SCIENTIFIC AMERICAN

FORAGE CROPS

$1.25

February 1976

For the SX-70, a film that developed in broad daylight was an impossible must.

This is how we made it.

The SX-70 Land camera was designed to eject the film into full sunlight within one-tenth of a second after exposure. However, we had decided that the user would have nothing to do but take the picture. There would be no timing and no peeling off of black paper. So it was necessary to find a new method of protecting the light-sensitive negative. We concluded that the way to do this was to integrate a chemical opacifying system into the film.

But here we were faced with a problem of formidable proportions:

The chemical shield within the film would have to be out of the way while in the camera so exposure could take place, yet be immediately effective during emergence of the film. It would have to be so opaque that a thin layer would protect the sensitive negative from direct sunlight. And it would again have to be out of the way once its protective function was completed, leaving a pure white background against which the developed image could be viewed.

The way in which we accomplished this task is a prime example of the use of pure and applied science to meet a complex technical need:

When the SX-70 shutter is snapped, light from the lens passes through a protective plastic layer on the face of the film plus three positive layers (all four layers are transparent) and exposes the negative beneath. (See Figure A.)

As the exposed film is ejected from the camera by motor-driven rollers, a thin layer of processing fluid, approximately 3-thousandths of an inch thick, is squeezed out of a pod at the leading edge of the film and spread evenly between the positive and negative layers.

This fluid has a remarkable composition. It is highly alkaline, since it has a high content of potassium hydroxide. It contains other chemicals to initiate and control the development of the negative. It also contains titanium dioxide and two novel,

FIGURE A

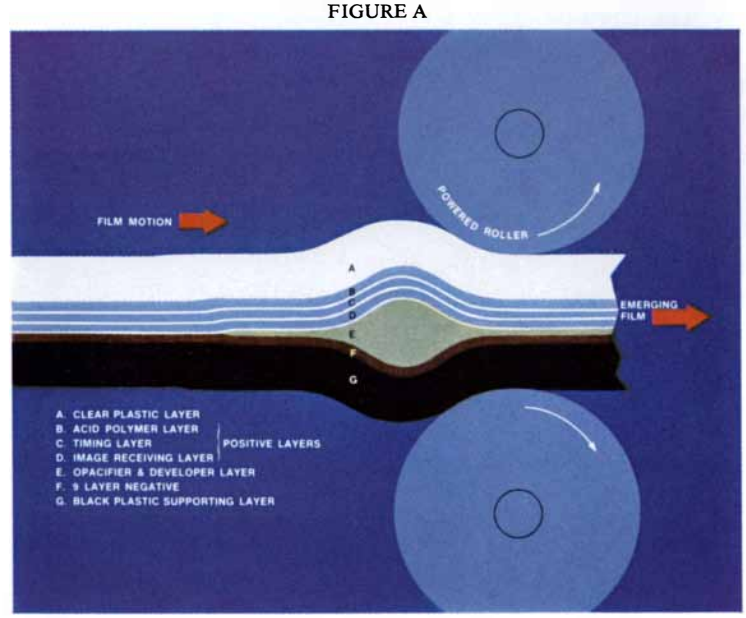
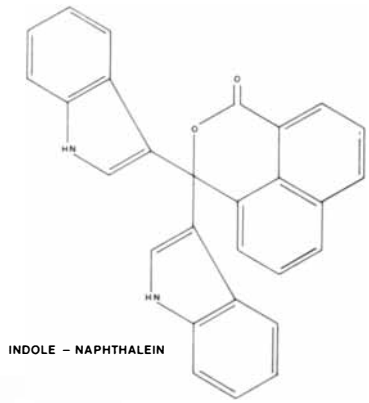


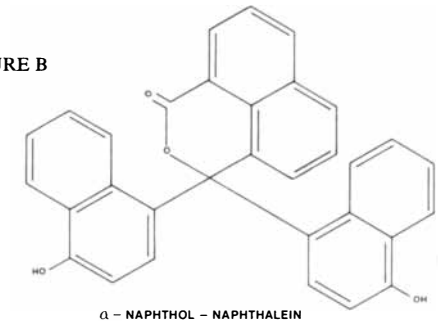
Diagram showing how instant darkness is spread between the positive and negative structures. Layer thickness is not to scale. The entire film is only .013″ thick as it emerges from the camera. The opacifier and developer layer is approximately .003″ thick. The top and bottom plastic layers comprise most of the thickness—.008″. Thus, aside from the opacifier and developer layer, all other picture development processes take place within a series of layers amounting to only .002″ in thickness.

Continued on next page



INDOLE – NAPHTHALEIN

FIGURE B



α – NAPHTHOL – NAPHTHALEIN

highly colored dyes. These last three substances acting together protect the developing negative from light after the exposed film emerges from the camera.

The two opacifying dyes were created over a period of intense organic chemical research to meet the specific requirements of the SX-70 system.

They are phthalein-type indicator dyes, and, like the indicator dyes employed by the analytical chemist, they have the ability to change color with changing pH.

The dyes are highly colored in strong alkali, but their color becomes less and less intense as the concentration of alkali decreases. They incorporate an important new achievement in that they become entirely colorless while the pH is still quite high (as high as the pH of washing soda). This permits rapid whitening of the image's background.

We provided for rapid decolorization of the opacifying dyes by introducing special hydrogen bonding groups into the dye molecules. These groups are placed in molecular proximity to the ionizable hydrogen atoms which are removed by potassium hydroxide to generate the colored form of the dyes. The hydrogen bonding groups raise the concentration of alkali needed to unlock the dyes' color. Conversely, the opacifier dyes are more readily decolorized by a decrease in alkalinity.

These dyes also have a molecular architecture that enables them to exist unchanged over long periods in their intensely alkaline environment.

In order to absorb blue-green light, we synthesized an indole-naphthalein which turns orange in alkali. (Indole is an unusual group in indicator dyes, and very little had been done with indole-naphthaleins until our basic research.) For red light, we synthesized a blue naphthol-naphthalein with an unusual structure. Together, these dyes absorb light throughout the entire region of silver halide sensitivity. (See Figure B for configuration of the basic groups.)

The synergistic effect of mixing finely divided titanium dioxide with opacifying dyes permitted us to use a lesser quantity of dyes than if we had used the dyes alone.

By itself, the TiO_2 in the thin layer between the positive and negative has a transmission density that will permit about one-fortieth of the impinging light to strike the negative. The transmission density of the small quantity of opacifying dyes used is also relatively low, allowing one-hundredth to one-thousandth of the light to pass through. But when the dyes and titanium dioxide are mixed, only about one-millionth of the light can penetrate the layer of dyes and TiO_2.

The effect of the TiO_2 is to increase the path length of the light rays within the opacifying layer. The photons bounce back and forth between the particles of highly reflective white pigment until nearly all of them are absorbed by the dyes. To the emerging negative, it's as though the sun turns black.

Once the chemical shield is in place, formation of the image begins as follows:

As the negative develops, the picture dyes migrate upward from the negative, penetrate through the opacifying layer and form a colored image in the image-receiving layer above it. (See Figure A.)

At the same time the alkali from the opacifying layer slowly penetrates layer "C". This is a timing layer designed to delay the encounter of the alkali with

the acid polymer layer above it until development has reached a predetermined point.

When the alkali passes through the timing layer, it reacts with the acid to form a little water and a salt of the acidic polymer. As this occurs, the alkalinity of the processing fluid is rapidly reduced, and the indicator dyes lose color until they are entirely colorless.

This reveals the full whiteness of the titanium dioxide, which forms a highly reflective background for the picture, giving the photograph extraordinary depth and brilliant highlights.

Thus, the seemingly incompatible goals of providing darkness and light with a thin chemical layer are achieved.

Since our film develops outside of the camera without human intervention, the SX-70 photographer is free to concentrate on his creative task. He can focus on his subject and take up to 10 pictures at 1.5-second intervals.

And he doesn't have to give a thought to print development.

Because we did.

The SX-70 System from Polaroid

The Saab Philosophy, No. 2 in a Series:

"WE BELIEVE HOW FAST A CAR GOES SLOW IS AS IMPORTANT AS HOW FAST A CAR GOES FAST."

When you're heading down the highway at 55 mph, you probably don't think about how fast you can stop. Until somebody stops in front of you.

We thought about it, however, when we first designed the Saab. And that's why we equipped it with big, power-assisted disc brakes on all four wheels (not just two like a lot of cars). They can take a Saab from 60 mph to 0 mph in only 175 feet.

Of course, braking is only important if you're going, so we gave the Saab a strong, two-liter, fuel-injected, overhead cam engine to get it going—from 0 mph to 60 mph in 9.5 seconds.

And with manual transmission, it gets an estimated 32 mpg on the highway and 22 mpg in the city based on EPA tests. The actual mileage you get will vary depending on the type of driving you do, your driving habits, your car's condition and optional equipment.

And to make the driver and passengers feel even more secure, the Saab has front wheel drive, rack-and-pinion steering and roll-cage construction, six solid steel posts, door impact panels, and a body designed to absorb impact in case you can't avoid an accident.

To really appreciate the Saab Philosophy, however, you must drive the result, the new Saab 99 GL. It will make you a believer...fast.

SAAB

IT'S WHAT A CAR SHOULD BE.

Established 1845 **SCIENTIFIC AMERICAN** February 1976 Volume 234 Number 2

ARTICLES

25 **THE ETHICS OF EXPERIMENTATION WITH HUMAN SUBJECTS, by Bernard Barber** A study of the attitudes of investigators is not reassuring.

32 **THE RECEPTORS OF STEROID HORMONES, by Bert W. O'Malley and William T. Schrader** A hormone-receptor complex acts on genetic material.

44 **IS GRAVITY GETTING WEAKER?, by Thomas C. Van Flandern** Measurements of the moon's orbital period are helping to provide the answer.

60 **FORAGE CROPS, by Harlow J. Hodgson** The grasses and legumes fed to livestock are the mainstay of U.S. agriculture.

76 **ROBOT SYSTEMS, by James S. Albus and John M. Evans, Jr.** Machines able to perform simple human tasks are now installed in many factories.

88 **THE FINAL PALEOLITHIC SETTLEMENTS OF THE EUROPEAN PLAIN, by Romuald Schild** The culture of hunters changed with the environment.

101 **CARBENES, by Maitland Jones, Jr.** Short-lived molecules are clues to the reaction mechanisms of organic chemistry.

114 **THE BIOLOGICAL CLOCK OF INSECTS, by D. S. Saunders** Insects anticipate the seasons by measuring changes in the length of the day.

DEPARTMENTS

8 LETTERS

16 50 AND 100 YEARS AGO

20 THE AUTHORS

54 SCIENCE AND THE CITIZEN

122 MATHEMATICAL GAMES

128 THE AMATEUR SCIENTIST

134 BOOKS

140 BIBLIOGRAPHY

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. SECOND-CLASS POSTAGE PAID AT NEW YORK, N.Y., AND AT ADDITIONAL MAILING OFFICES. AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $15 PER YEAR, U.S., POSSESSIONS AND CANADA; $18 PER YEAR, ALL OTHER COUNTRIES. EUROPEAN SUBSCRIPTION COPIES MAILED C/O PUBLICATIONS DISTRIBUTION SERVICE, ST. KATELIJNEVEST 14, B2000 ANTWERP, BELGIË.

THE COVER

The painting on the cover shows part of the foliage of a single alfalfa plant seen from above. Alfalfa is the principal forage crop in the U.S. (see "Forage Crops," page 60). Forage crops are the grasses and legumes fed to ruminant animals, mainly cattle and sheep. Their particular agricultural value is that only ruminant animals can digest them because of their high content of cellulose. Since they will grow in a wide range of soil conditions, they make productive much land that is not suitable for other crops. Alfalfa is a legume; like other members of the family Leguminoseae it provides its own nitrogen through its symbiotic relationship with *Rhizobium* bacteria, which convert nitrogen into a form the plant can utilize. Visible in the painting are several of the flowers produced by alfalfa as it nears maturity. Usually a crop of alfalfa is harvested when no more than 10 percent of it is in flower.

THE ILLUSTRATIONS

Cover painting by Peter McGinn

Page	Source
28–31	Total Communications Industries, Ltd.
33	Walter E. Stumpf and Madhabaranda Sar, University of North Carolina
34–41	George V. Kelvin
42	George V. Kelvin
43	P. M. Conn, Baylor College of Medicine
44–52	Jerome Kuhl
60	Remote Sensing Institute, South Dakota State University
61–64	Lorelle M. Raboni
65–66	Carl C. Lowe, New York State College of Agriculture and Life Sciences, Cornell University
67	Sperry New Holland Division, Sperry Rand Corporation
68	A. O. Smith Harvestore Products Inc.
76	Ralph Morse
78	Auto-Place, Inc.
79	Dan Todd
80	James S. Albus and John M. Evans, Jr., National Bureau of Standards
81	Jon Brenneis
82–83	Ralph Morse
84	Dan Todd
86	James S. Albus and John M. Evans, Jr., National Bureau of Standards
86B	Ralph Morse
89–99	Alan D. Iselin
101–113	Bunji Tagawa
114–115	Sarah Landry
116–126	Ilil Arbel
129	Fritz Goro
130–133	Jerome Kuhl

You can put it together. Together.

Meet the Dynaco Stereo 150 in kit form: 24 transistors, 2 zener diodes, 16 diodes, 2 thermal sensors and all the other components for building your own high-performance two-channel power amplifier. Assemble them over a weekend with a soldering iron, screwdriver, pliers, and the most readable step-by-step instructions in the industry. If you have a restless kid around the house, this kind of hands-on experience can add a happy new dimension to your relationship — and maybe open new horizons for him.

Completed, your Dynaco Stereo 150 is one of the most respected amplifiers on the market, with plenty of power for any loudspeaker. An output of 75 watts per channel* and less than one-quarter-of-one percent total harmonic distortion. (If you've never heard of Dynaco, ask an audiophile friend.) Add the Dynaco FM tuner (assembly time 6 hours), the Dynaco Preamplifier (12 hours), connect them to a pair of Dynaco Speakers (already assembled), hit the switch and enjoy a thrill neither of you will ever forget. You built it yourself, together, and we'll send you a personal nameplate to commemorate the venture. You've also saved $350 or more for the fun of doing it.

An excursion through our assembly instructions is the best way to try the project on for size.

