILLION CHARACTERS PER MINUTE. Currently servicing orders for more than 100,000 booklovers, it is the only such network of its kind in existence.

FOOTNOTE 3 Cardholders receive only genuine publishers editions. Conventional book clubs often send members "club editions"—i.e., special "economy" printings using cheaper paper, smaller type, flimsier bindings. Our discounts are never achieved this way. All books are publishers' editions and are guaranteed brand new and bindery fresh—the same item selling for full price in the "bookstore around the corner."

HOW THESE LARGE DISCOUNTS ARE ACHIEVED

Your orders need not pass through a conventional bookstore. Our Computer Control Network and large-volume purchasing helps slash the costs (and "mark-ups") of the normal book distribution chain. Resultant savings are passed on to cardholders as discounts of up to 81% off original publishers' prices. (See footnote 3 below)

YOU MUST BE SATISFIED OR MONEY REFUNDED

If for any reason after receiving it you feel your card is not as valuable as we have led you to believe (or even if you have simply changed your mind), you are free to return it within 10 days for a full refund of your Registration Fee. (Please tear card in half before returning it.) Naturally, any book you purchase with your card must also satisfy you completely or you may return it for a full refund—no questions asked.

To apply for an ALL-BOOK/ALL-PUBLISHER DISCOUNT CARD, complete and mail the form below. (Should your application not be approved, your Registration Fee will be returned to you at once.) Mail form to ALL-BOOK/ALL-PUBLISHER DISCOUNT CARD, American Book Club, Hazleton, Pa. 18201.

COMPUTER CONTROL NETWORK

Links thousands of current cardholders with all major publishers in the U. S. via electronic computer. Bypasses the conventional retail bookstore—and its "mark-up." Passes the savings of "direct-from-publisher" ordering (up to 81%) on to cardholder. Also permits cardholders (if they wish) to order any British books in print for a slight additional service charge.

FREE WITH YOUR CARD

Master Catalog of American Books (available to cardholders only) lists 16,000 books in print at guaranteed discounts up to 81% off original publishers' prices. An invaluable reference work, it is indexed by subject, author, title and publisher. You never pay full price for any book you order, under any circumstances. This 350 page Master Catalog remains yours—free and without obligation—even if you later decide to return your card for a refund.

APPLICATION FORM

Mail to: **ALL-BOOK/ALL-PUBLISHER DISCOUNT CARD** 196
American Book Club, Hazleton, Pa. 18201

Please send a registered ALL-BOOK/ALL-PUBLISHER DISCOUNT CARD in my name. (Also forward Master Catalog of American Books free of extra charge.) I understand that acceptance of this card does not obligate me to buy anything. But it entitles me to order books at discounts up to 81% off original publishers' prices, plus shipping. Enclosed is my one-time Registration Fee of $5 to defray computer programming. **There are no other charges for the rest of my life.** If not satisfied, I am free to return card within 10 days for a full refund of my Registration Fee. **But the Master Catalogue of American Books will be mine to keep free, whether I decide to remain a cardholder or not.**

☐ Chge. to my Master Chge. #_____

Name _____ _____ Initial here. I am not a book dealer, and will not resell books purchased with card at a higher price.

Address _____

City _____ State _____ Zip _____

• *If your card is lost or stolen please notify American Book Club at above address without delay.*

Siemens

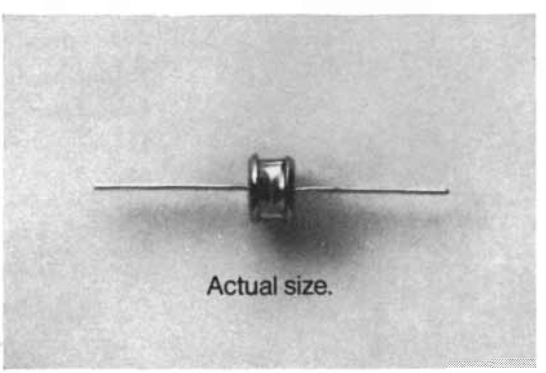
A $1.00 device can protect electronic equipment against damage from lightning and other voltage transients. We bring Qualitätsware to America.

Lightning strikes the earth about 100 times each second.

In a thunderstorm, the voltage potential between cloud and ground can build up to as much as 100 million volts.

The direct damage caused by lightning is estimated to be in the neighborhood of 500 million dollars a year. What cannot be easily measured are the additional losses resulting from equipment downtime.

Lightning is just one of the many voltage transients that can affect expensive electronic equipment. Some of the other sources are static discharges, internal switching of inductive components, short circuits, and inductive influences through grounded shorts.


Actual size.

Before the use of semiconductors or other delicate components, voltage transients did not present a serious problem to the design engineer. Today, however, solid-state circuitry and a greater degree of electronic sophistication demands greater attention to voltage protection.

To meet the need for transient protection, Siemens developed a line of miniature surge voltage protectors. These tiny, gas-filled devices react rapidly, reliably and repeatedly to voltage surges.

Widely used today in communications equipment, surge voltage protectors are fast gaining acceptance in other electronics areas. Control systems, processing equipment, testing and measuring instruments, data processing and peripheral equipment are some of the new applications for surge voltage protection.

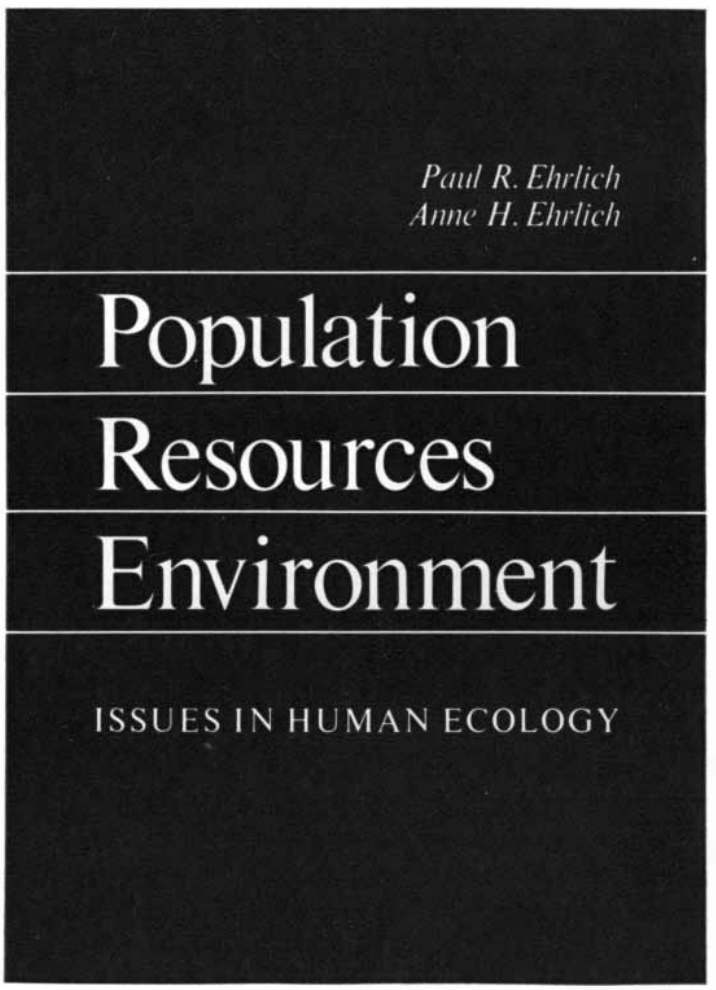
Siemens has also developed a similar device which can be triggered to achieve fast bi-polar switching of high currents up to 5,000 amps. This highly efficient device is designed to replace silicon controlled rectifiers (SCR's) in certain applications where the di/dt and dv/dt rating presents a problem. Energy required to trigger the Siemens spark gap is close to zero (10^{-6} Ws).

As a world-wide manufacturer and user of electronic components such as capacitors, resistors, transistors, electron tubes and rectifiers, Siemens has designed its surge voltage protectors with total circuit requirements in mind. They cover a dc striking voltage range from 90 volts to 6.5 kv (higher on request).

Siemens Corporation, 186 Wood Avenue South, Iselin, New Jersey 08830.

Siemens of West Germany. A worldwide company that sold 2.84 billion dollars of quality products last year.

