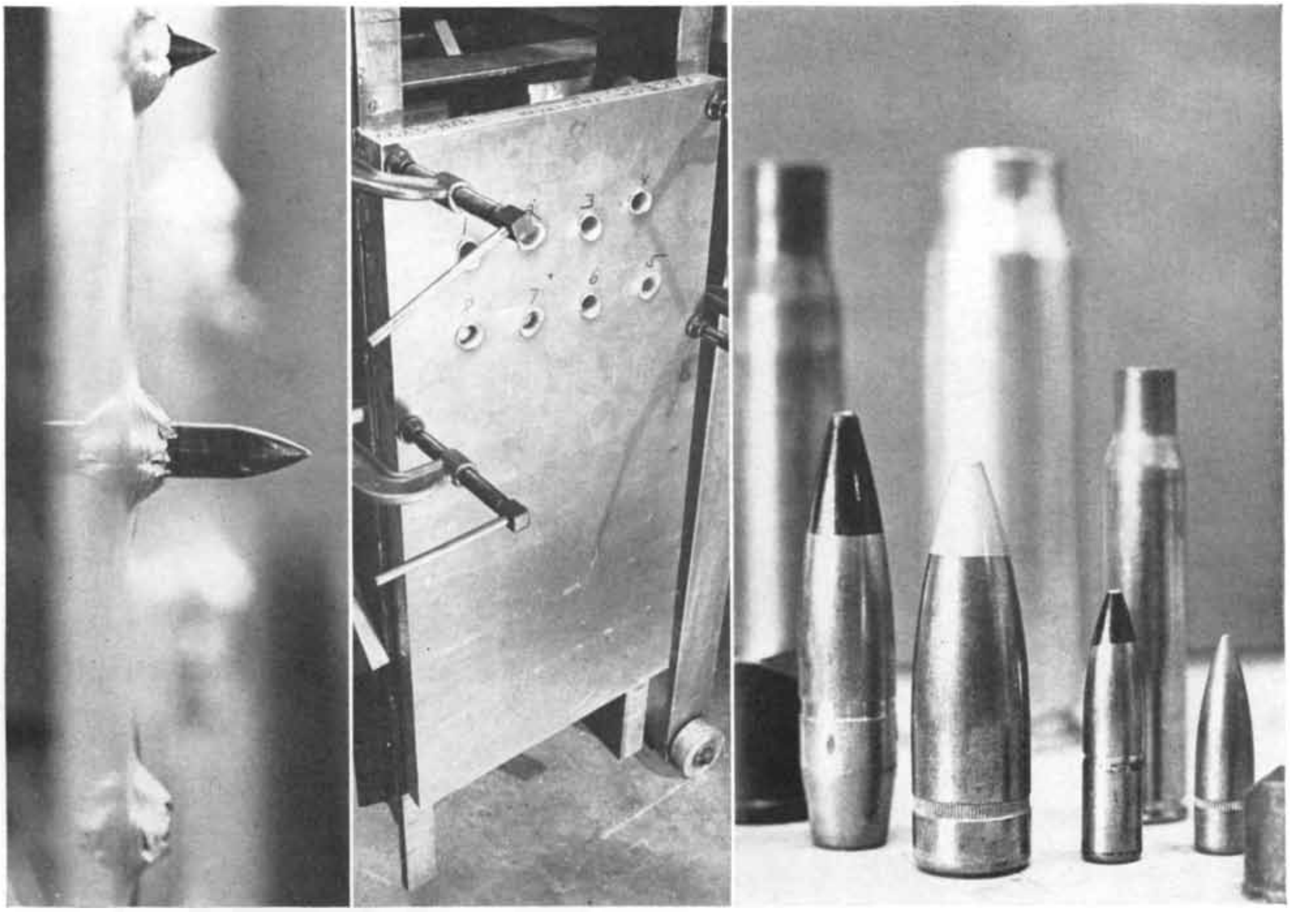
SCIENTIFIC AMERICAN

PULSE-CODE MODULATION

SEVENTY-FIVE CENTS

March 1968

London Citibankers with the Houses of Parliament in background.



Alcoa research helps make aluminum armor the logical choice.

We've been testing armor alloys and the effects of armor design mechanics for more than five years at Alcoa's Ballistic Laboratory. We've shot at hundreds of aluminum alloys and learned more than how to make holes in armor. Alcoa has learned how not to make holes in armor. We're helping develop new, lighter weight armor materials that defeat projectiles. We have learned to predict the resistance to penetration of plate, forgings, extrusions and castings based on the mechanical properties of the alloy, the thickness of the armor, and the angle of incidence.

Since building our laboratory to Army Ordnance standards, we've fired thousands and thousands of rounds of armor piercing and ball projectiles.

An electronic system is used to calculate impact velocity, and a computer program simplifies storing and correlating data.

Alcoa® Aluminum armor is used more and more, because when all the factors are considered (protection, weight, cost and ease of fabrication) it is the logical choice. It makes today's fast-moving, air-droppable military vehicles strong but light. For further information, write Aluminum Company of America, 550-C Alcoa Building, Pittsburgh, Pa. 15219.

Change for the better with Alcoa Aluminum

ALCOA

Established 1845 **SCIENTIFIC AMERICAN** March 1968 Volume 218 Number 3

ARTICLES

DEPARTMENTS

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. COPYRIGHT © 1968 BY SCIENTIFIC AMERICAN, INC. ALL RIGHTS RESERVED. 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: $8 PER YEAR.

THE COVER

The photograph on the cover can profitably be viewed from two distances: the normal reading distance and a distance of 10 feet or more. At close range the picture is a pattern of light and dark bands; at longer range it is a dazzling smile. (The small reproduction above has the latter effect.) The photograph is an enlargement of a small section of the television picture at left and second from the top on the opening page of the article "Pulse-Code Modulation" (*page 102*). It shows how the television picture is essentially made up of luminous lines that vary in intensity as the energy of the electron beam that paints them is modulated. Normally the beam is modulated by continuous changes in the amplitude of the voltage controlling the energy of the beam; in pulse-code modulation the amplitude of the voltage is changed in distinct steps that are specified by a digital code.

THE ILLUSTRATIONS

Cover photograph by Bell Telephone Laboratories

Page	Source
22–30	Jerome Kühl
31	John Bryson, Rapho-Guillumette
33	Paul S. Moorhead, Wistar Institute
34	Leonard Hayflick, Wistar Institute (*top and bottom left*); Fred C. Jensen, Wistar Institute (*bottom right*)
36–37	Allen Beechel (*top and bottom left*); Leonard Hayflick, Wistar Institute (*bottom right*)
38	University of Pennsylvania Museum
39	J. E. Dixon, University of Cambridge
40	P. M. Warren (*top*), J. D. Evans (*bottom*)
41	Eric Mose
42	Chieko Yoshida
43	Eric Mose
44–45	Chieko Yoshida
46	Eric Mose
56–57	Sol Mednick
58–62	Joan Starwood
67	Joan Starwood (*top*), Allen Beechel (*bottom*)
68–69	Allen Beechel and Chieko Yoshida
72	Joan Starwood
79	Chas. Slackman
80	Joan Starwood
81	Joan Starwood (*top*), Roy DeCarava (*bottom*)
82–84	Joan Starwood
86	Roy DeCarava
90	United Kingdom Atomic Energy Research Establishment
92–97	Bunji Tagawa
98	A. F. Tulinov, University of Moscow (*left*); Bunji Tagawa (*right*)
102	J. S. Mayo, Bell Telephone Laboratories
104–106	Jim Egleson
107	J. S. Mayo, Bell Telephone Laboratories
108	National Aeronautics and Space Administration
111	Norman R. Hatton, University of Toronto (*top*); J. J. B. Smith, University of Toronto (*bottom*)
112–113	Thomas Prentiss
114	E. Hitchcock, University College London
115	Thomas Prentiss
116	E. Hitchcock, University College London
118	Thomas Prentiss
121–122	Chieko Yoshida
123	John McKay, Atlas Computer Laboratory
124	University of Illinois (*top*), Chieko Yoshida (*bottom*)
128–134	Roger Hayward

No maybes about it.

First of all, there aren't a million gadgets to monkey with. One little knob controls both recording and playback.

A viewfinder sitting right on top of the camera tells you exactly what you're recording. So, unless you want to, you won't end up with people's heads cut off.

Audio and video signal levels are adjusted automatically. No watching meters or anything like that. (If you're really gung-ho and prefer stick shift, you can switch to manual operation. And go it alone.)

If you leave something out, all isn't lost. You can add material just by pressing a button.

The same goes for background music. Or narration.

In fact, the only thing you'll find hard to do with the new Sony Videocorder is botch things up.

The price? (We went easy on that too.) $1620 complete.*

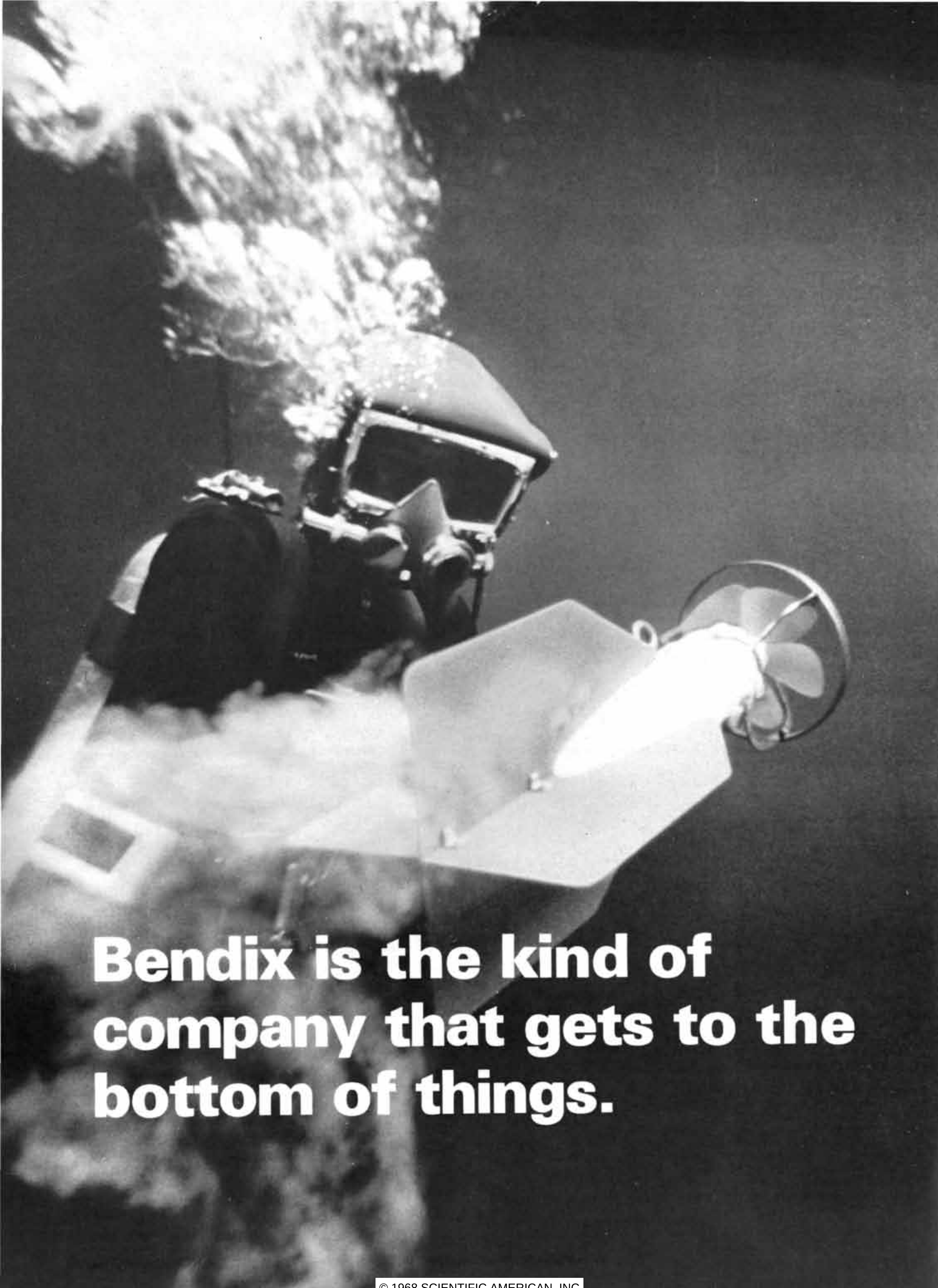
For more information, mail the coupon. What could be simpler?

SONY® VIDEOCORDER® (TCV 2110)

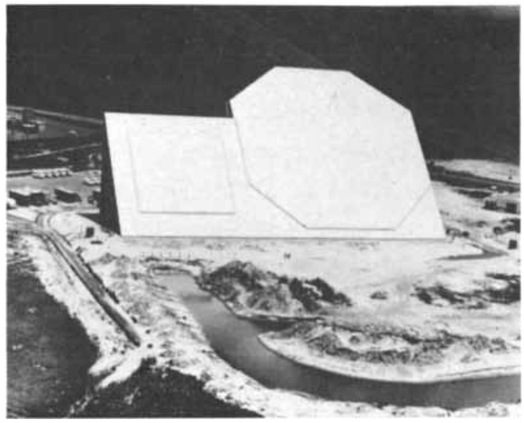
Name____________________
Title____________________
Organization____________________
Address____________________
City__________ State______ Zip______
Telephone____________________

Sony Corporation of America, B-68-31
47-47 Van Dam St., L.I.C., N.Y. 11101

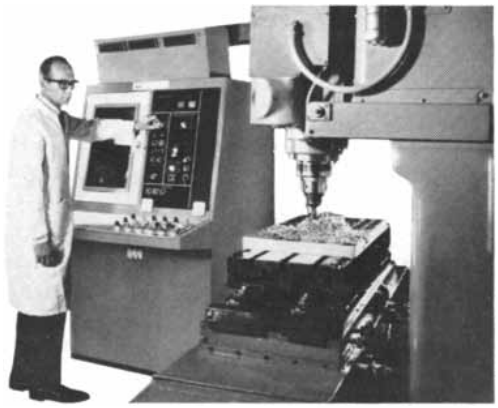
*Mfr's suggested retail price, includes camera ensemble with electronic viewfinder, monitor/recorder.


Bendix is the kind of company that gets to the bottom of things.

To get you to Europe in 3½ hours. Bendix will help guide the Anglo-French Concorde, when it starts carrying passengers across the Atlantic at 1,400 mph, sometime in 1971. The Concorde, likely to be the first supersonic transport, will have Bendix-designed automatic flight control systems, air data systems and other Bendix instrumentation.



To keep track of things in space. This 13-story structure houses a new kind of radar system that can detect, track, identify, catalog and report on hundreds of objects in space—simultaneously. It was designed and built by Bendix. Sponsoring agencies are the Electronic Systems Division and Rome Air Development Center of the Air Force Systems Command.



To make quality parts by the numbers. In action here is one of 12 new Bendix numerical control systems that operate drills, mills, lathes and other machine tools automatically, by interpreting numerical data on punched tape. Each of the systems will enable manufacturers to produce more and better parts in less time.

To give start-and-stop trucks another brake. Bendix is road-testing a new disc brake that will extend the advantages of disc brakes to trucks. It will mean proven performance, long life, self-adjustment and low-cost operation for both front and rear wheels.

To help tap ocean resources. This diver, installing Bendix equipment to measure the speed and direction of currents for harbor development, is one of a new breed of scientist/engineer at Bendix. He and others like him are helping industry and government utilize the wealth of the oceans. You'll find them making environmental studies for location of off-shore drilling, designing intercom units that allow divers to talk to each other under water, or searching for resources under the ocean floor. Our oceanics capability goes deep: from designing systems for sub-chasing torpedoes to showing ships how to rendezvous at an exact spot. In oceanics—as well as automation, aerospace, automotive and electronics—the heart of our business is creating new ideas and developing them to maximum usefulness. The Bendix Corporation / Fisher Bldg. / Detroit, Michigan 48202.

Bendix

Where ideas unlock the future

LETTERS

Sirs:

The facts on early maturation presented in J. M. Tanner's lucid article ["Earlier Maturation in Man"; Scientific American, January] have important implications for education to which I should like to draw attention.

While the age of physical maturity has fallen steadily, the age of university education has remained almost unchanged. As a result the campuses are now peopled by grown men and women yet are governed by traditions of organization and discipline which were made for adolescents. No wonder that those who remember themselves as striplings at college are outraged by the beards and the bosoms they see there now—and by the intransigence that is natural in the bearded and the bosomed, especially when they have to be treated as children. The university system as it is, historically, is two years out of step with the attitudes and emotions of contemporary students, and is only suited to the young who are now in high school.

A part of this difficulty will no doubt be repaired by giving university education more of the character of adult education. But the one change which is radical and inevitable is that students will go to college at least a year earlier, and in time two years earlier. This will require children to get through the present school curriculum one to two years faster than they do now, and the signs are that this is possible. Tanner's article of course is concerned only with body functions, and does not deal directly with brain and with intelligence. But even in the evidence for earlier physical maturation, these are grounds for seeing an implication that the brain is also maturing earlier. And there is at least a presumption from this (which has some support in tests) that intelligence quotient in its turn is rising in children.

At the least, Tanner's article should put an end to a classical fallacy about intelligence. This is the argument (popularized by Sir Ronald Fisher and copied by pundits ever since) which goes: *There is a positive correlation between the I.Q. of parents and the average I.Q. of their children. The average I.Q. of the children in large families is lower than in small families. Large families contribute more members to the next generation than do small families. Therefore the average I.Q. of the population must be falling.* There are several logical flaws in the statistical reasoning here. But what is more practical, Tanner's article proves in the most concrete way that the argument *must* be fallacious. For the argument can be repeated word for word about weight and about height, simply by putting the word "weight" or "height" in place of "I.Q." Yet as we see, the conclusion for weight and for height is demonstrably wrong.

J. Bronowski

Salk Institute for Biological Studies
La Jolla, Calif.

Sirs:

In their article "Non-Cantorian Set Theory" [Scientific American, December, 1967] Paul J. Cohen and Reuben Hersh note that set theory has not yet found any application in theoretical physics. It has, however, been applied in theoretical psychology. In *The Behavioral Basis of Perception* (Yale University Press, 1962) I have used set theory as a model for the process whereby three-dimensional perceptual space is generated from a multidimensional input to the receptor system. My model involves a hierarchy of sets of increasing complexity, culminating in one whose prop-

Scientific American, March, 1968; Vol. 218, No. 3. 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.

Subscription correspondence should be addressed to Jerome L. Feldman, Circulation 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.

Microfilm correspondence and orders should be addressed to Department SA, University Microfilms, Ann Arbor, Mich. 48107.

Subscription rates: one year, $8; two years, $15; three years, $21. These rates apply throughout the world. Subscribers in the United Kingdom may remit to Midland Bank Limited, 69 Pall Mall, London SW 1, England, for the account of Scientific American, Inc.: one year, two pounds 18 shillings; two years, five pounds eight shillings; three years, seven pounds 11 shillings.

Change of address: please notify us four weeks in advance of change. If available, kindly furnish an address imprint from a recent issue. Be sure to give both old and new addresses, including ZIP-code numbers, if any.

SCM breakthrough: Marchant Cogito 566 PR electronic calculator.

It programs.
It prints.
It's desk-top size.

The new Cogito® 566 PR™ is a total breakthrough. It comes with a full 66 program-step capacity, and has eight separate working and storage registers. Calculates and prints out answers in 1.5 seconds or less.

On manual or programmed operation, the Cogito 566PR prints out every step on a wide, easy-to-read paper tape. Special symbols indicate each factor, each operation, each result, each remainder. Like all other Marchant calculators, the Cogito 566 PR lets you Surecheck *every answer.*

The Cogito 566 PR can handle up to 16 whole digits and features automatic decimal selection to 0, 2, 4, 6, even 8 places.

Square Root is computed instantly and automatically by the touch of a single key. On multiplication it automatically rounds off the last decimal.

At any stop in a programmed operation, the Cogito 566 PR can be used for manual calculations, then returned to complete the original program.

The new Marchant Cogito 566 PR is available now. Call your SCM Marchant representative, or just mail our coupon.

Marchant makes you a better calculator.

SCM®

Smith-Corona Marchant
Division of SCM Corporation

Mail this coupon today.

MARCHANT CALCULATORS

299 Park Avenue
New York, New York 10017

I am interested in knowing more about the new Cogito 566 PR Electronic Calculator.

☐ Please send literature.

☐ Please have a representative call.

NAME ______ TITLE ______

COMPANY ______

STREET NO. ______

CITY ______ STATE ______ ZIP ______

MARCHANT® CALCULATORS

erties match the properties of the perceptual world. All the sets are finite, and therefore questions regarding the axiom of choice and the continuum hypothesis do not arise.

After I had completed my theoretical construction I found that Henri Poincaré, in two of his popular books, *La science et l'hypothèse* (1902) and *La valeur de la science* (1905), had already tackled the problem and constructed a model that in its essential features was similar to mine.

JAMES G. TAYLOR

Bovingdon, England

Sirs:

The article on perpetual motion machines [SCIENTIFIC AMERICAN, January] implies that the possibility of a "perpetual waterfall" was recognized as impossible only after the first and second laws of thermodynamics had been formulated in the 19th century. I can recall reading, more than 50 years ago, a fable (by Aesop?) that concluded with the moral, "The mill cannot grind with the water that is past." This must surely be the earliest statement of the second law, and it would be interesting to trace the original fable.

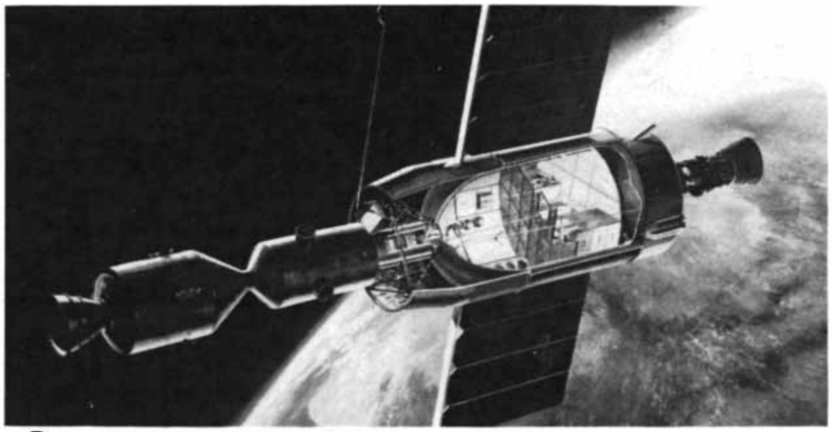
R. H. WRIGHT

Head
Olfactory Response Investigation
British Columbia Research Council
Vancouver, B.C.

ERRATA

In the article "Remote Sensing of Natural Resources" (SCIENTIFIC AMERICAN, January) no mention was made of the source of the three pictures on page 65. These pictures, showing the automatic analysis of tones in an aerial photograph, are the work of Philip G. Langley of the Pacific Southwest Forest and Range Experiment Station of the U.S. Forest Service.

In "Earlier Maturation in Man" (SCIENTIFIC AMERICAN, January) the scale of the opening illustration is incorrect. The unit of height on the scale at the left in the illustration should not be a foot but 10 inches.



50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

MARCH, 1918: "The transformation of one element into another by atomic disintegration has led to the theory that the whole of the elements have at one time been formed from a parent element, and that a slow but sure degradation of the elements is in progress. On this theory, calculations as to the probable life of the world have been made, but the data are too scanty to warrant much faith in the estimations. The slow transformation of one element into another, to effect which was the whole aim of alchemy, is now well demonstrated, but it does not yet seem feasible to put the phenomenon to a practical use. The transformation is attended by the release of a very large amount of energy, and could a means be found to hasten the process, the world would have a tremendous and, it must be feared, a terrible source of power. The logical conclusion to the discovery of such a means is well worked out by the imaginative powers of Mr. H. G. Wells in his book *The World Set Free*. Such a terrible agent of destruction would be at hand that the world would be forced to abandon its wars to protect itself from its own folly. Such results, however, are happily at present only in the realms of imagination."

"With increasing frequency during the past few weeks it has been stated that Germany is already giving evidence, in her thoughts if not in her deeds, that she is preparing for the next war, and nowhere has this evidence been more direct than in a book written by Lieut. General Baron von Freytag-Loringhoven, which bears the title *Deductions from the World War*. Although the General does not of course put it down in bald English that Germany has failed in the present war, it is impossible to read this work without realizing that the German General Staff understands that the great stake for which they played is lost—at least for the present. Be that as it may, the thing that stands out with most sinister significance in this book is the fact that the German high command is studying the lessons of the failures of this war, with a view to applying them in another attempt on an even more widely extended scale."

"The degree to which chemistry has entered into modern warfare is indicated by the use of poisonous gas. The Hague Convention prohibited the use of this infernal instrument, but nevertheless it has become a weapon of offense and defense horrible beyond description. Efforts at counteracting the malign effects of a gas wave collectively have not achieved satisfactory results and dependence is therefore placed upon the gas mask. Similarly, the use of liquid fire—incandescent gasoline or flaming phosphorus—is so deadly that protection along ordinary lines is without result. It is a sad commentary upon the diabolic instrumentality of war that the chemical industry, which has been of such constructive value in advancing human welfare, should become transformed into a hideous messenger of death against which chemistry is a weak defender. The use of poisonous gases and liquid fire are not to be condoned by saying, 'This is war.' The rights of combatants have always received some degree of consideration in order to preserve the semblance of humanity even in war. Some of the triumphs of chemistry in modern warfare represent the defeat of civilization."

"Several European observers of the total lunar eclipse of July 4–5, 1917, have reported that the brightness of the lunar disk appeared much greater around the limb than near the center. These observations lead M. A. Nodon of Bordeaux to revive a suggestion that has sometimes been made to account for the brilliancy of certain lunar craters; viz., that the surface of the moon may possess a luminosity of its own in the nature of phosphorescence. In that case, perspective would increase the apparent luminosity toward the limb."

SCIENTIFIC AMERICAN

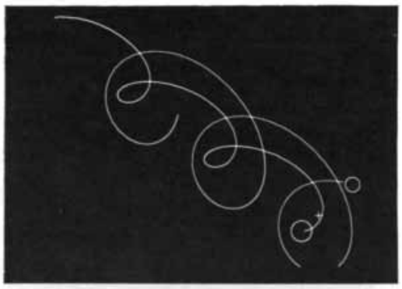
MARCH, 1868: "In *The Naturalist* Mr. Alfred Russel Wallace has published a very interesting paper on the 'Relation between Sexual Differences of Color and Nidification in Birds.' In some few species of birds the females can boast of a plumage more beautiful and brilliant than that of the male. In cases where the female has this conspicuous appearance the nest always conceals her, but in cases where the female is of a dull color the

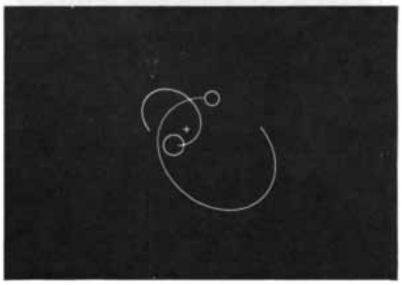
Report from

BELL LABORATORIES

Movies via computer

Abstract or complex concepts are difficult to communicate. Often they are best grasped visually—particularly through animated films. But making such movies has been tedious and expensive. At Bell Laboratories, therefore, we are experimenting with movies by photographing computer-controlled cathode-ray tubes. Not only is this more efficient than traditional methods, for many kinds of movies, but the computer can sometimes reveal motions and shapes which are otherwise concealed in masses of data. Here are examples of our work.





"Force, Mass, and Motion," an educational film by F. W. Sinden, shows how gravity-like forces and inertia affect bodies with various initial velocities. This interplay is hard to visualize, but is clear on the screen. It produces the curves in the upper picture, one frame from the movie. This film, costly with conventional animation, is inexpensive here because the computer makes pictures by solving equations.

In one sense, the computer movie is a "perfect laboratory"; it demonstrates exactly how our mathematical models would behave and helps us to look for imperfections in our experimental apparatus when we do go ahead in the laboratory.

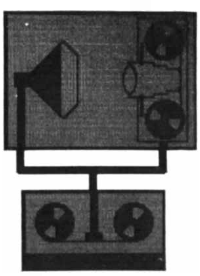
The lower picture is another frame from the movie. Here the program was slightly altered to view the system from a reference frame moving with the center of mass. The apparently complex curves traced by the bodies in the upper picture turn out in the lower one to be ellipses moving together linearly.

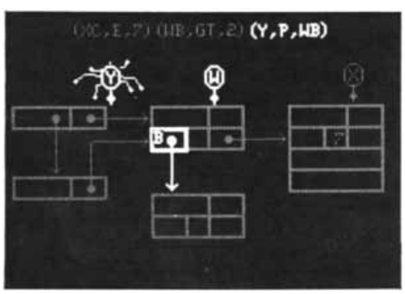
K. C. Knowlton's BEFLIX (Bell flicks) is a computer program whose input is a description of the desired movie in the language of the filmmaker: CAMERA, DISSOLVE, ZOOM. Its output is a magnetic tape containing an encoding of pictures. These are subsequently displayed on a cathode-ray tube where they are photographed.

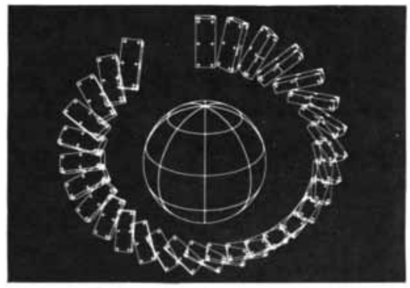
The BEFLIX picture is a rectangular array of dots; the intensity of each can vary through eight levels. The filmmaker can tell BEFLIX that lines or arcs should be drawn, areas "painted" various shades of gray, displayed shapes moved in various directions, and the like. There is an assortment of letter sizes and faces for titling.

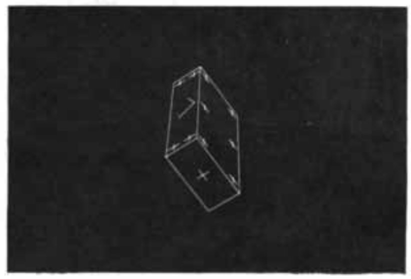
The frames below were produced in the BEFLIX language. The first is from a movie describing BEFLIX itself. The second is from a movie about a new programming language produced at Bell Laboratories. In this film, animated "bugs" demonstrate how information is moved around in the computer.

In this new method of animation, both film motion and display on the tube can be controlled automatically by information on a magnetic tape.









A movie by E. E. Zajac demonstrates the effects of gravity in keeping a communications satellite facing the Earth. Satellite motion is described by complicated differential equations. They can be solved on a computer, but the resulting list of numbers is almost incomprehensible. In the movie, however, the dynamics of satellite motion —stability, orientation, and time—are instantly visible.

The pictures show two parts of the movie. At top, the stylized satellite-Earth system is seen from a position fixed relative to Earth (thirty selected frames are superimposed). The lower picture shows the satellite from a position orbiting with it. This is an advantage of computer movie-making: the second viewpoint required only relatively minor program changes.

The film was "reshot" several times to show the effects of various stabilizing parameters.

Bell Telephone Laboratories
Research and Development Unit of the Bell System

RECENT FINDINGS

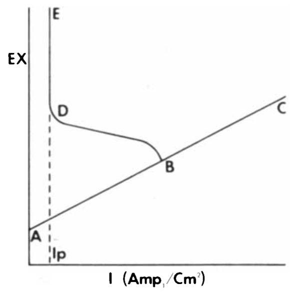
RESEARCH LABORATORIES

Ford

Mechanisms of corrosion reactions and models for the passive state

The study concerns the fundamental electrochemical behavior of metals in a corrosive environment. The reactions of metals with electrolytes is considered as a "dissolution" of electrons and positive ions. The rate of corrosion is measured in terms of a current density, and is a function of the electrical potential difference which develops across the metal-electrolyte interface, (the latter being measured as an electrode potential). A typical corrosion rate versus electrode potential curve is shown schematically in Figure 1. The curve $\overline{ABC}$, is typical for active metals such as iron, zinc, copper, etc., in many corrosive environments. This behavior leads either to rapid corrosion, or, in some cases, fairly slow corrosion.

Figure 1



Some metals and alloys such as aluminum, titanium, and stainless steels, which should also be thermodynamically active, show a passive behavior in many corrosive environments. The corrosion rate—potential curve is given by curve $\overline{ABDE}$. These metals assume a potential such as E_x and exhibit a very low corrosion rate I_p. This passive behavior is basically the cause for the corrosion resistance and non-tarnishing properties of these types, which are otherwise active metals.

Ford Motor Company scientists are studying the mechanisms of the reactions, both on the active and passive metals. Modern electrochemical techniques and instrumentation are used to determine the kinetics of the processes. Special attention is being given to surface structure, dislocation densities and point defects on the dissolution reaction.

The exact nature of the passive state is not completely understood, but Ford scientists are studying this passive state by means of both electrochemical and ellipsometric techniques. A model for the passive state has been developed which is based on the nucleation and growth of films on the surface—both processes being a function of potential and current density. The model yields the characteristic curve $\overline{ABDE}$ in Figure 1, and agrees with galvanostatic charge curves which have been obtained on some pure metals.

Of special importance is the vexing problem of the breakdown of passivity by aggressive anions, such as chloride ion. This type of corrosion leads to localized pitting attack. It has been shown that this pitting is a definite function of the electrode potential, as well as the chloride concentration; and the "pitting potential" is dependent upon the alloy composition. Curves showing the breakdown of passivity for several types of stainless steel are shown in Figure 2.

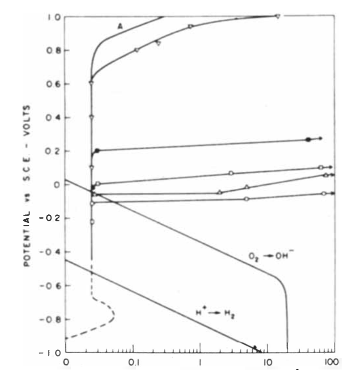


Figure 2

A critical analysis of some existing theories of passivity shows that theories do not allow for this potential dependency for anion effect. A modified theory has been suggested by our scientists.

The techniques and instrumentation used in these basic studies has led to the development of rapid and reliable test methods for the evaluation of the corrosion resistance of materials for specific applications, as well as problem solving.

PROBING DEEPER FOR BETTER IDEAS

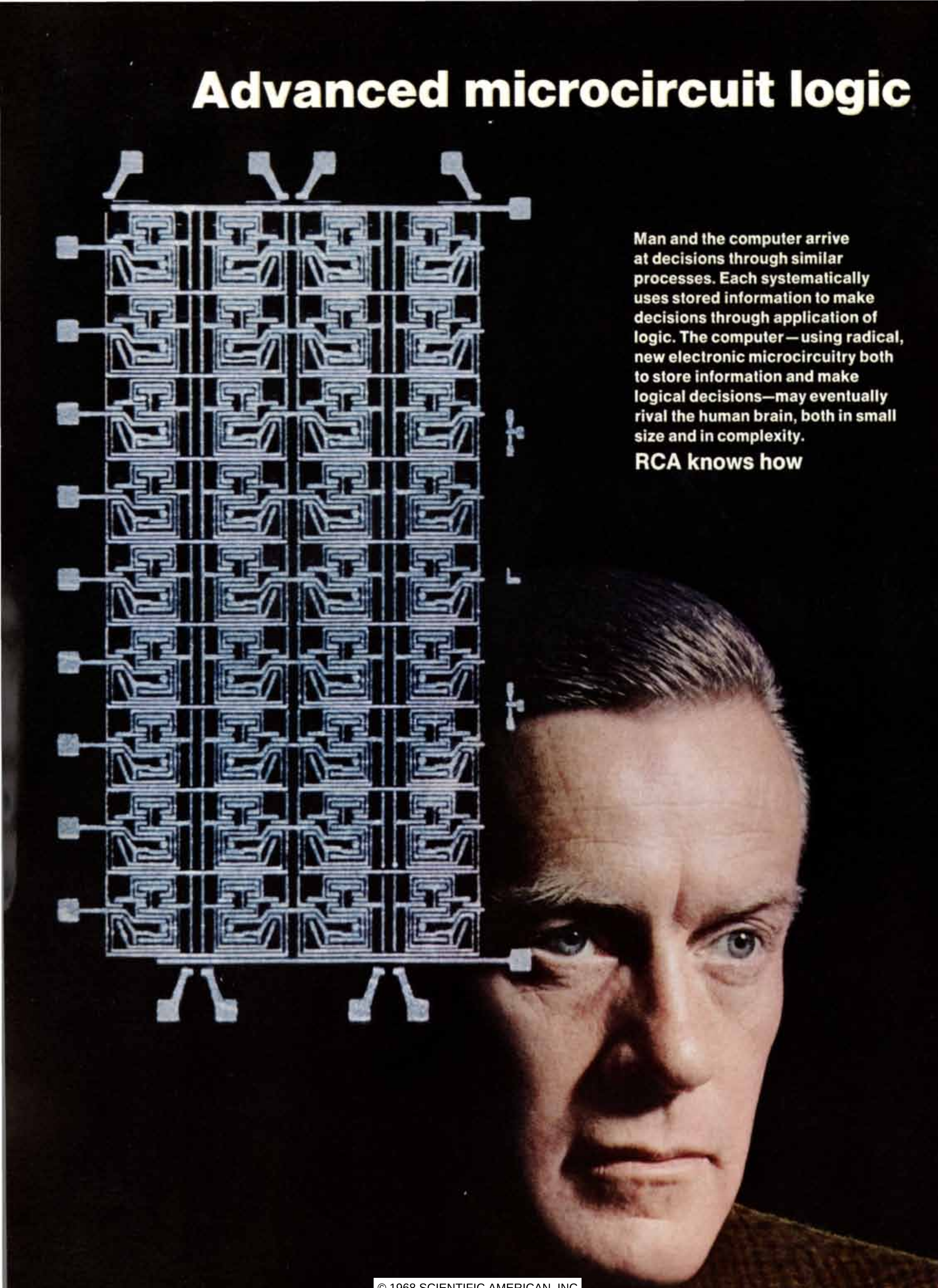
Ford

nest exposes a considerable portion of the sitting bird. When the male bird is less brilliant than his mate, it is found that he performs the duties of incubation. There seems, then, to be a connection between the color of the different sexes of birds and the sitting over the eggs. Mr. Wallace considers that Darwin's principle of natural selection most aptly explains this connection of color and nests."

"Hitherto transfusion of healthy blood into the veins of another being was applied only in extreme cases where loss of blood had rendered necessary some desperate means of replacing the loss of red corpuscles and oxygen in the arteries, and restoring respiration and circulation. Quite lately, however, two physicians, Drs. Eulenburg and Landers, have sent to the French Academy a treatise on Transfusion of Blood, in which by an extended series of experimental researches upon animals they seek to prove that this process, modified in a certain way and repeated if necessary, should be viewed as a sovereign mode of treatment in all cases of acute poisoning, viz., of such poisons which, after absorption by the blood, act injuriously upon the vital nervous centers."

"Prof. De la Rive of Geneva has contrived an instrument for measuring the transparency of the atmosphere. The inventor agrees with Pasteur, who supposes that the light dry fog which under certain conditions of the air intercepts the light is caused by myriads of organic germs floating near the earth, which are washed to the earth by the heavy rains or are destroyed by severe frosts, thus accounting for the clearness of the atmosphere at these times."

"The famous Thames tunnel, which for the 25 years since its completion has proved an indifferent speculation, is at last to be made of some practical use. It is stated that two railroads on opposite sides of the river propose forming a junction by means of this subaqueous passage-way, and will make gradual entrances a mile distant from either bank. The original cost of the tunnel was more than $2,000,000. It was sold a few years ago for half that amount, and even at this sacrifice the purchasers have found it to be a very unfortunate investment, the receipts, principally tolls from foot passengers drawn thither by curiosity, averaging but $125 per week, which have been entirely consumed by expenses."


Advanced microcircuit logic
Man and the computer arrive at decisions through similar processes. Each systematically uses stored information to make decisions through application of logic. The computer—using radical, new electronic microcircuitry both to store information and make logical decisions—may eventually rival the human brain, both in small size and in complexity.
RCA knows how

nearing human complexity?

As yet, no man-made instrument has approached the complexity of the human brain or its relatively small size and efficiency. However, much progress toward this goal has been made in data-processing technology since the first electronic computer was designed.

This early computer, in the light of today's electronic advancements, was a power-hungry machine assembled from a multitude of then-existing electronic components. It was a far cry from the "electronic brain" many have named it, both because of its huge size and relatively low packing density.

The continuing effort to simulate the human brain's capability and size has resulted in RCA's achievement of a new and different type of electronic microcircuitry—COS/MOS, the first commercially available COmplementary Symmetry Metal Oxide Semiconductors.

These integrated circuits—barely larger than this "O"—pack as many as a thousand transistors onto a single chip or substrate, to bring the "electronic brain" concept much closer to reality.

The computer "state of the art" has already come a long way since those early data-processing machines that used cumbersome banks of electron tubes and associated discrete components in the circuits. While these circuits were efficient in making rapid decisions, they were limited by sheer size; the tremendous heat generated by thermionic tubes; the vast number of discrete resistors and capacitors, and the many interconnections required.

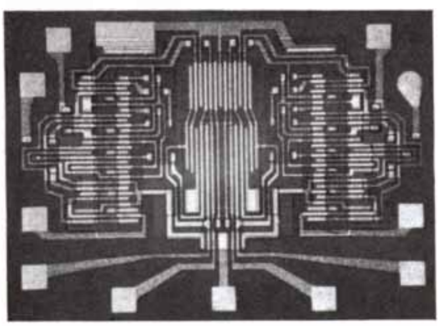


Fig. 1

To minimize these problems, computer and circuit designers first replaced electron tubes with transistors, mounted together with discrete components on circuit boards. Next, whole circuits were diffused into single small semiconductor substrates—the now familiar integrated circuit. *Fig. 1* is a photomicrograph of a COS/MOS dual flip-flop circuit containing 48 complementary MOS transistors on one minute chip.

Until the recent advent of the complementary symmetry MOS integrated circuit, however, each design advancement in computer circuitry was made through the use of conventional design techniques; discrete resistors and capacitors were replaced by built-in passive elements. But even these passive elements took up space, consumed significant power, and were generators of heat in appreciable amounts.

The new COS/MOS units, on the other hand, operate on nanowatts of power in the quiescent state and use greater power only during the instant of switching information. Of equal importance, they can pack hundreds of switching circuits into less than .01 square inch of silicon area. In time, with COS/MOS, the entire arithmetic and memory sections of a computer could be contained on only a few small chips of silicon.

Designers have long known the advantages of complementary symmetry circuit design, which uses only MOS transistors of opposite polarity and eliminates the use of other types of components. These circuits only require power when the switching cycle is actually taking place. But building such circuits onto a single substrate challenged the technical skills of the semiconductor industry.

RCA had already developed a semiconductor substrate material (silicon) and oxides of unsurpassed purity in producing the single polarity metal oxide semiconductors first made available to industry by RCA some years ago. This technology was sufficient to produce either n-channel or p-channel transistors separately. But to make COS/MOS circuits, an additional skill was required—the ability to produce transistors of both polarities on the same chip. RCA engineers have developed that capability.

This ability, combined with further advances in metallurgy and production techniques, has resulted in the currently available complementary symmetry circuits. *Fig. 2* shows—in diagram—some of the steps in producing one COS/MOS unit, with dimensions as critical as 10-millionths of an inch!

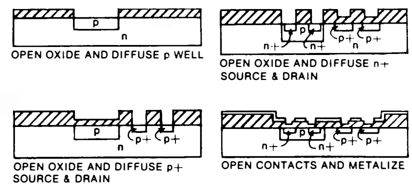


Fig. 2

RCA's new COS/MOS devices will bring an invaluable new circuit element to designers in their continuing effort to develop computers that approach the human brain in small size and extreme complexity. Computers built with these circuits will require far less air conditioning to dispose of heat (since COS/MOS circuits eliminate heat-generating passive components) and will take up considerably less space. These circuits may also make possible computers that operate from as little power as that now required for a portable radio!

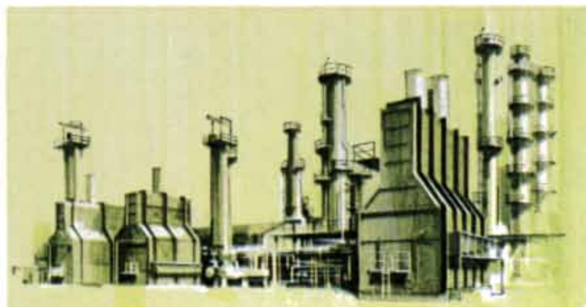
Totally new concepts in military and space computer applications will result from the reductions in size, power requirements, circuit interconnections and heat generation made possible by COS/MOS circuits. In addition, the low unit cost per circuit function anticipated for COS/MOS circuits promises wholly new applications of data-processing techniques and equipment—such as desk-top computers for science, business and industry, and the low-cost, low-power unit for use in the home to computerize anything from household records to checkbook balances.

Computer technology has now taken an important step toward rivalling the human brain!

For detailed information about currently available COS/MOS devices and additional application information, write Commercial Engineering, RCA Electronic Components and Devices, Section C95DC, Harrison, N. J. 07029.

RCA





THE AUTHORS

RICHARD L. GARWIN and HANS A. BETHE ("Anti-Ballistic-Missile Systems") are respectively director of applied research at the Thomas J. Watson Research Center of the International Business Machines Corporation and professor of physics at Cornell University. Garwin was graduated from the Case Institute of Technology in 1947 and received master's and doctor's degrees in physics from the University of Chicago. He has served as a consultant to the Los Alamos Scientific Laboratory and the President's Science Advisory Committee, of which he was a member from 1962 to 1965. Bethe, who won the Nobel prize in physics last year "for his contributions to the theory of nuclear reactions," has been at Cornell since 1935. He was chief of the theoretical physics division at Los Alamos during World War II and has continued to be a consultant to the Los Alamos laboratory as well as to several firms involved in atomic energy. From 1956 to 1959 he was a member of the President's Science Advisory Committee. In 1958 and 1959 he was a member of the U.S. delegation at the discussions in Geneva on the discontinuation of nuclear-weapons tests.

LEONARD HAYFLICK ("Human Cells and Aging") is professor of medical microbiology at the Stanford University School of Medicine. Until he went to Stanford recently he had spent most of his life in Philadelphia, where he was born and educated. He received all his university degrees from the University of Pennsylvania: a bachelor's degree in microbiology in 1951, a master's degree in medical microbiology in 1953 and a Ph.D. in medical microbiology and biochemistry in 1956. For several years he was a member of the faculty at the university and then was associated with the Wistar Institute of Anatomy and Biology in Philadelphia. In addition to his studies of cultured human cells Hayflick has worked with the smallest free-living microorganisms, the mycoplasmas. In 1961 he identified the agent causing primary atypical pneumonia in man as a mycoplasma; previously the organism had been thought to be a virus.

J. E. DIXON, J. R. CANN and COLIN RENFREW ("Obsidian and the Origins of Trade") are respectively at the University of Cambridge, the British Museum (Natural History) and the University of Sheffield. All three were graduated from Cambridge and did graduate work there. Dixon has remained at the university as a Science Research Council postdoctoral research fellow; his work deals with the glaucophane schists of Greece. Cann, who is a senior scientific officer at the British Museum (Natural History), finished his undergraduate work at Cambridge in 1959 and obtained a Ph.D. there in 1962. Until 1966 he was a research fellow at St. John's College, Cambridge, specializing in the geology of the ocean floor. He is continuing this work in London. Renfrew is a lecturer in prehistory and archaeology at Sheffield. He received a Ph.D. from Cambridge in 1965 and became a research fellow at St. John's College in the same year. He has made a special study of the prehistory of the Cycladic Islands of Greece, where he conducted archaeological excavations in 1964 and 1965.

R. B. MERRIFIELD ("The Automatic Synthesis of Proteins") is professor of biochemistry at Rockefeller University. He was graduated from the University of California at Los Angeles in 1943 and obtained a doctorate in chemistry there in 1949. He began work at Rockefeller University (then the Rockefeller Institute) in 1949 as a research assistant. Merrifield writes: "Virtually all my working time for the past five or six years has been on solid phase peptide synthesis. Before that I worked on peptide growth factors for bacteria and before that on the isolation and characterization of oligonucleotides that promote the growth of bacteria."

PAUL A. KOLERS ("Bilingualism and Information Processing") is a research associate in the Research Laboratory of Electronics and the department of electrical engineering at the Massachusetts Institute of Technology. He writes that he "went into psychology in order to study how the mind works" but that in graduate school he was "taught that one cannot study the mind." As a result, after receiving a Ph.D. in experimental psychology from New York University in 1957, he began studying how the human visual system operates on fairly simple targets. "The results kept implicating 'higher level' processes—what we now call cognitive processes—so from visual form perception I went to visual illusions, where I found that words and instructions were important variables in visual events. Again thrust 'upward,' I began studying the use of words and the relations between verbal and visual inputs. My current work is on the psychology of reading. It seems to have taken a long time, but here I am now studying the mind."

WERNER BRANDT ("Channeling in Crystals") is professor of physics at New York University and director of the university's Radiation and Solid State Laboratory. He received a doctorate in physics at the University of Heidelberg. On moving to the U.S. he worked in the Central Research Department of E. I. duPont de Nemours & Co. He went to N.Y.U. in 1961. Among his interests is the study of positrons. He writes: "We try to understand microscopic properties of matter by observing the ways positrons, the antiparticles to our electrons, are created and annihilated in our world or, as it were, in their antiworld. When I am not doing physics, I can sometimes be found in one of my canvas folding boats on a lake or a river in Maine, Florida, Lapland or Switzerland."

J. S. MAYO ("Pulse-Code Modulation") is director of the Ocean Systems Laboratory at the Bell Telephone Laboratories. Until recently he had been for several years head of the High-Speed Pulse-Code Modulation Terminal Department at Bell Laboratories. He joined Bell Laboratories in 1955, the year he obtained a Ph.D. in electrical engineering from the North Carolina State College of Agriculture and Engineering. Mayo also received bachelor's and master's degrees in electrical engineering at North Carolina State.

N. MROSOVSKY ("The Adjustable Brain of Hibernators") does research in the department of zoology at the University of Toronto; he says that "this department, having some 25 'cold rooms' in one building, must be one of the best equipped in the world for studies of hibernation." Mrosovsky studied physiology and psychology at the University of Cambridge and then went to the department of psychology at University College London, where he worked for the Medical Research Council and, in 1962, obtained a Ph.D. He writes that he "enjoys the challenges of widely different types of problems," ranging from the work he describes in his article to "investigations in the wild of the water-finding behavior of sea turtle hatchlings in the Caribbean."

VICTOR F. WEISSKOPF, who in this issue reviews *The Politics of Pure Science*, by Daniel S. Greenberg, is professor of physics at the Massachusetts Institute of Technology.

Established 1845 SCIENTIFIC AMERICAN March 1968 Volume 218 Number 3

Anti-Ballistic-Missile Systems

The U.S. is now building a "light" ABM system. The authors argue that offensive tactics and cheap penetration aids could nullify the effectiveness of this system and any other visualized so far

by Richard L. Garwin and Hans A. Bethe

Last September, Secretary of Defense McNamara announced that the U.S. would build "a relatively light and reliable Chinese-oriented ABM system." With this statement he apparently ended a long and complex debate on the merits of any kind of anti-ballistic-missile system in an age of intercontinental ballistic missiles carrying multimegaton thermonuclear warheads. Secretary McNamara added that the U.S. would "begin actual production of such a system at the end of this year," meaning the end of 1967.

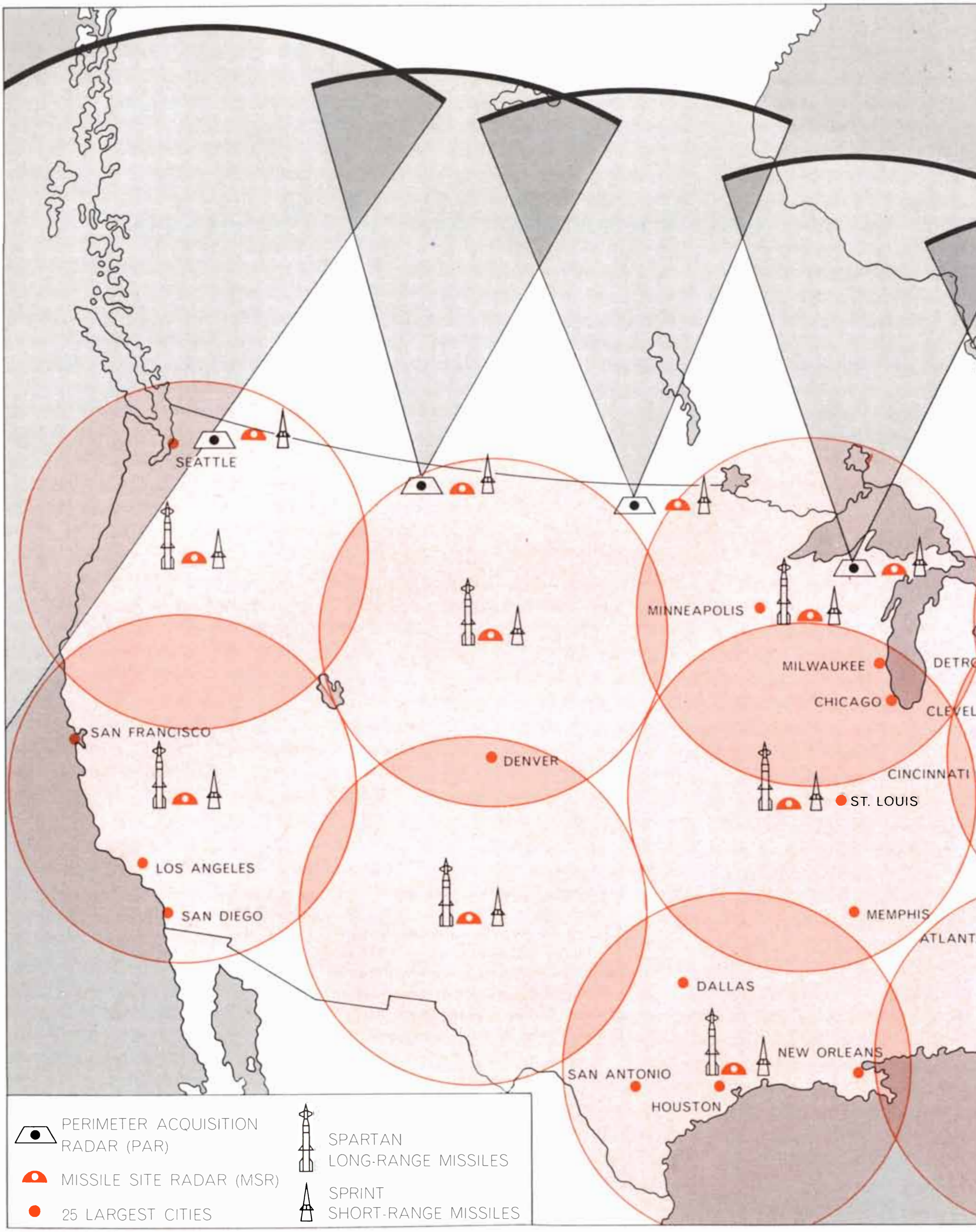
As two physicists who have been concerned for many years with the development and deployment of modern nuclear weapons we wish to offer some comments on this important matter. On examining the capabilities of ABM systems of various types, and on considering the stratagems available to a determined enemy who sought to nullify the effectiveness of such a system, we have come to the conclusion that the "light" system described by Secretary McNamara will add little, if anything, to the influences that should restrain China indefinitely from an attack on the U.S. First among these factors is China's certain knowledge that, in McNamara's words, "we have the power not only to destroy completely her entire nuclear offensive forces but to devastate her society as well."

An even more pertinent argument against the proposed ABM system, in our view, is that it will nourish the illusion that an effective defense against ballistic missiles is possible and will lead almost inevitably to demands that the light system, the estimated cost of which exceeds $5 billion, be expanded into a heavy system that could cost upward of $40 billion. The folly of undertaking to build such a system was vigorously stated by Secretary McNamara. "It is important to understand," he said, "that none of the [ABM] systems at the present or foreseeable state of the art would provide an impenetrable shield over the United States.... Let me make it very clear that the [cost] in itself is not the problem: the penetrability of the proposed shield is the problem."

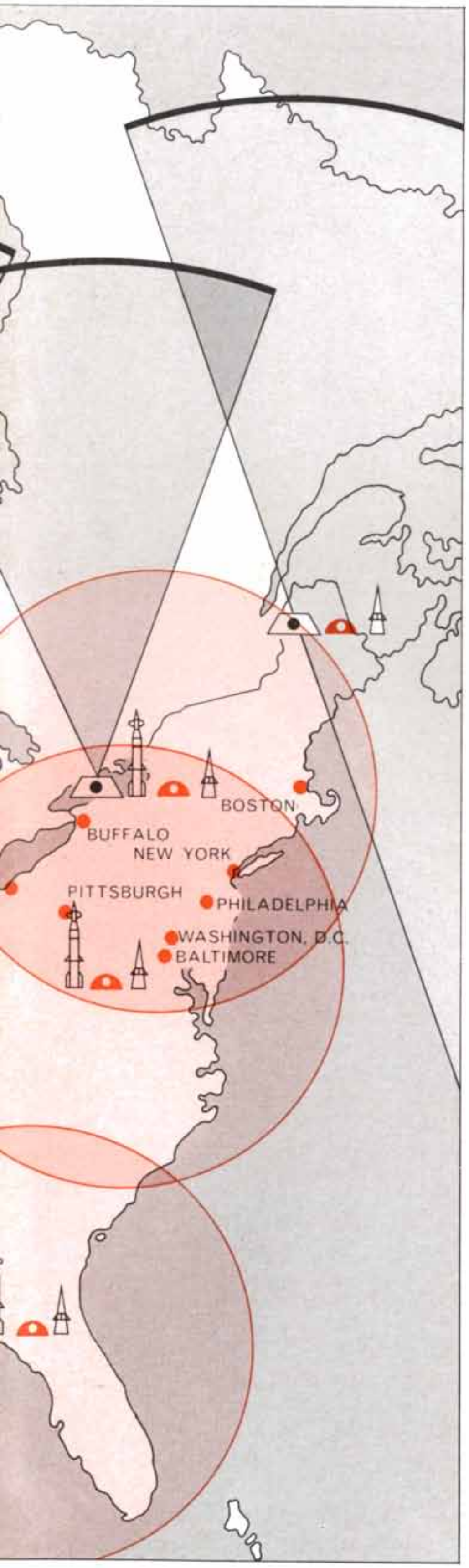
In our view the penetrability of the light, Chinese-oriented shield is also a problem. It does not seem credible to us that, even if the Chinese succumbed to the "insane and suicidal" impulse to launch a nuclear attack on the U.S. within the next decade, they would also be foolish enough to have built complex and expensive missiles and nuclear warheads peculiarly vulnerable to the light ABM system now presumably under construction (a system whose characteristics and capabilities have been well publicized). In the area of strategic weapons a common understanding of the major elements and technical possibilities is essential to an informed and reasoned choice by the people, through their government, of a proper course of action. In this article we shall outline in general terms, using nonsecret information, the techniques an enemy could employ at no great cost to reduce the effectiveness of an ABM system even more elaborate than the one the Chinese will face. First, however, let us describe that system.

Known as the Sentinel system, it will provide for long-range interception by Spartan antimissile missiles and short-range interception by Sprint antimissile missiles. Both types of missile will be armed with thermonuclear warheads for the purpose of destroying or inactivating the attacker's thermonuclear weapons, which will be borne through the atmosphere and to their targets by reentry vehicles (RV's). The Spartan missiles, whose range is a few hundred kilometers, will be fired when an attacker's reentry vehicles are first detected rising above the horizon by perimeter acquisition radar (PAR).

If the attacker is using his available propulsion to deliver maximum payload, his reentry vehicles will follow a normal minimum-energy trajectory, and they will first be sighted by one of the PAR's when they are about 4,000 kilometers, or about 10 minutes, away [*see illustration on page 26*]. If the attacker chooses to launch his rockets with less than maximum payload, he can put them either in a lofted trajectory or in a depressed one. The lofted trajectory has certain advantages against a terminal defense system. The most extreme example of a depressed trajectory is the path followed by a low-orbit satellite. On such a trajectory a reentry vehicle could remain below an altitude of 160 kilometers and would not



SENTINEL ANTI-BALLISTIC-MISSILE SYSTEM, described as a "relatively light and reliable Chinese-oriented ABM system," is now under construction at an estimated cost exceeding $5 billion. Designed to defend the entire U.S., Sentinel will depend on perhaps six perimeter acquisition radars (PAR's) along the country's borders to detect enemy missiles as they come over the northern horizon. The arcs at the end of the radar "fans" show where an enemy reentry vehicle (RV) would be detected if it were in a low, satellite-like orbit. The PAR's will alert Spartan interceptors located at some 10 or a dozen sites around the U.S. The sites shown on the map are not actual ones but indicate how the U.S. could be covered by a pattern of 10 Spartan sites, assuming that the Spar-



tans have an effective range of 600 kilometers. Each Spartan site will be protected by short-range Sprint missiles and will include missile-site radar (MSR) to help guide both types of missiles. Sprints and MSR's will also guard the PAR installations.

be visible to the horizon-search radar until it was some 1,400 kilometers, or about three minutes, away. This is FOBS: the fractional-orbit bombardment system, which allows intercontinental ballistic missiles to deliver perhaps 50 to 75 percent of their normal payload.

In the Sentinel system Spartans will be launched when PAR has sighted an incoming missile; they will be capable of intercepting the missile at a distance of several hundred kilometers. To provide a light shield for the entire U.S. about half a dozen PAR units will be deployed along the northern border of the country to detect missiles approaching from the general direction of the North Pole [*see illustration at left*]. Each PAR will be linked to several "farms" of long-range Spartan missiles, which can be hundreds of kilometers away. Next to each Spartan farm will be a farm of Sprint missiles together with missile-site radar (MSR), whose function is to help guide both the Spartans and the shorter-range Sprints to their targets. The task of the Sprints is to provide terminal protection for the important Spartans and MSR's. The PAR's will also be protected by Sprints and thus will require MSR's nearby.

Whereas the Spartans are expected to intercept an enemy missile well above the upper atmosphere, the Sprints are designed to be effective within the atmosphere, at altitudes below 35 kilometers. The explosion of an ABM missile's thermonuclear warhead will produce a huge flux of X rays, neutrons and other particles, and within the atmosphere a powerful blast wave as well. We shall describe later how X rays, particles and blast can incapacitate a reentry vehicle.

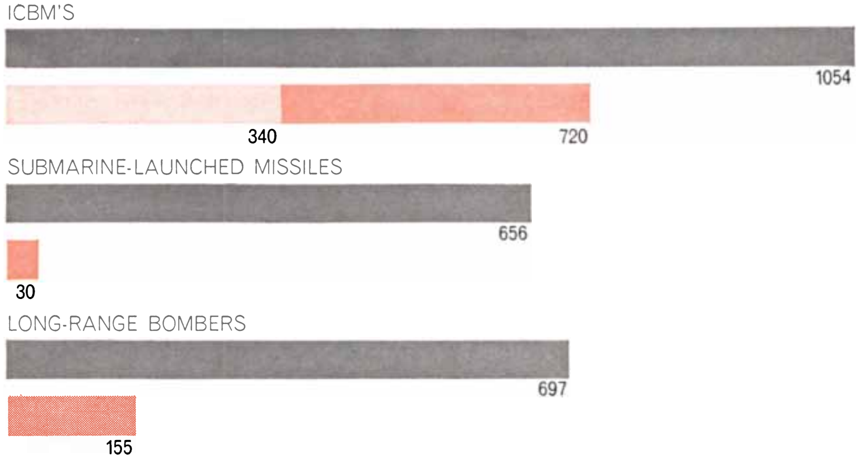
Before we consider in detail the capabilities and limitations of ABM systems, one of us (Garwin) will briefly summarize the present strategic position of the U.S. The primary fact is that the U.S. and the U.S.S.R. can annihilate each other as viable civilizations within a day and perhaps within an hour. Each can at will inflict on the other more than 120 million immediate deaths, to which must be added deaths that will be caused by fire, fallout, disease and starvation. In addition more than 75 percent of the productive capacity of each country would be destroyed, regardless of who strikes first. At present, therefore, each of the two countries has an assured destruction capability with respect to the other. It is usually assumed that a nation faced with the assured destruction of 30 percent of its population and productive capacity will be deterred from destroying another nation, no matter how serious the grievance. Assured destruction is therefore not a very flexible political or military tool. It serves only to preserve a nation from complete destruction. More conventional military forces are needed to fill the more conventional military role.

Assured destruction was not possible until the advent of thermonuclear weapons in the middle 1950's. At first, when one had to depend on aircraft to deliver such weapons, destruction was not really assured because a strategic air force is subject to surprise attack, to problems of command and control and to attrition by the air defenses of the other side. All of this was changed by the development of the intercontinental ballistic missile and also, although to a lesser extent, by modifications of our B-52 force that would enable it to penetrate enemy defenses at low altitude. There is no doubt today that the U.S.S.R. and the U.S. have achieved mutual assured destruction.

The U.S. has 1,000 Minuteman missiles in hardened "silos" and 54 much larger Titan II missiles. In addition we have 656 Polaris missiles in 41 submarines and nearly 700 long-range bombers. The Minutemen alone could survive a surprise attack and achieve assured destruction of the attacker. In his recent annual report the Secretary of Defense estimated that as of October, 1967, the U.S.S.R. had some 720 intercontinental ballistic missiles, about 30 submarine-launched ballistic missiles (excluding many that are airborne rather than ballistic) and about 155 long-range bombers. This force provides assured destruction of the U.S.

Secretary McNamara has also stated that U.S. forces can deliver more than 2,000 thermonuclear weapons with an average yield of one megaton, and that fewer than 400 such weapons would be needed for assured destruction of a third of the U.S.S.R.'s population and three-fourths of its industry. The U.S.S.R. would need somewhat fewer weapons to achieve the same results against the U.S.

It is worth remembering that intercontinental missiles and nuclear weapons are not the only means of mass destruction. They are, however, among the most reliable, as they were even when they were first made in the 1940's and 1950's. One might build a strategic force somewhat differently today, but the U.S. and the U.S.S.R. have no incentive for doing so. In fact, the chief virtue of assured destruction may be that it removes the need to race—there is no reward for



COMPARISON OF NUCLEAR WEAPON CARRIERS shows that the U.S. (*gray bars*) outweighs the U.S.S.R. in every category. These figures, representing U.S. intelligence estimates as of October 1, 1967, appear in the Secretary of Defense's latest annual report. The report notes that the number of Soviet ICBM's had increased from 340 a year earlier. Of the 1,054 U.S. ICBM's, 1,000 are Minutemen and 54 are Titan II's. The 30 submarine-launched ballistic missiles credited to the U.S.S.R. are believed to have a range considerably less than the range of some 2,500 kilometers for the 656 Polaris missiles aboard 41 nuclear-powered U.S. submarines. The report states that the combined U.S. force could deliver up to 4,500 nuclear warheads compared with about 1,000 larger warheads for the U.S.S.R. force.

getting ahead. One really should not worry too much about new means for delivering nuclear weapons (such as bombs in orbit or fractional-orbit systems) or about advances in chemical or biological warfare. A single thermonuclear assured-destruction force can deter such novel kinds of attack as well.

Now, as Secretary McNamara stated in his September speech, our defense experts reckoned conservatively six to 10 years ago, when our present strategic-force levels were planned. The result is that we have right now many more missiles than we need for assured destruction of the U.S.S.R. If war comes, therefore, the U.S. will use the excess force in a "damage-limiting" role, which means firing the excess at those elements of the Russian strategic force that would do the most damage to the U.S. Inasmuch as the U.S.S.R. has achieved the level of assured destruction, this action will not preserve the U.S., but it should reduce the damage, perhaps sparing a small city here or there or reducing somewhat the forces the U.S.S.R. can use against our allies. To the extent that this damage-limiting use of our forces reduces the damage done to the U.S.S.R. it may slightly reduce the deterrent effect resulting from assured destruction. It must be clear that only surplus forces will be used in this way. It should be said, however, that the exact level of casualties and industrial damage required to destroy a nation as a viable society has been the subject of surprisingly little research or even argument.

One can conceive of three threats to the present rather comforting situation of mutual assured destruction. The first would be an effective counterforce system: a system that would enable the U.S. (or the U.S.S.R.) to incapacitate the other side's strategic forces before they could be used. The second would be an effective ballistic-missile defense combined with an effective antiaircraft system. The third would be a transition from a bipolar world, in which the U.S. and the U.S.S.R. alone possess overwhelming power, to a multipolar world including, for instance, China. Such threats are of course more worrisome in combination than individually.

American and Russian defense planners are constantly evaluating less-than-perfect intelligence to see if any or all of these threats are developing. For purposes of discussion let us ask what responses a White side might make to various moves made by a Black side. Assume that Black has threatened to negate White's capability of assured destruction by doing one of the following things: (1) it has procured more intercontinental missiles, (2) it has installed some missile defense or (3) it has built up a large operational force of missiles each of which can attack several targets, using "multiple independently targetable reentry vehicles" (MIRV's).

White's goal is to maintain assured destruction. He is now worried that Black may be able to reduce to a dangerous level the number of White warheads that will reach their target. White's simplest response to all three threats—but not necessarily the most effective or the cheapest—is to provide himself with more launch vehicles. In addition, in order to meet the first and third threats White will try to make his launchers more difficult to destroy by one or more of the following means: by making them mobile (for example by placing them in submarines or on railroad cars), by further hardening their permanent sites or by defending them with an ABM system.

Another possibility that is less often discussed would be for White to arrange to fire the bulk of his warheads on "evaluation of threat." In other words, White could fire his land-based ballistic missiles when some fraction of them had already been destroyed by enemy warheads, or when an overwhelming attack is about to destroy them. To implement such a capability responsibly requires excellent communications, and the decision to fire would have to be made within minutes, leading to the execution of a prearranged firing plan. As a complete alternative to hardening and mobility, this fire-now-or-never capability would lead to tension and even, in the event of an accident, to catastrophe. Still, as a supplemental capability to ease fears of effective counterforce action, it may have some merit.

White's response to the second threat—an increase in Black's ABM defenses—might be limited to deploying more launchers, with the simple goal of saturating and exhausting Black's defenses. But White would also want to consider the cost and effectiveness of the following: penetration aids, concentrating on undefended or lightly defended targets, maneuvering reentry vehicles or multiple reentry vehicles. The last refers to several reentry vehicles carried by the same missile; the defense would have to destroy all of them to avoid damage. Finally, White could reopen the question of whether he should seek assured destruction solely by means of missiles. For example, he might reexamine the effectiveness of low-altitude bombers or he might turn his attention to chemical or biological weapons. It does not much matter how assured destruction is achieved. The important thing, as Secretary McNamara has emphasized, is that the other side find it credible. ("The point is that a potential aggressor must himself believe that our assured destruction capability is in fact actual, and that our will to use it in retaliation to an attack is in fact unwavering.")

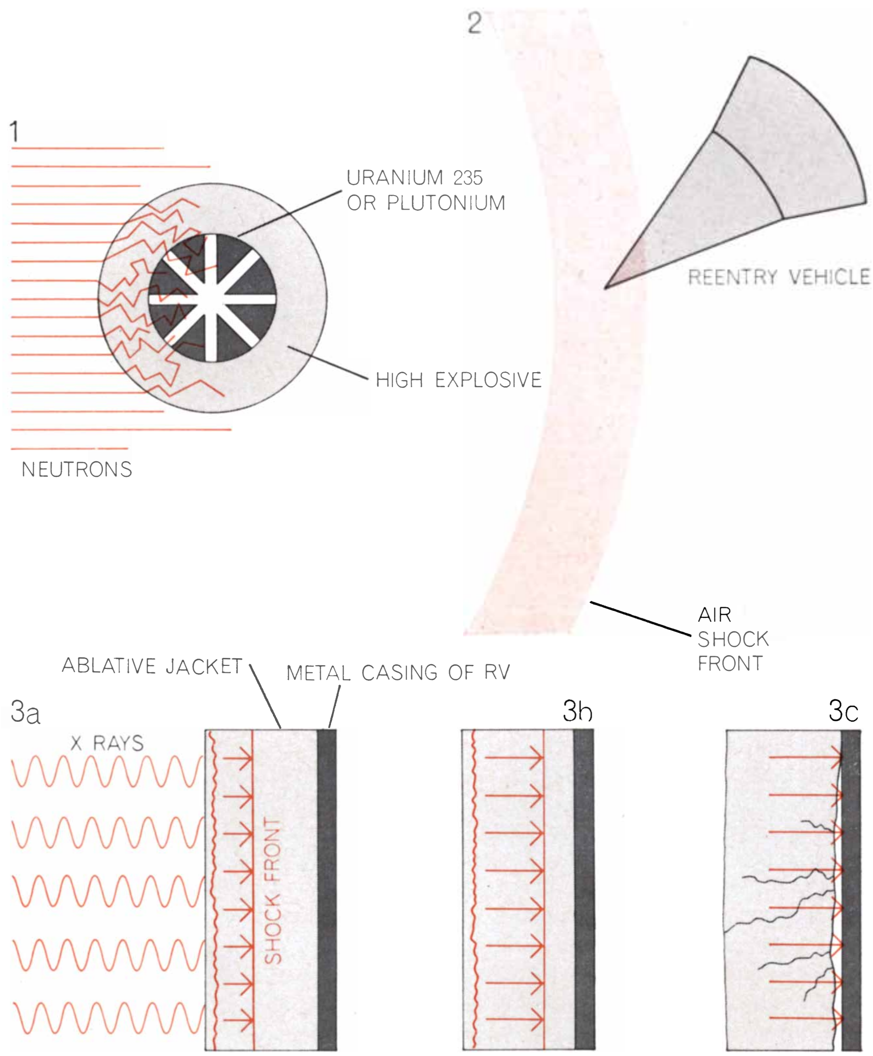
It is clear that White has many options, and that he will choose those that are most reliable or those that are cheapest for a given level of assured destruction. Although relative costs do depend on the level of destruction required, the important technical conclusion is that for conventional levels of assured destruction it is considerably cheaper for White to provide more offensive capability than it is for Black to defend his people and industry against a concerted strike.