*at 8 ohms, 20 Hz-20kHz.

dynaco

A DIVISION OF TYCO LABORATORIES

I'm interested. Send me instruction manuals for the following component(s) at $2.50 each:

____ Stereo 150 Amplifier
____ PAT 5 Preamplifier
____ FM Tuner
My check or money order is enclosed for $____
____ Send me a free brochure on the Dynaco product line and a list of dealers in my area.

Dynaco/Dynakit
Box 88 (Dept. SA) Blackwood, N.J. 08012

Name _____

Street _____

City _____ State _____ Zip _____

LETTERS

Sirs:

In your department "50 and 100 Years Ago" for last August there is an item that begins: "The Sicilian corner on sulfur, managed by an English syndicate, has been entirely broken through the success of chemical engineering in America." The American engineering achievement referred to is the Frasch process, in which sulfur is brought up from deep deposits by pumping hot water into it; the process is named after Herman Frasch, who developed it. The effect of the Frasch process on the Sicilian mines in the first 22 years of its application is indicated by the following figures. Up to 1903, when the Frasch process became a successful industrial operation, 95 percent of the world's sulfur was mined and prepared in Sicily; in 1903 there were 757 active mines with nearly 40,000 employees. By 1925 the number of mines worked had dropped to about 400 and the number of employees to about 20,000.

Even more dramatic than the final breaking of the Sicilian sulfur monopoly is the story of the first attempt to create a monopoly of Sicilian sulfur. This story is of particular interest today because of our experience with the Organization of Petroleum Exporting Countries (OPEC). By 1838 sulfuric acid had attained a key position in the British chemical industry. The raw material used in the production of the chemical was elemental sulfur, and it was obtained almost exclusively from the Sicilian mines. In 1838 the King of the Two Sicilies granted a monopoly for the export of Sicilian sulfur to MM. Taix & Cie. of Marseilles in an attempt, it was said, to stabilize prices. The result, however, was to raise the price of sulfur from 5.10 pounds sterling per ton to 15 pounds. This threefold increase was of course quite damaging to those industries that made a substantial use of sulfuric acid.

The British government regarded the granting of the Taix monopoly as a breach of the Convention of Commerce and Navigation, which had been signed in London in 1816. Apparently the situation was serious enough to lead to a Parliamentary inquiry into the conduct of the Foreign Secretary. For their part the British chemists worked out a process for producing sulfuric acid in which iron pyrite supplied the necessary sulfur, thereby eliminating the need for elemental sulfur. Iron pyrite is an iron sulfide ore and was abundantly available to British industry. With this competition Sicily lost nearly all the British trade. When it became clear that the monopoly was counterproductive, it was withdrawn and the price of five pounds per ton was reestablished.

A contemporary view of the sulfur monopoly, 1838 style, may be of interest here. In his *Familiar Letters on Chemistry* (1843) the great German chemist Justus von Liebig wrote: "We may fairly judge the commercial prosperity of a country from the amount of sulfuric acid it consumes.... As the price of sulfur has such an important influence on the cost of production of so many manufactured goods, we can understand why the British should have resolved to go to war with Naples in order to abolish the sulfur monopoly.... Nothing could be more opposed to the true interests of Sicily than such a monopoly; indeed, had it been maintained a few years it is highly probable that sulfur, the source of her wealth, would have been rendered perfectly valueless to her.... In commerce and industry every imprudence carries with it its own punishment, every oppression immediately and sensibly recoils upon the heads of those from whom it emanates."

ARTHUR F. SCOTT

Professor of Chemistry, Emeritus
Reed College
Portland, Ore.

Sirs:

In his article "Quarks with Color and Flavor" [SCIENTIFIC AMERICAN, October, 1975] Sheldon Lee Glashow remarks that physicists are upset by the apparent arbitrariness of the quarky rules for the formation of hadrons and for hadron interactions. I for one am not upset. Surely these rules are no more arbitrary than $1 + 1 = 2$, $\pi d = C$ or any of the other commonly accepted "laws" that govern our macroworld. And surely Dr. Glashow realizes that these rules require no explanation.

What might have been better said is that the quark hypothesis describes the behavior of the subatomic particles in a successful and concise manner. No "explanation" of these particles' behavior is, of course, ever possible through either pure or experimental science. "Why?" is a question left to philosophy and is not a proper part of any physical theory. Dr. Glashow has written, however, a very clear nonmathematical summary of the

PLYMOUTH VOYAGER

NO OTHER WAGON WORKS HARDER ON A GALLON OF GAS.

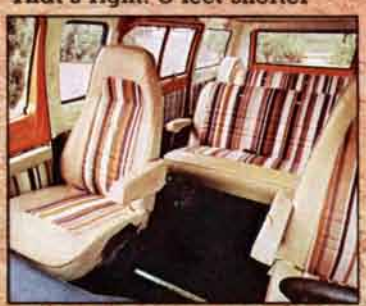
As a wagon buyer, you're a special breed with special needs. Try Voyager on for size. It comes with two different wheelbases, plus an extended body version that seats 15 people. Which is three more than any other wagon. Obviously, the more people room, the more cargo room.

26 HIGHWAY MPG, 18 CITY MPG.

That's according to the latest estimated EPA test results on Voyager's Slant 6 with manual transmission. If you don't think those are incredible figures, just compare them with those of any other wagon built anywhere in the world. Your actual mileage may differ depending on how and where you drive, the condition of your wagon and its optional equipment.

A TURNING DIAMETER 3-FEET SHORTER?

That's right. 3 feet shorter than Ford or Chevrolet. Which means tight U-turns and easier parking. Especially with optional power steering. Plus standard power front disc brakes and independent front suspension.



HERE'S "THE CLINCHER"

For the first 12 months of use, any Chrysler Corporation dealer will fix, without charge for parts or labor, any part of our 1976 passenger cars we supply (except tires) which proves defective in normal use, regardless of mileage. The owner is responsible only for normal maintenance like changing filters and wiper blades.

CHRYSLER
Plymouth

In California, see your dealer for mileage data for California equipped vehicles.

United States Steel asks a prominent American to speak out.

"What makes America work?

by Henry G. Parks, Jr.
Chairman, H.G. Parks, Inc.

If there is one lesson I've learned over the years, it is this: The American Dream is not dead.

Once, in the depth of my despair as a black man trying to make it in a white world, I thought it was. But I believe in it now. I lived it.

Exactly 25 years ago I managed to raise some capital partly by borrowing on my life insurance, and I rented an abandoned dairy in Baltimore. Afternoons, two workers and I ground meat and prepared sausage. Next morning, we sold them to stores from a used truck.

How we survived those first years I'll never know. But we did and today our company's sales approach $13 million annually.

Sure, I carry scars. I started out as poor as anybody could be and had endless frustrations and humiliations. But bitterness and hate are self-defeating. I believe in looking forward to see where America is going.

I know many in post-Watergate America are disillusioned and nobody has to tell me how minorities feel about their chances.

But the Dream lives and I'll tell you why. Doors are opening now for children of blue collar families, white and black. Important doors.

The old system of selecting business and industry leaders is changing. Belonging to the right clubs, marrying into the right families, will no longer guarantee advancement. Brains, talent, initiative—these will count most. Kids from across the tracks are on the way, loaded with all three.

Business leaders had better realize their own survival depends on getting the best people, or they'll lose the crucial race for executive talent.

Yes, the Dream lives—but it is being threatened, astonishingly, by Government itself!

We cannot return to laissez faire days, but must

...our freedom to succeed."

our economy be tied at both ends like my sausages?

We seem hell-bent on trying to cripple the economy by over-regulation. Price control was just one example of a classic disaster.

We have agencies upon agencies supervising all phases of business. It's like putting a battleship superstructure atop a rowboat. You wonder how business manages to stay afloat.

If we keep moving in that direction, the Dream dies and something else will take its place in our society.

We can't let that happen. Too many hopes of too many people ride on keeping that Dream alive.

"The old system of selecting business and industry leaders is changing. Brains, talent, initiative—these will count most."

Opportunity at U.S. Steel

Through our various recruiting, training and educational programs, hundreds of people are improving their qualifications and are going on to jobs of greater responsibility and greater personal reward at U.S. Steel. The idea of helping people to improve their lot is a policy with U.S. Steel and it will continue to be our policy.

United States Steel, 600 Grant Street, Pittsburgh, Pa. 15230.

USS TRADEMARK

We're involved.

SCIENCE/SCOPE

Main sensor aboard America's new weather-watching satellite, the GOES-A (Geostationary Operational Environmental Satellite-A), is the VISSR (visible/infrared spin-scan radiometer). It uses eight channels in the visible light band and two in the infrared band, making images by reflected sunlight during the day and by radiated energy from land, sea, and clouds at night.

VISSR's clear, sharp photos, transmitted to earth stations every 30 minutes, enable meteorologists to observe the growth and movement of weather patterns that may lead to hurricanes. Santa Barbara Research Center, a Hughes subsidiary, built VISSR for NASA's Goddard Space Flight Center.

GOES-A and two sister spacecraft, slated for launch in 1977, will be the first used in an international atmospheric research program whose goal is accurate prediction of weather up to two weeks in advance. Spacecraft to be orbited by the European Space Research Organization, Japan, and Russia will complete a global network.

A technique for producing laser mirrors with 99.9% reflectivity has been developed by Hughes research scientists under contract to the U.S. Air Force Weapons Laboratory. Their multilayer dielectric enhanced reflector technology gives the mirrors maximum resistance to damage by irradiation from high-power 10.6-micrometer CO_2 laser systems. It minimizes cooling requirements, reduces weight, and improves high-power laser thresholds.

While scuba divers from Jacques Cousteau's Calypso measured ocean floor reflectivity and water transparency in a recent experiment in the Central Bahamas, the multispectral scanner aboard NASA's Landsat 2 satellite measured water depths at the test site. The two sets of data -- later compared and analyzed -- could be of significant aid to maritime traffic and marine science. Communications for the experiment were relayed via NASA's Goddard Space Flight Center ATS-3 satellite, which has been in service since 1967. Both the ATS-3 and the Landsat multispectral scanner were built for NASA by Hughes.

Hughes needs Field Engineers to support advanced electronic systems. These growth opportunities include both domestic and overseas field engineering positions. Applicants must have a BSEE and U.S. citizenship. Graduates without experience will be considered. Please send your resume to: H.G. Staggs, Hughes Aircraft Company, Support Systems Division, P.O. Box 90515, Los Angeles, CA 90009. An equal opportunity M/F employer.

100 isolated villages in Alaska will receive satellite communications through small earth stations which incorporate traveling wave tube amplifiers built by Hughes. The lightweight, compact amplifiers provide more than 35 watts of RF power at 6 GHz. They are being built under contract to Alaska's Office of Telecommunications.

Creating a new world with electronics

HUGHES

HUGHES AIRCRAFT COMPANY

current state of elementary-particle theory and the state of the four-quark theory in particular.

R. L. McCalla

Instructor of Physics
Northwest Mississippi Junior College
Senatobia, Miss.

Sirs:

I feel I must take exception to Cornelius Van S. Roosevelt's identification of Habakkuk as the smallest man in the Bible because Habakkuk "stood on his watch" ["Mathematical Games," SCIENTIFIC AMERICAN, December, 1975].

Fabricated of the materials commonly available for such purposes in his day, baked clay and building stone (Israel was never a rich place like Egypt and Babylon), a watch was a formidable structure. That Habakkuk could carry such an instrument around with him clearly implies that he was an exceptionally big man. Even for him, however, putting his watch down to stand on it occasionally must have afforded welcome relief.

JOHN A. CHURCH

New York

Scientific American, February, 1976; Vol. 234, No. 2. 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 secretary; George S. Conn, treasurer; Arlene Wright, assistant treasurer.

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

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

Offprint correspondence and orders should be addressed to W. H. Freeman and Company, 660 Market Street, San Francisco, Calif. 94104. For each offprint ordered please enclose 30 cents.

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

NAME

NEW ADDRESS

OLD ADDRESS

Rediscover America

Price. Mileage. Maintenance. Engine guarantee. Pontiac's

MILEAGE—SMALL SPORTY CARS

MAKE & MODEL	ENGINE	TRANSMISSION	EPA ESTIMATES HIGHWAY	CITY
Pontiac Sunbird 2-Door Coupe	140-cu.-in. 4-cyl. 2-bbl.	Manual	35	22
Datsun 280Z Sport Coupe	168-cu.-in. 6-cyl. F.I.	Manual	27	16
Toyota Celica ST & GT Sport Coupes	133.6-cu.-in. 4-cyl. 2-bbl.	Manual	34	20
VW Scirocco Sport Coupe	97-cu.-in. 4-cyl. 2-bbl.	Manual	39	25
Fiat X1/9 Coupe	78.7-cu.-in. 4-cyl. 2-bbl.	Manual	31	21
Audi Fox 2-Door Sedan	97-cu.-in. 4-cyl. F.I.	Manual	37	24
BMW 2002 2-Door Sedan	121-cu.-in. 4-cyl. 2-bbl.	Manual	25	18
Mazda Cosmo Coupe	80-cu.-in. Rotary 4-bbl.	Manual	29	18
Porsche 914 Coupe	120-cu.-in. 4-cyl. F.I.	Manual	30	20
Triumph TR-7 Coupe	122-cu.-in. 4-cyl. 2-bbl.	Manual	30	21

Source: 1976 EPA Fuel Economy Guide.

PRICE—SMALL ECONOMY CARS

MAKE & MODEL	BASE PRICE*
Pontiac Astre 2-Door Coupe	$3120**
Toyota Corona 2-Door Sedan	3699
Datsun 710 2-Door Sedan	3614 plus dealer prep.
Fiat 131 2-Door Sedan	3958 plus dealer prep.
Mazda RX-3 Coupe	4049 plus dealer prep.
VW Dasher 2-Door Sedan	5195 plus dealer prep.
Volvo 242 2-Door Sedan	6295

*All dollar figures are manufacturers' suggested retail prices, including dealer new-vehicle preparation charges, except where noted. State and local taxes, available equipment and destination charges are additional.

**Includes available 2-bbl. engine.

PRICE—SMALL SPORTY CARS

MAKE & MODEL	BASE PRICE*
Pontiac Sunbird 2-Door Coupe	$3487**
Datsun 280Z Sport Coupe	6594 plus dealer prep.
Toyota Celica ST Sport Coupe	4145
Toyota Celica GT Sport Coupe	4499
VW Scirocco Coupe	4995 plus dealer prep.
Fiat X1/9 Sport Coupe	4808 plus dealer prep.
Audi Fox 2-Door Sedan	4875 plus dealer prep.
BMW 2002 2-Door Sedan	6670 plus dealer prep.
Mazda Cosmo Coupe	5800 plus dealer prep.
Porsche 914 Coupe	7145 plus dealer prep.
Triumph TR-7 Coupe	5100 plus dealer prep.