High Praise for a Major New Book:



1970, 383 pages, illustrated, $8.95

"For every American interested in understanding and solving the critical problem of overpopulation and our deteriorating environment, this thoughtful new book is a must."
Senator Joseph D. Tydings

"One of the year's most important books."
Saturday Review

"Paul Ehrlich is one of the most farsighted, outspoken prophets of the environment movement in this country. If this book does not arouse us to the dangers we face, none will!"
Stewart L. Udall

"The best single descriptive and analytic treatment of the subject that I have yet seen . . . Telling and courageous."
Robert L. Heilbroner, *The New York Review of Books*

"I have found this authoritative and well-documented discussion of today's great problems not only valuable but also very interesting."
Linus Pauling

"I have read it from cover to cover and am now busily recommending it to all my friends."
Pete Seeger

"A superb book—well written, well organized, informative, and well argued."
Preston Cloud

"A stimulating, timely, and important book. Its emphasis is on biological and ecological problems, but it has suggested to me an interdisciplinary approach for teaching physics to stress its relevance to the problems of human survival."
Albert V. Baez

"This is just a hell of a fine book."
Garrett Hardin

From your bookseller or from Dept. C,
W. H. FREEMAN AND COMPANY
660 Market Street, San Francisco, California 94104
Warner House, Folkestone, Kent, England

THE AMATEUR SCIENTIST

The effects of gravity on plant growth, and the deposition and use of thin films

Conducted by C. L. Stong

Soon after a potted plant has been laid on its side the stem turns up and the roots turn down. Experiments indicate that such changes in the direction of growth are induced by gravity: plants tend to align themselves in the direction of the earth's gravitational field. Botanists refer to this tendency as geotropism and have discovered by experiment that it arises from the influence of gravity on certain substances in plants, namely the organic compounds known as auxins, which stimulate the growth of the upper parts of plants but appear under some conditions to suppress the growth of roots.

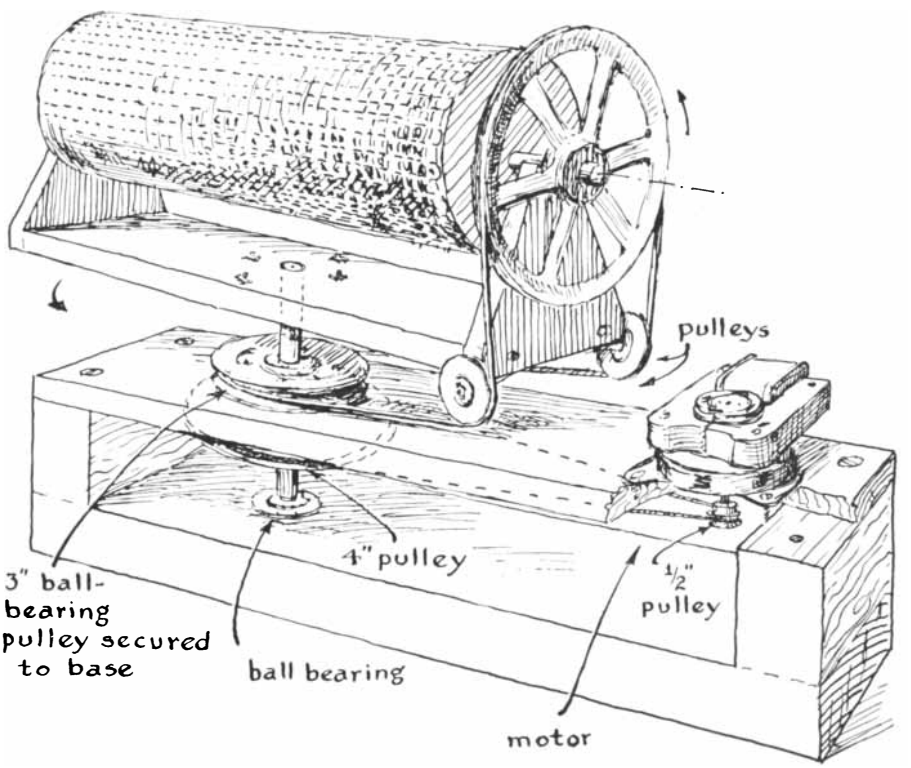
When a potted plant is inclined from the vertical, auxin concentrates in the lower sides of the stem and roots. The concentration causes the lower side of the stem to grow faster than the upper side. The stem bends upward. Conversely, auxin in the lower side of the roots retards growth, but normal growth continues in the upper side. The root turns downward.

These effects can be observed by making a simple experiment. With India ink draw a set of evenly spaced marks along the lower side of both the root and the stem of a seedling. Place the seedling horizontally in a moist container for 24 hours and then examine the marks. The spacing between the marks will have increased on the lower side of the stem where it bent upward but will not have increased on the lower side of the roots.

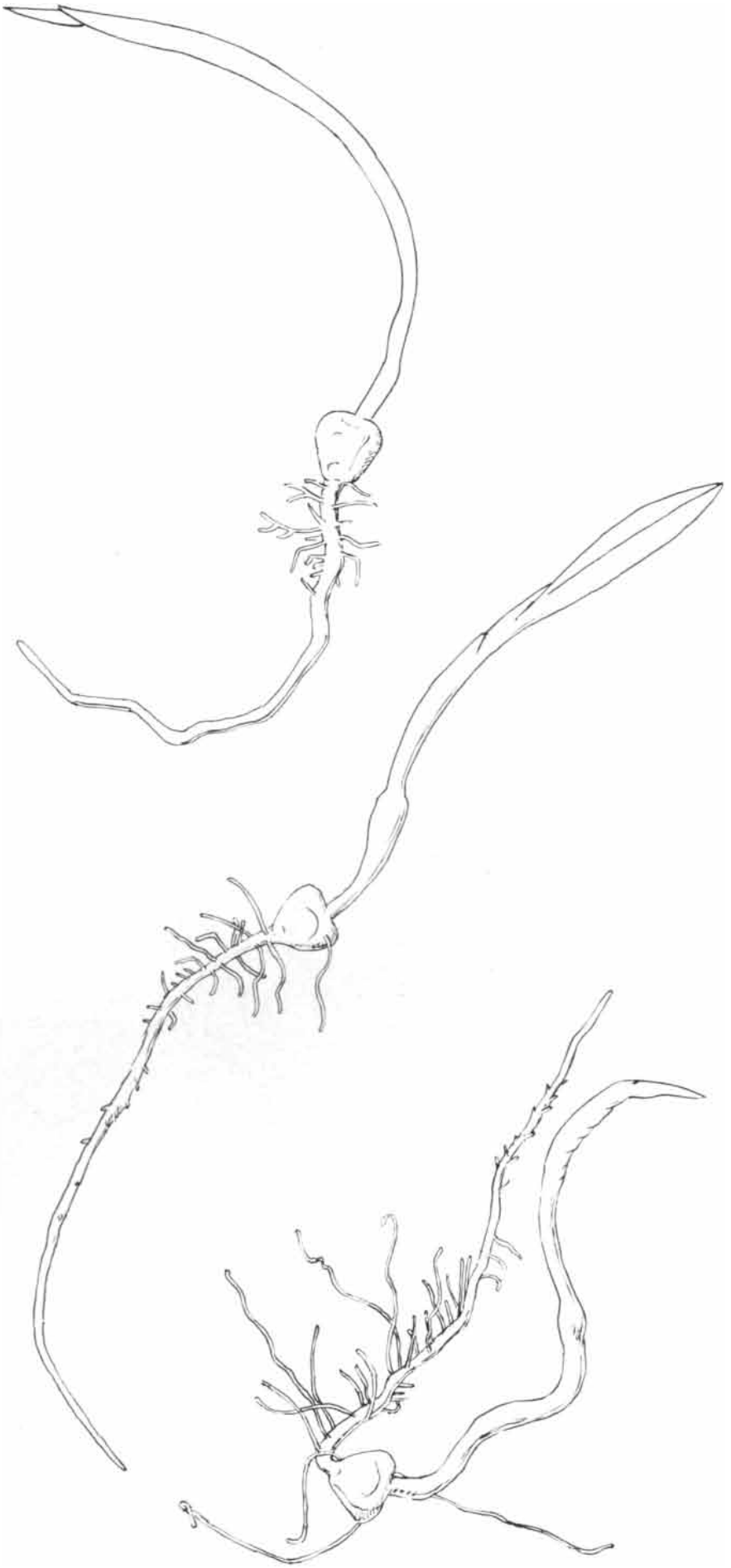
Of course this experiment does not prove that gravity is solely responsible for reshaping the plant. The tops of plants grow toward sources of light, and the leaves of many plants follow the sun. The unequal distribution of auxin is also responsible for this effect, which is known as phototropism, but botanists have not yet learned how light influences the substance. Three mechanisms have been suggested. Light may inhibit the production of auxin on the exposed side. Alternatively, it may denature a portion of the normal production. On the other hand, as in the case of gravity, light may cause auxin to move to the shady side of the stem. Whatever the mechanism, you can identify the side of the stem that grows faster by drawing evenly spaced rings of India ink around the stem of a potted plant, placing the pot upright near a window and measuring the spacing of the rings as the stem bends toward the light. This experiment casts doubt on the assumption that the stem of an inclined plant turns upward in response to gravity. Perhaps the stem is merely seeking light, which usually comes from above.