As an aside, it might be mentioned that scientists newly engaged in the evaluation of military systems often have trouble grasping that large systems of the type created by or for the military are divided quite rigidly into several chronological stages, namely, in reverse order: operation, deployment, development and research. An operational system is not threatened by a system that is still in development; the threat is not real until the new system is in fact deployed, shaken down and fully operative. This is particularly true for an ABM system, which is obliged to operate against large numbers of relatively independent intercontinental ballistic missiles. It is equally true, however, for counterforce reentry vehicles, which can be ignored unless they are built by the hundreds or thousands. The same goes for MIRV's, a development of the multiple reentry vehicle in which each reentry vehicle is independently directed to a separate target. One must distinguish clearly between the *possibility* of development and the development itself, and similarly between development and actual operation. One must refrain from attributing to a specific defense system, such as Sentinel, those capabilities that *might* be obtained by further development of a different system.

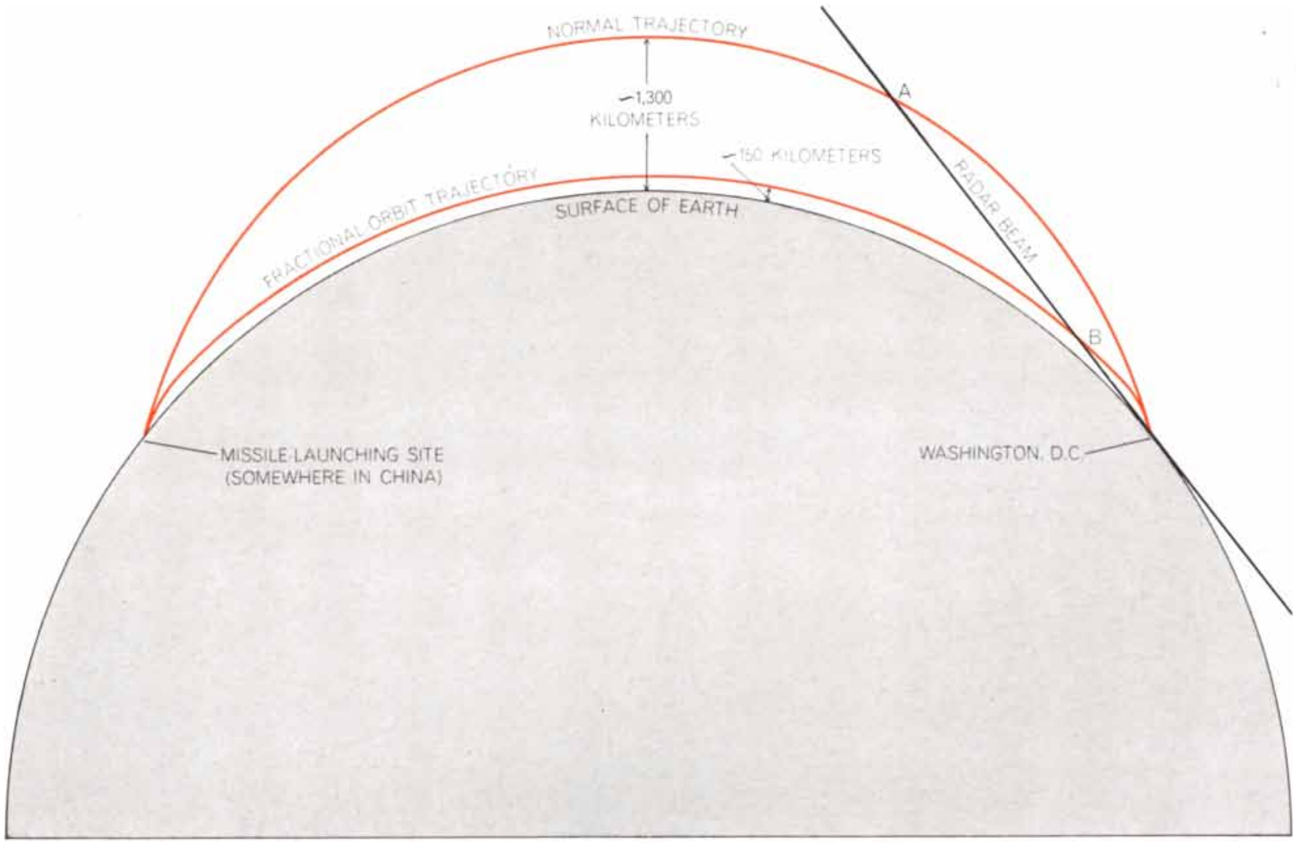
It follows that the Sentinel light ABM system, to be built now and to be operational in the early 1970's against a possible Chinese intercontinental ballistic missile threat, will have to reckon with a missile force unlike either the Russian or the American force, both of which were, after all, built when there was no ballistic-missile defense. The Chinese will probably build even their first operational intercontinental ballistic missiles so that they will have a chance to penetrate. Moreover, we believe it is well within China's capabilities to do a good job at this without intensive testing or tremendous sacrifice in payload.

Temporarily leaving aside penetration aids, there are two pure strategies for attack against a ballistic-missile defense. The first is an all-warhead attack in which one uses large booster rockets to transport many small (that is, fractional-megaton) warheads. These warheads are separated at some instant between the time the missile leaves the atmosphere and the time of reentry. The warheads from one missile can all be directed against the same large target (such as a city); these multiple reentry vehicles (MRV's) are purely a penetration aid. Alternatively each of the reentry vehicles can be given an independent boost to a different target, thus making them into MIRV's. MIRV is not a penetration aid but is rather a counterforce weapon: if each of the reentry vehicles has very high accuracy, then it is conceivable that each of them may destroy an enemy missile silo. The Titan II liquid-fuel rocket, designed more than 10 years ago, could carry 20 or more thermonuclear weapons. If these were employed simply as MRV's, the 54 Titans could provide more than 1,000 reentry vehicles for the defense to deal with.

Since the Spartan interceptors will each cost $1 million to $2 million, including their thermonuclear warheads, it is reasonable to believe thermonuclear warheads can be delivered for less than it will cost the defender to intercept them. The attacker can make a further relative saving by concentrating his strike so that most of the interceptors, all bought and paid for, have nothing to shoot at. This is a high-reliability penetration strategy open to any country that



MECHANISMS FOR KILLING REENTRY VEHICLES include the neutrons, blast and X radiation from a thermonuclear explosion. Neutrons (*1*) can penetrate the fission trigger of an enemy warhead, causing the uranium 235 or plutonium to melt and lose its shape. It can then no longer be assembled for firing. If the defensive warhead is fired inside the atmosphere, the resulting shock front of air (*2*) can cause the incoming reentry vehicle (RV) to decelerate with a force equivalent to several hundred times the force of gravity, thereby leading to its destruction or malfunction. If the explosion is outside the atmosphere, the X rays travel unimpeded to their target. On striking an RV (*3a*) they are absorbed by and intensely heat a thin layer of the RV's heat jacket. This creates a shock front that travels through the jacket (*3b*, *3c*) and may cause the jacket to break up or detach from the RV.



MISSILE TRAJECTORIES can be chosen by the attacker to reduce the effectiveness of the defender's radar. The normal trajectory, which requires the least fuel, carries an ICBM to an altitude of about 1,300 kilometers. On its return to the earth the missile would intersect the path of a horizon-search radar at a distance of about 4,000 kilometers (*A*), when the missile was about 10 minutes away. Longer-range but less precise radars may be able to detect the missile earlier. On a fractional-orbit trajectory the missile would stay so close to the earth that it would not cross the radar horizon (*B*) until it was about 1,400 kilometers, or about three minutes, away.

can afford to spend a reasonable fraction of the amount its opponent can spend for defense.

The second pure strategy for attack against an ABM defense is to precede the actual attack with an all-decoy attack or to mix real warheads with decoys. This can be achieved rather cheaply by firing large rockets from unhardened sites to send light, unguided decoys more or less in the direction of plausible city targets. If the ABM defense is an area defense like the Sentinel system, it must fire against these threatening objects at very long range before they reenter the atmosphere, where because of their lightness they would behave differently from real warheads. Several hundred to several thousand such decoys launched by a few large vehicles could readily exhaust a Sentinel-like system. The attack with real warheads would then follow.

The key point is that since the putative Chinese intercontinental-ballistic-missile force is still in the early research and development stage, it can and will be designed to deal with the Sentinel system, whose interceptors and sensors are nearing production and are rather well publicized. It is much easier to design a missile force to counter a defense that is already being deployed than to design one for any of the possible defense systems that might or might not be deployed sometime in the future.

One of us (Bethe) will now describe (1) the physical mechanisms by which an ABM missile can destroy or damage an incoming warhead and (2) some of the penetration aids available to an attacker who is determined to have his warheads reach their targets.

Much study has been given to the possibility of using conventional explosives rather than a thermonuclear explosive in the warhead of a defensive missile. The answer is that the "kill" radius of a conventional explosive is much too small to be practical in a likely tactical engagement. We shall consider here only the more important effects of the defensive thermonuclear weapon: the emission of neutrons, the emission of X rays and, when the weapon is exploded in the atmosphere, blast.

Neutrons have the ability to penetrate matter of any kind. Those released by defensive weapons could penetrate the heat shield and outer jacket of an offensive warhead and enter the fissile material itself, causing the atoms to fission and generating large amounts of heat. If sufficient heat is generated, the fissile material will melt and lose its carefully designed shape. Thereafter it can no longer be detonated.

The kill radius for neutrons depends on the design of the offensive weapon and the yield, or energy release, of the defensive weapon. The miss distance, or distance of closest approach between the defensive and the offensive missiles, can be made small enough to achieve a kill by the neutron mechanism. This is particularly true if the defensive missile and radar have high performance and the interception is made no more than a few tens of kilometers from the ABM launch site. The neutron-kill mechanism is therefore practical for the short-range defense of a city or other important target. It is highly desirable that the yield of the defensive warhead be kept low to minimize the effects of blast and heat on the city being defended.

The attacker can, of course, attempt to shield the fissile material in the offensive warhead from neutron damage, but the mass of shielding needed is substantial. Witness the massive shield required to keep neutrons from escaping from nuclear reactors. The size of the reentry vehicle will enable the defense to make a rough estimate of the amount of shielding that can be carried and thus to estimate the intensity of neutrons required to melt the warhead's fissile material.

Let us consider next the effect of X rays. These rays carry off most of the energy emitted by nuclear weapons, especially those in the megaton range. If sufficient X-ray energy falls on a reentry vehicle, it will cause the surface layer of the vehicle's heat shield to evaporate. This in itself may not be too damaging, but the vapor leaves the surface at high velocity in a very brief time and the recoil sets up a powerful shock wave in the heat shield. The shock may destroy the heat shield material or the underlying structure.

X rays are particularly effective above the upper atmosphere, where they can travel to their target without being absorbed by air molecules. The defense can therefore use megaton weapons without endangering the population below; it is protected by the intervening atmosphere. The kill radius can then be many kilometers. This reduces the accuracy required of the defensive missile and allows successful interception at ranges of hundreds of kilometers from the ABM launch site. Thus X rays make possible an area defense and provide the key to the Sentinel system.

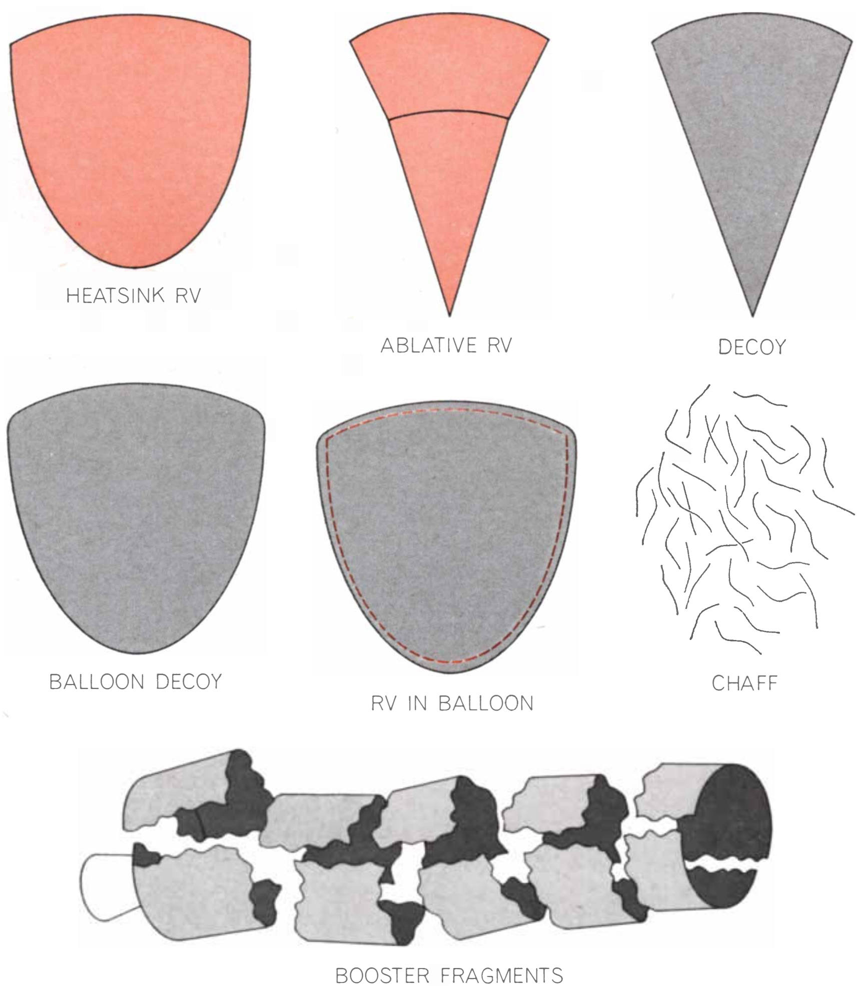
On the other hand, the reentry vehicle can be hardened against X-ray damage to a considerable extent. And in general the defender will not know if the vehicle has been damaged until it reenters the atmosphere. If it has been severely damaged, it may break up or burn up. If this does not happen, however, the defender is helpless unless he has also constructed an effective terminal, or short-range, defense system.

The third kill mechanism—blast—can operate only in the atmosphere and requires little comment. Ordinarily when an offensive warhead reenters the atmosphere it is decelerated by a force that, at maximum, is on the order of 100 g. (One g is the acceleration due to the earth's gravity.) The increased atmospheric density reached within a shock wave from a nuclear explosion in air can produce a deceleration several times greater. But just as one can shield against neutrons and X rays one can shield against blast by designing the reentry vehicle to have great structural strength. Moreover, the defense, not knowing the detailed design of the reentry vehicle, has little way of knowing if it has destroyed a given vehicle by blast until the warhead either goes off or fails to do so.

The main difficulty for the defense is the fact that in all probability the offensive reentry vehicle will not arrive as a single object that can be tracked and fired on but will be accompanied by many other objects deliberately placed there by the offense. These objects come under the heading of penetration aids. We shall discuss only a few of the many types of such aids. They include fragments of the booster rocket, decoys, fine metal wires called chaff, electronic countermeasures and blackout mechanisms of several kinds.

The last stage of the booster that has propelled the offensive missile may disintegrate into fragments or it can be fragmented deliberately. Some of the pieces will have a radar cross section comparable to or larger than the cross section of the reentry vehicle itself. The defensive radar therefore has the task of discriminating between a mass of debris and the warhead. Although various means of discrimination are effective to some extent, radar and data processing must be specifically set up for this purpose. In any case the radar must deal with tens of objects for each genuine target, and this imposes considerable complexity on the system.

There is, of course, an easy way to discriminate among such objects: let the whole swarm reenter the atmosphere. The lighter booster fragments will soon be slowed down, whereas the heavier reentry vehicle will continue to fall with essentially undiminished speed. If a swarm of objects is allowed to reenter,



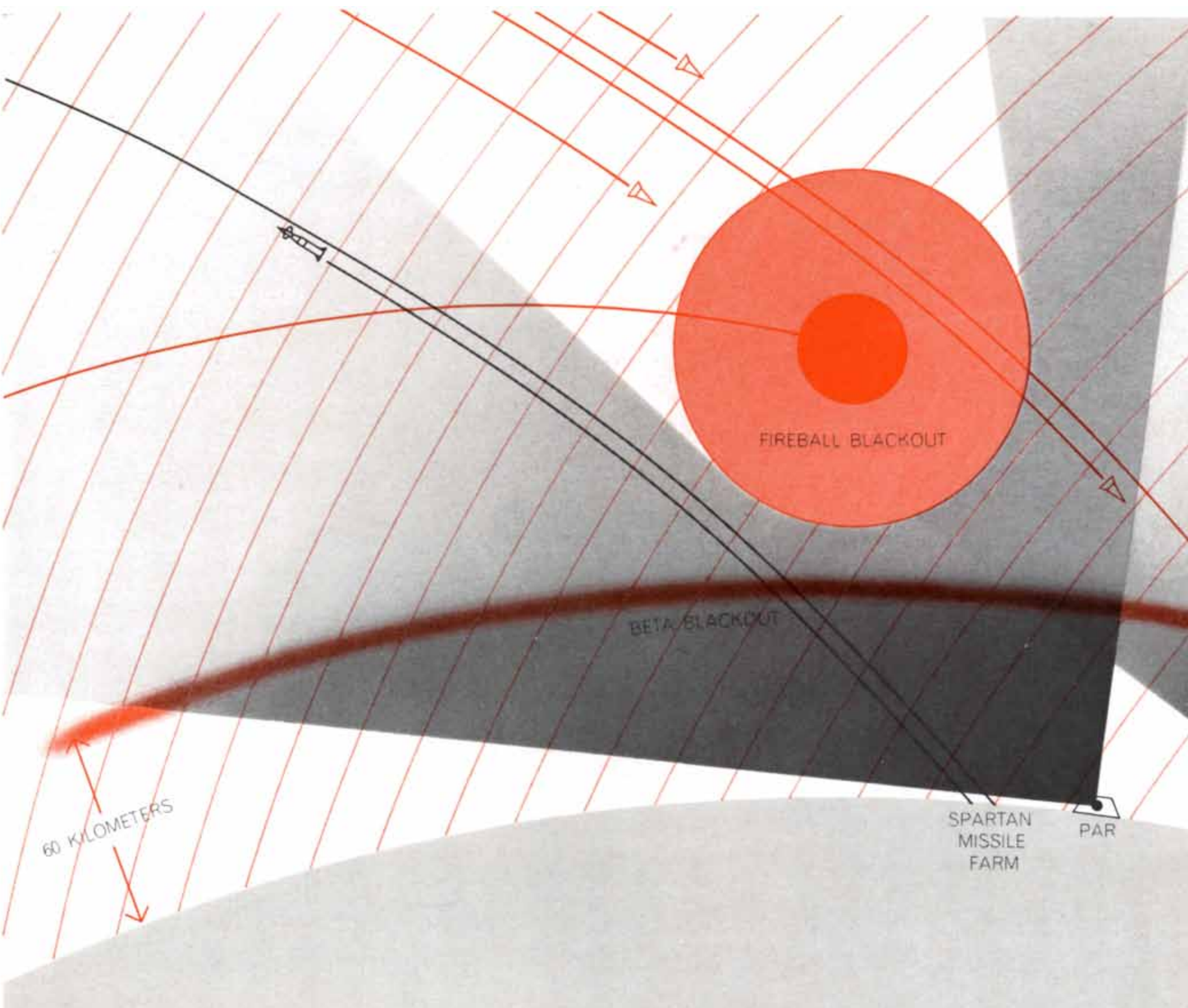
PENETRATION AIDS include objects that will reflect radar signals and thus simulate or conceal actual reentry vehicles (*color*). A decoy might be a simple conical structure or even a metallized balloon. RV's could be placed inside the same kind of balloon. Fragments of the launching vehicle and its fuel tank provide radar reflectors at no cost. Short bits of metal wire, called chaff, also make a cheap and lightweight reflector of radar signals.

however, one must abandon the concept of area defense and construct a terminal defense system. If a nation insists on retaining a pure area defense, it must be prepared to shoot at every threatening object. Not only is this extremely costly but also it can quickly exhaust the supply of antimissile missiles.

Instead of relying on the accidental targets provided by booster fragments, the offense will almost certainly want to employ decoys that closely imitate the radar reflectivity of the reentry vehicle. One cheap and simple decoy is a balloon with the same shape as the reentry vehicle. It can be made of thin plastic covered with metal in the form of foil, strips or wire mesh. A considerable number of such balloons can be carried uninflated by a single offensive missile and released when the missile has risen above the atmosphere.

The chief difficulty with balloons is putting them on a "credible" trajectory, that is, a trajectory aimed at a city or some other plausible target. Nonetheless, if the defending force employs an area defense and really seeks to protect the entire country, it must try to intercept every suspicious object, including balloon decoys. The defense may, however, decide not to shoot at incoming objects that seem to be directed against nonvital targets; thus it may choose to limit possible damage to the country rather than to avoid all damage. The offense could then take the option of directing live warheads against points on the outskirts of cities, where a nuclear explosion would still produce radioactivity and possibly severe fallout over densely populated regions. Worse, the possibility that reentry vehicles can be built to maneuver makes it dangerous to ignore objects even 100 kilometers off target.

Balloon decoys, even more than booster fragments, will be rapidly slowed by the atmosphere and will tend to burn up when they reenter it. Here again a terminal ABM system has a far better



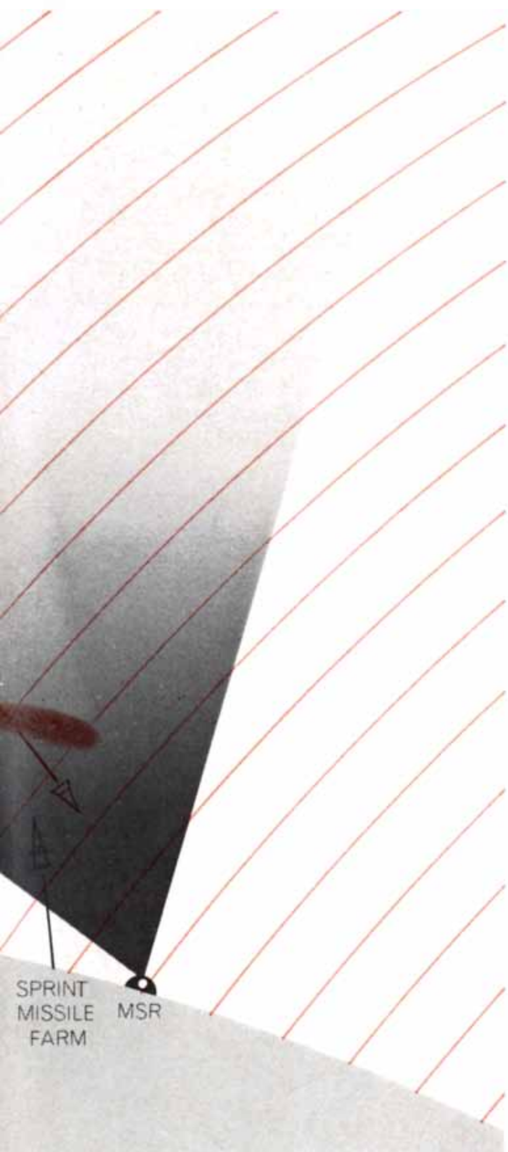
RADAR BLACKOUTS can be created if enough free electrons are released in a sizable volume of space. An attacker can use thermonuclear explosions to generate the electrons required. A fireball blackout results when the heat from a nuclear explosion strips electrons from atoms and molecules of air. In this diagram a high-altitude fireball has been created by an enemy missile launched in a low orbit. The beta rays (electrons) released in the decay of fission products can create another type of blackout. If the beta rays are released at high altitude, they travel along the lines of force in the earth's magnetic field (*parallel colored lines*) and remove electrons from molecules in the atmosphere below. An effective beta blackout could be produced by spreading the fission products of a one-chance than an area defense system to discriminate between decoys and warheads. One possibility for an area system is "active" discrimination. If a defensive nuclear missile is exploded somewhere in the cloud of balloon decoys traveling with a reentry vehicle, the balloons will either be destroyed by radiation from the explosion or will be blown far off course. The reentry vehicle presumably will survive. If the remaining set of objects is examined by radar, the reentry vehicle may stand out clearly. It can then be killed by a second interceptor shot. Such a shoot-look-shoot tactic may be effective, but it obviously places severe demands on the ABM missiles and the radar tracking system. Moreover, it can be countered by the use of small, dense decoys within the balloon swarms.

Moreover, it may be possible to develop decoys that are as resistant to X rays as the reentry vehicle and also are simple and compact. Their radar reflectivity could be made to simulate that of a reentry vehicle over a wide range of frequencies. The decoys could also be made to reenter the atmosphere—at least down to a fairly low altitude—in a way that closely mimicked an actual reentry vehicle. The design of such decoys, however, would require considerable experimentation and development.

Another way to confuse the defensive radar is to scatter the fine metal wires of chaff. If such wires are cut to about half the wavelength of the defensive radar, each wire will act as a reflecting dipole with a radar cross section approximately equal to the wavelength squared divided by 2π. The actual length of the wires is not critical; a wire of a given length is also effective against radar of shorter wavelength. Assuming that the radar wavelength is one meter and that one-mil copper wire is cut to half-meter lengths, one can easily calculate that 100 million chaff wires will weigh only 200 kilograms (440 pounds).

The chaff wires could be dispersed over a large volume of space; the chaff could be so dense and provide such large radar reflection that the reentry vehicle could not be seen against the background noise. The defense would then not know where in the large reflecting cloud the reentry vehicle is concealed. The defense would be induced to spend several interceptors to cover the entire cloud, with no certainty, even so, that the hidden reentry vehicle will be killed. How much of the chaff would survive the defensive nuclear explosion is another difficult question. The main problem for the attacker is to develop a way to disperse chaff more or less uniformly.

An active alternative to the use of chaff is to equip some decoys with electronic devices that generate radio noise at frequencies selected to jam the defensive radar. There are many variations on such electronic countermeasures, among them the use of jammers on the reentry vehicles themselves.



megaton explosion over an area 200 kilometers in radius. For several minutes the electron cloud would heavily absorb the long waves emitted by a PAR unit, here aimed to the north. The shorter MSR waves, however, would be attenuated only briefly.

The last of the penetration aids that will be mentioned here is the radar blackout caused by the large number of free electrons released by a nuclear explosion. These electrons, except for a few, are removed from atoms or molecules of air, which thereby become ions. There are two main causes for the formation of ions: the fireball of the explosion, which produces ions because of its high temperature, and the radioactive debris of the explosion, which releases beta rays (high-energy electrons) that ionize the air they traverse. The second mechanism is important only at high altitude.

The electrons in an ionized cloud of gas have the property of bending and absorbing electromagnetic waves, particularly those of low frequency. Attenuation can reach such high values that the defensive radar is prevented from seeing any object behind the ionized cloud (unlike chaff, which confuses the radar only at the chaff range and not beyond).

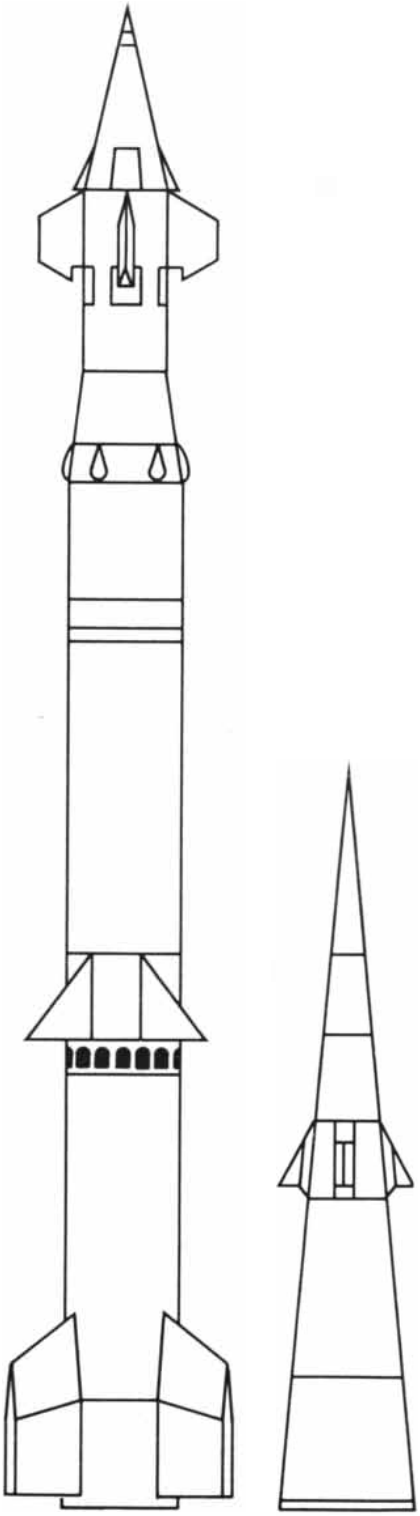
Blackout is a severe problem for an area defense designed to intercept missiles above the upper atmosphere. The problem is aggravated because area-defense radar is likely to employ low-frequency (long) waves, which are the most suitable for detecting enemy missiles at long range. In some recent popular articles long-wave radar has been hailed as the cure for the problems of the ABM missile. It is not. Even though it increases the capability of the radar in some ways, it makes the system more vulnerable to blackout.

Blackout can be caused in two ways: by the defensive nuclear explosions themselves and by deliberate explosions set off at high altitude by the attacker. Although the former are unavoidable, the defense has the choice of setting them off at altitudes and in locations that will cause the minimum blackout of its radar. The offense can sacrifice a few early missiles to cause blackout at strategic locations. In what follows we shall assume for purposes of discussion that the radar wavelength is one meter. Translation to other wavelengths is not difficult.

In order to totally reflect the one-meter waves from our hypothetical radar it is necessary for the attacker to create an ionized cloud containing 10^9 electrons per cubic centimeter. Much smaller electron densities, however, will suffice for considerable attenuation. For the benefit of technically minded readers, the equation for attenuation in decibels per kilometer is

$$\alpha = \frac{4.34}{3 \times 10^5} \frac{\omega_p^2}{\omega^2 + \gamma_c^2} \gamma_c .$$

Here ω_p is the plasma frequency for the given electron density, ω is the radar frequency in radians per second and γ_c is the frequency of collisions of an electron with atoms of air. At normal tempera-



SENTINEL MISSILES are the long-range Spartan (*left*), with a reported range of several hundred kilometers, and the Sprint (*right*), which will be used for terminal defense, usually below 35 kilometers. Both will be equipped with thermonuclear warheads designed to destroy or disable the attacker's bomb-carrying reentry vehicles. The Spartan, about 55 feet long, is a three-stage solid-fuel rocket. The smaller and faster Sprint missile uses two stages of solid fuel.

tures this frequency γ_c is the number 2×10^{11} multiplied by the density of the air (ρ) compared with sea-level density (ρ_0), or $\gamma_c = 2 \times 10^{11}\ \rho/\rho_0$. At altitudes above 30 kilometers, where an area-defense system will have to make most of its interceptions, the density of air is less than .01 of the density at sea level. Under these conditions the electron collision frequency γ_c is less than the value of $\omega = (2\pi \times 3 \times 10^8)$ and therefore can be neglected in the denominator of the equation. Using that equation, we can then specify the number of electrons, N_e, needed to attenuate one-meter radar waves by a factor of more than one decibel per kilometer: $N_e > 350\rho_0/\rho$. At an altitude of 30 kilometers, where ρ_0/ρ is about 100, N_e is about 3×10^4, and at 60 kilometers N_e is still only about 3×10^6. Thus the electron densities needed for the substantial attenuation of a radar signal are well under the 10^9 electrons per cubic centimeter required for total reflection. The ionized cloud created by the fireball of a nuclear explosion is typically 10 kilometers thick; if the attenuation is one decibel per kilometer, such a cloud would produce a total attenuation of 10 decibels. This implies a tenfold reduction of the outgoing radar signal and another tenfold reduction of the reflected signal, which amounts to effective blackout.

The temperature of the fireball created by a nuclear explosion in the atmosphere is initially hundreds of thousands of degrees centigrade. It quickly cools by radiation to about 5,000 degrees C. Thereafter cooling is produced primarily by the cold air entrained by the fireball as it rises slowly through the atmosphere, a process that takes several minutes.

When air is heated to 5,000 degrees C., it is strongly ionized. To produce a radar attenuation of one decibel per kilometer at an altitude of 90 kilometers the fireball temperature need be only 3,000 degrees, and at 50 kilometers a temperature of 2,000 degrees will suffice. Ionization may be enhanced by the presence in the fireball of iron, uranium and other metals, which are normally present in the debris of nuclear explosion.

The size of the fireball can easily be estimated. Its diameter is about one kilometer for a one-megaton explosion at sea level. For other altitudes and yields there is a simple scaling law: the fireball diameter is equal to $(Y\rho_0/\rho)^{1/3}$, where Y is the yield in megatons. Thus a fireball one kilometer in diameter can be produced at an altitude of 30 kilometers (where $\rho_0/\rho = 100$) by an explosion of only 10 kilotons. At an altitude of 50 kilometers (where $\rho_0/\rho = 1{,}000$), a one-megaton explosion will produce a fireball 10 kilometers in diameter. At still higher altitudes matters become complicated because the density of the atmosphere falls off so sharply and the mechanism of heating the atmosphere changes. Nevertheless, fireballs of very large diameter can be expected when megaton weapons are exploded above 100 kilometers. These could well black out areas of the sky measured in thousands of square kilometers.

For explosions at very high altitudes (between 100 and 200 kilometers) other phenomena become significant. Collisions between electrons and air molecules are now unimportant. The condition for blackout is simply that there be more than 10^9 electrons per cubic centimeter.

At the same time very little mass of air is available to cool the fireball. If the air is at first fully ionized by the explosion, the air molecules will be dissociated into atoms. The atomic ions combine very slowly with electrons. When the density is low enough, as it is at high altitude, the recombination can take place only by radiation. The radiative recombination constant (call it C_R) is about 10^{-12} cubic centimeter per second. When the initial electron density is well above 10^9 per cubic centimeter, the number of electrons remaining after time t is roughly equal to $1/C_R t$. Thus if the initial electron density is 10^{12} per cubic centimeter, the density will remain above 10^9 for 1,000 seconds, or some 17 minutes. The conclusion is that nuclear explosions at very high altitude can produce long-lasting blackouts over large areas.

The second of the two mechanisms for producing an ionized cloud, the beta rays issuing from the radioactive debris of a nuclear explosion, can be even more effective than the fireball mechanism. If the debris is at high altitude, the beta rays will follow the lines of force in the earth's magnetic field, with about half of the beta rays going immediately down into the atmosphere and the other half traveling out into space before returning earthward. These beta rays have an average energy of about 500,000 electron volts, and when they strike the atmosphere, they ionize air molecules. Beta rays of average energy penetrate to an altitude of about 60 kilometers; some of the more energetic rays go down to about 50 kilometers. At these levels, then, a high-altitude explosion will give

PROTOTYPE OF MISSILE-SITE RADAR is the multifunction array radar (MAR) at the Army's White Sands Missile Range in New Mexico. MSR, a smaller version of MAR, will be used to guide Spartans and Sprints to their targets. MAR provided an early demonstration that the fastest way to aim a radar beam at different parts of the sky is to switch the beam electronically.

rise to sustained ionization as long as the debris of the explosion stays in the vicinity.

One can show that blackout will occur if $y \times t^{-1.2} > 10^{-2}$, where t is the time after the explosion in seconds and y is the fission yield deposited per unit horizontal area of the debris cloud, measured in tons of TNT equivalent per square kilometer. The factor $t^{-1.2}$ expresses the rate of decay of the radioactive debris. If the attacker wishes to cause a blackout lasting five minutes ($t = 300$), he can achieve it with a debris level y equal to 10 tons of fission yield per square kilometer. This could be attained by spreading one megaton of fission products over a circular area about 400 kilometers in diameter at an altitude of, say, 60 kilometers. Very little could be seen by an area-defense radar attempting to look out from under such a blackout disk. Whether or not such a disk could actually be produced is another question. Terminal defense would not, of course, be greatly disturbed by a beta ray blackout.

The foregoing discussion has concentrated mainly on the penetration aids that can be devised against an area-defense system. By this we do not mean to suggest that a terminal-defense system can be effective, and we certainly do not wish to imply that we favor the development and deployment of such a system.

Terminal defense has a vulnerability all its own. Since it defends only a small area, it can easily be bypassed. Suppose that the 20 largest American cities were provided with terminal defense. It would be easy for an enemy to attack the 21st largest city and as many other undefended cities as he chose. Although the population per target would be less than if the largest cities were attacked, casualties would still be heavy. Alternatively the offense could concentrate on just a few of the 20 largest cities and exhaust their supply of antimissile missiles, which could readily be done by the use of multiple warheads even without decoys.

It was pointed out by Charles M. Herzfeld in *The Bulletin of the Atomic Scientists* a few years ago that a judicious employment of ABM defenses could equalize the risks of living in cities of various sizes. Suppose New York, with a population of about 10 million, were defended well enough to require 50 enemy warheads to penetrate the defenses, plus a few more to destroy the city. If cities of 200,000 inhabitants were left undefended, it would be equally "attractive" for an enemy to attack New York and penetrate its defenses as to attack an undefended city.

Even if such a "logical" pattern of ABM defense were to be seriously proposed, it is hard to believe that people in the undefended cities would accept their statistical security. To satisfy everyone would require a terminal system of enormous extent. The highest cost estimate made in public discussions, $50 billion, cannot be far wrong.

Although such a massive system would afford some protection against the U.S.S.R.'s present armament, it is virtually certain that the Russians would react to the deployment of the system. It would be easy for them to increase the number of their offensive warheads and thereby raise the level of expected damage back to the one now estimated. In his recent forecast of defense needs for the next five years, Secretary McNamara estimated the relative cost of ABM defenses and the cost of countermeasures that the offense can take. He finds invariably that the offense, by spending considerably less money than the defense, can restore casualties and destruction to the original level before defenses were installed. Since the offense is likely to be "conservative," it is our belief that the actual casualty figures in a nuclear exchange, after both sides had deployed ABM systems and simultaneously increased offensive forces, would be worse than these estimates suggest.

Any such massive escalation of offensive and defensive armaments could hardly be accomplished in a democracy without strong social and psychological effects. The nation would think more of war, prepare more for war, hate the potential enemy and thereby make war more likely. The policy of both the U.S. and the U.S.S.R. in the past decade has been to reduce tensions to provide more understanding, and to devise weapon systems that make war less likely. It seems to us that this should remain our policy.

HUMAN CELLS AND AGING

When normal cells are grown outside the body, their capacity to survive dwindles after a period of time. This deterioration may well represent aging and an ultimate limit to the span of life

by Leonard Hayflick

The common impression that modern medicine has lengthened the human life-span is not supported by either vital statistics or biological evidence. To be sure, the 20th-century advances in control of infectious diseases and of certain other causes of death have improved the longevity of the human population as a whole. These accomplishments in medicine and public health, however, have merely extended the *average* life expectancy by allowing more people to reach the upper limit, which for the general run of mankind still seems to be approximately the Biblical fourscore years. Aging and a limited life-span apparently are characteristic of all animals that stop growing after reaching a fixed, mature size. In the case of man, after the age of 30 there is a steady, inexorable increase in the probability of death from one cause or another; the probability doubles about every eight years as one grows older. This general probability is such that even if the major causes of death in old age—heart disease, stroke and cancer—were eliminated, the average life expectancy would not be lengthened by much more than 10 years. It would then be about 80 years instead of the expectancy of about 70 years that now prevails in advanced countries.

Could man's life-span be extended—or is there an inescapable aging mechanism that restricts human longevity to the present apparent limit? Until recently few biologists ventured to attempt to explore the basic processes of aging; obviously the subject does not easily lend itself to detailed study. It is now receiving considerable attention, however, in a number of laboratories [see "The Physiology of Aging," by Nathan W. Shock; SCIENTIFIC AMERICAN, January, 1962]. In this article I shall discuss some new findings at the cellular level.

No doubt many mechanisms are involved in the aging of the body. At the cell level at least three aging processes are under investigation. One is a possible decline in the functional efficiency of nondividing, highly specialized cells, such as nerve and muscle cells. Another is the progressive stiffening with age of the structural protein collagen, which constitutes more than a third of all the body protein and serves as the general binding substance of the skin, muscular and vascular systems [see "The Aging of Collagen," by Frederic Verzár; SCIENTIFIC AMERICAN, April, 1963]. In our own laboratory at the Wistar Institute we have addressed ourselves to a third question: the limitation on cell division. Our studies have focused particularly on the structural cells called fibroblasts, which produce collagen and fibrin. These cells, like certain other "blast" cells, go on dividing in the adult body. We set out to determine whether human fibroblasts in a cell culture could divide indefinitely or had only a finite capacity for doing so.

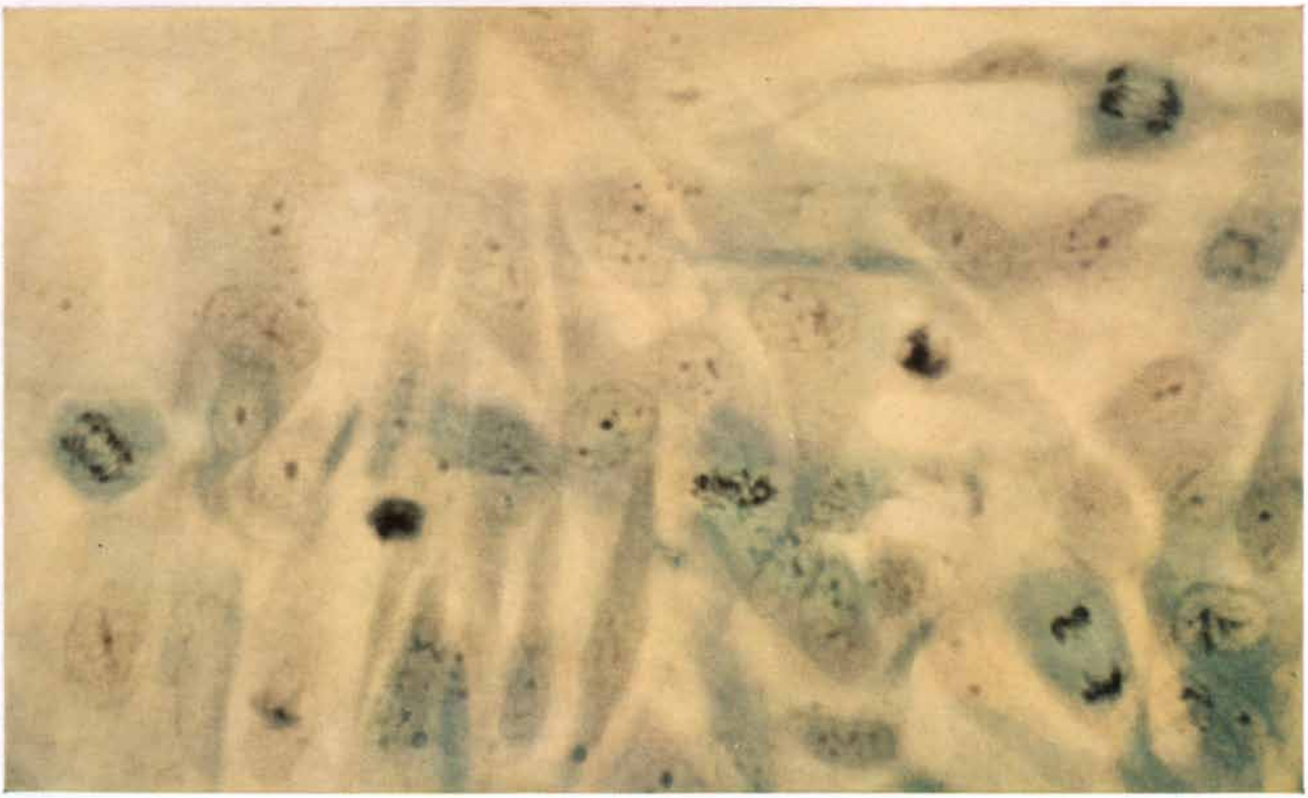
Alexis Carrel's famous experiments more than a generation ago suggested that animal cells per se (that is, cells removed from the body's regulatory mechanisms) might be immortal. He apparently succeeded in keeping chick fibroblasts growing and multiplying in glass vessels for more than 30 years—a great deal longer than a chicken's life expectancy. Later experimenters reported similar successes with embryonic cells from laboratory mice. It has since been learned, however, through improved techniques and a better understanding of cell cultures, that the conclusions drawn from those early experiments were erroneous.

In the case of chick fibroblasts it has been repeatedly demonstrated that, if care is taken not to add any living cells to the initial population in the glass vessel, the cell colony will not survive long. The early cultures, including Carrel's, were fed a crude extract taken from chick embryos, and it is now believed these feedings must have contained some living chick cells. That is to say, in all probability the reason the cultures continued to grow indefinitely was that new, viable fibroblasts were introduced into the culture at each feeding.

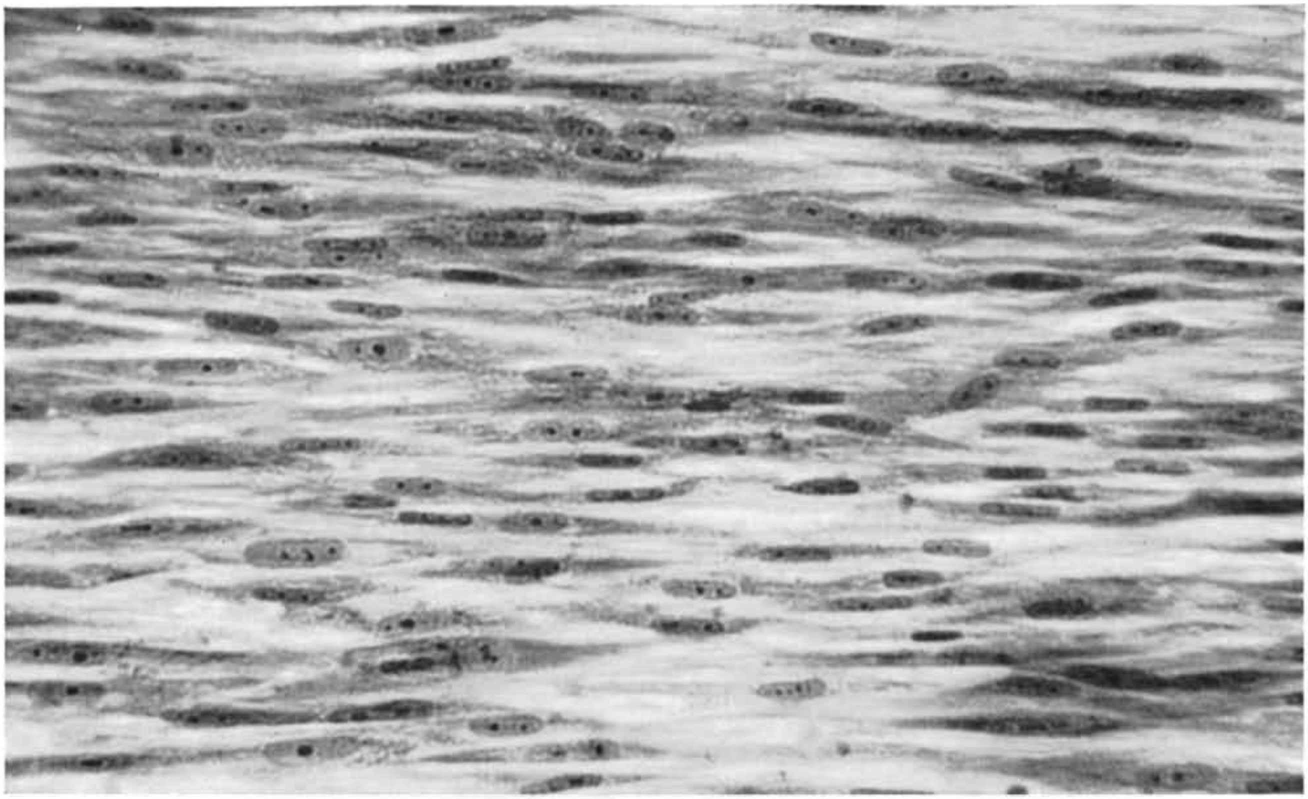
Restudy of the experiments in culturing mouse cells has brought to light a highly interesting fact. It has been found that when normal cells from a laboratory mouse are cultured in a glass vessel, they frequently undergo a spontaneous transformation that enables them to divide and multiply indefinitely. This type of transformation takes place regularly in cultures of mouse cells but only rarely in cultures of the fibroblasts of man or other animals. These transformed cell populations have several abnormal properties, but they are truly immortal: many of the mouse-derived cultures have survived for decades. Similarly, the famous line of transformed human cells called HeLa, originally derived from cervical tissue in 1952 by George O. Gey of the Johns Hopkins University School of Medicine, is still growing and multiplying in glass cultures.

On microscopic examination the transformed cells show themselves to be indeed abnormal. Instead of the normal number of 46 chromosomes in a human diploid cell, the "mixoploid" HeLa cells may have anywhere from 50 to 350 chromosomes per cell. They differ from normal chromosomes in size and shape and also stain differently. Moreover, they often behave like cancer cells: inoculated into a suitable laboratory animal, they can grow as tumors. This property of transformed cells has become an important tool in investigations of the

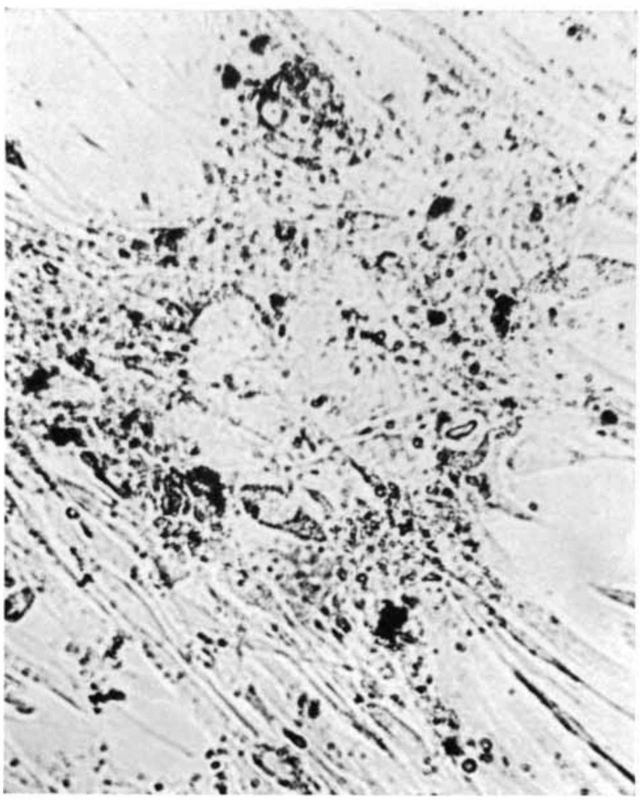
HUMAN CHROMOSOMES are seen magnified 3,000 diameters in a normal diploid cell grown in culture. The chromosomes assume this compact form during mitotic cell division. When grown in culture, human cells display a limited capacity to divide; their finite longevity may be related to the finite span of human life. The chromosomes in the picture are stained with a dye called Giemsa.



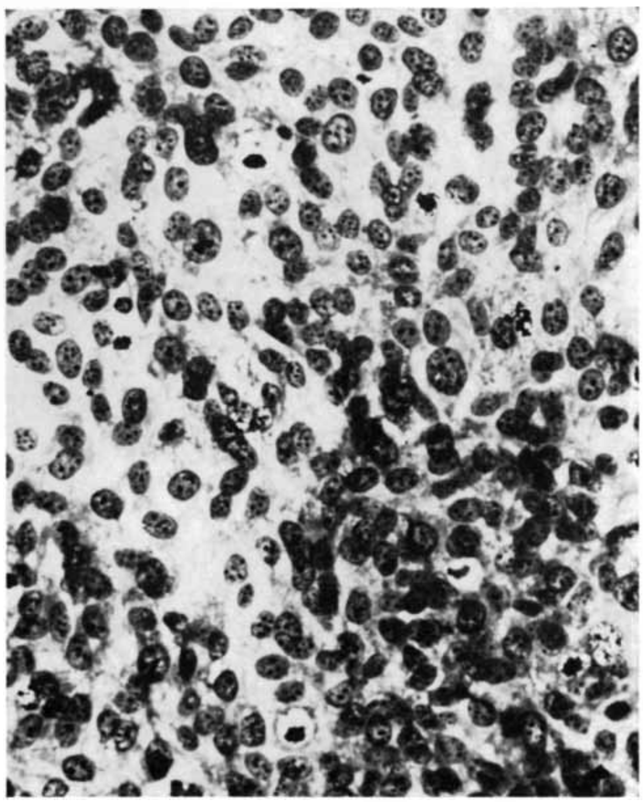
STAGES OF CELL DIVISION appear in this photomicrograph of human cells. The chromosomes are seen as dark clusters. Those in the nucleus at the center have not begun to divide. At left is a nucleus in which the chromosome cluster has separated into two parts; at right, top and bottom, are two nuclei in a later stage of the division cycle. The culture is stained with aceto-orcein and green counterstain. The magnification is 750 diameters. Both photomicrographs were made by Paul S. Moorhead of the Wistar Institute.



GROWING HUMAN CELLS cover the surface of the glass vessel in which they were planted. They form a layer one cell deep. Under proper conditions this normal cell population will continue to proliferate; it will not, except in unusual cases, multiply indefinitely.



AGED HUMAN CELLS are irregular in appearance; they no longer divide. The aging of such normal cell populations is apparently due to an intrinsic process, not a deficiency in growing conditions. The cells were taken from lung tissue and grown in glass. This photomicrograph and the one at top were made by the author.



IMMORTAL HUMAN CELLS appear in this photomicrograph made by Fred C. Jensen of the Wistar Institute. The cells, once normal like those at top, have been treated with the monkey virus SV-40; thus transformed, they can apparently multiply indefinitely. The magnification in these photomicrographs is 300 diameters.

genesis of cancer. Although the spontaneous transformation of human cells is rare, investigators of cancer are making use of the recent discovery that normal human cells can be routinely transformed into cancer cells by exposing them to the monkey virus known as SV-40.

A crucial consideration for the relevance of cells in culture to aging, of course, is that they are normal cells. Our interest is therefore directed not to abnormal cells but to the observation that normal cells do not divide indefinitely. Whereas a population of transformed cells will proliferate and survive for decades in cell culture, no one has succeeded in perpetuating a culture of normal animal cells. The same is true of cells implanted in a living animal. Transformed cells will go on growing indefinitely in a series of tissue transplants from animal to animal, but normal cells will not. Peter L. Krohn of the University of Birmingham has shown, for example, that normal mouse skin can survive only a limited number of serial grafts from one mouse to another in the same inbred strain.

Over the past seven years in our laboratory Paul S. Moorhead and I have been studying cell cultures of normal human fibroblasts. Unlike highly specialized cells, fibroblasts (which serve as the structural bricks for most body tissues) will grow and multiply in a nutrient medium in glass bottles. We have used lung tissue as our principal source of the cells. We break down the tissue into separated cells by means of the digestive enzyme trypsin, then remove the trypsin by centrifugation and seed the cells in a bottle containing a suitable growing medium. After a few days of incubation at 99 degrees Fahrenheit the fibroblasts have spread out on the glass surface and begun to divide. In a week or so they cover the entire available surface with a layer one cell deep. Since they will proliferate only in a single layer, we strip off the layer that has covered the bottle surface, again separate the cells with trypsin and plant half of the cells in each of two new bottles with fresh medium. In three or four days the inoculated fibroblasts grow over all the available surface in each bottle, thus doubling the number of cells taken from the original bottle. As this procedure is repeated at four-day intervals, the fibroblasts continue to proliferate in new bottles, doubling in number each time.

We found that fibroblasts taken from four-month-old human embryos doubled in this way about 50 times (the limit ranged between 40 and 60 doublings). After reaching this limit of capacity for division the cell population died. It could therefore be concluded that human fibroblasts derived from embryonic tissue and grown in cell culture have a finite lifetime amounting to approximately 50 population doublings (which in our culture covered a span of six to eight months).

Further study reinforced this conclusion. It turned out, for example, that if cell division was interrupted and then resumed, the total number of population doublings was not altered. In our experiments we did not, of course, double the number of bottles at each step; after only 10 doubling passages we would have had 1,024 bottles, and 50 doublings of the original seeding of fibroblasts could have produced about 20 million tons of cells! To keep the yield within reasonable bounds we set aside most of the cells from the subcultivations and put them in cold storage. We found they could be kept in suspended animation for apparently unlimited periods; even after six years in storage they proved to be capable of resuming division when they were thawed and placed in a culture medium. They "remembered" the doubling level they had reached before storage and completed the course from that point. For example, cells that had been stored at the 30th doubling went on to divide about 20 more times.

The geometric rate of increase of the cells in culture has made it possible to provide essentially unlimited supplies of the cells for experiments. Samples of one of our strains of normal human fibroblasts (WI-38), "banked" after the eighth doubling in liquid nitrogen at 190 degrees below zero centigrade, have been distributed to hundreds of research laboratories around the world. The stored cells presumably would be available for study far in the future. For instance, well-protected capsules containing frozen cells might be buried in Antarctica or deposited in orbit in the cold of outer space for retrieval many generations hence. Investigators might then be able to use them to study, among other things, whether or not time had brought about any evolutionary change in the aging of man or other animals at the cellular level.

We found further confirmation of the finite lifetime of human fibroblasts when we cultured such cells from adult donors. Samples of lung fibroblasts taken at the time of death from eight adults, ranging in age from 20 to 87, underwent 14 to 29 doublings in cell culture afterward. The number of doublings in these tests did not show a clear correlation with the age of the donor, but presumably this was because our method of measuring the doubling of cell populations in bottles is not sufficiently precise to disclose such a correlation in detail. Our current experiments do suggest, however, that consistent differences can be found between broad age groups. It appears that fibroblasts from human embryos will divide in cell culture 50 ± 10 times, those from persons between birth and the age of about 20 will divide 30 ± 10 times and those from donors over 20 will divide 20 ± 10 times.

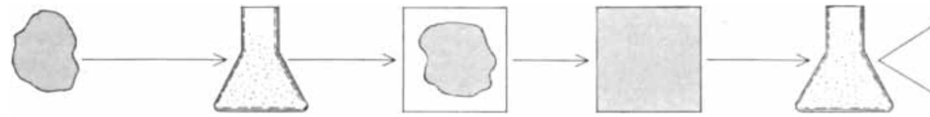
We have tested fibroblasts from several human embryonic tissues besides lung tissue and found that they too are limited to a total of about 50 divisions in culture. The doubling lifetimes of fibroblasts from animals other than man have also been studied. As one would expect, the cells of the shorter-lived vertebrates show less capacity for division. For example, normal fibroblasts from embryos of chickens, rats, mice, hamsters and guinea pigs usually double no more than 15 times in cell culture, and cells that have been taken from adults of the same species undergo considerably fewer than 15 divisions.

Early in our experiments it became evident that we had to examine the possibility that a lethal factor in the culture medium or a defect in the culture technique might be responsible for the limitation of division and ultimate death of the cells. Did the cells stop dividing because of some lack in the nutrient mixture or the presence of contaminating microorganisms such as viruses? We explored these possibilities by various experiments, one of which consisted in culturing a mixture of normal fibroblasts taken from male and female donors. Female cells can be distinguished from male cells either by the presence of special chromatin bodies (found only in female cells) or by the visible difference between the *XX* female sex chromosomes and the *XY* male chromosomes. As a consequence we were able to use these cell "markers" as a label for following the progeny of the respective original parents.

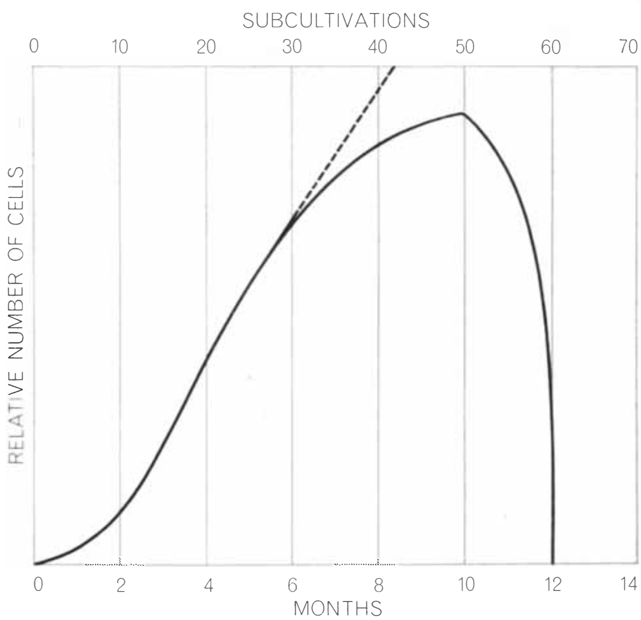
We seeded a bottle with a certain number of fibroblasts from a male population that had undergone 40 doublings and with an equal number of fibroblasts from a female population that had gone through only 10 doublings. If some inadequacy of the culture or the accumulation of a killing factor in the medium were the primary cause of cell death in our cultures, then in this experiment the number of doublings in culture should have been the same for both the male and the female cells after we had mixed them together; there is no reason to suppose that a nutritional inadequacy or a lethal factor such as a virus would act preferentially on the cells of a particular sex. Actually the male and female cells composing the mixture, presumably because of their difference in age, proved to have sharply different survival rates. Most of the male population, having already undergone 40 doublings before the mixture was made, died off after 10 more doublings of the mixed culture, whereas the "young" female population of cells in the same mixed culture, with only 10 previous doublings, went on dividing for many more than the male. After 25 doublings all the male cells had disappeared and the culture contained only female cells. These results appear to confirm in an unambiguous way that the life-span of fibroblast populations in our cell cultures is determined by intrinsic aging of the cells rather than by external agencies.

Does this aging result from depletion or dilution of the cells' own chemical resources? We considered the possibility that the eventual death of the cells might be attributable to the exhaustion of some essential metabolite the cells could not synthesize from the culture medium. If that were the case, however, the original store of this substance must be very large indeed to enable the cells to multiply for 50 generations. Simple mathematics showed that in order to provide at least one molecule of the hypothetical substance for each cell by the 50th doubling, the original cell would have to have at least three times its known weight even if the substance in question were the lightest element (hydrogen) and the original parent cell were composed entirely of that element!

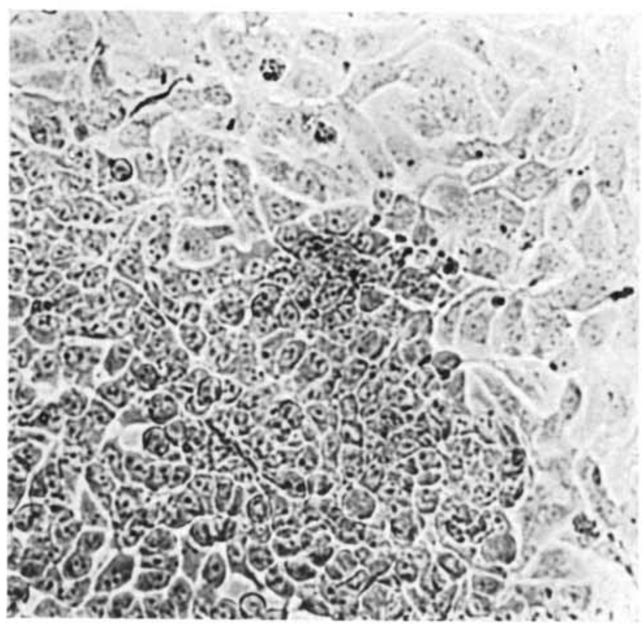
By isolating individual cells and developing clones (colonies) from them we have been able to establish that each human fibroblast from an embryo is capable of giving rise to about 50 doubling generations in cell culture. As the cell proliferation proceeds there is a gradual decline in the capacity for reproduction.



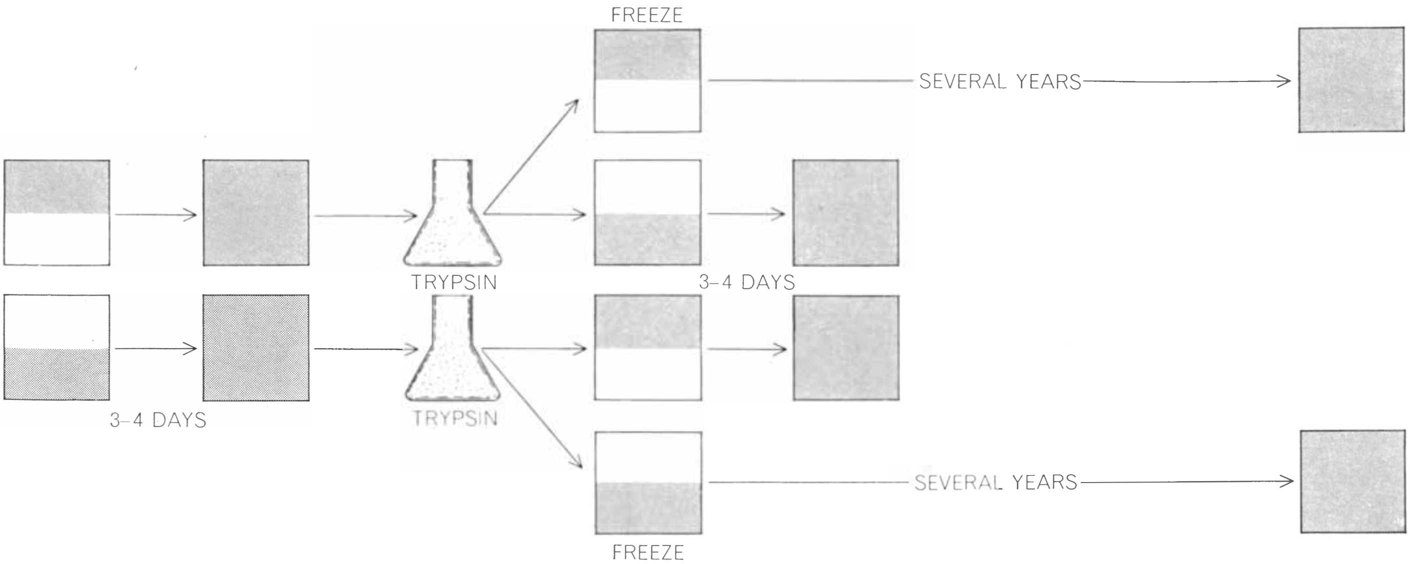
CULTIVATION OF HUMAN CELLS in the author's laboratory begins with the breakdown of lung tissue into separate cells. This is accomplished by means of the digestive enzyme



LIFETIME OF HUMAN CELLS was determined by allowing a population to multiply until it had doubled in size. After a culture of cells from embryonic tissue had grown to a particular point, it was divided in two (*see illustration at top of these two pages*). Cell division ceased after about 50 such subcultivations had doubled. It is possible at any time (although it is rare) for a spontaneous change to occur after which the cells multiply indefinitely (*broken line*).



TRANSFORMED HUMAN CELLS are distinguished by their morphology and reaction to staining. The darker amnion cells, forming a large island at lower left, have undergone a spontaneous transformation. They will continue to divide after the neighboring cells have died. Transformed, or "mixoploid," cells have more chromosomes than diploid cells do; they are utilized in cancer research. The magnification of this photomicrograph is about 180 diameters.



trypsin. After the cells, seeded in a bottle, have multiplied to cover its surface, they are again treated with trypsin, then divided into two halves and replanted. Most cells thus grown are placed in cold storage. Thawed and planted years later, they resume division.

Investigators in several laboratories have found that in successive generations of the multiplying cells a larger and larger fraction of the progeny becomes incapable of dividing until, by the 45th or 50th doubling generation, the entire population has lost this ability.

Curiously, as the population of human fibroblasts approaches the end of its lifetime, aberrations often crop up in the chromosomes. Chromosome aberrations and cell-division peculiarities related to age have also been observed in human leucocytes and in the liver tissue of living mice. The question of whether or not cell abnormalities are a common accompaniment of human aging remains a moot point, however; there is no clear clinical or laboratory evidence on the matter.

Our information on the aging of cells so far is limited to what we can observe in cell cultures in glassware. It has not yet been established that fibroblasts behave the same way in living animals as they do in an artificial culture. However, in man several organs lose weight after middle age, and this is directly attributable to cell loss. The human brain weighs considerably less in old age than it does in the middle years, the kidney also shows a large reduction in nephrons accompanying cell loss and the number of taste buds per papilla of the tongue drops from an average of 245 in young adults to 88 in the aged. We cannot be certain that fibroblasts stop dividing or divide at a lower rate as an animal ages or that the bodily signs of aging can be explained on that basis. It is well known, however, that certain cell systems in animals do stop dividing and die in the normal course of development. Familiar examples are the larval tissues of insects, the tail and gills of the tadpole and some embryonic kidney tissues (the pronephros and mesonephros) in the higher vertebrates.

Reviewing these phenomena, John W. Saunders, Jr., of the School of Veterinary Medicine of the University of Pennsylvania has suggested that the death of cells resulting in the demise of specific tissues is a normal, programmed event in the development of multicellular animals. By the same reasoning we can surmise that the aging and finite lifetime of normal cells constitute a programmed mechanism that sets an overall limit on an organism's length of life. This would suggest that, even if we were able to checkmate all the incidental causes of human aging, human beings would inevitably still succumb to the ultimate failure of the normal cells to divide or function.

In this connection it is interesting to consider what engineers call the "mean time to failure" in the lifetime of machines. Every machine embodies built-in obsolescence (intentional or unintentional) in the sense that its useful lifetime is limited and more or less predictable from consideration of the durability of its parts. By the repair or replacement of elements of a machine as they fail its lifetime can be extended, but barring total replacement of all the elements eventual "death" of the machine as a functioning system is inevitable.

What might determine the "mean time to failure" of an animal organism? I suggest that animal aging may result from deterioration of the genetic program that orchestrates the development of cells. As time goes on, the DNA of dividing cells may become clouded with an accumulation of copying errors (analogous to the "noise" that develops in the serial copying of a photograph). The coding and decoding system that governs the replication of DNA operates with a high degree of accuracy, but the accuracy is not absolute. Moreover, there is some experimental evidence that, as Leslie E. Orgel of the Salk Institute for Biological Studies has suggested, certain enzymes involved in the transcription of information from DNA for the synthesis of proteins may deteriorate with age. At all events, since the ability of cells to divide or to function is controlled by the inherited information-containing molecules, it seems likely that some inherent degeneration of these molecules may hold the key to the aging and eventual death of cells.

Pursuing the machine analogy, we might surmise that man is endowed with a longer life-span than other mammals because human cells have evolved a more effective system for correcting or repairing errors as they arise. Such an evolution would account for the generally progressive lengthening of the fixed life-span from the lower to the higher animals; presumably the march of evolution has developed improvements in the cells' error-repairing mechanisms. It is clear, however, that even in man this system is far from perfect. In the idiom of computer engineers we might say that man, like all other animals, has a "mean time to failure" because his normal cells eventually run out of accurate program and capacity for repair.

Obsidian and the Origins of Trade

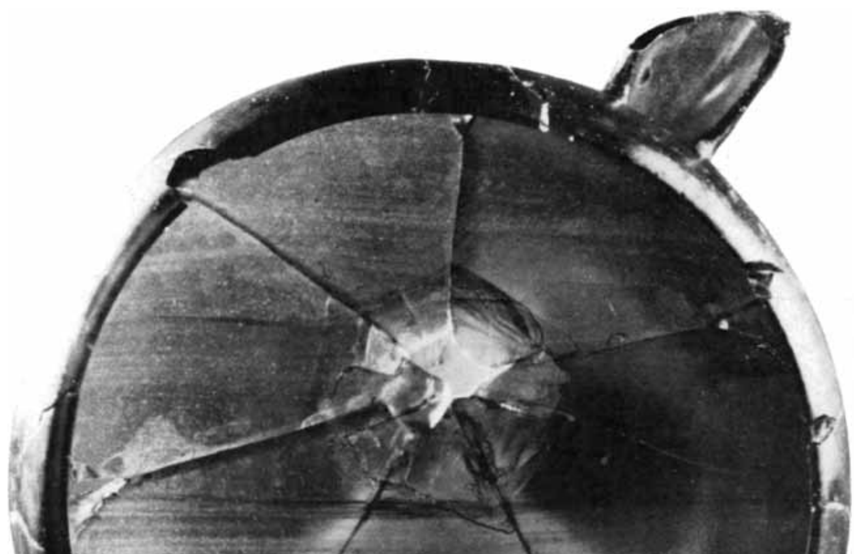
Objects made of this volcanic glass are found at many Neolithic sites around the Mediterranean. Spectroscopic analysis indicates that the raw material often came from hundreds of miles away

by J. E. Dixon, J. R. Cann and Colin Renfrew

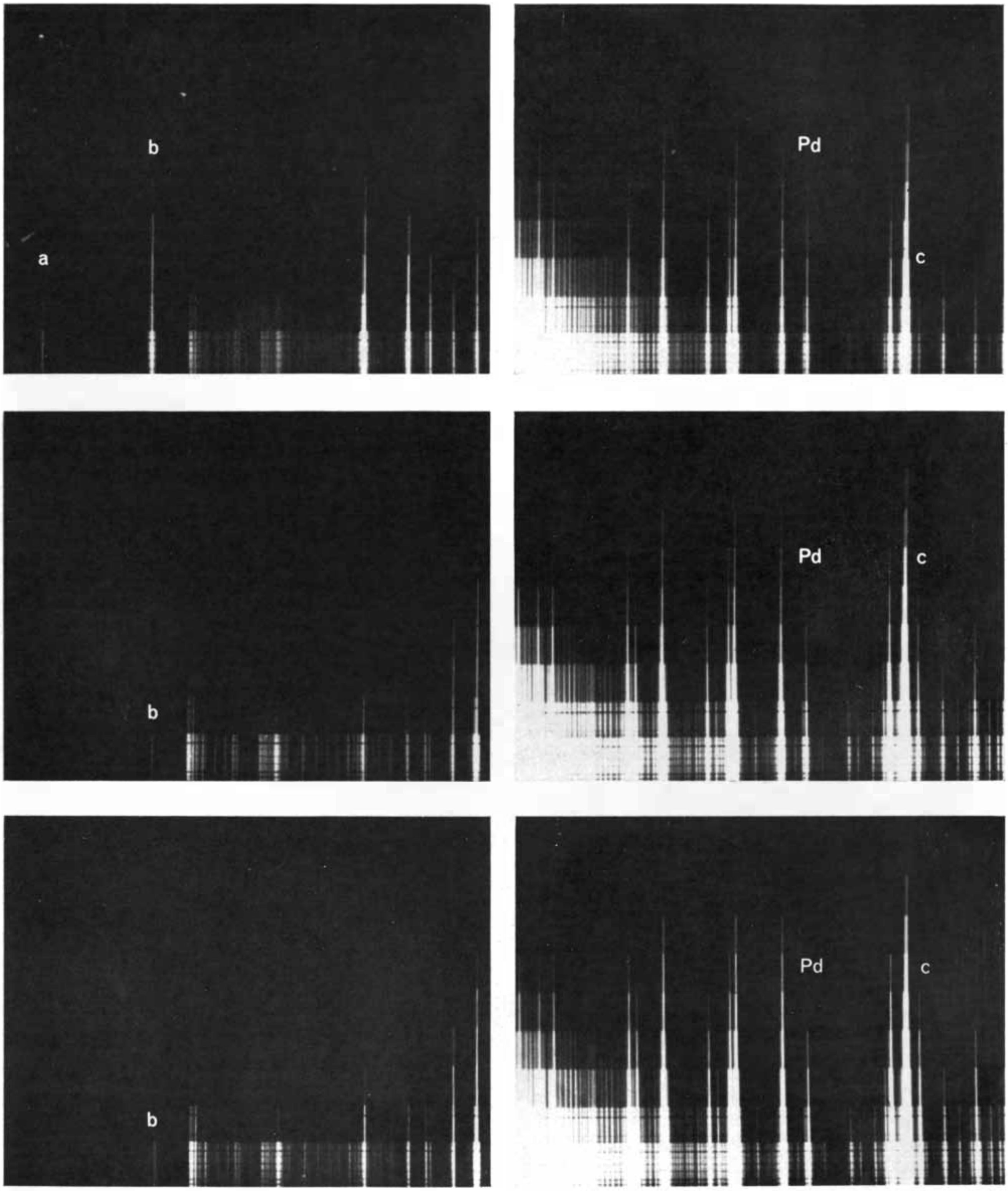
The transition from hunting to farming, which started mankind on the road to civilization, presents a number of interesting questions for investigators, not the least of which is the extent of communication among early human settlements. Archaeological explorations in recent years have unearthed the sites of prehistoric villages that were widely scattered in southwestern Asia and around the Mediterranean. The earliest communities—for example Jarmo in what is now Iraq and Jericho in Jordan—apparently were settled some 10,000 years ago [see "The Agricultural Revolution," by Robert J. Braidwood; SCIENTIFIC AMERICAN, September, 1960]. One might suppose the primeval villages, separated by hundreds of miles and often by mountains or water, were isolated developments, not even aware of one another's existence. There have been reasons to suspect, however, that this was not the case, and we have now found definitive evidence that the prehistoric communities throughout the Near East and the Mediterranean region were in active communication.