*All dollar figures are manufacturers' suggested retail prices, including dealer new-vehicle preparation charges, except where noted. State and local taxes, available equipment and destination charges are additional.

**Includes available 2-bbl. engine.

MILEAGE—SMALL ECONOMY CARS

MAKE & MODEL	ENGINE	TRANSMISSION	EPA ESTIMATES HIGHWAY	CITY
Pontiac Astre 2-Door Coupe	140-cu.-in. 4-cyl. 2-bbl.	Manual	35	22
Toyota Corona 2-Door Sedan	133.6-cu.-in. 4-cyl. 2-bbl.	Manual	34	20
Datsun 710 2-Door Sedan	119-cu.-in. 4-cyl. 2-bbl.	Manual	33	23
Fiat 131 2-Door Sedan	107-cu.-in. 4-cyl. 2-bbl.	Manual	29	18
Mazda RX-3 Coupe	70-cu.-in. Rotary 4-bbl.	Manual	30	19
VW Dasher 2-Door Sedan	97-cu.-in. 4-cyl. F.I.	Manual	37	24
Volvo 242 2-Door Sedan	130-cu.-in. 4-cyl. F.I.	Manual	27	17

Source: 1976 EPA Fuel Economy Guide.

Remember these mileage figures are estimates. The mileage you get will vary according to the kind of driving you do, your driving habits, your car's condition and available equipment.

In California, see your Pontiac dealer for applicable prices, EPA mileage figures and engine/transmission combinations available on California emission-equipped cars.

in a 1976 Pontiac.

subcompact Astre and Sunbird take on the foreign made cars.

You decide who wins.

RECOMMENDED MAINTENANCE SCHEDULES

MAKE	SPARK PLUG REPLACEMENT	OIL CHANGE	OIL FILTER CHANGE
Pontiac Astre & Sunbird	**22,500**	**7,500**	**7,500**
Datsun (All)	12,500	6,250	6,250
Toyota (All)	12,500	6,500	6,500
Fiat (All)	12,500	6,000	6,000
VW (Rabbit, Dasher, Scirocco)	15,000	5,000	15,000
Mazda (Rotary)	12,500	6,250	12,500
Audi (All)	15,000	7,500	15,000
BMW (2002)	12,500	6,500	6,500
Porsche (914)	15,000	5,000	5,000
Triumph (All)	12,500	6,000	6,000
Volvo (All)	15,000	7,500	7,500

Data for foreign made cars based on 1975 models.

ENGINE GUARANTEES

MAKE	GUARANTEE
Pontiac Astre & Sunbird with 140-cu.-in. 4-cyl. engines	**5 Years or 60,000 Miles**
Datsun	12 Months or 12,500 Miles
Toyota	12 Months or 12,500 Miles
Fiat	12 Months or 12,000 Miles
VW	24 Months or 24,000 Miles
Mazda	36 Months or 50,000 Miles
Audi	12 Months or 20,000 Miles
BMW	12 Months or 12,000 Miles
Porsche	12 Months or 20,000 Miles
Triumph	12 Months or 12,000 Miles
Volvo	12 Months

Data for foreign made cars based on 1975 models.

Here's how Pontiac protects you for 5 years or 60,000 miles.

Pontiac guarantees to the owners of 1976 Astres and Sunbirds with 4-cyl. 140-cu.-in. engines that any authorized Pontiac dealer will make repairs without charge to the owner, for 5 years or 60,000 miles (whichever comes first), to the cylinder block, cylinder head, all internal engine parts, the intake and exhaust manifolds and water pump made necessary because of defects in material and workmanship.

This guarantee, an added value feature on 1976 Astres and Sunbirds, is in addition to the New Vehicle Warranty but does not apply to repairs required because of misuse, negligence, alteration, accident or lack of reasonable or proper maintenance. See your Pontiac Dealer for complete guarantee.

GM

Astre 2-Door Coupe

Sunbird 2-Door Coupe

PONTIAC

The Mark of Great Cars

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

FEBRUARY, 1926: "Senator Marconi's latest development, the short-wave beam transmitter, concentrates radio energy as a searchlight reflects in a definite and desired direction. Marconi has pointed out that a tremendous waste of power occurs in the modern trans-Atlantic commercial stations, which he believes are destined to be scrapped and replaced by the short-wave beam installations. A ten-kilowatt station projecting its radiation in a narrow ten-degree beam becomes, so far as a distant receiver is concerned, the equivalent of a 360-kilowatt transmitter radiating broadcasts in all directions. A reflector can also be used at the receiver, giving a further increase in the total power intercepted. The masts of the beam transmitters are arranged in a straight line at right angles to the direction in which communication is to be established. The reflector consists of a number of vertical wires suspended from cross arms on the mast."

"Erosion or soil wash is impoverishing our sloping farm lands at a very much faster rate than the crops that are being taken from them are. Recently the Missouri Agricultural Experiment Station showed by actual measurement that within 24 years erosion has removed seven inches of the surface soil from an important type of Missouri farm land, which is plowed regularly to a depth of four inches. In bluegrass sod, however, the same type of soil erodes at the rate of seven inches in 3,547 years. Surely something should be done to check this enormous wastage."

"The recent announcement of the development of the Zworykin thermionic tube presents one of those opportunities to speculate on future developments that are impossible to resist. A combination of the photoelectric cell and the radio vacuum tube, its sensitiveness is infinitely greater than anything yet developed. Variations of light falling on this tube instantly become variations of electric current, amplified approximately 100,000-fold. An alkaline metal, such as potassium, is coated on the inside of the radio tube. When light falls thereon, a shower of electrons is thrown off. Variations of the strength of the light vary the shower, infinite gradations being possible. This shower is tremendously amplified and produces a current sufficient to operate an ordinary relay, thus making it applicable for any purpose."

"Sodium and potassium are practically interchangeable in most experimental processes in which they take part. One long-known and striking instance where they are quite different, however, is in their penetration into plant cells. Whereas theory suggests that they would be about equally absorbed, experiment shows that under comparable conditions 60 times as much potassium as sodium is taken up and held by the cell. Now comes the artificial cell, made by Dr. Daniel T. MacDougal, director of the Carnegie Institution of Washington's Laboratory for Plant Physiology. On subjecting his cell to appropriate experiments, Dr. MacDougal discovered that about five times as much potassium as sodium was absorbed. At first one might be inclined to say that after all there is a big difference between 60 times and five times. Biologists who have worked with the artificial cell are convinced, however, that once the true explanation is found for the five-fold deviation from the theory in this case, the same explanation will hold for the living cell."

SCIENTIFIC AMERICAN

FEBRUARY, 1876: "It is a well-known fact that many of the artesian wells sunk for oil, in Pennsylvania and other parts of the country, failed to produce oil but emit great quantities of gas. Some measurements and analyses lately made by Mr. O. Wuth, the well-known chemist of Pittsburgh, go to show that in the gas that is thus passing uselessly into the atmosphere we have a vast store of fuel, and that in allowing it to run to waste we have for years presented the most striking instance of extravagance in the world. The well he measured is named Burns, and is on the Duffy farm, thirty-five miles from Pittsburgh. He found the composition of the gas to be C_4H_6, or 80 per cent of carbon and 20 per cent of hydrogen, with but little oxide of carbon and carbonic acid. The well has a tube of five-eighths of an inch diameter, and the pressure of gas is no less than 200 pounds to the square inch. On these data Wuth estimated that the well yielded 1,000,000 cubic feet of gas, weighing 58 tuns, *per hour*."

"It is a curiously suggestive fact that very few of the weekly journals that go directly into families can nowadays be examined but that somewhere in their advertising columns are found announcements offering radical cures for the opium habit. It appears that, while alcoholic intoxication is decreasing throughout the United States, opium drunkenness is increasing, and the dealers in crude opium and the manufacturers of its alkaloids assert that the importation of the one and the production of the other are rising rapidly year by year, so much so regarding morphia that at one of our largest manufacturing centers the supply is said to be insufficient for the demand. Dr. J. B. Mattison, in an article in the *Medical Record*, considers that legislative enactments prohibiting the refilling of an opiate prescription or the dispensing of opium in any form, unless in pursuance of a prescription from the attending physician, would accomplish an immense amount of good."

"At the present time it is probable that the three cities of New-York, Brooklyn and Jersey City contain more than 1,700,000 inhabitants, occupying no more dwellings than are found in the single city of Philadelphia, with less than half as many inhabitants. The tenement-house population of New-York and Brooklyn probably exceeds 250,000. In Boston there are now 2,700 tenement houses, containing 43,000 rooms and inhabited by more than 60,000 people. A portion are new and well built, even costly, houses, but nearly 50,000 of this tenement-house population in Boston are living in discomfort from overcrowding, many of them in houses quite unfit to be occupied. The death rate is found in Boston, as in New-York, to be largest in those districts where there are the most tenement houses, and there also crime and vice and prostitution most abound."

"The new British man-of-war *Inflexible* is now in course of construction at Portsmouth. Her iron armor is to be two feet thick. The ship is 320 feet long and 75 feet wide and is to carry two 8-tun guns. These guns will have 24 feet length and 16 inches caliber. The projectile weighs 1,650 lbs., and over a barrel of powder (300 lbs.) is the firing charge. The vessel's engines will be of 7,000 horse power operating on twin screws. Altogether the *Inflexible* is the most wonderful specimen of naval architecture ever undertaken."

As your introduction to membership in the BOOK-OF-THE-MONTH CLUB®

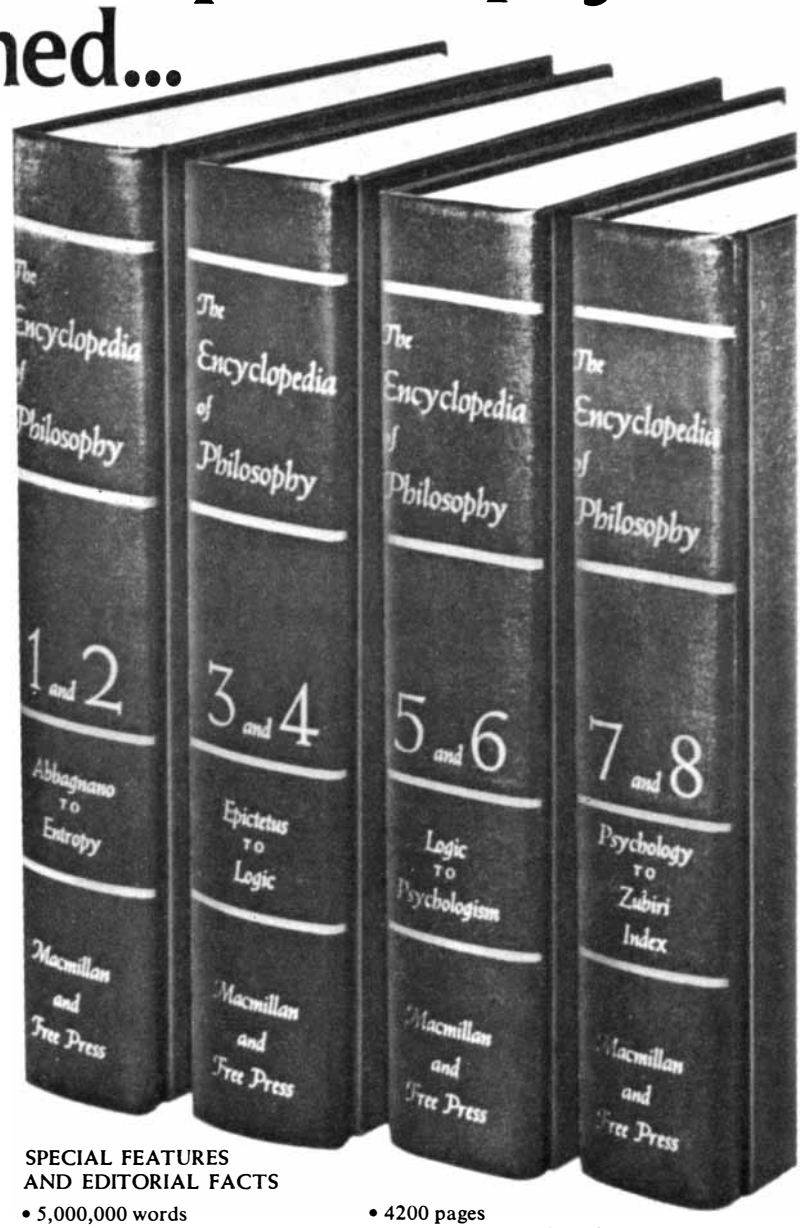
The most comprehensive encyclopedia of philosophy ever published...

FOR ONLY $17.50

THE SUGGESTED TRIAL
You simply agree to buy four Club choices within a year at substantial savings on most books you choose.

PUBLISHER'S
LIST PRICE: $99.50

The complete contents of the original 8-volume $219.50 edition now in a handsome 4-volume reader's edition



ONE OF THE GREATEST repositories of philosophical information available anywhere, *The Encyclopedia of Philosophy* is the only major philosophical reference work published in English since 1901.

In this unique 5,000,000-word work can be found clear, authoritative answers to almost any question concerning philosophy — ancient, medieval and modern, Eastern as well as Western. And because the work represents the collaboration of more than 500 philosophers and philosophical scholars from twenty-four countries, every shade of opinion is represented. No thinker or movement is dismissed as too radical, no issue is scanted as too controversial. There are articles, for example, that discuss such contemporary matters as the ethics of suicide and euthanasia and the benefits and dangers of computing machines.

Clearly, *The Encyclopedia of Philosophy* belongs in every personal library, particularly in homes where there are high school or college students. Up to now its high price has kept it out of the reach of most families. But, as a new member of the Book-of-the-Month Club, you may own it for less than a tenth of the cost of the original eight-volume edition. Because only a limited number of sets have been set aside for this offer, you are advised to act quickly.