All doubt concerning the role of geotropism in plant growth can be resolved by another experiment that was first performed about 150 years ago. In this experiment upright pots that contain seeds or seedlings are uniformly illuminated on all sides, but the gravitational field is tilted from the vertical by mounting the pots upright on the rim of a wheel that turns in the horizontal plane. (Each pot is enclosed in a transparent container for protection against currents of air.) When the wheel turns, the pots are acted on by two components of inertial force: a horizontal component arising from the circular motion of the wheel and a vertical component resulting from the acceleration of gravity. The resultant force acts at an intermediate angle that is determined by the speed of the wheel.

The roots of plants that are grown on the continuously rotating wheel extend outward and downward at precisely the angle of the resultant force. The stems grow inward and upward in exact align-



Don Graham's apparatus for experiments with geotropism



Corn seedlings grown in a constantly changing gravitational field

ment with the roots. The lines of resultant force along which the plants grow trace a cone in space as the wheel rotates. The altitude of the cone varies inversely with the speed of the wheel, an experimental result that can be explained only on the assumption that inertial force strongly influences the direction in which plants grow. Having established this fact, experimenters appear to have closed the books on geotropism and shifted their attention to other matters.

Interest in geotropism has now been dramatically rekindled, at least for one amateur, by the development of space vehicles. About a year ago Don Graham, who is a commercial artist in Petrolia, Calif., began to wonder how a plant might react if it were grown in a weightless state. Graham decided to undertake the experiment but could think of no way of eliminating gravity without putting plants in a space vehicle. Instead he devised an apparatus that interferes with the natural response of auxin to the gravitational field. Plants that are grown in the apparatus apparently lose their sense of direction. Graham built a pot that rotates slowly but continuously in all coordinates of three-dimensional space and undertook to grow corn in it. He describes the experiment as follows:

"My apparatus consists essentially of a cylindrical pot that rotates simultaneously on its axis and in the horizontal plane [*see illustration on preceding page*]. The hollow cylinder, made of the wire mesh known as hardware cloth, is a foot long and about four inches in diameter. The ends of the cylinder are closed by two wood disks. The cylinder is supported by a shaft that passes through snugly fitting holes in the center of the disks and is rotated on its axis by a pulley on one end. The shaft is supported at its ends by a pair of vertical brackets that are fastened to a horizontal wood base.

"The base is rotated in the horizontal plane by a vertical shaft that is coupled to a slow-speed motor by a belt. The motor turns at eight revolutions per minute. A 1:8 pulley ratio reduces the speed of the vertical shaft to one revolution per minute. The pulley that drives the cylinder is coupled by a belt to a fixed pulley attached to the frame on which the motor is mounted. None of the dimensions are critical, but the diameter of the fixed pulley should not be a multiple of the diameter of the driven pulley, because this ratio would generate a cyclical pattern of cylinder positions. A 7:11 ratio works well.

"The cylinder is filled with a mixture of four parts of sphagnum moss to six

parts of rich loam. I moistened the soil and packed the cylinder as though it were an ordinary pot. Sweet corn was selected for the experiment because the seedlings of corn develop in the form of a series of concentric whorls that appear to be stronger and sturdier than most plants are during the first few days of germination.

"Seven uniformly spaced openings, each 1/2 inch square, were cut in the wire mesh to form a helical path of one full turn that extends to within an inch of the ends of the cylinder. With tweezers I pushed a seed through each opening and into the soil to a depth of two inches, which is to say to the middle of the cylinder. The cylinder was wrapped with a single sheet of clear polyethylene to conserve moisture.

"The apparatus was placed on the ground in the backyard, where it would receive full sunlight. The motor was turned on and operated continuously for 14 days, except during brief intervals when it was stopped for a check on the temperature and moisture of the potted soil. During this entire period seven additional seeds of the same stock were growing in an adjacent garden area that contained identical soil. These plants served as controls.

"On the 14th day all seedlings (both the experimental ones and the controls) were removed from the soil, washed gently, measured and replanted in the garden. All seven control seedlings had grown to an average height of 2½ inches. They appeared to be normal in every respect. The most vigorous measured seven inches from the tip of the root to the tip of the longest leaf.

"The experimental seedlings had grown as vigorously as the controls. The largest measured nine inches from root tip to leaf tip. There the similarity ended. Whereas the controls grew straight up and down, most of the seedlings were sadly misshapen. Only one plant had found daylight; it grew about 2½ inches beyond the wire. The root, which was about 3½ inches long, bent randomly through the soil. One seedling grew in reverse: the root penetrated the wire and the stem remained in the soil. The root and stem of another seedling grew parallel in the same direction! One seed failed to germinate. Another produced a short root and an even shorter parallel stem. No experimental seedling had grown in the normal up-down direction. As the plants were removed I made a record of the direction in which each had grown with respect to its position in the cylinder. The record indicated that the direction of growth had been random.

Seed ear produced by an experimental plant

"All seedlings matured in the garden, where they were cultivated and weeded regularly. The controls grew to heights ranging from five to nine feet and yielded an average of four ears of corn per stalk. In contrast, the confused seedlings matured at a height of less than three feet. Only one experimental plant produced an ear, and it was a distorted, underdeveloped runt [*see illustration above*].

"Other experiments involving geotropism come to mind. For example, how long can a germinating plant survive without damage in the absence of a normal gravitational field? My plants were rotated for 14 days. How much damage might have been evident if I had transplanted the seedlings after the fourth or the eighth day? How would a plant react to an increase or a decrease in the intensity of the gravitational field?

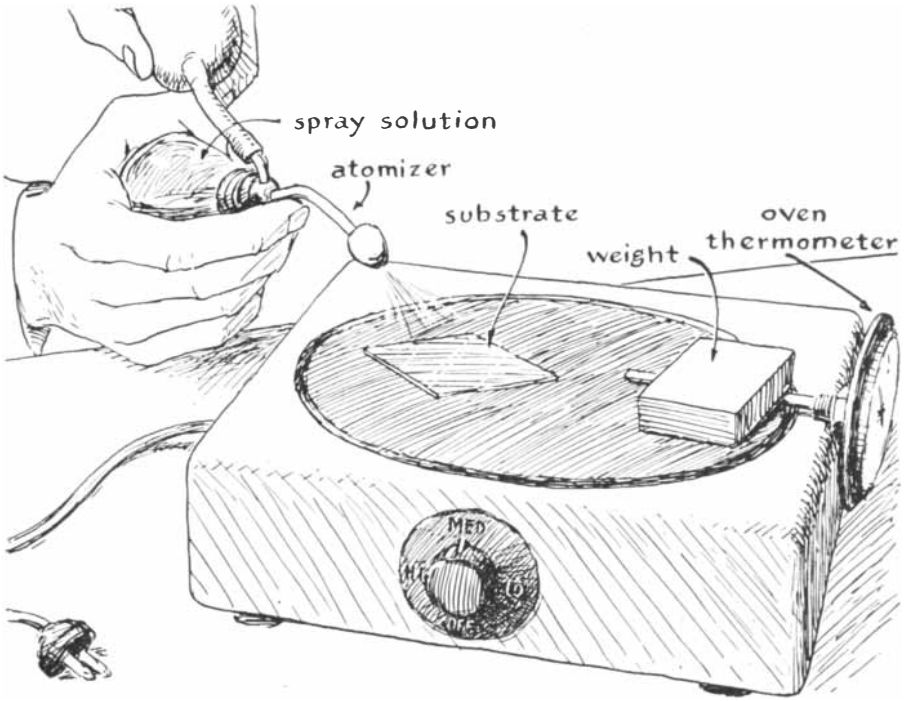
"I can think of no practical apparatus for lowering the strength of gravity on the earth to, say, that of the planet Mars. On the other hand, it is easy to investigate the influence on germinating seeds of an inertial force greater than the earth's gravity by growing plants on the rim of a wheel that is spinning. It might be interesting to find out how sweet corn would grow on Jupiter, where gravity at the surface is 2.6 times stronger than it is on the earth.

"One should not place too much confidence in the outcome of a single experiment. Nonetheless, having observed the reaction of my confused corn, I suspect that no plant in an advanced stage of evolution can grow normally in a weightless environment. Nor can such a plant reproduce itself for more than a few generations, notwithstanding the fact that one of mine developed seeds. Perhaps lower marine organisms such as algae, corals or fungi could multiply in the absence of a gravitational field. So far as higher plants are concerned, however, gravity appears to be as essential to growth as sunlight. In my opinion, an orbiting spacecraft would make a poor garden."

When one turns on a television set, makes a photograph, looks through a telescope or sends a party of astronauts to the moon, one calls into service thin films of metal or metallic compounds. The films are essential elements in fluorescent screens, photocells, thermistors, transistors, antireflection lenses and scores of other devices with which amateurs can experiment. Not many amateurs bother to make their own films or to experiment with them, perhaps because most thin-film devices are inexpensive. Why make a poor transistor when you can buy a good one for less than $1? Roger Baker of Austin, Tex., explains that he makes his own transistors (and other thin-film devices) largely because he enjoys tinkering with small devices that occasionally behave in unexpected ways.