What kinds of evidence of communication between geographically separated peoples might one look for? Obviously in the case of prehistoric peoples the only materials available for study are the remains of the objects they made or used. In the search for signs of possible contact between two cultures archaeologists have generally depended on a comparative examination of the artifacts. If the two cultures show strong similarities in knowledge or technique—say in the method of working flint or the style of pottery—this is taken to signify mutual contact and perhaps actual trade in objects. At best, however, such evidence is only suggestive and is always subject to doubt; it leaves open the possibility that the similarities, no matter how close they are, may be mere coincidence, the two peoples' having hit independently on a natural and obvious way of doing things.

The raw materials of which the objects were made, on the other hand, may offer an opportunity for a more decisive inquiry. If a material used by a community does not occur locally in the raw state, one must conclude it was imported, and the possibility exists that it was obtained in trade with another population. One can then start on the task



OBSIDIAN BOWL made during the fourth millennium B.C. is seen from above, its spout jutting out at the right. The translucent bowl, which is nearly eight inches in diameter, comes from the Mesopotamian site of Tepe Gawra (*see illustration on page 46*). Trace-element analysis shows that the obsidian comes from Acigöl, 400 miles to the west in Turkey.



SPECTROGRAMS of three obsidian samples, reproduced here in part, show that differing proportions of trace elements make possible the identification of obsidian from different volcanic deposits. At top are two parts of the spectrogram of raw volcanic glass from Melos in the Aegean Sea. Matching parts of a spectrogram of a piece from an obsidian blade unearthed at Ali Kosh in Iran are at center. Matching parts of a spectrogram of volcanic glass from Nemrut Dağ, near Lake Van in Armenia, are at bottom. In the top specimen the spectral lines of strontium (*"a" at left, 4,607.4 angstrom units*) and barium (*"b" at left, 4,554 angstroms*) are relatively long, indicating proportions of some 200 parts per million and 700 parts per million respectively. The same lines are almost invisible in the center and bottom specimens. The spectral line of zirconium (*"c" at right, 3,438 angstroms*) is relatively short in the top specimen, indicating a concentration of about 50 parts per million. In the other two specimens the long zirconium lines indicate a concentration of about 700 parts per million. Palladium (*spectral lines labeled "Pd" to the left of "c"*) is not a normal trace element in obsidian; the element is added to provide a standard for calibration. The trace-element differences show that the Ali Kosh artifact (*center*) could not have been made of obsidian from Melos but instead is chemically similar to the obsidian found at Nemrut Dag.

OBSIDIAN FROM MELOS, in the form of a core from which blades have been flaked, shows the characteristic glistening surface of this volcanic glass. Obsidian was probably traded in the form of glass lumps or cores that the final purchaser turned into tools himself.

SCULPTURED SEASHELL, carved from a variety of obsidian with distinctive white spots, was found in a Minoan site on Crete. Sir Arthur Evans, the pioneer student of Cretan prehistory, thought the obsidian had come from Lipari, off Sicily, where similarly spotted volcanic glass is common. Analysis now proves that it came from nearby Giali in the Aegean.

of tracing the material to its source.

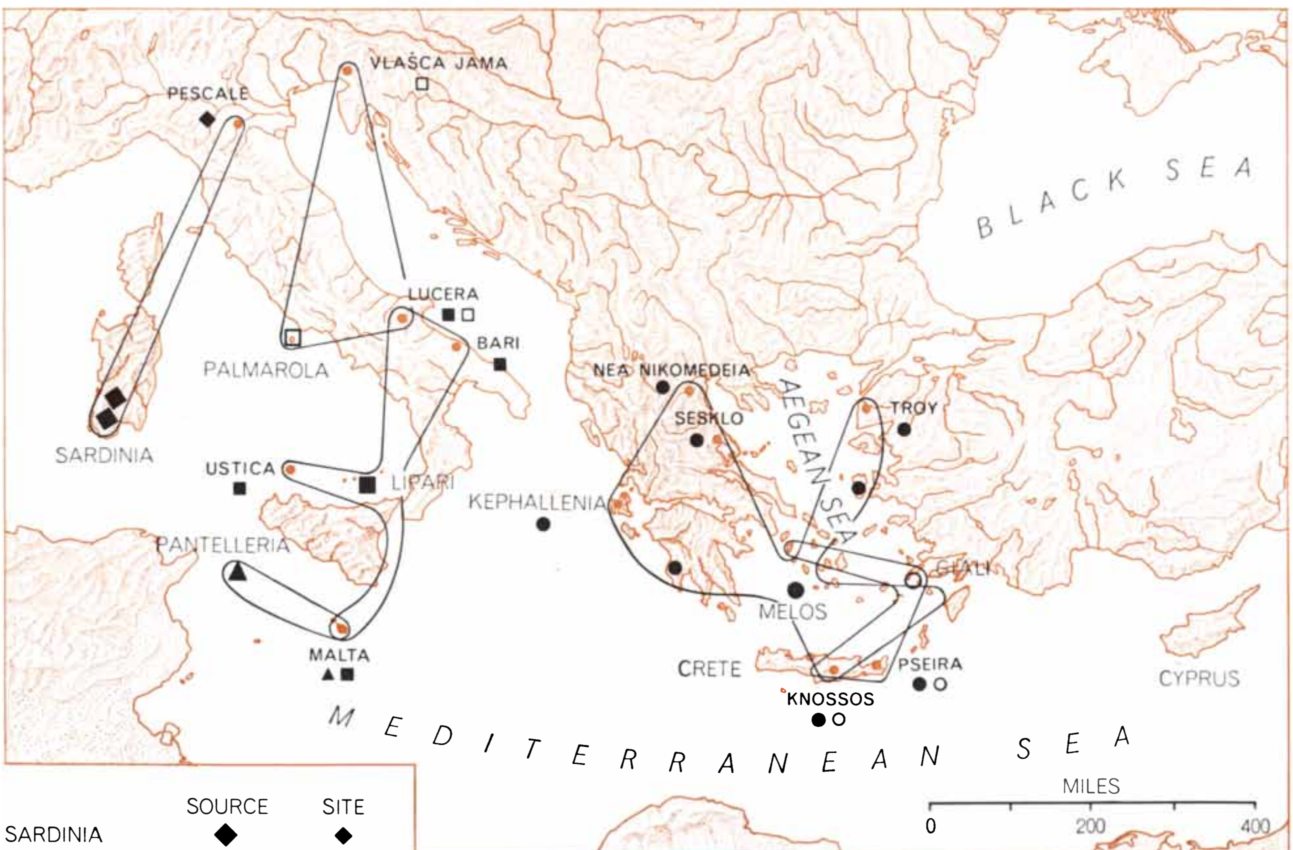
It occurred to us that obsidian might be an ideal material for a tracer investigation of this kind. Obsidian is a hard, brittle volcanic glass that can be chipped like flint and fashioned into a sharp tool. It is known to have been used for knives and scrapers by prehistoric men as early as 30,000 years ago. Obsidian tools have been found in nearly every early village site in the Near East and the Mediterranean region. Yet for most of these sites it was an indubitably foreign material; it can be obtained only in certain areas of recent volcanic activity, which in that part of the world means the region around Italy, some islands in the Aegean Sea and certain areas in modern Turkey and Iran. Some of the ancient villages where obsidian tools were used were many hundreds of miles from the nearest natural source of the material.

How could one identify the particular source from which the obsidian was obtained in each case? Clearly our first task was to determine whether or not obsidian samples showed distinguishable differences that could be connected with their source. We considered several possible criteria. Physical appearance obviously would not be a reliable guide, because obsidian samples from a single volcanic deposit may vary greatly in visible characteristics such as color. Microscopic examination was not helpful: the obsidian tools were generally made of material that is uniform in structure, without crystalline inclusions. Chemical analysis of the main components was also of no avail, because all samples of obsidian are substantially alike from this point of view. We decided finally on a chemical test based on the presence of trace elements. Perhaps the obsidian samples would show distinct differences in their trace-element content that could be identified with the deposits from which they came.

In order to explore this possibility we began by analyzing samples of obsidian collected from various well-known volcanic sources in the Mediterranean area: Lipari, a volcanic island north of Sicily, two areas in Sardinia, the islands of Pantelleria and Palmarola and the island of Melos in the Aegean Sea.

For chemical analysis of the samples we used a convenient spectrographic method that archaeologists have long employed on metal objects. Every element emits characteristic wavelengths of light when it is heated to incandescence; a familiar example is the yellow light of burning sodium. By passing the light

SOURCES OF OBSIDIAN in the Mediterranean and the Near East during Neolithic times included volcanic areas in Sardinia and on Palmarola, Lipari and Pantelleria in the central Mediterranean, the islands of Melos and Giali in the Aegean Sea, two central Anatolian sites, Acigöl and Çiftlik, and several places in ancient Armenia, including Bingöl in eastern Turkey and Nemrut Dağ near Lake Van.

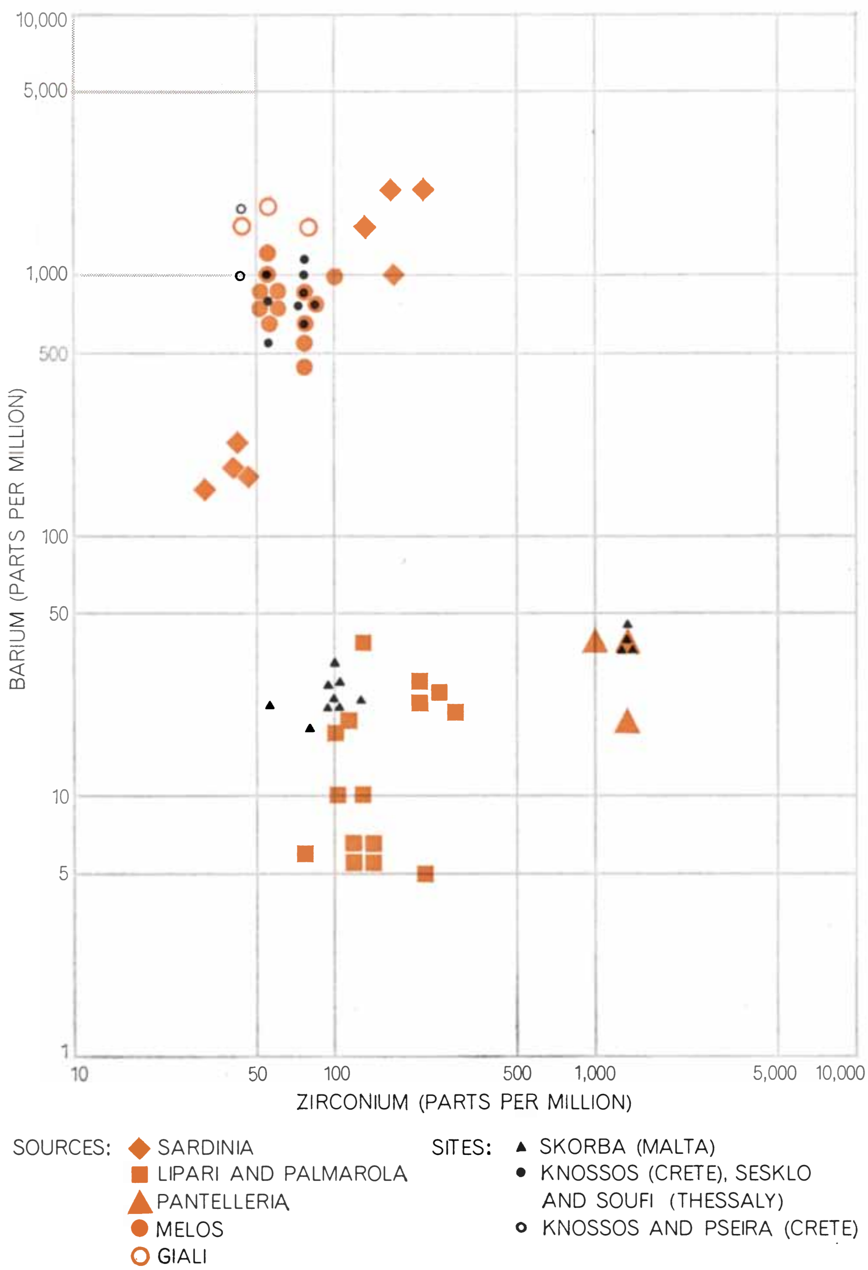


	SOURCE	SITE
SARDINIA	◆	◆
LIPARI	■	■
PALMAROLA	□	□
PANTELLERIA	▲	▲
MELOS	●	●
GIALI	○	○

OBSIDIAN TRADE PATTERNS in the central Mediterranean and the Aegean regions show that, although the volcanic glass was often shipped long distances from its sources, trade apparently did not take place between the two regions. Within each region, however, obsidian from two sources is often found at one site. Not all the sites indicated are named.

from a mixture of elements through a prism or diffraction grating that spreads out the wavelengths in a spectrum, one can separate the emissions of the various elements and detect trace elements that are present even in the amount of only a few parts per million. The beauty of the method for studying archaeological specimens is that accurate measurements can be obtained from very small amounts of material. Sixty milligrams taken from a sample is sufficient. Ground to a fine powder, mixed with an equal amount of carbon and ignited in a carbon arc, this material yields a spectrographic picture that gives a measure of the quantity of each trace element in the sample; the quantity is indicated by the intensity (photographically the height) of the element's spectral lines.

We obtained readings for 16 elements in our samples of obsidian. Among the trace elements found to be present, the two that showed the greatest quantitative variation over the range of samples were barium and zirconium. We therefore tried using the relative concentrations of these two elements as a test for identification of the source [*see illustration on this page*]. To our immense satisfaction these quantities were found to indicate the geographical source quite well. Samples from various flows and outcrops at Lipari, for instance, all showed much the same proportion of barium to zirconium; those from Melos had characteristic contents of these elements different from those at Lipari. The Pantelleria and Sardinia samples likewise could be distinguished on the same basis. The Palmarola samples turned out to be similar to the Lipari ones on the barium-zirconium graph, but we found we could distinguish them from the Lipari type by their content of another trace element, cesium.

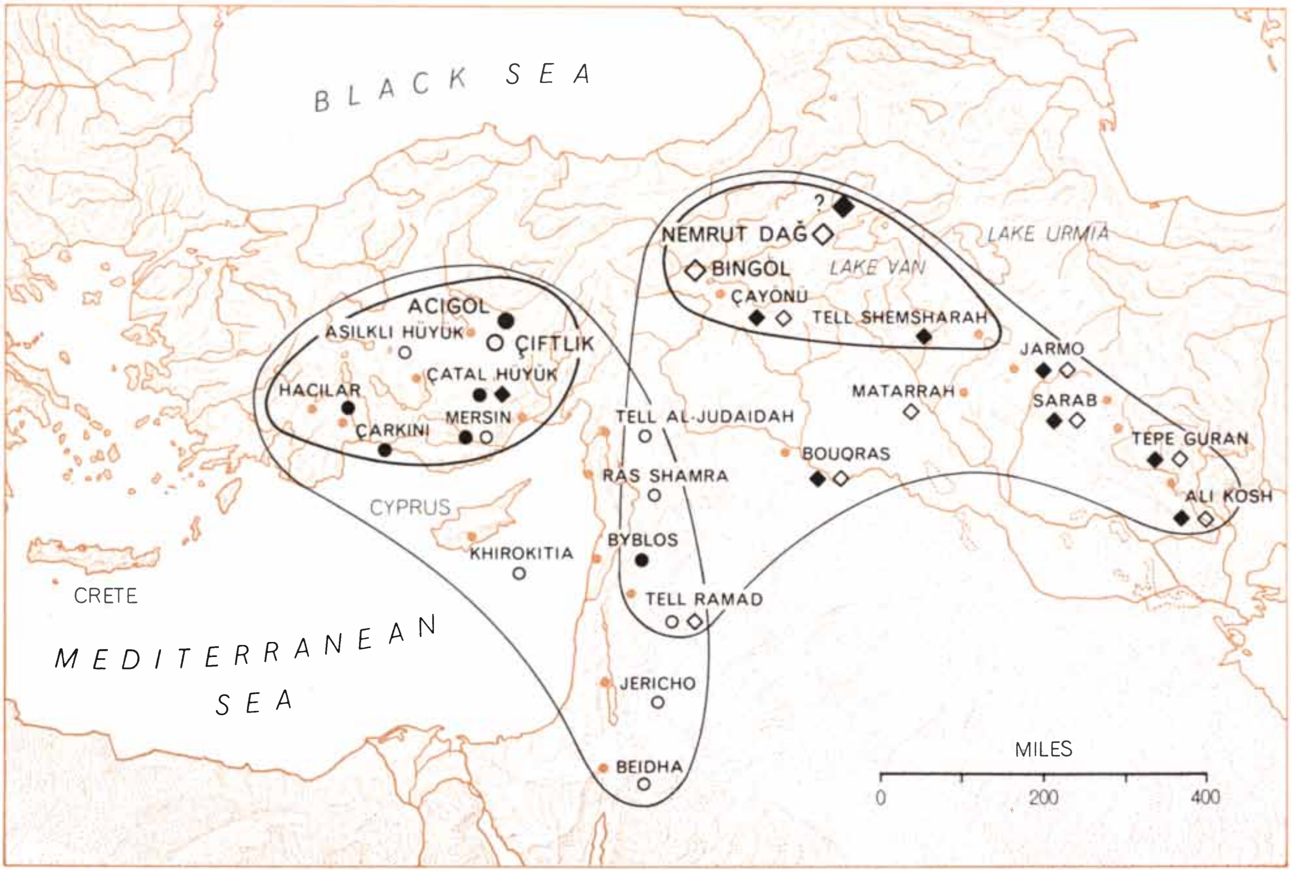


TWO TRACE ELEMENTS, barium and zirconium, provide the principal means of identifying the sources of obsidian artifacts from central Mediterranean and Aegean sites. Obsidian from both Sardinian sources is richer in barium than obsidian from Pantelleria, Lipari and Palmarola, whereas Pantellerian obsidian is the richest of all in zirconium. Other trace elements, not plotted here, allow additional distinctions to be drawn. Although their barium and zirconium contents are similar, the high calcium content of Gialian obsidian distinguishes it from Melian, whereas Palmarolan obsidian is much richer in cesium than Liparian.

Having established these markers for identifying the obsidian sources, we were in a position to determine the raw-material origins of obsidian tools found at the sites of ancient settlements. The little island of Malta, south of Sicily, offered objects for a clear-cut test. The remains of a remarkable prehistoric society of 5,000 years ago, marked by colossal stone temples, have been unearthed on this island. The finds include small obsidian tools. There are, however, no natural obsidian deposits on the island. Where, then, did the material come from? Some archaeologists had suggested that it might have been brought there by Minoan traders from the island of Melos, 600 miles to the east. Our trace-element analysis of the Malta tools disclosed that this conjecture was incorrect: the obsidian was of the types found on Lipari and on Pantelleria, a tiny island 150 miles northwest of Malta.

The findings revealed two important facts about the ancient Maltese settlement. They showed that the island people in that early Neolithic period were already accomplished seafarers, traveling frequently to Sicily, Lipari and Pantelleria. On the other hand, the obsidian evidence also indicated that the Maltese people had little or no contact with the contemporary Minoan settlements of the Aegean area. Their stone temples may well have been their own invention.

Our tests resolved another question involving the Minoan culture. Sir Arthur Evans, the illustrious archaeologist who excavated the palace at Knossos on Crete more than half a century ago, found a number of finely carved objects there that were made of a variety of obsidian marked by prominent white spots. He concluded that this material came from



NEOLITHIC NEAR EAST was another scene of active obsidian trade. Cypriot, Anatolian and Levant villages obtained obsidian mainly from two sources in Anatolia: Acigöl and Çiftlik. Mesopotamian villages, in turn, depended on sources in Armenia. The locations of two, Nemrut Dağ and Bingöl, are known. A third variety of obsidian, found at many Mesopotamian sites, is also probably Armenian but its source is not yet known. A heavy line surrounds a nuclear zone within each trade area. These are designated "supply zones" by the authors: more than 80 percent of the chipped-stone tools at supply-zone sites are obsidian.

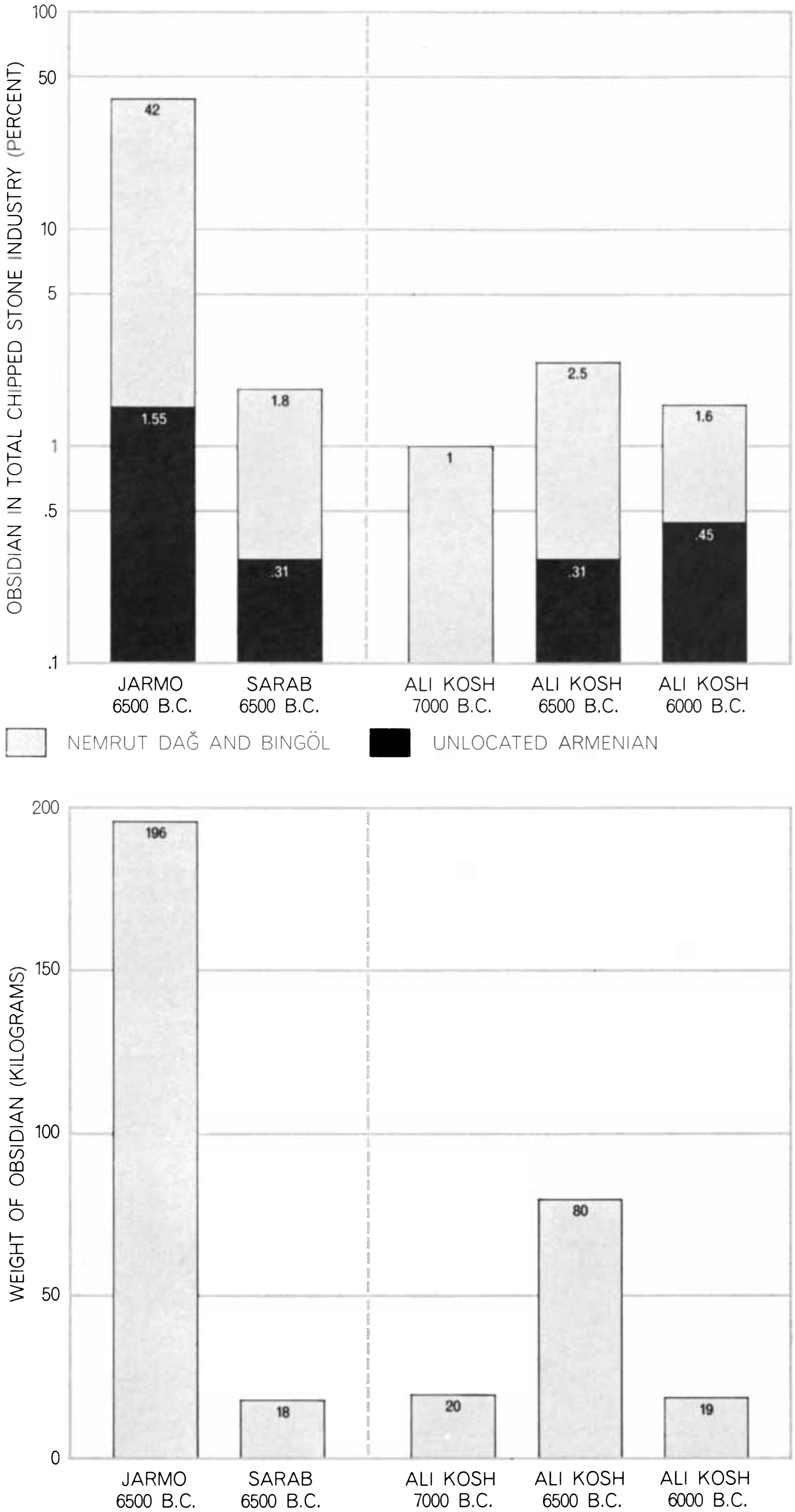
Lipari, some of whose obsidian also has white spots. On trace-element examination, however, it now turns out that the white-spotted obsidian at Knossos came not from Lipari but from the small island of Giali some distance north of Rhodes. The Knossos remains also include tools made of unspotted obsidian. Analysis shows that this material came from Melos, which is to be expected, since Melos is the nearest obsidian source to Crete. In general, the obsidian evidence establishes the early Aegean islanders as skilled sailors and traders, disseminating the material not only among the islands but also to settlements in Greece and Turkey.

The obsidian tracer work, begun only six years ago, has given rise to investigations at various institutions in Britain, the U.S. and elsewhere, and studies are going forward on material from early settlements in Europe, in the Middle East, in Mexico, California and the Great Lakes region of the New World, in New Zealand and in Africa, where early man made hand axes of obsidian as long as 100,000 years ago. Our own group at the University of Cambridge, having verified the validity of the method by the Mediterranean tests, has proceeded to apply it to an investigation of the origins of trade among the earliest settlements of man in the Near East, that is to say, in Mesopotamia and in Turkey, Palestine and Egypt.

The first problem in the study of this region was to locate the natural sources of obsidian. A number of sources (all volcanic, of course) have now been identified in mountainous areas of Turkey and northern Iran [*see top illustration on page 41*] and in Ethiopia to the south. All the available geological evidence indicates that the entire region between these places, including Egypt, is devoid of natural deposits of obsidian. This means that every prehistoric village in the "fertile crescent," where farming began, had to import its obsidian.

Samples from all the natural deposits and from the obsidian artifacts found in the village sites were analyzed. They were found to be definable in eight different groups, or types, according to their barium-zirconium content, and in some groups the presence of another trace element (like that of cesium in the Palmarola obsidian) served additionally to distinguish the particular source of the material. Some of the obsidian artifacts could not be matched to any known natural source. This necessarily called for searches for the missing sources. The composition of the natural obsidian samples suggested a certain geographical pattern of distribution, and this clue led to the discovery of at least one missing source. A few sources have not yet been located precisely, but the basic pattern of sources and destinations is now clear enough to provide a good picture of the movement and trade routes of obsidian in the period when the first steps toward civilization (variously called the "agricultural revolution" or the "neolithic revolution") were taking place.

By about 9000 B.C. groups of people



INFLUENCE OF GEOGRAPHY on trade in tool materials is apparent from the total obsidian and its proportion to other stone tools at three Zagros villages that imported Armenian obsidian during the seventh to sixth millenniums B.C. Jarmo, although outside the Armenian supply zone (*see illustration on preceding page*), was nonetheless well supplied with obsidian in terms both of percentage (*top graph*) and of estimated total weight of the material (*bottom graph*). The reduction in both percentage and weight at two more distant towns, Sarab and Ali Kosh (*plotted for three periods*), is drastic. Both distance and the difficult terrain apparently contributed to the scantiness of trade in the case of Sarab, a hill town.

in the Near East had begun to practice an incipient agriculture: selectively hunting and perhaps herding sheep and goats and harvesting wild prototypes of wheat and barley. They were not yet using obsidian to any appreciable extent. By the time the first farming villages were founded, probably a little after 8000 B.C., obsidian had come into rather general use. Naturally the extent of adoption of the material varied with distance from the sources of supply, and this is clearly traceable in the obsidian objects found at the sites of the ancient villages.

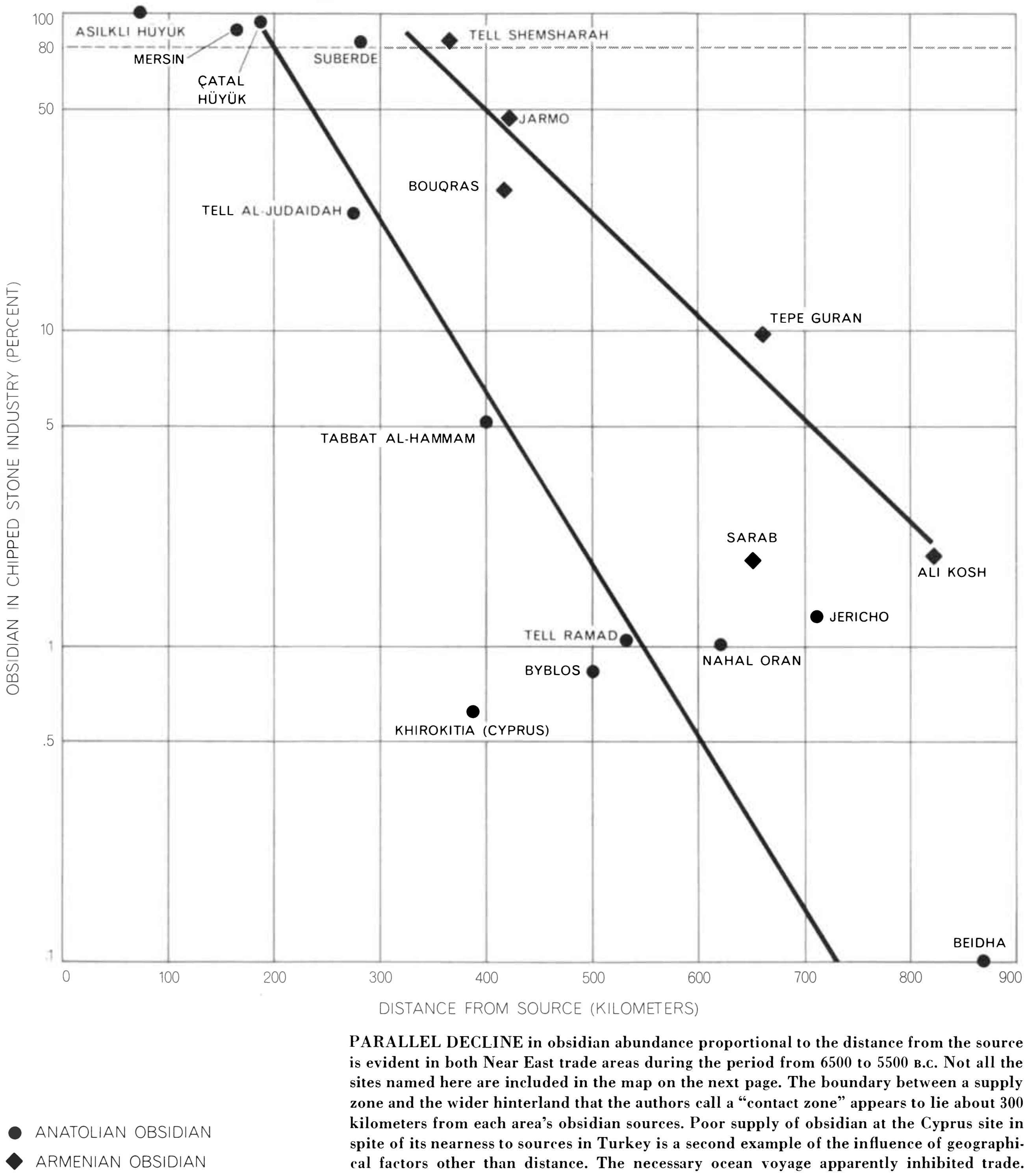
Within 150 to 200 miles of the obsidian deposits in Turkey and Armenia most of the chipped-stone tools found in the early prehistoric village sites are of obsidian: 80 percent, as against only 20 percent of flint. From that zone the proportion of obsidian falls off nearly exponentially with distance. This is clearly illustrated by the distribution of obsidian from a natural source in the volcanic area around Çiftlik, near the present town of Niğde in Turkey. At Mersin, the site of an early village on the Mediterranean coast not far from Çiftlik, obsidian was the most common chipped-tool material. From there its use diminished rapidly down the Levant coast, until at Jericho, 500 miles from the source, it is found only in very small quantities, most of the chipped tools there being made of flint. A few pieces of Çiftlik obsidian have been found, however, even at a Neolithic settlement on Cyprus, which indicates at least a trading contact across the water. Some of the obsidian from Turkish sources was distributed over distances of more than 600 miles in the early Neolithic period.

The early villages near the sources in Turkey developed a rich art and craftsmanship in obsidian. Particularly impressive are the objects found at Çatal Hüyük, a 6000 B.C. village that was so large it can be called a town. Among its obsidian products were beautifully made daggers and arrowheads and carefully polished mirrors, as sophisticated as any of those made 7,000 years later in Aztec Mexico.

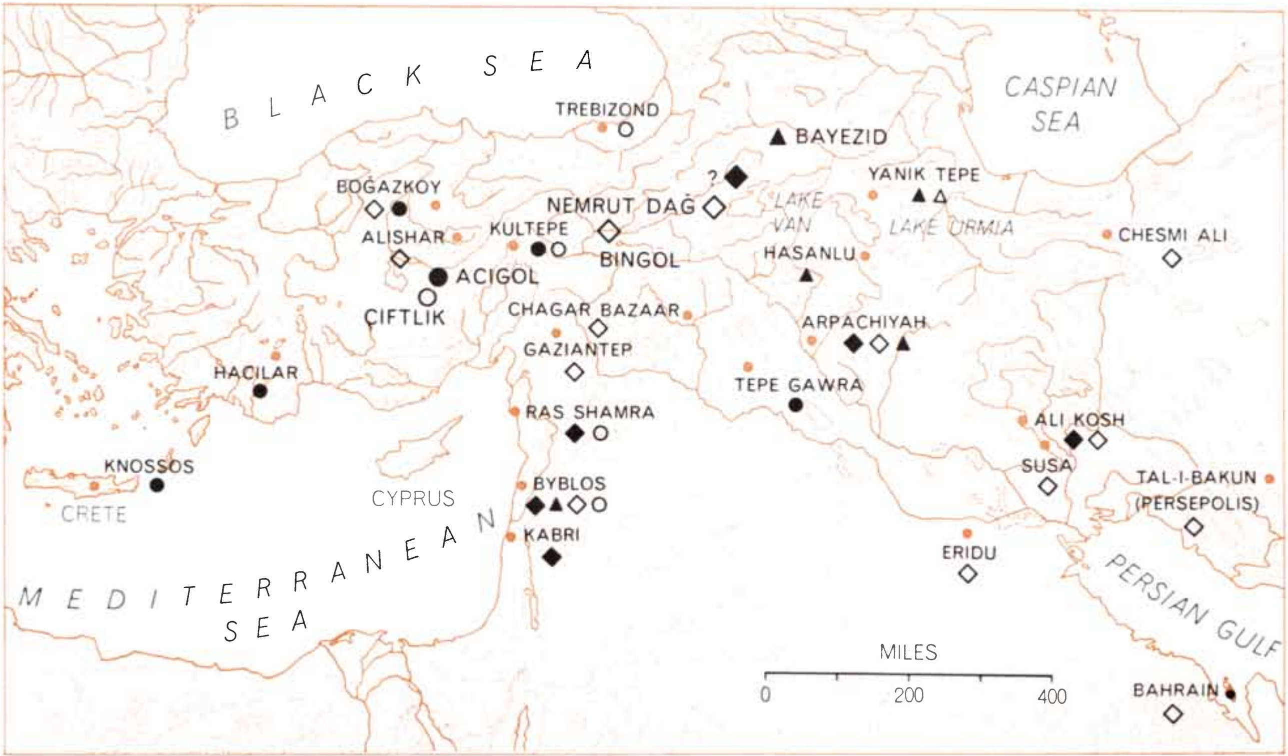
Jarmo, one of the earliest-known villages in the fertile crescent, was favored by proximity to several obsidian sources. At least two of these sources furnished considerable amounts of the material, estimated to total 450 pounds or more, to Jarmo in very early times. From this it appears that well before 6000 B.C. Jarmo must have been conducting a thriving trade across the mountains that brought it into contact with the communities to the north in Armenia. There were then no wheeled vehicles (they were not invented until 3,000 years later) and not even pack animals. Hence all the traded goods, including obsidian, must have been transported on foot, or perhaps part of the way in boats down the Tigris River.

Tracing the varieties of obsidian from their sources to the villages where they turn up in manufactured objects, we can reconstruct the trade routes of that early time in man's economic and social history. The routes, crossing mountains, deserts and water, connect the early settlements with a network of communications that must have influenced their development profoundly. No doubt goods besides obsidian were traded over these routes; indeed, it seems likely that there was a trade in perishable commodities that was much greater in size and economic significance than that in obsidian. Clearly, however, the most important traffic must have been in ideas. The network of contacts arising from the trade in goods must have been a major factor in the rapid development of the economic and cultural revolution that within a few thousand years transformed mankind from a hunting animal to a builder of civilization.

Obsidian now furnishes us with a tool for retracing the communications at the beginning of the revolution, more than 3,000 years before the invention of writ-



PARALLEL DECLINE in obsidian abundance proportional to the distance from the source is evident in both Near East trade areas during the period from 6500 to 5500 B.C. Not all the sites named here are included in the map on the next page. The boundary between a supply zone and the wider hinterland that the authors call a "contact zone" appears to lie about 300 kilometers from each area's obsidian sources. Poor supply of obsidian at the Cyprus site in spite of its nearness to sources in Turkey is a second example of the influence of geographical factors other than distance. The necessary ocean voyage apparently inhibited trade.



		SOURCE	SITE
ARMENIAN OBSIDIAN	1G	◆	◆
	3A	▲	▲
	4C	◇	◇
ANATOLIAN OBSIDIAN	1E-F	●	●
	2B	○	○

POST-NEOLITHIC TRADE, its directions often traceable by means of luxury items made from obsidian, was cosmopolitan in its extent. Two new sources of supply in the Lake Urmia area of Armenia were developed and Armenian obsidian was traded as far west as the Levant and as far south as Bahrain on the Persian Gulf. Obsidian from Turkey was carried westward to Crete and was transported for the first time across the desert to Mesopotamia. Ethiopian obsidian holds the Near East travel record; a slab of this material, bearing a 16th-century B.C. Egyptian inscription, has been discovered at Boğazköy, a Hittite site in Turkey.

ing. In addition to revealing the pattern and range of contacts among the prehistoric settlements, it gives us a rough picture of trade statistics (through the amounts of material involved) that indicates the strength of the communication links between particular communities. Furthermore, the development of communications over the millenniums after the villages were first established can be traced in the record of the obsidian trade.

As time went on and transportation, aided by the domestication of the ass, improved, the trade in obsidian expanded, both in the number of sources mined and in the distances of distribution. Trade routes developed across the Syrian desert in both directions, and more obsidian began to appear on the Levant coast. Obsidian from Armenia was exported to villages as far distant as Bahrain on the Persian Gulf and the Teheran area near the Caspian Sea. The use of obsidian for tools declined with the coming of the metal ages after 4000 B.C., but it continued to be prized for ornamental objects such as bowls, statuettes and even small articles of household furniture, such as tables. By that time the Egyptians, apparently obtaining the material from Ethiopia, also had begun to use obsidian for this purpose. A remnant of a little toilette table of obsidian made in Egypt and bearing a hieroglyphic inscription of Pharaoh Chian, of the 16th century B.C., has been found at Bogazköy, the capital of the ancient Hittite kingdom in Turkey. It may have been a gift sent by the Pharaoh to the Hittite king.

By then obsidian itself was no longer an important material of commerce. Nonetheless, the great trade routes that had developed between the cities of the Near East may well have followed the same paths that had first been blazed by the obsidian trade thousands of years earlier.

The analysis of the early obsidian objects now throws new light on the revolution, some 10,000 years ago, that led to man's emergence from the hunter's way of life. There has been a tendency to think of this beginning as an isolated, small-scale phenomenon—of a little tribal group of people settling down somewhere and developing an agricultural system all by itself. In recent years an intensive search has been pursued for the "birthplace" of this event: Did the first village spring up in the Levant or in the Zagros Mountains on the rim of the fertile crescent or in Turkey? That question now becomes less interesting or significant than it was thought to be. The farming way of life, it appears, originated not at some single location but over whole regions where the peoples of various settlements exchanged ideas and the material means of sustenance.

Throughout the 2,000 years or more during which agriculture was first developing in the Near East the communities dispersed through the region were in more or less continual communication with one another, primarily trading goods but also inevitably sharing their discoveries of agricultural techniques and skills. There is every reason to believe the region functioned essentially as a unit in moving along the road of technological advance. The early villages show considerable diversity in the customs and beliefs that make up what is called a society's "culture," but there can be little doubt that their mutual contact greatly influenced not only their material progress but also their social development and world view.

In hope of doing each other some good

The way into print

Publish or perish. The doctrine still holds sway despite bitter words spoken against it over ethanol after midnight. For validity one's work must be consecrated by a typesetter. Only after ink has been applied to paper by a decent number of revolutions of a metal cylinder can a piece of scholarship be considered accomplished. This is the custom. It dates from a day when learned men with news to spread were few, which is no longer so.

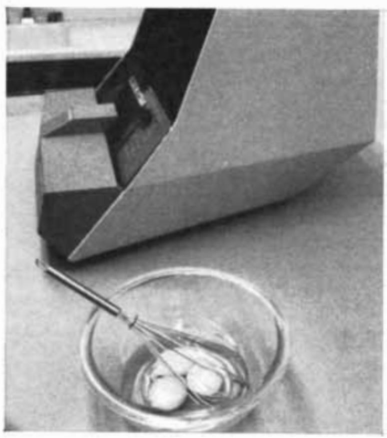
Now that myriads of learned men and women have to get into print and get into it frequently, the technology by which this is done had better be improved. On top of Gutenberg's 15th century concept of movable type and Mergenthaler's 19th century concept of casting liquid metal into lines of type, photography now further facilitates the printing of the printed word. This involves us.

Companies that built typesetting machines around a pot of simmering Pb-Sb-Sn alloy now build them around computers and certain light-sensitive litho plates of our manufacture. The neatness with which the characters are arrayed on the page—which has come over the centuries to signify commitment and authority—is attended to by the computer. With less asked of human judgment and fingers, the flux of words and symbols onto our plates can be maintained at superhuman levels. As the plates emerge from the machine, they are dunked in solutions that change them to actual printing plates, ready to run on offset presses in certain printing applications today and probably more tomorrow.

Dunking sounds like a bottleneck, and it is. That bottleneck, too, is being broken. To the printing establishments prepared to cope best with the publishing needs of the day we are supplying another machine (called KODAK Litho Plate Processor, Model 1-N). It sits beside the typesetting machine, smoothing the way into print by bridging the dunking step. (What might happen to the words and symbols thus at length properly arrayed in a good black on a respectable paper is hinted at below.)

If we can be of any assistance, please inquire of Eastman Kodak Company, Graphic Arts Sales Development, Rochester, N.Y. 14650.

Vision of a well-equipped home



The low price of the RECORDAK EASAMATIC Reader PFCS is significant. It gives quick access to data such as parts lists, catalog pages, and buyers' guides, many pages of which can be compacted on a 4 x 6″ or 3¼ x 7⅜″ microfiche. In truth, very few of this particular model are likely to be purchased for use in the home. Nevertheless, as microfilm readers become mass-produced items and the unit price continues to drop with rising volume, the more pervasive will microfilm become. When microfiche readers are found in as many homes as dishwashers are today, there may be quite an effect on the manner of publishing anything of a nature at all specialized.

Questions about the RECORDAK EASAMATIC *Reader or about any other aspect of microfilming are answered by Business Systems Division, Eastman Kodak Company, Rochester, N.Y. 14650.*

Hungry?

From three eggs in a mixing bowl can grow a masterpiece of self-fulfillment to fill the self, family, and guests full of purring satisfaction. All it takes are talent, time, inclination, and a comfortable income.

There is a magazine called *Food Technology* whose very name suggests a colder view of food preparation, where eggs get beaten in 300-pound batches because there are several billion people waiting to be fed and relatively few of them keeping their own laying hens. One of our chemists, Gerald T. Luce, has an article in *Food Technology*. He writes not of eggs but of acetylated monoglycerides.

These were first proposed by the U.S. Department of Agriculture and were put into commercial production by us. Metabolizable and nutritious themselves, they form coatings of low permeability to oxygen and water vapor. Consequently foods coated with them stay appetizing and wholesome longer. This works particularly well with fish steaks cut from the frozen round fish—a promising development with the switch from hunting the sea to farming the sea.

Luce's paper goes into detail on nuts, raisins, meats, cooked lobster, breast of chicken, and many other good things to eat. For a reprint write Distillation Products Industries, Rochester, N.Y. 14603 (Division of Eastman Kodak Company).

Data on the drapes

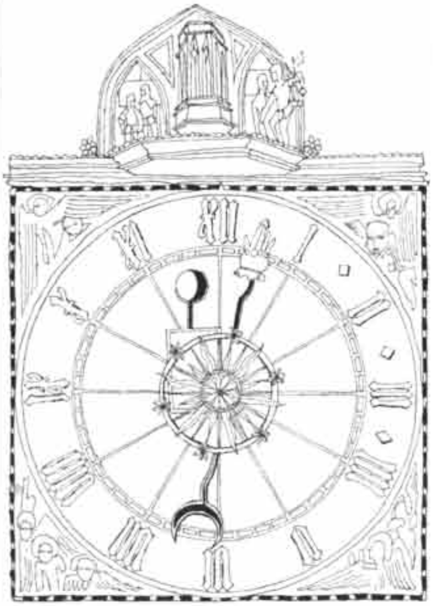
Draperies greatly affect heat- and light-exchange. To compute heating, cooling, and lighting requirements, planners need "fenestration data" on the drapery fabrics, such as solar transmittance, solar reflectance, openness factor, and shading coefficient, as understood in the Guide and Data Book of the American Society of Heating, Refrigerating, and Air Conditioning Engineers. To obtain fenestration data on a fabric made of VEREL *Modacrylic Fiber, send request with suitable sample of the fabric to Merchandising Services Department, Eastman Chemical Products, Inc., 260 Madison Avenue, New York City 10016 (Subsidiary of Eastman Kodak Company).*

George Eastman ran an ad for the glass photographic plates he was making at home after work at the bank. The job at the bank was all right, but he enjoyed photography more. After a while he quit the job. The bank survived. It thrives today at the same location, in a building that is much brighter and more comfortable to work and bank in—summer and winter—than when Eastman worked there. And here, less than 90 years after the first ad, is another where the company that the first ad started is offering not only glass photographic plates* but also help with heating and air-conditioning. A logical if lengthy thread connects the plates to the air-conditioning.

Flexible film pleased the photographers more than plates, so Eastman made film and the film made more photographers. Cellulose nitrate film base was hazardous, so the company switched over eventually to cellulose acetate. To make cellulose acetate efficiently we needed to make a lot of it, so what we didn't need for film, we put into fibers for textiles. The textile experts and customers thus acquired broadened us from cellulosics to fully synthetic fibers. Of these VEREL Modacrylic Fiber has won tremendous favor for draperies. Draperies greatly affect heat- and light-exchange.

*Which may be selected with the aid of Pamphlet P-140, obtainable from Industrial Photo Methods, Eastman Kodak Company, Rochester, N.Y. 14650.

SCIENCE AN



What Is U.S. Science Policy?

The science establishment of the U.S. has grown without plan or policy over the past three decades, largely as the result of "initiatives precipitated by urgency" from one phase to the next in the cycle of war and cold war. Suddenly, within the present fiscal year, Federal outlays for research and development have leveled off and have even declined as a percentage of the national budget. The downturn is attributed to stringencies imposed by the war in Vietnam and the completion of major programs of "the defense–space–atomic-energy group." The need to develop a more positive science policy than "mere response to external pressures" is becoming clear. At this historic turning point "the fronts seem to have changed completely from those days, a quarter of a century ago, when farsighted scientists convinced political and military leaders that they had a vital contribution to offer to the war effort. Today, the call for better and fuller use of the potential of science in the service of man comes mainly from the politicians and the administrators."

This appraisal of American science and its leadership constitutes a major theme of a report issued by the Organization for Economic Cooperation and Development, the consortium of European nations organized in 1947 as the party of the second part for Marshall Plan transactions. The 600-page report presents a comprehensive staff study of the "Government-industry-university tripod" that sustains American science today. Personal appraisals are supplied by four "examiners": H. G. B. Casimir, director of the Philips Research Laboratories in the Netherlands; Théo Lefèvre, former prime minister of Belgium; Pierre Massé, board chairman of Electricité de France, and C. H. Waddington, professor of animal genetics at the University of Edinburgh. Although the authors were charged to explore the "technology gap" between the Old World and the New on behalf of their European colleagues, they have much to say to the scientific community of America.

The "pluralist" arrangements that bring $15 billion from the Federal budget into research and development—including the primary $1.8 billion for science in the universities—have avoided "the danger of bureaucratic control over the world of science." These funds flow on short-term contracts and grants from numerous agencies and principally for highly specialized "mission-oriented" ends. The universities respond with "great flexibility and freedom of action . . . and rapidly inaugurate new programs and departments." It is a system that encourages "excellence," backs "the young man who is doing first-class work" and suppresses "gerontocracy."

"The amplitude of Federal aid" has brought significant changes, on the other hand, in the constitution of American universities, particularly the 21 largest public and private universities that receive nearly 60 percent of the funds. "The university is no longer . . . a single vast intellectual enterprise. . . . Researchers turn toward their colleagues in other institutions interested in the same problems rather than to their immediate university environment." In a "biological analogy," Waddington likens the "old ideal of a university to . . . a whole nervous system . . . concentrated in a single central brain" and contrasts this to "the multiversity ideal . . . a highly developed peripheral nervous system . . . deficient in central organization."

The European observers found that competition for project funds makes it difficult for American scientists to commit time and energy to "contemplative, imaginative thought." For all the Nobel prizes, their work is not the kind that "changes the paradigms of science." It is the political rather than scientific leadership that has pushed new measures, such as institutional grants, to reduce the centrifugal pressure of project-funding and to redress its "concentration at the summit." The entrepreneurs of science have avoided equally the difficult question of priorities. Against the military and geopolitical challenges that so largely determined the lines of recent scientific growth, political pressures have been advancing the claims of a wider range of public interest. The scientific community has offered little in the way of "positive and constructive initiatives."

Doctors' Dilemmas

In George Bernard Shaw's *The Doctor's Dilemma* (1906) the decision that confronts Sir Colenso Ridgeon is which of two patients should receive a lifesaving tuberculosis vaccine. As medical techniques have grown more sophisticated the dilemmas presented to physicians have multiplied, since each extension of capabilities presents a new set of choices. In the past few years such questions have been sharpened by the increasing feasibility of operations in which an organ is transplanted from one person to another, and in recent months news of the first heart transplants has brought the whole issue dramatically to public attention. One question that has been raised is: Might there be a tendency toward haste in deciding that a potential donor is indeed dead? "The doctor who decides that a patient is sufficiently dead for him to remove the heart while it is still functional takes a difficult decision, but this problem is not new to medicine," says an editorial in the British journal *Nature*. "There is no sharp dividing line between life and death"; the best safeguard is to make sure that different doctors care for the potential donor and the prospective recipient of an organ transplant.

Broader problems arise from continued advances in life-prolonging technology, none of them easily answered. Some questions are fairly direct and technical: Can this life be saved? By what procedures can it best be saved? Are those procedures feasible here and now? At what risk? Other questions involve social judgments and ethical standards: Should life be prolonged in the presence of terminal disease or of gross disability? Whose life should be prolonged if a choice must be made? At what cost—human and financial—should a given life be prolonged? The issue of monetary cost has always been implicit in many medical decisions; it becomes explicit as a society assumes more responsibility for health, making the cost of medical care a social and political question rather than a purely individual or family problem.

A number of these ethical and social considerations were explored last year in a symposium in England, sponsored by the Royal Society of Medicine, titled "The Cost of Life." In one paper David N. S. Kerr of the Royal Victoria Infirmary at Newcastle upon Tyne assessed the cost of regular hemodialysis—the maintenance, by intermittent treatment with an artificial kidney, of a patient who lacks kidney function and would otherwise die of uremia. The treatment works: Kerr pointed out that three of the four original patients started on regular dialysis in 1960 by Belding H. Scribner of the University of Washington School of Medicine were still alive. It is expensive, however; Kerr estimated that if all potential British candidates for dialysis up to age 54 were maintained on artificial kidneys, the direct costs would be some £33 million (about $80 million) a year and the indirect costs would be far greater. He concluded: "I see no alternative to the selection of a small proportion of the potential candidates for many years to come." Kerr's group now makes that selection in part on the basis of what he calls a "trial by ordeal": only those patients are accepted who have shown that they can stick to the stringent diet that is required because dialysis is not as efficient as the kidneys in removing waste products from the blood.

M. F. A. Woodruff of the University of Edinburgh Medical School considered ethical problems involved in kidney transplants. He said it was "right and proper" to allow a person with two good kidneys to donate one to a relative "or possibly to a good friend" if transplantation from a cadaver is not feasible, if the degree of tissue compatibility is acceptable and if the probable result of the operation and the risk to the donor have been fully explained. In cadaver transplants, Woodruff pointed out, the kidney must be removed quickly. In his hospital the kidney-transplant team is informed of potential donors who are under the care of other doctors and "are thought likely to die"; blood samples are taken from such patients for typing, "but nothing more is done until the prospective donor is pronounced dead by the doctor responsible for him... in particular, in the case of patients on respirators, the transplantation team takes every possible care not to influence the decision as to whether the machine should be turned off and when this should be done."

A different kind of ethical problem was raised by Walpole Lewin of Addenbrooke's Hospital in Cambridge. He discussed severe head injuries, an increasing problem as a result of highway accidents. When a person is unconscious for a long time and seems to have suffered irreversible brain damage, how hard should doctors try to keep him alive? Lewin's conclusion was that aside from a few patients whose death is inevitable, it is "impossible to forecast with certainty whether an injury is recoverable or not. There has to be, therefore, one standard of treatment for all patients."

Other participants dealt with a variety of questions: the care of defective children, the very old and the completely paralyzed; euthanasia and abortion; the effect of research in providing new methods of treatment. In summing up, Lord Cohen of Birkenhead, past president of the Royal Society of Medicine, said that doctors face essentially two kinds of dilemma: first, the determination of priorities when limited resources make it impossible to provide all necessary medical care, and second, the decision as to when to discontinue artificial aids to survival. In this connection he quoted the 19th-century poet Arthur Hugh Clough: "Thou shalt not kill; but need'st not strive/Officiously to keep alive." He suggested that "the physician (and indeed the legislator and the lay citizen) should in reaching decisions pose three questions: (1) Has this person [or his guardian] been given *all* the necessary knowledge, and freedom to decide whether he accepts the course of action recommended...? Is my recommendation based on the best scientific probabilities? Is my proposed course of action one that I would advise for someone I love, and if I were similarly circumstanced would I wish this done for me?"

Speed Limit for Computers

How fast can a computer compute? To put the question another way, how far is it possible to reduce the number of sequential steps a computer must take in carrying out an elementary arithmetic operation? Now designers of electronic logic circuitry have an exact answer to guide their efforts in this direction, in the form of a set of general formulas put forward by Shmuel Winograd, a mathematician on the staff of the International Business Machines Corporation.

According to Winograd, the limit depends on the type of basic circuit one is dealing with. His formulas are composed of only three nontemporal factors: (1) the number of different signals that can be sent over a line between two basic circuits; (2) the "fan-in," or the number of input leads to a circuit; (3) the "word size," or the number of digital positions that are available in an arithmetic unit to represent a single number.

For most digital computers in use today the first factor is fixed by the initial choice of a binary system of logic. In this situation only two signals—a 1 or a 0—can be sent over any line. (A binary logic circuit is designed so that its output signal represents a logical relation among its input signals. For example, an "and" circuit will produce a 1 if all input signals are 1, and a 0 if they are not.)

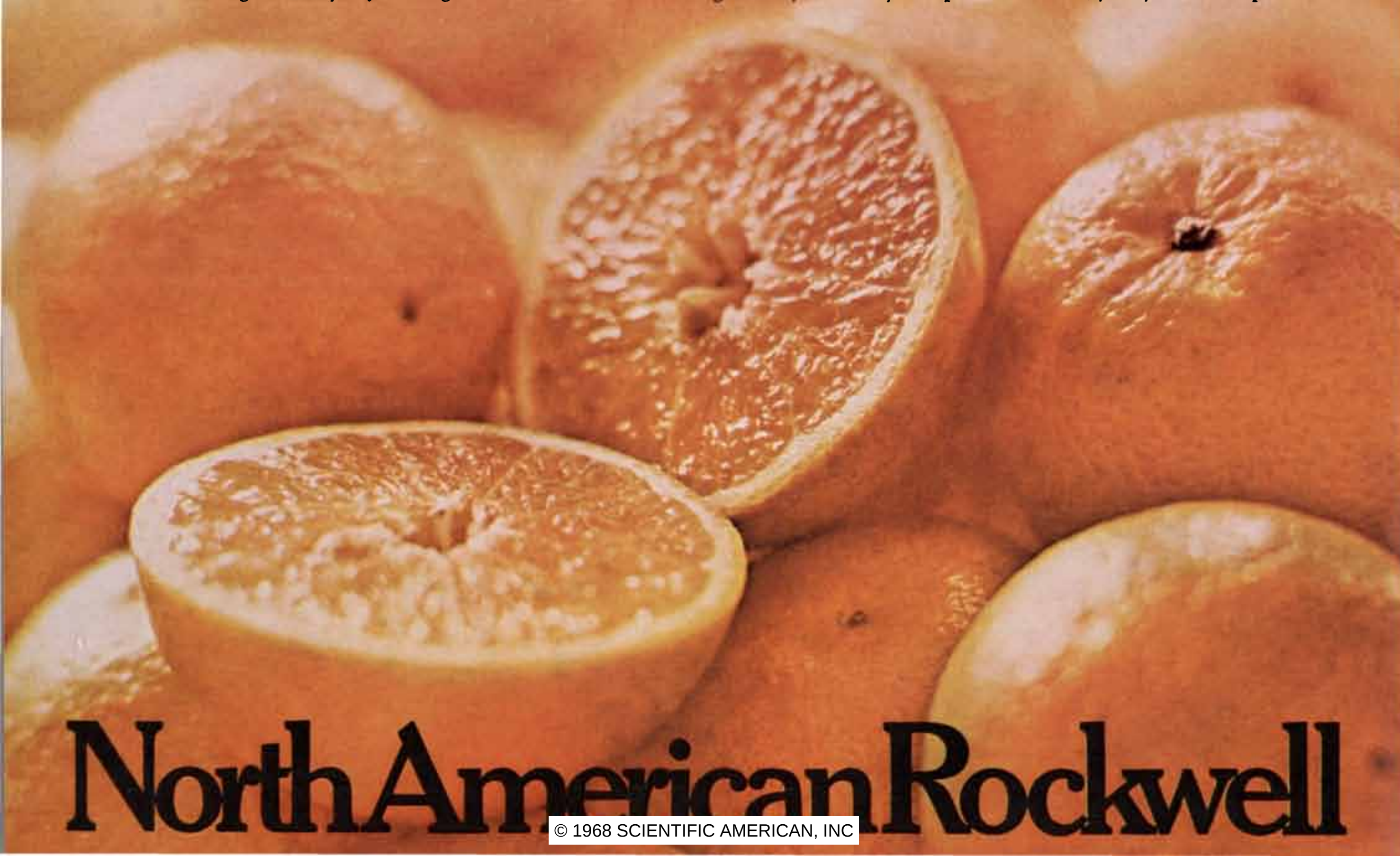
Furthermore, most computers use what is called a modulo-2^n arithmetic, in which n is the word size. For instance, if n is three, then only the first three digital positions of the binary number system are available in the arithmetic unit to represent a number. Hence the unit can represent the numbers 0 through 7, since there are only eight possible combinations of 1's and 0's that can be made with three available positions. Such a unit is said to perform modulo-8 arithmetic.

Applying Winograd's general formula for addition to the typical case, in which only two signals can be sent over any line and the computer uses a modulo-2^n arithmetic, one obtains the simplified formula $t \geq [\log_r 2n] \Delta t$, where t is time, r is the fan-in, n is the word size and Δt is the time delay of each binary circuit. The expression inside the brackets is the minimum number of steps in the computational sequence of an arithmetic unit. If the quantity inside the brackets is not an integer, it evaluates to the next higher integer. The formula shows that the ultimate speed of an assembly of basic circuits increases with increasing fan-in and decreases with increasing word size.

For example, assuming a fan-in (r) of three inputs, a time delay (Δt) of 10 billionths of a second and a word size (n) of 48 binary digits, the formula predicts that addition cannot be performed in fewer than five steps, and that the minimum time in which two 48-binary-

Every day a large part of California gets moved to New York and vice versa.

Americans have never let a little thing like a few thousand miles stand between them and the things they want. You live in New York and you want California oranges? Aisle three, right next to the pineapples from Hawaii. You live in Los Angeles and you just bought your wife a dre s? Look at the label. New York, right? Getting things from where they are to where they're supposed to be is a big, important business in this country. And trucks are a big part of the business. Last year, trucks hauled 390 billion ton-miles of goods. By 1970 they'll be hauling 500 billion ton-miles a year. But you don't add 33% more ton-mile unless you add something like 33% more trucks. 33%. You know, that's a lovely figure to look at. Especially when you're the world's largest producer of heavy-duty truck components.



North American Rockwell

digit numbers can be added is 50 billionths of a second.

Surprisingly, Winograd's corresponding formula for multiplication, $t \geq [\log_r 2(n-2)]\Delta t$, shows that the ultimate speed of multiplication is slightly faster than that of addition. In order to build adders and multipliers that approach these formulas, however, the code used to represent numbers in the multiplier must be different from the one used in the adder.

The validity of Winograd's general formulas does not depend on the type of coding system used to represent the numbers inside the computer, as long as the code is nonredundant (each number must be represented by one, and only one, code word). Hence the number of steps indicated by the formulas cannot be reduced further, even by radical improvements in the organization of logic circuitry.

Ubiquitous DDT

DDT is the most widely used insecticide and it is toxic, in high enough concentrations, to many higher animals, including man. A number of studies have shown that DDT can be spread by wind and water far from its sites of application. Even more important, it is concentrated by natural ecological cycles: as plankton filter water that contains DDT, as the plankton are then ingested by fish and the fish by marine birds, the concentration of DDT in the tissues of successive predators rises [see "Toxic Substances and Ecological Cycles," by George M. Woodwell; SCIENTIFIC AMERICAN, March, 1967]. Now two investigators from the Freshwater Fisheries Laboratory at Pitlochry in Scotland have reported in *Nature* that seals and porpoises taken in waters far from any source of pesticides contain particularly high concentrations of DDT.

Other studies had shown that residues of DDT and related insecticides are routinely found in marine fishes from Scotland's waters. The insecticide concentration was highest in tissues with high lipid, or fat, content. Seals and porpoises feed on fish, and a high proportion of their body weight is in blubber, or subcutaneous fat. High pesticide levels could therefore be expected in these animals as a result of the process of food-chain concentration. A. V. Holden and K. Marsden therefore subjected samples of the blubber of seals from waters off Scotland and Canada to analysis by gas-liquid chromatography. Their expectation was confirmed. The concentration of DDT alone averaged 7.8 parts per million in seals from the east coast of Scotland and was as high as 23.3. In porpoises from the same area it averaged 21.1 parts per million; in one porpoise the total concentration of the DDT group and dieldrin, another pesticide, was 73.3 parts per million. Holden and Marsden conclude: "This degree of accumulation of residue in an environment not deliberately contaminated, and in species ecologically far distant from the target organisms of persistent pesticides, underlines the impossibility of confining such chemicals to the areas of application."

One-Atom Microscope

A microscope capable of examining a single atom has been developed by Erwin W. Müller and his associates at Pennsylvania State University. The instrument, called the atom-probe field-ion microscope, represents a major modification of the field-ion microscope invented by Müller in 1956. Müller, John A. Panitz and S. Brooks McLane describe the new instrument in *The Review of Scientific Instruments.*

A field-ion microscope consists essentially of an extremely fine needle sealed into one end of an evacuated tube and pointed toward a fluorescent screen at the other end of the tube. The needle, which can be made of various pure or alloyed metals, is the sample being analyzed. If a small amount of a gas such as helium is admitted to the tube, and a high positive voltage is applied to the tip of the needle, the helium atoms are ionized in the extremely high field region above each metal atom protruding from the evenly curved end of the tip. The ions are repelled from the positive tip and diverge in a radial direction to project an image of the crystalline surface of the tip on the fluorescent screen. The instrument magnifies atoms from two to five million times so that they can be seen on the screen as clusters of dots, but it does not enable the observer to differentiate between species of atoms. The atom-probe field-ion microscope gives him that capability.

The new instrument differs from the old one in that there is a tiny "probe hole" in the center of the viewing screen. The observer can select one atom that looks interesting to him; he lines it up over the probe hole by tilting the tip mount, which can be done by means of a flexible joint in the microscope. A high-voltage pulse is used to rip the selected atom from the tip of the needle. The particle travels as a positive ion through the probe hole into a mass spectrometer. A detector at the end of the spectrometer's yard-long drift tube signals the arrival of the ion to an oscilloscope, measuring the ion's time of flight. With the ion's energy given by the voltage pulse, its atomic mass and thereby its chemical identity are derived from the time of flight. Heavier particles travel more slowly. Müller and his colleagues express the belief that the instrument will be highly useful to metallurgy and surface physics in providing information about crystal structure and the location and nature of metallic impurities.

The Young and the Dead

Since about 1930 the rate of death from motor vehicle accidents in the U.S. has risen markedly for infants and people aged 15 through 44 while declining or remaining fairly steady for other age groups, according to a group of statisticians in the U.S. Public Health Service. The investigators, headed by Robert E. Markush, describe in the *Journal of the American Medical Association* their study of vehicular death rates by age for the period from 1906 to 1964. They find that for the entire period the death rate was highest for people aged 65 and older, but they note that "if present long-range trends continue, the young-adult group will have the highest rates by about 1970."

Markush and his colleagues offer several hypotheses as possible explanations of the findings. One is that the overall trend can be explained in terms of experience: inexperienced drivers have high death rates, and the rates decline with increasing experience. Dividing the 15-through-44 group into three parts (15 through 24, 25 through 34 and 35 through 44), the investigators found that "the rate of increase diminishes as age increases." They suggest that the rising death rate of infants may reflect the fact that their mothers, who tend to be in the young-driver group, use cars increasingly for domestic activities.

Another possibility cited by Markush and his colleagues is that "mass social phenomena influence motor vehicle mortality by changing individual concern for self-preservation." They note a tendency for vehicular death rates to increase in periods of unrest preceding wars. A related possibility, they say, is that "young drivers may be driving more poorly than before because of psychological factors, such as less concern for self and others and increasing insecurity."

Hormone from the Heart

The heart can now be added to the list of tissues that produce hormones: "chemical messengers" secreted in one part of the body that affect the function of another part. Using cats and dogs as experimental animals in a study of kidney function, Mary F. Lockett of the University of Western Australia has found that when cats suffer a sudden drop in blood pressure, their hearts react by releasing a hormone that acts on the kidneys to minimize the loss of body fluids. Traveling from the heart to the kidneys in the blood, the substance produces an immediate decrease in sodium concentration and urine flow.

Reporting in a series of papers in *The Journal of Physiology,* Miss Lockett notes that the substance is a steroid, that it resembles aldosterone in action and that it can be extracted from heart muscle. Neither the messenger nor its mode of action can be detected in dogs. This finding leaves open the question of whether a similar response to lowered blood pressure is already at work in other animals, including man.

A Radical Proposal

During the past two years radio astronomers have reported the discovery of a number of galactic radio sources with emission spectra characteristic of radiation from hydroxyl radicals (OH). Two recent developments indicate that these previously unidentified sources may coincide with protostars—extremely low-density stars in the precollapse stage of formation. The evidence is discussed in two recent articles in *The Astrophysical Journal.*

In the first article T. K. Menon of the National Radio Astronomy Observatory in Green Bank, W.Va., interprets the results of his study of the distribution and motion of a number of OH-emitting sources, including several in the Orion nebula. He finds that the sources in Orion appear to partake of the motion of a cluster of visible stars within the nebula, rather than of the motion of nebula as a whole. On the basis of this finding he concludes that the OH sources may be protostars of the cluster. He points out that in this case the OH emission could be excited by a maser-like optical-pumping mechanism powered by either ultraviolet or infrared radiation from a central star.

In the second article Ernst Raimond and Baldur Eliasson of the Owens Valley Radio Observatory of the California Institute of Technology report the results of their detailed interferometric measurements of one of the OH sources in Orion. They find that the position of the source is very close to that of an infrared point source that has already been nominated as a possible protostar. If the OH source and the infrared source turn out to be related, this finding would lend support to the theory that hydroxyl radicals are among the first molecules formed in a contracting interstellar cloud.

Mammoth-Hunter's Wrench

An 11,000-year-old wrench, evidently used by the North American big-game hunters of the Clovis culture to straighten their spear shafts, has been found in association with mammoth bones and stone projectile points at Murray Springs in southeastern Arizona. The foot-long tool, made of mammoth bone, is the first of its kind to be found at any prehistoric site in the New World. In the Old World shaft-wrenches and arrow-straighteners are found in the Upper Paleolithic strata of Europe dating from 28,000 to 9000 B.C.; they are usually made of bone or antler, are a foot or so in length with a bevel-edged hole drilled through the wide end, and are often decorated with engraved animal figures. They were grandiosely called *bâtons de commandement* by the 19th-century archaeologists who first discovered them, a name they retained until it was noted that living hunters, notably Eskimos and some American Indians, used similar wrenches to straighten out-of-true spear shafts and arrows. This they did by slipping a crooked shaft through the hole in the wrench and applying appropriate pressure, using the wrench handle as a lever.

The Murray Springs shaft-wrench was found adjacent to a level at the site with a carbon-14 date of about 9000 B.C. The discoverers of the site, C. Vance Haynes and E. Thomas Hemmings of the University of Arizona, have analyzed the size, shape and beveling of the hole in the wrench. Reporting in *Science,* they conclude that, if the tool was indeed used to straighten shafts, the most suitable shaft diameter would have been about 3/4 inch. In view of the Clovis hunters' preference for mammoth and other big game, Haynes and Hemmings suggest that a shaft that slender was probably only the front part of a heavier compound spear or dart.

All this for $995. The new compact 130S from Canon.

Automatic minus sign shows credit balance directly—no long, confusing string of 9's. (And only Canon has it at anywhere near this price.)

Reverse key interchanges dividend and divisor.

13-digit display with ⅝" illuminated numbers you can read without bifocals; automatic overflow indicator, plus automatic floating decimal—no lost or cut-off digits.

Accumulates and totals results of successive operations.

Backs one digit at a time off the display, for error correction or rounding off.

"K" key locks in constant for multiplication or division.

You don't have to go into training to operate it... or carry it! Small and light enough to follow the work from desk to desk. And makes no noise at all. (We can't promise that for the operator. Ours, incidentally, isn't for sale.)