BOOK-OF-THE-MONTH CLUB, INC.
Camp Hill, Pennsylvania 17012

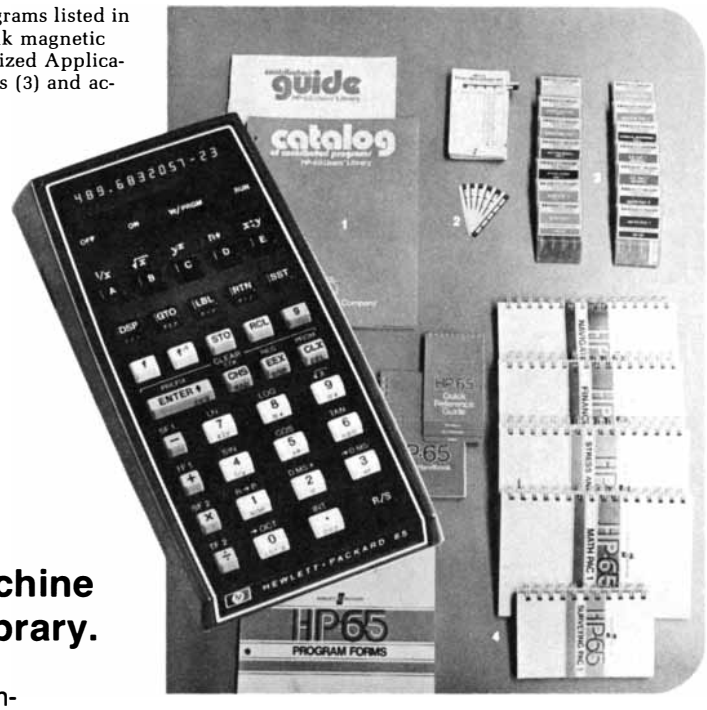
SPECIAL FEATURES AND EDITORIAL FACTS

- 5,000,000 words
- 158-page index
- Paul Edwards, editor in chief
- Nearly 1500 articles, including 900 on individual philosophers
- 4200 pages
- Editorial Board of 153 international scholars
- Contributors: 500 specialists from twenty-four countries

A SMALL SAMPLING OF THE IDEAS AND PHILOSOPHERS IN THESE VOLUMES

Atheism	Semantics	Good and Evil	Nonsense	Pain	Nothing
Free Will	The Absolute	Existentialism	Humanism	Truth	Linguistics

CONFUCIUS	SANTAYANA	DESCARTES	KANT	RUSSELL
ERASMUS	SPINOZA	HEGEL	NIETZSCHE	EINSTEIN
SARTRE	THOMAS AQUINAS	CAMUS	ARISTOTLE	NEWTON
LOCKE	WITTGENSTEIN	KIERKEGAARD	PLATO	PIAGET



HP-65 software includes 4000 programs listed in the Users' Library catalog (1), blank magnetic programming cards (2), 14 specialized Application Pacs including magnetic cards (3) and accompanying manuals (4).

The HP-65 hand-held fully programmable: the $795* machine with the priceless program library.

Without programs (software), a fully programmable calculator can be a time-consuming novelty. On the other hand, with its extensive software library, the HP-65 is immediately available to solve an astonishing range of computational problems.

Like a computer, the HP-65 memorizes programs of any size that are fed into it on tiny magnetic cards, 100 program steps to a card. With just a few keystrokes by the user, it executes the program and gives the answer to complex problems with an accuracy of up to 10 digits.

For many, a large part of the utility and pleasure of owning the HP-65 centers on its ease of programming. You don't have to know computer languages to do it well. Just push the keys in a normal problem-solving sequence, and record the program on a magnetic card for future use.

For many others, want of time or desire makes programming a chore.

During its two years of existence, the HP-65 has been accumulating programs: both those developed by Hewlett-Packard, and those submitted by thousands of enthusiastic users.

As a consequence HP-65 owners in the U.S.A. have immediate access to more than 400 prerecorded programs, each of which quickly transforms the HP-65 into a dedicated computer to solve a specific problem. These problems are grouped in 14 Application Pacs, with up to 40 programs per pac, for such disciplines as statistics, math, medicine, chemical engineering, electrical engineering, and finance. They are priced at $45* per pac.

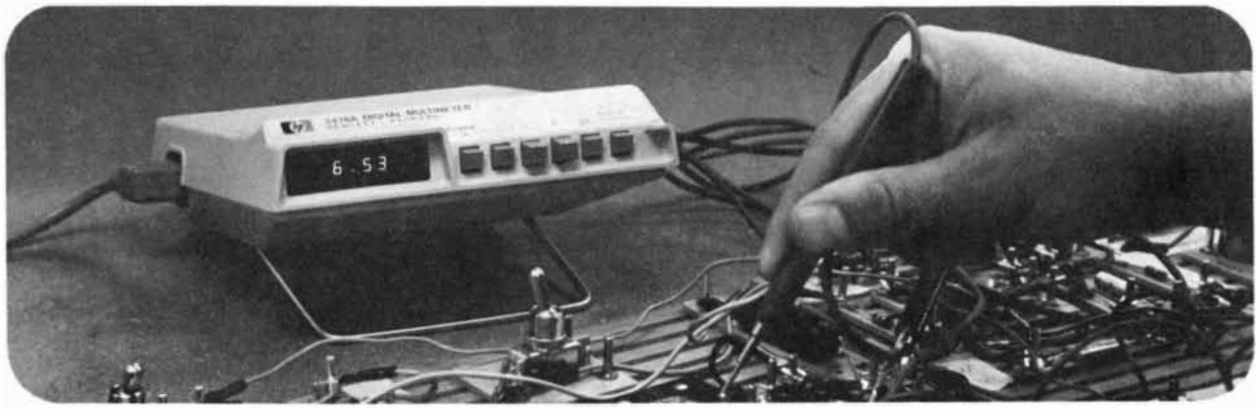
In addition, more than 4000 contributed programs are available through the U.S. HP-65 Users' Library. These are not prerecorded, but are fully documented and described step by step, so that it requires only a few minutes to key one into your HP-65 and permanently record it on magnetic cards. All programs are listed and described in the *Catalog of Contributed Programs*, free to HP-65 owners and to readers who request it (see coupon at right). These library programs with full documentation are available at a nominal charge.

HP-65 owners also receive the quarterly newsletter *Key Note*** to keep them informed of programming advances and other useful facts. For many, a fourth pillar of support for HP-65 owners is the most important one: personal assistance to help you solve individual computational problems, either by calling a toll-free number in the U.S.A., or by mail.

An offer you shouldn't refuse**

Now, to start the HP-65's third year, $795 will buy, between February 1 and April 30, 1976, four application pacs of your choice, five of the most popular programs from the Users' Library, and an HP-65 on which to run them.

No matter how you calculate it, the HP-65 remains the most powerful hand-held computing system available, with thousands of programs carefully crafted to save you time, sweat, and tears. You can get it at quality department stores and campus bookstores, or from HP. Call 800-538-7922 in the U.S.A. (in California, 800-662-9862) for the name of the HP retailer nearest you.



Rugged, five-function portable digital multimeter lets you concentrate on making good measurements.

With complete autoranging, autopolarity, and autozero on every one of its five functions, the new HP 3476 DMM combines high quality, portability, and economy. It's available for $225*.

Once again, a new HP-developed microcircuit process has paved the way to a fundamental improvement in instrument design—this time in the DMM, one of the most indispensible and familiar tools used by those who service electronic equipment.

The 3476 was made possible by a new, fine-line tantalum nitride-on-sapphire resistor pack. This chip, which contains 19 precision resistors matched to .02 percent by a laser-trim process, is combined with an N-MOS control chip on a single custom hybrid—greatly decreasing instrument size, interconnections, and cost, while improving temperature stability and reliability.

How good is the new 3476? With complete autoranging, autopolarity, and autozero, the 3476 gives you the right measurement quickly and easily. Push the button to select AC or DC voltage, AC or DC current, or resistance—then simply read the measurements on the LED display, directly in volts, amps, or kilohms. The 3476 lets you concentrate on your work rather than fussing with range knobs, looking for the right scale, figuring out where the decimal points belong, reversing test leads to change polarity, or wondering if you should have made a zero adjustment. If there's an overload, a bright red horizontal bar on the display tells you so unambiguously, preventing an erronous reading. You can even lock the 3476 on a specific range, to speed and simplify repetitive measurements and diode polarity checks.

How rugged is it? Because the 3476 has built-in safeguards, you don't have to be concerned about common overloads putting it into the shop for expensive repairs. All five functions are fuse-protected; and when you do blow a fuse, replacement is easy. The fuses pop out when you slide a cover plate back; replace them with standard fuses and you're back in business: no recalibration is required. The instrument is housed in a tough, high-impact polycarbonate case.

Small enough to fit in your briefcase, this new DMM is available in two models: the 3476A for AC operation at $225*; and the 3476B for AC/rechargeable-battery operation at $275*.

For more information on these products, write to us. Hewlett-Packard, 1505 Page Mill Road, Palo Alto, California 94304.

HEWLETT hp PACKARD

Sales and service from 172 offices in 65 countries

Mail to: Hewlett-Packard, 1505 Page Mill Road, Palo Alto, CA 94304.
Please send me further information on

( ) HP-65 Hand-Held Programmable and software
( ) HP 3476 Digital Multimeter
( ) Please send me a free Catalog of Contributed Programs for the HP-65.

Name _____
Company _____
Address _____
City _____ State _____ Zip _____

*Domestic U.S.A. prices only.
**Available only in continental U.S.A., Alaska, and Hawaii.
Software offer void where prohibited by law.

00641

THE AUTHORS

BERNARD BARBER ("The Ethics of Experimentation with Human Subjects") is professor and chairman of the department of sociology at Barnard College; he is also a member of the graduate faculty of Columbia University. Barber was educated at Harvard University, obtaining his B.A. in 1939, his M.A. in 1942 and his Ph.D. in 1949. A former member of the National Research Council's Drug Research Board, he currently serves on the Health Advisory Panel of the Congressional Office of Technology Assessment, the Council of the Society for the Social Studies of Science and the Human Subjects Review Committee (a Columbia group). In addition to his work on the sociology of science Barber has written extensively on the subject of social stratification. His books include *Science and the Social Order* (1952), *Drugs and Society* (1967) and (with others) *Research on Human Subjects* (1973).

BERT W. O'MALLEY and WILLIAM T. SCHRADER ("The Receptors of Steroid Hormones") are at the Baylor College of Medicine. O'Malley, who is professor and chairman of the department of cell biology at Baylor, received his B.S. in 1959 and his M.D. in 1963, both from the University of Pittsburgh. On completing his internship and residency at the Duke University Medical Center he joined the staff of the National Cancer Institute. Five years later he was appointed professor of reproductive biology at the Vanderbilt University School of Medicine, a post he held until he moved to Baylor in 1973. Schrader joined the department of cell biology at Baylor that same year, having spent two postdoctoral years working with O'Malley at Vanderbilt after obtaining his Ph.D. from Johns Hopkins University in 1969.

THOMAS C. VAN FLANDERN ("Is Gravity Getting Weaker?") is an astronomer at the U.S. Naval Observatory. As an undergraduate at Xavier University in Cincinnati he majored in mathematics; his Ph.D., from Yale University, is in astronomy. An expert on celestial mechanics, Van Flandern reports that his chief current research interest focuses on "evidence that the asteroids and the comets originated from a former planet between Mars and Jupiter. I also engage in research in nutrition and biochemistry and am a medical member of the American Schizophrenia Association."

HARLOW J. HODGSON ("Forage Crops") is an agricultural consultant and writer who currently serves as chairman of a task force conducting a study titled "Ruminants in Support of Man" for the Winrock International Livestock Research and Training Center in Arkansas. Before his retirement last year from the U.S. Department of Agriculture he was principal agronomist for the USDA's Cooperative State Research Service. Hodgson was born and raised on a dairy farm in southern Wisconsin. He has an undergraduate degree in agronomy from the University of Wisconsin and graduate degrees (an M.S. and a Ph.D.) in plant breeding from Iowa State University. He has worked at various times as a soil conservationist for the USDA, a plant breeder in Georgia, a pioneer in subarctic crop research in Alaska, an owner-operator of a diversified livestock farm in southern Iowa, a professor of biology and a hospital administrator. In his last stint at the USDA he was responsible for, among other things, planning cooperative efforts in agricultural research involving the U.S. and other countries, including the U.S.S.R. and Japan.

JAMES S. ALBUS and JOHN M. EVANS, JR. ("Robot Systems"), work in the Office of Developmental Automation and Control Technology at the National Bureau of Standards. Albus, who was associated with the National Aeronautics and Space Administration for 15 years before he joined the National Bureau of Standards, studied physics as an undergraduate at Wheaton College and electrical engineering as a graduate student at Ohio State University and the University of Maryland; he received his Ph.D. from the latter institution in 1972. He has written several theoretical papers on the neurological mechanisms of the brain and has just finished a book describing a possible future society based on robot labor. Evans, who joined the Bureau of Standards in 1970, is a graduate of Yale University and has a Ph.D. in physics from the University of Colorado. In addition to his primary responsibilities at the bureau, he has in the past year served as chairman of a session at the Fifth International Symposium on Industrial Robots, as executive secretary of a Federal Task Group to consider a standard programming language for machine tools, as chairman of the Numerical Control Society's Standards Committee and as a member of the Department of Commerce's Technical Advisory Committee on Numerical Control.

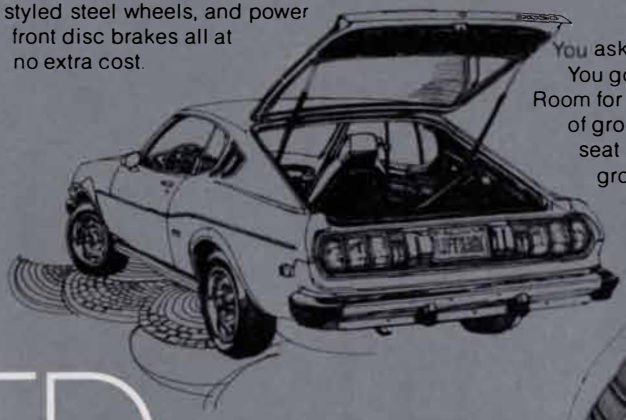
ROMUALD SCHILD ("The Final Paleolithic Settlements of the European Plain") is a senior researcher at the Institute of the History of Material Culture in the Polish Academy of Sciences. A native of Lvov, he received his M.A. from the University of Warsaw in 1957 and his Ph.D. from the Academy of Sciences in 1962. In addition to his interest in the Stone Age prehistory of Europe, he has done fieldwork in Egypt, Ethiopia and the Sudan as a member of the Combined Prehistoric Expedition, an international research body jointly sponsored by the Institute for the History of Material Culture, Southern Methodist University and the Geological Survey of Egypt.