"Recently," Baker writes, "I learned of a governing principle known as Murphy's first law of biology. It states: 'Under any given set of environmental conditions an experimental animal behaves as it damn well pleases.' The same law appears to govern the behavior of thin metallic films, at least those I make. Some of my 'transistors' make dandy thermistors, and an occasional photocell works better as a fluorescent screen. There might be fewer surprises if I used better tools and had more experience, but some of the fun might also be lost. The techniques used by industry for making thin films are not beyond the reach of amateurs, but they require vacuum pumps, electronic heating devices and controlled sources of high voltage that are costly and inconvenient to use. Thin films can also be deposited chemically. I use this method.

"Most of my films are deposited on substrates of glass. Usually I heat the glass and spray the surface with a solution of selected chemicals. The sprays



Roger Baker's apparatus for depositing metallic films on glass

react immediately to form the film.

"Films can be annealed in various atmospheres and at various temperatures that alter their composition, structure and properties. The properties of a deposited film can also be modified by recrystallization, by solid-state diffusion or by a vapor-phase displacement reaction. These procedures are much simpler to perform than their imposing names suggest. The properties of films can be radically altered by the addition of minute amounts of impurities, either when they are formed or by subsequent diffusion. The microstructure of the substrate can also influence the properties of a film. For example, calcium sulfide forms an amorphous film when sprayed on a metal surface, but on glass it becomes a crystalline film.

"The required tools include an electric hot plate, a diamond point for cutting thin glass, an atomizer and a microammeter. Desirable accessories are a fume hood, which can be improvised if you have an exhaust fan, an oven thermometer for measuring the temperature of the hot plate, a triple-beam chemical balance, a pair of tweezers and chemical glassware for preparing solutions. For substrates I use mostly cover glasses of 35-millimeter Kodak slides, which I cut into small rectangles with the diamond point. These glasses can be heated (up to 600 degrees Celsius) and sprayed without breaking. Thin disks of alumina can be used at higher temperatures. I salvage them from discarded vacuum tubes. I immerse the glass slides for three days in a solution of one part by volume of nitric acid to 12 parts of distilled water. The acid leaches sodium and calcium ions from the glass, exposing a surface layer of relatively pure silica.

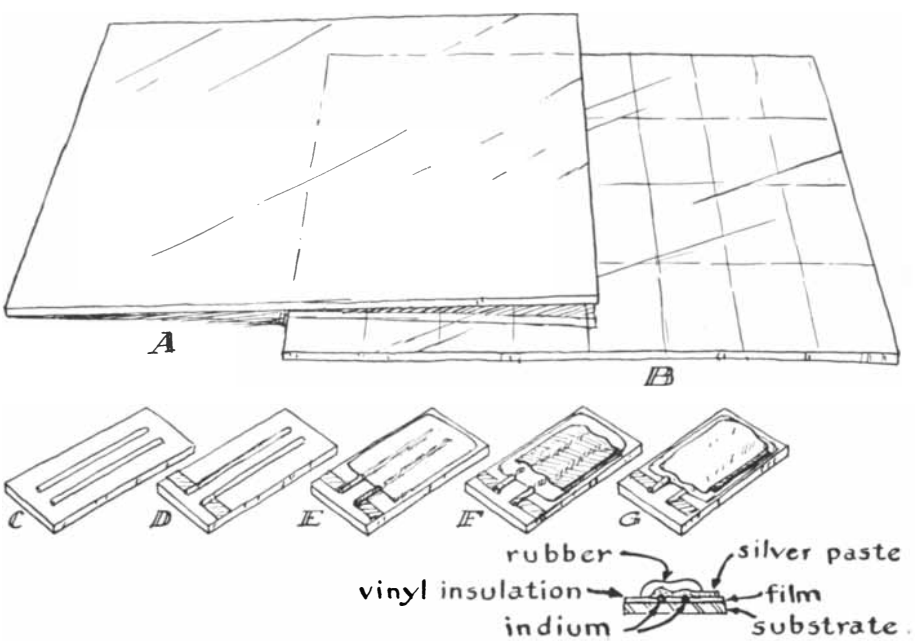
"A great variety of semiconducting oxide films can be made by thermally decomposing the resinate salts of metals. Resinate salts are prepared by stirring an excess of pure granulated resin into a one-normal (1 N) solution of sodium hydroxide. The solution turns milky as it cools. Pour off and retain the milky solution. To make a metal resinate, reheat the milky sodium-resinate solution and combine it with a weak solution of the metal salt, stirring the mixture vigorously.

"A relatively large volume of sodium resinate reacts with a small volume of metal salt. An excess of sodium is indicated by a *p*H of 8 or more. Add metal salt to lower the *p*H. The desired metal resinate appears as a thick precipitate.

"Filter the solution to recover the precipitate and wash it thoroughly with hot distilled water. Spread the moist filter cake and dry it at a temperature of about 50 degrees C. Dissolve the dried material in an organic solvent such as turpentine. Allow the sediment to settle. Use the clear upper layer for experiments.

"With a disposable capillary tube apply a few drops of the clear fluid to the center of a prepared cover glass and rock the glass to spread the fluid into a uniform film that extends to the edges. Heat the coated glass on the hot plate. The film will smoke and turn dark. In time, at a temperature that depends on the nature of the resinate, the dark color will clear, leaving a thin film of metallic oxide. The cover glass can then be scribed with the diamond point and broken into rectangles of convenient size for further processing and experimentation.

"Sulfide films can be formed directly from a number of oxide films. Sprinkle a few milligrams of sulfur on the back side of the coated substrate, wrap it with several layers of aluminum foil, fold the



Sequential steps in making a thin-film transistor

ends of the foil over the package and heat the package. The hot vapor of the sulfur will react with many oxides to form adherent films with interesting electrical properties. Two drops of different resinates can be allowed to diffuse partially together so that the properties of various ratios of the two can be explored.

"So far I have experimented with gold, nickel, cobalt, copper, iron, manganese, silver, indium, chromium, zinc and cadmium resinates. The salts of noble metals decompose into metallic films instead of oxides. They can be used for making electrical connections between various films of oxide previously applied to a substrate.

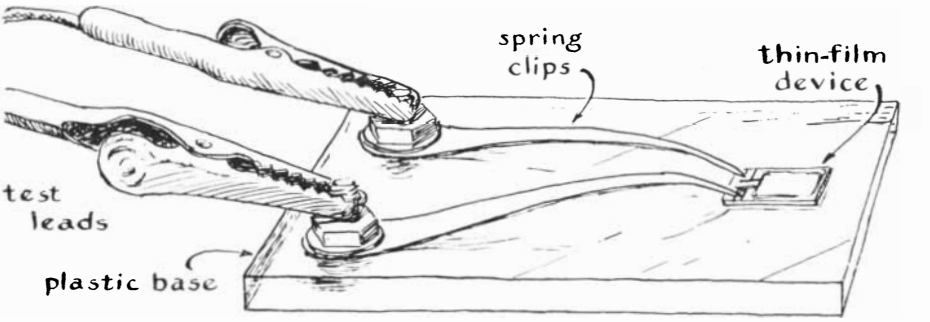
"The preparation of a field-effect transistor illustrates a typical experimental procedure. The substrate, which has been treated with nitric acid solution, is first coated with a film of cadmium sulfide. With distilled water prepare 500 milliliters of a stock solution containing .01 molar (.01 M) thiourea and .01 M cadmium chloride.

"Place the substrate inside a 250-milliliter beaker so that it rests diagonally against the side of the beaker. Cover the substrate with stock solution and slowly add concentrated ammonium hydroxide until the mixture turns faintly cloudy and then clears. Cover the beaker and put it in a double boiler. Heat the vessel slowly and boil for about 15 minutes. The contents of the beaker will turn yellow-orange, indicating the precipitation of cadmium sulfide.

"Pour off the contents and replace them with distilled water. Swab the substrate lightly with a tuft of absorbent cotton to remove adhering particles of cadmium sulfide. Rinse the substrate with distilled water. Repeat the entire procedure to double the thickness of the film, after which you can clean the beaker with hydrochloric acid.

"Bake the substrate in air at 500 degrees C. for 30 minutes. The color of the hot substrate will gradually change from yellow to red and, as it cools, to a deeper shade of orange. With the diamond point scribe and break the cooled glass into rectangular chips 1/4 inch wide and 1/2 inch long.

"The transistor requires two contacts that function as electrodes, one a source and the other a drain. The electrodes are conveniently made of indium, a soft metal that can be pressed into firm contact with the film. Indium is available from dealers in chemicals. Place a small pellet of indium on clean plate glass and roll it into thin foil with a short length of clean glass tubing. Transfer the foil to a yielding surface such as glossy white cardboard and, by pressing straight down with a sharp razor blade, cut the metal into strips about 1/32 inch wide and 1/4 inch long.