☐ Please show me the Canon 130S.
☐ Please send literature.
Name_______ Title_______
Firm_______
Address_______
City_______ State_______ Zip_______

CANON, U.S.A., Inc., 64-10 Queens Blvd., Woodside, N.Y. 11377

Canon ELECTRONIC CALCULATORS

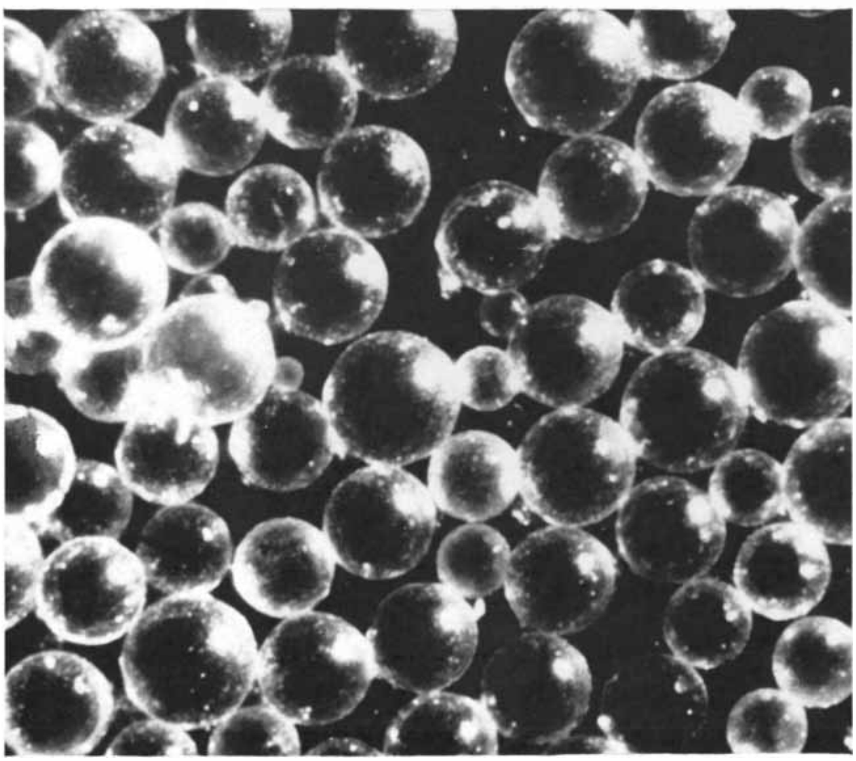
THE AUTOMATIC SYNTHESIS OF PROTEINS

By anchoring an amino acid to a plastic bead one can add other amino acids one by one in automatically controlled steps. This method has already been used to make the small protein insulin

by R. B. Merrifield

The synthesis of proteins is one of the primary functions of the living cell, and the intricate series of operations by which the cell accomplishes the task has recently become known in considerable detail. Long before the biosynthesis of proteins was at all understood, at the turn of the century, the great chemist Emil Fischer believed proteins could be synthesized in the laboratory. It took a long time to accumulate the knowledge and techniques required to put together one of these enormously complex substances. Beginning in 1963, as the result of concerted work by many individuals, three groups of chemists—one in the U.S., one in Germany and one in China—succeeded in making in the laboratory a comparatively simple protein: the pancreatic hormone insulin.

The conventional process for the chemical synthesis of proteins, or of the smaller chains of amino acids called peptides, is a slow and painstaking affair. In our laboratory at Rockefeller University we set out some years ago to look for a simpler and more efficient process—one that might lend itself to automatic operation. The result was a new technique, "solid phase" peptide synthesis, in which the peptide chains are assembled not in solution but on small, solid beads of polystyrene. In 1965 the solid phase method was successfully applied to the synthesis of insulin.



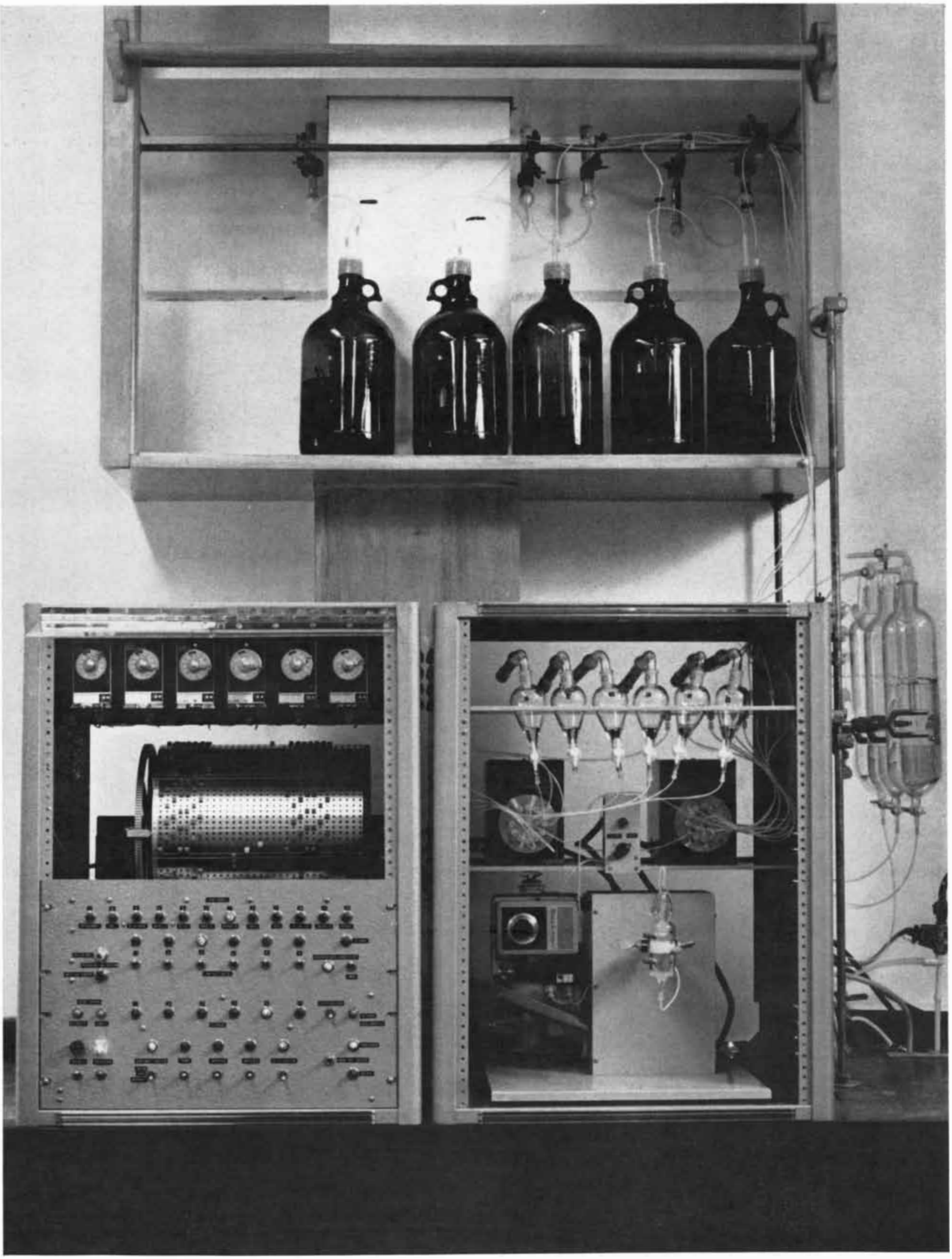
BEADS OF POLYSTYRENE, on which amino acid subunits are assembled in the "solid phase" method to make a peptide chain, are enlarged 300 diameters. They average 50 microns (.002 inch) in diameter but become about twice as large when swollen in a solvent to make them more reactive. About a trillion peptide chains can be "grown" on a single bead.

The availability through controlled chemical synthesis of insulin with all the properties of the natural hormone is of great value because the road is now open for making related molecules that differ from the parent compound in precisely known ways. Such analogues should help to clarify the manner in which the natural hormone functions and may in time lead to insulin derivatives that exhibit greater or more prolonged activity for the treatment of diabetes.

Even more exciting than the synthesis of insulin is the possibility of synthesizing an enzyme—one of the proteins that catalyze metabolic processes. That goal is now in sight and will surely be attained. It will add significantly to our understanding of this important class of proteins and of the mechanisms by which the living cell carries out its essential functions.

Peptide Synthesis

To understand how the chemical synthesis of a protein can be approached it is best to look first at how peptides are synthesized. Peptides are simpler models of proteins; they consist of the same structural elements (but fewer of them) linked together in the same way. The linkage between the amino acid subunits is known as the peptide bond; a series of such bonds constitutes the primary "backbone" structure of peptides and

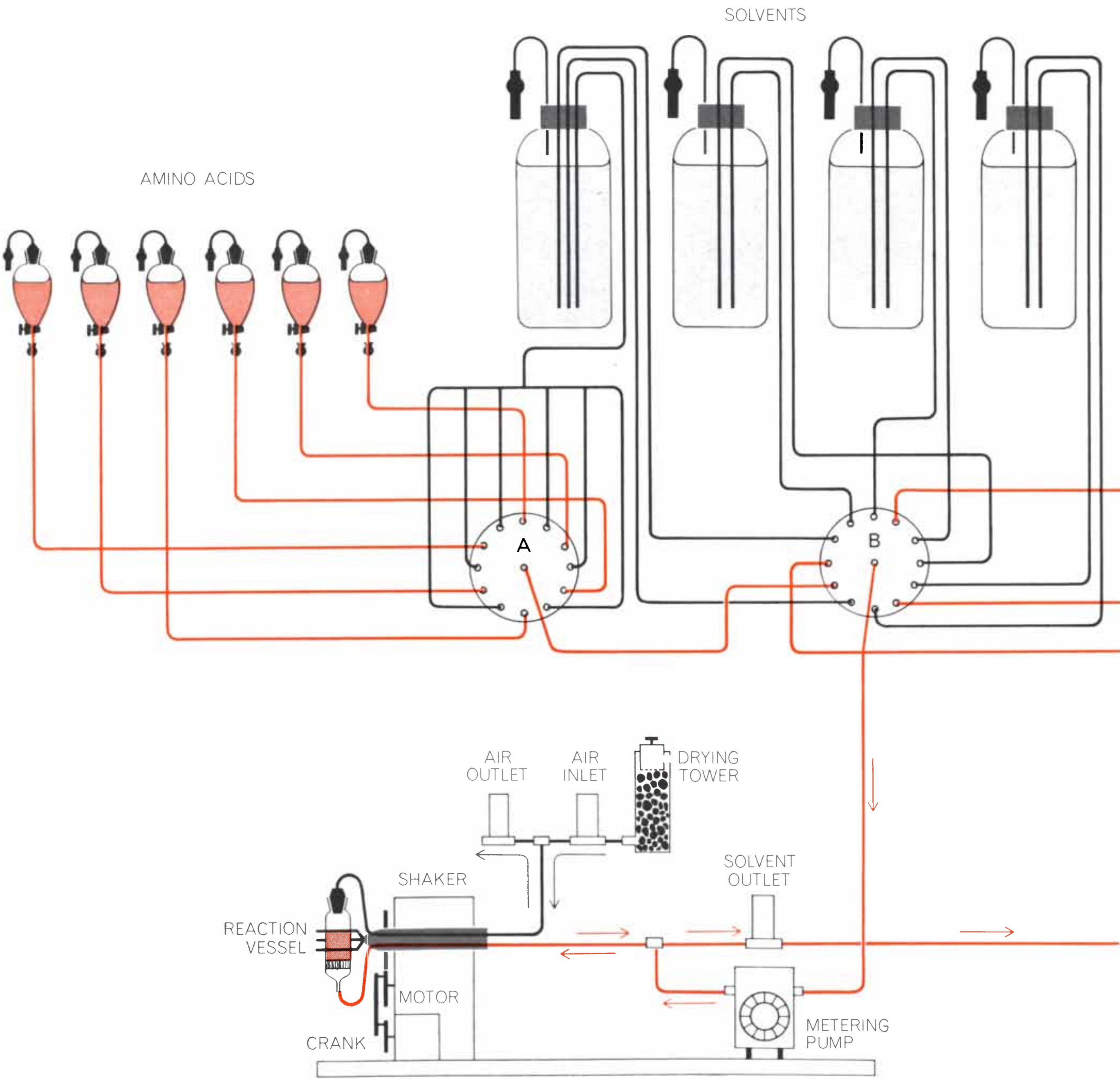


APPARATUS for the automatic synthesis of peptide chains is seen in the author's laboratory. It includes the small glass reaction vessel (*lower right*) with its attendant "plumbing" and a programming unit (*left*). The rectangular pins on the rotating drum operate switches that control the pump, valves, timers and shaker that fill and empty the vessel and mix the reagents. Amino acids are supplied from the six glass vessels (*middle right*). Solvents and other reagents are supplied from the larger containers above and at right.

proteins. The formation of these bonds is the principal problem in peptide synthesis and is also the first step in protein synthesis. In the case of proteins, however, there are also secondary and tertiary bonds that control the cross-linking and folding of the molecule and are responsible for its three-dimensional shape. Because peptides are shorter and lack these complicating additional bonds, they are simpler compounds to study. They have been used to develop the chemistry needed to begin the synthetic work on proteins.

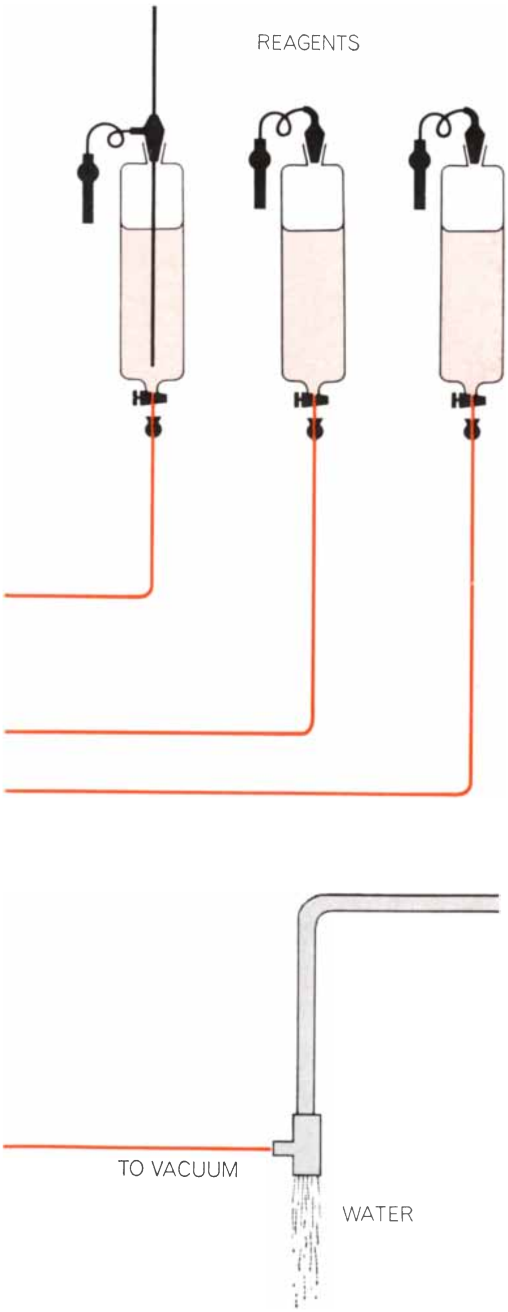
Amino acids are compounds containing several reactive groups: one amino group (NH_2), one carboxyl group (COOH) and in many instances another reactive group located on a side chain [*see top illustration on page 60*]. In general all but one of these groups must be protected against undesired combinations during a chemical reaction if a specific, pure product of known structure is to be obtained. In order to prepare even the simplest chain of two amino acid units (a dipeptide) the basic amino group of one unit and the acidic carboxyl group of the other must be blocked. It is then possible to activate the carboxyl group of the first amino acid—that is, to increase its energy level—so that it will couple with the free amino group of the second one to form the peptide bond.

Now let us consider extending the chain to form a longer peptide. There are two general approaches. In the "fragment" method short peptide chains are built up and are combined to form the larger final molecule; in the "stepwise" method single amino acid units are added one at a time until the final molecule



SCHEMATIC DIAGRAM of the automatic apparatus shows the "plumbing" circuits. The proper amino acid, other reagent or solvent is pumped from its reservoir through a selector valve (*A or B*) into the reaction vessel while air is displaced at the top of the vessel. A mechanical shaker rocks the vessel to mix the reactants. Solvents and by-products are removed by vacuum through the filter in the bot-

is completed. (In both methods the successive additions can in principle be made at either end of the molecule, although in practice there are certain limitations.) The fragment technique is the older and until recently was the more common approach. Its advantages are that more of the intermediate peptides are of small size and that there are greater differences between the properties of the reactants and of the products than there are in the stepwise procedure. On the other hand, the coupling yields are generally lower in the fragment method



tom of the vessel while dry air is admitted at the top. One cycle of the synthesis requires the selection of 12 reagents in sequence.

and there is a greater chance of unwanted side reactions.

Before any of the chain-lengthening processes can be carried out it is necessary to remove one of the blocking groups from the initial dipeptide. To "deprotect" selectively in the presence of the other protecting group (or of several such groups) and without damage to the peptide chain requires careful planning. The choice of the protecting groups and the activating, or coupling, agent for each amino acid has been a major concern of peptide chemists.

At each step of the synthesis it is usually necessary to isolate, purify and characterize the products of the reaction, and it is at this point that the greatest difficulties are often encountered. Crystallization is the classical procedure for purifying peptides, as it is for most other organic compounds. It depends on the formation of an orderly array of molecules that grows in size until it precipitates from solution. Ideally only molecules of one kind will be in the crystalline precipitate and all undesired substances will remain in solution and be washed away. Sometimes one can obtain from a reaction mixture quite pure peptides that do crystallize readily. Particularly when one is working with long peptide chains, however, the yield may often be amorphous material or crude crystalline precipitates contaminated with various by-products. One must then resort to special purification procedures that may require many days each. Consider for a moment the time and effort involved in the synthesis of a 100-unit peptide if such coupling and purification steps must be performed 99 times!

The Solid Phase Approach

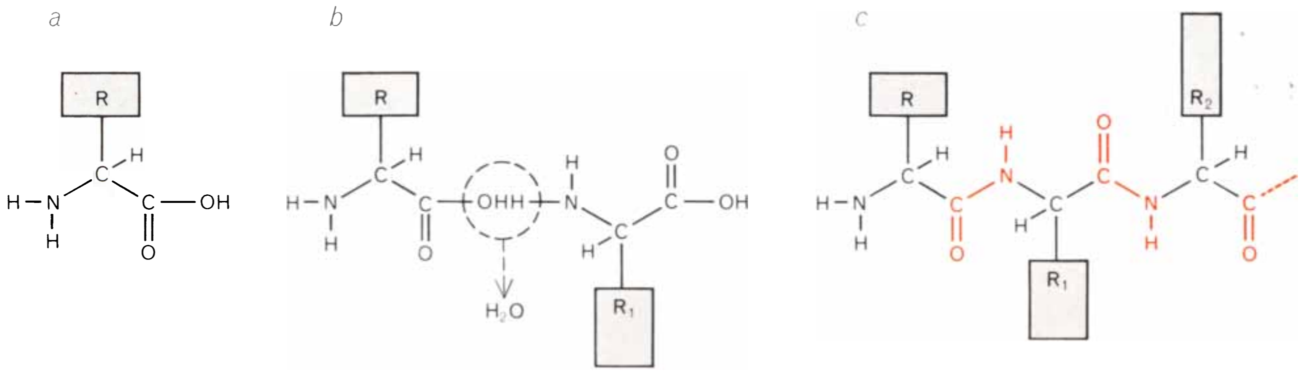
To synthesize molecules of the size and complexity of proteins, it seemed clear, methods of the greatest efficiency and simplicity would have to be developed. In 1959, with these requirements in mind, a new approach to peptide synthesis was conceived. The new idea was to synthesize the long chains one unit at a time, but without stopping to isolate the individual intermediate peptides; to make this feasible the plan was to anchor the chain to an insoluble solid support.

The first amino acid would be chemically bonded to a solid particle and the rest of the amino acid units would be added to it stepwise in the proper order. Since the solid support would be completely insoluble in the various solvents, all the intermediate peptide products would also be held in an insoluble state; they could therefore be purified simply by dissolving the unwanted by-products and reagents and washing them away. This would involve only an elementary filtration step, but it would accomplish essentially the same kind of purification as the classical recrystallization: the growing peptide would be an insoluble precipitate, whereas the undesired reagents would be in solution. Filtration is much easier and faster than crystallization. Most important, it can be done in the same way at each step, whereas crystallization is necessarily an individualized procedure that is different for each new intermediate peptide.

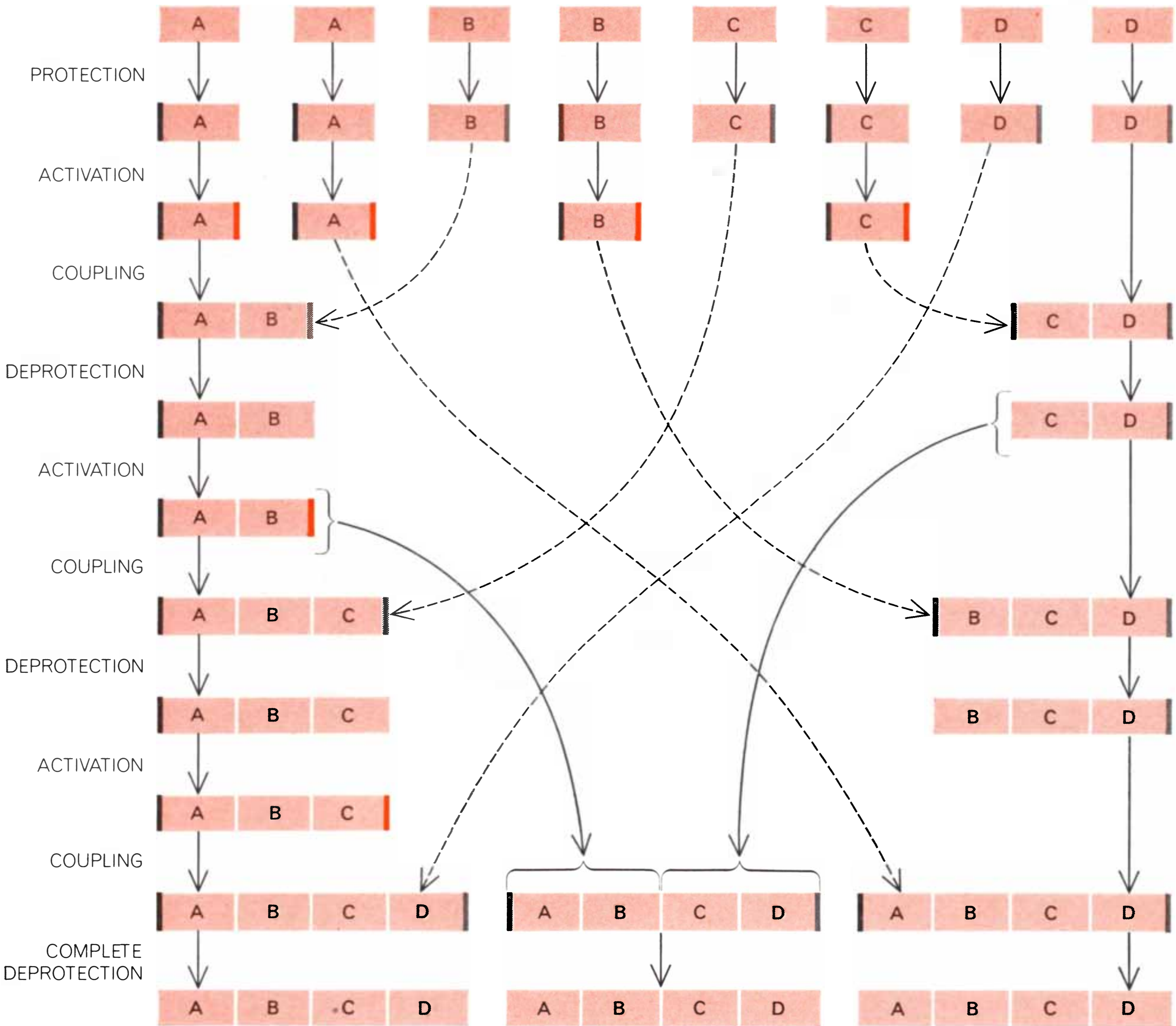
The general scheme for solid phase peptide synthesis is straightforward. A suitable solid support is selected and a reactive site is produced on it. The first amino acid—actually the terminal amino acid of the proposed peptide chain—is attached by its carboxyl group to the reactive site. Now the second amino acid, with all but one of its reactive groups protected, is activated and coupled to the first amino acid, leaving a protected dipeptide firmly bound to the support. The solid can be filtered and washed thoroughly to remove all the excess reagents and any by-products without the slightest danger of losing the desired peptide.

Next the protecting group on the amino end is removed and the whole process is repeated exactly as before but with a new amino acid. After the required sequence of amino acids has been assembled in this manner the peptide chain is finally removed from the support by selectively breaking the bond that has been holding the two together throughout the synthesis. Now for the first time the peptide chain is free and can be dissolved and separated from the solid support. Once it is in solution conventional purification procedures can be carried out.

Before this scheme could be developed into a workable procedure a number of rather severe requirements had to be met, having to do with the nature of the solid support, the type of bond linking the peptide to the support and the choice of protecting groups and coupling reagents. The solid support had to be completely insoluble in all the solvents that might be used in the synthetic reactions or in the washing steps; it had to be physically stable and in a convenient form to permit filtration and other manipulations; it had to have a reactive site at which the peptide chains could be attached but should otherwise be chemically inert and stable; finally, in order to



AMINO ACID (*a*) has an amino (NH_2) and a carboxyl (COOH) group separated by a carbon atom that carries a side chain (*R*). The peptide bond (*b*) forms between the carboxyl and the amino ends of two units with the elimination of a molecule of water. A series of amino acid subunits held together by such bonds (*color*) constitutes the primary ("backbone") structure of peptides and proteins (*c*).



TWO GENERAL APPROACHES to peptide synthesis are diagrammed. In the "stepwise" method amino acid units are added one at a time starting at the amino end of the peptide (*left*) or the carboxyl end (*right*) until the final peptide (of four amino acids, *A*, *B*, *C* and *D*, in this case) is assembled (*bottom left and right*). In the "fragment" method small peptides are prepared stepwise and are then combined (*long solid arrows*) to form the final peptide (*bottom center*). In each case it is necessary to protect the amino ends (*black bars*) and carboxyl ends (*gray bars*) against unwanted reactions, to activate the carboxyl ends (*dark-color bars*) for coupling, to "deprotect" selectively in preparation for the next step and finally to deprotect completely. The peptide can then be purified.

allow the synthesis of a sufficient quantity of peptide, it should either have a very large surface-to-volume ratio or be readily permeable to the soluble reagents.

After considerable exploration a substance that met these requirements was found. It is a polystyrene resin, a linear polymer of styrene in which the styrene chains are loosely linked together with divinylbenzene, and it is in the form of beads about 50 microns (.002 inch) in diameter. The amount of cross-linking agent was selected to give a resin of high molecular weight that would be completely insoluble but at the same time free to swell in organic solvents. This makes the beads permeable to reagents dissolved in the solvent, so that reactions can occur not only on the surface but also within the interstices of the gel-like matrix. Although we are talking about beads that are barely visible as separate particles to the unaided eye, they are actually enormous compared with the dimensions of amino acids or even of proteins: each bead can support some 10^{12} (one trillion) peptide chains!

The Anchor Bond

Polystyrene itself has no convenient reactive site for anchoring the peptide chain, but it can be readily modified in many ways to make such a site. The choice of the modification was dictated by the kind of bond needed to hold the peptide to the resin. The bond must be easy to form, it must be completely stable during the dozens of reactions involved in assembling the peptide chain and it must be readily cleaved under relatively mild conditions at the end of the synthesis. The anchor we chose was prepared by attaching chloromethyl groups ($ClCH_2$) to the six-carbon rings of the polystyrene and then reacting them with the first amino acid to form what is known as a benzyl ester [*see illustration on next page*]. The benzyl group also served to protect many of the reactive side chains.

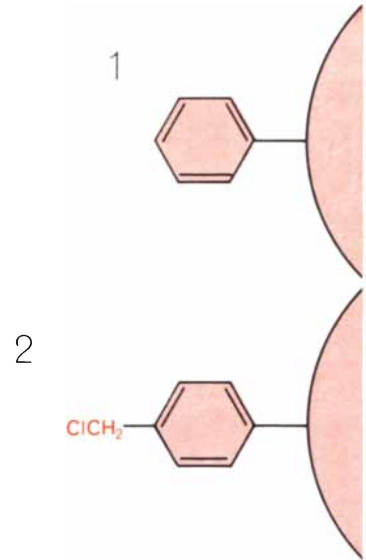
The choice of the benzyl group in turn influenced the choice of the protecting group for the amino ends of the successive amino acids. That is, it had to be possible to remove the amine protection selectively at every cycle of the synthesis without detaching the peptide from the resin or deprotecting the side chains. The protecting agent we chose was tertiary butyloxycarbonyl ("Boc"), a group that had been developed a few years earlier for use in conventional syntheses. It was sensitive to certain anhydrous acids that would not affect the ester and it could therefore be removed without disturbing the anchor.

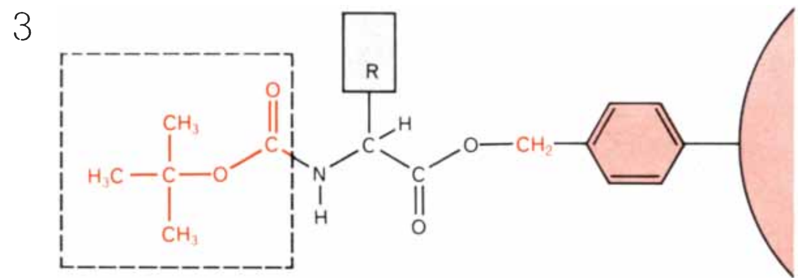
The principal remaining problem was the peptide-forming reaction itself, which had to be rapid and must not give rise to side reactions. Most important of all, it had to go to completion; it was absolutely crucial to the success of the process that this step give essentially 100 percent yields. Suppose, for example, the second amino acid were to couple to the extent of only 90 percent, which is considered a most acceptable yield in organic chemistry. What will happen when the third amino acid is coupled? It will react with the amino end of the dipeptide to form a tripeptide, but it will also couple with the unreacted 10 percent of single amino acids to produce dipeptides that lack the second amino acid. In an ordinary synthesis the intermediate products are isolated at each stage; in the solid phase method they are simply washed free of soluble impurities, and the abnormal chain will be carried through the entire synthesis, giving at the end a mixture consisting of 90 percent of the correct peptide and 10 percent of a peptide with a missing link. In this simple case the two products could probably be separated and purified, but if incomplete reactions were to occur several times during the synthesis of a long peptide, a complex mixture would result that might not be so easy to separate [*see upper illustration on page 67*].

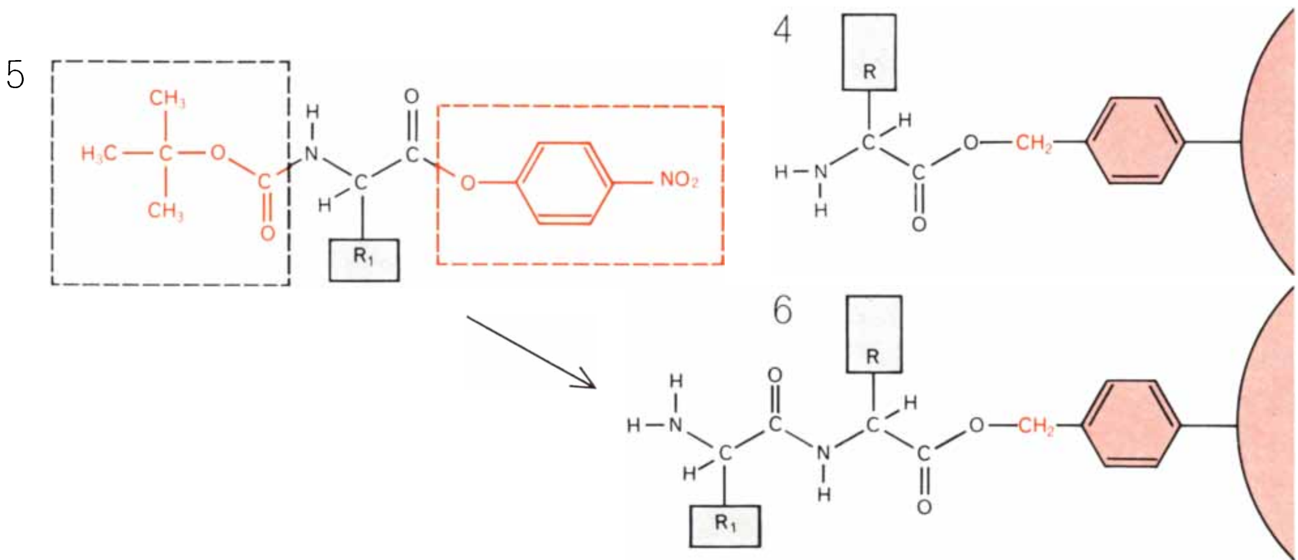
We therefore put great stress on finding reagents and conditions that would lead to complete coupling reactions, and we have fortunately been able to achieve that goal in practice. Many ways to activate amino acids have been developed for use in peptide synthesis. For the solid phase method the most successful procedures have been to activate with the reagent dicyclohexylcarbodiimide or to use the nitrophenyl esters of the amino acids. These activated forms are highly effective, but only if they can reach the proper sites. To ensure their rapid, unimpaired penetration into the resin, where most of the peptide chains are located, a solvent of high swelling capacity, such as methylene chloride, is necessary.

This solvent causes the beads to swell to approximately twice their original diameter, which means that the polymer chains are then distributed in eight times the initial volume. The polymer molecules occupy only about 12 percent of the total space of each swollen bead, and the remaining 88 percent is filled with solvent containing the activated amino acid molecules. The diffusion of these small molecules is relatively little inhibited by the polymer, so that the reactions take place almost as fast as they would in solution.

The efficiency of coupling also depends on the concentrations of the reactants: the peptide chain and the amino acid being added to it. If one begins with equal amounts of the reactants, their concentrations will decrease to very low levels as the reaction nears completion and the reaction rate, which is propor-







SOLID PHASE METHOD is carried out stepwise from the carboxyl end toward the amino end of the peptide. An aromatic ring of the polystyrene (*1*) is activated by attaching a chloromethyl group (*2*). The first amino acid (*black*), protected by a butyloxycarbonyl (Boc) group (*black box*), is coupled to the site (*3*) by a benzyl ester bond and is then deprotected (*4*). Subsequent amino acid units are supplied in one of two activated forms; a second unit is shown in one of these forms, the nitrophenyl ester of the amino acid (*5*). The ester (*colored box*) is eliminated as the second unit couples to the first. Then the second unit is deprotected, leaving a dipeptide (*6*).

The trouble with polyester cord tires is they look like ordinary tires.

With the difference polyester cord has made in tires, they ought to come out looking like works of art.

The new polyester cord tires, for instance:

1. Ride as smooth as rayon. (No thump, thump, thump with a cold start.)
2. Are as tough as nylon.
3. Run cooler.
4. Wear more evenly; fight squirm.
5. And last longer.

At last, a safer, smooth riding, hard driving, long distance tire has been born.

We're not tire manufacturers. We don't even make polyester tire cord. All we do is supply two key ingredients used in making polyester: high purity DMT (Dimethyl Terephthalate) and TA (Terephthalic Acid).

(We were the first company in the U.S. to produce a commercial fiber grade TA, and the only company to this day that supplies both chemicals.)

That doesn't give us the liberty to tell anybody how to design tires.

And maybe no one would buy one that looked like this. But, we thought we ought to do something to make the point.

Amoco Chemicals Corporation, 130 East Randolph Drive, Chicago. Illinois 60601.

AMOCO CHEMICALS

Where what's happening gets its start.

Although computers aren't new to the industry, only now has it been demonstrated how completely computer technology can serve every phase of an airline's organization.

This happened when United Air Lines commissioned UNIVAC to design and build an on-line computerized information system.

It represents the largest such investment the business world has seen thus far. The first to utilize cathode ray tube sets (input/output devices resembling TV monitors) on a nationwide basis.

Over 2800 of these UNISCOPE™ visual communication terminals will link 116 United cities to a centralized complex using three UNIVAC® 1108-II computers.

To the United customer this will mean better service.

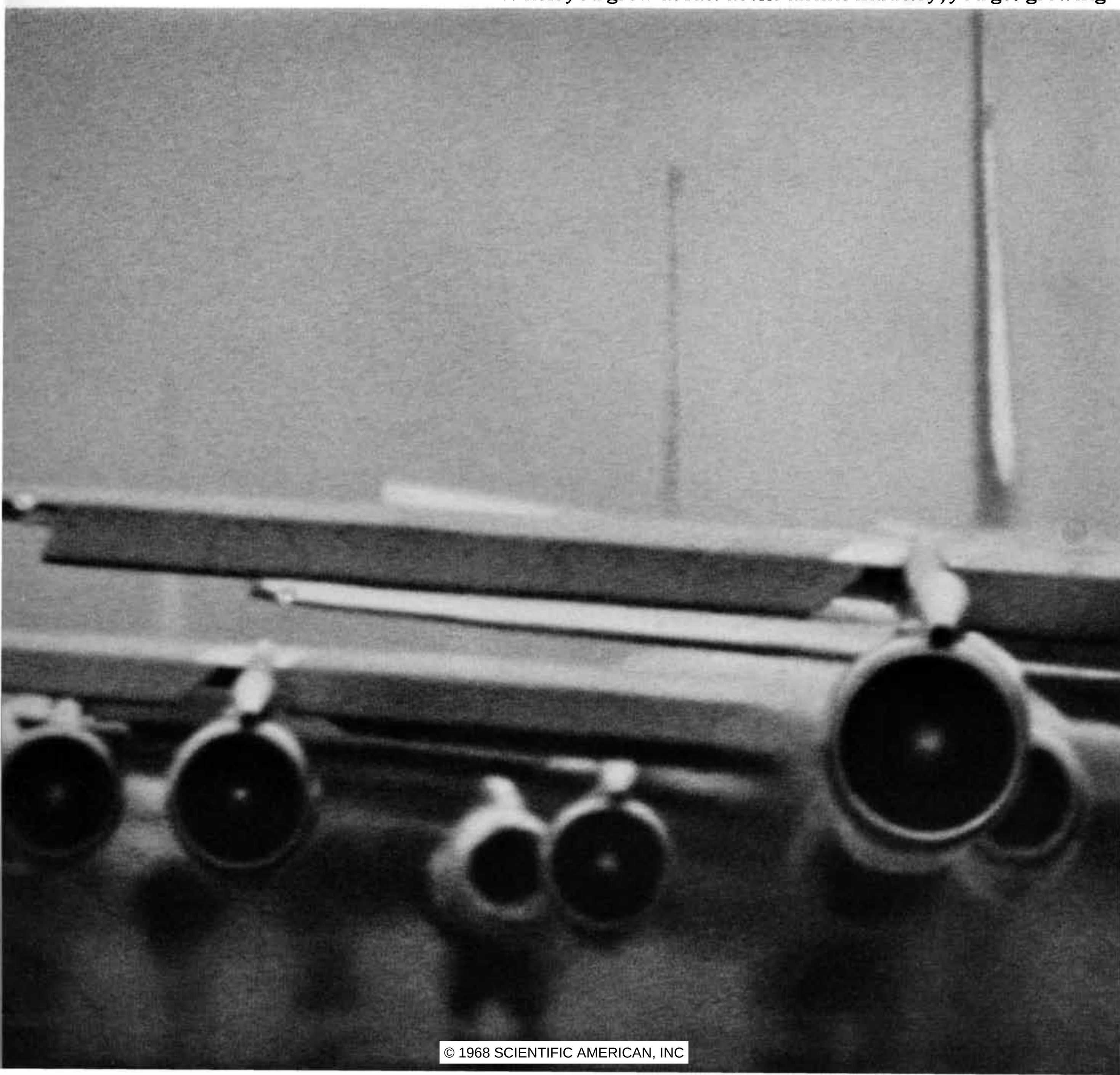
When a United agent queries his set on seat availability he'll get a response in one second.

And in that time he'll also get a complete readout on his screen with availabilities for five flights in addition to the one he requested.

Agents at ticket counters and in air freight terminals will have the unique advantage of automatic ticket or airbill printing provided by any one of 700 printer units that are integral to this

Univac

When you grow as fast as the airline industry, you get growing

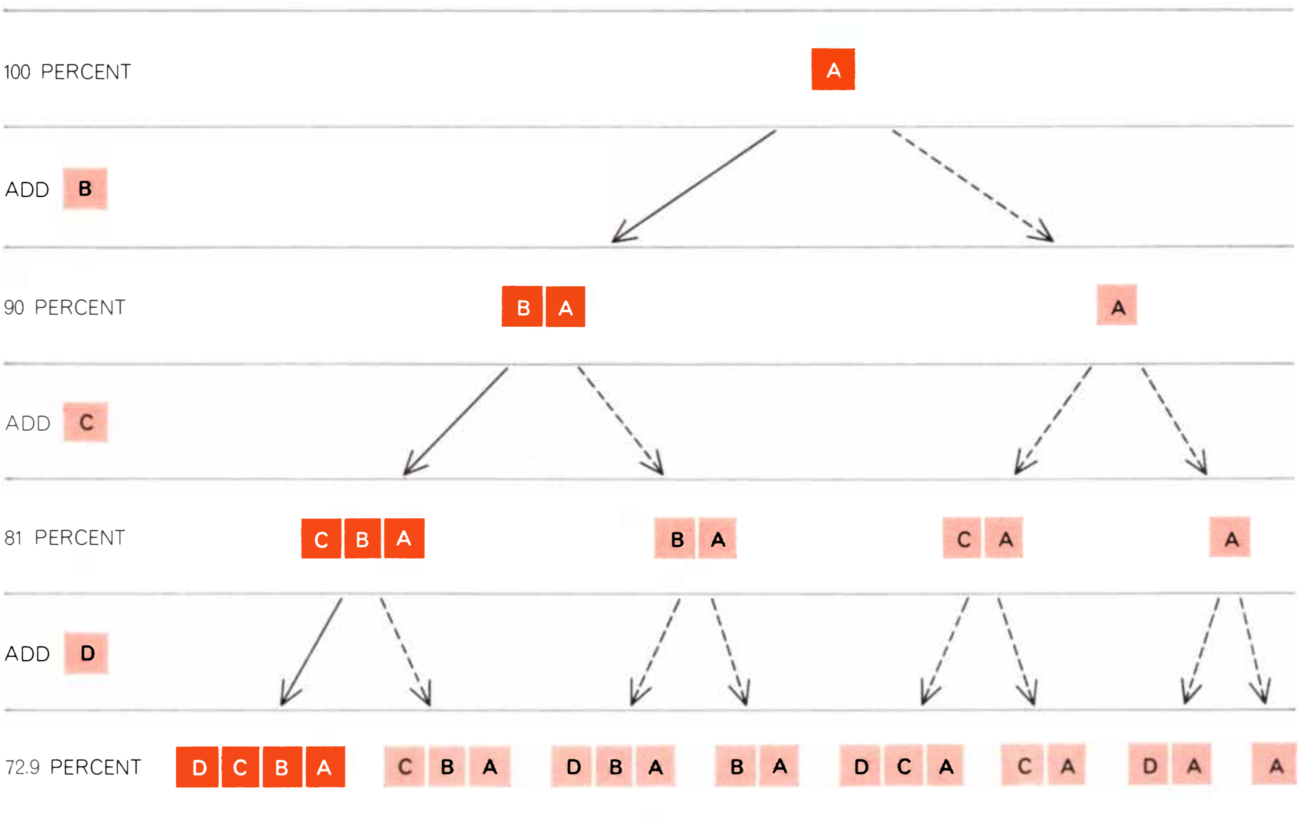


tional to the product of the concentrations, will gradually approach zero. The practical consequence is that some of the amino acids never do link up, and the reaction never quite goes to completion. This is the usual situation in conventional syntheses.

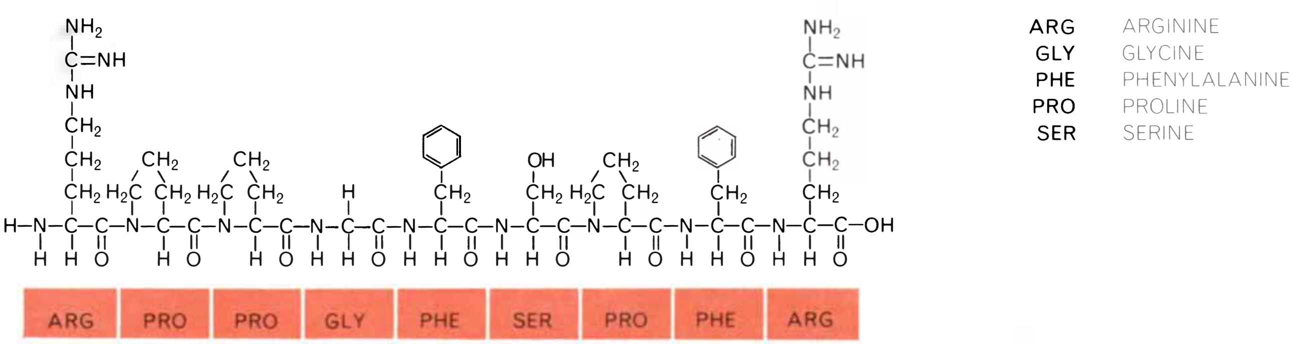
If, on the other hand, there is a rather large excess of one of the reactants, a significant rate can be maintained until essentially all the limiting reactant (the peptide chain) has entered into bond formation. To illustrate, suppose we begin with 100 parts of each reactant and call the initial reaction rate 10,000 (100 × 100). After the reaction is 99 percent completed, one part of each reactant would remain, and the relative rate would be only 1, or 1/10,000 as fast as at the beginning. Even with a very fast initial rate it would take a long time to complete the reaction.

Suppose instead we begin with 100 parts of the peptide chain as before but with 400 parts of the activated amino acid. Now after 99 percent of the chain is used up the relative rate would still be 301 (1 × 301). In the presence of the fourfold excess of the amino acid, one can calculate, the reaction can go to 99.99 percent of completion in the same time it would take to go to only 75 percent if equal amounts of the reactants had been used. One of the important advantages of solid phase peptide synthesis is that such an excess of the amino acid derivative can be used without complicating the subsequent purification procedure, since at the end of each reaction the excess is simply removed by filtration and washing. Thus we can force the reaction to completion and leave essentially no free, unreacted peptide chains.

The deprotection and coupling reac-

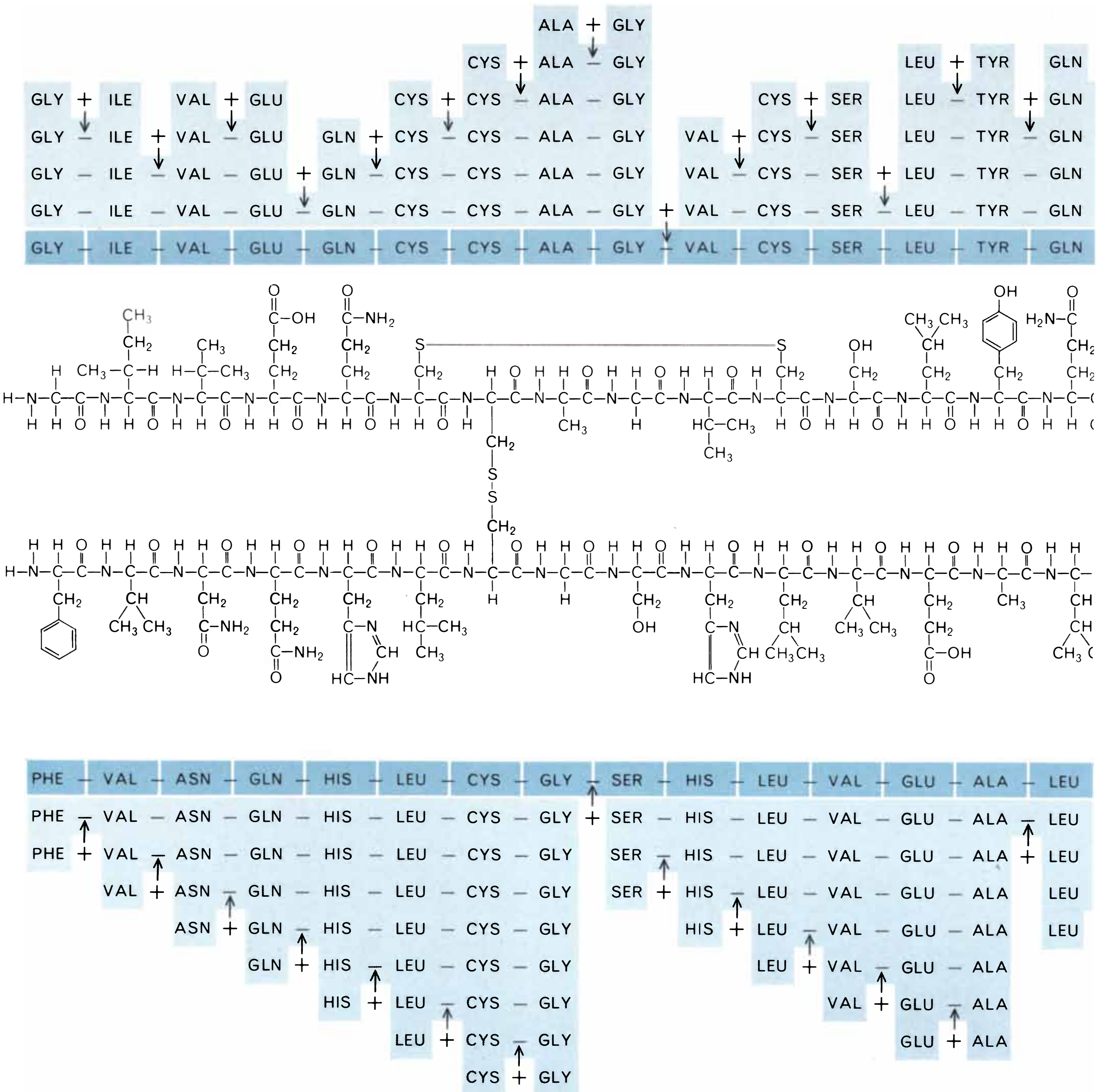


HIGH YIELD from each coupling reaction is important in the solid phase method because products of incomplete reactions persist through the filtering steps. If each amino acid were to couple with even a relatively high efficiency, say 90 percent, the yield of pure peptide (*dark color*) would be down to 72.9 percent by the time the fourth unit was added. More important, the seven different peptide fragments that lack one amino acid unit or more would have to be separated chemically. In practice, yields are close to 100 percent.



BRADYKININ, a small peptide hormone, was one of the first peptides synthesized in the author's laboratory by the solid phase method. Its nine amino acid subunits (five different amino acids) were assembled stepwise from the carboxyl (*right*) end of the chain.

tions just described can be repeated alternately until the desired peptide chain has been assembled on the resin beads. The final step is the cleavage of the benzyl ester bond that has been holding the chain to the resin throughout the synthesis. As indicated earlier, this bond was chosen because it is stable during the synthesis but can be selectively split at the right time without damaging the peptide chain. The resin is suspended in anhydrous trifluoroacetic acid, and dry hydrogen bromide gas is bubbled through to effect the splitting. (More recently anhydrous liquid hydrogen fluoride has been successfully employed for this step.) These reagents also remove the protecting groups on side chains. The peptide, now in a free and soluble state, is separated from its resin support by filtration and is purified. It



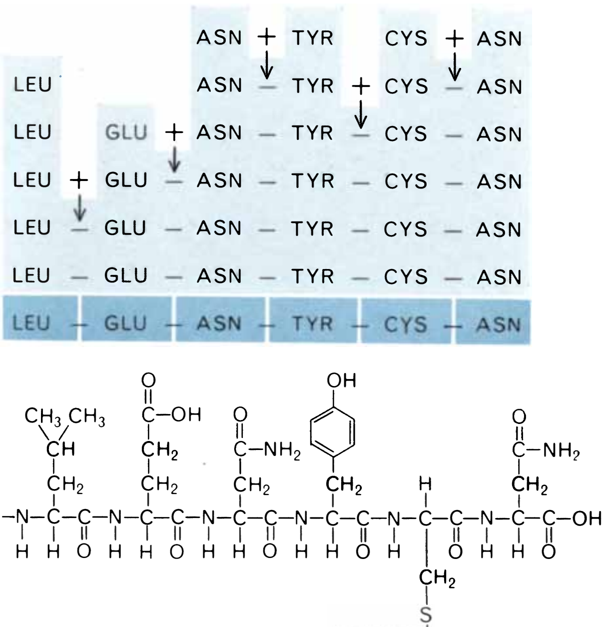
FIRST SYNTHESIS of insulin chains was accomplished by the fragment method, as illustrated here. The primary structure of the two chains is diagrammed at the center, together with the names of the amino acid units (*see key at top right*). Note the disulfide (*S—S*) bonds between cysteine units, two of which link the *A* (*top*) and *B* (*bottom*) chains. The general pattern of the synthesis of the is then ready for analysis and, where possible, for biological assay.

Those who are familiar with the mechanism of protein synthesis in the living cell will see some superficial resemblance between it and the system just described. Both depend on a particulate support (in the cell the support is the ribosome), both involve activation of the amino acid (in the cell the amino acid is activated by the energy-rich molecule adenosine triphosphate, or ATP) and both are stepwise processes. It even has been learned recently that in bacteria the synthesis starts with one end of the chain protected and that an enzyme later carries out a deprotection step just as we do in the laboratory [see "How Proteins Start," by Brian F. C. Clark and Kjeld A. Marcker; SCIENTIFIC AMERICAN, January].

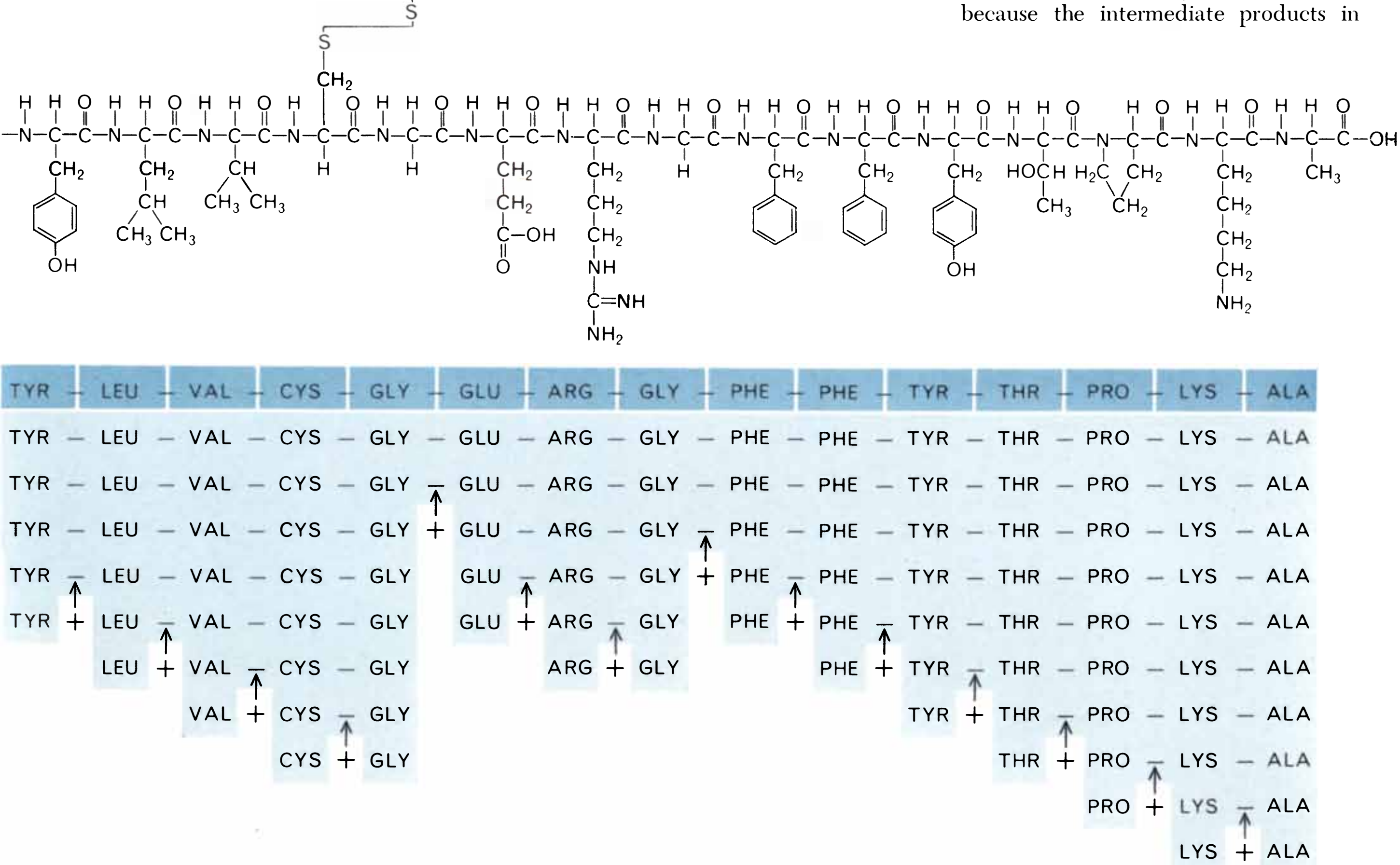
The analogy should not be pushed too far, however. The natural synthesis is much more elegant and efficient than the laboratory one, and no one presumes to duplicate or even approach the complexities of the cell's scheme at this point. As a matter of fact, the chemical synthesis was not even patterned after nature; it was only in retrospect that the similarities became evident. Nevertheless, it could well be that in the future organic chemists may benefit from a more complete understanding of the way in which living systems perform their tasks.

An Automatic System

It was clear to us from the outset that an automatic mechanized process was needed for making large peptides, and that the solid phase approach would be well suited to such a process. This is so because the intermediate products in



ALA	ALANINE
ARG	ARGININE
ASN	ASPARAGINE
CYS	CYSTEINE
GLN	GLUTAMINE
GLU	GLUTAMIC ACID
GLY	GLYCINE
HIS	HISTIDINE
ILE	ISOLEUCINE
LEU	LEUCINE
LYS	LYSINE
PHE	PHENYLALANINE
PRO	PROLINE
SER	SERINE
THR	THREONINE
TYR	TYROSINE
VAL	VALINE



A chain by P. G. Katsoyannis' group at the University of Pittsburgh is shown above the insulin formula. The 21 amino acid units were assembled stepwise into intermediate fragments of from two to five units and the fragments were then coupled in stages to assemble the complete chain. The similar synthesis of the *B* chain by Helmut Zahn's group in Germany is shown below the insulin formula.

the synthesis need not be isolated but are purified by simple filtration and washing reactions that can be carried out in a single vessel; the manipulations required to transfer products from one container to another have been eliminated. Once the resin beads with an amino acid attached are placed in the vessel it is only necessary to introduce the appropriate liquid solvent or reagent, allow it to react, remove the excess reagent and by-products by filtration and then repeat the process with the next reagent.

It is easy to visualize how all these steps can be accomplished automatically by a rather simple device. We constructed a machine that consists essentially of two parts: a reaction vessel with the plumbing necessary to introduce and remove the solvents in the right order at the right times and a programmer that controls these operations. The solvents and the reagents, contained in a series of reservoirs, are selected one at a time by a specially designed rotary valve. The solvents are introduced into the bottom of the reaction vessel by a metering pump while air is displaced at the top. Valves to the vessel are then closed and a mechanical shaker mixes the reactants for a predetermined time. Next a vacuum withdraws solvent through a porous glass filter disk in the bottom of the vessel while dry air enters at the top; the beads, with the peptide attached, remain in the vessel. One cycle of the synthesis (the lengthening of the peptide chain by one amino acid) requires 12 different reagents, one of which is a protected amino acid. The next cycle calls for the same series of reagents except for a different amino acid, which is selected by a second rotary valve.

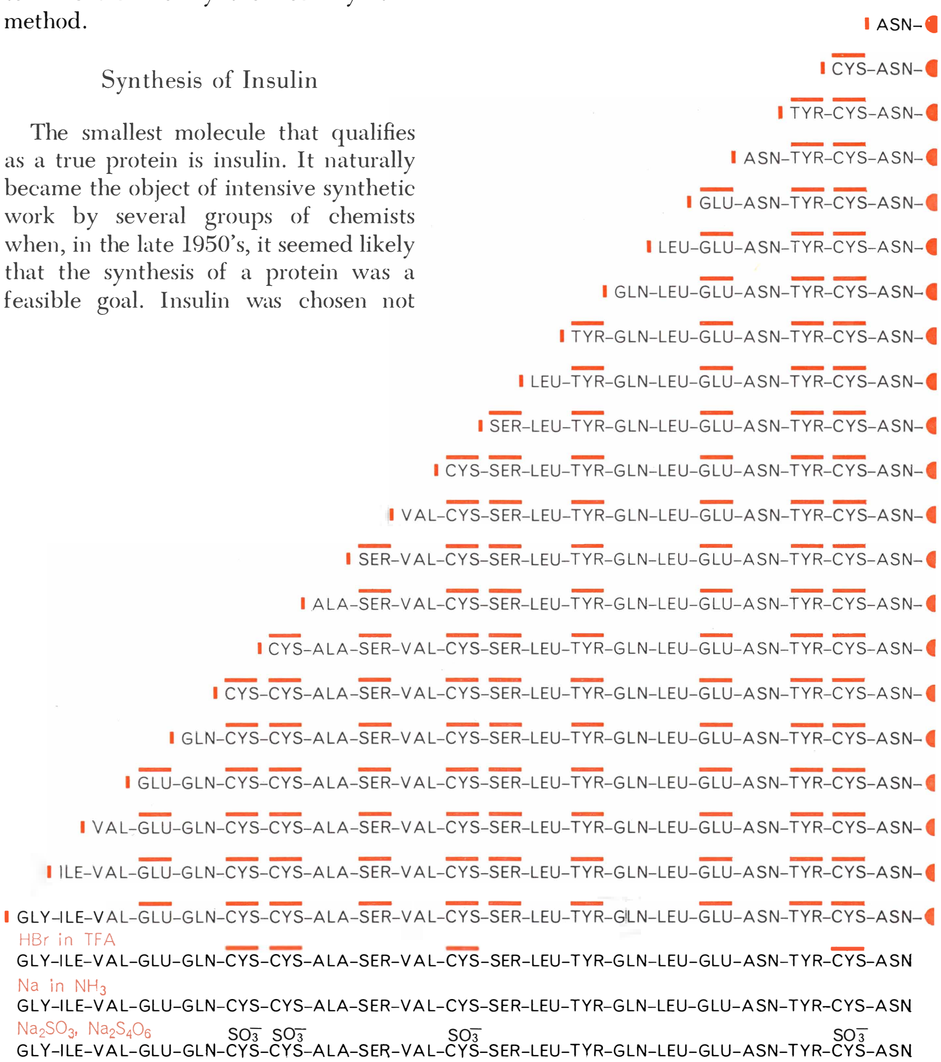
All the steps just described are controlled by a "stepping-drum" programmer. It is like an old-fashioned music box. Once the pins have been positioned on the drum to play the proper tune the machine takes over and directs the chemical synthesis. The pins activate switches that turn the pump and shaker on and off and open and close the valves at the proper times and in the proper sequence. One cycle of the synthesis requires 100 steps of the drum and takes about four hours, so that it is now possible to carry out automatically all the operations required for the assembly of a peptide chain at the rate of six amino acids a day.

After the details of the synthetic scheme had been worked out by the synthesis of small peptides the procedure was given a more demanding test: the preparation of the hormone bradykinin, a nine-amino-acid peptide with several physiological activities that served as sensitive criteria to demonstrate the identity and purity of the final product. The synthetic bradykinin was identical in all respects, both chemically and biologically, with the natural hormone. During the past four years the solid phase technique has been applied to the preparation of bradykinin analogues (nearly 100 of which have been made by J. M. Stewart at Rockefeller University) and other small peptide hormones such as angiotensin and oxytocin. It has also been used in our laboratory and others for the synthesis of the antibiotics gramicidin-S and tyrocidin, and for synthetic studies of the immunological determinants of hemoglobin and tobacco mosaic virus. It was clear that peptides containing 10 or 20 amino acids could be made by the solid phase method as well as by classical procedures. The important question then became whether or not molecules as large and complex as proteins could be synthesized by this method.

Synthesis of Insulin

The smallest molecule that qualifies as a true protein is insulin. It naturally became the object of intensive synthetic work by several groups of chemists when, in the late 1950's, it seemed likely that the synthesis of a protein was a feasible goal. Insulin was chosen not only for its size but also for several other important reasons. The availability of synthetic hormone would help to answer many questions about its mechanism of action. Most important, the composition and complete primary structure (the sequence of amino acid units) had become known a few years before through the work of the group led by Frederick Sanger at the University of Cambridge [see "The Insulin Molecule," by E. O. P. Thompson; SCIENTIFIC AMERICAN, May, 1955].

The insulin molecule is much more complex than a simple peptide such as bradykinin [*see illustration on preceding two pages*]. It not only has nearly six times as many amino acids but also has a greater variety of them: 17 rather than five. This introduces many new prob-

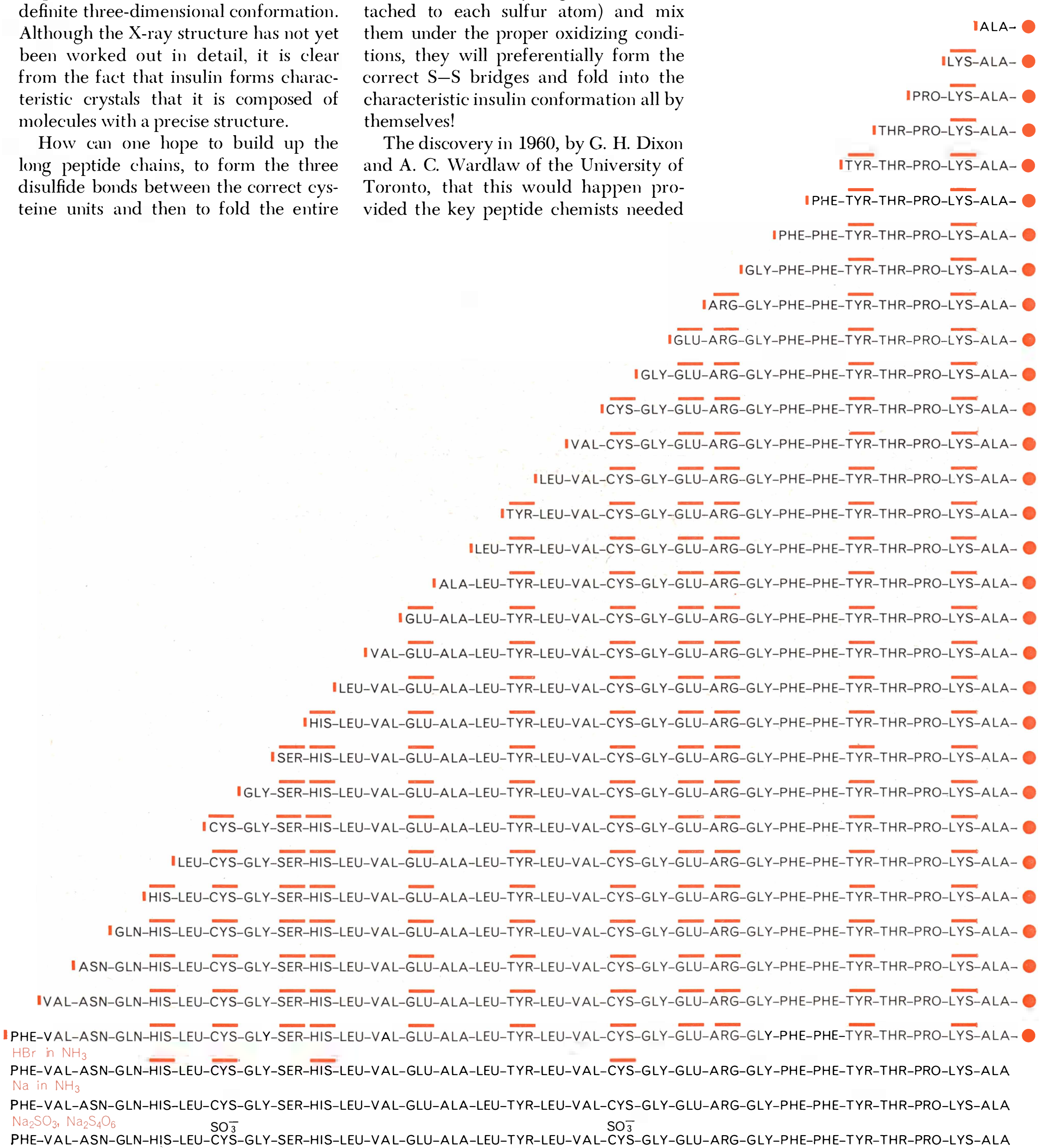


SOLID PHASE SYNTHESIS of the *A* chain (*left*) and the *B* chain (*right*) is diagrammed. An amino acid protected by a Boc group (*vertical bar*) is coupled to a polystyrene bead (*top*), then deprotected. Activated amino acids, protected at the amino end and if necessary

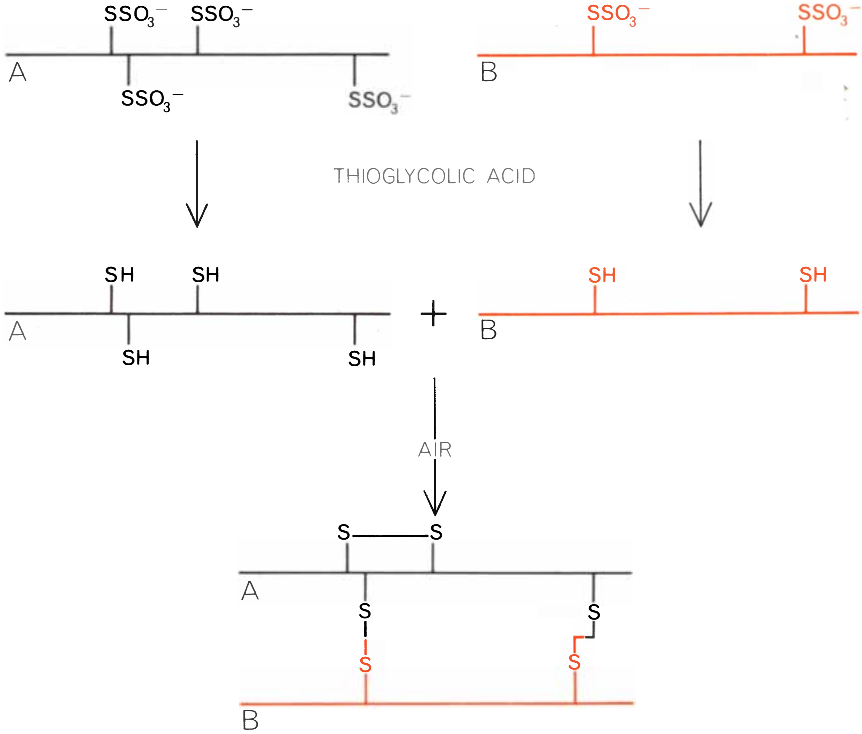
lems of side-chain protection. Particularly complicating is the presence of three disulfide-bond (S—S) cross-links between cysteine units. Insulin consists of two linear peptide chains: an *A* chain with 21 amino acids and a *B* chain with 30. They are held together by two interchain disulfide bridges, and in addition one of the chains has an intrachain disulfide loop. Furthermore, the molecule has a definite three-dimensional conformation. Although the X-ray structure has not yet been worked out in detail, it is clear from the fact that insulin forms characteristic crystals that it is composed of molecules with a precise structure.

How can one hope to build up the long peptide chains, to form the three disulfide bonds between the correct cysteine units and then to fold the entire assembly into its proper shape? This is asking a lot, because there are many possible ways for the S—S bonds to form, and the possible variations in the conformation of the molecule are enormous. It is only possible at this time because nature comes to the chemist's aid. If we simply make the two chains with the six cysteine units all in the reduced (SH) form (in which a hydrogen atom is attached to each sulfur atom) and mix them under the proper oxidizing conditions, they will preferentially form the correct S—S bridges and fold into the characteristic insulin conformation all by themselves!

The discovery in 1960, by G. H. Dixon and A. C. Wardlaw of the University of Toronto, that this would happen provided the key peptide chemists needed to undertake the synthesis of insulin. All four laboratories that have made insulin have depended on this fact and have made the two chains separately. From the point of view of the chemist this is really peptide synthesis rather than protein synthesis, but when the two chains



at the side chain (*horizontal bars*), are then coupled stepwise. When the chain is complete, it is cleaved from the bead, and most protecting groups are removed, by hydrogen bromide treatment. Sodium treatment removes groups protecting the cysteine and histidine side chains. Then the cysteines are changed from the sulfhydryl form to the stable S-sulfonate form in preparation for purification.



COMBINATION of the *A* and *B* chains, now purified and in the S-sulfonate form, was carried out by the method developed in China. The chains were mixed and reduced to the sulfhydryl form with thioglycolic acid. On exposure to air in an alkaline solution the sulfhydryl groups oxidized slowly and the three disulfide bonds characteristic of natural insulin were formed. By-products formed by incorrect cross-linking could be removed by extraction.

are combined, the final product meets all the usual criteria for a protein and justifies the conclusion that a real protein has been synthesized in the laboratory.

The first published synthesis of an individual insulin chain was made by P. G. Katsoyannis and his colleagues at the University of Pittsburgh School of Medicine in 1963. They made the 21-residue *A* chain of sheep insulin by the fragment method. When the synthetic *A* chain was linked with the natural *B* chain, the combination gave rise to a small but definite amount of insulin activity. Later that same year a large group of chemists at the Technische Hochschule at Aachen in Germany, under the direction of Helmut Zahn, reported the synthesis of both the *A* and the *B* chain and the successful combination of the two for the first total synthesis of insulin. The overall yields (2.9 percent for the *A* chain and 7 percent for the *B*) and the extent of the combination (.2 to 1 percent) were still low, but true insulin activity was obtained. These syntheses, which also followed the fragment approach, required 89 reaction steps for the *A* chain, 132 steps for the *B* chain and three more steps for the combination of the two. Each step, of course, required numerous operations.

During the same period a third group was working on insulin at the Academy of Science in Shanghai and the University of Peking. Their first important contribution to the problem was the development of improved methods for the separation and recombination of natural insulin chains. Their yield was eventually increased to about 50 percent, which meant that the combination was far from a random process. Their major contribution was the preparation in 1965 of the first crystalline, all-synthetic insulin. The crystals were obtained in low yield, but they had the same form as the native molecule and, most important, the full biological activity (more than 20 units per milligram). This was a crucial element of the proof that insulin had in fact been synthesized.