MAITLAND JONES, JR. ("Carbenes"), is professor of chemistry at Princeton University, where he and his wife, the artist Susan Hockaday, serve as comasters of Stevenson Hall, an undergraduate college. He writes that he turned to chemistry "only after the discovery of the curve ball by my contemporaries made it clear that I would never be a major-league center fielder." His enthusiasm for the subject was kindled when, as a high school student, he worked with William Doering of Yale University at the Hickrill Chemical Research Foundation in Katonah, N.Y. Jones went on to study chemistry at Yale, earning his B.S. in 1959, his M.S. in 1960 and his Ph.D. in 1963. He joined the Princeton faculty in 1964 and has been there ever since, except for sabbatical leaves as a visiting professor at Columbia University and the Free University in Amsterdam.

D. S. SAUNDERS ("The Biological Clock of Insects") is a zoologist at the University of Edinburgh. A graduate of King's College of the University of London, he received his Ph.D. in medical entomology from the London School of Hygiene and Tropical Medicine. He writes: "For the past 12 years or so my research interests have been in the time-measuring and timekeeping abilities of animals. Before that, and for a time concurrently, I was interested in the reproductive physiology of viviparous flies, particularly the tsetse fly, an interest that took me on four occasions to Africa. Currently I am chairman and convenor of the Edinburgh University Expeditionary Society and am actively involved in planning a major scientific expedition to Ecuador in the summer of 1976."

You asked for a quality sporty car that you can afford to run. You got it. An interior which surrounds you with gauges. A 2.2 liter hemi-head coupled to a 4-speed plus overdrive fifth gear to move you on the highway at 36 miles per gallon. (1976 EPA 36 mpg on highway, 20 city. These mileage figures are estimates. The actual mileage you get will vary, depending on your driving habits, your car's condition and equipment.) MacPherson Strut front suspension, wide steel-belted radials on styled steel wheels, and power front disc brakes all at no extra cost.

You asked for GT styling with room. You got it. The Celica GT Liftback.™ Room for four people plus nine bags of groceries, or flip down the rear-seat and two of you can haul the groceries plus a side of beef.



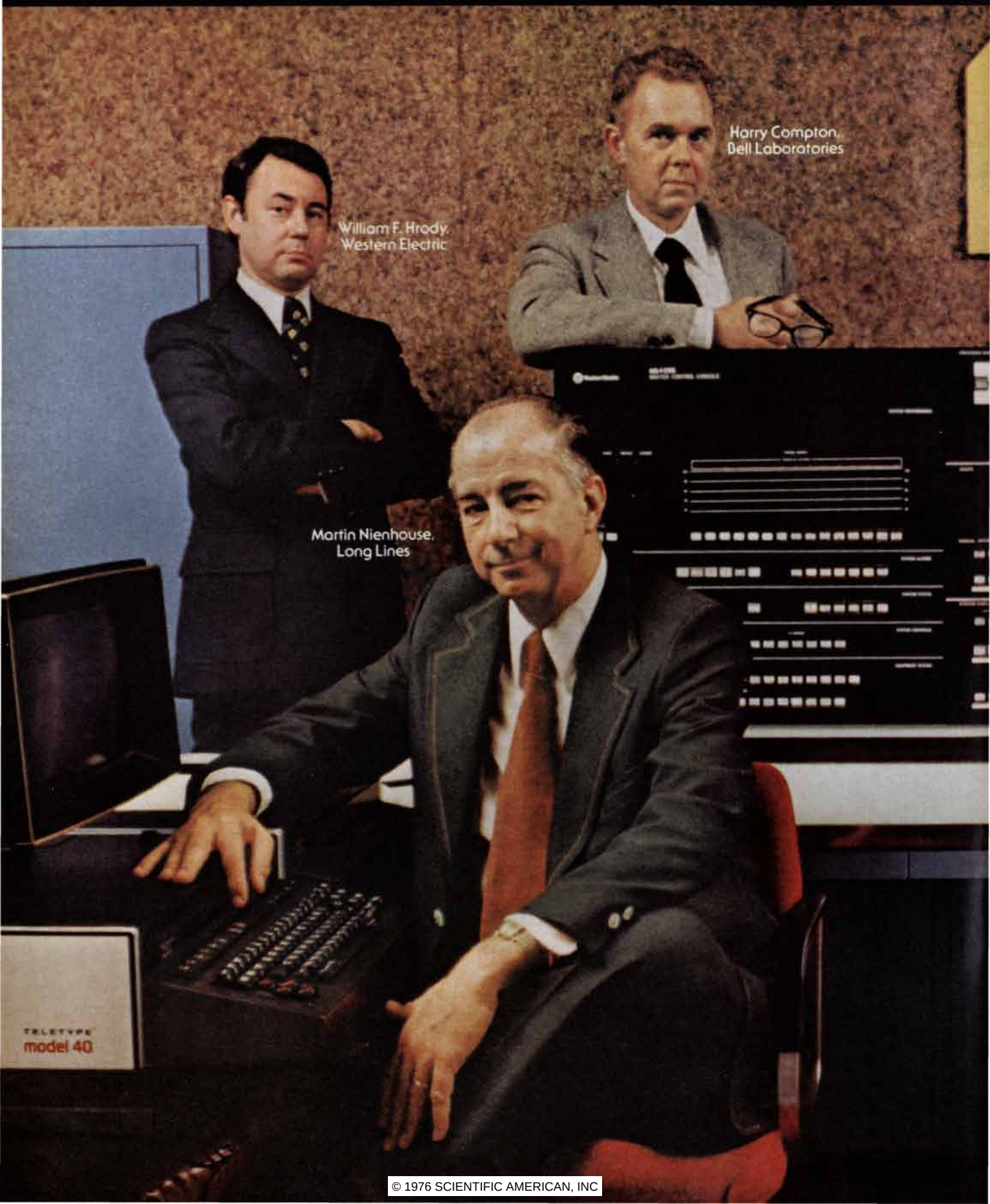
YOU ASKED FOR IT.

You asked for comfort. You got it. Glove-soft bucket seats with 120 different positions, tinted glass and AM/FM Stereo so you can boogie as you buggy.

YOU GOT IT.

TOYOTA THE NEW CELICA GT LIFTBACK

HOW WE'RE KEEPING AMERICA'S



PHONE SYSTEM THE BEST IN THE WORLD.

Electronic Long Distance Switching: A revolutionary new way to keep calls moving quickly and trouble-free.

In the years to come, we expect the number of long distance calls to increase enormously. So a team from Bell Labs, Western Electric, AT&T Long Lines and the local Bell Telephone Companies has been working on a new way to keep up with demand.

It's an Electronic Long Distance Switching System. The first unit will go on line in Chicago in early 1976.

It can handle over four times as many calls as the equipment it replaces—up to 550,000 an hour. And, if there are problems, the system can diagnose and correct many of them—automatically.

By the end of 1978, twenty units like this one in Chicago will be put in service by other Bell Telephone Companies throughout the country. It's a big project. But so is keeping America's phone system the best in the world.

Bell System

Guess what Scotch is about to be served.

CHIVAS REGAL · 12 YEARS OLD WORLDWIDE · BLENDED SCOTCH WHISKY · 86 PROOF · GENERAL WINE & SPIRITS CO., NEW YORK, N.Y.

Established 1845 SCIENTIFIC AMERICAN February 1976 Volume 234 Number 2

The Ethics of Experimentation with Human Subjects

Research with human subjects has produced advances in medicine but also some instances of ethical abuses. Studies of the attitudes and practices of investigators suggest that better controls are required

by Bernard Barber

The power, scope and funding of biomedical research have expanded enormously in the past 40 years. So also, inevitably, has clinical research with human subjects. That expansion has led in the past decade to widespread reflection on what is increasingly perceived as a new social problem: the abuse of human subjects of medical experimentation. In particular it is alleged that human subjects are not always protected from undue risk and do not always have the opportunity to voluntarily give their adequately informed consent to participation in experiments.

A social problem is defined in part by the concern it arouses, and this one has clearly aroused concern. Members of the medical profession itself led the way, with increasing numbers of journal articles, books and seminars on the issues. The public has become aroused, largely through popular accounts of dramatic incidents—genuine scandals in certain cases—involving the violation of the dignity and rights of patients. And the Federal Government has moved to protect human subjects, potential or actual. Beginning in 1966 the National Institutes of Health, the Food and Drug Administration and the Department of Health, Education, and Welfare have issued increasingly detailed regulations governing experimentation with human subjects in projects they support, which means in most of the biomedical research done in the country. In 1974 a National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research was established to advise the Department of Health, Education, and Welfare, and it is to be replaced by a long-term National Advisory Council that is to deal with the same issues.

The regulations, commissions and councils and the very fact of interference in medical activities by outsiders are viewed by many investigators as being onerous and even dangerous. On the other hand, many outsiders believe far more social control is required. The debate on the issue has been conducted without much reference to objective evidence. In 1970 our Research Group on Human Experimentation undertook two studies of investigators' attitudes and practices. On the basis of our results I would argue that there is indeed inadequate ethical concern among biomedical investigators, that it is reflected in excessively risky procedures and that better internal and external controls are essential.

There are two major reasons for the general recognition that experimentation with humans is a subject for concern, one of which I alluded to at the outset: the increased power, scope and funding of biomedical research. The other reason is a change in values: increased emphasis on equality, participation and the challenging of arbitrary authority.

It is easy to forget how new scientific medicine is. The revolutionary advances based on knowledge of physiology and biochemistry have come in the past 40 years, and they came from research. The basic work could be done with test-tube preparations and laboratory animals, but eventually human subjects had to be involved. Man is "the final test site," as Henry K. Beecher, a pioneer among physicians concerned about the ethics of research, once put it. Unfortunately there are no statistics on the number of people who are subjects in medical experiments or even on how many projects involve human subjects; the National Institutes of Health keeps records according to area of research (a disease or a physiological process, for example) rather than according to species of experimental subject; the NIH can say only that recently about a third of the projects it approves involve human subjects. It is clear, however, that the number of human subjects is larger than it used to be and that some small but significant minority of those subjects are involved in risky experiments. If more people have been put at more risk, then there is a rational basis for concern about the satisfactory balancing of risks and benefits, about adequate protection from unnecessary risk

45. A researcher plans to study bone metabolism in children suffering from a serious bone disease. He intends to determine the degree of appropriation of calcium into the bone by using radioactive calcium. In order to make an adequate comparison, he intends to use some healthy children as controls, and he plans to obtain the consent of the parents of both groups of children after explaining to them the nature and purposes of the investigation and the short and long-term risks to their children. Evidence from animals and earlier studies in humans indicates that the size of the radioactive dose to be administered here would only very slightly (say, by 5-10 chances in a million) increase the probability of the subjects involved contracting leukemia or experiencing other problems in the long run. While there are no definitive data as yet on the incidence of leukemia in children, a number of doctors and statistical sources indicate that the rate is about 250/million in persons under 18 years of age. Assume for the purpose of this question that the incidence of the bone disease being discussed is about the same as that for leukemia in children under 18 years of age. The investigation, if successful, would add greatly to medical knowledge regarding this particular bone disease, but the administration of the radioactive calcium would not be of immediate therapeutic benefit for either group of children. The results of the investigation may, however, eventually benefit the group of children suffering from the bone disease. Please assume for the purposes of this question that there is no other method that would produce the data the researcher desires. The researcher is known to be highly competent in this area.

45A. Hypothetically assuming that you constitute an institutional review "committee of one," and that the proposed investigation has never been done before, please check the lowest probability that you would consider acceptable for your approval of the proposed investigation. (Check only one)

( ) 1. If the chances are 1 in 10 that the investigation will lead to an important medical discovery.

( ) 2. If the chances are 3 in 10 that the investigation will lead to an important medical discovery.

( ) 3. If the chances are 5 in 10 that the investigation will lead to an important medical discovery.

( ) 4. If the chances are 7 in 10 that the investigation will lead to an important medical discovery.

( ) 5. If the chances are 9 in 10 that the investigation will lead to an important medical discovery.

( ) 6. Place a check here if you feel that, as the proposal stands, the researcher should not attempt the investigation, no matter what the probability that an important medical discovery will result. (IF YOU CHECKED HERE, please explain): _____

_____

45B. Which of the above responses comes closest to what you feel the existing institutional review committee in your institution would make? ____ (Please write in the number of the response.)

45C. Which of the above responses comes closest to what you feel the majority of the researchers in your institution would make, acting in their role as researcher rather than as a "committee of one"? ____ (Please write in the number of the response.)

HYPOTHETICAL EXPERIMENT described here was one of six experiments submitted to investigators and administrators in hospitals and other research centers in a mailed questionnaire. In each case respondents were asked whether, under specified conditions, they would approve of the experiment. This proposal involved giving radioactive calcium to children with a bone disease and to a control group and measuring its uptake by bone.

and about some groups being put at more risk than other groups.

Over and beyond this utilitarian basis for the new social concern with medical experimentation is the value factor, which arises from recent social changes. All over the world individuals have been demanding more equality of treatment and the right to be informed about and to participate in decisions affecting them and have been challenging the right of experts to make those decisions unilaterally. People who define themselves as being unequal, underprivileged or exploited are demanding better treatment and better protection, whether it is underdeveloped countries as against developed ones, blacks as against whites, women as against men, young as against old, patients as against doctors—or subjects as against investigators. This moral revolution of rising value-expectations has combined with the revolution in medicine to focus attention on the ethics of experimentation with human subjects.

Public awareness of the problem is too much the result of headlined scandals, but the scandals do illustrate some of the possible abuses. In the 1960's two respected cancer investigators who were studying the immune response to malignancies injected live cancer cells into a number of geriatric patients at the Jewish Hospital and Medical Center of Brooklyn without first obtaining the patients' informed consent. A few years later a leading virologist conducted an experiment at Willowbrook, a New York State institution for the severely retarded. Reasoning that a serious liver infection, hepatitis, was in effect endemic in the hospital anyway, he deliberately exposed some children to hepatitis virus in an attempt to achieve controlled conditions for testing a vaccine. The accusation was that the children's parents were not given enough information on which to base informed consent, and that in some cases consent was given perfunctorily by administrators of the institution.