Fixture for testing thin-film devices

"With a sewing needle maneuver two of the strips to a clear portion of the paper so that they are parallel and spaced about 1/16 inch apart. Lay one of the coated chips over the strips so that the ends of the strips are even with one end of the chip. Press the chip firmly and evenly against the metal. The strips will adhere lightly to the film. Burnish them firmly into place by turning the chip strip side up on the plate glass, covering it with a glossy magazine cover and rubbing it with a fingernail. Place a small dab of conductive silver paste on the outer end of each indium strip [*see bottom illustration on opposite page*]. The dabs serve as terminals for connecting the electrode device to a power source.

"A layer of insulation is applied to the film and the indium strips in preparation for adding the third electrode, which is known as the gate. With a sewing needle apply a thin, uniform layer of vinyl cement by stroking the cement across the upper surface of the device. Do not coat the silver terminals. When the cement dries, apply a coat of silver paste over the insulation. Do not let the silver make contact with the cadmium sulfide film, the indium foils or the source and drain terminals. This completes the gate electrode.

"Finally, to protect the active region of the device coat the upper surface with a layer of silicone rubber that cures in air. This material is available from dealers in hardware. Leave one small region of the gate electrode exposed. This small area will be used for making electrical contact with the gate. Do not coat the source and drain terminals with rubber.

"To operate the device improvise a test fixture such as the one shown in the accompanying illustration [*above*] for holding the transistor and connecting it to a battery. Power is applied to the source and drain electrodes by a nine-volt transistor battery that is connected in series with a 10,000-ohm resistor and a 0–50 microammeter. If the transistor is reasonably good, the meter will indicate a current of about 10 microamperes. This is called the leakage current.

"Connect a one-megohm resistor between the gate electrode and the positive terminal of the battery. The positively charged gate will attract free carrier electrons into the cadmium sulfide film. Current through the film should rise to about 50 microamperes, indicating that the transistor is a so-called *N*-channel device and that it is operating in the enhancement mode. The gate electrode draws little current.

"If negative charge is now applied to the gate by transferring the one-megohm resistor to the negative terminal of the battery, current in the source-drain circuit should fall below 10 microamperes. The transistor is now operating in the depletion mode. I do not know why some homemade transistors work better than others. I suspect that their performance may be related to the crystalline structure of the films.

"Capacitors can be made by sandwiching insulation between films, resistors by etching away portions of film to form narrow conducting paths, photocells by doping cadmium sulfide with trace amounts of silver, copper or manganese. Films of zinc sulfide fluoresce strongly. Of course, devices are available on the market that work better than those one can make at home, but mine are better playthings.

"Certain hazards must be mentioned. Metallic salts and acids are toxic. Work either in a fume hood or outdoors when you spray chemicals onto a hot substrate. Wear gloves and an apron of neoprene when you handle acids. Remember that chemicals are hazardous and handle them accordingly."

BOOKS

Three lives in the palmy days of modern physics: Gamow, Fermi and Oppenheimer

by Philip Morrison

MY WORLD LINE: AN INFORMAL AUTOBIOGRAPHY, by George Gamow. The Viking Press ($5.95). ENRICO FERMI, PHYSICIST, by Emilio Segrè. The University of Chicago Press ($6.95). THE SWIFT YEARS: THE ROBERT OPPENHEIMER STORY, by Peter Michelmore. Dodd, Mead & Company ($6.95). Read the titles again, with care. They are sharply revealing. The Gamow book, his own words, is a fragmentary, brief, delightful book, "a collection of short stories, all of them pertaining to me, and all of them completely true." (But don't bet on it.) The stories are deeper than true, because they are his own, and they are well told. The Fermi book, both documented and intimate, a thorough extension of Professor Segrè's fascinating annotations to the collected works of Fermi, is a physicist's book about a physicist, knowing, analytical, an indispensable piece of history. The Oppenheimer book, like its title, is an expert journalist's mixture, partly beautiful and partly banal. But of the whole corpus of books about this man, so loved and so hated, it comes closest to understanding him. Michelmore has never heard, it seems, of Bernhard Riemann, and he is naïve about 1938 America and about the nature of research. But he knows men. His Oppenheimer is not the grand myth we see on the stage. Nor is he the man in travail we see in the admirable Philip Stern study in depth and in the transcribed 1954 remake of the notorious 1633 hearing in Rome before the Commissary-General of the Inquisition, which admonished the greatest Italian physicist of them all. The most moving words in the book were quoted by "Opje" himself to Vannevar Bush a fortnight before the Trinity test, from the Sanskrit of Bhartrihari:

In the forest, in battle, in the midst
of arrows, javelins, fire
Out on the great sea, at the precipice
edge in the mountains
In sleep, in delirium, in deep trouble
The good deeds a man has done
before defend him.

Michelmore's book, based on many interviews and on telling use of the personal papers of Oppenheimer, makes those good deeds real.

What can we learn about these men? Their first introduction to physics? For Gamow, born in Odessa by Cesarean section (they had decided to cut him to pieces to save his mother, but a surgeon from Moscow intervened), it was perhaps Jules Verne, and trying to make a small bell "like the jingle bells on Santa Claus's reindeer" work from a battery. He led his school class, as was expected of the grandson of a Metropolitan. By the time he was 13 or 14, Odessa was a much used battleground, and when "reading a book on Euclidean geometry near a window in our apartment... the window pane was suddenly shattered by the shock wave of an artillery shell." But "the experiment which made me a scientist" was an investigation of Russian Orthodox dogma made with a five-and-ten-cent-store microscope, examining a crumb of bread that had been dipped into the already transubstantiated red wine during Holy Communion. It matched a control breadcrumb more than it did a piece of his own skin, a control for the body of Christ.

Fermi was easily the best student in his *ginnasio* class, with a prodigious memory and Latin strong enough to study an 1840 two-volume text in physics he bought in the secondhand stalls at the age of 14. He had built electric motors, model airplanes and other toys with his brother, a year his senior. By 13 his genius was evident. An engineer colleague of Fermi's railway-accountant father found the boy "truly a prodigy" first when young Fermi in two months finished a projective-geometry text the man had lent him, solving all the problems. The harder problems "would have taken me too much time," Ingegnère Amidei recollects, but the boy could give him all the proofs, although he used only the little free time left from the arduous compulsory studies of the day.

Oppenheimer too was a star pupil. The frail, overprotected, bookish boy was a devoted and serious rock collector (he wrote a mineralogical paper at age 12 or so), then a home chemist. By 16 a skillful and dashing yachtsman on Great South Bay, he named his 28-foot sloop—the elder Oppenheimers were wealthy and indulgent—*Trimethy*, after the fishy-smelling substance trimethylamine. He entered Harvard as a chemistry major but switched as a sophomore to physics: "the study of order, of regularity, of what makes matter harmonious."

All three men studied at Max Born's institute at Göttingen, Fermi in 1923, Oppenheimer in 1926, Gamow in 1928. Fermi worked there nearly alone. Born was cordial enough to him, but somehow the shy Fermi made no mark in that heady world. Oppenheimer fared better. Max Born and the American, who was a cultivated as well as brilliant young man, worked so closely together that other students were put off by the dominance of the visitor.

Gamow tells the most remarkable story. He somehow "was not engulfed in this whirlpool" of the quantum theory of atoms and molecules. It was too mature for him even then, "requiring oodles and oodles of 'exing.'... I was always losing my way in [the calculus]." But on his first day in the Göttingen physics library he read a new paper of Rutherford's on alpha-scattering by uranium. Even the high-energy alpha particles emitted by polonium are scattered according to the laws of pure Coulomb electrostatic repulsion, although they have more energy than the alphas emitted by uranium. How is it that high-energy particles cannot get across the Coulomb barrier from outside whereas the low-energy ones from inside are able to leak out, albeit slowly? Rutherford had a curious explanation involving electron capture and loss. Gamow knew "before I closed the magazine" what

happens. The barrier-tunneling effect was in the wave mechanics, and it was well known. Typically Gamow explains that he could not do the very simple integral that arose and sought help from another Russian at Göttingen. (Edward Condon and Ronald Gurney were to publish the same generally correct account a few days before Gamow's.) Gamow went to Copenhagen for a day on the way back to Leningrad to meet Bohr and stayed for the year.

What of the bomb? It is not necessary to give an account here of Oppenheimer's profound involvement, nor of Fermi's. The famous scene of Fermi at the first bomb test releasing torn bits of paper to meaure the air displacement as the shock wave passed is now classic. He had prepared a simple numerical table to give him the yield of the bomb immediately from the motion of the paper. There is the man: a mind that touched the essence of the phenomenon in utter simplicity, a hand able to make and use the instrument just fit for the task, a detachment—even a coldness—of the spirit.