Automatic Synthesis of Insulin

There remained a very real problem. Large numbers of chemists had to work for several years to produce tiny quantities of the peptides. In order to produce useful amounts of insulin and to be able to make modifications in the structure in a more efficient way, we undertook in 1965 to apply the solid phase method to the task. The results have been very encouraging. Although more than 5,000

separate operations were required to assemble the 51 amino acids into the two chains of bovine insulin, most of these were performed automatically under the control of the drum programmer, so that it was possible for one man to carry out the synthesis of both chains in only a few days.

Beginning with three grams of resin, Arnold Marglin of Rockefeller University was able to prepare approximately two grams of protected *A* chain. A total of eight grams of the protected *B* chain was made on eight grams of the resin. The reaction that detached the peptide chains from their polystyrene support also removed most of the side-chain protecting groups, leaving only the benzyl groups on the cysteine and histidine side chains. These could be removed by reduction with metallic sodium dissolved in liquid ammonia, a reaction that was discovered many years ago by Vincent du Vigneaud of the Cornell University Medical College and was the key to his historic syntheses of the pituitary hormones oxytocin and vasopressin. Applied to insulin, however, the sodium treatment at first broke some of the bonds between the amino acids threonine and proline in the *B* chain. Once we recognized what was happening it was possible to keep the chains from splitting by careful modification of the conditions of the reaction.

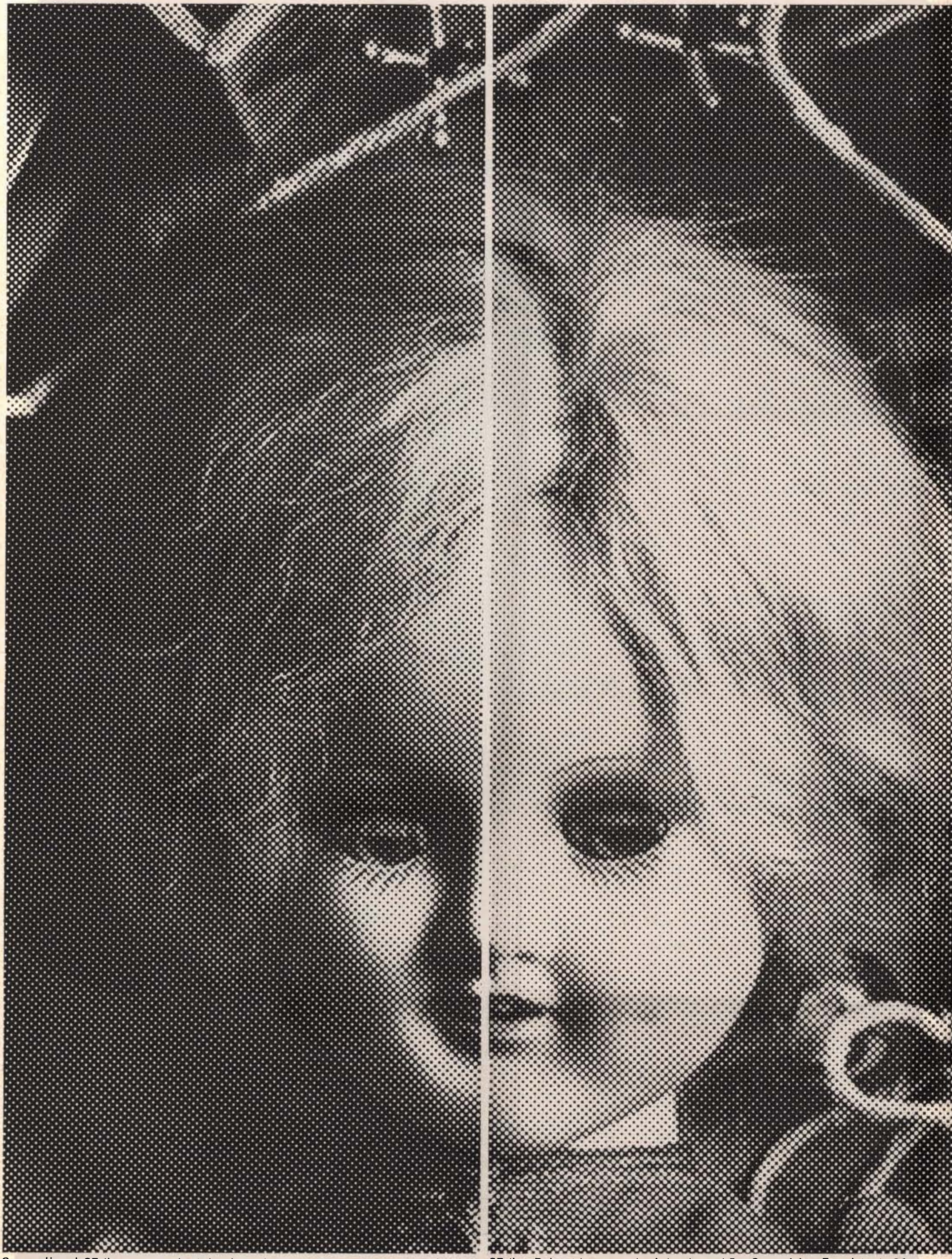
This deprotecting step left the cysteine groups in the reduced (SH) form. Although it was just this SH form that we would later want for the final oxidation step to link the two chains, the SH groups were too unstable to undergo the purification procedures that were now necessary; they were stabilized by conversion to S-sulfonates (SSO_3^-). Then the two peptide chains could be purified by three methods: filtration, which depends on molecular size; countercurrent distribution, which depends on differential solubility, and free-flow electrophoresis, which depends on electric charge. The resulting products could be shown to be homogeneous by other electrophoretic and chromatographic criteria. Amino acid analyses showed that the chains had the compositions characteristic of the *A* and *B* chains of insulin.

The final step in the synthesis of insulin was the combination of the two purified chains. First the cysteine sulfonates were converted back to the SH form. Then the chains were combined by the method developed by the Chinese, which involves the slow oxidation by air of the SH forms of both chains [*see illustration on page 72*]. We were able to show that either of our synthetic chains could be combined with the complementary natural chain to produce biologically active, semisynthetic insulin. Then the two synthetic chains were combined to form all-synthetic insulin. The synthetic hormone was active in the standard biological assay, which is based on the amount that must be injected to lower the blood sugar enough to cause convulsions in 50 percent of a group of experimental mice. The response to the synthetic preparations was shown to be due to low blood sugar rather than to some nonspecific toxic effect because the animals recovered rapidly after they were given glucose. In addition, the synthetic material behaved like insulin in various physical and chemical tests. For example, its mobility in paper electrophoresis was the same as that of the natural hormone.

The yield and the purity of the chains themselves, which constitute the peptide synthesis portion of the work, are quite good, but our combination yields are still poor. Important progress in that regard has been made recently by Katsoyannis and his co-workers. They have modified one of the earlier Chinese methods (using the SH form of the *A* chain and the SSO_3^- form of the *B* chain) and can now obtain much improved yields in the final step of the synthesis. It is fair to conclude that the synthesis of one protein has been accomplished and that another major hurdle has been cleared by chemists in their continuing struggle to duplicate nature.

Current developments in peptide chemistry should bring within reach other small protein molecules that are of great biological interest. The structures of myoglobin, cytochrome *c*, ferredoxin and growth hormone are known, for example, and we can expect some of them to be synthesized. A major step toward the synthesis of living systems will come with the synthesis of virus-coat protein; that of the tobacco mosaic virus may be the first.

It is the enzymes that are probably of greatest current interest. It is important to learn how these complex protein molecules function in their control of biochemical reactions. Why are they such enormously active catalysts and why are they so specific in their action? What factors are responsible for the "active centers" of enzymes and for their specific binding sites? How does the primary structure of the protein control its three-dimensional structure and its function; how will changes in the amino acid sequence influence biological activity? Automatic solid phase syntheses should help to answer some of these questions.



Conventional 65-line screened print enlarged 3x. Cost of 4" x 5" original: $4.00; time: 3 hours.

65-line Polaroid screened print enlarged 3x. Cost of 4" x 5" original: 60c; time: 6 minutes.

Bilingualism and Information Processing

A person who can speak two languages has clearly mastered two sets of symbols. Experiments that cause the two sets to interact provide important clues to how the mind works

by Paul A. Kolers

Is the human mind too complex to be a profitable object of study? Many investigators have felt that it is, and yet one approach to it has always seemed promising. One of the principal activities of the human mind is the manipulation of symbols; might not an investigation of the way people use symbols yield some insights into the workings of the mind?

If so, a person who can speak two languages with reasonable fluency is of particular interest, because he works with two distinct sets of symbols. By presenting a bilingual subject with information in one language and then testing him in the other, the investigator should be able to learn much about the mental operations involved in the acquisition, storage and retrieval of the information. This has been the objective of experiments my colleagues and I have conducted with bilingual subjects in the Research Laboratory of Electronics at the Massachusetts Institute of Technology and in the Center for Cognitive Studies at Harvard University.

At the outset a qualification is in order. The experiments were concerned only with words, whereas the mind also receives and manipulates information in many other forms. One can remember the appearance of an object, the tonal quality of a musical instrument, the texture of a surface or the smell of a flower without being able to describe them precisely in words. The reader can remind himself of this fact by trying to find words for the smell of a rose. Nonetheless, much of a human being's thinking is expressed in words; they are clearly his principal means of receiving, storing, manipulating and transmitting information. The question of how words are involved in these processes is now the subject of intensive inquiry.

Let me proceed to our own work with an anecdote. Once when I was visiting Belgrade I set out with a colleague to buy a certain kind of decorated shoe—a part of the national costume of Yugoslavia—that had caught his eye. We tried several shops, where, with a combination of German, French, guidebook Serbian and gestures, he tried to get what he wanted. Finally we found a shop that had the shoes, but not in the right size.

As we started to leave, two other men came into the store speaking Italian. My friend listened and then said in Spanish, "They don't have that size; I just asked." One of the newcomers said, "Why do you speak Spanish to us? We're speaking Italian." "I know," said my friend in Spanish, "but I don't speak Italian. Can't you understand me? I understand you." "Well then," said the other man in French, "it is not so good. How is your French?" My friend answered in French, "I don't understand why you don't understand Spanish when you know Italian. My French is poor. Do you speak German?" "But yes, all right, let us speak German. Where do you come from?" My friend replied in German, "The United States. And where are you from?" The reply—in English—was "We're from New York," and everyone laughed. The entire exchange, involving the use of five languages, lasted for less than a minute.

I tell this story not only because it illustrates a number of aspects of the skilled use of languages but also because it was in thinking about the implications of the episode that I became interested in bilingualism. One point the story makes about the skilled user of two or more languages is that he can switch readily from one language to another. A second point is that the changeover is usually total: the people in the episode did not speak a mixture of Italian, Spanish, French, German and English; they spoke one or another exclusively.

Let us consider what such switching entails. In some languages, such as English and French, the meaning of a sentence is strongly dependent on the sequence of the words. The point is well made by the contrast between "The dog bit the man" and "The man bit the dog." The individual words are identical; the meanings are not. In other languages, such as German and Latin, meaning is less dependent on word order because the subjects and objects of sentences are indicated by case endings and the declension of articles. The difference in German between "Der Hund biss den Mann" and "Den Mann biss der Hund" is more one of emphasis than of meaning. Even though the order of words is different, both sentences translate as "The dog bit the man," although the second sentence might be taken to indicate a particular man.

There are of course many rules that characterize the use of a language. The body of rules is the grammar of the language; the individual words are its lexicon. The two men speaking in the Belgrade store did at least three things when they switched languages. They selected words from five different lexicons. They used words in different order, that is, they used different grammatical rules to generate meaningful sequences of words. They also made sounds in different ways, that is, they used a German accent for German, a French accent for French and so on. Moreover, although they were performing a complicated psychological task in switching among linguistic codes, they did not really have to think about the process.

One of our experiments was aimed at assessing the psychological cost of such code-switching. We were interested not only in the mental processes of a bilingual person when he hears or reads either of his languages but also in what is involved when he speaks or writes either of them. Our approach made use of passages of connected discourse, some of which violated normal grammatical rules. In one session bilingual subjects read such passages silently and then were tested for comprehension of what they had read. In a second session they read the passages aloud.

Four abbreviated passages are shown in the illustration on the next page. Two of them are wholly unilingual—one in English and one in French. The other two are mixed; both are made up of some English words and some French ones, but in the first the word order is English and in the second it is French. All four passages convey the same message.

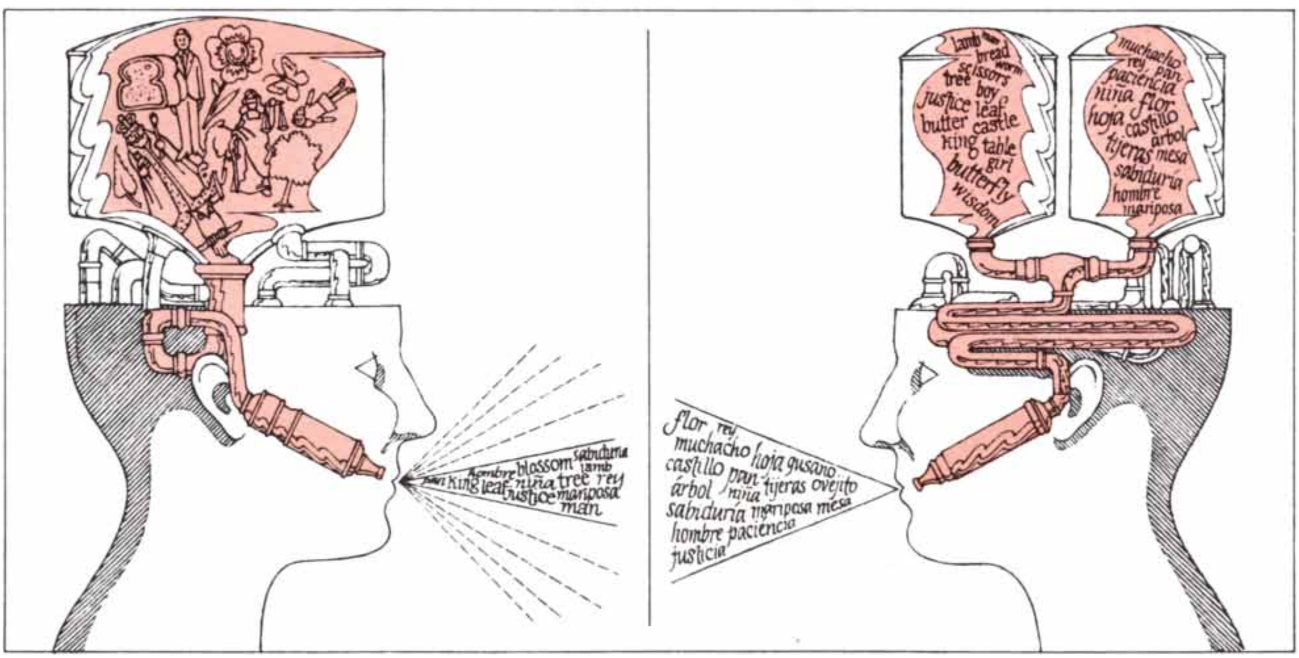
Before testing our subjects we had established how much time other subjects needed to read unilingual passages of the same length as the experimental passages and get a score of 75 percent correct on a comprehension test. Our experimental subjects were then asked to read the various unilingual and mixed passages in exactly that length of time. One might think that in order to understand a mixed passage a subject would have to translate all the words into one language or to switch between linguistic codes in some other way. If so, one might expect that the subjects would be so busy translating and switching that they would have less time to consider the meaning of the passage. Hence they would get a lower score on a comprehension test of a mixed passage than on one of a unilingual passage.

Our findings, however, were that the subjects had almost identical scores on comprehension tests following the silent reading of unilingual and mixed passages. I concluded that a skilled reader of two languages can—in reading silently—comprehend a passage readily no matter to what extent words from either language are mixed in the passage. He apparently does not have to do any switching between linguistic codes when the passages are read. (We have not yet done the experiment to test if the same ease of comprehension is evident when a bilingual person listens to a message in which words from his two languages are mixed.)

The results were markedly different when we had our subjects read various passages aloud instead of silently. They needed more time to read the mixed passages than to read the unilingual ones. Evidently reading aloud entailed some kind of code-switching between languages; the reader could not move as smoothly through "his horse, followed de deux bassets" as he could through "his horse, followed by two hounds."

We had constructed the passages in such a way that the unilingual ones contained an average of 110 words of English or French. The mixed passages contained 55 words from each language. We therefore were in a position to measure the amount of time required for code-switching by seeing how long it took a subject to read one passage in English and one in French and then subtracting the average of those times from the amount of time it took to read a mixed passage. Dividing the difference by the number of linguistic transitions in a mixed passage—the number of times a switch occurred between English and French—we determined that the average amount of time required for each switch in code was a third of a second. That is, it took a subject a third of a second longer, on the average, to read something like "his horse, followed de deux bassets" than to read "his horse, followed by two hounds."

Doubtless some of the difference is attributable to mechanical effects: the subject must physically adjust his vocal apparatus in switching from the sounds of one language to the sounds of another. We are not sure how much of the difference is due to such adjustments, but from control experiments involving code-switching in English alone we have concluded that a significant portion of the



TWO HYPOTHESES about the way a bilingual person handles information are represented by two arrangements of tanks. One hypothesis (*left*) is that all his information is stored centrally, or in one tank, and that he has access to it equally with both languages, which are represented by the various taps. The other (*right*) is that his information is stored in linguistically associated ways, or in separate tanks. Experiments by the author indicated that the actual situation of a bilingual person combines parts of both hypotheses.

His horse, followed by two hounds, made the earth resound under its even tread. Drops of ice stuck to his cloak. A strong wind was blowing. One side of the horizon lighted up, and in the whiteness of the early morning light, he saw rabbits hopping at the edge of their burrows.

Son cheval, suivi de deux bassets, en marchant d'un pas égal faisait résonner la terre. Des gouttes de verglas se collaient à son manteau. Une brise violente soufflait. Un côté de l'horizon s'éclaircit; et, dans la blancheur du crépuscule, il aperçut des lapins sautillant au bord de leurs terriers.

His horse, followed de deux bassets, faisait la terre résonner under its even tread. Des gouttes de verglas stuck to his manteau. Une violente brise was blowing. One side de l'horizon lighted up, and dans la blancheur of the early morning light, il aperçut rabbits hopping at the bord de leurs terriers.

Son cheval, suivi by two hounds, en marchant d'un pas égal, made resound the earth. Drops of ice se collaient à son cloak. A wind strong soufflait. Un côté of the horizon s'éclaircit; et, in the whiteness du crépuscule, he saw des lapins sautillant au edge of their burrows.

SIMILAR PASSAGES of connected discourse were used in experiments with bilingual persons. The passages present the same message in four ways: unilingually in English and French and bilingually in mixed form, one favoring English word order and the other French. Subjects reading mixed passages silently lost no time switching between languages, whereas in reading aloud they took longer with mixed passages than with unilingual ones.

code-switching interval is occupied by a mental operation. The operation can be described as a "call time," meaning the amount of time the mind needs to organize a set of procedures for handling a piece of information. The length of call time probably varies from person to person. It may vary also with the procedure being called. In reading a science textbook, for example, one sees words, pictures, formulas and other kinds of symbol and uses different procedures for each of them. The length of time required to call the appropriate procedures may also differ.

The experiments with mixed passages involved the important matter of context. Clearly the fact that each of the passages had a context—a thematic continuity—made it easier for the subjects to comprehend the passages. In some instances context is created by the system of symbols itself, as when a writer uses words to tell the reader what topic he is discussing. Other systems of symbols are different. Computer programmers and engineers, for example, cannot usually understand each other's programs or circuit drawings until they are told separately what the program or the drawing is designed to do—what its context is. (How subtle one's dependence on context can be is illustrated by a recent newspaper story that described the bewilderment of a foreign visitor to New York when he saw a sign saying, "BUS STOP. NO STANDING." Lacking the context that would be familiar to any New Yorker driving a car, he at first took the sign to mean that he was supposed to sit down while waiting for the bus.)

Words and other symbols, however, are not always embedded in a context. I wanted to investigate how the mind dealt with words that were isolated from context. To that end I undertook two other experiments. Before I describe them I need to supply some context.

Many bilingual people say that they think differently and respond with different emotions to the same experience in their two languages. For example, reading a poem or a play in French and reading its translation in English are said to create markedly different feelings and impressions. It is difficult to assess these introspective statements, if only because emotive texts are notoriously difficult to translate well. As Robert Frost once remarked, when a poem is translated, the poetry is often lost.

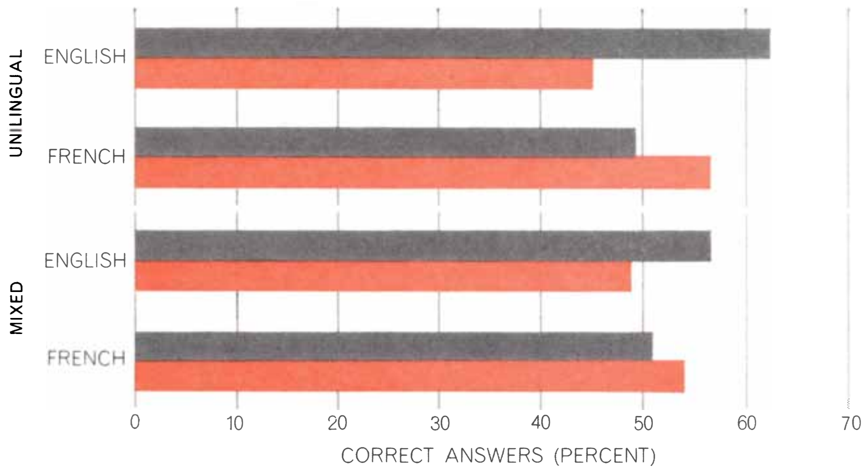
Nonetheless, if one accepts the premise that such statements reflect a genuine mental experience, one wonders about its nature. In particular we wondered whether the difference in impression arises from the difficulty of translating words accurately or from some overall property of languages and the contexts in which they are used. To put the question another way, we wondered how verbal symbols are stored in the mind.

Perhaps a metaphor will help to clarify the issue. Regard the mind as a storage tank and languages as taps. Is all the information that words represent stored in some central tank in the mind, so that if a person is bilingual he has access to the same information even though he is using two different taps? If so, one could expect a variety of taps: some could be large and some small; some could release the contents of the tank as a spray and some as a stream. That is, the taps might be regarded as the rules of grammar that affect the translation of information in the mind into sentences. The information being tapped would always be the same, but its appearance and form would differ (according to the grammar being used) in such characteristics as word order, tense agreement and the like.

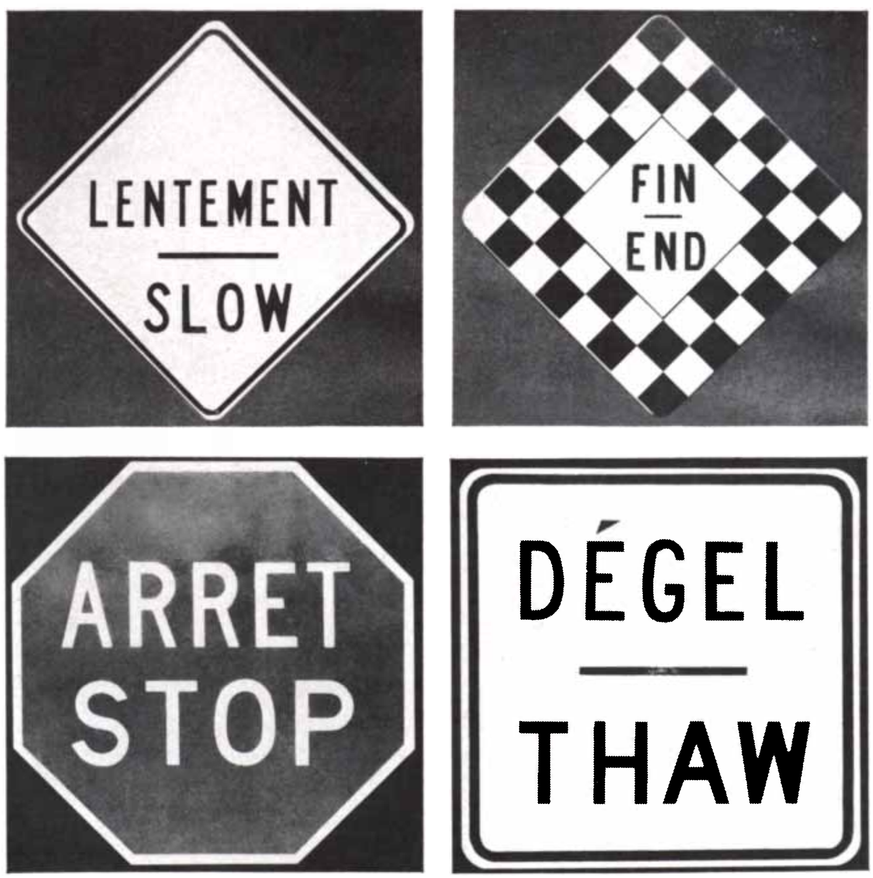
Another possibility is that the information in the mind of a bilingual person depends fundamentally on the language that was used to put it there. To continue the metaphor, such a person would have two tanks in his mind, each with its own tap. The tanks would reflect a situation in which the rules for using a language are indelibly stamped on the information stored, so that the bilingual person has access to different information when he uses the different taps.

The first of the two alternatives can be described as common storage of information. The second entails separate storage [*see illustration on page 79*]. The alternatives define two extreme ways of characterizing the issue. If common storage were the case, the differences in reading a poem in two languages would be due entirely to the difficulty of translation. If separate storage were the case, the difference would be due to other kinds of experience. The fact is, as I shall show with a description of the experiments, that neither extreme alternative correctly describes the mental storage of information. A third arrangement that combines features of the other two seems to be required.

The method I chose for examining the extreme alternatives was a word-association test in which the subject is required to say the first word that comes to mind in response to a stimulus word. For example, a large percentage of English-speaking adults respond to "table" with "chair" and to "black" with "white." My subjects were students whose native languages were German, Spanish or Thai but who were also fluent in English. In my tests the subjects responded in their native language to a list of words in that language; they responded in English to the same list in English, and they responded in one language to stimulus words presented in the other. A typical selection of words is



WRITTEN TESTS of comprehension of unilingual and mixed passages produced these results for Americans (*black*) and Europeans (*color*). The scores are roughly equivalent for all conditions, although the subjects did somewhat better with texts favoring their native syntax.



BILINGUAL SIGNS, a common sight in Quebec, are indicative of a situation in which use of two languages is an ordinary matter. Quebec often uses pictorial signs as an alternative.

	ENGLISH	GERMAN	SPANISH	THAI
EVOCATIVE	man	Mann	hombre	poo chai
	table	Tisch	mesa	dto
	bread	Brot	pan	ka-nom bpung
	boy	Junge	muchacho	dek poo chai
	blossom	Blüte	flor	dauk mai barn
	girl	Mädchen	niña	dek poo ying
	butter	Butter	mantequilla	nur-ie
	scissors	Schere	tijeras	gkan gkrai
ABSTRACTIONS	freedom	Freiheit	libertad	say-ree parp
	justice	Gerechtigkeit	justicia	yoo-dti tum
	law	Gesetz	ley	gkot mai
	honor	Ehre	honor	gkee-at-dti
	patience	Geduld	paciéncia	kwam ot-ton
	wisdom	Weisheit	sabiduría	kwam raub roo
	duty	Pflicht	el deber	nah tee
	civilization	Zivilisation	civilisación	ah-ra-ya tum
THINGS	lamb	Lamm	ovejito	look gkaa
	thorn	Dorn	espina	nam
	butterfly	Schmetterling	mariposa	pee sur-ah
	worm	Wurm	gusano	naun
	smoke	Rauch	humo	kwan
	castle	Schloss	castillo	bprah-sart
	tree	Baum	árbol	dton mai
	Norway	Norwegen	Noruega	nor-way
FEELINGS	pain	Schmerz	dolor	chjep bpoo-at
	hate	Hass	odio	kwam gklee-at
	jealousy	Eifersucht	celos	heung
	fear	Furcht	miedo	kwam gklau
	love	Liebe	amor	kwam ruk
	guilt	Schuld	culpa	kwam pit
	sadness	Traurigkeit	tristeza	kwam sow
	pity	Mitleid	piedad	song sarn

LEXICON OF WORDS used in word-association tests of the kind shown in the upper illustration on the opposite page was in four categories. The first contains words that evoke similar responses when used as stimulus words; for example, most people hearing "man" will respond "woman." Other categories are self-explanatory. Thai words are presented here in transliteration, whereas subjects familiar with the language saw them in Thai alphabet.

shown in all four languages in the illustration above.

Consider the German words *Haus* (house) and *Tisch* (table). Suppose a person fluent in German and English who was taking the German-German association test responded to *Tisch* with *Haus*. Would he respond to *table* with *house* or with some other word, such as *chair*? And how would he respond when the stimulus was in one language and he was asked to react in the other?

If the hypothesis of a common store of information were correct, one would expect a large percentage of responses to be similar in all the tests, since the concepts with which the subject was dealing would be essentially the same regardless of the language he was speaking. On the other hand, if information were stored according to language, one would expect the percentage of such direct translations to be low; for example, the subject might respond to *Tisch* with *Haus* in the German-German test and to *table* with *chair* in the English-English test.

Our finding was that about a fifth of the responses were the same in a bilingual subject's two languages. That is too large a percentage to warrant the belief that the meanings of words were stored completely in linguistically separate tanks. On the other hand, the large number of responses (about a quarter of the total) confined to one language or another enabled us to reject the idea that the meanings existed in a single tank for which the languages were merely taps. The bilingual person does not have a single store of meanings in his mind that he taps with his two languages. What it comes down to is that access to the information one has in one's mind is in some cases restricted to the language by which—or, more broadly, the context in which—it was encoded.

What are these cases? An indication is provided by the different responses we received to different categories of words. Some of the words we used referred to concrete objects; examples are *lamb, thorn, tree*. Other words were more abstract: *freedom, justice, wisdom, materialism*. Still other words—*hate, jealousy, love, guilt*—referred to feelings.

Our results revealed that words referring to concrete, manipulable objects were more likely to elicit similar responses in the bilingual person's two languages than abstract words. The abstract words in turn elicited a larger number of similar responses than the words referring to feelings. To put the matter another way, *love* and *Liebe* or *democracy* and *Demokratie* do not mean the same thing to someone familiar with English and German, even though they are dictionary translations of one another. He has different contexts and different expectations for each of the two words in the pairs. In contrast, words that refer to objects that people in various countries manipulate in similar ways—objects such as pencils, books and desks—have very similar meanings in the two languages. The idea that there are operational definitions of terms, as many philosophers of science put it, seems to have some psychological reality as one basis of meaning.

Our work showed that some information can be stored in such a way that it is readily accessible in either of two languages. Other information is, in terms of its accessibility, closely bound to the language by which it was stored in the mind. In another set of experiments we explored the way in which words are stored and retrieved. The question was: Are words perceived and then stored in the memory as individual items or does the process take place in terms of their meanings?

Our experiment was based on a phenomenon first studied in detail by Nancy C. Waugh, a former colleague of mine who is now at the Harvard Medical School. She found that if a subject was presented with a unilingual list of words, some of which were repeated, his ability to recall a given word was directly proportional to the number of times it had been repeated. If a subject is shown, say, 120 words one at a time for about a second each, and if a few of the words are repeated on the list, he is twice as likely to recall a word presented four times as one presented twice.

My colleagues and I wondered what the result would be if a list were presented with some words appearing in two languages. Taking as an example the English word *fold* and its French translation *pli,* would a bilingual subject seeing each of them two or three times in a long list of words presented singly recall *fold* and *pli* according to the frequency with which each appeared or would his recall reflect the frequency of occurrence of the common meaning of the two words? An English-French list typical of the ones we used appears in the top illustration on the next page; the reader must remember that the subject saw the words one at a time and not in a complete array as in the illustration. Among the words that translate each other are *fold* and *pli* and *ten* and *dix;* among the words that are not translated are *herd* and *fonds* (funds).

The results showed that the percentage of recall increased linearly with the frequency of occurrence of meaning [*see bottom illustration on next page*]. Presenting *fold* twice and *pli* twice produces the same effect on the recall of either word as presenting either one four times. Since *fold* and *pli* neither look alike nor sound alike, it cannot be the words themselves that interact in perception and memory. Our subjects did not see and store the words individually as visual or phonetic objects; they stored them in terms of their meaning.

The implication is clear that the subjects were able to code and store verbal items in some form other than the language in which the items appeared. A further implication is that information repeated in different languages (different symbol systems) is as well retained as information repeated in a single language. The amount of information that can be retained, however, is not increased by using different symbol systems for storing it, but access to the information is increased.

To put the point more concretely, suppose one wanted to give a student two lessons in geography. If the student knew two languages, he would retain as much geography from one lesson in each language as from two lessons in one of them. Moreover, he would be able to talk about geography readily in both languages. On the other hand, teaching him geography in one language and also teaching him a second language would not necessarily enable him to express his knowledge of geography in the second language without some kind of additional instruction. The information one has and the mechanisms or rules used to acquire it are clearly separate aspects of memory.

I have so far described two aspects of the use of verbal symbol systems: the mental switching that characterizes the successful use of different languages and one of the ways language limits access to information stored in the memory. A third aspect involves the set of rules a person learns for employing a language. In some of our experiments we found that such rules affect his linguistic performance in subtle ways.

Earlier I described how our bilingual subjects switched between their languages. Such switching is not always per-

INTRALINGUAL

ENGLISH	ENGLISH
table	dish
boy	girl
king	queen
house	window

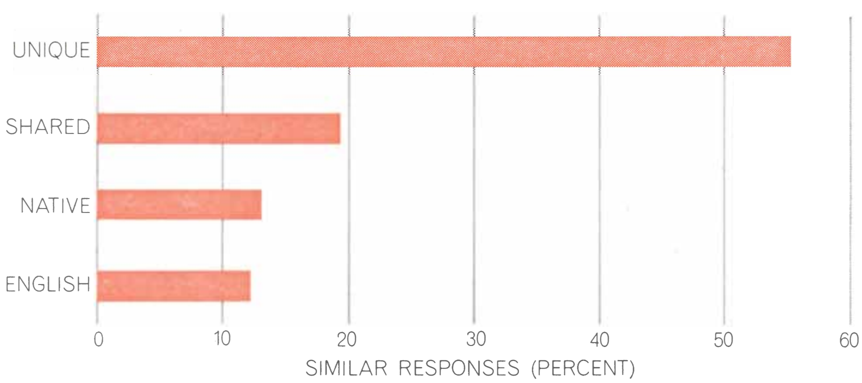
SPANISH	SPANISH
mesa	silla
muchacho	hombre
rey	reina
casa	madre

INTERLINGUAL

ENGLISH	SPANISH
table	silla
boy	niña
king	reina
house	blanco

SPANISH	ENGLISH
mesa	chair
muchacho	trousers
rey	queen
casa	mother

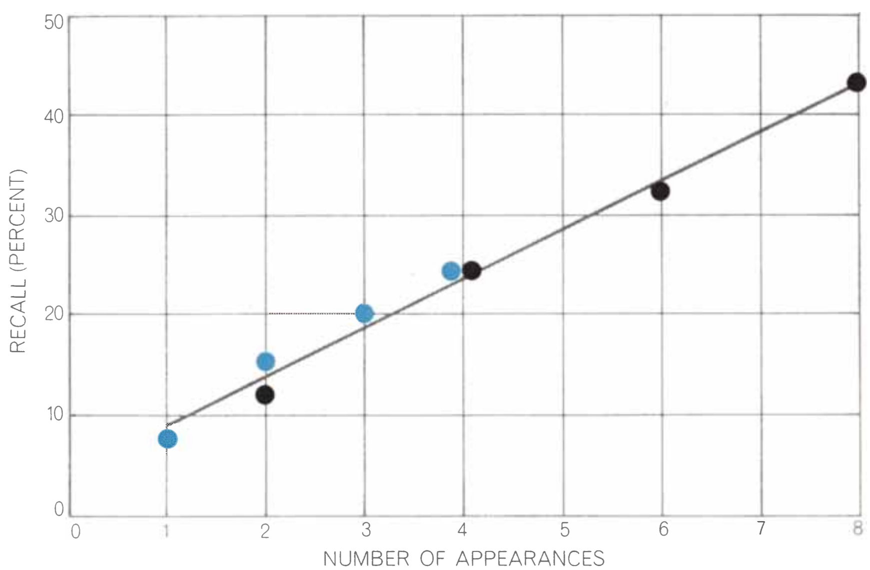
TYPICAL RESPONSES in a word-association test were given by a subject whose native language was Spanish. He was asked to respond in Spanish to Spanish stimulus words, in English to the same words in English and in each language to stimulus words in the other.



SUMMARY OF RESPONSES to interlingual word-association tests shows that more than half of the responses were unique, that is, not shared between languages. For example, in the upper illustration on this page the response *blanco* to *house* in the English-Spanish test differed from the response *window* to *house* in the English-English test. In contrast, *silla* was the response to both *table* and *mesa* and would be scored as a "native" response; answers of *reina* and *queen* to *king* are translations and are scored as "shared" responses.

ten	nerve	riz	tique
herd	truffe	isthme	preux
fold	paste	dix	fouet
soul	ice	glace	game
spout	gust	riz	clash
fonds	soul	âme	deux
jeu	truffe	fonds	game
tain	dix	crook	leaf
deux	preux	pli	bulk
pli	seing	bonne	golf
stub	ten	pli	bonne
bonne	nerf	spout	rampe
herd	rampe	golf	seing
âme	maid	two	gust
fold	maid	whip	clash
tain	jig	pâte	two
fold	truffe	psaume	whip
pli	gust	maid	âme
gust	preux	leaf	maid
bulk	cook	bonne	rampe
fouet	preux	clash	soul
fold	bulk	leaf	tain
riz	ice	glace	deux
riz	jeu	leaf	golf
two	juge	whip	fouet

STORAGE OF WORDS was tested with lists in which a subject saw, one at a time, words in both his languages. Some were repeated in the same language, others in both languages by means of translations; for example, *ten* and *dix* translate each other. Recall is improved by repetition. The question was whether recall of translated words would reflect the frequency with which their common meaning appeared or only the frequency with which the words themselves appeared. The results showed that words are stored in terms of their meanings.



PERCENT OF RECALL was essentially the same for words repeated in one language (*colored circles*) and for words repeated as translations (*black circles*). Since most of the translations, such as *snow* and *neige*, do not resemble one another, the results show that the concept is the decisive factor in recall and that two languages increase access to concepts.

fect, particularly in the daily use of language. Linguists use the word calque, which is the French word for "imitation," to describe the interference of one linguistic system with another. Examples of calques appear in the semi-Germanic sentences "Throw the baby out the window a bottle" and "Throw mama from the train a kiss." I have heard Hebrew-speaking people (Israeli students in the U.S.) inadvertently say "spoontea" for "teaspoon" and "cuptea" for "teacup." (In Hebrew the adjective always follows the noun.) Once I heard such a student say "washdisher" for "dishwasher." The last example is of particular linguistic interest because the Hebrew for "dishwasher" translates literally as "washer-dish"; the speaker, however, combined the Hebrew word order with the English sequence of syllables.

In sum, speakers of a language develop linguistic habits, or characteristic ways of ordering words. One effect of these habits was revealed when our bilingual subjects were asked to read aloud linguistically mixed passages of the kind described at the beginning of this article [*see illustration on page 80*]. The rules we had used for constructing the passages gave rise to many cases in which the normal word order of English or French was violated. Two examples in the illustration are "made resound the earth" and "une violente brise."

Subjects reading the passages aloud sometimes said "made the earth resound" and "une brise violente." Thus they showed that their experience with the normal syntactic forms affected their way of speaking words presented visually. In effect, the students were producing calques, but in a direction opposite to the normal one. Usually a calque distorts a verbal expression by applying a syntactic form of one language to words in another. In our experiment the subjects' ingrained skills in using the rules of English and French induced them to rectify word sequences that had been distorted deliberately.

The various experiments I have described embody some significant implications for both education and the study of the mind. Education entails the acquisition of information and the use of mental skills. Languages, as I have shown, can train one's mind in the way it orders and uses information. The phenomenon of bilingualism enables us to give people information or teach people a skill in one language and find out if the information or the skill can be expressed in another language. In this way we can separate for study the mental

Freeing Coatings for Examination

Sophisticated electrolytic techniques separate metallic coatings from base steel for closer study. The result is a better understanding of coating characteristics and product applications.

by Samuel M. Purdy, Research Supervisor

Enormous amounts of steel are today being produced with metallic coatings. The Youngstown Sheet and Tube Company alone makes 700,000 tons of it each year. Most of it goes into sheet steel, which amounts to approximately 100,000,000 square feet of coated steel, not counting steel used in electroplating.

Coated steel has many uses. A common example is the so-called tin can, actually made of tin coated steel. Galvanized steel is familiar as roofing and siding for farm and industrial buildings, such as the Quonset hut. The list of uses is almost endless.

Newly created conditions demand new varieties of coated steels. For example, the high price of tin and the political instability in tin-producing countries have led to a search for safe substitutes for tin plate in food containers. Cost conscious users of galvanized steel have helped motivate the development of varieties which don't need primer coats in painting.

New technological improvements such as vapor deposition suggest new methods of applying coatings to metals – coatings heretofore impossible. A thin film of Al, Ti or stainless steel on steel now seems possible. Some of the coatings now foreseen may be extraordinarily thin.

These new coatings and processes require new techniques for characterization of coatings. What will work for tin plate will not work for vapor deposited Al or stainless steel. At the same time, the effort to improve conventional products requires a closer look at today's coatings. And that look must be made from a different viewpoint. That is, while properties like color, reflectance and corrosion resistance have been readily observed on the coated steel itself, newer studies show that more information can be obtained after the coating has been separated from the steel. This is especially true of thin films made by electro-deposition, or by the new vapor deposition techniques.

Properties to be seen through deeper scrutiny include chemical composition, grain size, micro structure, etc. But, properties that make a good film, such as tight adherence, make separation difficult.

Because the coating can't be peeled from the steel, unless the coating was faulty to begin with, and since chemically dissolving the coating defeats the purpose, the steel must be taken away from the coating. This can be done by taking advantage of chemical differences between coating and steel.

One such difference is the phenomenon called "passivation." If a metal specimen is the anode in an electrolytic cell, its corrosion rate is proportional to the current flowing through the cell. Most metals corrode faster as the voltage is increased because more current flows.

Some metals, like Fe and Cr, don't behave this way. At first, as voltage increases, current increases. But at some critical voltage, the current through the cell drops suddenly to a low value. The metal becomes passive and does not corrode rapidly.

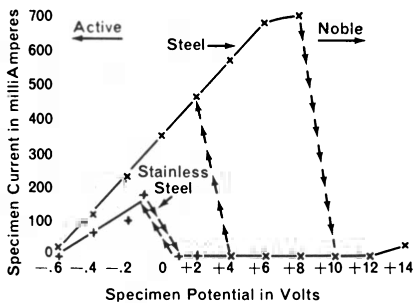


Fig. 1 Anodic Corrosion Behavior of Iron and Stainless Steel

Plots of current vs. specimen potential in 1N H_2SO_4 for steel and stainless steel are shown in the accompanying graph. Specimen potential is measured on the anode, showing accurately what is happening there without becoming entangled with cathode reactions. Steel corrodes actively, i.e., there is a large current flow at potentials up to + .2V. But above this value, up to 1.0V, there is an unstable region in which the steel becomes passive. The rate of passivation depends on how far above + .2V the specimen potential is set. The stainless steel coating shows a similar behavior, but with a different set of values well below those of steel.

Now we have the principle for a technique to separate steel from a stainless steel coating. The specimen is made the anode in an electrolytic cell containing 1N H_2SO_4. When the anode potential is set between − .1V and + .3V, steel corrodes and stainless steel becomes passive and is not attacked.

Interestingly, the technique reveals poor adhesion. Poor films flaked off readily across the specimen's face when the stainless face was electrolyzed, indicating a relatively greater amount of exposed Fe in the poor coating and implying that the coating was porous. Good specimens did not flake off, but came off in sheets. Electron microscopic examination showed that good films were free of observable porosity up to 100,000 × and that poor films had numerous pores less than ½ micron (20 micro inches) in diameter.

Coatings produced by vacuum evaporation appear extremely fine grained. Grain size for good coatings ran from 1400 down to 800 Å mean diameter and down to 500 Å in one poor one.

The technique has been extended to extremely thin (up to 1000 Å thick) coatings of electroplated Cr. The Cr plates were extremely fine grained, estimated at 100Å mean diameter or less. Diffraction patterns showed a preferred orientation in that the (110) plane was parallel to the substrate surface but that the direction varied from place to place. The Cr plate showed no obvious pores up to 100,000 ×.

The work in the development of these techniques has been only a small part of the constant research at Youngstown. If you believe Youngstown can help you, call at your convenience. Or, write Department 251D8.

Youngstown Steel

THE YOUNGSTOWN SHEET AND TUBE COMPANY · YOUNGSTOWN, OHIO 44501

IMPORTANCE OF CONTEXT in understanding a language is indicated by a foreign visitor's reaction to a sign such as this one in New York. Not being an American motorist, he thought the sign meant that he should sit on the curb while he waited for a bus.

processes used in acquiring or manipulating information from the information itself.

One example of the difference between mental skills and information is found in mathematics. Nearly all our bilingual subjects remarked during interviews that they did mathematical operations in the language in which they were taught the operations. They could always tell us the results of their operations in either language, and they could even describe what operations they had performed and how they had performed them, but the operations could be performed in only one way. Indeed, a bilingual colleague once told me that, having moved from France to the U.S. at the age of 12, he does his arithmetic in French and his calculus in English.

The point to be made is that mental activities and information learned in one context are not necessarily available for use in another. They often have to be learned anew in the second context, although perhaps with less time and effort. The fact is, however, that relatively little is known about how the activities of the mind affect one another. The study of bilingualism, being a study of the interaction of symbol systems and the way they affect one's acquisition and use of information, promises to provide valuable information on these questions.

This man's job didn't exist ten years ago.

What's he doing today at IBM?

IBM's John Goodrich is probably the loftiest systems engineer in the country. He works at the Climax Molybdenum Mine, located 12,000 feet up in the Colorado Rockies.

His job—and thousands like it—is one of many new specialties that owe their existence to the data processing industry's remarkable growth over the past ten years.

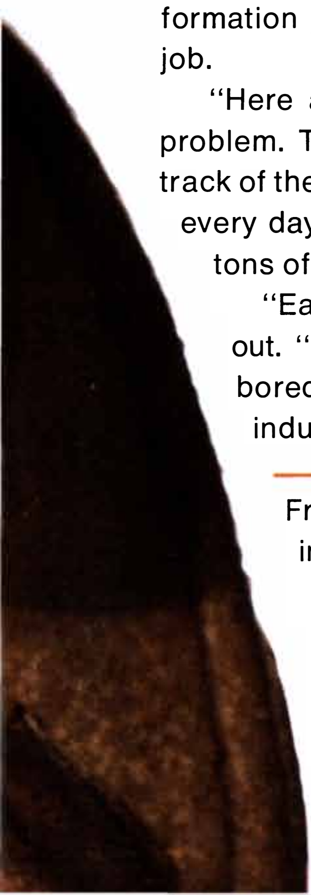
What kind of career is systems engineering? "A bit like being a doctor," Goodrich confesses. "You've got to diagnose your customer's information problem, design a remedy—then make sure that it does the job.

"Here at their mine, American Metal Climax had a big inventory problem. To solve it, we've just installed a computer system to keep track of the huge flow of materials needed to produce 43,000 tons of ore every day. It takes everything from 10 tons of blasting powder to 35 tons of steel.

"Each customer has a different problem," John Goodrich points out. "But that's what makes this an exciting career. You're never bored when it's your job to learn the ins and outs of many different industries. From molybdenum mines to you name it."

IBM®

From a beginning little more than a decade ago, the computer industry and its suppliers now employ hundreds of thousands of people. John Goodrich's career is only one example of the new job opportunities that have sprung from this growth. The future promises even greater opportunities.



CHANNELING IN CRYSTALS

The recent discovery that a beam of charged particles can pass through a crystal lattice in certain directions with surprising ease is examined for its theoretical and practical implications

by Werner Brandt

If you have ever toyed with a ball-and-stick model of a single crystal—the basic architectural unit of most solids—you may have noticed that it is very transparent in certain directions. Rotate the model slowly and you will see complex arrays of overlapping atoms give way to symmetrical patterns in which open "channels" alternate with planes or straight columns of atoms.

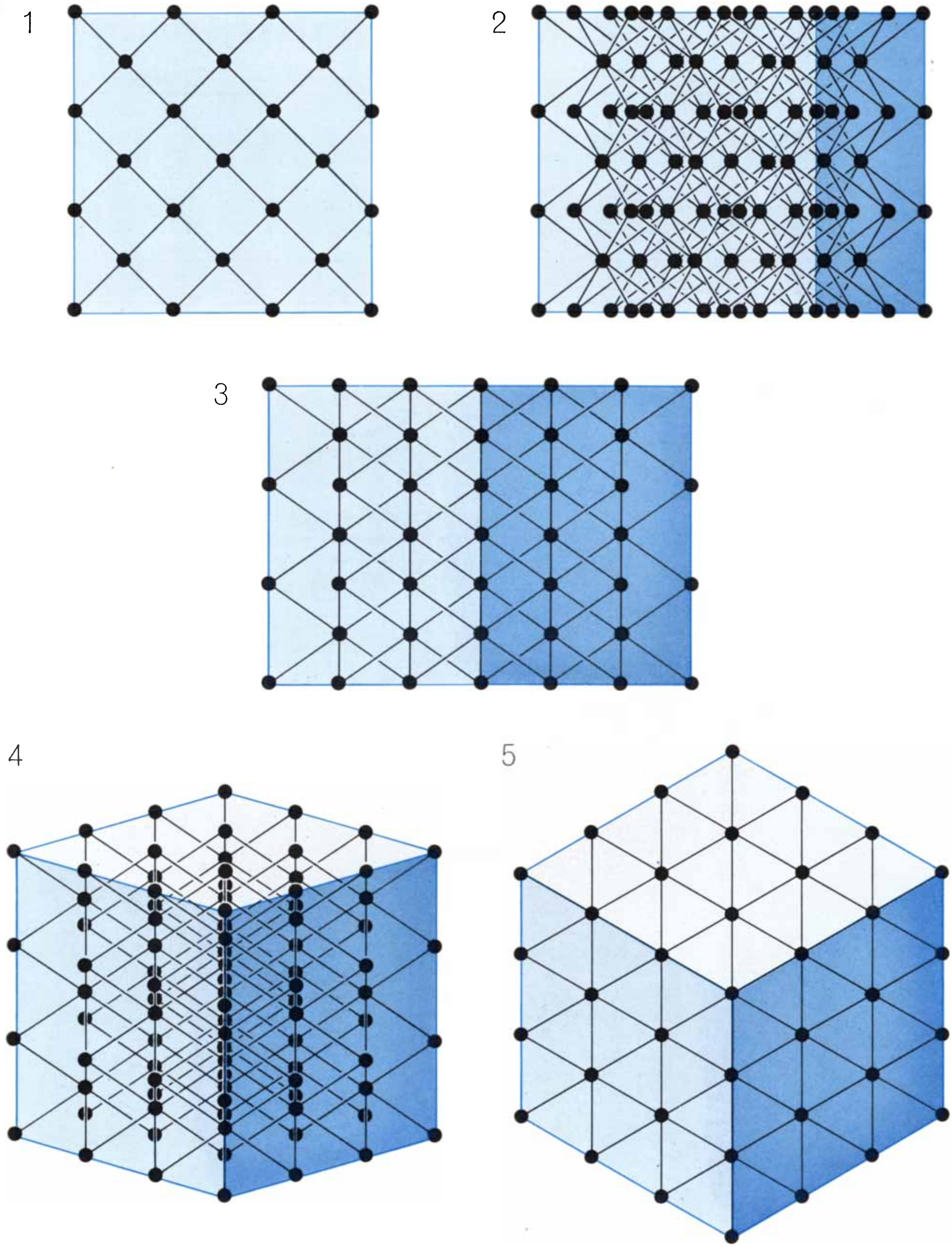
As you might expect, when a beam of ions, or charged particles, is aimed at a real crystal in the direction of these channels, the particles can penetrate some distance into the crystal without colliding with the atoms of the crystal. It has been discovered only quite recently, however, that such particles travel much farther through the crystal than simple geometric transparency can explain. In the past few years the investigation of this phenomenon, called channeling, has provided physicists with a new probe for studying the complex interactions of corpuscular radiation with matter. It is also opening the way for some important technical innovations. For example, channeling enables electrical engineers to implant a desired impurity in a semiconducting crystal by bombarding the crystal with a beam of ions, causing a minimum of radiation damage to the crystal. In this way the distribution of impurities can be controlled in ways not possible with conventional diffusion techniques.

The concept of channeling was first put forward in 1912 by Johannes Stark of Germany. In a paper on the interaction of accelerated hydrogen ions (protons) with solid targets, he took into account the existence of open channels in crystalline solids and arrived at the conclusion that ions striking such targets from certain directions had a better chance of penetrating deep into the target than they would in amorphous solids.

As it happened, Stark's prediction coincided almost exactly with the discovery of X-ray diffraction by another German physicist, Max von Laue. The link between these two approaches was recognized almost immediately by W. H. Bragg of Britain, but the experimental techniques of the day were inadequate to demonstrate the feasibility of channeling experiments. Years later Bragg's son and colleague W. L. Bragg recalled: "In the summer of 1912 my father showed me von Laue's paper, which had aroused his intense interest because of his work on the exciting of cathode rays by X rays, which pointed to the projectile-like properties of X rays, and he discussed with me possible alternative explanations for the effects which von Laue had found. I undertook some experiments at Leeds that summer to see whether we could explain von Laue's spots by shooting of particles down avenues in the crystal lattice rather than by the diffraction of waves, experiments which were of course abortive." In the wake of the triumphant progress of the X-ray-diffraction technique in elucidating the atomic structure of solid matter, Stark's idea was swept aside and forgotten.

It was not until about 1950 that techniques for studying the interactions of charged particles with solids became refined to the point where physicists began to speculate anew on the effect of crystallinity on the path of a particle moving through a solid. The current spate of experiments on channeling phenomena dates from 1963, when Mark T. Robinson and Ordean S. Oen at the Oak Ridge National Laboratory published the results of an ingenious "computer experiment" that dealt with the channeling of energetic copper atoms through a copper crystal. The crystal lattice was simulated by feeding into the computer the electric potentials of a properly spaced array of copper atoms. The computer was then instructed to calculate the trajectories of the injected energetic copper atoms as they moved through the crystal. As expected, most of the atoms (which had been started in random directions) came to rest after a few hundred collisions with the lattice atoms. In some cases, however, they kept moving as far as it was in the capacity of the computer to follow them. The story is told that the first sign of these mavericks was the alarmingly large computing bill presented by the Oak Ridge computer

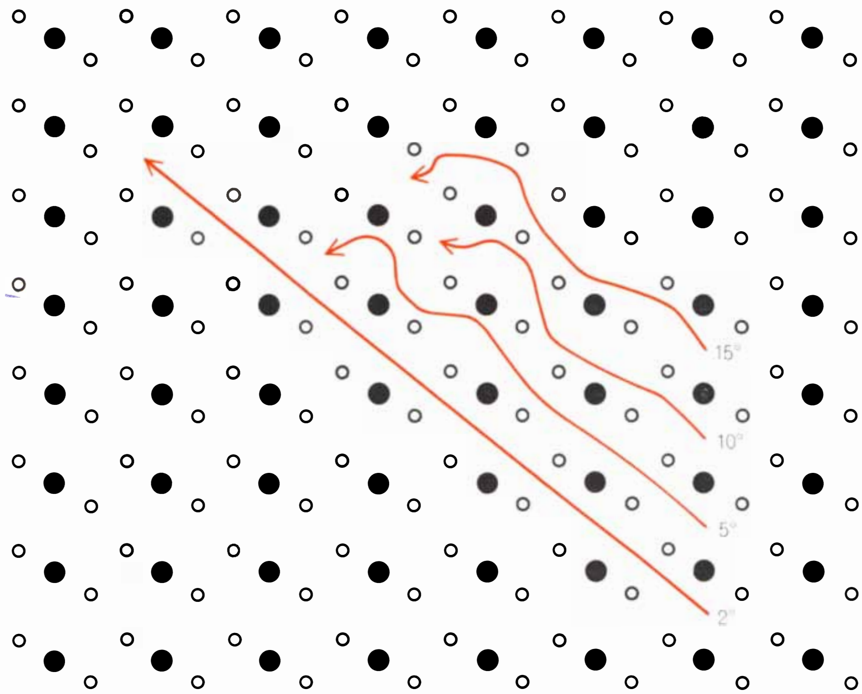
DEMONSTRATION of the effect of channeling on a beam of protons passing through a crystal of silicon is given in the color photograph on the opposite page. The proton beam was directed at the crystal along one of its open lattice directions. The energy of the beam was adjusted so that the protons that were not channeled had just enough energy left after passing through the crystal to penetrate the uppermost emulsion layer of the photographic film, turning it yellow. The protons that were channeled down the crystal planes, on the other hand, lost less energy and were stopped at an intermediate emulsion layer, producing the four-pronged red star. The bluish-black dot at the center of the star corresponds to the intersection of the open crystal planes and indicates that the protons moving exactly along this channeling direction lost the least energy and therefore penetrated to the lowest emulsion layers of the film. The demonstration is the work of Richard S. Nelson and his colleagues at the Atomic Energy Research Establishment at Harwell in England.



CHANNELING DIRECTIONS in a crystal coincide with the crystallographic planes and their intersections. These drawings show a ball-and-stick model of a body-centered-cubic crystal lattice four atoms thick in five different orientations. The best channeling directions are those with the highest degree of symmetry. Actual crystals are at least several hundred atoms thick. Hence their channels have the relative dimensions of the space between pages in a book or the inside of a long, thin pipe (*see illustration on page 94*).

center whenever certain initial directions were specified for the atoms' motion. On analyzing this curious correlation the investigators found that the atoms that penetrated farthest into the crystal were those that moved in the open channels between the dense columns of lattice atoms.

Since 1963 channeling in crystals has been the subject of intensive study. The theory of channeling, developed mainly by Jens Lindhard of the University of Aarhus in Denmark, indicates that channeling is a very general phenomenon, and a number of experiments have been performed to verify some of the detailed predictions of this theory. Most of these experiments have been done with energetic heavy particles such as protons or heavier ions. These particles are so massive that the fact that they can also behave as waves—in accordance with a celebrated requirement of the quantum theory of matter—is practically irrelevant for their behavior in crystals. Only a few experiments have been done so far with light particles such as electrons and positrons, for which quantum effects can become an important factor. Here the discussion will be confined to the channeling of energetic heavy ions, where quantum-mechanical qualifications are unnecessary.



RESULTS OF COMPUTER EXPERIMENT conducted by Lewis T. Chadderton at the University of Cambridge showed that the particles traveling farthest in a crystal channel are those that start at the smallest angle with respect to the axis of the channel. The experiment simulated the channeling of low-energy lead ions in a lead-iodide crystal lattice.

The ball-and-stick model of a crystal lattice suggests that channeling is nothing more than the straight flight of particles through the atom-free regions of the crystal. If such were the case, it is a simple problem in geometry to show that particles can penetrate the crystal only if their angle of incidence with respect to the channeling direction is less than a certain critical angle that decreases as the thickness of the crystal increases [*see top illustration on page 96*]. At angles larger than this critical angle the particles would collide with the lattice atoms and hence be scattered out of the straight channeling path. In short, the simple transparency model predicts that the critical angle for channeling through a crystal is inversely proportional to the thickness of the target and is independent of the energy of the moving particle.

Actually the apparent openness of the crystal model is deceptive. Whereas a typical model may be three or four atoms thick, the thinnest single crystal that can be prepared in the laboratory is hundreds of atoms thick. Under these circumstances the critical angle for channeling would be only about .01 degree or less. Because it is difficult to collimate and aim a particle beam to such rigorous specifications, one would expect channeling to be an extremely elusive phenomenon in a thin crystal and almost impossible to detect in a thick one.

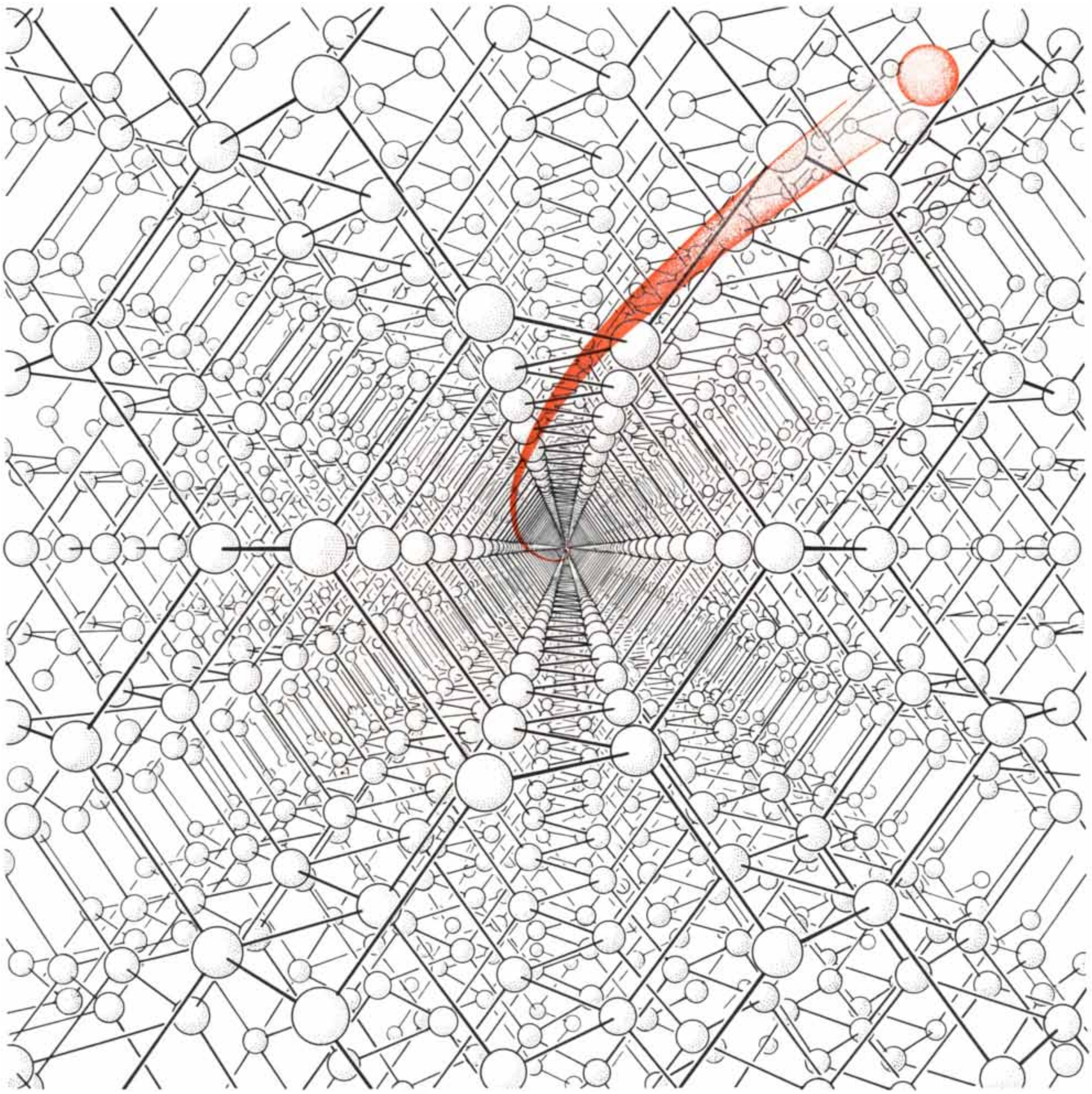
And yet channeling turns out to be a fairly rugged phenomenon. It is readily observed in both thin and thick crystals. One of the earliest experiments proving this was done in 1963 by Richard S. Nelson and Michael W. Thompson of the Atomic Energy Research Establishment at Harwell in England. They directed a beam of energetic protons at a gold foil roughly 1,000 atoms thick—just thick enough so that if the gold had not been in the form of a single crystal, it would have stopped nearly all the incoming particles [*see illustrations on page 95*]. The protons that penetrated the foil could be picked up as a transmitted current by a collecting electrode. The crystal was rotated around an axis perpendicular to its surface, causing different crystal orientations to coincide successively with the direction of the incoming beam. The experimental record of the transmitted current showed that whenever the particle beam penetrated the crystal along a direction of high symmetry (even with an angular beam width much larger than .01 degree), the particles moved freely through it and were detected by the collector electrode.

Clearly another model is necessary to explain channeling. The transparency model breaks down because particles and atoms do not act like hard spheres. Instead, when a positively charged particle approaches the nucleus of an atom in the rows that define a channel, it is repelled by electrostatic forces and bent back toward the axis of the channel. The closer the particle comes to the nucleus, the stronger the repulsion. If a particle's trajectory deviates only slightly from the channel axis, it is nudged back toward the axis, whereas if it deviates strongly, it is slammed back. In short, a charged particle moves in a crystal lattice not like a ball in a pinball machine but rather like a bobsled on its walled track. As long as the angle between the path of the sled and the axis of the track is small enough, the sled stays on the track by climbing the walls and rebounding. A particle similarly channels its way through a crystal in a wavy motion consisting of a succession of oblique ricochets.

If the angle between the particle and the axis becomes too large, the particle will leave the channel. The angle at which this occurs depends not on the length of the channel (that is, on the thickness of the crystal) but rather on the particle's energy and on the electrostatic potential of the rows of atoms. Calculations based on this model show the critical angle for channeling to be about a degree—a much more reasonable result than the one derived from the transparency model. A particle with low kinetic energy does not cross over from one channel to another, because it cannot overcome the potential wall set up by the atoms comprising a row. It is forced to make its way along one channel. A high-energy particle, on the other hand, can overcome this potential barrier and is free to roam from one channel to another as it moves forward.

The way particles move through crystals also depends on their size. If the particles are small compared with the cross section of the channel (for instance if they are protons or alpha particles), they move on an essentially flat channel "bottom," rattling back and forth like steel balls in a chute. On the other hand, if the size of the particles is comparable to the size of the channel width (for instance if they are slow copper atoms moving in a copper crystal), they swing to and fro between the channel walls like a pendulum. During each lateral oscillation the particle would move down the channel approximately 100 atomic distances.

If particles can be eased through a crystal in channels, it follows that in other directions they must be blocked by the atoms that form the walls of the channels. Furthermore, if an atom of the crystal lattice is set in motion by a collision with an incoming particle, it is always barred from flying off in the direction of the row from which it starts, because the atoms that form the row are in the way. This complementary phenomenon to channeling, called blocking, has been



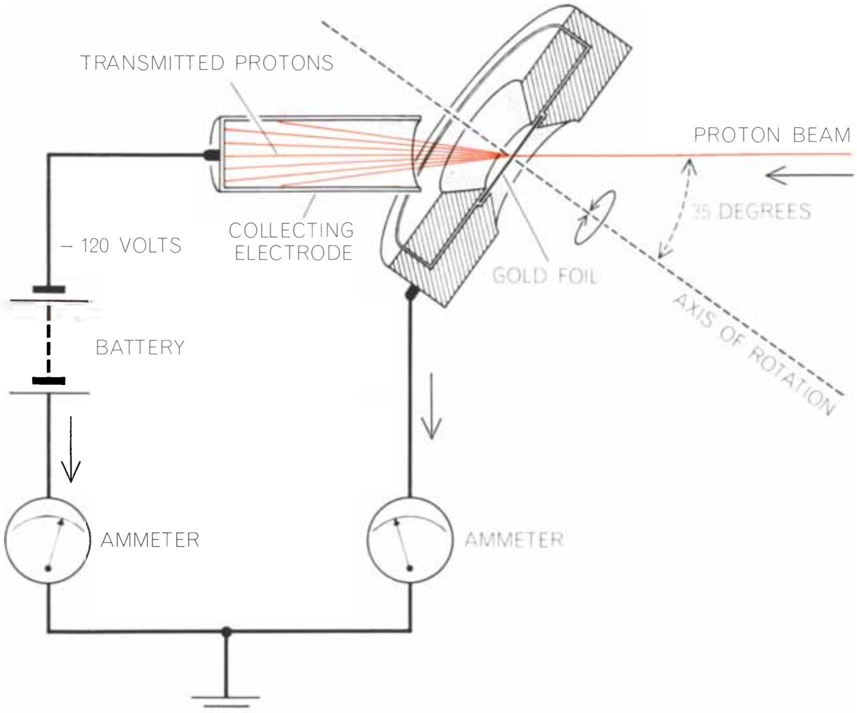
POINT-BLANK VIEW down the axis of a channel in a diamond-like crystal lattice shows the spiral path followed by a typical channeling particle. The distance covered by the particle in one turn of the spiral path is on the order of 100 interatomic distances.

used to study the structure of crystals by shooting beams of particles at the crystals and then recording on a photographic plate the pattern produced by the emerging particles [*see illustration on page 98*]. Such pictures resemble those obtained in X-ray analysis of crystal structures. Indeed, if Stark's idea had been proposed a little earlier and pursued vigorously, the channeling and blocking of charged particles might well have given the first direct experimental proof of atomic order in crystals.

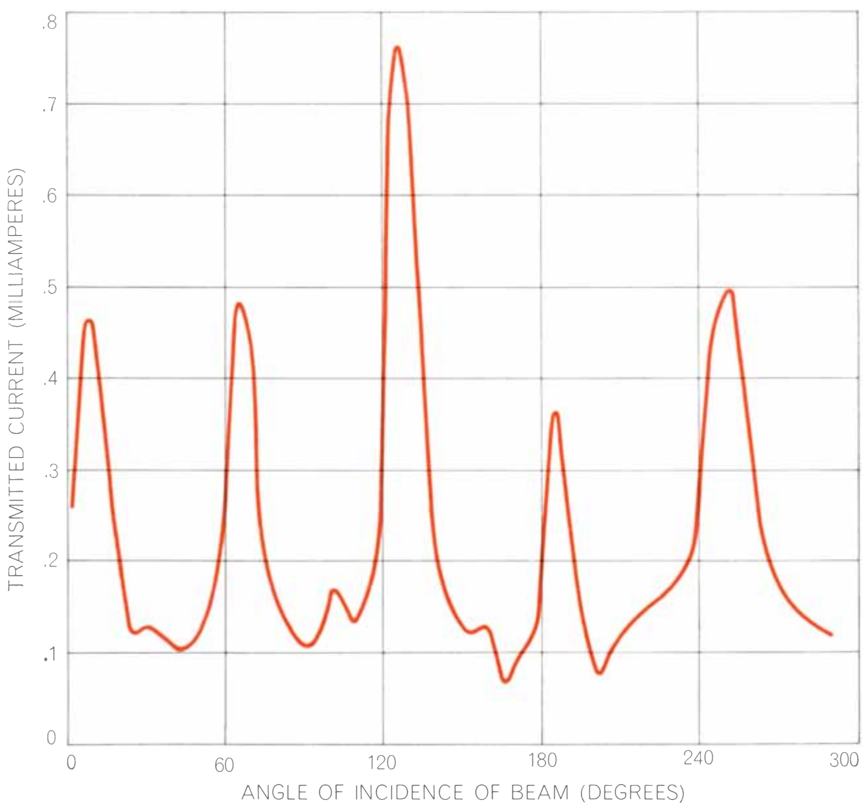
How can one test the mechanism that has just been described to explain channeling? Imagine the atomic landscape a channeling particle encounters in making its way through a crystal. Each atom is composed of a heavy nucleus located at a lattice site, with electrons arranged in shells around it. The electron density decreases sharply from the nucleus outward, and it is lowest at the midpoint between adjacent atoms. Only the outermost electron shell—the valence shell—is engaged strongly in binding the crystal together. In some instances the binding distorts the valence shell so completely that the valence electrons can no longer be identified with any one atom. Instead they belong to the entire crystal, and they can roam freely through it. In metals such itinerant valence electrons are called conduction electrons; they are free to carry a current under an applied voltage. The stability of a channeling trajectory depends on the depth to which the moving particles penetrate the electron-shell structure of the lattice atoms.

Consider a series of experiments in which we shoot particles into a crystal along some channeling direction with an ever increasing angle of incidence between the beam and the axes of the channels. We can then distinguish three ranges of particle-crystal interaction: (1) at angles that are small compared with the critical angle the particles encounter mainly valence electrons, (2) at intermediate angles the particles penetrate the inner electron shells of the crystal atoms in the course of their wavy motion in a channel and (3) at angles approaching the critical angle the particles can make contact with the nuclei of the crystal atoms [*see bottom illustration on next two pages*].

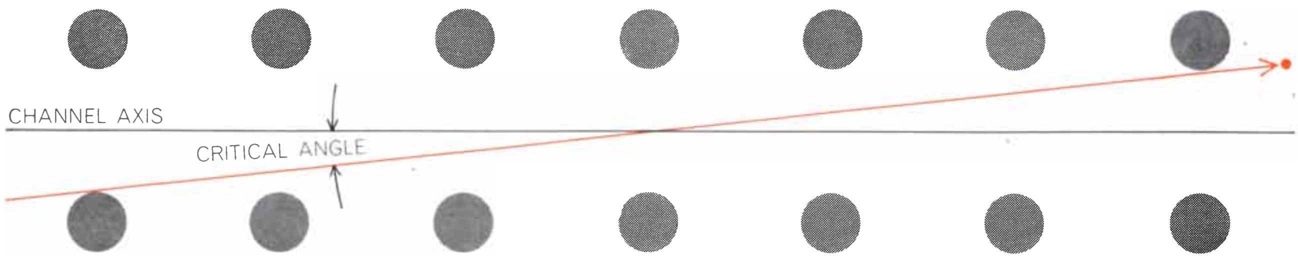
Particles that move down the center of a channel at small angles to the channeling axis plow through regions of the lowest mean electron density in the crystal, and therefore their loss of kinetic energy to excitation of the crystal elec-



EARLY CHANNELING EXPERIMENT, performed at Harwell in 1963, demonstrated that channeling is fairly easy to observe in thin crystals. A beam of protons was directed at a gold foil about 1,000 atoms thick—just thick enough so that if the gold had not been in the form of a single crystal, it would have stopped nearly all the incoming particles. Protons that penetrated the foil could be detected as a transmitted current picked up by a collecting electrode. The crystal was rotated around an axis perpendicular to its surface, causing different crystal orientations to coincide successively with the direction of the incoming beam.