More recently there was the exposure by the press of the ongoing syphilis experiment in Tuskegee, Ala. Since the 1930's a group of black subjects with syphilis had been kept under observation in an effort to study the course of the disease. That was not considered wrong in the 1930's, when the known treatments for the disease were only marginally effective, but by 1945 penicillin had become available as a safe and extremely effective cure for syphilis. Yet somehow the experiment was continued, and presumably some men died of the disease who could have been cured.

How significant are such scandals?

We do not know, because no one has been doing the kind of social bookkeeping about numbers of subjects, degree of risk, adequacy of consent and efficacy of protective mechanisms that would yield an overall view of experimentation with human beings and that might contradict the more extreme allegations of abuse elicited by the publicized scandals. In the absence of such intensive record keeping it remains for social research to fill the gap by sampling the total range of experimentation with human subjects. To that end our group conducted first a national mail survey of nearly 300 biomedical research institutions and then an intensive interview study of 350 individual investigators at two institutions.

Our national survey questionnaire was answered by 293 teaching and nonteaching hospitals and other research institutions that, our analysis showed, constituted a nationally representative sample of all such institutions. Those who filled out the questionnaire were generally themselves active researchers and members of their institution's review committee, set up to pass on research proposals. We asked the investigators to give us their response to six simulated proposals such as those that might come before a review committee. The proposals were detailed research protocols designed to measure the degree of the investigators' concern about informed consent and their willingness to approve of studies involving various levels of risk. We could be confident that the protocols were "hypothetical-actual" rather than "hypothetical-fantastic" because we constructed them with careful attention to the research literature, checked them with specialists and pretested them with a dozen chiefs of research at medical centers, who found them to be convincingly real.

One protocol described a study of chromosome breakage in users of hallucinogenic drugs; blood samples (for chromosomes) and urine samples (for evidence of drug use) were to be taken, at no risk but also without notification of the experimental purpose, from students routinely visiting the university health center. Another protocol proposed that the thymus gland, which is a component of the immune system, be removed unnecessarily from a random sample of children undergoing heart surgery; the objective was to learn the effect of the thymectomy on the survival of an experimental skin graft made at the same time. The other protocols dealt with a random test of alternative treatments for a congenital heart defect in children; with an evaluation of the efficacy of a new drug for severe depression (placebos were given to some patients); with a study of lung function in patients kept under unnecessarily prolonged anesthesia after undergoing a routine hernia repair, and with an investigation of the effect of radioactive calcium on bone metabolism in children [*see illustrations on these two pages*].

44. It has been shown that the thymus has an important bearing on the development and maintenance of immunity. For this reason the researcher proposes an investigation to determine the effect of thymus removal on the survival of tissue transplants, a very timely and important problem. In a sample of children and adolescents admitted for surgery to correct congenital heart lesions, he would randomly select an experimental group for thymectomy. Though the thymectomy will prolong the heart surgery by a few minutes, there is otherwise extremely little additional surgical risk from this procedure. At the conclusion of each heart operation, a full-thickness skin graft, approximately one cm. in diameter and obtained from an unrelated adult donor, would be sutured in place on the chest wall of both the experimental and control groups. He would then compare the survival of the skin grafts in each of the groups. It has been shown in a number of investigations of neonatal rats and other animals that those whose thymus had been removed were much less likely to reject skin grafts. The possible long-term immunological problems that might result are as yet not completely known, but a number of studies in animals indicate significant immunological deficiencies after thymectomy. Studies done in humans with myasthenia gravis, some of whom had undergone thymectomy, have not definitively demonstrated that the immunological abnormalities discovered in these patients were the result of thymectomies. To quote one authority: "There were no immunologic abnormalities that could be attributed to the effect of thymectomy per se."

The research will result in no therapeutic benefits for the patients involved. The researcher plans to obtain the consent of his potential patient-volunteers and/or their parents after explaining the procedures involved in the investigation as well as the possible short-term surgical and long-term immunological hazards for the subjects.

REMOVAL OF THYMUS GLAND during heart surgery was the experimental procedure proposed in another protocol in the questionnaire. Respondents were asked if they would approve of the experiment, given various probabilities that it would show thymectomy "considerably increases the probability of tissue-transplant survival in children and adolescents."

The answers to the thymectomy, anesthesia and radioactive-calcium protocols in particular gave us measures of the respondents' attitudes toward the balancing of risks and benefits. A clear pattern emerged. In the case of the high-risk thymectomy, for example, 72 percent of the respondents said the project should not be approved no matter how high the probability was that it would establish the efficacy of thymectomy in promoting transplant survival. On the other hand, 28 percent of the respondents said they would approve the experiment; 6 percent said they would approve it even if the chance of significant results was no better than one in 10 [*see illustration on next page*]. Similarly, 54 percent were against doing the calcium study at all—but 14 percent said they would approve it even if the odds were only one in 10 that it would lead to an important medical discovery. Our basic finding was that whereas the majority of the investigators were what we called "strict" with regard to balancing risks against benefits, a significant minority were "permissive," that is, they were much more willing to accept an unsatisfactory risk-benefit ratio.

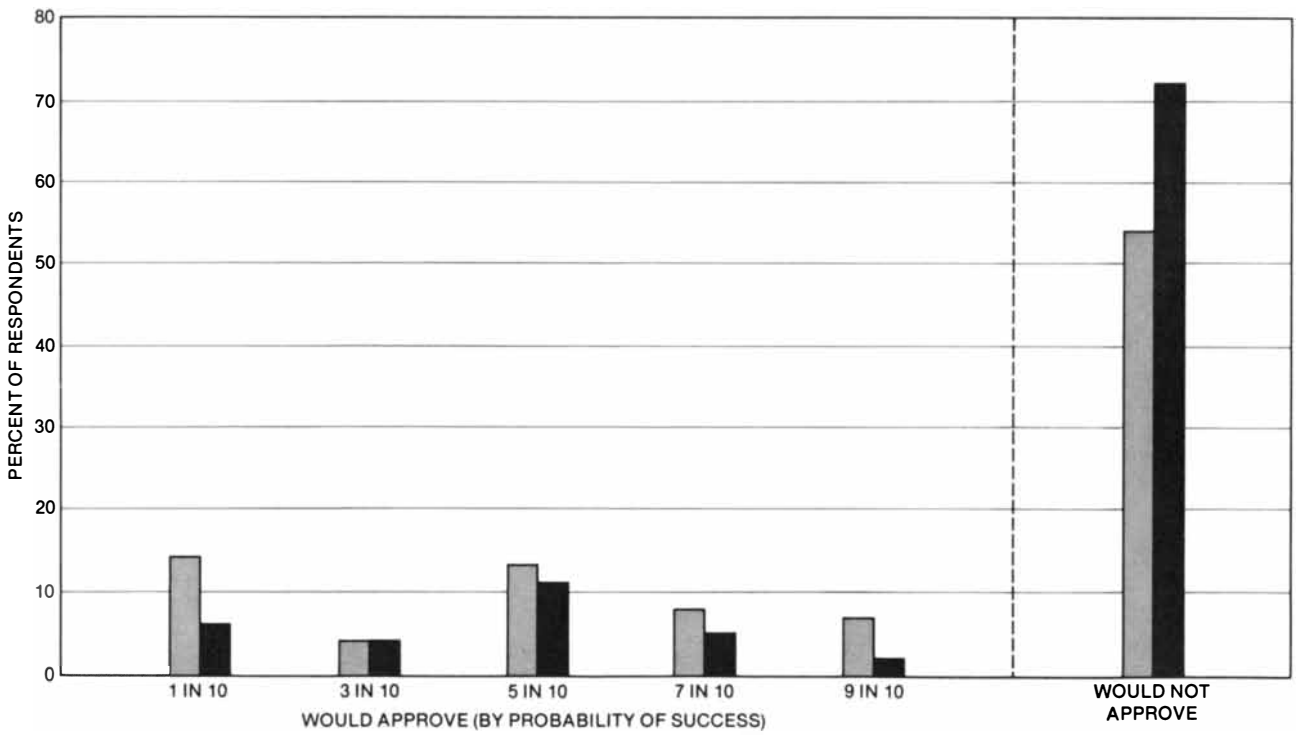
The same general pattern of a strict majority and a permissive minority emerged from our second study, in which we interviewed 350 investigators actively engaged in research with human subjects. The investigators were at institutions to which we gave the synthetic names University Hospital and Research Center and Community and Teaching Hospital. The institutions were picked (by a technique known as cluster analysis) as being representative of two kinds of medical center that do considerable amounts of research. The interviewees told us about 424 different studies involving human subjects, and for each study they estimated the risk for subjects, the potential benefit for subjects, the potential benefit for future patients and the potential scientific importance of the study. It was reassuring to find that the investigators considered that only 56 percent of the clinical investiga-

tions graded for risk and benefits involved any risk for the subjects. We went on, however, to cross-tabulate the estimated risks and benefits [*see illustration on opposite page*], and we concluded that in 18 percent of the studies the risk was not adequately counterbalanced by the benefits. We called those studies the "less favorable" ones, and we proceeded to classify them further according to their potential benefits for other patients or for medical science. Even when these compensating justifications were taken into account, tabulation revealed a "least favorable" category of studies in which the poor immediate risk-benefit ratio was not compensated for by possible future benefits. These "least favorable" investigations constituted 8 percent of the investigations in our analysis.

The concept of informed consent is a troublesome one. The investigator wants to have enough subjects and is afraid of scaring them off. Patients are likely to be concerned about their own condition, may feel powerless with respect to the physician or hospital and often have difficulty understanding medical language or concepts. Even established medical procedures can have somewhat unpredictable consequences, so that physicians feel there is a limit to how completely "informed" a patient can be. The fact remains that regulations of Government funding agencies and most institutions now require that the human subject of an experiment (or his guardian, in the case of small children and mentally incompetent patients) understand that something is being done (or some treatment is being withheld) for reasons other than immediate therapeutic ones; the subject or guardian must be informed of any risks and must give consent voluntarily.

With regard to informed consent, our questionnaires and interviews again revealed a minority with "permissive" views and practices, although that minority was smaller than it was for unfavorable risk-benefit ratios. For example, 23 percent of the questionnaire respondents said they would approve the chromosome-break proposal, which presented the informed-consent issue clearly and in effect by itself. The situation was more complex in the heart-defect protocol. Here other dubious elements competed with the fact that the investigator would not inform the parents that his decision whether or not to operate would be a random one, not based on therapeutic considerations. Only 12 percent of our respondents said they would approve of the study without requiring any revisions, but only 65 percent specifically mentioned the lack of informed consent as a problem.

The best available research evidence on informed consent comes from a study conducted by Bradford H. Gray, who was then a graduate student at Yale University, at a distinguished university hospital and research center (not the one in our interview study). With the consent of the responsible investigator, Gray interviewed 51 women who were the subjects in a study of the effects of a new labor-inducing drug. Although the women had signed a consent form, often in the hectic course of the admitting procedure or in the labor room itself, 20 of them (39 percent) learned only from Gray's interview, which was held after the drug infusion had been started or even after the delivery, that they were the subjects of research. Among those who did know, most of them did not understand at least one aspect of the study: that there might be hazards, that it was a double-blind experiment, that they would be subjected to special monitoring and test procedures or that they were not required to participate; four of the women said they would have refused to participate if they had known there was any choice. Many of the women had been



REACTION OF RESPONDENTS to the two hypothetical-experiment questions illustrated on the preceding two pages is shown: the calcium study (*light gray bars*) and the experimental removal of the thymus gland for a skin-transplant study (*dark gray*). "As the proposal stands," 54 percent of the respondents would refuse to approve the calcium study and 72 percent would refuse to approve the thymectomy, regardless of the probability of success. Substantial minorities were much more "permissive," however.

referred for the study by their private physician, but instead of being informed that an experimental drug was to be administered they were told that it would be a "new" drug; they trusted their doctor and assumed that "new" meant "better."

How does it happen that the treatment of human subjects is sometimes less than ethical, even in some of the most respected university-hospital centers? We think the abuses can be traced to defects in the training of physicians and in the screening and monitoring of research by review committees, and also to a fundamental tension between investigation and therapy. We have data bearing on each of these causative factors.

It is in medical school that the profession's central and most serious concerns are presumably given time and place and that its basic knowledge and values are instilled. Yet the evidence from our interviews shows that there is not much training in research ethics in medical school. Of the more than 300 investigators who responded to questions in this area, only 13 percent reported they had been exposed in medical school to part of a course, a seminar or even a single lecture devoted to the ethical issues involved in experimentation with human subjects; only one respondent said he had taken an entire course dealing with the issues. Another 13 percent reported that the subject had come to their attention when, as students, they did practice procedures on one another; for 24 percent it was in the course of experiments with animals; 34 percent remembered discussion of ethical issues in specific research projects. One or more of these learning experiences were reported by 43 percent of the respondents—but the remaining 57 percent reported not a single such experience. The figures were about the same whether the investigators were graduates of elite U.S. medical schools, other U.S. schools or foreign schools. The figures were a little better, however, for those who had graduated since 1950 than for older investigators.

THERAPEUTIC BENEFIT FOR SUBJECTS	RISK		
	NONE	VERY LITTLE	SOME, MODERATE OR LARGE
MINOR, LITTLE OR NONE	11	14	2
SOME	14	12	2
GREAT	10	19	7

RISKS AND BENEFITS were cross-tabulated for some 400 current research projects reported by investigators in two hospitals. Studies falling on or below the diagonal were considered to have risks for subjects that were more or less counterbalanced by benefits for subjects. (In 9 percent of the studies respondents reported no risk and were not asked about benefits.) Studies above the diagonal (*colored boxes*) were classified as "less favorable" for their subjects: they contained risks for subjects and, according to the investigators, offered relatively low benefits. These cases, 18 percent of the total, were further subdivided (in a table not reproduced here) according to benefit for others or for science. Studies that were low in those justifications (8 percent of total) were called "least favorable."

What little ethics training there is is apparently not very effective: the investigators who reported having learned something about research ethics were only slightly less permissive in response to protocols presenting the risk-benefit issue than those who reported no such experiences. It would appear that both the amount and the quality of medical-school training in the ethics of research could be improved. In this connection it is worth remembering that the many physicians who are not engaged in investigation at all also need some background in experimentation ethics, if only so they can evaluate requests that they direct their patients toward a colleague's research project.