Surprisingly, Gamow comes off a bit cold as well, in spite of his constant wit, his jocular view of life, his ingenuousness and his modesty. He calmly relates one catastrophe after another to friends or teachers—such as the cosmologist Alexander Friedmann, who died of pneumonia after a flight in a free meteorological balloon. "This ruined my plans to continue my work on relativistic cosmology." No doubt of it, Russia was a grim place to grow up in from 1905 to the 1920's. Gamow, working in America from 1934 on, was not cleared for the atomic bomb project; he recalls independently inventing explosive lenses for implosion as consultant to some Navy wartime studies, and finding that no one seemed very anxious to give him any priority for the work. For the fusion bomb he came to Los Alamos in 1948, to work beside his good friends Edward Teller and Stanislaw Ulam, "the mother and father of the H-bomb.... Well, so much about military activities."

All three of these books have many photographs that are both evocative and revealing. George Gamow's own drawings add special interest. His little book, incomplete though it is, is irresistibly readable.

SEEDS OF CHANGE: THE GREEN REVOLUTION AND DEVELOPMENT IN THE 1970's, by Lester R. Brown. Praeger Publishers ($6.95). THE HUNGRY FUTURE, by René Dumont and Bernard Rosier. Translated from the French by

Rosamund Linell and R. B. Sutcliffe. Frederick A. Praeger, Publishers ($6.95). If the greatest problem of our time—perhaps thermonuclear war is less cosmopolitan—is too many people with too little to eat, it has at least received adequate treatment in books. A flood of books in the past few years have discussed the topic, some ably, one even forecasting U.S. use of famine relief as a strategic weapon, a volume unmatched for inhumanity since the first widely circulated thinking of Dr. Herman Kahn. Here are the two that have seemed best: they are each reasoned, informed accounts. The two confront each other in many ways.

Professor Dumont, of the National Institute of Agronomy in Paris, has for 35 years seen and studied the problems of farmers all over the world, particularly in developing countries. His young colleague is professor at the Institute of Political Studies. Their book is wide in scope, giving the background of population history and the statistics of food production. Brown's is a more topical, less data-filled, briefer volume of argument. Dumont is *engagé*, angry at "the filthy war" in Vietnam, explicitly aiming for social justice. He has been an adviser to governments of the left and the center, from the Central African Republic to Tunisia, Cuba and North Vietnam, and to the United Nations. But he is a realist. He seeks the end of hunger rather than show. He has visited more than one little-used jet airfield whose cost would pay for helping many a peasant to better farming.

Brown, still in his thirties, is the American Establishment at its technical best. He was an adviser and planner in the Kennedy-Johnson Department of Agriculture. He too is concerned with results, a pragmatist. Like Dumont, he recognizes that science and technique are necessary but that any true solution involves not merely a technical "quick fix" but rather an entire system made to function. Dumont sees, for example, the failure of China's Great Leap Forward as a "disparity between the enormous land... and water installations, and the low increase in the production of chemical fertilizer." Yet China has gone beyond the danger of famine, whereas India has not quite.

Within the past two or three years the new strains of rice and wheat, developed in Mexico and the Philippines in research institutes financed by American foundations, have been adopted widely in Asia, except in China. By the end of the crop year just past, about 15 percent of the wheatlands and ricelands in those countries, particularly in India and Pakistan, was planted to the new varieties. The crop plant that grows traditionally in those fields is tall and thin-strawed; it can compete with weeds and keep above the floodwaters. The new strains are dwarfed, a Japanese innovation. Given plenty of fertilizer, the traditional strains do set heavy heads of grain. Then, however, they are top-heavy, and they lodge, or fall over, before their grain is ripe. The new strains yield 20 pounds of new grain for each pound of nitrogen, and will absorb three times as much nitrogen before lodging.

The results are spectacular. Conservatively the non-Chinese Asian food supply has grown by something like two or three times the American annual export of grain during the years of monsoon failure, by say 17 million tons per year. More new grain than new mouths! That is the green revolution. There is a "yield takeoff" in a whole set of countries, particularly Mexico, India, Pakistan and Turkey, and at least tentatively in Indonesia. The farmers using the new grains make a bigger net income. They spend much more for fertilizer and water supply, but they more than make it up in product. The new grains are not yet wholly acceptable as food: the wheat is too red, the rice too brittle. They sell at a lower price. They still face the danger of some single widespread disease like the Irish potato blight; they do require investment in tube wells and diesel pumps and more irrigation. They promise multiple crops on many acres, and therefore imply more mechanization to avoid delays due to labor shortage at peak periods. These are the very problems that fade before technical change.

Yield, however, is not a panacea. The full bowl must be brought to the hungry. Generally they must be able to pay. They must have jobs. Otherwise the result will be merely a struggle between the rich farmers and the poor displaced ones, between landowners and city poor. Market surpluses will grow, ironically hand in hand with incipient famine. The green revolution can buy mankind time, if the next 15 years somehow manage to deal not with the food and population problem but with the jobs and population problem.

It doesn't look easy. Yet for Brown it seems within reach. He views the multinational corporations, able swiftly to transfer new technology across borders, as a key to development. Did these great firms not bring nitrogen fertilizer to India? Did they not use the new low-pressure centrifugal process of the M. W. Kellogg Company to build large, cheap plant for cheaper nitrogen? They did. Yet they did so only after the U.S. Government had at their insistence broken the statutory determination of famine-stricken India to keep fertilizer production in the public sector.

There is another worry. The U.S.S.R. and China do not appear in even one single index entry in *Seeds of Change* (although there are a couple of unindexed sentences about poor Chinese statistics). Dumont, however, treats the problems of agriculture in the Socialist world with much concern, and a good deal of sharp criticism, particularly for the Stalinist mode of heavy-handed compulsory and premature collectivization that impaired production without much compensatory social gain. Brown instead ignores a third of the world, a fact that, although understandable, lowers one's confidence in the reliability of his calm, objective expertise.

Dumont has a bitter dedication: "To the children of backward countries who never attain their full promise, or who have died of kwashiorkor, because the fish meal which might have saved them has fed the chickens gorged by the rich." Yet chicken in America today is not the food of the rich; it is the best-buy protein of the working people. This charge seems a little wide of the mark. The system is more ironic and more complex than that. From Dumont's book one comes away realizing that he shares with Brown the view that profound agrarian reform is less a moral issue than an economic measure.

Brown feels that "American foundations and the largest and best-equipped multinational agribusiness corporations" uniquely qualify the U.S. to help eradicate global hunger. Yet nationalism and ownership, foreign troops and foreign laws and foreign investments are issues of life and death today, and not only in Southeast Asia. A kind of agricultural Marshall Plan by way of the World Bank is Brown's U.S.-centered scheme for the future. The French experts hope rather for a growth through an extended World Food Program to a controlled international market in the cash crops, and eventually to a $40-billion-a-year International Solidarity Fund, a universal war on hunger, with an arbitration scheme for sanctions against conditions of inefficient use. Some arbitrators are even named, for example Mahalanobis, Myrdal and Balogh.

There is hope in both books. What will happen very likely lies in between.

U.S. interests are probably too strong to be resisted except by fortunate heroes, but their selfishness can be moderated. The revolution of varieties and pumps is the easier one; the equally requisite sharing of power, the harder one, and further off.

Brown criticizes almost no one. Only the beet-sugar industry receives his censure. It sells its uneconomic but identical product at from three to five times the world sugar price, protected from the poor cane-growing competitors by a tariff wall we consumers pay for each year in all the developed countries. There is now too much sugar in our world, too much both for those who overeat it and for those who must eat by selling it. War made the strange sugar anomaly; we come back to where we started. Perhaps war, after all, remains the greatest problem.

ANTIBODIES AND IMMUNITY, by G. J. V. Nossal. Basic Books, Inc., Publishers ($5.95). In 1961 a Melbourne immunologist, then working in London, was on the track of leukemia. The thymus gland at the base of the throat was known to be involved in leukemic mice (another Melbourne finding). He proceeded to remove the organ from some tiny newborn mice, virtuoso veterinary surgery in patients not an inch long and weighing less than a gram. Adult mice were known to show no ill effects from thymus removal. His newborn mice, however, did poorly. They became runted and sickly and died of a wasting disease in a month or two. A thymus graft would save them. The function of the thymus, long an enigma, was ready to be revealed.

Those mice died of a want of internal know-how. Germ-free mice, born and reared in a completely germ-free environment, do not waste away without a thymus. What besets the mouse is infection by microorganisms. The thymus, together with the spleen, the lymph nodes and the bone marrow, forms a remarkable system, the system that produces the arsenal of well-informed globulins: smallish protein molecules of the bloodstream that stand ready to coat the working surfaces of a foreign cell or virus or otherwise defuse all novel molecular structures introduced into the body. Any beast at least as complex as the ancient lamprey has some kind of immune system. Otherwise how can such coordinated organisms survive in a world awash with single cells, all capable of being multiplied in a few days by 10 powers of 10 in the rich media

that bathe the living cells of their host?