RESULTS of the experiment depicted at top showed that whenever the protons penetrated the crystal along a direction of high symmetry (*peaks in curve*), most were transmitted.



TRANSPARENCY MODEL predicts that the critical angle (the largest angle between the trajectory of the particle and the axis of the channel at which channeling can occur) is inversely proportional to the length of the channel. For the thinnest crystal that can be prepared in the laboratory the critical angle would then be only about .01 degree, which would make channeling an extremely difficult phenomenon to detect even in a thin crystal. And yet channeling turns out to be easy to observe in both thin and thick crystals.

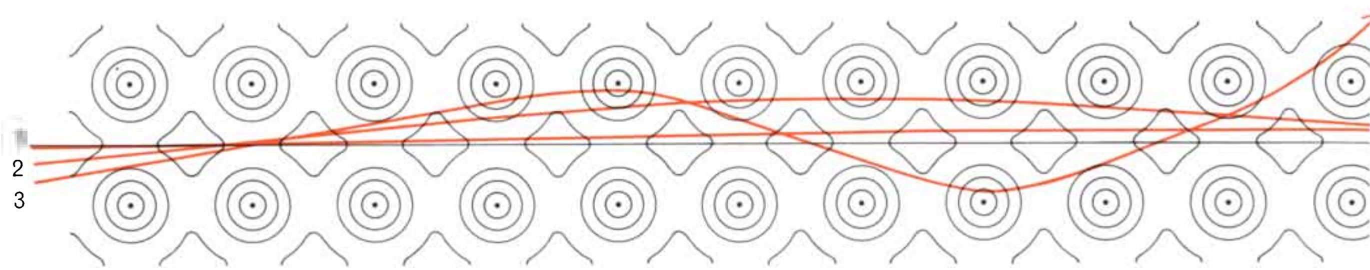
trons is at a minimum. They can cover long ranges before being stopped.

The first direct evidence of the low energy loss associated with this first range of small angles came in 1964 from experiments conducted by Geoffrey Dearnaley at Harwell. At the time Dearnaley was engaged in investigating solid-state devices for detecting particle beams. Usually such detectors consist of a crystal of silicon or germanium acting as a diode in an applied electric field. Whenever a swift charged particle penetrates the detector, an electrical pulse is counted in the associated electronic equipment. For such detectors to work reliably they must produce a pulse of a specific magnitude for each particle of the same type and energy. Dearnaley foresaw the possibility that if some particles were to channel through the detector, they would probe only the lowest electron density in the crystal and would not produce as much electronic excitation as unchanneled particles; they would give rise to anomalously small pulses open to erroneous interpretation. He directed a beam of energetic protons into very thin single crystals of silicon in an experiment much like the one performed by Nelson and Thompson. Analysis of the energies of the transmitted protons showed that they lost less energy when the beam direction coincided with an open crystal direction.

Later work by Cavid Erginsoy and his colleagues at the Brookhaven National Laboratory and researchers at other institutions confirmed quantitatively that the energy loss of channeled particles reaches a minimum when the angle of incidence approaches zero. Incidentally, now that channeling is recognized as a potential source of error in solid-state counters, it can be avoided by orienting the open crystal directions out of the acceptance angle of the counter.

If one works with a thick crystal as a target, one would expect to find the channeled particles at much longer ranges, or depths of penetration, than the rest of the particles. This effect can be demonstrated and traced in detail by shooting radioactive atoms into a crystal along the channeling directions. After irradiation the crystal is gently peeled away layer by layer, and the decrease of the radioactivity remaining in it is measured. This is not an easy task, since heavy ions often do not penetrate much deeper than from about 100 to 1,000 layers of atoms. The "profile" of the embedded ions can be resolved only if each peeling step does not remove more than, say, 10 atomic layers, or a thickness of 20 angstrom units. So far two methods have been devised to accomplish this remarkable feat.

The first method, developed by John A. Davies at the Canadian Government Research Establishment at Chalk River, Ont., uses the electrolytic oxidation of aluminum in a weakly basic solution. If the solution is maintained at a constant temperature, the thickness of the oxide layer at the surface of the aluminum crystal is dependent only on the applied voltage; the depth of the layer ranges from about 20 angstroms at a few volts to a few hundred angstroms at several hundred volts. After a layer of known thickness is formed, it is dissolved away by a judiciously selected solvent that does not attack the rest of the crystal. The radioactivity in the dissolved layer or the radioactivity remaining in the crystal, or both, are measured and then the next layer is removed. Although the electrolytic peeling method has been extended to several metals and semiconductors, its application is limited by its exacting chemical requirements.



REVISED MODEL of channeling takes into account the fact that particles and atoms do not act like hard spheres. Instead, when a positively charged particle approaches the nucleus of an atom in the crystal lattice, it is repelled by electrostatic forces and bent back toward the axis of the channel. Around each nucleus electrons are arranged in shells, with the electron density decreasing sharply from the nucleus outward and reaching a minimum at the midpoint between the atoms. Electrons in the outermost electron shell (the valence shell) interact to bind the crystal together. In channeling one can distinguish three ranges of particle-crystal interaction:

The second peeling method, developed by Rudolf Sizmann of the University of Munich, relies on the erosion of the crystal by the impact of ions with energies so low that, although they "sputter" the target material away, they do not penetrate far enough to disturb the outcome of the measurement. The sputtering method has the advantage of not being limited to selected materials or to definite temperature ranges.

Both methods reveal in detail how particles penetrate deeper along open crystal directions than in a multicrystalline or noncrystalline target of the same average density. Until very recently all channeling profiles were measured by peeling the target at room temperature. In some instances "supertails" were found in the profile, indicating that a small percentage of all channeled particles move into the crystal much farther than is consistent with any conventional theoretical estimate. Since then experimental evidence has made it plausible that the supertails occur not because of some new low-loss penetration mechanism specific to channeling but as a result of the rapid diffusion of particles in channels. This diffusion proceeds with ease at, or even somewhat below, room temperature. Supertails can be suppressed completely by keeping the target at a very low temperature during irradiation and peeling.

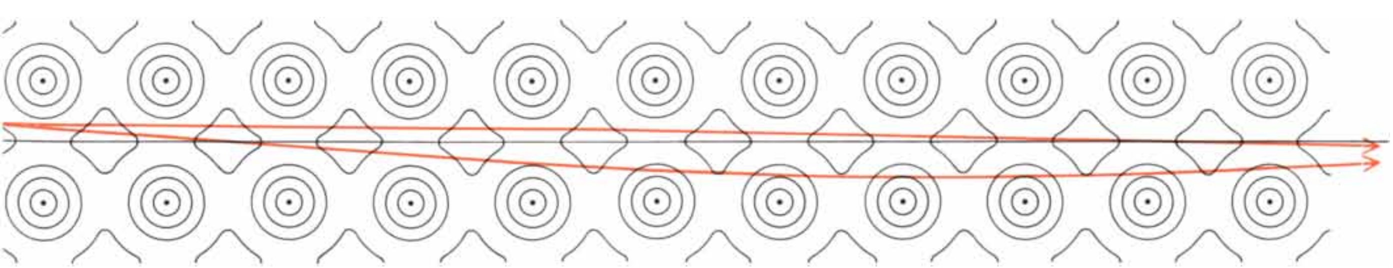
The results of various other experiments are consistent with these findings. For example, the reflection, or backscattering, of ions from a single crystal is minimal along the channel directions, because here the particles plunge into the target rather than bouncing back. Similarly, radioactive atoms deposited in a target crystal emit their decay products with equal probability in all directions, but these products emerge freely only along the channeling directions and are blocked in other directions. And of course the damage normally inflicted on a crystal by energetic particles knocking against its atoms is reduced if the particles impinge along open crystal directions.

These experiments have pointed to a number of promising practical applications of channeling. The blocking pattern produced by a beam of protons impinging on a single crystal provides one of the most direct methods for orienting crystals with high accuracy. If defects or impurities distort a channel, some of the impinging particles are backscattered and can be detected. In this way, one can analyze with good resolution the nature and distribution of crystal imperfections near the surface.

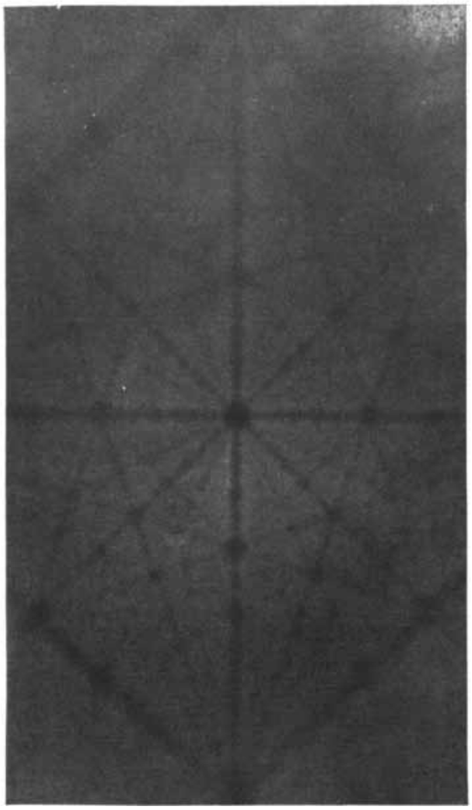
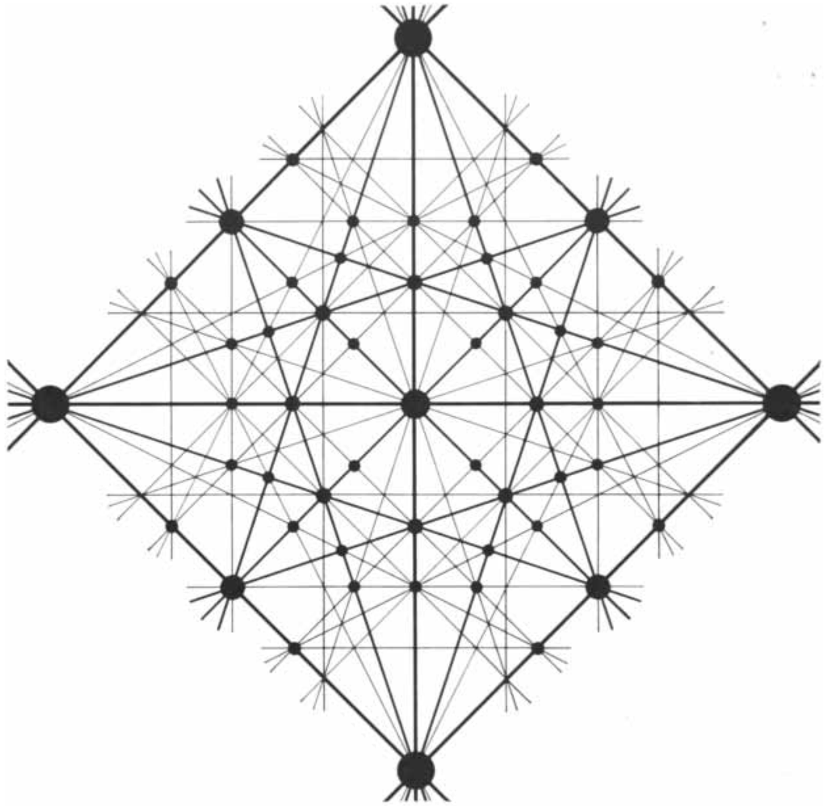
Channeling is becoming important for opening up the vast and still mysterious frontier of surface physics. As a corollary, channeling affects the technique of impurity implantation in solids by ion bombardment. Pure semiconductors must normally be "doped" with traces of other elements to endow them with desirable electronic properties. Conventional doping by diffusion is restricted by the solubility of the trace element in the semiconductor material. If the material is bombarded with energetic ions of the trace element, the ions usually stay lodged in it. With ion bombardment one can implant in the surface layers of a semiconductor whole patterns of trace elements simply by deflecting the ion beam or by shadowing and masking the surface as one does in the fabrication of microelectronic circuits. Moreover, one can combine such techniques with programmed changes in the energy and composition of the ion beam so as to dope materials in new ways by implanting the impurities with unusual depth profiles. Channeling promises to become an important variation on this theme, since ions can penetrate deep enough for this purpose along channeling directions, even at comparatively low energies, thus reducing radiation damage to the target crystal during implantation.

In the second range of incident angles—the intermediate one—the particles start piercing the atomic shell structure beyond the valence electrons as the turning points in their trajectories approach the channel walls more closely. In passing through a given shell they can knock an electron out of the shell, leaving a hole in the electronic structure of the atom. The atom reacts by rearranging electrons until an electron falls into the hole to fill the shell. In so doing the atom emits a quantum of X radiation with an energy characteristic of the shell in which the hole was created. The X ray can then be detected and its energy measured. In this way the passage of particles through different atomic shells along channeling trajectories can be traced by correlating the characteristic X-ray yield with the angle of incidence. When particles are close to the channel axes and therefore do not encounter the inner shells of target atoms, X-ray production should be at a minimum. As the angle of incidence increases, the X-ray production should increase accordingly. In fact, X-ray production should provide a versatile method for studying channeling conditions, since energetic particles excite characteristic X rays over a wide range of particle velocities.

Only 10 days after I had developed these notions at New York University, this effect was found and measured with protons in aluminum and copper crystals, thanks to the expert collaboration of Jhan Kahn and his experimental group at the University of California's Lawrence Radiation Laboratory. Along the channeling direction the X-ray yield can



(1) at angles small compared with the critical angle the particles encounter mainly valence electrons, (2) at intermediate angles the particles penetrate the inner electron shells of the atoms and (3) at angles approaching the critical angle the particles can make contact with the nuclei of the crystal atoms. Light and X rays can be produced by the particle-crystal interactions that take place in ranges 1 and 2, whereas gamma rays can be produced as a result of the nuclear reactions that take place in range 3. These radiations can be analyzed for incident particles at various energies to determine the channeling conditions at different depths in the crystal.

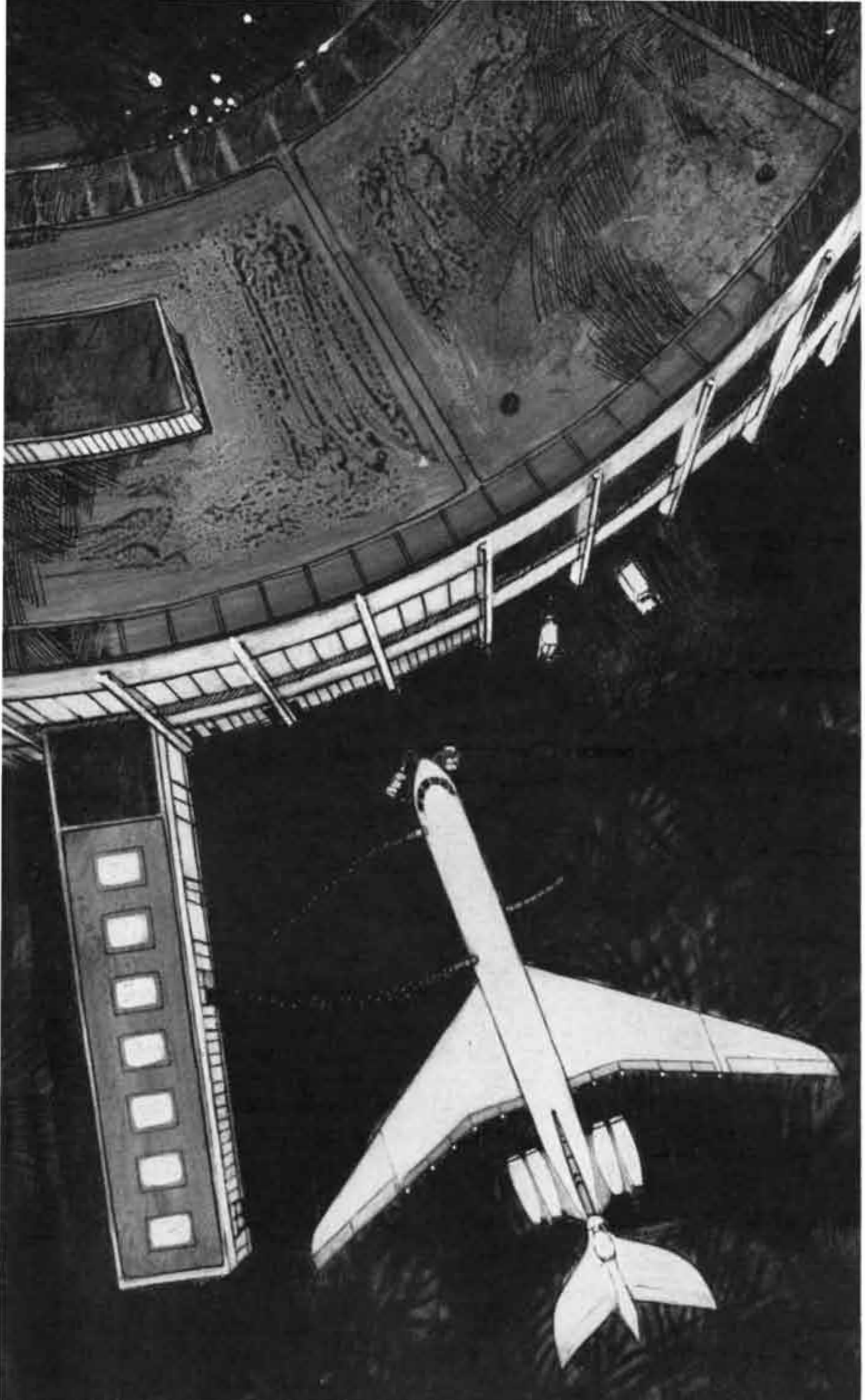
BLOCKING of charged particles by the atoms that form the walls of the channels can be used to study the structure of a crystal by recording on a photographic plate the pattern produced by the emerging particles. In the photograph at left, made by A. F. Tulinov of the Institute of Nuclear Physics of the University of Moscow, the shadows cast by the blocking of a beam of energetic protons correspond to the columns and planes of atoms in a single crystal of tungsten. In this experiment the beam entered the crystal in a nonchanneling direction and the angular distribution of the backscattered protons was recorded on a plate parallel to the surface of the crystal. The spot at center is the shadow of the columns of atoms that are at right angles to the plate. The drawing at right shows the geometric projection of a tungsten-like body-centered-cubic lattice on a plane parallel to its surface. The diameter of the dots and the thickness of the lines represent roughly the density of atoms in the corresponding columns and planes of the crystal.

drop by as much as 80 percent from the normal yield of polycrystalline targets. It rises sharply within a few degrees of the channeling direction. Since then this method has provided very detailed information on the channeling process in aluminum and copper.

As the funneling angle enters the third range—the critical one—the incoming particles begin to collide with the nuclei of the crystal atoms and nuclear reactions become possible. In particular, gamma rays can be elicited from nuclei by protons passing them at definite "resonance" energies. This was first demonstrated in 1964 by Karl Ove Nielsen and his colleagues at Aarhus. Working with protons in aluminum and silicon crystals, they observed a drastic reduction in the yield of gamma radiation from nuclear reactions along channeling directions, and a rapid increase as the angle of incidence approached the critical angle. Similar but less pronounced effects have been observed with nuclear reactions in which neutrons are emitted. An interesting aspect of the Aarhus experiment was the exploitation of the fact that the gamma rays are emitted only when the protons have a distinct resonance energy. By gradually increasing the energy of the bombarding particles above this resonance level, the channeling conditions could be studied at various depths where the protons were just coming into resonance. It turns out that as the energy is increased the gamma-ray yield in the channeling direction rises. What this means is that the protons are scattered out of the beam everywhere along the channel and in the process come close enough to nuclei to produce nuclear-resonance radiation. This attenuation of the channeled beam can be caused by imperfections in the crystal lattice or by thermal vibrations of the atoms in the lattice, which change the dimensions of the channel at random. The possibility of detecting such subtle crystal properties is drawing attention to channeling as a means of gaining new insights into the solid state.

Many sophisticated channeling experiments have now been performed with a wide range of crystals and with various particles at different energies. All these experiments support the view that channeling results from the correlated deflections of the particles in the electrostatic-force field of the orderly array of atoms in the crystals. It appears, after only a scant four years, that the exciting first phase of discovery is drawing to a close. Ahead lies the careful quantitative mapping of all possible forms of channeling and the clarification of poorly understood details.

The advent of nuclear technology has made the study of the interaction of corpuscular radiation with matter an important part of materials research. The somewhat belated appearance of channeling is certain to deepen our understanding of the solid state and of the processes that can be induced by radiation. Channeling promises to be a useful tool in many areas of physics and technology, but it is still too early to say where it will foster the most important advances.



BOAC wants confirmation in print of every passenger reservation made anywhere in the world. Right now.

Easy, when you ask Kleinschmidt.

Kleinschmidt is the language of modern telecommunications.

Home offices talk to division headquarters. Computers talk to production control. And in the case of BOAC, Kleinschmidt data printers in ticket offices provide passengers with fast, convenient flight reservations and confirmation.

Working in communications systems like BOAC's, the Kleinschmidt 311™ Data Printer "talks" with unbelievable speed and simplicity. It prints-out up to four times as fast as most other teleprinters. Operates with far fewer moving parts and far more reliability. *Outperforms* all competition.

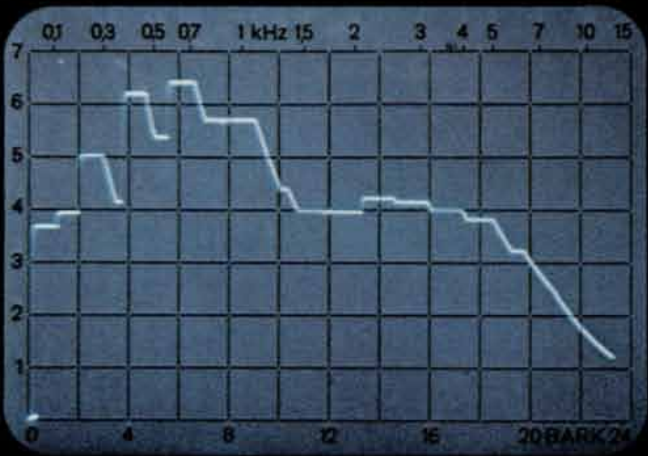
And like other Kleinschmidt data printers, the 311 is compatible with telecommunications equipment of all makes. It can fit directly into the system you now have, or the one you are having designed for you.

Of course, Kleinschmidt is a language most original systems manufacturers already understand.

But we're always glad to repeat the message.

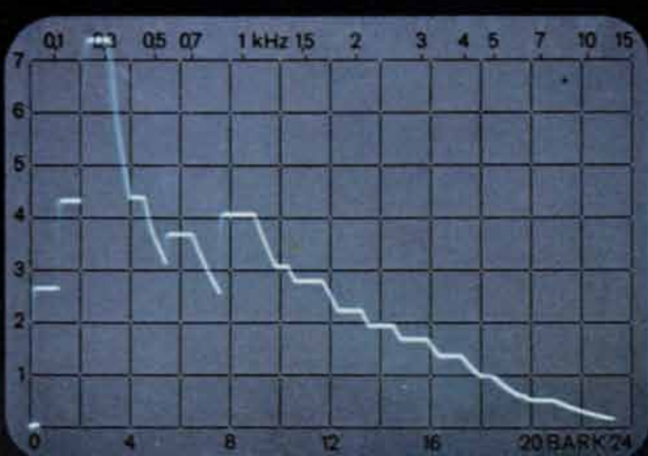
KLEINSCHMIDT
DIVISION OF SCM CORPORATION
DEERFIELD, ILLINOIS 60015 · TELEPHONE: 312 945-1000

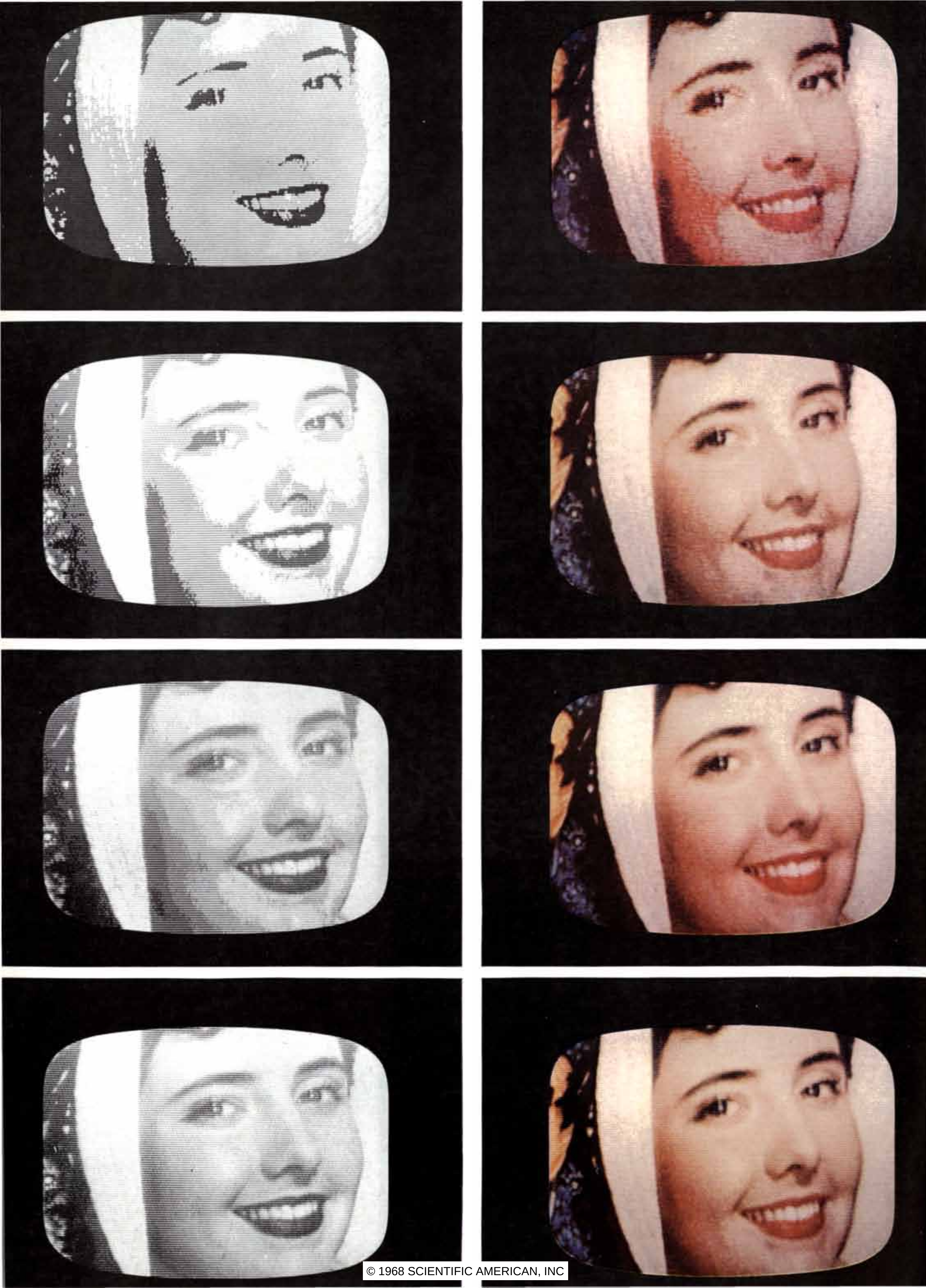
SCM®


0,1 0,3 0,5 0,7 1 kHz 1,5 2 3 4 5 7 10 15
7 6 5 4 3 2 1
0 4 8 12 16 20 BARK 24


0,1 0,3 0,5 0,7 1 kHz 1,5 2 3 4 5 7 10 15
7 6 5 4 3 2 1
0 4 8 12 16 20 BARK 24


0,1 0,3 0,5 0,7 1 kHz 1,5 2 3 4 5 7 10 15
7 6 5 4 3 2 1
0 4 8 12 16 20 BARK 24


0,1 0,5 0,7 1 kHz 1,5 2 3 4 5 7 10 15
7 6 5 4 3 2 1
0 4 8 12 16 20 BARK 24



PULSE-CODE MODULATION

The growth of communications traffic has enhanced the attractiveness of this method, which converts messages into binary digits that can then be transmitted in the form of closely spaced electrical pulses

by J. S. Mayo

This year marks the 20th anniversary of two events that have deeply transformed the theory and practice of electrical communication. The first was the publication of a paper by Claude E. Shannon titled "A Mathematical Theory of Communication"; the second was the invention of the transistor by John Bardeen, Walter H. Brattain and William Shockley. The effect of Shannon's theoretical paper on the one hand and the solid-state device on the other was to focus the attention of communication engineers on the convenience of transmitting messages in the form of closely spaced pulses rather than in the form of continuous waves that vary in amplitude or frequency. The result was the rapid development of pulse-code modulation (PCM), a concept then some 10 years old, as a rival to amplitude modulation (AM) and frequency modulation (FM). An important stimulus was a second paper written in 1948 by Shannon, in collaboration with Bernard M. Oliver and John R. Pierce, called "The Philosophy of PCM."

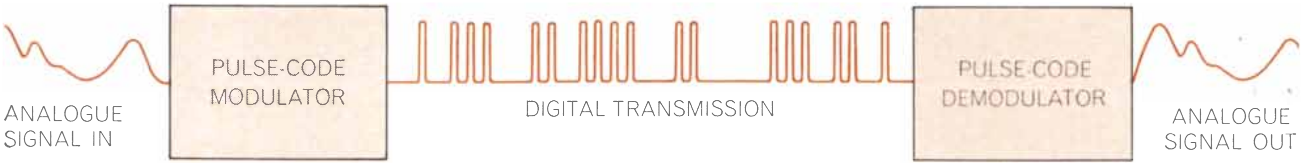
The philosophy, in brief, is that messages of various kinds—speech, television pictures, instrument readings and columns of figures—can be expeditiously transmitted through a general-purpose communication network, such as the telephone system, if they are first converted to digital form. In digital communication, information is transmitted in the form of a sequence of numbers. The numbers (called code words) are represented by electrical pulses (called pulse codes). Code words are usually expressed as binary numbers, that is, as a sequence of 0's and 1's, the same language used in modern computers. Although other number systems are sometimes used, binary digits have the advantage in that they can be expressed simply by the presence or absence of an electrical pulse. The transistor and similar devices are admirably suited to a binary communication system because they can be rapidly switched from being fully conducting to being fully nonconducting.

In spite of the rapid growth of computer networks, the great bulk of communications traffic still consists of speech and television pictures, both of which originate as analogue signals. This means that the signal information is represented by continuous fluctuations of electric voltage or current. To convert an analogue signal into a digital one a device called a pulse-code modulator is used to sample and measure the analogue signal voltage at frequent intervals and to assign each sample a numerical value. A simple mathematical theory assures us that no information is lost if the analogue signal is sampled at a rate that is at least twice the highest frequency component in the signal. This means that voice signals must be sampled and measured about 10,000 times a second and television signals about 10 million times a second.

As in every measuring operation, one must decide how precise to make each measurement. A certain amount of rounding off is inevitable. For instance, a sample value of 3.215 volts might be encoded as the binary number equivalent to 3.2. The remainder, .015 volt, is the quantizing error. In encoding speech or television pictures one must determine experimentally how many binary digits must be used in the code words so that the noise due to quantizing error cannot be detected by the listener or viewer, or at least does not add appreciably to the noise already present in the analogue signal. High-quality transmission can usually be achieved with about seven binary digits, corresponding to 128 distinct levels. It is also found when encoding speech with seven digits that there should be a nonuniform relation between the sample value and the binary number. The desired relation is approximately logarithmic. This allows faint speech to activate an appreciable number of pulses and prevents loud speech from overloading the pulse-code modulator.

When black-and-white television is transmitted by PCM, the use of too few digits per sample results in appreciable contouring [*see illustration on opposite page*]. This happens because there can only be as many shades of gray in the picture as the binary code word can specify. Thus a two-bit code can specify only four levels and a three-bit code only eight levels. However, when codes reach

QUALITY OF TELEVISION PICTURES encoded by pulse-code modulation (PCM) improves as one increases the number of digits in the binary number used to express the amplitude of a sample of the analogue signal that represents the picture. In the sequence of black-and-white and color television pictures on the opposite page the pair at top were made with two binary digits per sample; the succeeding pairs were made with three digits (*second from top*), four digits (*third from top*) and five digits (*bottom*). A two-digit code is capable of specifying four different voltage levels, which control the energy of the electron beam used to "paint" the picture; a three-digit code specifies eight levels, a four-digit code 16 levels and a five-digit code 32 levels. In each case some of the available levels are not visible because they are used for the synchronizing part of the signal. The effect of using too few digits is "contouring" in black-and-white television and "snow" in color television.



PCM SYSTEM consists of a pulse-code modulator, which converts an analogue signal into a stream of digits, a transmission facility, which delivers the digits to the receiver, and a pulse-code demodulator, which converts the digits back into an analogue signal.

five bits (32 levels) and six bits (64 levels), the eye is fooled into seeing a nearly continuous gradation of tones. In color television the signal voltage at any instant is related to color as well as to brightness; therefore the use of too few digits does not result in simple contouring but instead produces a noisy picture, that is, one marred by "snow."

In a PCM system the pulse-code modulator converts an analogue signal whose bandwidth is W cycles per second into a stream containing $2mW$ pulses per second, where m is the number of pulses per code word. Since high-quality speech and television require m to be about 7, the pulse rate (in pulses per second) needed to transmit such signals is 14 times the original signal bandwidth (in cycles per second). The bandwidth required to transmit $2mW$ pulses per second can theoretically be as low as mW cycles per second, but under practical conditions it is about $2mW$ cycles per second. Thus PCM requires about $2m$ times as much bandwidth as the original analogue signal. Conventional modulation schemes (AM and FM) do not usually need as much bandwidth as PCM. As we shall see, however, digital transmission has advantages in many applications that more than offset the bandwidth penalty.

When the binary numbers of a PCM signal arrive at the receiver, they must be converted to their original analogue form. This is readily accomplished by passing the train of pulses through an electrical filter. The filter transmits the low frequencies (those that are less than half the sampling frequency) but practically eliminates the high frequencies. The low-frequency content of the pulse train is, except for small quantizing errors, essentially identical with the original signal that was fed into the pulse-code modulator.

One may ask: If PCM is so straightforward and simple in concept, why is it considered new? Why was it not widely applied long ago? The idea of transmitting analogue signals in pulse form is indeed not new: several early patents embody one or more of the basic concepts. These concepts were not easy to reduce to practice, however, until rather advanced vacuum tubes were available. The recognized "father" of PCM is A. H. Reeves, whose basic pulse-coding patent was issued in 1938. It was Reeves who pulled the concepts together and first produced drawings that showed the fundamental features of a complete PCM system.

Although a successful PCM system was built for the military in World War II, it had severe limitations that made it unattractive for commercial purposes. These limitations were almost wholly associated with the use of vacuum tubes, which are too big, consume too much power and are too prone to fail to be used in large numbers. Moreover, vacuum tubes tend to be good amplifying devices but not good switching devices. A PCM system requires switching devices that can be turned on and off very rapidly. This is precisely the same requirement presented by the electronic computer. Designers of both PCM systems and computers found the device they needed in 1948, with the invention of the transistor. That same year Shannon's work on information theory made it possible to assign quantitative values to the merits of PCM systems, thereby confirming intuitive and qualitative evaluations.

By 1955 serious development of PCM for widespread commercial use was under way. Preliminary models of PCM systems for telephone communication were tested in 1958 and commercial operation of production systems began in 1962. The first large-scale application of PCM was used to transmit 24 simultaneous telephone conversations over a pair of telephone wires. This system is now widely installed in the U.S. with the result that millions of people have heard speech transmitted by PCM without being aware of it. Recently PCM telephone systems have been introduced in Europe and Japan. Work is continuing in the U.S. and abroad on more advanced systems, particularly systems of much higher capacity that simultaneously transmit thousands of telephone conversations, or their equivalent, over a single communication channel.

Existing PCM systems generally use pairs of wires in a regular telephone cable, which contains up to 900 pairs. These wire pairs attenuate the digital signal so severely that transmission would be limited to very short distances if amplifying devices (called repeaters) were not located at intervals along the cable to restore the signal strength. The repeaters are placed either underground or on telephone poles. Because they must be installed every 6,000 feet one can appreciate the importance of small size, low power consumption and high reliability. Transistor technology has made such PCM systems possible.

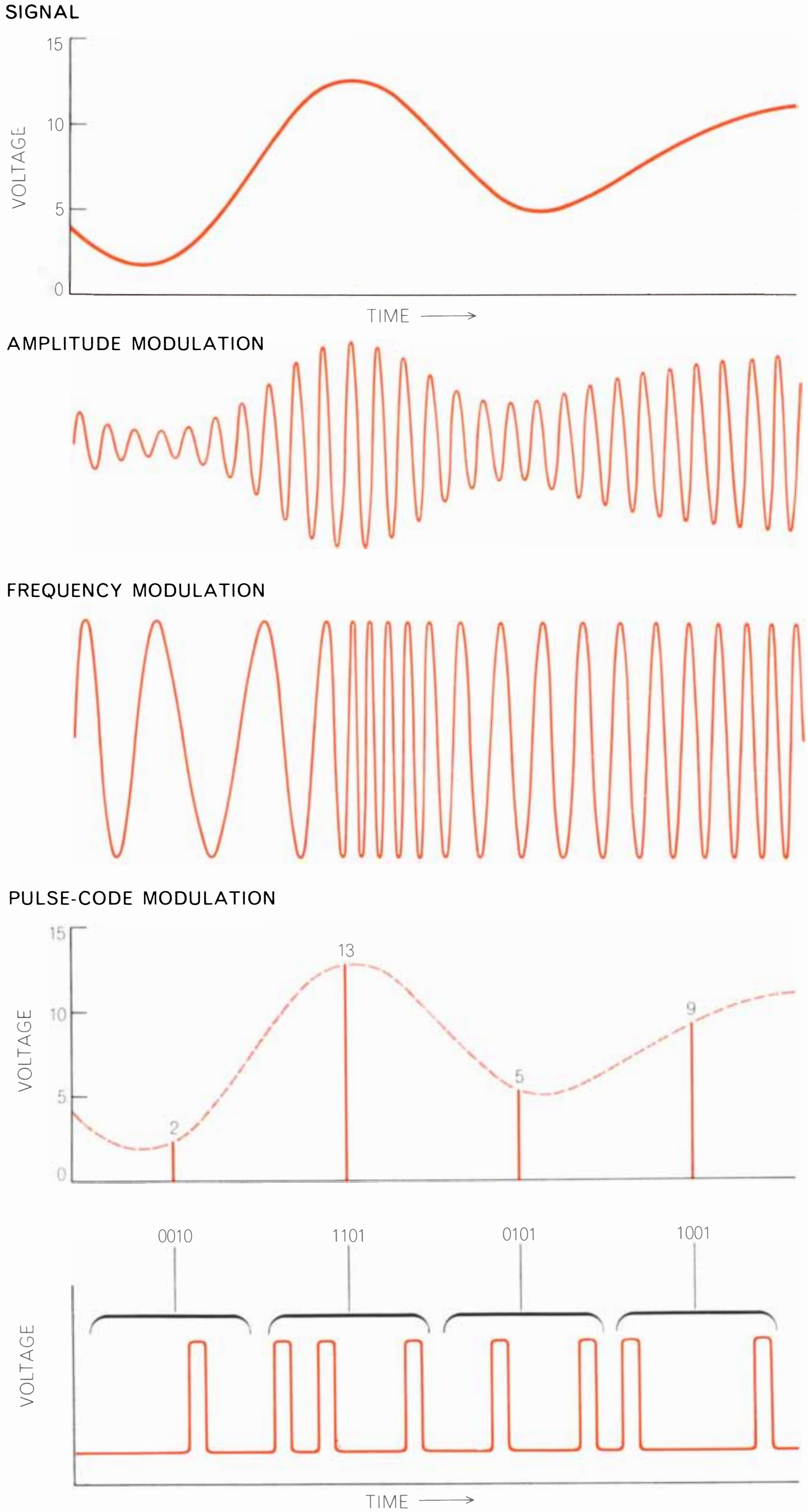
The fundamental advantage of PCM over analogue transmission for telephony arises directly from the nature of the repeater. In conventional AM systems the repeater is really nothing more than a high-quality amplifier. It therefore amplifies noise of all kinds along with the signal proper. Sources of noise include extraneous noise picked up by the cable, noise coupled in from other conductors ("cross talk") and thermal noise due to the random motion of electrons in the cable and amplifier circuit. The analogue amplifier has no mechanism for separating signal from noise.

Contrast this with the repeater that can be designed for a digital system. The PCM repeater is not a conventional amplifier but rather a circuit designed to detect whether an electrical pulse is present or absent. A moderate amount of noise added to a pulse does not alter the basic fact that a pulse is present. Conversely, if a pulse is absent from a particular time slot, a considerable amount of noise must be present before there will be a false indication of a pulse. In short, a repeater for digital signals (commonly called a regenerative repeater) can receive a digital pulse train that has been highly contaminated by noise, extract a clean signal and pass it along to the next repeater [*see top illustration on page 106*].

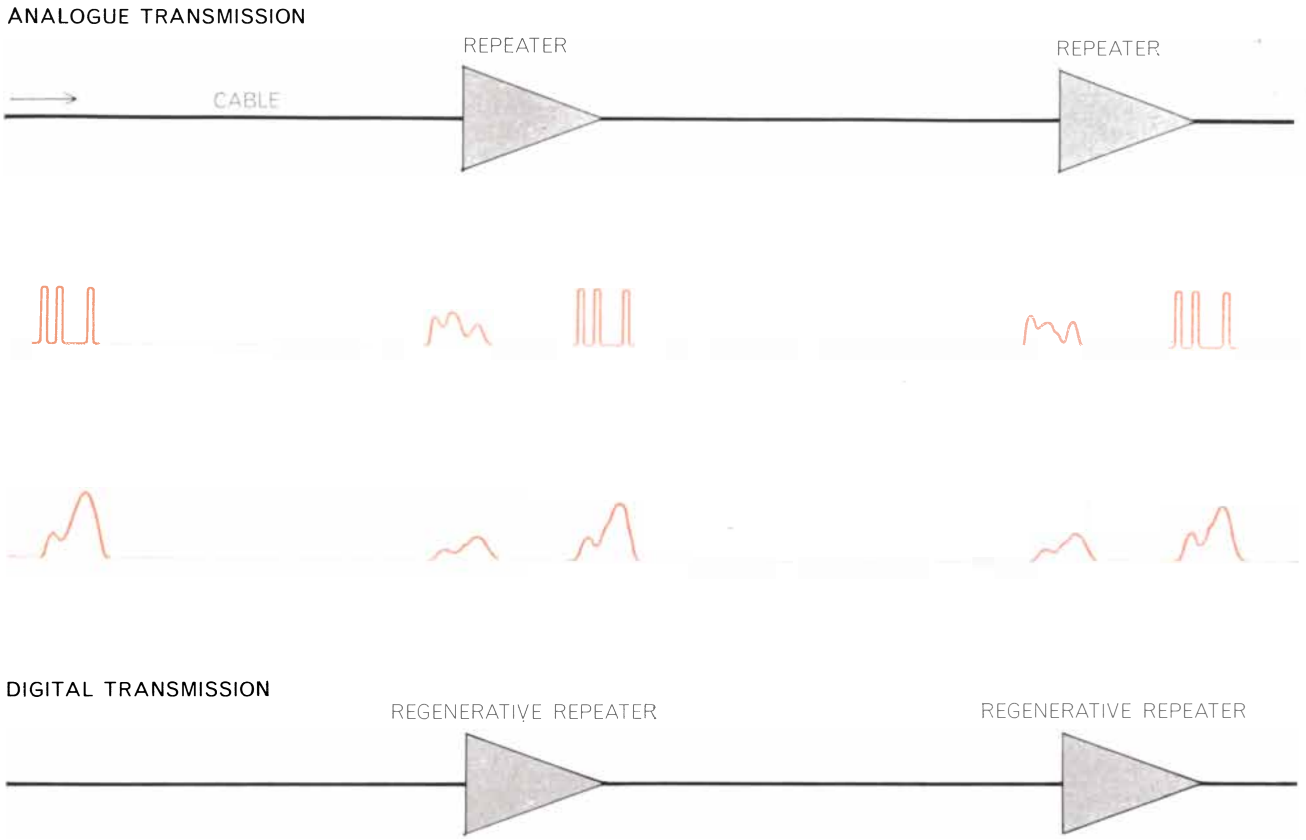
Occasionally a regenerative repeater will insert an extra pulse or drop out an existing one. If such errors are frequent, a speech signal will be accompanied by a background resembling the sound of frying eggs and a television picture will exhibit a random pattern of snowlike dots. If the error rate is no worse than 10^{-6} (less than one error per million pulses), the impairment is usually undetectable. If the noise contaminating the pulse stream is "white" noise (meaning noise with a "flat" power spectrum and a Gaussian amplitude distribution) a 10^{-6} error rate for an ideal repeater corresponds to a signal-to-noise ratio of about 100 to one, or 20 decibels. Repeaters, of course, are not perfect, and noise may not be entirely white. Nevertheless, an overall error rate of 10^{-6} is routinely achieved in transmission circuits containing thousands of repeaters in tandem. If the noise is white, the error rate is very sensitive to the signal-to-noise ratio. An improvement of two decibels in the signal-to-noise ratio (which could be achieved, for example, by placing repeaters slightly closer together or improving their design) will decrease the error rate from one in a million to one in a billion. To achieve a comparable performance conventional AM and FM systems generally require a much better signal-to-noise ratio.

The PCM system now used in telephone communication does not fully utilize the transmission capability of telephone cable. Recent work has shown that it is possible to transmit 96 conversations simultaneously on a single pair of wires. This is at least four times the amount of information that can be achieved with other modulation methods. Much higher capacities can be achieved on coaxial cables, the type of conductor in which a wire of small diameter is suspended in the center of a tube of larger diameter. Coaxial cables provided the original means for transmitting television signals. It has now been shown that by using PCM more than 3,000 telephone channels can be accommodated by a single coaxial cable. In demonstrating this capability engineers at Bell Laboratories used a pulse transmission rate of 224 million pulses per second. The duration of each pulse was about 4×10^{-9} second, or the time required for light to travel about four feet. This high-frequency pulse stream could have carried several color-television signals or various mixtures of television and voice [*see illustration on page 107*]. The system met all the stringent requirements for high-quality transmission.

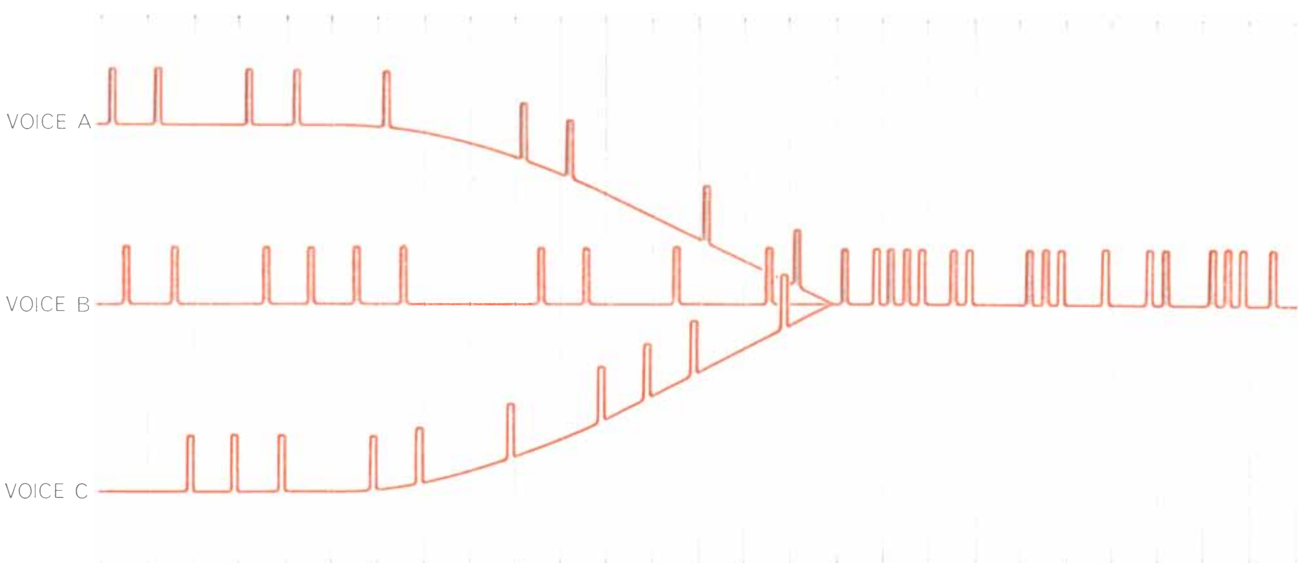
Engineers are currently investigating whether or not it would be desirable to employ digital signals in microwave radio links, which now usually carry FM signals. Such microwave systems are extensively used in land communication and for relaying voice and television



ANALOGUE SIGNAL CAN BE MODULATED for transmission by at least three different methods. In amplitude modulation (*top*) the voltage of the signal wave is represented by continuous variations in the amplitude of a higher-frequency carrier wave. In frequency modulation (*middle*) the signal is represented by variations in the frequency of a carrier wave. In pulse-code modulation (*bottom*) the voltage of the analogue signal is first sampled and the values of the samples are then expressed as binary numbers. The binary numbers are in turn transmitted in the form of a pulsed, or digitized, signal wave.



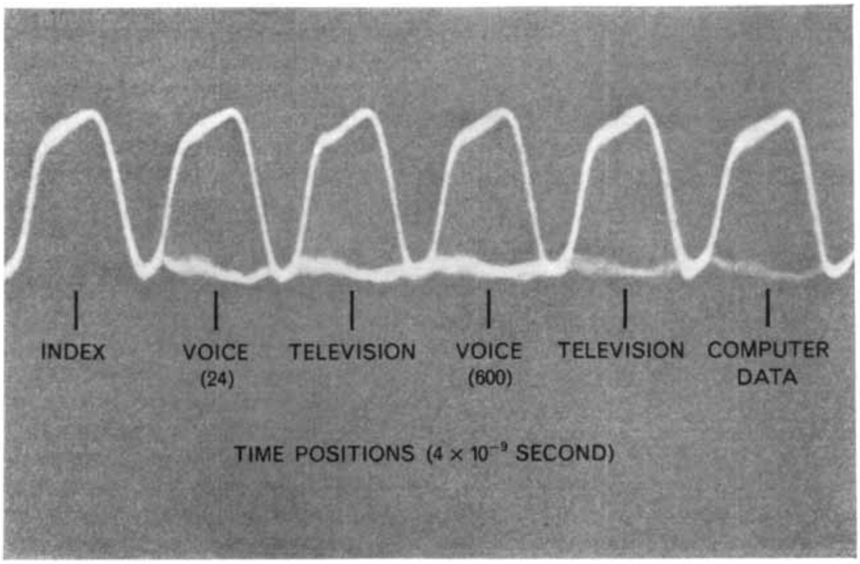
USE OF REGENERATIVE REPEATERS constitutes a main advantage of PCM transmission over AM or FM transmission. Both analogue and digital signals are severely attenuated and contaminated by "noise" in traveling long distances by wire, necessitating in both cases the use of some kind of amplifying device. The amplifiers, or repeaters, used in analogue systems (*top*) have the disadvantage of amplifying noise along with the signal proper. In PCM systems (*bottom*) regenerative repeaters examine the contaminated input signal in each time position to see if the voltage is higher or lower than a certain reference voltage (*broken horizontal line*). The repeater then regenerates an output signal that is identical with the original signal, with all the noise of the received signal removed. Thousands of regenerative repeaters can be connected in tandem to transmit high-quality signals over very long distances.



TIME-DIVISION MULTIPLEXING is a method for combining low-speed digital signals into higher-speed pulse trains. In this schematic representation narrow pulses from three pulse-code modulators are interleaved to form a single pulse train. The fact that the signals are all shown as encoded voice is unimportant; digitized signals of all types can be interleaved in the same way. Only the pulse rate (not the nature of the original signal) is a factor. The three pulse streams can be easily separated at the receiver, provided that a reference timing signal (called the framing signal) is available there. The framing signal in this case identifies the *A* pulse positions; the *B* and *C* pulse positions are generated by an electronic counter that is synchronized to the framing signal.

signals to satellites. As I have indicated, a disadvantage of ordinary PCM systems is that they require greater bandwidth than conventional FM systems. This disadvantage can be offset, however, by using a number system more complex than the binary one. Fewer pulses will suffice if each pulse is designed to have more than two amplitudes. Such a multilevel pulse train could modulate the amplitude, frequency or phase of the radio carrier signal.

Additional bandwidth saving results when the original signal is in digital form (such as the output of a computer) or when the nature of the analogue signal is such that it can be efficiently coded into digital form. Efficient coders are more complicated than simple pulse-code modulators because they are designed to remove some of the redundancy normally present in the analogue signal and thus achieve an output pulse rate much lower than that usually obtainable. Considerable effort has gone into the efficient coding of television and videotelephone pictures. There are many other variants of PCM that are more conservative of bandwidth than the straightforward sampling-and-coding system. Only further study and development will clarify the arguments for and against using digital modulation techniques in terrestrial and satellite microwave systems.



TRANSMISSION CAPABILITY of a PCM system is demonstrated by this oscilloscope picture of a PCM signal that is capable of carrying 224 million bits of information per second. The signal was obtained by interleaving pulses from several pulse-code modulators with other sources of digital signals. The pulse stream simultaneously transmits a television picture, computer data and two groups of voice signals, one containing 24 one-way conversations and the other containing 600 one-way conversations. The index pulse allows the various types of signals to be separated at the receiver. The duration of each pulse is four billionths of a second. If the entire pulse stream were devoted to data communication, it would be capable of transmitting letter by letter the contents of a 100,000-book library in about 15 minutes.

Because PCM signals normally require generous bandwidths they are well matched to "guided wave" transmission mediums: cable pairs, coaxial cables, intercity wave guides and light beams. The last two are extremely promising for the more distant future. Moreover, the peculiarities of wave guides and light beams are such that it is difficult to employ them except with digital-modulation schemes. Because these two mediums have enormous capacity, they can carry many digital signals simultaneously over different portions of their total frequency band. Their transmission characteristics are such, however, that there can be appreciable interference among the many signals being simultaneously transmitted. Removing such cross talk is a specialty of regenerative repeaters. For this reason almost all development work on intercity wave guides and light beams has taken for granted the use of PCM.

An intercity wave guide consists of a round pipe about two inches in diameter whose inner surface is precisely engineered. A favored design is the helix guide: a structure consisting of a continuous wrapping of fine copper wire on the inside of a steel pipe. A wave guide ordinarily supports several modes of propagation; the function of the wire inner lining is to suppress all but the desired mode that has minimum loss. In addition, the helix guide need not be fabricated as precisely or laid as straight as a wave guide consisting of a plain copper tube. The helix guide is neither very complex nor very costly. Such a wave guide can transmit millimeter-wave signals in the frequency band between 40 and 110 gigacycles (40,000 to 110,000 megacycles). A bandwidth of 70 gigacycles provides approximately six times the transmission capacity of the whole radio spectrum. Using PCM, the wave guide could simultaneously transmit at least 180,000 two-way telephone conversations or 300 color-television programs. Attenuation is so low that repeaters can be spaced 20 miles apart. The intercity wave guide will be economically attractive, however, only on routes with exceedingly heavy traffic.

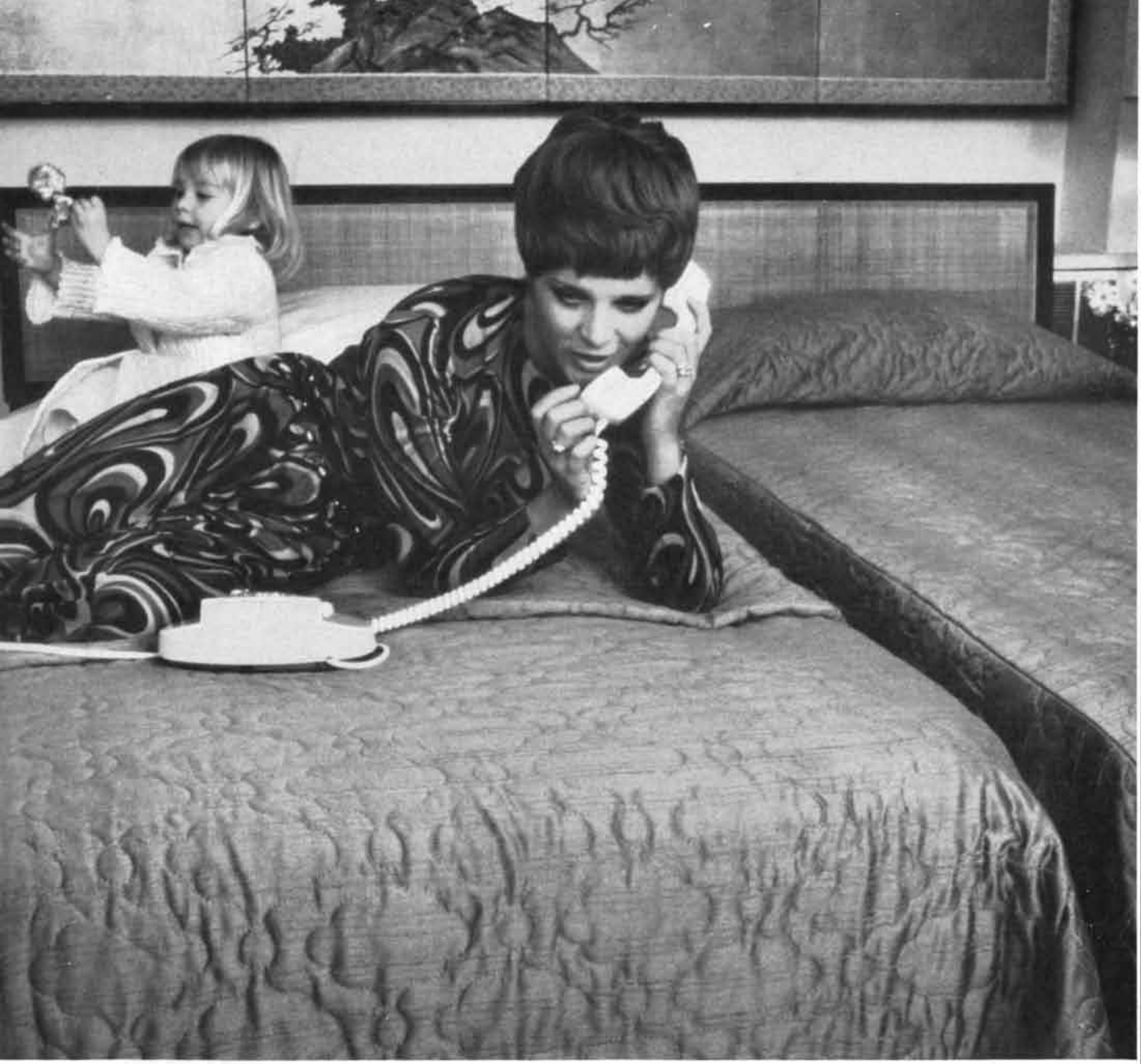
The other high-capacity communication channel under intensive study, the light beam, became possible with the invention of the laser. The laser emits high-directional light of a single frequency that can be modulated at enormous information rates [see "Communication by Laser," by Stewart E. Miller; Scientific American, January, 1966]. Although the feasibility of pulse modulation of laser beams has been demonstrated, many refinements will be needed before commercial operation can be contemplated. It does not seem likely that high-reliability communications can be achieved by using laser beams in the open atmosphere; they are too readily blacked out by rain, snow and fog. An alternative is to shine the beam through a pipe fitted with special lenses to keep the beam from spreading. Good progress has been made in using the refractive properties of gases at carefully controlled temperatures to act as lenses. Like the intercity wave guide, laser-beam communication systems will require heavy traffic to be economic.

At the other end of the demand spectrum PCM has a unique ability to fill the special needs of many low-traffic communication systems. For example, the first photographs of Mars were transmitted across 130 million miles of space at the rate of 8⅓ digits a second. This low rate was necessitated by the low power of the transmitter aboard the spacecraft *Mariner IV* and the high level of background noise. Digital transmission has also made secure communication readily available. A simple encrypting device can scramble the transmitted digits so that only a properly keyed receiver can unscramble and extract the original message.

If PCM develops as vigorously as expected, one can visualize a vast nationwide communication network in which all kinds of signals—voice, television, videotelephone and data—are transmitted in the form of binary digits. Once digitized, all signals are in the same electrical format and can be intermixed indiscriminately by time-division multiplexing without incurring penalties in cost or performance.

DIGITIZED VIEW OF MARS was transmitted back to the earth by the spacecraft *Mariner IV* in July, 1965. A PCM signaling system was employed to enable the low-power transmitter (10 watts) to send a virtually noise-free signal across 130 million miles of space. The *Mariner IV* pictures of Mars were transmitted to the earth at a rate of 8⅓ digits per second.

DETAIL OF MARS PICTURE shows the subtle gradation of tones produced by the 64-step intensity scale that was specified by the six-digit code used to transmit the *Mariner IV* pictures. The craters in this photograph are the ones at lower left in the photograph at top.



The Adjustable Brain of Hibernators

The unique state of torpor that some mammalian species can attain is the result of a complex chain of events. It now appears that changes in the usual functioning of the brain underlie these events

by N. Mrosovsky

Many animals face an annual crisis in their struggle for existence: the winter shortage of food. To meet this challenge they have evolved two quite different biological strategies. Some species meet the food shortage by the complex chain of events we call migration; among the links in the chain are awe-inspiring feats of navigation. A quite different strategy is the chain of events we call hibernation. Although this behavior is perhaps not as dramatic as migration, it is in many ways even more remarkable.

Consider an animal actually in a state of hibernation, say a ground squirrel of the genus *Citellus*. The temperature of the animal's body hovers only a few degrees above the freezing point (about five degrees centigrade, more than 30 degrees C. below the normal level). Curled up, immobile, the squirrel appears to have withdrawn from nature. In comparison with its active state nearly every system in its body is radically altered. The composition of the blood, the functioning of the digestive tract, the state of the fat reserves, the output of hormones—these and much else are changed so that a new physiological state is achieved.

Some aspects of this state are astonishing. The animal's heart is beating; perhaps only once a minute, but still beating. How is this ultraslow rhythm maintained? How indeed is cardiac muscle capable of contracting at all at five degrees C.? In nonhibernators the heart generally will not beat below 20 to 10 degrees. The same question arises concerning the nervous system; at temperatures at which mammalian nerve normally does not function the hibernator's nerves remain active. Disturb the ground squirrel and it will respond. Its body heaves, its ears may prick up and sometimes it emits a piercing shriek. The animal remains alive and responsive at five degrees; a nonhibernating mammal kept at this temperature for any length of time would be dead.

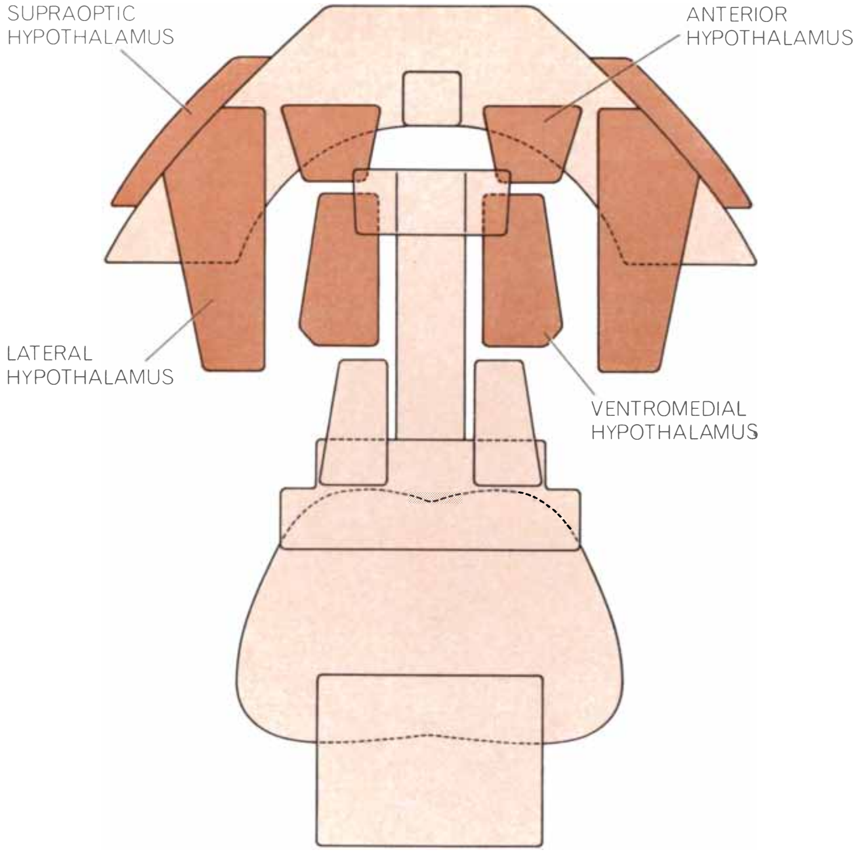
One should ask at the start what facts have been established about hibernation. Evidently the hibernating animal becomes cool and torpid, but what are the details of the process? Exactly how does the reduced body temperature compare with the temperature of the surroundings? Does the body temperature remain constant? In 1959 Felix Strumwasser, then working at the Walter Reed Army Institute of Research, made continuous records of the temperatures of hibernating animals that answer these questions. He implanted thermocouples in ground squirrels of the species *Citellus beecheyi* and found that when the animal was hibernating, the temperature of the brain remained constant in spite of fluctuations in the temperature of its skin. Moreover, its brain temperature consistently remained a few degrees warmer than the ambient level [*see bottom illustration on page 113*]. This small but consistent difference is enough to show that the hibernator is not thermally passive but can regulate and maintain its temperature within closely defined limits. The limits are quite different from those of the summer months, but the animal's "thermostat" (to use an analogy proposed by C. Ladd Prosser of the University of Illinois) has not been turned off; it has merely been turned down.

Coming out of hibernation is quite as remarkable an event as entering it. Whether the arousal is spontaneous (another way of saying that the cause is unknown, although it is presumably internal) or is in response to external disturbances, the main occurrence is a rapid rise in body temperature. Charles P. Lyman and his colleagues at the Harvard Medical School, working with hamsters, have recorded temperature rises of 30 degrees C. or more over a period of three hours. During this interval there is a dramatic transformation from torpor to activity. The self-rewarming will take place in spite of such handicaps as a cold-room environment only a few degrees above freezing. This is a considerable achievement in heat production; it appears to be made possible largely by a special tissue called brown fat [see "The Production of Heat by Fat," by Michael J. R. Dawkins and David Hull; Scientific American, August, 1965].

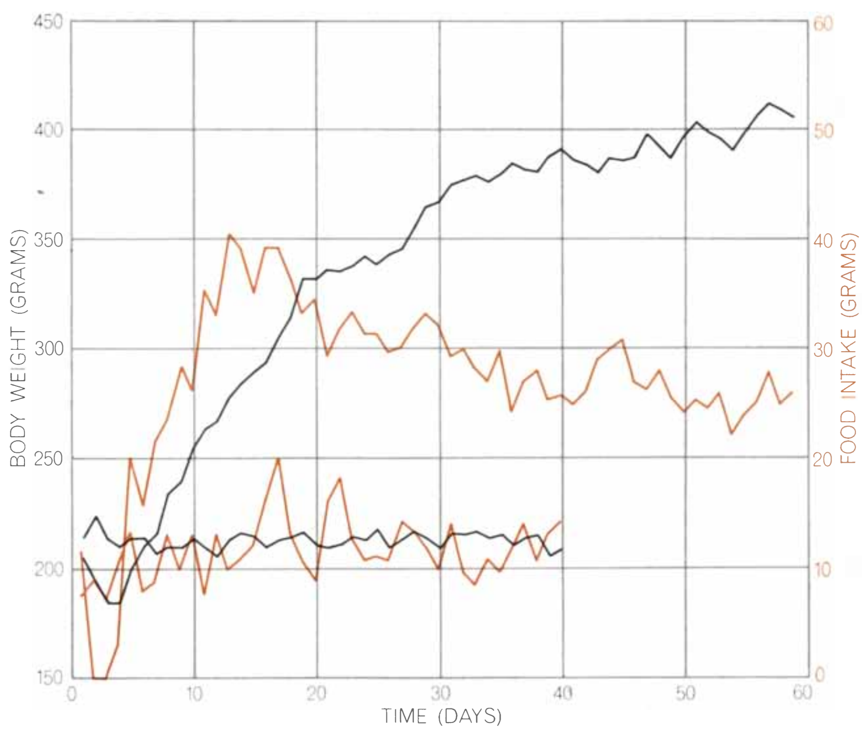
It is thus evident that the hibernator is in charge of its temperature and not completely at the mercy of its environment. But how is this arranged? Strumwasser's recordings provide a clue; he studied ground squirrel temperatures not only during hibernation but also during the period when the animals were entering the state. Like other animals that are active during the day, ground squirrels show diurnal fluctuations in temperature: they are somewhat cooler when they are asleep at night than they are during the daylight hours. As fall comes to an end, however, the nightly dips in body temperature are increasingly accentuated until finally the coolness and torpor of the night hours continue through the following day. The ground squirrel then remains in hibernation (except for occasional spontaneous arousals) for the rest of the winter [*see top illustration on page 113*]. Viewing this pattern in terms of control-system analogies, it seems apparent that the transition from an active state to a hibernating one involves an exaggerated swing of the same thermostat that ordinarily manages the daily temperature cycle.

HIBERNATING GROUND SQUIRREL, curled and undisturbed, lies in an investigator's hand. The body temperature of a hibernating ground squirrel in the wild often stays at five degrees above zero centigrade for periods that would kill a nonhibernating mammal.

ALERT GROUND SQUIRREL has recovered from its state of hibernation. The animal can raise its body temperature by 30 degrees C. or more in a three-hour period by boosting heat production of its brown fat. Both photographs show the 13-striped ground squirrel.



HYPOTHALAMIC REGION of a rodent's brain is seen from above, divided schematically. Names identify four areas that are apparently implicated in some aspects of hibernation.



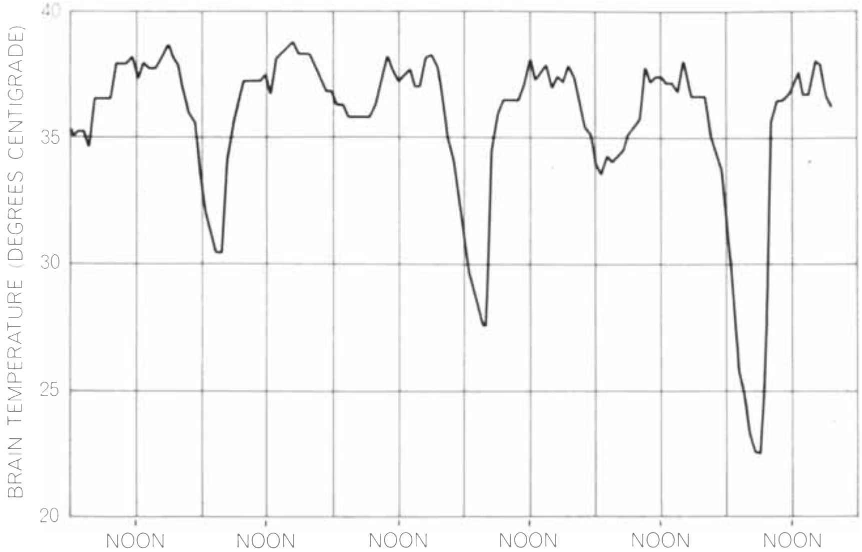
DESTRUCTION OF SATIETY CENTER in the ventromedial area of a rat's hypothalamus produces a sudden increase in consumption of food (*color*) and a resulting rise in body weight (*black*). Food intake soon levels off and the brain-damaged animal's weight reaches a new plateau in a few weeks. The records for brain-damaged rats and for normal rats (*bottom pair*) were obtained by Philip Teitelbaum of the University of Pennsylvania in 1955.

But what exactly controls the normal daily cycle? Although we are a long way from having the complete answer, it is known that certain areas of the mammalian brain, located in the anterior region of the hypothalamus, are particularly sensitive to temperature changes. When this region in a nonhibernating animal—say a rat, a dog or a goat—is locally cooled, heat-producing mechanisms such as shivering are activated. When the same region is warmed, heat-loss mechanisms such as panting and sweating start to operate. Furthermore, if this part of the brain is surgically destroyed, the animal loses its ability to keep warm in the cold. The anterior hypothalamus, then, contains one (whether or not it is the only one) temperature-sensitive structure in a kind of internal thermostatic control system.

The facts presented so far favor the possibility that hibernators are able to alter the setting of their thermostat. The animals' temperature-control system remains active during hibernation. Their mode of entry into hibernation suggests an exaggeration of the normal temperature cycle. Finally, areas in the anterior hypothalamus are known to control the temperature-regulating activities of some mammals. It is therefore certainly possible—even plausible—that hibernation involves the alteration of the animal's hypothalamic temperature controls. The existence of such a possibility is a far cry from established proof. But what if several other characteristic features of hibernation could be shown with equal plausibility to involve adjustments in the same general brain areas? If this were so, at least the attention of investigators would be focused on these areas, and perhaps a more general and unifying conception of hibernation might emerge.

One of the most striking seasonal changes among hibernators is the autumnal development of obesity. The animals' food consumption soars far above normal as winter approaches; the obesity is a kind of physiological hoarding that protects the animal when food supplies run out. Now, the regulation of food intake involves many factors, but as far as obesity is concerned one brain system has been implicated with particular clarity. The key area is in the ventromedial region of the hypothalamus; it constitutes a satiety center. The system works as follows. Signals from various receptors in the body that are activated by eating impinge on nuclei, or groups of brain cells, in the ventromedial hypothalamus. The ventromedial nuclei then send inhibitory messages to eating centers nearby in the lateral hypothalamus. When the inhibitory impulses reach the lateral areas, eating decreases or stops. If the nuclei of the satiety center are surgically destroyed or chemically blocked, however, obesity is induced. The signals normally associated with satiety no longer have their effect; the animal does not know when to stop eating and soon becomes grossly fat.