Scientific "peer review" is a keystone of scientific inquiry, operating implicitly in many ways and explicitly in the case of professional journals, grant-awarding committees and many institutional reviewing boards such as the "tissue committees" that assess the results of surgery in hospitals. Ethical peer review of experimentation with human beings should be the counterpart of scientific peer review, but until the mid-1960's such activity received limited support among biomedical researchers. Even after 1966, when the NIH mandated ethical peer review for all its grantees, effective review did not become universal. Our questionnaire went to hospitals and other research centers that had filed with the NIH formal assurances that the required institutional review committee had been established, but 10 percent of the respondents said their institution's committee reviewed only proposals for outside funds and 5 percent reported that only formal proposals to the NIH were reviewed. The two institutions in our interview study were among the 85 percent that stated they were reviewing all research proposals, and yet 8 percent of our interviewees volunteered the information that at least one of their own investigations with human subjects had not been reviewed.

How effective are the review committees in handling the protocols that do come before them? Our questionnaire respondents told us that in 34 percent of the institutions the committees had never required any revisions, rejected any proposals or had any proposals withdrawn in anticipation of rejection for ethical reasons; 31 percent reported revisions, 32 percent outright rejections and 19 percent withdrawals. Either some of these committees have very few ethical problems coming before them or they are ineffective. Gray's study in an institution with an active and strong committee suggests that they are ineffective rather than underworked. The committee whose performance he examined found relatively few proposals that did not need some kind of modification, and he thinks "a record of few actions by committees is an indication that their members are indifferent or that their standards are loose."

The peer-review groups seemed weak in other ways. In some institutions there was no face-to-face discussion among the reviewers. Only 22 percent of the committees had members from outside the institution, something that was then recommended and has since been mandated by the Department of Health, Education, and Welfare. In practically none of the institutions was there continuous monitoring of studies that were approved, although this was even then required by Government regulations. In general ethical peer review is hampered by the fact that each committee operates in isolation and must consider every new issue on its own and without benefit of precedent. A case-reporting system, such as operates in the law, would make that

unnecessary and would promote both equity among institutions and high standards. The major weakness in the system is the lack of keen interest in and support of the review committees on the part of most working biomedical investigators. Research is their business; research is their mission and predominant interest, not applied ethics or active advocacy of patients' rights.

Most biomedical investigators are, however, interested in taking care of patients and making them well. As a result medical institutions and individual investigators operate today with two powerful sets of values and goals. On the one hand there is the pursuit and advancement of scientific knowledge. On the other there is the provision of humane and effective therapy for patients. Through a broad range of complex interactions these two sets of values and goals are harmonious, even complementary and mutually reinforcing. Occasionally, however, scientific research and humane therapy can be in conflict. When that happens, there is sometimes a tendency to choose the pursuit of knowledge at the expense of the ethical treatment of patients. An irreducible minimum of conflict may be inevitable. The ethical task now is to come as close as possible to that minimum—and to resolve unavoidable conflict in favor of humane therapy.

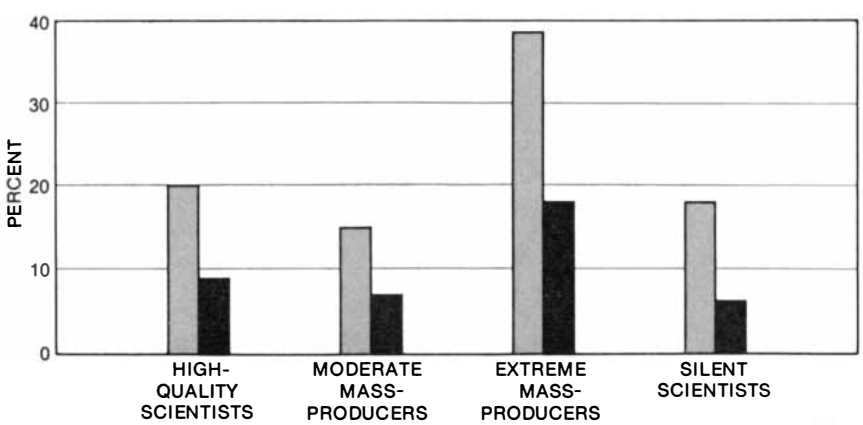
There is evidence that the enhanced excitement attending scientific achievement and the rewards bestowed on it in recent decades have skewed the decision-making process in many cases of conflict. As our data show, the medical schools have been largely indifferent to training their students in the ethics of research. Moreover, their record in peer review has been inferior to that of other institutions. Answers to our questionnaire showed they were less likely than other research centers to have set up a review committee before the NIH required one, less likely to have one that met the first NIH guidelines in 1966, less likely to have a committee that reviews all clinical research and less likely to include on their committee medical or nonmedical members from outside the institution. Medical schools, the Association of American Medical Colleges and professional associations of clinical investigators have been much quicker to seek research funds or to protest funding cuts than to organize seriously for the purpose of studying the ethics of research and making policy in that area.

The same emphasis on the pursuit of knowledge rather than on ethics is apparent among individual biomedical investigators. Ethical concern for the subjects of their research is not a major factor when they select their collaborators; at least it is not often mentioned as a characteristic they look for in collaborators. Scientific ability is a major concern. When we asked our 350 interview subjects, "What three characteristics do you most want to know about another researcher before entering into a collaborative relationship with him?" 86 percent of the respondents mentioned scientific ability, 45 percent mentioned motivation to work hard and 43 percent mentioned personality. Only 6 percent of them listed anything we could classify as "ethical concern for research subjects."

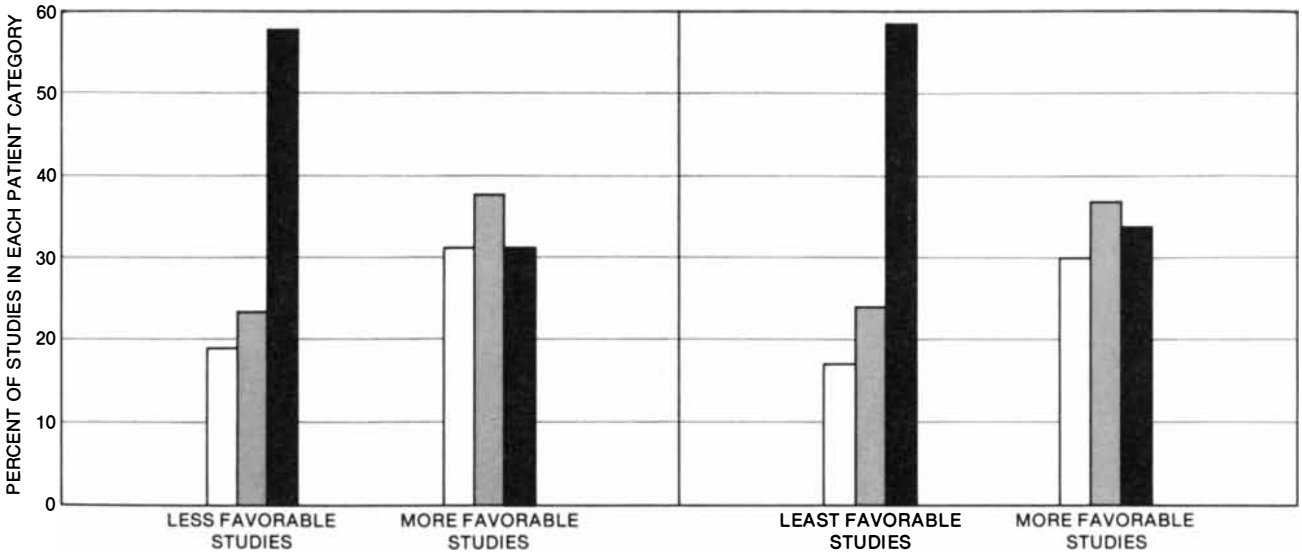
The tension between investigation and ethical concern is perhaps best illustrated by indications that the struggle for scientific priority and recognition exerts pressure on ethical considerations. Our data show that the social structure of competition and reward is one of the sources of permissive behavior in experimentation with human subjects; the relatively unsuccessful scientist, striving for recognition, was most likely to be permissive both in his approval of hypothetical protocols and in his own investigative work. We divided our respondents into four categories based on the number of papers they had published and the number of times their work had been cited by other workers; the frequency of citation has been shown to be a good measure of scientific excellence. We called the most-cited investigators the "high quality" scientists and those who had published a great deal but were never cited the "extreme mass-producer" scientists. It was the extreme mass-producers who were most often engaged in investigations with less favorable risk-benefit ratios, who approved of the protocols with poorer risk-benefit ratios and who least often expressed awareness of the importance of consent. Caught up in the socially structured competitive system of science, unsuccessful in it but still pursuing the prize of peer recognition, they appear to be more likely to overvalue scientific work as against humane therapy.

It is not only the mass-producers, contending for recognition among peers in their discipline, who are apt to be more permissive. We also weighed the rank achieved by each worker within his own institution against various measures of his effectiveness compared with that of his colleagues. We found that the "underrewarded" investigators tended to be the more permissive. There is also a quite different kind of medical investigator who we think is likely to be pushed toward permissive practices by scientific competition: some of the professionally esteemed, highly successful medical scientists who are engaged in intense competition for priority and recognition in well-publicized areas of research. There are not many of those people, and they did not emerge in our sample, although some workers who refused to be interviewed may belong in that category. In the absence of real data we can only point to such evidence as published discussions concerning the worldwide heart-transplant competition of a few years ago, which raised questions about the premature exposure of human subjects to what were then still experimental procedures.

Given the fact that there are ethical defects in current medical-research standards and practices, do the resulting



PRESSURE TO PRODUCE leads to "permissiveness." Investigators were classed as "high-quality scientists" (most cited), "moderate mass-producers" (many papers, few citations), "extreme mass-producers" (many papers, no citations) or "silent scientists" (few papers and citations). Extreme mass-producers were twice as likely as high-quality scientists to have a role in one of the less favorable (*light gray*) or least favorable (*dark gray*) studies.



DIFFERENTIAL TREATMENT of various patient categories is evident when studies identified as having the less favorable or the least favorable risk-benefit ratios were classified according to whether fewer than 50 percent of their subjects were ward or clinic patients (*white bars*), between 50 and 75 percent were (*light gray*) or more than 75 percent were (*dark gray*). The less favorable studies were almost twice as likely to have subjects a large majority of whom were ward or clinic patients (*dark gray bars*) as the more favorable studies (*left*). About the same thing was true of studies that had been identified as having least favorable ratios (*right*).

abuses strike particularly, as is often alleged, at certain social groups: at the poor, at children and at institutionalized patients (prisoners in particular)?

The evidence from our interviews with 350 investigators indicates that the poorer patients in hospitals are indeed at a disadvantage as subjects of research. For each of the 424 studies our respondents reported, they told us whether fewer than 50 percent, between 50 and 75 percent or more than 75 percent of the subjects were ward or clinic patients (as opposed to patients in private or semiprivate rooms and under the care of their own physician). We found first of all that ward and clinic patients were more likely to be subjects of experiments. Moreover, when we examined the cases we had previously identified as having "less favorable" and "least favorable" risk-benefit ratios, we found that both categories were almost twice as likely to involve subjects more than three-quarters of whom were ward and clinic patients as the studies with the more favorable ratios were.

The ward and clinic patients are, of course, vulnerable to that kind of discrimination. They can most readily be channeled into an experimental group by admitting physicians and clerks without interference from a personal physician. They tend to be less knowledgeable about hospitals, more readily intimidated and less likely to understand what they are told about an experimental project, and therefore less likely to be able to withhold their consent or to give genuinely informed consent. In sum, they are the least likely to be able to protect themselves.

Many institutionalized patients are poor and perhaps incompetent, and they may feel completely dependent on the institution's administrators and physicians. Prisoners are a special case: they are institutionalized in an implicitly coercive situation, so that genuinely informed consent may be a logical impossibility. On the other hand, a prison population is by definition a good source of experimental and control subjects living under controllable conditions, and there have been instances where prison studies have been conducted humanely, with good scientific results and apparently with good effect on the prisoners' morale. Experimentation with prisoners is nevertheless subject to grave abuses. Last summer the head of the Food and Drug Administration told a Senate committee that a review of experimentation in 19 prisons revealed abuses ranging from unprofessional supervision of drug tests to inadequate medical care and follow-up treatment.

Children constitute still another special group. Small children cannot give consent for their own participation in experiments; older children, who could, are often not asked. As the Willowbrook incident demonstrated, parents are not always adequately protective of their children's interests. In the case of institutionalized patients, prisoners and children, new regulations of the Department of Health, Education, and Welfare call for special protective committees and procedures. These will only be effective, however, in a context of better ethical training for investigators and more effective peer review.

The ethical problems that attend medical research with human subjects are representative of an entire class of problems created by the impact of professionals and professional power on the general public and on public policy. In the area of research with human subjects the medical investigators are not alone; there is a tendency in other fields too for humane concerns to be left at the laboratory door. Psychologists and sociologists have often been accused of circumventing the requirement for consent and of applying unethical manipulative techniques in their investigations of human behavior, and neither profession has welcomed scrutiny from outsiders or restrictive regulation. The issue goes beyond research ethics, however. Many professions now command knowledge that has great potential usefulness for human welfare but bestows power that can be abused. Because professional power is largely based on knowledge that has not yet diffused to the general public it must to a considerable degree be self-regulated, but because professional power is of such major public consequence it must also be subject to significant public control. The medical-research profession does not have a proud record of self-regulation or acceptance of public controls.

THE RECEPTORS OF STEROID HORMONES

The cellular response to these hormones depends on the presence of protein molecules called receptors. The complex formed by the hormone and the receptor acts on the genetic material of the cell

by Bert W. O'Malley and William T. Schrader

Hormones secreted by the endocrine glands can influence the functioning of cells in distant and apparently unrelated tissues. In the subtle effects of these regulatory molecules there is a puzzling specificity: all cells are exposed to hormones, yet only a few respond to them. The effects of one class of hormones—the steroid hormones—have been particularly perplexing. The steroids are small molecules, and it is not obvious how they can incorporate enough information in their structure to ensure adequate specificity and to account for their diverse influences.

Interest in the steroid hormones extends beyond endocrinology because it has been found that the hormones stimulate cells by controlling the synthesis of particular proteins. Protein synthesis is a central process in the metabolism of the cell, and it directly expresses the information in the genes. One of the principal preoccupations of molecular biology has been the search for the mechanisms that control gene expression and thereby determine overall patterns of growth and development.