It used to be thought that the ingeniously made double strings of amino acids that make up the five classes and the million or so distinct forms of antibody molecule were a protective system that learned by doing. Each antibody somehow folded around a new antigen (the name given the inducing molecular invader) and in its folded shape became a template for making more specific antibody. Then in 1957 Sir Macfarlane Burnet of Melbourne and David Talmage of the University of Colorado saw independently that it was different. The immune system is born with the genetic ability to synthesize all the antibodies it can ever make. Indeed, one can find a very small amount of antibody against every one of a large variety of bacteria and viruses in any healthy uninfected animal.

The induction is a matter of natural selection: an antigen stimulus picks out just that cell which knows how to form that particular antibody. The cell is coated and somehow stimulated within the thymus to reproduce vigorously. Its entire clone of descendants retains that skill. The protein formed has an amino acid sequence that depends on the cell selected. This is a deathblow to any template theory, but it is a natural part of our present picture of information storage in living cells. Antibody can be overwhelmed by massive doses of antigen, or in an embryo so young that the immune system has not fully developed. One has killed, or at least damaged, the theory says, the parent cells of the proper clone. Those cells are small, wandering lymphocytes, born in bone marrow, spread in lymph, grown in the spleen. They are often found with "cytoplasm so scanty that under the light microscope the cell looks like a naked nucleus." The parent antibody cell is only an inherited recipe for the inhibition of specific antigen structures.

"From its very beginnings, immunology has been an outstanding example of the interwoven nature of so-called pure and applied research." Of all the domains of molecular biology, this one has the deepest present impact on patients. We could expect diseases of the entire immune system: outlaw clones, antibodies that attack the very self. A few such autoimmune diseases are known; even diabetes and some forms of schizophrenia may turn out to be examples. There may be antigens to specific tumors. Organ transplantation, whether of kidney, liver, heart or skin, depends on the gross means now known for chemically disabling antibody formation by slowing lymphocyte division in order to prevent rejection of the foreign tissue. A more specific approach aimed at inducing tolerance to the new graft is the present goal.

This brief, clear, candid book—uncertainties and puzzles abound—reads like a set of particularly exciting lectures, complete with blackboard diagrams, given by a devoted and expert friend to a small, lucky group. Dr. Nossal is indeed an expert, the present director of the Walter and Eliza Hall Institute of Medical Research in Melbourne, his office "just 10 yards" from the laboratory where Macfarlane Burnet, his predecessor and still a leader in the field, works. This book is Publication No. 1217 of that institute; it continues the high tradition of good writing and humane inference set long ago by Sir Macfarlane himself. Australian science has and deserves a good press these days: the kangaroos bounding down the long lines of radio dishes make excellent television pictures. Yet virology and immunology at the Melbourne institute, with their sensible, prompt and energetic application to public health, and the responsibility for popular exposition taken by its leading figures, set a standard now four decades old which matches that on any continent.

ANATOMY OF AN EXPEDITION, by Henry W. Menard. McGraw-Hill Book Company ($6.95). "One of the great delights of a small ship is that it is almost completely dark. When you come on deck only the great stars... can be seen by light-sated eyes.... In ten minutes ... stars are everywhere. The Milky Way deserves its name, the Southern Cross shines brightly.... We saw several galaxies... with the naked eye and verified them with binoculars. The sky was cloudless, the air warm, the sea calm, and the ship glided in beauty."

Professor Menard is a submarine geologist of the Scripps Institution of Oceanography who narrates most joyfully his personal account of a 1967 expedition, called Nova for its destination, the sea floor around New Caledonia, New Britain and New Guinea, in that oceanographically neglected and romantic region called Melanesia. Nova used two ships: *Horizon*, a curiously happy ship, welded together in wartime Texas to displace 800 tons, run by a crew of about 18 men with a party of a dozen research people; *Argo*, about as old as *Horizon*, more than twice as big, much roomier, "a more professional ship but also a colder one." They were after big geological fish—the evidence for the spreading of the sea floor in that unique geography, where a jumble of islands stands "part continent and part ocean basin." They trawled for their truths, streaming magnetometers and hydrophones at the end of 1,000 feet of steel cable, where the attack of sharks deeply grooves the steel instrument case. They trawl and they listen, while they steam ahead at five miles per hour to reduce their own noise. The ships are not passive; they snap for bottom echoes, pop air guns and make giant sparks for subbottom profiles. They sometimes stop and dredge the bottom just as the *Challenger* did a century ago, and they work by pairs for deep seismic surveys, one shooting and one listening, with high-explosive charges whose size grows from half a pound to 300 pounds as the ships move apart. They found plenty too. Such an expedition is faster at publishing its results these days; they are out within about a year with much software and computer aid. (The books are not as beautiful as they once were.) Neat sketch maps show the "high and hot" Fiji Plateau, which Nova surveyed in some detail for the first time, finding that it loses inner heat in places at rates up to five times the widespread terrestrial mean, "the highest heat flow of any known large area on earth."

The whole story of continental drift, sea-floor spreading, atolls and seamounts adds up to a triumph for the oceanographer. Professor Menard shares a quiet excitement in the modest Nova addition to this Scheherazade's account, far from ended. His book is not, however, a regional version of that remarkable subject matter, nor even a history of oceanographic methods and traditions. Those things are there, with graphs, Victorian savants and all, in word and picture. His book is rather a display of the life of a scientist, a real man who invents, plans, enjoys, endures, teaches, fails, guesses and sails on. Nova, like most science and some children, was conceived well before the proposal. Nova cost about $4,000 per ship-day for the best part of a year. Its budget was heavily cut, but they could make do. (Captain Cook sailed in the *Endeavour* at about the same real cost.) A sedentary scientist must admire even the very first travel in support of Nova. It was a month's jaunt by Menard (and his wife) to Suva and Sydney and other places, to consult the local experts about any unpublished charts and to line up support, friends and ports of call. The planning and

preparation take longer than the voyage, even at an experienced center such as Scripps. Why? Well, Brisbane Harbor wouldn't admit explosives. Magnetometers, the best in the world and perfect when tested, fail at sea. (Almost everything fails at sea: steel rusts to weakness, circuits vibrate to bits, streams of water "from the drinking fountain or into the urinal . . . behave in defiance of gravity.") An ordinary traveler, even a tanker sailor, "cannot imagine what life is like on a small ship. . . . The ship moves. The contents move. . . . Every aspect of work is affected by constant motion."

About 180 people at sea "caught the fever called Nova" at one time or another. They were cooks and graduate students, a reporter from *Paris-Match*, oilers and captains, several women scientists. As for graduate students, "on average they are the cheapest, highest-quality and most eager workers in the world." They do much, they learn much, even how to be "chief scientist for a day." A $10,000 deep-sea camera of Russian make lies for the grabbing somewhere in the Mariana Trench, and now there is another one from Scripps. The endless problem of knowing your place at sea forms the leitmotiv of almost every chapter, although "the satellite navigation system still has not been delivered—but there is no point in starting this book over again." *Argo* had already started around the world, on an expedition named Circe, by the spring of 1968.

If *The Double Helix* was a venomous analogue to a cruel and witty play of Oscar Wilde's, with an endearing rascally hero, this book is like some happy novel of a big family—cheerful, affirmative, crowded with people lively in thought. The title fails: "anatomy" evokes scalpel and formaldehyde. In this enjoyable story Nova is alive and well.

MACHINA EX DEO: ESSAYS IN THE DYNAMISM OF WESTERN CULTURE, by Lynn White, Jr. The M.I.T. Press ($5.95). Modish subtitle aside, this brief 1969 collection of 11 essays written over three decades by a thoughtful historian is a book too valuable to overlook. Witches as a periodic social and individual necessity, da Vinci as an engineer who painted when he was broke, the origins and effects of alphabetization, stirrups, counterweighted artillery and the nomination of St. Francis of Assisi as patron saint of ecologists are a sample of what is here. Professor White is provocative, clear and deep. (A list of references is sorely missed.)

INDEX OF ADVERTISERS

JUNE 1970

BIBLIOGRAPHY

Readers interested in further reading on the subjects covered by articles in this issue may find the lists below helpful.

"SILENT MAJORITIES" AND THE VIETNAM WAR

THE NATURE OF BELIEF SYSTEMS IN MASS PUBLICS. Philip E. Converse in *Ideology and Discontent*, edited by David E. Apter. The Free Press of Glencoe, 1964.

THE RANDOM PROBE: A TECHNIQUE FOR EVALUATING THE VALIDITY OF CLOSED QUESTIONS. Howard Schuman in *American Sociological Review*, Vol. 31, No. 2, pages 218–222; April, 1966.