Studies of experimentally induced obesity in rats by various research groups have made clear one important element in obesity. After their satiety centers are destroyed rats undergo an initial phase of rapid weight gain, but the rate of gain gradually decreases and eventually the animals reach a new weight plateau. When they are temporarily starved and then once more given unlimited access to food, they soon regain the same plateau. Interference with the satiety signals thus produces a defect not only in eating but also in whatever weight-level control maintains an animal at its normal weight. The obese rat still regulates its weight, but the control has been set at a much higher level than the one that prevailed before surgery.



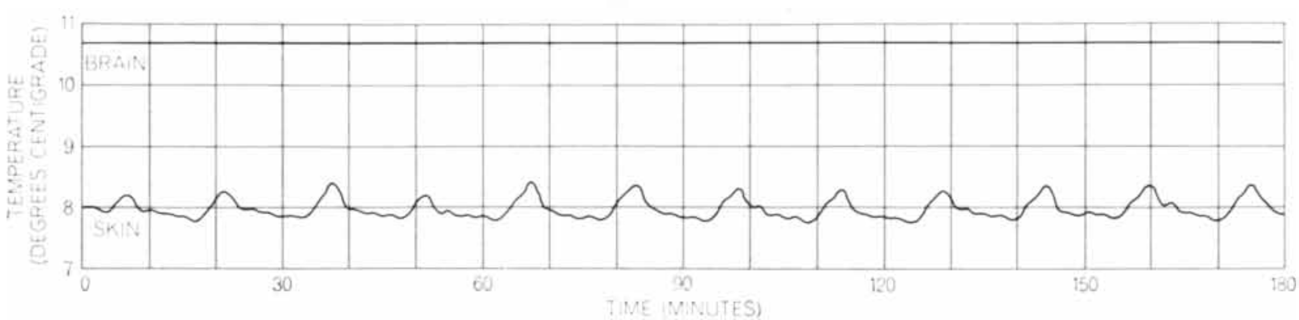
ONSET OF HIBERNATION is signaled by larger than normal drops in the animal's temperature at night. Eventually night cooling produces a torpor from which the animal does not recover the next day. Felix Strumwasser recorded these progressive drops in the brain temperature of a ground squirrel in studies at the Walter Reed Army Institute of Research.

It is thus evident that, in rats at least, areas in the ventromedial hypothalamus function to keep the weight of the animal fairly constant over long periods of time. Rats, of course, are not hibernators. But what if among hibernating animals some depression or functional quiescence occurred naturally in the ventromedial hypothalamus in late summer or early fall? With their satiety signals upset, the animals could be expected to emulate the surgically altered rats—to overeat and grow obese. There is no direct evidence that such quiescence occurs, but some details of autumnal obesity in hibernators bear a striking resemblance to the details of artificial obesity.

In addition to the overeating that follows destruction of the satiety center in rats there are changes in the strength of the animals' urge to reach food. When one sees an immediately postoperative rat shoveling in food as it staggers around even before the anesthetic has worn off, one might conclude that the animal was very hungry indeed. This is not so. If reaching the food is made difficult in some way—such as placing it in a pot with a heavy lid or requiring the rat to press a bar many times in order to receive a food pellet—a rat with its satiety center destroyed not only eats less than when its access to food is free but also takes even less food than an unoperated control rat will.

In effect the components of the hunger drive in the brain-damaged rat have been dissociated. Appetite is gone and only the capacity to consume remains. If food is readily available, such a rat overeats and becomes fat. If the food is hard to obtain, however, the animal remains thin. It is not hard to imagine behavior of this kind having survival value for a hibernator in the wild at the onset of winter. Available food would be consumed voraciously, but the animal would be reluctant to expend energy in a food quest that might go unrewarded.

In fact, hibernators sometimes do adopt such a paradoxical but useful attitude toward food. While working at University College London, I conducted an experiment in which the dormouse, a European species of hibernator, and the North American ground squirrel *Citellus lateralis* were given daily access to a food bar; when the animals pressed the bar down, a pellet of food was delivered [*see illustration on page 116*]. Over a period of days the number of presses required to release one food pellet was increased until the point was reached at which the animals would no longer work the bar. I then compared my results with some obtained for both artificially obese rats and unoperated control rats by



ANIMAL'S TEMPERATURE shows surface fluctuations during hibernation (*bottom*). In spite of the surface changes, however, the animal's slightly higher internal temperature stayed steady (*top*). These ground squirrel records are also from Strumwasser's work.

Philip Teitelbaum of the University of Pennsylvania. When only a few presses produced a reward, my hibernators were eager for food, but when the work required of them was increased, they soon gave up. Their laziness in this respect resembled that of the obese rats. In contrast, Teitelbaum's normal rats still worked the bar even when 256 presses were required to obtain a single pellet.

In 1965 Elizabeth Cornish and I undertook a series of tests at University College London to compare the activity of hibernating and nonhibernating rodent species that were temporarily deprived of food. Dormice and ground squirrels comprised the first group, guinea pigs and rats the second. The deprived nonhibernators became hyperactive—as if they were searching for food or reacting to stimuli that might lead them to some. The dormice and ground squirrels remained relatively unmoved and showed only slight changes in activity. Two points of similarity thus exist between normal hibernators and brain-damaged rats. Both possess a potential for getting fat and both exhibit a paradoxical laziness with respect to working for or searching for the food on which they will gorge themselves when it is available. These constitute circumstantial evidence that the same weight-control systems, involving the satiety centers of the ventromedial hypothalamus, are involved in both instances.

One important difference exists between the naturally obese hibernator and the artificially obese rat: the rat remains obese for life but the hibernator begins to lose weight as soon as it grows torpid. To put it another way, brain damage has produced a new and permanent weight plateau in the rat—a new, fixed setting of the weight-level control. In the hiber-

OBESE DORMOUSE has overeaten for weeks and is now ready to hibernate. The animal weighs more than half a pound, in contrast to a more usual weight of four ounces. The fat accumulated before hibernation provides the energy needed to rewarm the cold body.

nator, on the other hand, the control must move flexibly from a high setting in the fall, which leads the animal to become increasingly fat, to a relaxed setting (after the programmed weight level is reached), which enables the animal to hibernate and lose weight.

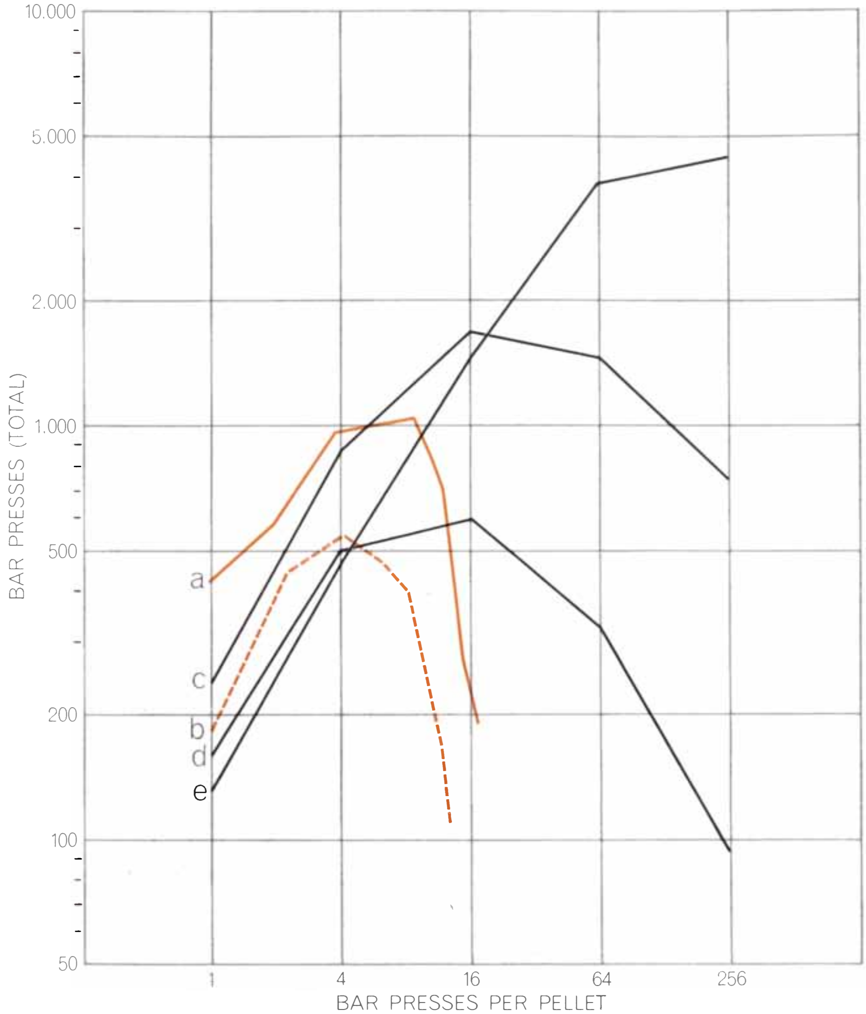
Students of hibernation are indebted to Eric T. Pengelley (now at the University of California at Riverside) and Kenneth C. Fisher of the University of Toronto for years of detailed observations of a number of long-term cycles in hibernating animals. They have worked mainly with the ground squirrel *C. lateralis*, measuring these animals' body weight, food intake and state of torpor for months on end. Their data show that all three indexes fluctuate in cycles of about a year's duration [*see upper illustration on page 118*]. Pengelley and Fisher have also found that the animals' hibernation is by no means as simple as a direct response to winter's shortened daylight hours and lower temperatures. The ground squirrels hibernated when they were kept in rooms illuminated 24 hours a day and held at a constant temperature. They even hibernated at normal room temperature (21 degrees C.). Pengelley and Fisher have come to the conclusion that some internal biological clock with a periodicity of about a year is responsible for the animals' hibernation; Pengelley now calls this a "circannian" rhythm (from the Latin for "about a year") to emphasize its apparent kinship with various "circadian" rhythms (meaning "about 24 hours") noted by many investigators of biological clocks.

My observations of dormice, which are limited when compared with the work of Pengelley and Fisher, nonetheless suggest a different explanation of this cyclical phenomenon. The dormice were kept in a warm room and weighed every day. The act of picking up the animals for weighing was sufficient to arouse them at times when they were torpid. Under these conditions the cycle of body weight ran its course in a few weeks rather than many months [*see lower illustration on page 118*]. In both dormice and ground squirrels heavy eating and weight gain was followed by reduced eating and activity and by weight loss. With respect to duration, however, the body-weight cycles in dormice differed from those in ground squirrels. The short cycles in dormice could in no way be termed circannian. I prefer the following interpretation.

Each time I roused one of the torpid dormice, the animal was forced into an extra expenditure of energy; as I have indicated, the act of rewarming requires intensive heat production. This abnormal spending of energy, in turn, caused the dormice to lose weight more rapidly than they would have otherwise and soon brought them down to a critical minimum weight. At the point where any further weight loss would have endangered their survival, the animals' hibernation ended. In my view the various phases of the body-weight cycle are triggered by the state of the animal's fat deposits rather than by any biological clock; the sooner a critical minimum is reached, the sooner the next stage in the cycle—active weight gain—will begin.

On this basis a hibernation cycle is a chain of ordered events; each event or stage ushers in the next one. Accordingly hibernation has similarities with the sequences of events studied by ethologists. If one event in the sequence does not take place, progress along the chain should be blocked. There are instances in which this happens. Thin ground squirrels, for example, seem to hibernate less readily than fat ones; the lack of fat evidently obstructs the animal's movement onward to the torpor phase of the cycle.

What have we learned about hibernation up to this point? At least some facts are definite: first, hibernators are not thermally passive; second, hibernators possess a highly developed mechanism for heat production; third, hibernation can occur even in a warm and lighted room; fourth, hibernators can consume enormous amounts of food in the fall but at the same time are only



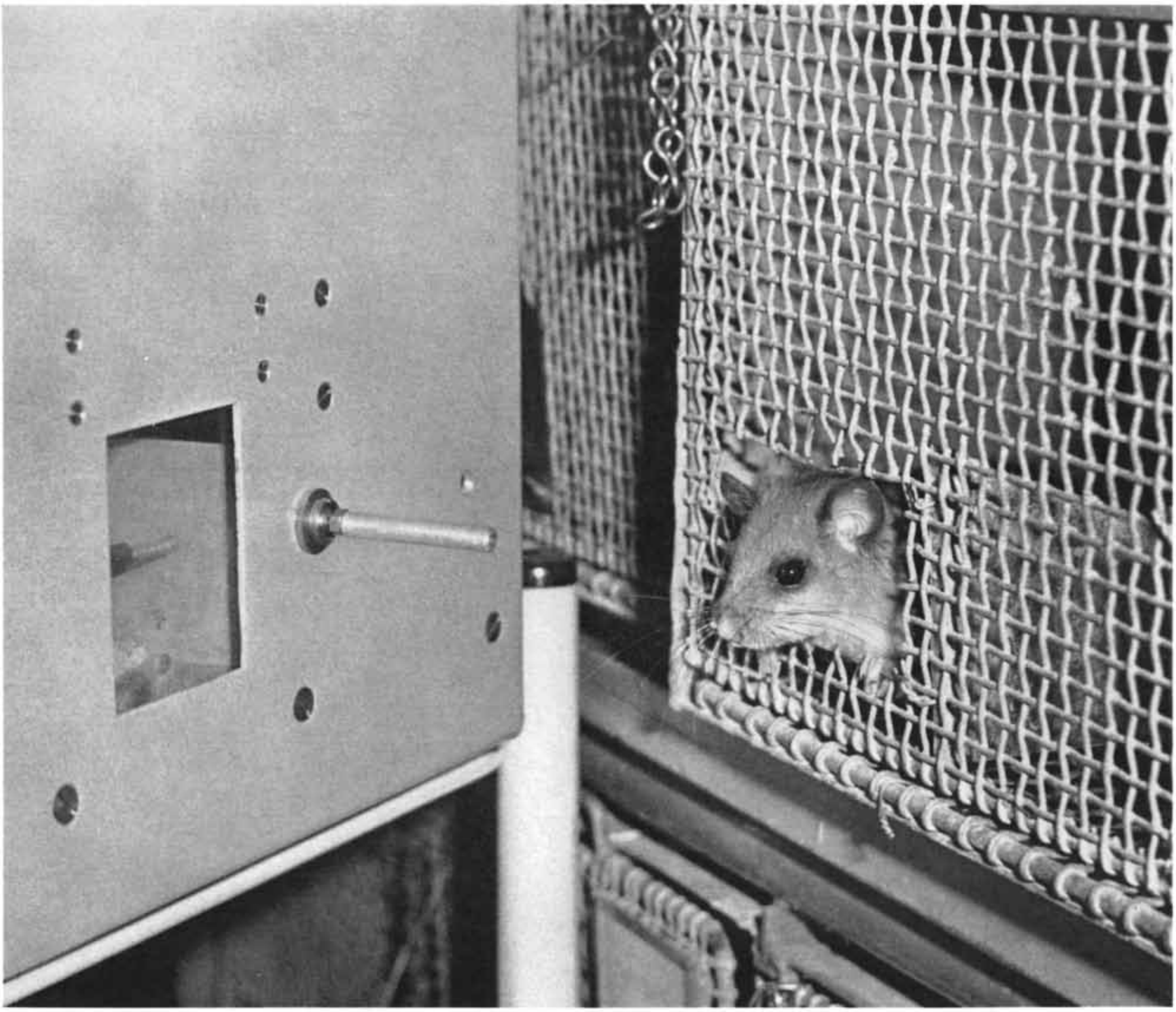
FEEDING "LAZINESS" exhibited by two hibernating species (*color*) during tests of hunger motivation is compared with the performances of normal and brain-damaged rats. All five groups received food pellets after they had pressed a bar; the number of presses required to obtain a pellet was steadily increased. The dormice (*a*) and ground squirrels (*b*) stopped working sooner than already obese rats (*c*) and rats still in the weight-gaining phase (*d*). All four groups had quit, however, before effort by normal rats (*e*) began to flag.

weakly motivated to obtain it. It is now time to add one more fact to our list.

During hibernation there is an obvious need for the conservation of water. Now, the body's capacity for water retention depends on what amount of an antidiuretic hormone is released by the pituitary gland. This release is controlled by groups of brain cells in the anterior part of the hypothalamus that are known as the supraoptic nuclei. This is one more case in which a region of the hypothalamus is implicated in the hibernation process. Here, however, the implication is more than speculative. Paavo Suomalainen and his colleagues in Finland have found that the appearance of the supraoptic nuclei in hedgehogs (which are hibernators) fluctuates with the seasons. In winter there is an increase in the size of the nucleus; this indicates intensified secretion by the cells. Here is a concrete example in the brain of seasonal plasticity related to hibernation. When we now suggest that changes in other hypothalamic regions could provide most or all of the mechanisms required to produce the entire chain of hibernation behavior, we are at least no longer without a precedent.

Exactly what can be reasonably postulated? The hibernation cycle involves enormous fluctuations both in food intake and in temperature regulation. Leaving aside the question of temperature regulation, I think it reasonable to assume that the food-intake changes in hibernators involve a quiescence in the ventromedial nuclei of the hypothalamus. The quiescence alters the balance between the satiety centers and the adjacent eating centers in favor of the latter. No new properties need be invoked to account for this aspect of hibernation behavior; all that is required is a rebalancing of existing components. The key problem that remains is learning how this rebalancing of components—and of other components possibly responsible for other aspects of hibernation—is achieved. A question that is probably less important, but that must be faced, lies in the fact that there are many different species of hibernator. Each species may behave quite differently.

If brain changes do occur in all hibernators, a fuller understanding of the hypothalamus of hibernators could be instructive and also relevant to man. The hypothalamus of hibernators evi-



FOOD DISPENSER ***(left)*** **was used by the author to measure the extent to which hibernating species of rodents would work to get extra nourishment. The results proved to resemble those obtained by Teitelbaum in a similar test of brain-damaged and normal rats (*see illustration on preceding page*). This suggests that quiescence of a satiety center in hibernator's brain occurs before hibernation.**

SCIENCE/SCOPE

The seventh Surveyor, sent to explore the unknown rock-strewn ridges of the moon's south polar region, landed only 1½ miles from its bullseye near the crater Tycho. Its TV camera shortly revealed the boulders it had barely missed. Later, laser beams aimed at the moon from stations in California and Arizona were recorded by Surveyor 7 and transmitted back to earth. One of the lasers was built by Hughes Research Laboratories.

So complete had four earlier Surveyors tested and photographed all proposed landing areas for the first manned moon mission that Surveyor 7 could be freed for this important scientific mission. NASA scientists hoped to learn more about the character of the moon, which could lend clues to its origin and that of the universe, by landing near the fallout from this large new crater.

Surveyor 7's only glitch -- its temporary inability to lower the tiny box containing its alpha backscattering instrument all the way to the surface -- was turned into a brilliant demonstration of the commandable-spacecraft concept by a 40-man team of JPL and Hughes scientists. After an all-night session of delicate maneuvering, they successfully used the surface scoop to free the box and guide it to the soil.

The West German Ministry of Defense has awarded Hughes a contract for two prototype computer-controlled flight-line testers for their F-104G Starfighter's inertial navigation system. The tester, using pre-programmed test sequences, enables relatively unskilled operators to make fast, accurate tests. It uses a Hughes H3118M computer, and can be adapted for testing other avionics.

Hughes has openings for engineers in these fields: circuit and logic design, telecommunications systems design, display systems, space/airborne command system design, computer memory systems. Requirements: accredited degree, U.S. citizenship, and at least two years related experience. Please write: Mr. J.C. Cox, Hughes Aircraft Co., Culver City, Calif. An equal opportunity employer.

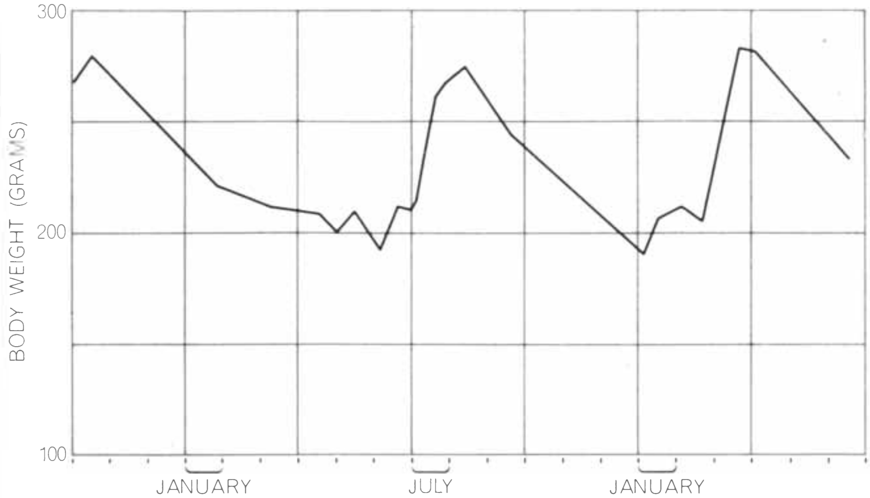
It's the TV camera in Walleye's nose that gives this U.S. Navy and U.S. Air Force air-to-surface weapon its phenomenal accuracy. After launch, its electro-optical guidance system keeps it on course as it glides to its target. Walleye is now being manufactured at the Hughes Tucson, Arizona, plant under a second-source contract.

First working laser, developed in 1960 at Hughes Research Laboratories, was the ruby laser. Last November, Hughes was awarded a U.S. patent covering all ruby laser systems until 1984. Although many patents have been granted on other lasers since 1960, the ruby laser is one of the most important for practical applications.

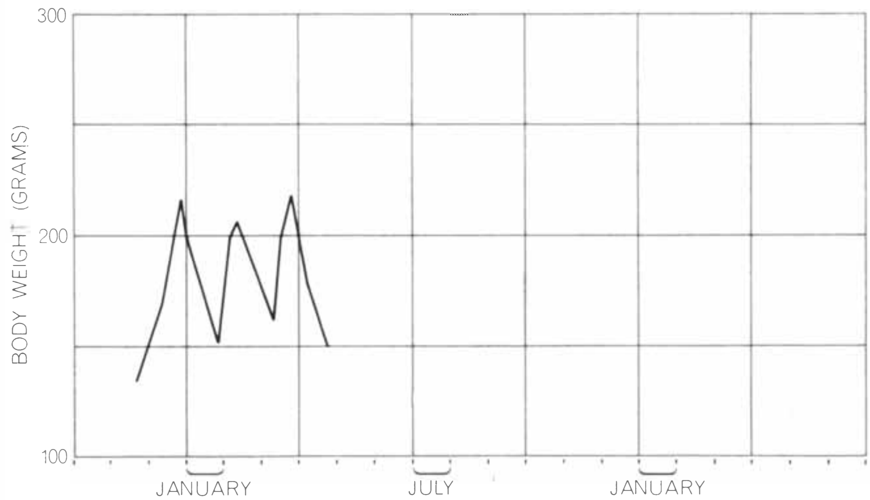
Creating a new world with electronics

HUGHES

HUGHES AIRCRAFT COMPANY



BIOLOGICAL CLOCK, in control of an annual rhythm, is suggested by the cyclical nature of ground squirrels' preparatory overeating, their resultant obesity and subsequent loss of weight during hibernation. Both the observations and the proposal of a biological clock were made by Eric T. Pengelley and Kenneth C. Fisher at the University of Toronto.



ANNUAL RHYTHM is brought into question by one of the author's experiments. Daily handling of dormice prevented continuous torpor. This caused an overexpenditure of energy and the animals repeated a cycle of weight gain and loss in weeks rather than in months.

dently has the same basic components that are found in the hypothalamus of nonhibernators. For example, Evelyn Satinoff of the University of Pennsylvania has found that when the anterior hypothalamic region of ground squirrels is destroyed, the animals' thermoregulation is impaired just as it is in nonhibernators. I have electrically stimulated areas in the lateral hypothalamic region of ground squirrels and thereby produced feeding behavior just as similar stimulation does in nonhibernators. J. Robert Troyer of the Temple University School of Medicine has shown that the anatomy of the supraoptic-pituitary area of the hypothalamus in bats, which are notable hibernators, is essentially the same as in the brain of nonhibernators.

Hibernators, with their vast obesity, their torpor and their huge fluctuations in physiological function, are extremists. Other mammals, however, also experience fluctuations in weight, temperature and activity. If hibernators achieve certain extremes by means of systems that are essentially the same as those in other mammals, then problems of weight, temperature and activity in such mammals could be explained by smaller changes in the same systems. The possibility that knowledge obtained from hibernators can eventually be related to a wide spectrum of behavior in other animals, ourselves included, makes this a particularly exciting field of investigation.

What gives this speeding squad car extra safety during emergency stops?

A life-saving R/M composite.

Lights flash. Sirens scream. Emergency! Rapid decelerations in quick succession overheat ordinary brake linings, causing them to grab, pull, or fade.

For greater safety, we added copper to the brake linings on this emergency vehicle, speeding dissipation of heat from lining to shoe to drum . . . preventing heat and water fade.

This copper/asbestos/resin composite is only one of literally hundreds of brake linings and disc brake pads we manufacture for every type of service. A high percentage of all autos, trucks, buses, and off-highway vehicles are equipped with our brake, clutch, and transmission materials.

Using combinations of asbestos, rubber, sintered metals, specialized plastics, and other natural and man-made materials, we create life-saving composites for the automotive industry.

Got a friction material problem? Get in touch with us now.

Specializing in creative composites

Raybestos-Manhattan, Inc.

Passaic, N.J. 07055

R/M

The end of the world

To early man, *ocean* was a great river endlessly circling the earth. His world ended at the horizon in a limbo of sky and water.

As civilization advanced, so did man's field of vision. The earth grew round. The ocean took on real dimensions. We explored and charted its surface. We discovered that it covers nearly three-fourths of the globe and contains 80 percent of all life on earth. We turned our scientific probing to the ocean's third dimension: the world beneath the waves, an alien environment almost as boundless as outer space, and as full of mystery.

LTV Aerospace Corporation is in the van of those high-technology enterprises involved in this new era of discovery.

We're applying over 50 years of proven aerospace technology to developing specialized systems for undersea communications and acoustics studies, surveillance, mapping, vehicle instrumentation and hardware recovery. And we're conducting basic ocean research and engineering support services.

Yet, this is only part of our total environmental capability. LTV Aerospace builds advanced products and systems to take up the challenges of *all* environments: land, air, space . . . as well as the forbidding sea.

There is a tremendous range of scientific and engineering talent required throughout our company. Whether your interest is in research, design, development, production or management . . . we have a need and a place for your special skills.

Write Professional Placement, Dept. O, P.O. Box 5907, Dallas, Texas 75222. An equal opportunity employer.

LTV AEROSPACE CORPORATION
OF LING-TEMCO-VOUGHT, INC DALLAS

MATHEMATICAL GAMES

A short treatise on the useless elegance of perfect numbers and amicable pairs

by Martin Gardner

One would be hard put to find a set of whole numbers with a more fascinating history and more elegant properties, surrounded by greater depths of mystery—and more totally useless—than the perfect numbers and their close relatives the amicable (or friendly) numbers.

A perfect number is simply a number that equals the sum of its divisors (including 1 but not the number itself). The smallest such number is 6, which equals the sum of its three divisors, 1, 2 and 3. The next is 28, the sum of $1 + 2 + 4 + 7 + 14$. Early commentators on the Old Testament, both Jewish and Christian, were much impressed by the perfection of those two numbers. Was not the world created in six days and does not the moon circle the earth in 28? In *The City of God,* Book 11, Chapter 30, St. Augustine argues that although God could easily have created the world in an instant, he preferred to take six days because the perfection of 6 signifies the perfection of the universe. (Similar views had been advanced earlier by the first-century Jewish philosopher Philo Judaeus in the third chapter of his *Creation of the World.*) "Therefore," St. Augustine concludes, "we must not despise the science of numbers, which, in many passages of Holy Scripture, is found to be of eminent service to the careful interpreter."

The first great achievement in perfect number theory was Euclid's ingenious proof that the formula $2^{n-1}(2^n - 1)$ always gives an even perfect number if the parenthetical expression is a prime. (It is never a prime unless the exponent n also is prime, although if n is prime, $2^n - 1$ need not be, indeed rarely is, prime.) It was not until 2,000 years later that Leonhard Euler proved that this formula gives *all* even perfects. In what follows, "perfect number" will mean "even perfect number" because no odd perfects are known and they probably do not exist.

To get an intuitive grasp of Euclid's remarkable formula and see how closely it ties perfect numbers to the familiar doubling series 1, 2, 4, 8, 16 . . ., consider the legendary story about the Persian king who was so delighted with the game of chess that he told its originator he could have any gift he wanted. The man made what seemed to be a modest request: he asked for a single grain of wheat on the first square of the chessboard, two grains on the second square, four on the third and so on up the powers of 2 to the 64th square. It turns out that the last square would require 9,223,372,036,854,775,808 grains. The total of all the grains is twice that number minus 1, or a few thousand times the world's annual wheat crop!

If each square of a chessboard is marked according to the number of grains it would hold [*see illustration below*], then, if we remove one grain from a square, the number remaining on that square will fit the parenthetical expression in Euclid's formula. If that number is a prime, multiply it by the number of grains on the preceding square, the 2^{n-1} of the formula. *Voilà,* we have a perfect number! Primes of the form $2^n - 1$ are now called Mersenne primes after the 17th-century French mathematician who studied them. The colored squares in the illustration mark the cells that become Mersenne primes after losing one grain and that consequently provide the first nine perfect numbers.

From Euclid's formula it is not difficult to prove all kinds of weird and beautiful properties of perfect numbers. For example, all perfects are triangular. This means that a perfect number of grains can always be arranged to form an equilateral triangle like the 10 bowling pins or the 15 pool balls. Put another way, every perfect number is a partial sum of the series $1 + 2 + 3 + 4 + \ldots$. It is also easy to show that every perfect

2^{0}	2^{1}	2^{2}	2^{3}	2^{4}	2^{5}	2^{6}	2^{7}
2^{8}	2^{9}	2^{10}	2^{11}	2^{12}	2^{13}	2^{14}	2^{15}
2^{16}	2^{17}	2^{18}	2^{19}	2^{20}	2^{21}	2^{22}	2^{23}
2^{24}	2^{25}	2^{26}	2^{27}	2^{28}	2^{29}	2^{30}	2^{31}
2^{32}	2^{33}	2^{34}	2^{35}	2^{36}	2^{37}	2^{38}	2^{39}
2^{40}	2^{41}	2^{42}	2^{43}	2^{44}	2^{45}	2^{46}	2^{47}
2^{48}	2^{49}	2^{50}	2^{51}	2^{52}	2^{53}	2^{54}	2^{55}
2^{56}	2^{57}	2^{58}	2^{59}	2^{60}	2^{61}	2^{62}	2^{63}

Powers of 2 on a chessboard. Colored squares yield Mersenne primes number except 6 is a partial sum of the series of consecutive odd cubes: $1^3 + 3^3 + 5^3 + \ldots$. Still more surprising, the sum of the reciprocals of the divisors of a perfect number, including the number itself as a divisor, is always 2. For example, the reciprocals of the divisors of 28:

$$\frac{1}{1}+\frac{1}{2}+\frac{1}{4}+\frac{1}{7}+\frac{1}{14}+\frac{1}{28}=2.$$

The digital root of every perfect number (except 6) is 1. (To obtain the digital root add the digits, then add the digits of the result, and continue until only one digit remains. This is the same as casting out nines, and thus to say that a number has a digital root of 1 is equivalent to saying that the number has a remainder of 1 when divided by 9.) The proof involves showing that Euclid's formula gives a number with a digital root of 1 whenever n is odd, and since all primes except 2 are odd, perfect numbers belong to this class. The one even prime, 2, provides the only perfect number, 6, that does not have 1 as its digital root.

Because perfect numbers are so intimately related to the powers of 2, one might expect them to have some kind of striking pattern when expressed in the binary system. This proves to be correct. Indeed, given the Euclidean formula for a perfect number, one can instantly write down the number's binary form. Readers are invited first to determine the rule by which this can be done and then to see if they can show, before the simple proof is given next month, that the rule always works.

The two greatest unanswered questions about perfects are: Is there an odd perfect number? Is there a largest even perfect number? No odd perfect has yet been found; nor, although it looks so easy, has anyone proved that such a number cannot exist. The second question hinges, of course, on whether there is an infinity of Mersenne primes, since every such prime immediately leads to a perfect number. When each of the first four Mersenne primes (3, 7, 31 and 127) is substituted for n in the formula $2^n - 1$, the formula gives a higher Mersenne prime. For more than 70 years mathematicians hoped this procedure would define an infinite set of Mersenne primes, but the next possibility, $n = 2^{13} - 1 = 8{,}191$, let them down: in 1953 a computer found that $2^{8,191} - 1$ was not a prime. No one knows whether the series of Mersenne primes continues forever or has a highest member.

Øystein Ore, in his *Number Theory and Its History*, quotes an amusing prediction from Peter Barlow's 1811 book *Theory of Numbers*. After giving the ninth perfect, Barlow adds that it "is the greatest that will ever be discovered, for, as they are merely curious without being useful, it is not likely that any person will attempt to find one beyond it." In 1876 the French mathematician Édouard Lucas, who wrote a classic four-volume work on recreational mathematics, announced the next perfect to be discovered, $2^{126}(2^{127} - 1)$. The 12th Mersenne prime, on which it is based, is one less than the number of grains on the last square of a *second* chessboard, if the doubling plan is carried over to another board. Years later Lucas had doubts about this number, but eventually its primality was established. It is the largest Mersenne prime to have been found without the aid of modern computers. The illustration at the left lists formulas for the 23 known perfects and the number of digits in each, and gives the numbers themselves until they get too large. The last perfect—which has 22,425 divisors—came to light in 1963 when a computer at the University of Illinois discovered the 23rd Mersenne prime. (The university's mathematics department is so proud of this that its postage meter has been stamping on envelopes a rectangle bearing the statement: "$2^{11213} - 1$ is prime.") The printout of this largest prime known [*see opposite page*] is from John McKay, who used a computer at the Atlas Computer Laboratory of the British Science Research Council and a fast program devised by Fred Lunnon of the University of Manchester.

The end digits of perfect numbers present another tantalizing mystery, appropriately discussed in the year 1968 since it concerns the digits 6 and 8. It is quite easy to prove from Euclid's formula that any even perfect must end

	FORMULA	NUMBER	NUMBER OF DIGITS
1	$2^1(2^2-1)$	6	1
2	$2^2(2^3-1)$	28	2
3	$2^4(2^5-1)$	496	3
4	$2^6(2^7-1)$	8,128	4
5	$2^{12}(2^{13}-1)$	33,550,336	8
6	$2^{16}(2^{17}-1)$	8,589,869,056	10
7	$2^{18}(2^{19}-1)$	137,438,691,328	12
8	$2^{30}(2^{31}-1)$	2,305,843,008,139,952,128	19
9	$2^{60}(2^{61}-1)$		37
10	$2^{88}(2^{89}-1)$		54
11	$2^{106}(2^{107}-1)$		65
12	$2^{126}(2^{127}-1)$		77
13	$2^{520}(2^{521}-1)$		314
14	$2^{606}(2^{607}-1)$		366
15	$2^{1,278}(2^{1,279}-1)$		770
16	$2^{2,202}(2^{2,203}-1)$		1,327
17	$2^{2,280}(2^{2,281}-1)$		1,373
18	$2^{3,216}(2^{3,217}-1)$		1,937
19	$2^{4,252}(2^{4,253}-1)$		2,561
20	$2^{4,422}(2^{4,423}-1)$		2,663
21	$2^{9,688}(2^{9,689}-1)$		5,834
22	$2^{9,940}(2^{9,941}-1)$		5,985
23	$2^{11,212}(2^{11,213}-1)$		6,751

The 23 known perfect numbers

```
EXECUTION STARTED ON 25/09/67   AT   17.17.29  WITH  THE INSTRUCTION COUNTER READING      248

                                       2 81411 20136 97373 13339 31529 75842 58419 18186 62382 01360 07878 92419 34934 55151 76682
27631 38107 15094 74563 32570 74198 78930 85350 71537 34244 50164 18881 80178 93905 48709 41439 18572 57571 56575 87064
78418 35674 70706 74633 49718 80530 50875 41682 16243 25680 55582 60711 10691 94660 74608 73056 96536 08305 71590 24277
49342 26866 18396 63091 85433 46251 45374 84258 65598 23862 35046 02922 75078 01410 90716 33484 39547 78109 33972 60096
90967 70918 43944 55575 42211 15477 34376 02069 79650 06708 78849 93478 01297 72778 78532 80743 22345 54020 93157 18023
10429 92316 75884 32457 03610 41108 50960 43976 90384 50365 51402 23496 25383 66575 12071 69661 69735 27322 36111 92684
64547 51701 73452 70113 79148 17510 78208 21297 62894 67956 31098 96076 74922 50494 83425 40733 34414 12162 78339 39461
53921 25289 32010 72613 66892 93688 81566 54916 71395 17471 04526 63709 17575 36037 74156 85576 65153 13827 61372 72816
96692 63352 96663 63787 28653 97699 41609 10777 71835 93336 00268 01245 17633 45149 04395 98324 82383 64572 51219 40639
14326 35639 22560 45560 42396 00430 77993 61927 37990 05864 00420 76309 23208 13392 26249 29420 76312 93326 80338 18471
55525 58206 39308 88994 86655 70202 40381 58563 13578 94977 97670 46261 84532 79567 25767 28920 52623 11752 01478 62478
13331 83401 50844 75386 76052 66122 17340 57972 12374 14485 80372 53554 63022 00953 63010 08145 86752 47046 04618 86203
90935 55206 19532 82409 51895 10704 07932 84825 09546 25301 51872 82399 71717 64140 66331 58043 09008 61194 25783 80931
06474 89915 94407 47632 84377 85848 82542 39211 70614 93829 40294 83257 16297 92993 88940 69587 73754 48948 08110 83452
93394 32780 84527 29789 83413 51401 93912 41966 17994 88795 21032 82381 12742 21870 06345 41149 74365 72872 32843 42636
93488 04878 99347 19624 03393 96785 76761 50371 60019 66502 52168 25011 77931 78488 01200 05054 22821 36255 05205 09209
72445 98958 52366 82747 78516 19190 50325 48531 15029 40313 21789 89005 19575 11943 01340 27728 27303 90683 65112 05878
95060 19875 31218 82187 78865 70240 07291 78418 65185 89977 78851 03067 43945 89610 86452 58766 41569 28256 64174 47061
61533 05144 85227 38845 49635 05925 54106 06458 42732 38641 09506 68763 63144 47514 26909 49329 53219 92421 25946 95157
65500 91585 21173 42092 32758 82063 32762 54086 17963 03296 20335 72563 55360 40560 97832 11154 75359 08988 43381 69197
47615 81716 16066 20557 30700 03771 94730 01343 18155 60750 15902 78421 64901 42254 45712 24546 93679 32349 70894 95466
84254 36412 34778 53761 94310 03013 90805 68383 42077 26286 18722 64610 97075 06566 92810 28000 33941 70434 39919 62002
05979 45655 27774 91388 32377 56792 72006 55437 68640 79217 74415 59278 27235 08230 92843 68353 43966 79150 22967 61018
34243 78782 04200 87274 02861 72126 84576 38873 36057 69491 22410 98665 92577 36066 62414 67280 15898 86055 23486 34588
08822 27855 50570 63092 76349 41503 45476 77180 61829 63528 66263 00550 92222 54318 45976 81941 26727 60304 74603 44175
58102 92983 20171 22635 52344 39676 81630 99191 27574 20633 48077 19021 87541 38915 80871 52904 91878 29308 41213 34009
10419 75631 30215 40478 43660 41784 46757 73899 86320 83586 20799 22340 85162 63437 54067 71169 70732 32139 88284 94377
91221 71985 95360 58979 02291 78176 82865 48287 87818 04150 60635 46004 71641 04095 48377 72017 37468 87332 40685 50430
69582 62103 04316 33638 53113 84093 49002 13323 72463 45337 39774 27405 89667 38275 44203 12857 48745 81960 33523 20056
37229 31959 23692 88171 37527 67022 60450 91173 50695 04025 01666 77552 14932 07364 36541 99488 47701 03639 09372 00575
78999 89580 77577 51266 21113 05790 57174 49417 22201 60705 30243 91611 67059 90451 30425 62063 18289 29773 83030 95152
43054 97722 39514 96482 16018 38628 86144 63019 36017 71054 67775 03109 26303 09947 47397 61857 62073 73447 72544 14271
35362 42636 08636 69327 15763 59830 45447 97181 67188 01639 86954 75251 46305 65557 18437 17916 87566 91403 20724 97856
85867 18527 58660 24396 02335 28351 39449 80064 32703 02781 04224 14497 18836 80541 68978 47962 67391 47608 76963 92191

END OF JOB.
```

The 23rd Mersenne prime, $2^{11,213} - 1$

in 6 or 8. (If it ends in 8, the preceding digit is 2; if it ends in 6, the preceding digit must be 1, 3, 5 or 7 except in the cases of 6 and 496.) The ancients knew the first four perfects—6, 28, 496 and 8,128—and from them rashly concluded that the 6's and 8's alternated as the series continued. Scores of mathematicians from ancient times through the Renaissance repeated this dogmatically, without proof, particularly after the fifth perfect number (first correctly given in an anonymous 15th-century manuscript) turned out to end in 6. Alas, so does the seventh. The series of terminal digits for the 23 known perfects is

6, 8, 6, 8, 6, 6, 8, 8, 6, 6, 8, 8,
6, 8, 8, 8, 6,6, 6, 8, 6, 6, 6.

The sequence contains infuriating hints of order. The first four digits alternate 6 and 8, then 66 and 88 alternate for eight digits, followed by a meaningless 6, then 888 and the palindromic 6668666. Are the digits trying to tell us something or is all this accidental? If we partition the series into triplets, starting at the left, no triplet contains three of a kind. If this pattern continues, the next perfect should end in 28, but so far no one has found a reliable rule for predicting the last digit of the next, undiscovered perfect. It is easy, of course, to determine the terminal digit of any perfect number if you know its Euclidean formula. Can the reader work out a simple rule?

Amicable numbers derive from an obvious generalization of the perfects. Suppose we start with any number, add its divisors to obtain a second number, then add the divisors of *that* number and continue the chain in the hope of eventually getting back to the original number. If the first step immediately restores the original number, the chain has only one link and the number is perfect. If the chain has two links, the two numbers are said to be amicable. Each is equal to the sum of the divisors of the *other*. The smallest such numbers, 220 and 284, were known to the Pythagoreans, who considered them symbols of friendship. Biblical commentators spotted 220, the lesser of the pair, in Genesis 32:14 as the number of goats given Esau by Jacob. A wise choice, the commentators said, because 220, being one of the amicable pair, expressed Jacob's great love for Esau. During the Middle Ages this pair of numbers played a role in horoscope casting, and talismans inscribed with 220 and 284 were believed to promote love. One poor Arab of the 11th century recorded that he once tested the erotic effect of *eating* something labeled with 284, at the same time having someone else swallow 220, but he failed to add how the experiment worked out.

It was not until 1636 that another pair of amicable numbers, 17,296 and 18,416, were discovered by the great Pierre de Fermat. He and René Descartes independently rediscovered a rule for constructing certain types of amicable pairs—a rule they did not know had previously been given by a ninth-century Arabian astronomer. Using this rule, Descartes found a third pair: 9,363,584 and 9,437,056. In the 18th century Euler drew up a list of 64 amicable pairs (two of which were later shown to be unfriendly). Adrien Marie Legendre found another pair in 1830. Then in 1867 a 16-year-old Italian, B. Nicolò I. Paganini, startled the mathematical world by announcing that 1,184 and 1,210 were friendly. It was the second lowest pair and had been completely overlooked until then! Although the boy probably found it by trial and error, the discovery put his name permanently into the history of number theory.

Today more than 600 amicable pairs are known, many of the numbers running to more than 30 digits. J. Alanen, Øystein Ore and J. Stemple, in their paper "Systematic Computations on Amicable Numbers" (*Mathematics of Computation*, Vol. 21, No. 98, pages 242-245; April, 1967), list the 42 pairs of amicables in which the smaller number is less than 1,000,000. The 66 amicables between 1,000,000 and 10,000,000 were tabulated last year by Paul Bratley and John McKay with a computer at the University of Edinburgh, but their list (which includes 38 new pairs) has not yet been published. The list at the bottom of the next page is of all pairs smaller than 100,000. The last pair was a 1964 computer discovery by Howard L. Rolf of Baylor University.

Postage-meter stamp honoring the largest prime

All known amicable pairs consist of two even numbers or (more rarely) two odd numbers. No one, however, has yet proved that a pair of mixed parity is impossible. Bratley and McKay conjecture that all odd amicable numbers are multiples of 3, but this too has not yet been proved. All even pairs have a sum that is a multiple of 9, another strange fact no one seems to understand. There is no known formula for generating all amicable pairs, nor has it been established whether the number of such pairs is infinite or finite.

If the chain that leads back to the original number has more than two links, the numbers have been called "sociable." Only two sociable chains are known. In 1918 a French mathematician, P. Poulet, announced a chain of five links (12,496, 14,288, 15,472, 14,536, 14,264) and a truly astounding chain of 28 links (a perfect number) that starts with 14,316 (move 3 to the front and you have pi to four decimals!).

The big unsolved question is whether a "crowd"—a three-link chain—exists. Alanen, Ore and Stemple, in the paper cited above, disclose that if there is a crowd, its lesser number must be more than 1,000,000. At Edinburgh, Bratley and McKay are making extensive computer sweeps for crowds, so far without success. Such searches will surely continue, useless though crowds may be, until such a triple chain is encountered or some clever number theorist proves its impossibility.

In last month's illustration of free trees with seven points, the two duplicates are trees 5 and 8. The second starting position for clock solitaire is a loss. Its graph is not a tree; not only is it disconnected but also one part contains a circuit.

The illustration at the top of page 126 shows how to draw minimal-length trees joining the corners of a square and of a regular pentagon. The dotted angles are 120 degrees. It might be thought that the two diagonals of a square would provide an "economy tree" (length $2\sqrt{2} = 2.828+$) connecting a unit square's corners, but the network shown has a length of $1 + \sqrt{3} = 2.732+$. A proof without calculus that this is minimal is given in Hugo Steinhaus' *100 Problems in Elementary Mathematics*, Problem 73 (Basic Books, 1964). The minimal path inside a pentagon of unit side is 3.891+.

The minimum tree inside an equilateral triangle has a fourth point in the triangle's center. Minimum trees on regular polygons of six or more sides are simply the perimeter with one side removed. For the general problem of finding minimal networks connecting n points on the plane—and a technique for finding such networks by capitalizing on the surface tension of a film of soap solution—see Chapter 7 of *What Is Mathematics?*, by Richard Courant and Herbert E. Robbins (Oxford, 1941).

It is impossible to do justice to the many readers who commented on the short problems given in November. Several dozen letters pointed out that the question about the size of the average family in Fertilia (where parents stop having children after the first boy) can be answered without summing an infinite series as explained in December. After one has proved that the ratio of boys to girls always remains one-to-one, it follows that in the long run there will be as many boys as there are girls. Since each family has exactly one son, there will be, on the average, also one girl, making an average family of two children in Fertilia.

In giving C. Dudley Langford's problem of arranging n pairs of objects in a row so that one object separates one pair, two objects separate another pair, and so on to n objects between the nth pair, I said that for $n = 7$ there were 25 distinct solutions, not counting reversals as being different. This figure, taken from a journal reference cited in December, is short by one. James Bartow and Gerald K. Schoenfeld, each working by hand, found 26 solutions. This was confirmed by five separate computer programs written by Malcolm C. Holtje, John Miller, Robert A. Smith, Glenn F. Stahly and Thomas Starbird. The first four programs also found 150 solutions for $n = 8$. There are no solutions unless n is a multiple of 4 or one less. E. J. Groth, an engineer with Motorola, Inc., at the Aerospace Center in Scottsdale, Ariz., used a program coded in FORTRAN to extend the count to 17,792 sequences for $n = 11$ and to 108,144 for $n = 12$. The printout for the two sets of solutions is a stack of about 3,000 sheets.

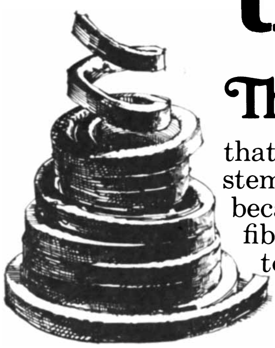
The problem of cutting and folding a three-inch square to cover the largest possible cube was answered correctly if overlapping is not allowed. I did not add this proviso, however, not dreaming that overlapping would permit better solutions. John H. Halton, a mathematician at the University of Wisconsin, was the first to send a cut-and-fold technique by which one can approach as closely as one wishes to the ultimate-size cube with a surface area equal to the area of the square sheet! (Three readers, David Elwell, James F. Scudder and Siegfried Spira, each came close to such a discovery by finding ways to cover a cube larger than the one given in the answer, and George D. Parker found a complete solution that was essentially the same as Halton's.)

Halton's technique involves cutting the square so that opposite sides of the cube are covered with solid squares and the rest of the cube is wrapped with a ribbon folded into one straight strip so that the amount of overlap can be as small as one wishes. In wrapping, the overlap can again be made as small as one pleases simply by reducing the width of the ribbon. Assuming infinite patience and paper of infinitely small molecular dimensions, as Halton put it, this procedure will cover a cube approaching as

1	220	284
2	1,184	1,210
3	2,620	2,924
4	5,020	5,564
5	6,232	6,368
6	10,744	10,856
7	12,285	14,595
8	17,296	18,416
9	63,020	76,084
10	66,928	66,992
11	67,095	71,145
12	69,615	87,633
13	79,750	88,730

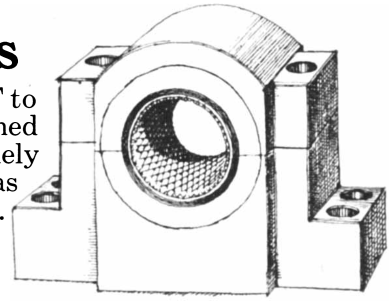
Amicable pairs with five or fewer digits

Teflon® is the resin that makes the fiber that makes:



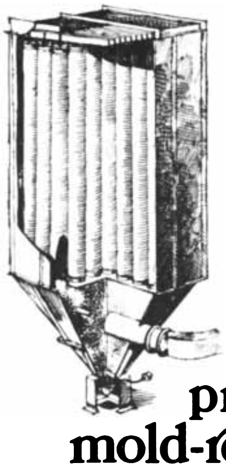
The pump packing

that seals rotating and reciprocating pump shafts and valve stems with minimum leakage and no glazing. It's self-lubricating because it's 100% TEFLON fiber, not just a coating on another fiber. TEFLON fiber is soft and flexible, and it holds its shape—to insure perfect functioning from –400°F to 550°F.



The bearings

that take up to 60,000 psi. at temperatures from –220°F to 500°F—without lubrication! Like Teflon* resin, bearings lined with a fabric of Teflon† fiber have an extremely low coefficient of friction—which gets even lower as load and/or temperature increase.



The filters

that work far longer than the filters they replace. No wonder. Filter fabrics of TEFLON TFE fiber are inert to virtually all chemical attack over the entire pH range. They withstand temperatures up to 500°F, have excellent cake release, high tensile strength and zero moisture absorption. Bags of TEFLON fiber are now used for dust collection in many industries. Applications include manufacture of carbon black, pigments, organic chemicals and fly-ash collection. Because of its outstanding properties—chemical inertness, temperature resistance, slipperiness—TEFLON fiber is also ideal for making

protective clothing, wicking felts, mold-release fabrics, gaskets, sewing thread, electrical lacing cord and leader tapes.

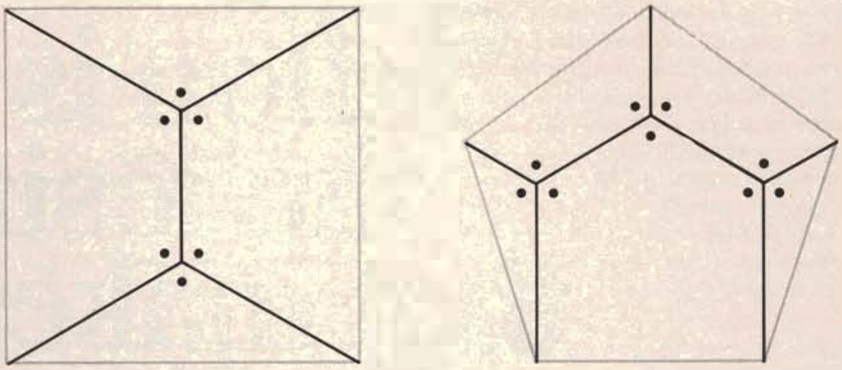
What could TEFLON fiber do for you? For more information write Du Pont Company, Teflon Fiber Marketing, Department B-1, Centre Road Building, Wilmington, Delaware 19898.

*Du Pont registered trademark for TFE fluorocarbon resin. †Du Pont registered trademark for TFE fluorocarbon fiber.

TEFLON® FIBER

DU PONT REG. U.S. PAT. OFF.

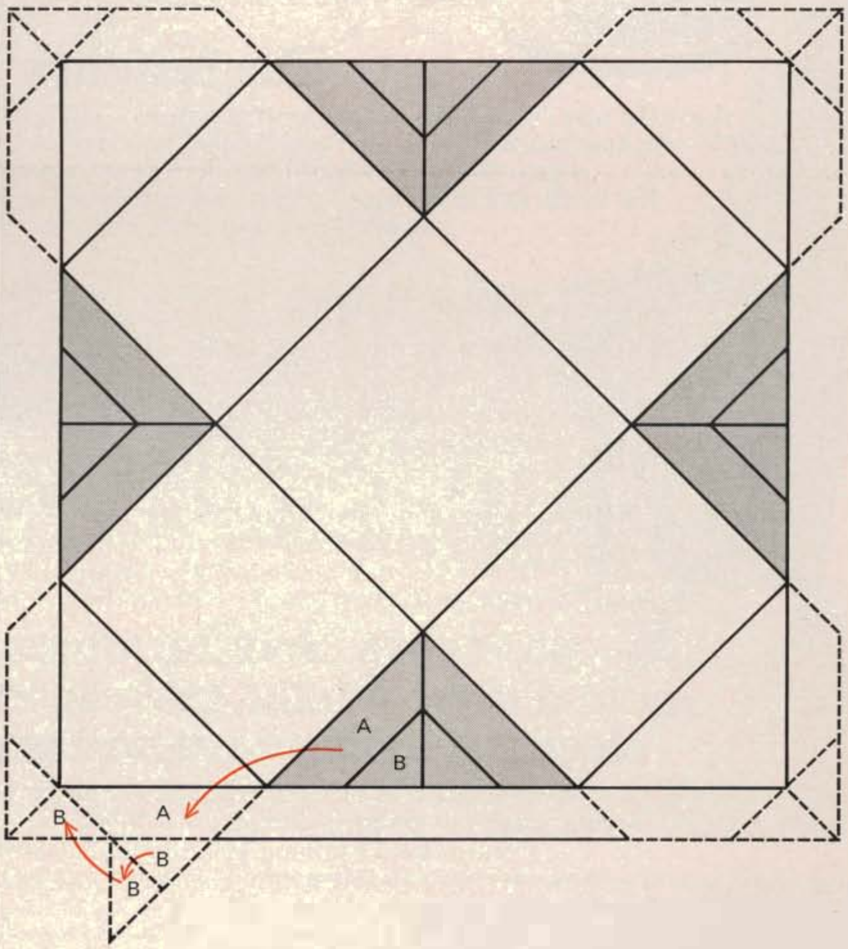
Better things for better living...through chemistry



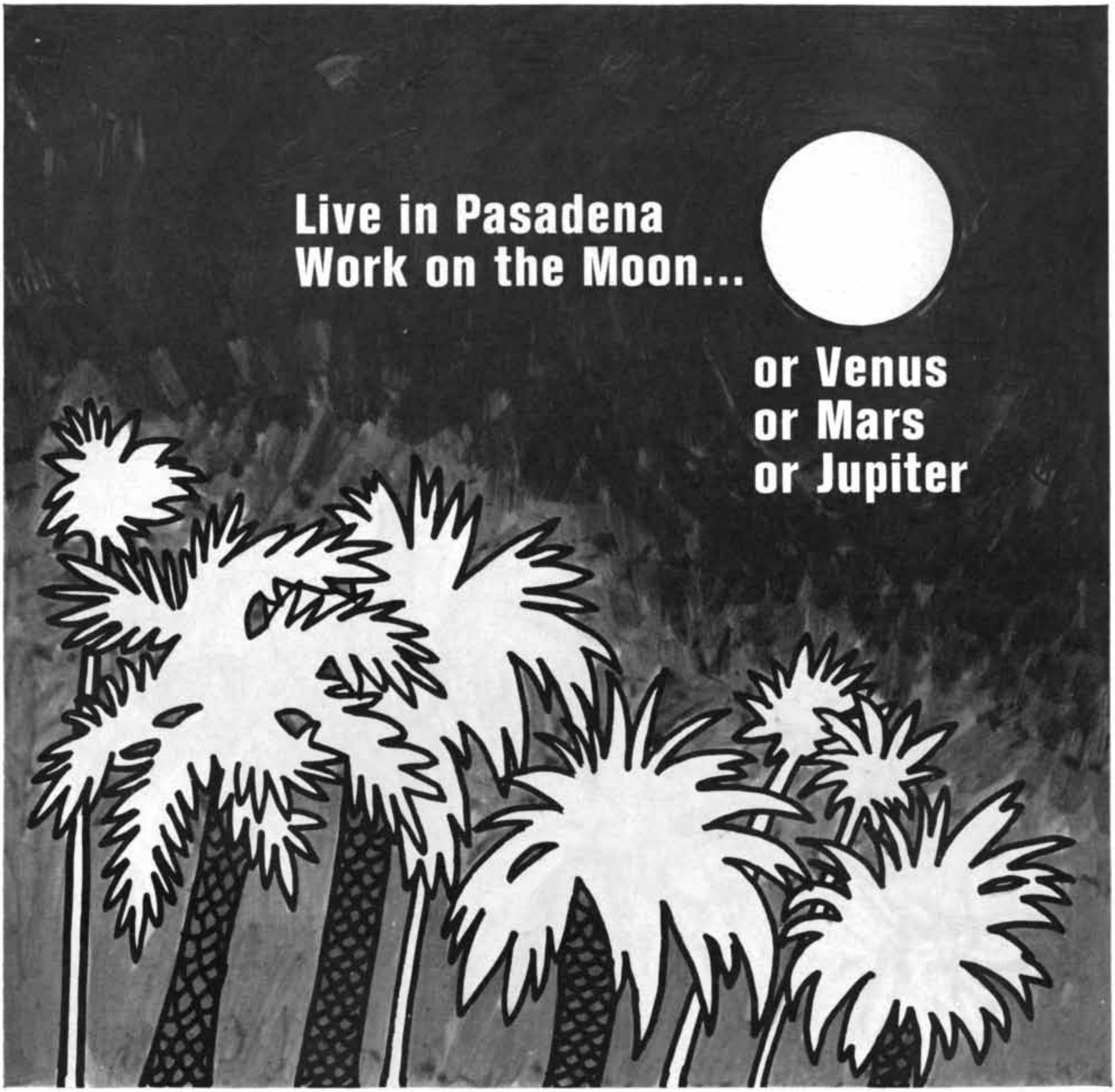
Economy trees connecting corners of square and pentagon

closely as desired to the limit, a cube of side $\sqrt{3}/\sqrt{2}$.

Fitch Cheney, a mathematician at the University of Hartford, found another way to do the same thing by extending the pattern given in December for covering the maximum cube without overlapping. By enlarging the center square as shown below, he makes the four surrounding squares rectangles and the four corner triangles correspondingly larger. The shaded areas, cut as shown, can be folded as indicated to extend each corner triangle. (*A* turns over once, *B* three times.) The result: a pattern exactly like the one given in December except that it is larger. Since the unavoidable overlap can be made as small as one wants, it is clear that this method also approaches the $\sqrt{3}/\sqrt{2}$ cube as a limit.



Cheney method of folding a square to cover a cube



And Surveyor VII has helped solve the commuting problems.

Right now, the men at Jet Propulsion Laboratory in Pasadena, California, are busy analyzing the Moon-data collected and sent back to Earth by Surveyor VII. Within minutes after its successful soft-landing on January 9, JPL's unmanned research spacecraft was shooting TV pictures under new lighting conditions and scooping up lunar soil from the most southerly landing site yet.

This mission was the last in the series of seven Surveyor flights and was probably the most difficult lunar surface probe ever made by the United States. Landing less than two kilometers from the target area near the crater Tycho in the rugged highlands, Surveyor VII was the first to investigate the Moon's "Deep South." Four of the previous successful Surveyors landed in the relatively smooth equatorial belt. They proved the lunar soil in that area is in many respects just like Earth soil and that it will support later astronaut landings. The most recent flight indicates that the southern soil should also support astronaut landings.

Admittedly, space experts don't know when men will actually go to the Moon. It's missions like Surveyor VII that land, dig, probe, photograph, analyze and accurately report the findings back to Earth which bring this date nearer. And it's men like those at JPL in Pasadena who work with, probe and analyze these findings (to solve outer space commuting problems) who bring flying to the Moon, Venus, Mars or anywhere else out there a lot closer. Send your resume, in confidence, to Mr. Wallace Peterson, Supervisor, Employment. He'll personally help solve your commuting problems.

JET PROPULSION LABORATORY

jpl 4802 Oak Grove Dr., Pasadena, Calif. 91103
Attention: Professional Staffing Dept. 3

"An equal opportunity employer." Jet Propulsion Laboratory is operated by the California Institute of Technology for the National Aeronautics and Space Administration.

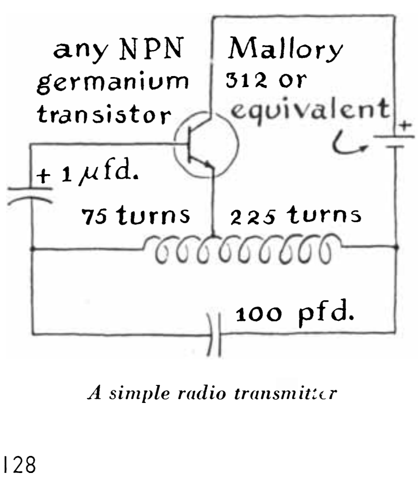
THE AMATEUR SCIENTIST

Little radio transmitters for short-range telemetry

Conducted by C. L. Stong

An amateur who builds a radio transmitter is likely to measure his success and pleasure in terms of the transmission range he achieves. In such terms there is more pleasure in talking with another "ham" across the country than with one in the next county, even if the more distant operator turns out to be a rather dull fellow. The aim of reaching out ever farther has spurred amateur radio operators to extend the range of their apparatus steadily; it was about five miles in 1900, rose to 5,000 miles by 1920 and reached 12,000 miles in 1930. The ultimate in distance was approached recently when an amateur in Finland communicated with one in California by means of signals bounced off the moon. The round trip was some 500,000 miles.

What next? R. Stuart Mackay, a physicist and biologist at Boston University, suggests that radio experimenters shift their enthusiasm to short-range telemetry. Here apparatus the size of a vitamin capsule radiates signals over distances measured in feet or even inches. The challenge lies in making the most of the least.



A simple radio transmitter

Most miniature transmitters carry devices for sensing one or more variables that may include temperature, pressure, displacement, vibration, acceleration, acidity or alkalinity, bioelectric potential, fluid flow, radioactivity, light, color and so on. Miniature transmitters, equipped with appropriate sensors, have been shot into the tops of trees, dropped into deep water, swallowed by various animals, enclosed in vacuum systems, ovens, reaction vessels and the rotors of centrifuges, cemented to the backs of turtles and tied around the necks of grizzly bears. Variations in the design and use of the little transmitters appear to be limited only by the imagination of experimenters.

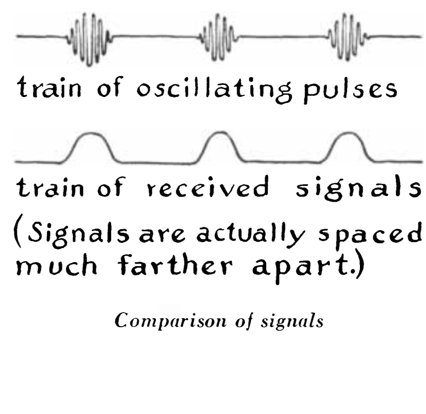
Short-range radio telemetry is fairly new. When I asked Mackay for the best way to get started, he lent me the manuscript of his forthcoming book: *Bio-Medical Telemetry* (John Wiley & Sons, 1968). The following discussion is based on that information and on my own subsequent experiments. I learned that one needs little experience in electronics to build successful transmitters but that the difficulty of construction varies inversely with the size of the units. When you attempt to make pea-sized transmitters, you develop admiration for the kind of craftsman who makes and repairs watches.

The electrical circuits and basic parts of the small transmitters are like those in conventional equipment in all respects except size. The equipment includes resistors, capacitors, transistors, dry cells, quartz crystals, coils of copper wire and so forth. All parts except the coils can be bought in miniature sizes. The coils must be wound by the experimenter. Extensive use is made of plastics for forms on which the coils are wound and for enclosing the apparatus. Transmitters that are to be immersed in fluids or swallowed by animals are first coated with wax and then covered with silicone rubber. The units are energized by mercury cells of the kind used in small hearing aids and electric wristwatches.

I chose the simplest transmitter for my first experiment. The device is easy to make, has interesting applications and usually works on the first try. The circuit is known to electronic engineers as a Hartley oscillator [*see illustration at bottom left on this page*]. The coil that is illustrated in the lower part of the drawing and the capacitor shown at the bottom form a resonant circuit in which electrical oscillations occur at a characteristic frequency. A charge of electrons, when placed on one side of the capacitor, tends to distribute itself equally on both sides by moving through the coil from the charged to the uncharged side. The flow creates a magnetic field around the coil.

It would seem reasonable to suppose that the flow would stop after the charge became equally divided. The magnetic field is now present, however, and it must collapse. In collapsing the field cuts through the turns of the coil and generates an electric force that keeps the electrons moving toward the formerly uncharged side. The situation soon becomes as unbalanced as it was in the first place, so that the action reverses. The electrons rush back to their initial position and then are ready for the next cycle.

Were it not for the loss of some energy during each transit, the electrons would slosh back and forth endlessly. However, some of the energy is spent during each transit in heating the coil and the capacitor. Another portion of energy is radiated by the coil in the form of radio waves. Still another part is



Comparison of signals

lost when the undulating magnetic field that surrounds the coil generates alternating currents in nearby conductors.

It is this last part of the energy that most interests those who experiment with short-range telemetry, because the energy is strong enough to actuate a nearby radio receiver. The oscillations gradually die away, much as a playground swing eventually comes to rest. A swing vibrates at a characteristic frequency that depends on the length of the swing and the strength of the force of gravitation. Similarly, electrical oscillations in a resonant circuit vibrate at a frequency that is determined by the size of the coil and the capacitor.

Swings can be kept going by a series of periodic pushes. Electrical oscillations can likewise be kept going in a resonant circuit by adding charge to the capacitor each time the current sloshes back and forth. Periodic charge can be added by connecting the capacitor momentarily to a dry cell with a fast-acting, automatic switch. A transistor can function as a switch of this kind.

The dry cell is connected to the capacitor through two electrodes of the transistor—the emitter and the collector. Normally the path between these electrodes resembles an open circuit. If a weak current is made to flow in one direction through the emitter and the third electrode of the transistor, called the base electrode, the switch is "on," but when the voltage reverses, the switch turns "off." The dry cell is then connected to the capacitor for a moment until the current again reverses and turns the battery off. In the transmitter under discussion the alternating current that activates the transistor is taken from part of the coil (through a tap that is connected to the emitter) and is fed into the base of the transistor through a one-microfarad capacitor.

This circuit is designed to oscillate for a brief interval, turn off for a much longer interval, oscillate again and so on. The pulsing action arises from another characteristic of the transistor, which is that the transistor contains an internal junction between the base and the emitter; the junction acts as a one-way valve for electric current. Current can pass from the base to the emitter, but it encounters resistance in the opposite direction.

Because of this rectifying action, charge gradually accumulates in the one-microfarad capacitor. After some hundreds of oscillations the charge becomes so large that it suppresses the switching action. Oscillation then stops while the accumulated charge slowly leaks through the transistor.

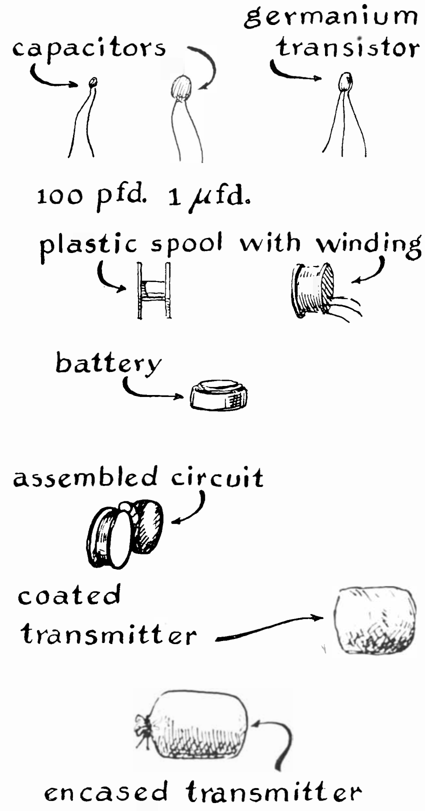
Some transistors are leakier than others. Those made of germanium leak at a rate that varies with the temperature of the device. Indeed, germanium transistors are so responsive to changes in temperature that they make excellent thermometers. The circuit under discussion includes a thermometer of this kind. It has been used for telemetering the internal temperature of things as diverse as the rotor of a centrifuge, a batch of setting concrete and the gastrointestinal tracts of animals.

The signals can be picked up by an ordinary transistor radio placed within a few feet of the coil of the transmitter, where the pulsating magnetic field carries the information. Each burst of oscillations is heard as a distinct click. The frequency of the clicks increases with the leakage of the transistor, and the leakage increases with temperature. The unit can be calibrated in terms of temperature by placing the transmitter and a thermometer in a beaker of cold water and making a two-column table that lists the frequency of the clicks and the corresponding temperature as the water is warmed. A calibration graph is then drawn by plotting click rate against temperature.

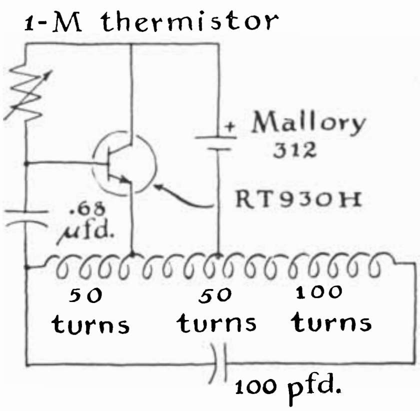
The clicks can be timed with a stopwatch. When a count is being made, accuracy will be improved if the observer swings his arm up and down in unison with the clicks. This trick enables him to continue the count if he is momentarily distracted by a burst of incidental noise. Always assign "zero" to the first click and simultaneously start the watch: count "zero," "one," "two" and so forth.

My first transmitter was about the size of a pack of cigarettes. Smaller versions were made later. Any germanium transistor of the NPN type will work in the circuit illustrated at bottom left on the opposite page. Any germanium transistor of the PNP type can also be made to work by reversing the connections to the dry cell. The coil can be made of insulated copper wire of any size and should be slop-wound. I just wrapped it around a pencil. Do not wind the coil as a neat, single layer.

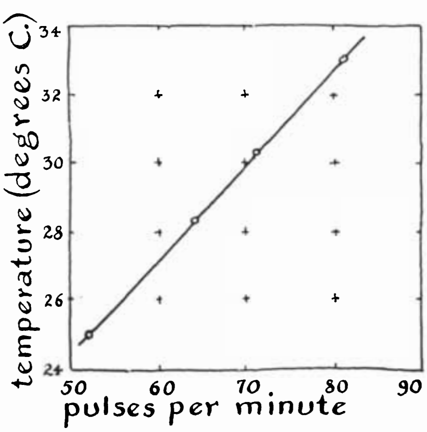
The tap to which the emitter connects can be brought out of the coil as a short loop. The completed coil can be coated with epoxy and should be slipped from the form on which it was wound. I wound several coils on tiny spools, but it was a needless waste of time. When soldering transistors into a circuit, clamp the leads between the jaws of a pair of



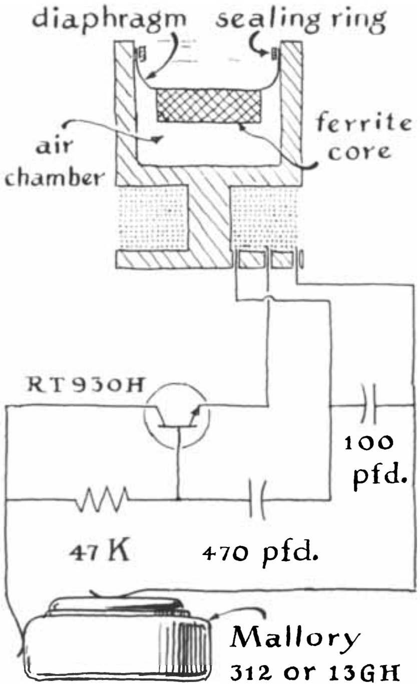
Elements of transmitter



Temperature-sensing transmitter



Transmitter calibration graph



Pressure-sensing transmitter

long-nosed pliers at a point between the device and the hot joint. The jaws act as a heat sink and keep the transistor cool.

Soldered connections can be insulated with a dab of quick-drying cement. The physical layout of the parts is not critical. Transmitters of the smallest size are made by sandwiching the capacitors and the transistor between the coil and the dry cell [*see top illustration on preceding page*].

Coils of the smallest size are wound on a Teflon form. One end of a Teflon rod, approximately 1/4 inch in diameter, is machined into a cone that is about 1/4 inch long and 1/16 inch in diameter at the base and tapers toward the end about one degree. The conical projection should be polished. A perforated disk cut from the same stock slips onto the tapered projection and should make a tight fit at a point about 3/32 inch from the shoulder of the rod. The coil is wound in the space between the inner face of the disk and the shoulder. Epoxy cement is applied as the wire is wound. The coil is slipped off the form when the epoxy has hardened. (Epoxy does not stick to Teflon.)