AMERICAN PUBLIC OPINION ABOUT VIETNAM. John P. Robinson and Solomon G. Jacobson in *Peace Research Society (International) Papers*, Vol. 10, pages 63–79; June, 1968.

THE ORIGIN OF GALAXIES

THE FORMATION OF STARS AND GALAXIES: UNIFIED HYPOTHESES. David Layzer in *Annual Review of Astronomy and Astrophysics: Vol. II*, edited by Leo Goldberg, Armin J. Deutsch and David Layzer. Annual Reviews, Inc., 1964.

THE BLACK-BODY RADIATION CONTENT OF THE UNIVERSE AND THE FORMATION OF GALAXIES. P. J. E. Peebles in *The Astrophysical Journal*, Vol. 142, No. 4, pages 1317–1326; November 15, 1965.

COSMIC BLACK-BODY RADIATION AND GALAXY FORMATION. Joseph Silk in *The Astrophysical Journal*, Vol. 151, No. 1, Part 2, pages 459–471; February, 1968.

THE FORMATION AND EARLY DYNAMICAL HISTORY OF GALAXIES. G. B. Field in *Stars and Stellar Systems, Vol. IX: Galaxies and the Universe*, edited by A. Sandage and M. Sandage. The University of Chicago Press, in press.

SOME CURRENT PROBLEMS IN GALAXY FORMATION. M. J. Rees in *Italian Physical Society Proceedings of the International School of Physics Enrico Fermi, Course 47: Relativity and Cosmology*, edited by R. K. Sachs. Academic Press, in press.

GENETIC REPRESSORS

ISOLATION OF THE *LAC* REPRESSOR. Walter Gilbert and Benno Müller-Hill in *Proceedings of the National Academy of Sciences*, Vol. 56, No. 6, pages 1891–1898; December, 1966.

ISOLATION OF THE λ PHAGE REPRESSOR. Mark Ptashne in *Proceedings of the National Academy of Sciences*, Vol. 57, No. 2, pages 306–313; February, 1967.

SPECIFIC BINDING OF THE λ PHAGE REPRESSOR TO λ DNA. Mark Ptashne in *Nature*, Vol. 214, No. 5085, pages 232–234; April 15, 1967.

THE *LAC* OPERATOR IS DNA. Walter Gilbert and Benno Müller-Hill in *Proceedings of the National Academy of Sciences*, Vol. 58, No. 6, pages 2415–2421; December, 1967.

COMPUTER DISPLAYS

SKETCHPAD—A MAN-MACHINE GRAPHICAL COMMUNICATION SYSTEM. I. E. Sutherland in *AFIPS Conference Proceedings: Joint Computer Conference*, Vol. 23, pages 329–346; Spring, 1963.

ON THE DESIGN OF DISPLAY PROCESSORS. T. H. Myer and I. E. Sutherland in *Communications of the ACM*, Vol. 11, No. 6, pages 410–414; June, 1968.

A CLIPPING DIVIDER. Robert F. Sproull and Ivan E. Sutherland in *AFIPS Conference Proceedings: Joint Computer Conference*, Vol. 33, Part I, pages 765–776; Fall, 1968.

A DISPLAY PROCESSOR DESIGN. T. H. Myer, I. E. Sutherland, M. K. Vosbury and R. W. Watson in *AFIPS Conference Proceedings: Joint Computer Conference*, Vol. 35, pages 209–217; Fall, 1969.

HOW SNAKES MOVE

LOCOMOTION WITHOUT LIMBS. Carl Gans in *Natural History*, Vol. 75, No. 2, pages 10–17; February, 1966.

LOCOMOTION WITHOUT LIMBS: PART II. Carl Gans in *Natural History*, Vol. 75, No. 3, pages 36–41; March, 1966.

ANIMAL LOCOMOTION. Sir James Gray. George Weidenfeld & Nicholson, Ltd, Publishers, 1968.

SIDEWINDING AND JUMPING PROGRESSION OF VIPERS. Carl Gans and H. Mendelssohn in *Proceedings of the Second International Symposium on Animals and Plant Toxins*, in press.

NEOGLACIATION

GLACIER VARIATIONS AND CLIMATIC FLUCTUATIONS. H. W:son Ahlmann. The American Geographical Society, 1953.

CHRONOLOGY OF NEOGLACIATIONS IN THE NORTH AMERICAN CORDILLERA. Stephen C. Porter and George H. Denton in *American Journal of Science*, Vol. 265, No. 3, pages 177–210; March, 1967.

AN ARCHAIC INDIAN CEMETERY IN NEWFOUNDLAND

THE CULTURAL AFFINITIES OF THE NEWFOUNDLAND DORSET ESKIMO. Elmer Harp, Jr., in *National Museum of Canada*, Bulletin No. 200, Anthropological Series No. 67, pages 1–183; 1964.

EVIDENCE OF BOREAL ARCHAIC CULTURE IN SOUTHERN LABRADOR AND NEWFOUNDLAND. Elmer Harp, Jr., in *National Museum of Canada*, Bulletin No. 193, Contributions to Anthropology 1961–62, Part I, pages 184–261; 1964.

FIVE PREHISTORIC BURIALS FROM PORT AU CHOIX, NEWFOUNDLAND. Elmer Harp, Jr., and David R. Hughes in *Polar Notes*, No. 8, pages 1–47; June, 1968.

THE CLOCK OF THE MALARIA PARASITE

RHYTHMIC FUNCTIONS IN THE LIVING SYSTEM. William Wolf, Conference Editor, in *Annals of the New York Academy of Sciences*, Vol. 98, Art. 4, pages 753–1326; October 30, 1962.

MALARIA PARASITES AND OTHER HAEMOSPORIDIA. P. C. C. Garnham. Blackwell Scientific Publications, 1966.

24- AND 48-CYCLES OF MALARIA PARASITES IN THE BLOOD: THEIR PURPOSE, PRODUCTION AND CONTROL. Frank Hawking, Michael J. Worms and Kenneth Gammage in *Transactions of the Royal Society of Tropical Medicine and Hygiene*, Vol. 62, No. 6, pages 731–760; 1968.

MATHEMATICAL GAMES

GEOMETRY REVISITED. H. S. M. Coxeter and Samuel L. Greitzer. Random House, 1967.

5-CON TRIANGLES. Richard G. Pawley in *The Mathematics Teacher*, Vol. 60, No. 5, pages 438–443; May, 1967.

THE AMATEUR SCIENTIST

BOTANY: AN INTRODUCTION TO PLANT SCIENCE. Wilfred W. Robbins, T. Elliot Weier and C. Ralph Stocking. John Wiley & Sons Inc., 1964.

THIN FILM MICROELECTRONICS: THE PREPARATION AND PROPERTIES OF COMPONENTS AND CIRCUIT ARRAYS. Edited by L. Holland. John Wiley & Sons Inc., 1965.

Unite your business.

Now you can centralize — *and control* — the data your managers need to run every facet of your business. You can make your data processing system work smoothly and effectively in the mainstream of your companywide operations. How? Through General Electric United Systems.

The management information systems of the 70's.

Your computers will no longer serve just the needs of single functions, units and locations. They will unite through powerful data communications to serve your total business. (About 75% of all GE computer systems in operation have communications capability, vs. less than 50% for the industry as a whole.) And proven data base management techniques will let you organize and analyze millions of facts from many operations into meaningful business information. Result: faster, more profitable business decisions.

Choose the GE-600 Line System that's the right size for your business . . . then grow it in modules as your needs for it grow. Unite your business, your computers and your people now by calling your nearest GE Information Systems Sales Representative. Or write to Section 290-78, General Electric Co., 1 River Road, Schenectady, N.Y. 12305.

GENERAL ELECTRIC

General Electric Computers: The United Systems of the 70's

What's new from Monsanto and AstroTurf?

THE NO-SWEAT LAWN.

Picadilly Square apartment complex, Denver, Colorado.

No mower will ever touch this lawn. No hose will water it. No one will ever feed, weed or seed it. That's because you're looking at AstroTurf landscaping surface, the care-free, man-made lawn.

This newest member of the AstroTurf family is made of durable polyethylene. Tested for over three years under the searing Florida sun, AstroTurf lawns stay as inviting as the day they were installed.

And AstroTurf landscaping surfaces can be installed at almost any time of the year. So new buildings can look their best as soon as they are finished. No wonder more and more smart developers are turning to AstroTurf landscaping surfaces.

What's good for the rapidly growing apartment and condominium market is great for service stations, motels, hotels, mobile home parks, marinas, restaurants—and even highway median strips. The city of Lawndale, California recently installed nearly two miles of AstroTurf on its traffic median strips.

Find out how AstroTurf landscaping surfaces can serve you. Write: Dept. 327D, Monsanto Company, 800 N. Lindbergh Blvd., St. Louis, Missouri 63166.

AstroTurf®

LANDSCAPING SURFACES BY Monsanto