When winding coils of 40-gauge wire (.003 inch in diameter) or less, uncoil the wire from the end of the spool. Attempts to pull off the wire by rotating the spool invite a break. When this happens, it is almost impossible to locate the broken end. Insulated wire of 40 gauge and other components for the transmitter can be bought from Allied Radio, 100 North Western Avenue, Chicago, Ill. 60612.

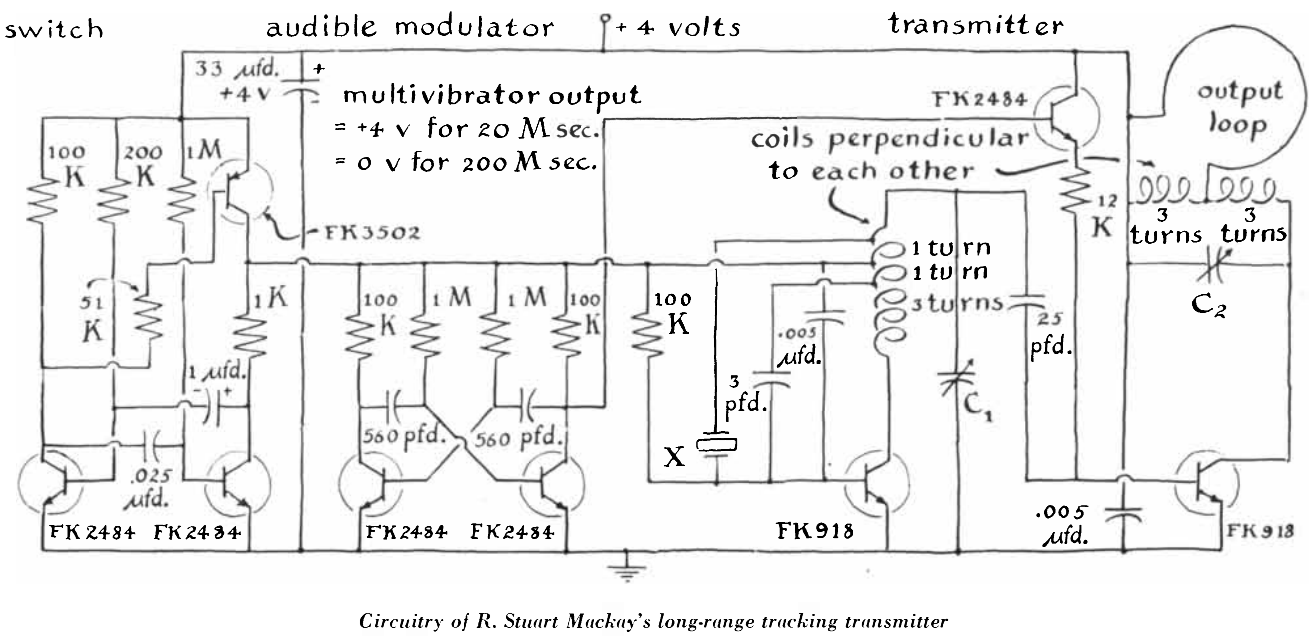
Connections can be spot-welded directly to mercury cells (see "The Amateur Scientist"; SCIENTIFIC AMERICAN, November, 1966). A miniature soldering iron for making minute splices can be improvised by brazing a short length of 12-gauge copper wire to the replaceable tip of an ordinary electric soldering iron. A hot knife, useful for opening epoxy housings, can be similarly improvised by brazing the blade from a pocketknife to a replaceable soldering tip. Fully assembled transmitters are first embedded in epoxy and then coated with wax. If they are to be swallowed, they should also be enclosed in the tip cut from the finger of a rubber glove and tied with nylon thread, as illustrated.

The performance of the transmitter can be altered in a number of ways. For example, the frequency of oscillation can be increased by substituting a smaller capacitor for the 100-picofarad unit specified. (A picofarad is a millionth of a microfarad.) The frequency can also be increased by winding a coil of fewer turns. (The position of the tap should be shifted in proportion toward the near end of the coil.)

The frequency of the clicks can be increased or decreased by using a smaller or larger capacitor in the position of the specified one-microfarad capacitor. The shape of the magnetic waves emitted by the oscillator can be altered either by shifting the position of the tap on the coil or by inserting a variable resistor of approximately 600 ohms in series with the lead that connects to the emitter. The strength of the emitted signal can be increased by applying a higher battery voltage to the unit.

The time required for the transmitter to respond to changes in temperature can be reduced by adding a separate thermistor to the circuit. The thermistor must be added if a silicon transistor is used [*see middle illustration on preceding page*]. In the transmitter shown the battery was connected to the center of the coil to increase the strength of the signal.

For the quickest response to changes in temperature the thermistor is mount-



Circuitry of R. Stuart Mackay's long-range tracking transmitter

The Hasselblad EL
it could be one of the most important research tools you ever buy.

In previous advertisements we have discussed the many and varied applications for which the Hasselblad System can be used in the scientific and industrial field. We would now like to discuss a unique combination of Hasselblad components and some of the rather unique applications to which they can be put.

The camera in question is the Hasselblad EL, an electrically driven 2¼″ square, single lens, reflex camera, powered by one or two rechargeable batteries, each battery good for 1000 exposures on a single charge. The film is wound on and the shutter cocked automatically after each exposure. Exposures can be made manually or by remote control, using either long release cables or a radio release.

The Hasselblad EL accepts practically all the accessories that are available for the Hasselblad 500C (the standard body in the the Hasselblad System), including the 120-12 exposure magazine, the 220-24 exposure magazine, and, of particular interest with the EL, the 70mm-70 exposure magazine. The Hasselblad EL also accepts all seven lenses available in the Hasselblad System, from the extreme wide angle Zeiss Distagon of 40mm focal length, 88° angle of view, with maximum aperture f/4, to the Zeiss Tele-Tessar of 500mm focal length, 9° angle of view, maximum aperture f/8.

Listed below are five particular and diverse applications for the EL.

General Instrument Recording

Many Hasselblad EL cameras are already proving their worth, in industrial and research institutions all over the world, as recording devices for the constant surveillance of instrument banks and oscilloscope screens on a 24 hour basis.

By the use of the EL with a lens of the appropriate focal length, and the 70 exposure 70mm film magazine, banks of cameras, using the Hasselblad remote control timer, can make a number of exposures between 2 and 60 intervals for each of 3 time ranges—seconds, minutes or hours.

Thus, many valuable man hours can be saved which would otherwise be wasted making manual photographic records.

Hydraulic Engineering and Fluid Flow Research

The Hasselblad EL is particularly suited to many forms of fluid flow research and in the solving of river current and flow location problems. Banks of up to 20 Hasselblad ELs are suspended over a scale model of the river bed or sections of the ocean floor to be studied. By floating numbers of white polystyrene balls down the model and illuminating them by mercury vapor lamps, a series of tracks is formed on the negatives against the black of the river bed. By computing the distance of the tracks against a speed scale included in the photograph, flow speeds can be calculated.

By using much smaller plastic chips and the same photographic techniques, current patterns are formed at mouths of rivers, in bays and around structures in the river.

Obviously, the remote control features of the Hasselblad make it extremely useful for this kind of work, and the use of either the 70mm-70 exposure, or 220-24 exposure magazine, allows the researcher to make many exposures before bringing the camera down from the roof of the building. And, unless the building has an extremely high roof, (in which case the 80mm-Planar could be used) the 40mm Distagon will allow the maximum area to be covered by each camera.

Materials Testing

Other than the more regular forms of material testing which are usually carried out under ideal laboratory conditions, there are certain times when photographs of fractures or breakages of materials are needed. Yet, the structures are inaccessible to a photographer e.g., the inspection of blast furnaces or large capacity wine storage casks, both containing large quantities of toxic gas. In these instances, the lowering of the EL into the structures to be tested and the operation of the camera by remote control, provides the solution to the problem. Once again, the use of the 70mm magazine is desirable if numerous exposures are required.

Because of the confined space of the structures, the wide angle lenses available for the Hasselblad, the 40mm Distagon, with its 88° angle of view, or the 50mm Distagon, with its 75° angle of view, would be most useful.

Cave Photography & Speleology

A great deal of photography in cave and cavern research and its related sciences, palaeontology, anthropology and prehistory, is being done with the Hasselblad EL.

Working conditions are usually so bad—mud, water and of course, constant darkness—that film changes are not only undesirable but usually impossible. Use of either the 220-24 exposure or the 70mm-70 exposure magazine will reduce the number of film changes to the absolute minimum. Because of the spacial limitations of a cave, a wide angle lens is indispensable. Either the 40mm Distagon, with its 88° angle of view, or the 50mm Distagon, with its 75° angle of view, cannot be bettered.

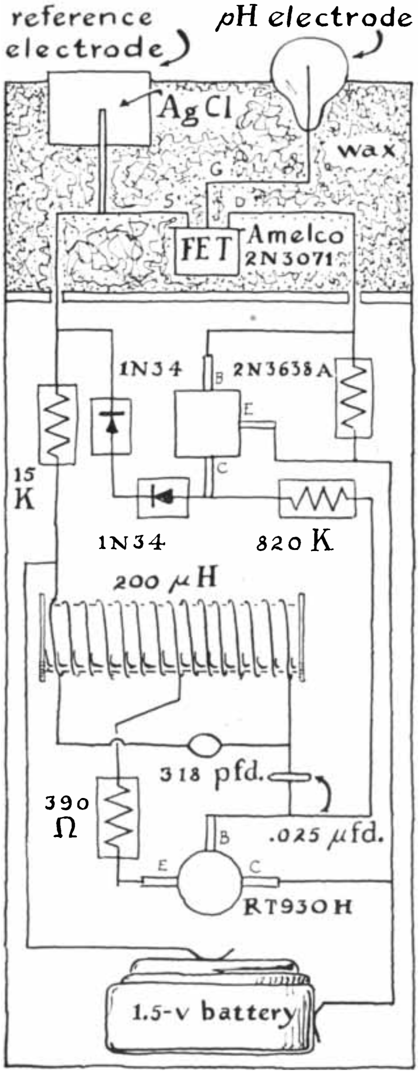
Aerial Photography

The Hasselblad EL is also ideally suited for many phases of aeronautical research. In most aerial research photography the camera is not handled by a skilled photographer, but usually by a flight test engineer or even, in the case of a single seat airplane, by the pilot himself. These people do not have the time to be concerned with manual operations such as exposing, winding on or changing film.

Flight instrumentation recording by remote control operation of the pre-focused Hasselblad EL allows for the reconstruction of flight conditions at pre-determined intervals during the test flights.

Numerous accessories are available for the Hasselblad EL. These are fully outlined in a 40 page illustrated booklet which we would be happy to send to you on request. If you also have any particular problems of a photographic nature, please address your enquiries to our Technical Director . . . Paillard Inc., 1900 Lower Road, Linden, New Jersey 07036.

HASSELBLAD



Transmitter for sensing p*H*

ed either on the surface of the transmitter or externally as a separate unit. The transmitter can be made responsive to changes in pressure by suspending close to one end of the coil a piece of magnetic material, such as a thin disk of ferrite. The disk is supported by a flexible diaphragm [*see top illustration on page 130*].

The disk can be sawed from the end of a cylindrical "loop stick." Such sticks are available from dealers in radio supplies. The diaphragm can be cut from any thin sheet of plastic (such as Saran Wrap) that is relatively impermeable and assembled in the cavity of a coil form machined from a plastic rod.

Increases in the pressure of a fluid in which the transmitter is immersed move the ferrite disk toward the coil, thus increasing the inductance of the coil and lowering the frequency of the radiated signal. The inward motion of the diaphragm also compresses the trapped air, creating a restoring force that moves the ferrite away from the coil when the external pressure falls. Changes in pressure are sensed by the receiver as changes in frequency of the radiated signal. The tuning dial of the receiver can be calibrated accordingly. The transmitter, when so constructed, can be used simultaneously for telemetering both temperature and pressure.

Although almost any radio receiver will pick up the signals, some receivers work better than others. The best ones must be found by trial. Take a transmitter into a radio store and try it on an assortment of receivers. Select for additional test those receivers that tune most sharply and appear to be most sensitive.

Take these receivers outside the building. In large cities some of the receivers will pick up several interfering broadcasting stations at the same time. Buy the one that is most selective. Usually little correlation is found between size, price and quality. Many experiments can be conducted in which the receiver is within two or three feet of the transmitter. In such experiments signals of adequate strength are picked up by the antenna built into the receiver. Depending on the physical arrangement of the experiment, however, it may be desirable to equip the receiver with a loop antenna consisting of one or more turns of wire connected to the input terminals of the set. When one is making studies of caged animals, for example, a loop antenna can be installed inside the cage, or below it if the bottom of the cage is made of insulating material.

Occasionally the experimenter may be interested only in the location of a free-ranging animal—for example such burrowing animals as gophers and moles. Here wires can be stretched over the ground in an even pattern that will ensure a wire's being near the animal at all times, so that a clear signal is consistently obtained. (It is not difficult to improvise a small, comfortable harness for attaching the transmitter to a captured animal that is then released.)

To pinpoint the position of a transmitter within a given area the experimenter can stretch insulated wires in the form of slender, hairpin loops that form a pattern of crossed grids over the area. The loops should be about two feet wide; they can be hundreds of feet long. A transmitter that is within a foot or so of the intersection of two loops will generate a clear signal in each one. The observer pinpoints the location of the transmitter by connecting the receiver to each of the loops in sequence. The connections can be made by hand or, preferably, with an automatic switching mechanism.

Large animals that range over many square miles are tracked by transmitters more powerful than those that send out conventional radio signals. The frequency of these transmitters should be controlled by quartz crystals. The accompanying illustration [*bottom of page 130*] shows the circuit of a transmitter that has an effective range of several miles. The quartz crystal, which is designated *X* in the illustration, determines the frequency of the radiation.

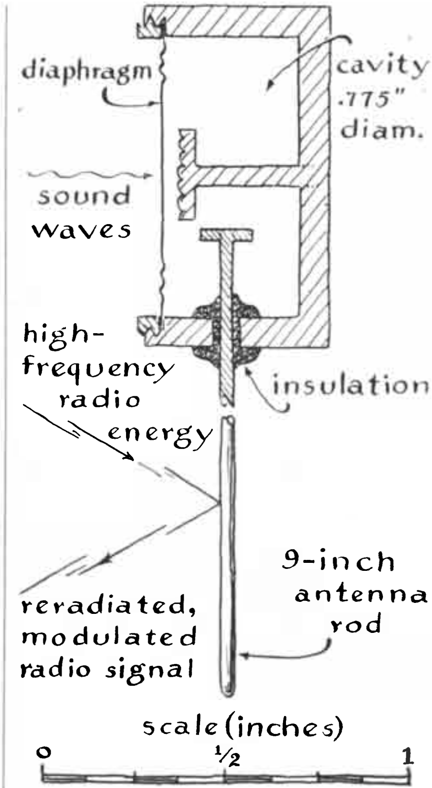
Highly directional antennas that resemble television receiving antennas can be used to follow an animal that carries a transmitter of this type. Two such antennas, spaced a known distance apart, are turned to the direction from which the signal is loudest. The position of the animal with respect to the observer is then calculated by triangulation.

A number of legal considerations must be observed by experimenters who operate radio transmitters of any kind or size. The regulations are altered from time to time, so that it is not possible to summarize them here. Information about them can be obtained from Ben F. Waple, Secretary, Federal Communications Commission, Washington, D.C. 20554.

Sensors of many kinds can be built into transmitters. Their variety is suggested by two examples. One unit, which can be made small enough to pass through the gastrointestinal tract of some large animals, measures and transmits *p*H—the relative acidity or alkalinity of surrounding fluids.

The active sensing elements of the device include a small but otherwise conventional glass electrode and a silver chloride reference electrode. Potentials developed across the electrodes are amplified by an electrometer circuit that employs a field-effect transistor. The output of the electrometer modulates the transmitter. This device is somewhat complex and not easy to construct.

At the opposite extreme is a transmitter that senses sound. Requiring no battery, it is powered by an external transmitter. Radio energy picked up by the antenna enters a small metallic cavity that is tuned, much like an organ pipe, to resonate at the frequency of a selected beam of radio waves. Sound waves that impinge on the thin metallic



A passive transmitter

diaphragm closing the cavity cause the diaphragm to vibrate and thus alter the natural resonant frequency of the cavity. This change, in turn, alters the strength of the signal that is reradiated by the antenna.

Devices of this kind that need no local source of energy are known as passive transmitters. Sounds that are within the transmission range of the reradiated signal can be heard in a radio receiver. A minute transmitter of this type, with an impressive transmission range, made international news in 1960 when Ambassador Henry Cabot Lodge reported to the United Nations its discovery in a wall hanging of a U.S. embassy.

My experiments have been made mostly with transmitters of the Hartley type. I have had only one brush with failure, but even this experiment turned out well. With the cooperation of a local veterinarian, I fed a temperature-sensing transmitter to my dog. The animal emitted gratifying clicks for two days, and I started to draw a graph of variations in temperature along his gastrointestinal tract. On the third day the transmissions mysteriously ceased. I immediately assumed that the transmitter had failed. Later, however, I went around my yard carrying my receiver. In a far corner of the front lawn the signals came in loud and clear!

IMPERIAL

If you want more than luxury in your luxury car

Watch the Bob Hope Show on NBC-TV.

Imperial Crown Four-Door Hardtop Interior

Welcome to a totally new driving environment.

In Imperial, you don't just see luxury. You *experience* it.

You feel it in a front seat that adjusts to you, not vice versa. It is actually three seats in one—a 5-foot wide sofa which adjusts into twin armchairs, or an armchair plus recliner.

You feel it in the response of a 440 cubic inch V8—largest ever offered by Chrysler Corporation.

You feel it in the incomparable torsion-bar ride.

What you *don't* feel is boredom. You can order an 8-track stereo tape player that plays through five speakers. Or an AM/FM radio that tunes itself. Or Auto-Temp air conditioner that maintains the temperature you select—winter and summer.

On long drives, set the optional Auto-Pilot speed control, readjust the seat, assume a new position. Relax. The tensions of mind and muscle are soon forgotten.

This year, get *more* than luxury in your luxury car. Drive the 1968 Imperial, finest automobile built by Chrysler Corporation.

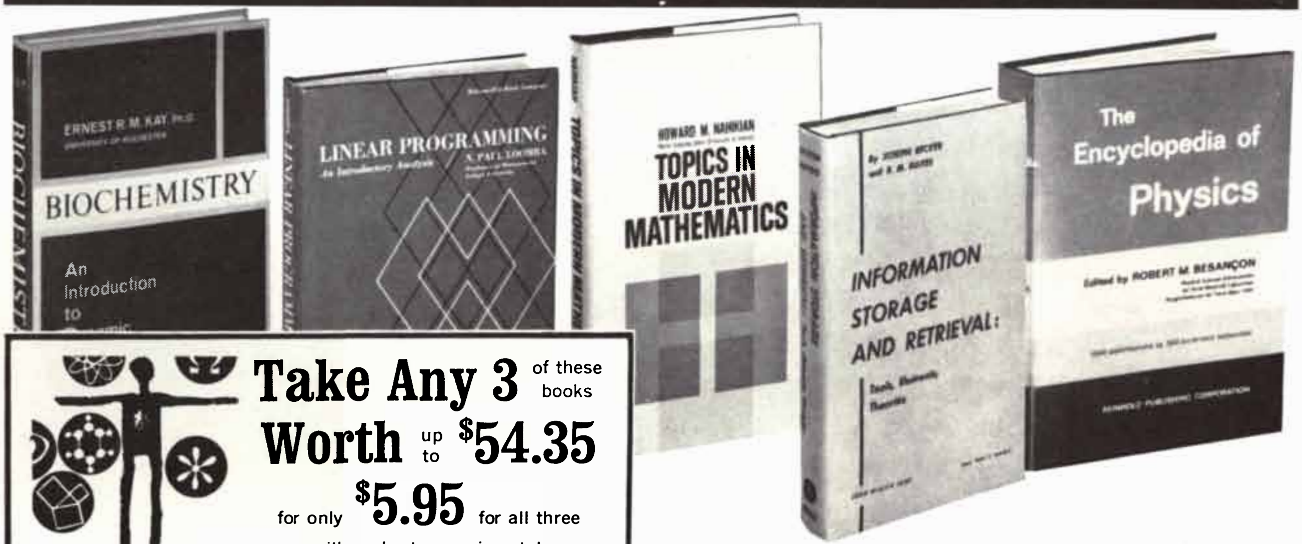
IMPERIAL

CHRYSLER MOTORS CORPORATION


N8157G


TRIM
CLEAN
REMINGTON

Keep Up with the Changing World of Mathematics / Physics / Chemistry / Astronomy / Biology Archaeology / Genetics / Anthropology / Technology / Space Flight / Computers / Evolution



Take Any 3 of these books Worth up to $54.35 for only $5.95 for all three

with a short, experimental membership in

THE LIBRARY OF SCIENCE

The Suggested Trial: Simply agree to purchase 3 additional Selections during the coming year, all at special reduced Member's Prices **at savings up to 40%.** (Then get a **free** Bonus Book as well!)

3416. APPLIED COMBINATORIAL MATHEMATICS. Edwin F. Beckenbach, Editor. The "who's who" of combinatorics contribute to this useful and lucid volume on the mathematics of organized-complexity problems, as based on algebra and number theory. Publisher's Price $13.95

3623. BIOCHEMISTRY. Ernest R. M. Kay. Brilliant demonstration of the life process in terms of the chemical events that underlie the remarkable interdependence of structure and function in living cells. Publisher's Price $7.95.

4207. DEVELOPMENT OF MODERN CHEMISTRY. Aaron J. Ihde. Panoramic history of the past two centuries, as chemistry became the science we know today. 851 pages. Publisher's Price $13.50

4320. THE EARTH. Carl O. Dunbar. The dean of American geologists reviews all that we know about the development of the earth and the perennial mystery of its origins. Publisher's Price $12.50

4436. ELEMENTS OF ABSTRACT ALGEBRA. Richard A. Dean. Superb, newly published presentation of the challenging branch of modern mathematics whose foundations lie in the theory of groups. Publisher's Price $7.95

4494. ENCYCLOPEDIA OF PHYSICS. Robert M. Besancon, Editor. Indispensable one-volume reference work covering all standard and advanced topics in physics. 316 articles by 320 prominent contributors. Publisher's Price $25.00 *Counts as two choices.*

4620. ESP: A SCIENTIFIC EVALUATION. C. E. M. Hansel. The first book to scientifically appraise the mounting evidence for psychic phenomena, especially the "laboratory" experiments that apparently make the case for ESP. Publisher's Price $6.95

4667. EXCURSIONS IN NUMBER THEORY /MATH DIVERSIONS. A well-selected collection of fascinating aspects of number theory by C. S. Oglivy and John T. Anderson, and delightful, wide-ranging and historically famous math problems compiled by J. A. H. Hunter and Joseph S. Madachy. Publisher's Price $9.95 *Two books count as one choice.*

4834. FAMOUS PROBLEMS OF MATHEMATICS. Heinrich Tietze. Brilliant presentation of modern and classic problems in mathematics, including Fermat's last problem, trisecting an angle and many more. Publisher's Price $10.00

4887. FOUNDATIONS OF PLASMA DYNAMICS. E. M. Holt and R. E. Haskell. Extremely lucid outline of the fundamentals of this spectacular new branch of physics. Publisher's Price $12.95

4890. FRIENDSHIP AND FRATRICIDE. Meyer A. Zeligs, M.D. Psychoanalytic insights applied to the protagonists in the Hiss-Chambers case by an eminent clinician. Publisher's Price $8.95

5210. GUIDE TO MATHEMATICS FOR THE INTELLIGENT NON-MATHEMATICIAN. Edmund C. Berkeley. Leads the non-mathematician painlessly through definite and indefinite numbers, calculus, the "New Math" and more. Publisher's Price $6.95

5338. HISTORY OF TECHNOLOGY, Volume IV: The Industrial Revolution. Charles Singer, E. J. Holmyard, A. R. Hall, Trevor I. Williams, Editors. Panoramic view of the Industrial Revolution from 1750 to 1850 which ushered in our modern era. Publisher's Price $26.90

5529. INFORMATION, COMPUTERS, AND SYSTEM DESIGN. Ira G. and Marthann E. Wilson. Comprehensive overview of systems concepts in the light of information theory. Publisher's Price $12.50

5530. INFORMATION STORAGE AND RETRIEVAL. Becker and Hayes. Outstanding guide that covers information system theory and design, man-machine relationships, much more. Publisher's Price $11.95

5554. IN SEARCH OF THE PRIMITIVE. Lewis Cotlow. Fascinating glimpses of the customs and lore of primitive peoples whose way of life is fast vanishing under the onslaught of civilization. Publisher's Price $8.95

5628. INTRODUCTION TO LINEAR ALGEBRA. Marvin Marcus and Henryk Minc. Rigorous, lucid treatment of fundamental concepts of linear algebra and matrix theory which introduces practical applications. Publisher's Price $7.95

5631. INTRODUCTION TO PROBABILITY AND STATISTICS. B. W. Lindgren and G. W. McElrath. Clearly presents the basic ideas of statistics and probability, illustrated with many easily understood problems from many disciplines. Publisher's Price $7.95

5634. INTRODUCTION TO MATHEMATICAL STATISTICS. Robert V. Hogg and Allen T. Craig. Answers the question that crops up in all fields of science: "What are the odds against it?", and gives the mathematical methods for reaching that answer. Publisher's Price $8.95

5773. NUMBER: THE LANGUAGE OF SCIENCE. Tobias Dantzig. "The evolution of mathematical thought from the earliest times to the latest constructions is presented here with admirable consistency and in a wonderfully lively style." —*Albert Einstein*. Publisher's Price $6.95

5810. LECTURES ON ELEMENTARY NUMBER THEORY. "A highly original, exciting, and readable introduction to the many aspects of number theory."—*Melvin Hausner*. Publisher's Price $7.75

5864. LINEAR PROGRAMMING. N. Paul Loomba. Definitive explanation of today's most widely used programming technique. Publisher's Price $8.50

5940. MACHINERY OF LIFE. Dean E. Wooldridge. Brilliant exposition of the thesis that the biological development of life is an inevitable result of the operation of the laws of physics. Publisher's Price $7.95

6036. MAN'S PLACE IN THE DYBOSPHERE. Richard R. Landers. A challenging exploration of the brave new world which machines that "think" have created — and will create — for modern man. Publisher's Price $5.95

6153. ELEMENTARY MATRIX ALGEBRA. Franz E. Hohn. Provides a concrete understanding of matrix theory and of the applications to fields from industrial "transport" problems to atomic physics and game theory. Publisher's Price $8.50

6396. THE NATURE OF SOLIDS. Alan Holden. Lucid treatment of heat, symmetry, crystal structure, quantum mechanics, electrical conduction, and magnetic phenomena as they concern the solid state. Publisher's Price $6.95

6436. THE NEW AGE IN PHYSICS. Sir Harrie Massey. Gives a clear understanding of the complex advances in twentieth-century physics and the resulting changes in our concept of matter. Publisher's Price $10.00

6860. PHILOSOPHY OF MATHEMATICS. Paul Benacerraf and Hilary Putnam. Two dozen of the greatest mathematical minds — brought together for the first time — focus on the fascinating problems to which modern mathematics has given birth. Publisher's Price $11.35

8383. TOPICS FROM THE THEORY OF NUMBERS. Emil Grosswald. Historical approach to the presentation of number theory, including divisability, congruences, quadratic residues and arithmetic functions. Publisher's Price $8.50

8436. TOPICS IN MODERN MATHEMATICS. Howard M. Nahikian. Ideal introduction to some of the most exciting and basic concepts in modern mathematics: sets and subsets, Boolean algebra, probability, vector spaces, matrices, linear transformation, etc. Publisher's Price $7.50

ENROLL NOW TO RECEIVE UP TO $54.35 WORTH OF BOOKS ON SCIENCE FOR ONLY $5.95

The Library of Science
Front and Brown Streets, Riverside, N.J. 08075

Please enroll me as a Trial Member and send the 3 books, whose numbers I have circled below, for only $5.95 plus shipping.

As a Trial Member, I need accept as few as 3 more Selections during the next 12 months, always at reduced Member's Price plus shipping, and I may cancel Membership any time thereafter. Each month I will receive a Newsletter describing the forthcoming Selection, along with a convenient form for requesting Alternate Selections or no book at all.

I also understand that I may choose a *free* Bonus Book for every 4 Selections purchased. (Introductory Offer counts as the first Selection.)

Three Books for only $5.95 *(circle numbers):*

3416	4494	4890	5554	5810	6396
3623	4620	5210	5628	5864	6436
4207	4667	5338	5631	5940	6860
4320	4834	5529	5634	6036	8383
4436	4887	5530	5773	6153	8436

GUARANTEE: If, after examining the volumes of your choice, you are not completely satisfied, you may return all books within 10 days and your Membership will be cancelled. Pay nothing—owe nothing.

NAME...

ADDRESS..

CITY.......................... STATE.................... ZIP CODE.............

Please be sure to indicate correct zip code. 2-626

BOOKS

Is pure science doomed to decline in the U.S.?

by Victor F. Weisskopf

THE POLITICS OF PURE SCIENCE, by Daniel S. Greenberg. The New American Library ($7.95).

The pursuit of pure science has become an expensive activity. The U.S. annually spends roughly $3 billion on pure science, that is, science that is not directed toward any practical application. The line between pure and applied science is somewhat arbitrary; the amount is the official one quoted by Government sources, and it may include a good deal of money spent for research that some people would consider to be applied. Let us not, however, be too alarmed by the figure. It is less than .4 percent of the gross national product and a rather small investment in a highly successful venture. Two-thirds of the money comes from the Federal Government in one form or another. This heavy dependence of basic science on financial support from Washington has introduced a new sociological element into the life of the scientific community. Only some younger scientists have the privilege of devoting all their time to pure science or teaching. Most of the senior people and an increasing number of the younger ones cannot concentrate exclusively on science. They must spend part of their time—sometimes most of it—on the difficult, complicated and perhaps disagreeable task of getting money from the Government. This activity is almost synonymous with "politics." Thus we have a new phenomenon: the politics of science. That is the topic of the book reviewed here.

The review is divided into two parts. The first deals with the book itself, the second with the problems it raises. The author is an editor of *Science*, the official magazine of the American Association for the Advancement of Science. He is responsible for the magazine's news section, which reports on political aspects of science. Under his editorship this section has become lively reading. It tends to emphasize what one might call the scandals in the relationship between scientists and the rest of society. In the news section of *Science* this emphasis is not necessarily a bad thing; at least it is counterbalanced in the rest of the magazine by a solid fare of purely scientific articles and also by editorials that, whether or not one agrees with them, thoughtfully consider important issues. Greenberg's book similarly emphasizes the frailties of the scientific community. It is not, however, much counterbalanced; positive achievements are conceded only grudgingly. Indeed, Greenberg is scornful of assertions that basic research is important: "When the statesmen of pure science proclaim that pure science is the locomotive of contemporary society, their sincerity is equaled only by their lack of evidence."

The book begins with a description of the scientific community. Greenberg regards this community with the air of a man who has studied a strange tribe that is quite different from the rest of humanity, and he finds the members of the tribe terribly impressed with themselves. As he puts it: "Let this be observed of scientists: Greater chauvinism, xenophobia, and evangelism hath no professional."

It should be said that Greenberg does know a lot about the scientific tribe and its institutions. His description of how science has been supported in the U.S. is quite interesting. He gives a vivid account of how before World War II research was carried out on a shoestring, and how during the war physical science suddenly became "important." He is somewhat confused in a chapter on postwar events: the efforts of scientists to secure international control of atomic energy abroad, civilian control at home and financial support of basic science in the universities. He is more successful in telling the amazing story of how American science gained financial salvation through the armed forces. One never ceases to marvel that the Navy and the Air Force lavishly support pure science of all kinds, including some of the most esoteric, and yet regardless of possible military applications leave the scientist complete freedom of research. One feels that it is wrong to have basic science supported by the military, but it is difficult to find evidence of bad effects. An interesting chapter.

The following chapters, however, confirm one's feeling that Greenberg is emphasizing the negative. As he describes the financial support of science today he loves to tell us about the quarrels, the ambitions and the weaknesses of scientists. He is most eloquent when he has dug out some unhappy tale. For example, one of the main themes in his chapter "High Energy Politics" is the tragic story of the "MURA" project, an attempt by a group of Middle Western university physicists to get financial support for the construction of a large particle accelerator of high intensity rather than high energy. The attempt was fruitless, and the story is not one of the most attractive episodes in science politics. High-energy physics did not, however, suffer too much from the project's lack of success. It may even be that the failure of the MURA group is an example of how well the clumsy system of forces and counterforces, committees and countercommittees can function.

Greenberg has a long chapter on the fiasco of the "Mohole" project, the attempt to organize the effort of drilling a hole through the ocean floor. This project came to naught because of an unusual display of shortsightedness and thoughtlessness; it is therefore an attractive subject for Greenberg. He tells the story with great zest, and he makes sure that geophysicists appear to be an irresponsible group that dreamed up a multimillion-dollar project on the spur of the moment simply to attract attention to their field. In actuality the project was carefully examined before it was proposed, and the conclusion was drawn that it would lead to important fundamental knowledge about the structure and history of our planet. Unhappily the project was thoroughly mismanaged by those who should have carried it out. Its demise is a great loss.

Perhaps Greenberg should not be

blamed too much for his critical attitude. It is always instructive to study shortcomings and failures. In general his stories seem to be factually correct (apart from a few glaring exceptions). His interpretation of motives, however, is often biased and misleading. It gives a distorted picture of the scientific community and its mode of operation. This is all the more dangerous because the community is currently in a precarious position before a justly critical public that is worried about the lack of controls in the use of tax money for the basic sciences. After all, the only people who can judge the value of a project in pure science are the scientists themselves. Hence the only control on the effectiveness of projects financed by the taxpayer is exerted by the recipients of the money. Nonetheless, it is remarkable how well this peculiar system has worked. Today the achievements of science in the U.S. are more impressive than they are anywhere else, a fact that may be related to generous spending. The percentage of mediocre work is also lower here than anywhere else, a result that is independent of the amount of money spent.

Ambitions, pressures and shady politics in Washington are not a monopoly of scientists. Still, there is one difference between scientists and other pressure groups. Scientists surely fight not only for the sake of science but also for their own ambitions and for influence and power. They do not, however, fight for financial profits. Greenberg concedes this point; indeed, he admires the effectiveness of science's disorganized process of self-control. Thus it is all the more puzzling that he is so fascinated by failures, machinations and scandals and less interested in achievements.

His last chapter deals with the problems the scientific community will face in the future in its demands for continuing support. He speaks of a "new politics of science" in which the going will be much rougher, not only because of the war in Vietnam but also because of a mounting dissatisfaction with the way funds are distributed to science. As an example he tells the story of the difficulties the National Institutes of Health have encountered in recent years, when their methods of distributing funds have been sharply criticized. This institution has an excellent record of constructive work and patronage; nevertheless, it came under bitter attack after its budget had risen to more than $1 billion a year (not all of which went to pure science). At the end of the book Greenberg gives the scientific community some advice on how to improve the image of science before the public. Much of his advice is worth heeding, in particular his remark that scientists should try harder to formulate "just what it is that science contributes to society."

Let me try in the second part of this review to set forth a few thoughts in this direction. Why should the Government use the taxpayer's money to support basic science? In most other areas of public spending Congress and the people can judge (or believe they can judge) the value and the implications of what public money is spent for; they believe they can find out whether the money was spent well or not. When it comes to basic science, however, they are not able to judge the quality and relevance of what they get for their money, and they are well aware of it. Only scientists themselves can gauge a project's importance and its success or failure. They are recipients, administrators and judges all at the same time. The trouble comes from the fact that the layman cannot understand science well enough to make responsible judgments. He may have a rough idea of what is going on, but he is far from able to evaluate new scientific developments. It is a most unusual and most irritating situation.

The problem has become acute only during the past two decades. Before the war the cost of basic research was so low that it could be borne by private sources and by public money earmarked for education. That is no longer so. The figure I have cited for the annual expenditure on basic science in the U.S., $3 billion, is too much for such sources. To be sure, it is not much compared with the $13 billion per year we spend on alcoholic beverages, the $8.4 billion we spend on tobacco and the more than $3 billion we spend on "durable toys, sports equipment, pleasure boats and pleasure aircraft." I shall leave aside the $30 billion per year we now pay for our tragic involvement in the war in Vietnam.

I mention these figures in order to suggest why we should do away with one popular argument against basic science, which goes as follows: It is immoral to spend so much money on something that has no immediate use and is of interest only to a very small community when money is needed for the antipoverty program, for social services and for the rehabilitation of cities. This argument derives its strength from the fact that we spend even less on the war against poverty than we do on basic science. This is a shameful state of affairs that should be changed. The problem must, however, be put in a different form. Clearly the support of basic science is less urgent than the war against poverty. That is not the question; all we need to show is that basic science is more important than, say, 15 percent of drinking, smoking and "durable toys."

Even this may be a difficult task. We should distinguish between two groups of basic sciences. One group I would call the "obviously applicable" kind of sciences. The other is the not so obviously applicable kind for which I shall use the term "frontier" sciences, without meaning to imply that the sciences of the first kind are not also scientific frontiers. The first group includes molecular biology, solid-state physics, plasma physics and so on. These disciplines function as pure sciences, that is, they are pursued with the aim of discovering basic phenomena and without any practical applications in mind. Still, it is obvious that any progress in molecular biology will be relevant to some medical problem, that any deeper understanding of the solid state will be helpful for the production of metals and synthetic materials, and that plasma physics will eventually enable us to exploit the energy of nuclear fusion. The second group includes elementary-particle physics (usually called high-energy physics), galactic and extragalactic astronomy (today the astronomy of the solar system is obviously applicable!), cosmology and so forth. These sciences deal with distant objects; elementary particles in the modern sense are also "distant," since mesons and baryons appear only when matter is subjected to extremely high energies that are normally not available on the earth but are probably found somewhere else in the universe. The "distant" feature of these sciences is what makes them expensive. It costs money to create conditions in the laboratory imitating conditions that may exist only in some exploding galaxy. It costs money to build instruments for studying the limits of the universe. In short, these sciences do not have obvious applications because they deal with phenomena that do not occur on the earth.

The problem of public support is quite different for the two groups of sciences. Support for the first group is no problem at all. It is true that even in these sciences only scientists are able to judge the results of their work; therefore the scientific community has been both the recipient of funds and the distributor, and will have to remain so. The system has worked fairly well and has produced remarkable results. It is well known that these results have had a heavy impact on medicine and technology. Everyone wants cures for cancer, everyone wants an economic method for desalting seawater, everyone understands the necessity of basic research of the first kind for these results and others, such as more effective methods of birth control, the production of new foodstuffs, cheap energy and, last but unfortunately not least, new weapons of war. There will always be money for basic science of the first kind. It will sometimes be too little and sometimes even too much, but there is no fundamental problem in convincing the nonscientist of the importance of these sciences for many practical aims.

The situation is different with the second group of sciences. There are two aspects to these sciences. Let us take high-energy physics as an example. One aspect is that the research is directed toward the innermost structure of matter and therefore is looking for the most fundamental laws of nature that govern the behavior of matter. It seeks the reasons for the existence of electrons, protons and neutrons and for their having the properties they exhibit. The other aspect is that sciences such as high-energy physics uncover a new world of natural phenomena unknown to us before. When matter is bombarded with particles that have energies of billions of electron volts, it exhibits phenomena of a completely novel character. The proton and the neutron are transformed into new and formerly unknown states (often referred to as unstable new elementary particles) and a great variety of different mesons are created. A world of new particles, reactions and interactions is about to be discovered that represents a behavior of matter very unlike phenomena on the atomic and nuclear scale. The same is true in regard to the other "frontier" sciences such as astronomy and cosmology.

One must grant that the first aspect is exciting primarily to the scientists involved, since only they can sense and evaluate the fundamental nature of the discoveries. To the outsider the second aspect also has a rather esoteric character. It seems to have no connection with the rest of our activities and vital interests. To be sure, these discoveries are interesting, sometimes even fascinating, as when cosmology touches on questions such as "Did the world begin with a 'big bang'?" Still, the ordinary citizen may feel that he can live without this excitement and that it does not seem worth several hundred million dollars a year.

This general feeling was expressed by Alvin M. Weinberg of the Oak Ridge National Laboratory when he suggested that we weigh the importance of a branch of pure science by its relevance for other sciences and technology. Clearly on Weinberg's scale the "obviously applicable" sciences weigh much more than the "frontier" sciences. The sciences of the first group deal with objects that have been under study for a long time and belong to the human environment. That is why these sciences are relevant to each other and to technology. The argument against the sciences of the second group is that they deal with problems far removed from the human environment and are therefore of minor social relevance.

I believe this point of view is thoroughly inconsistent. What is the human environment? Ten thousand years ago there were no metals in the human environment. Metals are rarely found in pure form in nature. After man discovered how to create them from ores, however, they played an important role in his environment. The first piece of copper must have looked very esoteric and useless. In fact, man for a long time used it only for decoration. Later the introduction of this new material into man's ken gave rise to interesting possibilities that ultimately led to the dominant role of metals in his environment. In short, we have created a metallic environment. To choose another example, bulk electricity in nature is rare; it appears only in lightning and in frictional electrostatics. It was not an important part of the human environment. After long years of esoteric research into minute effects it was possible to recognize the nature of electrical phenomena and to find out what a dominant role they play in the atom. The interaction of these new phenomena with the human world created the completely new electrical environment in which we live today.

The newest example is nuclear physics. In the early days prying into the nucleus was considered a purely academic activity, directed only toward the advancement of knowledge about the innermost structure of matter. In 1933 Lord Rutherford said (he is quoted in Greenberg's book): "Anyone who expects a source of power from the transformation of these atoms is talking moonshine." I imagine his conclusion was based on the same kind of reasoning I have cited above: Nuclear phenomena are too far removed from the human environment. It is true that, apart from the comparatively small contribution of natural radioactivity, nuclear reactions must be artificially induced with costly fluxes of energetic particles. Large-scale nuclear phenomena on the earth are strictly man-made; they occur naturally only in the center of stars. Here again the introduction of these man-made phenomena into the human environment has led to a large number of interactions: artificial radioactivity has revolutionized many branches of medicine, biology, chemistry and metallurgy, and nuclear fission is a steadily enlarging source of energy. Nuclear phenomena too are an important part of the human environment.

Physics is now in the process of discovering another new world of phenomena at even higher energies. The discovery of this new world is one of the most remarkable feats of science. Its importance is usually not recognized by the layman, who is more likely to be impressed by the discoveries concerning the surface of the moon. What is the surface of the moon compared with phenomena that in nature are probably happening only in the interior of quasars and exploding galaxies? The lack of public understanding is the fault of scientists; they try to explain SU_3 symmetries to newspaper reporters instead of telling them what the whole thing is about. They themselves cannot see the forest for the trees!

When this new world of subnuclear-particle phenomena is introduced into the human environment, it is bound to interact with the environment in new ways. Nobody knows what will happen when it becomes possible to produce beams of particles thousands of times more intense than those of today—beams in which all the new particles will be mass-produced to react with the environment. Even now some people are speculating about the special clinical effects of pion beams, effects that cannot be produced by ordinary radiation. It reflects a rather conservative attitude when our high-energy physicists say that their field will never be of any practical use. Greenberg quotes Hans Bethe as saying, just as Rutherford did, "No... practical application has appeared, or is likely to appear, for particle physics." I myself would consider it highly unlikely that among the many new phenomena that will result from the interaction of intense high-energy beams with ordinary matter there would not be one with any practical application.

Here we see the relevance of "frontier" sciences: they open up new realms of natural processes and deepen the pool of phenomena we can use for scientific and technological purposes. This, I believe, is part of the answer to Greenberg's question: "When a handful of particle physicists, at the expense of several hundred millions of dollars per year in public funds, explore the 'particle zoo' are they fulfilling a purpose of society?

Or are they merely pursuing their own curiosity in virtually total disengagement from the society which supports them?"

This, however, is not the only relevant aspect of the frontier sciences. Another aspect lies in their importance for scientific education and for the state of science as a whole. The frontier sciences are directed toward the investigation of phenomena that are distinctly different from what is known. At the same time they are directed toward problems concerning the foundations of our present natural laws and the limits of their validity; they get at the deeper significance of these laws and their still unknown connections. The expansion of the universe, the origin of the universe, the relations among electrodynamics, nuclear forces and gravity, the origin of the elementary electric charge, the conservation of baryons, the nature of weak nuclear interactions—all are problems of this character. Science cannot shun these questions; it is the essence of scientific inquiry to proceed toward the bottom of things. New phenomena cannot be ignored; they must be investigated. Laws and regularities cannot be accepted; they must be understood.

The value of fundamental research does not lie only in the ideas and results it produces. The spirit that prevails in the basic sciences affects the whole scientific and technological life of a community because it determines its way of thinking and the standards by which its creations are judged. An atmosphere of creativity is established that penetrates to every frontier. The applied sciences and technology adjust to the intellectual standards that are developed in the basic sciences. This influence works in many ways: Some fundamental-research students go into industry, where the techniques that have been applied to meet the stringent requirements of fundamental research serve to create new technological methods. One example that comes to mind is the technique of measuring very short intervals of time, which was developed for the purposes of high-energy physics. In a variety of ways the style and the level of scientific and technical work are determined in pure research; it is what attracts productive people and what brings productive scientists to those countries where science is at its highest level. That is why so many good scientists have moved to the U.S. from Europe in recent decades. This is one of the important social functions of pure science: it establishes the climate in which all scientific and technological activities flourish; it pumps the lifeblood of ideas and inventiveness into laboratories and factories.

But it is not only the scientific and technological world that is influenced by basic science. The deeper insights into the workings of nature that pure science provides are also relevant for the nonscientific part of the community, even if the nonscientist has little knowledge of the underlying ideas and concepts. It is an awareness of the far-reaching results of science that counts, an awareness that we know about the age of the earth, the evolutionary development of life, the origin of the elements, the indestructibility of matter—above all, an awareness that nature works according to exact laws that exclude magic and demonstrate that man is not at the mercy of a capricious universe. In many instances science is more relevant when it predicts what cannot happen than it is when it makes positive predictions.

There are similar reasons why science must play an important role in education. Here the "frontier" sciences are of especial significance. When students are introduced to the workings of nature, the open questions and the unsolved fundamental problems are bound to be the center of interest. There is more to it than just the teaching of science. There is no scientific education without the active pursuit of research. One cannot learn science without concrete participation in the process of analyzing facts, sifting evidence and recognizing new phenomena. The scientists and engineers of the future will have had to be immersed in the spirit of inquiry, they will have had to have tasted the atmosphere of continuous search at the frontiers of knowledge if they are to be successful in any field—pure or applied. Basic research, therefore, is an essential part of higher education in science and engineering.

Such considerations may help in answering one of the most difficult questions of scientific policy: Given that pure science must be supported, how much money should be spent on it? What determines the right quantity of pure science and how fast pure science should develop? I believe the answer lies in the educational role of basic science. There should be enough support of pure science—particularly the frontier sciences—to ensure that scholars and students at our major universities can actively participate in such research. This criterion does not allow an exact determination of the amount of money needed, but it furnishes a sound order of magnitude. On this basis the support of pure science in the U.S. has been adequate up to the present. Over the past 40 years a climate has been created in our universities and research institutions that has been unusually favorable for the development of basic science. Now, however, all signs point toward a steady decline; the future looks dark. On the one hand we face a rapid expansion of our system of higher education, with many new universities appearing all over the country; we also face a steady increase in the cost of pure research because of the greater sophistication of our research tools and the general increase in prices. On the other hand the present political trend is against any expansion of Federal support of science; we may consider ourselves lucky if the basic science budgets remain at a constant level. Hence only a few of the newly created universities will be able to participate actively in basic research, and most of the old ones will have to reduce their programs. In a few years this will lead to a sharp decline of basic science in the U.S. Today the main problem in the politics of science is: How can we protect this important sector of modern society from the onslaught of our current moral, social, financial and military crisis?

Short Reviews

THE PARTICLE ATLAS, by Walter C. McCrone, Ronald G. Draftz and John Gustav Delly. Ann Arbor Science Publishers, Inc. ($125). CRIME AND SCIENCE: THE NEW FRONTIER IN CRIMINOLOGY, by Jürgen Thorwald. Harcourt, Brace & World, Inc. ($8.95). Human senses perceive the details of the world on a scale from samples with a mass of tons down to samples of a milligram. Below that, at least without artificial aids, we enter a generalized realm, one that no longer consists of apples and ashtrays, automobiles and poppy seeds but that is merely vague and dusty. Yet atoms are very much smaller, and bits and pieces of the world on the scale from micrograms to nanograms can still hold in their tiny compass an entire history and provenance. The acrid molecules of smog, say, have little history; the atoms of which the molecules consist cannot betray their origins very specifically. Any smokestack might have been their source. On the other hand, the dust particles in smog can tell a tale, often a quite candid one. *The Particle Atlas,* compiled by a pioneer group in the microscopy of air pollutants, is so beautiful in form that it takes a while to grasp the depth and ingenuity of its content. It displays some 500 striking photomicrographs in color, altogether resembling the catalogue of a New York art show in the last years before op and pop had taken over. They illustrate a taxonomy of wild dusts: the results of natural wind erosion and of combustion and industrial processes—the "mechanical excreta" of our cities. These specimens and their classification comprise the weight of the atlas, but with them comes a remarkably clear account of how to set up a laboratory competent to do such analysis (at a cost about equal to that of a large limousine) and a detailed explanation of such special identifying tricks as "dispersion staining"—the measurement of refractive index by noting under dark-field illumination the colors at the edges of particles embedded in a liquid with a known dispersion curve. When the indexes match, the edge is colorless; the color fixes the degree of mismatch. Other techniques, even the most modern ones such as the electron microprobe, are also described, but the work is keyed to the results of optical study: transparency, color, birefringence, refractive index and shape. Domestic incinerator dust shows tarry, partly burned paper, unfused ash and bits of metal can; a well-operated pulverized-coal-fired boiler yields glassy spheres of colorless ash plus purple-black ones of magnetite. Carborundum and chicken feathers, paint spray and petroleum-cracking catalyst are all exhibited and concisely described.

Crime and Science is a continental look at the study of microscopic incriminating traces in the years since Sherlock Holmes and Alphonse Bertillon first spoke of the art. Filled with cases tried everywhere from Zurich to Sydney, a nontechnical narrative recounts the growth of modern techniques for identifying human blood from crude beginnings in the simple measurement of red cells. A second theme is the study of dusts and hairs with microchemical techniques such as neutron activation. That the Sydney kidnaper owned a Pekingese and lived in a brick house joined with pink mortar set in a garden where two particular cypress species grew was enough to find the man and to convict him, although he had been turned up earlier by a widespread search for the implicated blue Chevrolet and had talked his way off the list of 5,000 suspect owners of such a car.

PROPERTIES OF LIQUID AND SOLID HELIUM, by J. Wilks. Oxford University Press ($24). There is a tradition of monographs on helium that has been formed by two generations of fine summaries of the properties of this strange atom, so abundant in the stars and so rare on the

A handful of people like Mary Carnwath are trying to keep our promise to the Indians. But they won't make it without you.

The Hopi Indians' village of Shipaulovi in Arizona sits on land so poor, infertile and inhospitable that so far nobody has tried to take it away from them.

Electricity has not yet reached the Hopis. Water must be hauled from three miles away. Jobs are few and far away. Only poverty and despair are close-by and in abundance.

Yet for the first time in generations, Mary Carnwath and people like her are stirring hope among the Hopis.

Mary Carnwath works and lives two thousand miles away, in Manhattan. Her own daughter is now grown-up, and through *Save the Children Federation* she is sponsoring one of the village girls, 8-year-old Grace Mahtewa.

The Mahtewas (two parents, three children, one grandmother and a sister-in-law) live tightly packed in a tiny rock and mud house. The father who knows ranch work but can't find any most of the year, isn't able to provide the family with even the bare necessities.

Grace, bright, ambitious and industrious, would possibly have had to quit school as soon as she was old enough to do a day's work. But, because of Mary Carnwath, that won't be necessary.

The $15.00 a month contributed by Mary Carnwath is providing a remarkable number of things for Grace and her family.

Grace will have a chance to continue schooling. The family has been able to make its home a little more livable. And with the money left over, together with funds from other sponsors, the village has been able to renovate a dilapidated building for use as a village center. The center now has two manual sewing machines that are the beginnings of a small income-producing business. It's only a small beginning. More money and more people like Mary Carnwath are needed. With your help, perhaps this village program will produce enough money to end the Hopi's need for help. That is what Save the Children is all about.

Although contributions are deductible, it's not a charity. The aim is not merely to buy one child a few hot meals, a warm coat and a new pair of shoes. Instead, your contribution is used to give the child, the family and the village a little boost that may be all they need to start helping themselves.

Sponsors are desperately needed for other American Indian children—who suffer the highest disease rate and who look forward to the shortest life span of any American group.

As a sponsor you will receive a photo of the child, regular reports on his progress and, if you wish, a chance to correspond with him and his family.

Mary Carnwath knows that she can't save the world for $15.00 a month. Only a small corner of it. But, maybe that is the way to save the world. If there are enough Mary Carnwaths. How about you?

Save the Children Federation is registered with the U.S. State Department Advisory Committee on Voluntary Foreign Aid, and a member of the International Union of Child Welfare. Financial statements and annual reports are available on request.

Save The Children Federation—
Founded in 1932
National Sponsors (partial list):
Andy Griffith, Gene Kelly, Mrs. Eli Lilly, Paul Newman, Mrs. J. C. Penney, Frank Sinatra, Mrs. Earl Warren

Save The Children Federation
NORWALK, CONNECTICUT 06852

I WISH TO SPONSOR AN AMERICAN INDIAN CHILD.
ENCLOSED IS MY FIRST PAYMENT OF:
☐ $15 MONTHLY ☐ $45 QUARTERLY
☐ $90 SEMI-ANNUALLY ☐ $180 ANNUALLY

I CAN'T SPONSOR A CHILD, BUT I'D LIKE TO HELP.
ENCLOSED IS A CONTRIBUTION OF $________
☐ SEND ME MORE INFORMATION.

NAME________________

ADDRESS________________

CITY__________ STATE__________ ZIP__________

CONTRIBUTIONS ARE U.S. INCOME TAX DEDUCTIBLE SA 3/

earth. The present up-to-date (to the spring of 1966) research review, which treats only of the low-temperature wonders of this unique substance, is a worthy successor. Giving due weight to both experiment and theory, the author, a learned cryogenist of the Clarendon Laboratory at Oxford, has produced a work that is clear and complete and yet never oppressive in detail. He divides the work logically into a study of first the normal fluid and the superfluid of the isotope helium 4, then those of helium 3 (which are strikingly different) and then the solids of the two isotopes. His account of efforts to found the theory of such quantum fluids on first principles, without the crutch of the ingenious approximation of two fluids or of special quantum excitations as invented by Fritz London and by Lev Landau, is perhaps too compact to be very helpful. Still, it is a great tale. Consider that today physicists have found in these quantum substances four analogues to "first," or ordinary, sound: second sound, which is an oscillatory wave not of pressure but of temperature; zero sound, which is what normal sound waves become in a Fermi-particle quantum liquid such as helium 3—a pressure wave with anomalous velocity and absorption; third and fourth sounds, which are respective modifications of second and first sound peculiar to thin superfluid films. Perhaps even more marvelous are the flow of liquid helium superfluid in a superfluid "wind tunnel," which at low velocities has zero effect on a propeller, and the existence of tiny vortex rings in the superfluid, made detectable not by tobacco smoke but by inducing the rings to drag along a charged ion. The energy and velocity of these rings conform to the classical theory of smoke rings in air but on such a scale that the flow of fluid around the ring amounts to a single quantized unit of the fluid circulation, even though 100 million atoms may take part!

A HANDBOOK OF LIVING PRIMATES, by J. R. Napier and P. H. Napier. Academic Press ($21.50). PRIMATE ETHOLOGY, edited by Desmond Morris. Aldine Publishing Company ($10.75). In a reference volume no less engaging than it is authoritative the Napiers, man and wife, have given us a compact and yet critical review of the biology of the order of primates. The center of gravity of the book is in its "profiles." Each of the 60-odd living genera (with the unhappy exception of *Homo*) is listed in laconic form, giving the data on range, habitat, diet, weight, morphological features, chromosomes, blood groups, reproduction and behavior (wild and captive). Every genus has at least a page; the more interesting are spread over half a dozen pages. There are excellent photographs representing most genera. Here is a winsome child of the South African chacma baboon, there a chimp in the very attitude, finger to brow, of puzzled recollection. The senior chimp in captivity died at the age of 41; the dusky titis of the Amazon sit for hours side by side in pairs with their long tails charmingly entwined in a double helix.

The second book is a collection of 10 papers by men and women who study field behavior in the mode made famous by Konrad Lorenz and Niko Tinbergen. They approach primate behavior with a tradition of success in the study of wild fish and birds behind them, not with the experimentalist bias of the rat-maze psychologist nor yet with the baffling intricacy of anthropology. The work is only beginning, and in some cases it sounds even naïve. It is nonetheless stimulating and genuinely rich with interesting hypotheses. It reminds the general reader of the excitement of first reading the psychoanalysts, and yet its data are far clearer than those out of that shadowy domain. One paper, by Wolfgang Wickler, treats of the way in which signals so obviously sexual—such as the bright skin colors of the anogenital region in some monkeys—have come to have a far more general use and function. The hindquarters of some male monkeys, and in the mandrill even the face, imitate the female sexual skin and are displayed in ways that are continuous with the prototype but go far beyond it. Such continuity can be traced from the familiar urine-marking of territory by many mammals to ithyphallic herms, the curious carved stone pillars used by men in antiquity as "house guards." Spanking with pants down is a forced presentation in the genus *Homo* that still humiliates.

DRUG RESPONSES IN MAN, edited by Gordon Wolstenholme and Ruth Porter. Little, Brown and Company ($13). No drug, save perhaps aspirin, has all three desiderata: safety, effectiveness and low price. This London symposium (sponsored by Ciba) spells out the reasons. It is hard to extrapolate from laboratory animals to man; it is harder still to predict effects that occur only once in 100,000 times. Hardest of all is it to find rare, even if serious, side effects that "the medical profession . . . does not look for." Phenacetin—as in APC tablets—can cause kidney damage after very long use; it took 75 years of clinical experience to realize it. Allergies and genetic differences are hidden dangers for many. One anesthetist saw a young boy turn rigid and die under the anesthetic. Persuaded by statistics of the rareness of this condition—it happens perhaps once in a million times—he was reassuring when the boy's sister needed an operation. She too died; the two cases were thus not independent. The papers in this volume span the topic well, but best of all is the give-and-take of discussion here reported. The colloquy presents a frank and illuminating study of the difficult path from plausible biology to clinical medicine. A well-organized early-warning system in which the entire profession cooperates to spread the word about new drugs is a social mechanism that will help; eight million Americans were exposed to chloramphenicol before it was recognized that the drug affects the bone marrow in about one out of 50,000 cases. Not hundreds of persons but only a few would have suffered if there had been a warning system. It is also clear that the brave men, convicts, who willingly expose themselves to risky procedures as test subjects, without the promise but in the hope of lenient treatment thereafter, deserve recognition and reward. Perhaps not only the prisons need be used to find such volunteers. That even official negative data on unmarketed drugs remain trade secrets is one more peculiarity of the competitive pharmaceutical industry.

INDOCHINA, by Bernard Philippe Groslier. The World Publishing Company ($10). In pages full of scholarly candor and informed criticism the expert author—most of the photographs in the book were taken on the spot by him—gives us a brief history of this region, an account of its prehistory, its art and the growth of Western knowledge of these subjects. A hundred million people live and die today between the valley of the Irrawaddy River and the Gulf of Tonkin. Their history has been one of ebb and flow between the influence of India, its gods, its canal-builders and its trade, and China, its rice culture, its generals and its marvelous crafts. Behind it all there lies a long Neolithic period, just uncovered, and out of that past there speaks a family of rich and strange cultures. Their beautiful works are displayed here in splendid photographs, both in black-and-white and in color. Everyone knows of Angkor and the Khmers of Cambodia, but the great Burmese temples of Pagan and Mandalay, the Buddhist art of Tonkin, the Thai spires of the 13th century many can see here more or less for the first time. Hué and Vientiane, Dong-son and Vinh rise from this handsome book as something more than contested bases or angry targets.

Penguins, by John Sparks and Tony Soper. Illustrated by Robert Gillmor. Taplinger Publishing Company ($8.95). These flightless birds of antarctic waters, sleekly feathered in a down parka over the modeled layer of insulating blubber, are masters of underwater flight. How they make their living from squid and krill (perhaps profiting a little from the terrible toll that has now been taken of the blue whale), how they navigate by the sun and an inner clock, how they have probably evolved from the family of albatrosses, the finest fliers of all, and just about everything else of a semitechnical nature one might want to know about these endearing creatures is the tale of this enthusiastic book. There are fine photographs aplenty and attractive casual marginal sketches, but the real treat is the strong, simple, droll paintings of all the penguin species.

The Harper Encyclopedia of Science, edited by James R. Newman. Revised edition. Harper & Row, Publishers ($35). Measures for Progress: A History of the National Bureau of Standards, by Rexmond C. Cochrane. Editorial consultant, James R. Newman. U.S. Department of Commerce ($5.25). These two useful works, an improved one-volume version of a four-volume reference for the general reader and an interesting narrative of our national laboratory since its founding after the Spanish-American War, are mentioned here together because they are the past published testimony to the unfailing taste and depth of thought of the late editor of Scientific American's book section, James R. Newman.

A Mathematician's Apology, by G. H. Hardy, with a foreword by C. P. Snow. Cambridge University Press ($2.95). This touching, candid and sharp self-portrait by an aging artist—one can describe Hardy in no other way—appears after 10 years out of print, with a brief new account of the man by an old friend.

L'Énergie des Marées, by Robert Gibrat. Presses Universitaires de France ($3.60). A French paperback of authority and charm that treats of the tides, the economics of their use for power and the engineering problems of such big plants as the recently dedicated at Rance.

INDEX OF ADVERTISERS

MARCH 1968

BIBLIOGRAPHY

Readers interested in further reading on the subjects covered by articles in this issue may find the lists below helpful.

ANTI-BALLISTIC-MISSILE SYSTEMS

ADDRESS BY HONORABLE ROBERT S. MCNAMARA, SECRETARY OF DEFENSE, BEFORE UNITED PRESS INTERNATIONAL EDITORS AND PUBLISHERS, SAN FRANCISCO, SEPTEMBER 18, 1967. Office of Assistant Secretary of Defense (Public Affairs), 1967.

THE EFFECTS OF NUCLEAR WEAPONS. Edited by Samuel Glasstone. United States Department of Defense and the United States Atomic Energy Commission, April, 1962.

STATEMENT OF SECRETARY OF DEFENSE ROBERT S. MCNAMARA BEFORE THE SENATE ARMED SERVICES COMMITTEE ON THE FISCAL YEAR 1969–73 DEFENSE PROGRAM AND 1969 DEFENSE BUDGET. U.S. Government Printing Office; January 22, 1968.

HUMAN CELLS AND AGING

SENESCENCE AND CULTURED CELLS. Leonard Hayflick in *Perspectives in Experimental Gerontology: A Festschrift for Doctor F. Verzár.* Charles C Thomas, Inc., 1966.

THE SERIAL CULTIVATION OF HUMAN DIPLOID CELL STRAINS. L. Hayflick and P. S. Moorhead in *Experimental Cell Research,* Vol. 25, No. 3, pages 585–621; December, 1961.

TOPICS IN THE BIOLOGY OF AGING. Edited by Peter L. Krohn. Interscience Publishers, 1966.

OBSIDIAN AND THE ORIGINS OF TRADE

OBSIDIAN AND EARLY CULTURAL CONTACT IN THE NEAR EAST. Colin Renfrew, J. E. Dixon and J. R. Cann in *Proceedings of the Prehistoric Society for 1966,* New Series, Vol. 32, No. 2, pages 30–72; December, 1966.

OBSIDIAN IN THE AEGEAN. Colin Renfrew, J. R. Cann and J. E. Dixon in *The Annual of the British School at Athens,* No. 60, pages 225–242; 1965.

THE AUTOMATIC SYNTHESIS OF PROTEINS

THE CHEMISTRY AND BIOCHEMISTRY OF INSULIN. H. Klostermeyer and R. E. Humbel in *Angewandte Chemie: International Edition in English,* Vol. 5, No. 9, pages 807–822; September, 1966.

INSTRUMENT FOR AUTOMATED SYNTHESIS OF PEPTIDES. R. B. Merrifield, John Morrow Stewart and Nils Jernberg in *Analytical Chemistry,* Vol. 38, No. 13, pages 1905–1914; December, 1966.

SOLID PHASE PEPTIDE SYNTHESIS, I: THE SYNTHESIS OF A TETRAPEPTIDE. R. B. Merrifield in *Journal of the American Chemical Society,* Vol. 85, No. 14, pages 2149–2154; July 20, 1963.

THE SYNTHESIS OF BOVINE INSULIN BY THE SOLID PHASE METHOD. A. Marglin and R. B. Merrifield in *Journal of the American Chemical Society,* Vol. 88, No. 21, pages 5051–5052; November 5, 1966.

BILINGUALISM AND INFORMATION PROCESSING

LANGUAGES IN CONTACT: FINDINGS AND PROBLEMS. Uriel Weinreich. Linguistic Circle of New York, 1953.

PROBLEMS OF BILINGUALISM. Edited by John Macnamara. *The Journal of Social Issues,* Vol. 23, No. 2; April, 1967.

READING AND TALKING BILINGUALLY. Paul A. Kolers in *The American Journal of Psychology,* Vol. 79, No. 3, pages 357–376; September, 1966.

WORDS AND THINGS. Roger Brown. The Free Press, 1958.

CHANNELING IN CRYSTALS

INFLUENCE OF CRYSTAL LATTICE ON MOTION OF ENERGETIC CHARGED PARTICLES. Jens Lindhard in *Det Kgl. Danske Videnskabernes Selskab. Mathematisk Fysiske Meddelelser,* Vol. 34, No. 14, pages 1–63; 1965.

MOTION OF ENERGETIC PARTICLES IN CRYSTALS. Sheldon Datz, Cavid Erginsoy, Gunther Leibfried and Hans O. Lutz in *Annual Review of Nuclear Science,* Vol. 17, pages 129–188; 1967.

PASSAGE OF CHARGED PARTICLES THROUGH CRYSTAL LATTICES. Cavid Erginsoy. *Brookhaven Lecture Series No. 46,* BNL 944 (T-390). Brookhaven National Laboratory, March, 1966.

PULSE-CODE MODULATION

CERTAIN TOPICS IN TELEGRAPH TRANSMISSION THEORY. H. Nyquist in *Quarterly Transactions of the American Institute of Electrical Engineers,* Vol. 47, pages 617–644; April, 1928.

THE MATHEMATICAL THEORY OF COMMUNICATION. Claude E. Shannon and

Warren Weaver. University of Illinois Press, 1963.

The Philosophy of PCM. B. M. Oliver, J. R. Pierce and C. E. Shannon in *Proceedings of the I.R.E.*, Vol. 36, No. 11, pages 1324–1331; November, 1948.

Telephony by Pulse Code Modulation. W. M. Goodall in *Bell System Technical Journal*, Vol. 26, No. 3, pages 395–409; July, 1947.

THE ADJUSTABLE BRAIN OF HIBERNATORS

The Effect of Temperature and Photoperiod on the Yearly Hibernating Behavior of Captive Golden-mantled Ground Squirrels (Citellus lateralis tescorum). Eric T. Pengelley and Kenneth C. Fisher in *Canadian Journal of Zoology*, Vol. 41, No. 6, pages 1103–1120; September, 1963.

The Performance of Dormice and Other Hibernators on Tests of Hunger Motivation. N. Mrosovsky in *Animal Behavior*, Vol. 12, No. 4, pages 409–415; October, 1964.

The Physiology of Natural Hibernation. C. Kayser. Pergamon Press, 1962.

Regulatory Mechanisms, Brain Activity and Behavior during Deep Hibernation in the Squirrel, Citellus beecheyi. Felix Strumwasser in *American Journal of Psysiology*, Vol. 196, No. 1, pages 23–30; January, 1959.

MATHEMATICAL GAMES

Amicable Numbers. Edward Brind Escott in *Scripta Mathematica*, Vol. 12, No. 1, pages 61–72; March, 1946.

Friendly Numbers. Howard L. Rolf in *The Mathematics Teacher*, Vol. 60, No. 2, pages 157–160; February, 1967.

Solved and Unsolved Problems in Number Theory. Daniel Shanks. Spartan Books, 1962.

THE AMATEUR SCIENTIST

Bio-medical Telemetry: Sensing and Transmitting Biological Information from Animals and Man. R. Stuart Mackay. John Wiley & Sons, Inc., 1968.

The Radio Amateur's Handbook. Byron Goodman. American Radio Relay League, Inc., 1967.

Transistor Circuit Engineering. Edited by Richard F. Shea. John Wiley & Sons, Inc., 1957.

start with the need

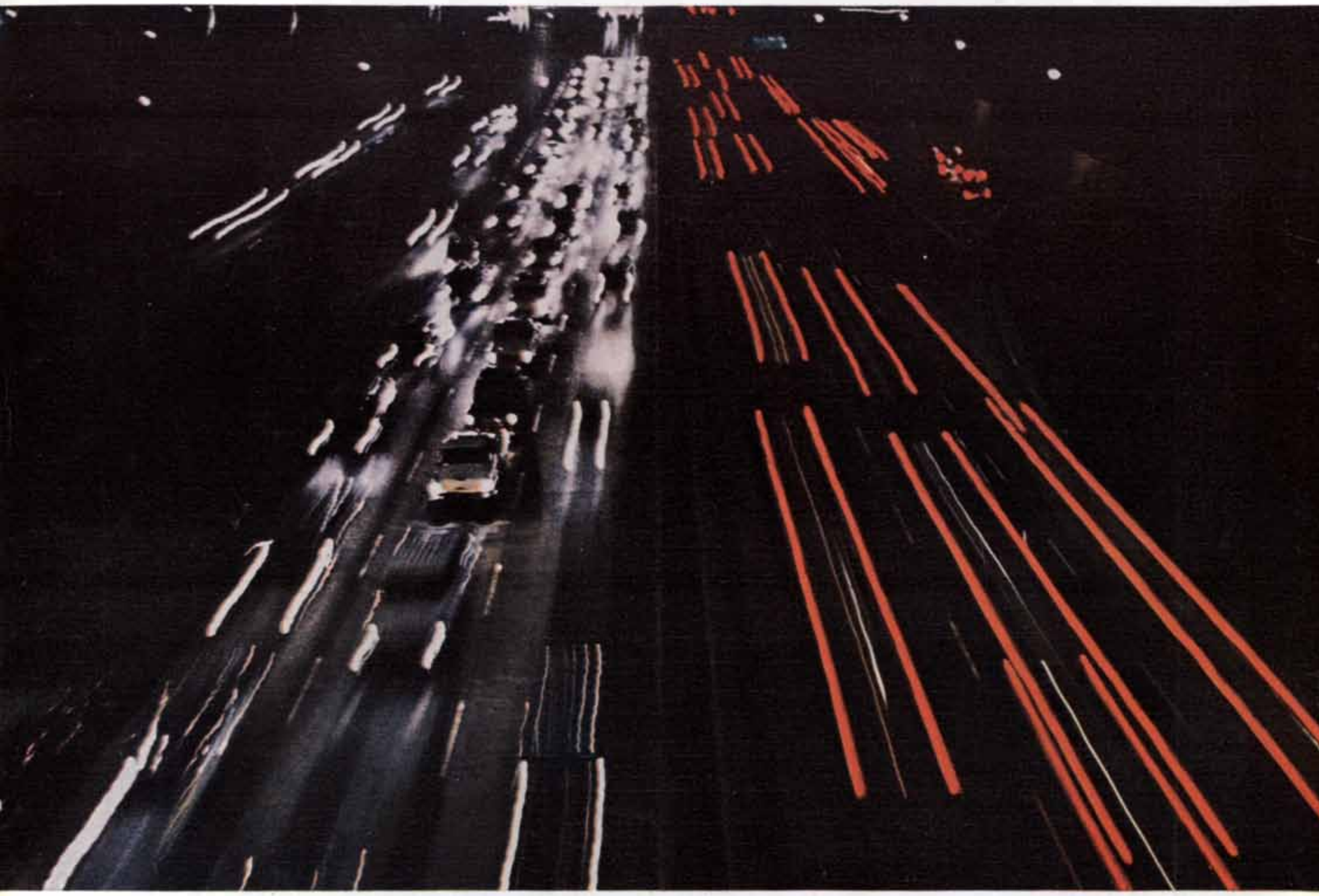
Markets are people and people have needs. When these needs are fulfilled, a remarkable cycle begins. A rising standard of living generates new hopes, new ambitions and new needs, until growth becomes a normal condition of the economy. The injection of capital creates new profits and new employment and an ever-widening circle of benefits for the nation's people at large, including, of course, those who share in ownership. Celanese considers its job to be a contribution to higher living standards throughout the world by filling human needs through our engineering, technical and marketing skills. This we do in chemicals, plastics, fibers, petroleum, paint and forest products, in 100 plants employing 50,000 people in 27 countries throughout the free world.

Celanese®

CELANESE

SAFER CARS

Injuries that don't have to happen, won't.



Allied Chemical makes the materials that make cars safer. Caprolan® nylon for tire cord protects your car against blowouts by resisting heat, moisture, and flex breaks. Caprolan nylon for seat belts reduces the chance of injury if an accident occurs. Urethane foam chemicals for dashboards help cushion the impact from sudden stops. They're all typical examples of how Allied Chemical puts chemistry to work.

ASK Allied Chemical

RIGHT IN THE CENTER OF THINGS AT NO. 1 TIMES SQUARE

Divisions: Agricultural, Fabricated Products, Fibers, Industrial Chemicals, International, Plastics, Semet-Solvay, Union Texas Petroleum, Allied Chemical Canada, Ltd.