In recent years the part played by the steroid hormones in the regulation of protein synthesis has been worked out in considerable detail, although a few crucial events have remained obscure. The functioning of the hormones depends on their interaction inside the cell with protein molecules designated receptors. Each of the steroid hormones affects only a few tissues because only the cells of those tissues contain the appropriate receptor proteins. The complex formed by the hormone and its receptor controls protein synthesis by acting directly on the genetic material of the cell.

There are two large classes of hormones, the peptides and the steroids, and the two kinds of molecule seem to function through quite different mechanisms. The peptide hormones can be regarded as small proteins; each consists of a sequence of amino acids linked by characteristic chemical bonds called peptide bonds. A typical peptide hormone is insulin, which is secreted by specialized cells in the pancreas; human insulin consists of 51 amino acid units.

The peptide hormones are comparatively large molecules, typically with a molecular weight of about 10,000. Because of their size they cannot readily pass through the plasma membrane that surrounds all cells, and indeed they probably never enter the cells they stimulate. Instead they act on molecules on the surface of the plasma membrane, which then transmit signals to the interior of the cell. The peptide hormones can influence a great variety of intracellular processes, but only indirectly.

The steroid hormones are the subject of the remainder of this article. They include the male sex hormones (collectively called the androgens), the female sex hormones (the estrogens and the progestins) and the hormones secreted by the cortex, or outer layer, of the adrenal glands (the corticosteroids). They typically have a molecular weight of about 300, which is roughly the same as that of common table sugar.

All the steroids are based on a common central structure of four interconnected rings of carbon atoms; three of the rings have six carbon atoms and the other ring has five atoms. The differences between the various steroids are determined by the pattern of chemical bonds within the rings and by the nature and orientation of the side groups attached to the rings [*see illustration on pages 34 and 35*]. Although the differences may appear to be superficial, they alter the molecules' shape and can thereby completely change their biological activity. For example, testosterone and progesterone are identical except for one side group attached to one of the rings, yet the two hormones differ radically in function: testosterone governs the development of the male secondary sex characteristics and progesterone contributes to the maintenance of pregnancy.

The structural similarities of the steroid hormones are reflected in their origin. All of them are synthesized from the same chemical precursor, cholesterol. Although cholesterol has no known hormonal function, it is a steroid and has the characteristic four-ring structure. The hormones are manufactured by altering the side groups of cholesterol; the modifications are made by enzymes in the cells of endocrine glands.

Since the steroid hormones are small molecules, they can easily diffuse into cells of many kinds, and they can just as easily diffuse back out into the bloodstream. They are effective even in very small quantities; their concentration in the blood reaches only about 10^{-9} mole per liter. (One mole is an amount of a substance in grams equal to its molecular weight.) If the palate were as sensitive to flavors as the target cells are to hormones, we would be able to detect a pinch of sugar dissolved in a swimming pool.

Any theory explaining the action of the steroid hormones must account for their effectiveness in those minuscule concentrations. It must also explain how the relatively simple steroid molecules, which have few distinctive chemical groups and which resemble one another in structure quite closely, can have such a broad spectrum of biological effects. Finally, it must explain how molecules that enter virtually all cells can selectively stimulate a small population of them and pass through the rest with no apparent effect.

The first important clue to the functioning of the steroid hormones came from studies of bacteria. It was in bacteria that a mechanism regulating the expression of genes was first discovered.

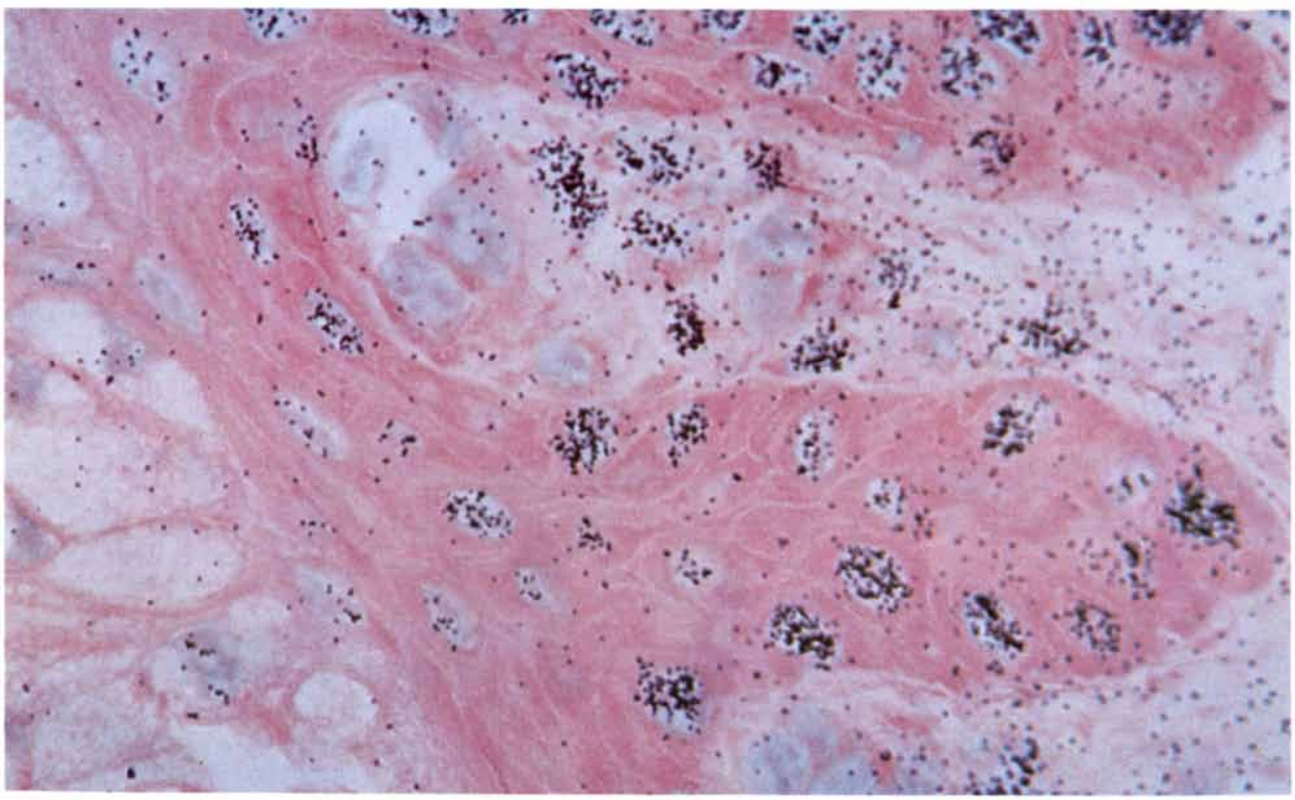
The sequence of events in protein synthesis is essentially the same in bacteria and in higher organisms: it is the process in which information is transferred from the DNA, the genetic "master tape" of the cell, to a molecule of messenger RNA, a segment of "tape" carrying instructions for the synthesis of the protein. The DNA molecule, as is now well known, is a double-strand helix held together by bonds between pairs of nucleotide bases. The four kinds of base are always paired in the same way: adenosine bonds only to thymidine and guanosine bonds only to cytidine. The two strands are therefore complementary; given the sequence of bases in one strand, the sequence in the other is completely defined.

The genetic information is expressed when a segment of DNA representing a gene is transcribed to yield a corresponding length of messenger RNA, a process controlled by the enzyme RNA polymerase. The RNA, which is a single-strand molecule (and which also differs from DNA in that all the thymidine units are replaced by uridine), then becomes associated with the organelles called ribosomes. At the ribosomes the RNA is translated, each group of three bases specifying one amino acid, which is added to the protein molecule. Through that mechanism the sequence of bases in a segment of DNA specifies a sequence of amino acids, and this sequence determines all the properties of the protein.

The genetic code explains how a gene can specify the structure of a protein, but it cannot explain the particular assortment of proteins that is synthesized by a cell. In a bacterial cell there are thousands of genes, but only a few function at any given time. Systems regulating the expression of genes must therefore exist.

Our understanding of gene regulation is still far from complete, but one mechanism by which it is accomplished, at least in bacteria, was explicated in 1959 by the elegant studies of Francois Jacob and Jacques Monod of the Pasteur Institute in Paris. Working with the colon bacterium *Escherichia coli,* Jacob and Monod showed that a system of genes concerned with the metabolism of the sugar lactose is regulated by a specific protein. When the protein is present on the DNA, transcription is prevented: the synthesis of new messenger RNA from that gene cannot take place. They therefore called the protein a repressor. They found, however, that when lactose was added to the culture medium in which the bacteria were grown, the sugar in some way removed the repressor from the DNA, allowing transcription to begin. They called lactose an inducer for this system of genes, and called the regulatory process derepression.

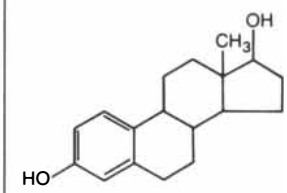
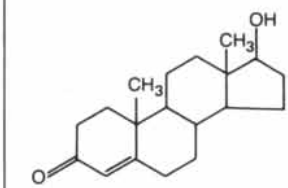
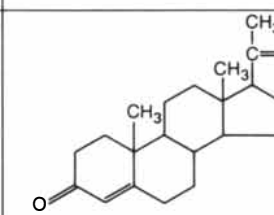
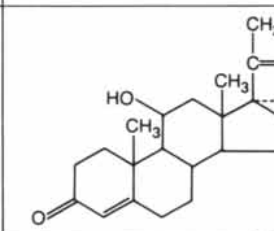
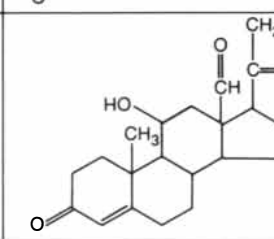
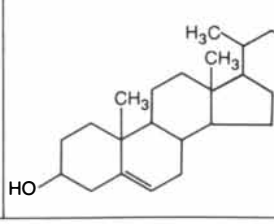
The model devised by Jacob and Monod represents one system through



SPECIFICITY OF STEROID HORMONES for particular cells and the localization of the hormones within the nuclei of those cells are demonstrated by the technique of autoradiography. The black dots reveal the location of radioactively labeled molecules of the steroid hormone progesterone in a section of tissue from the uterine cervix of a guinea pig. The dots are clustered over the nuclei of cells and are sparsely distributed elsewhere. The effect is a result of the activities of progesterone receptor molecules. The hormone binds to the receptors in the cytoplasm of the cell, forming a complex that is then sequestered in the nucleus. In the making of the autoradiograph the guinea pig was first injected with progesterone labeled with tritium, the radioactive isotope of hydrogen. After 15 minutes the animal was killed and thin sections of frozen tissue were mounted on glass slides coated with a photographic emulsion. The slides were kept in darkness and cold for about three months, then the emulsion was developed and the tissue was stained. The radioactive decay of the tritium atoms exposed grains in the emulsion where the hormone was concentrated. The autoradiograph was made by Walter E. Stumpf and Madhabaranda Sar at the University of North Carolina at Chapel Hill.

which a relatively small and simple molecule (the inducer) can regulate the expression of a gene. The model is of particular interest to endocrinologists because there are a number of apparent similarities between the action of an inducer on a bacterial cell and the action of a steroid hormone on the cells of higher animals. In 1964 Peter Karlson of the University of Marburg made the comparison explicit: he proposed that the steroid hormones are in fact inducers and that they act by combining with repressor proteins. In the absence of hormone the repressors were supposed to block transcription.

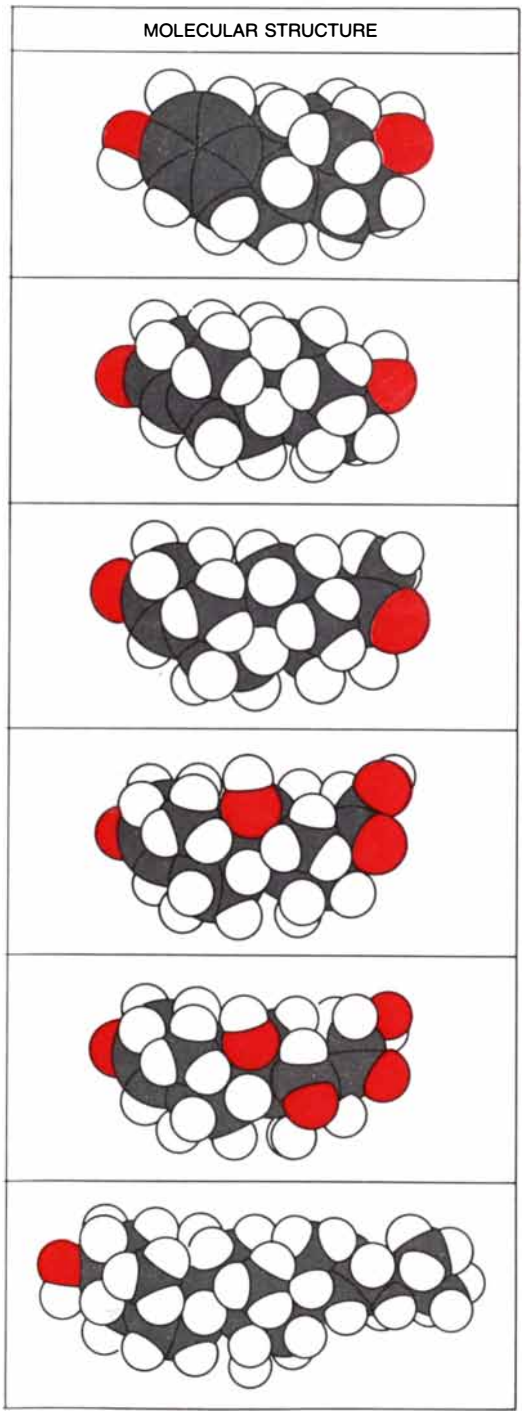
It was soon confirmed that the administration of hormone to an animal results in an increase in the synthesis of RNA in the target cells, followed by an increase in protein synthesis. These findings were entirely consistent with the hypothesis that the steroid hormones act at the level of the gene [see "Hormones and Genes," by Eric H. Davidson; SCIENTIFIC AMERICAN, June, 1965]. It had not yet been demonstrated, how-

STEROID	PRINCIPAL SOURCE	PRINCIPAL TARGET TISSUES	HORMONAL FUNCTION	CHEMICAL FORMULA
ESTROGEN 18 CARBON ATOMS	OVARY	UTERUS, VAGINA, BREAST, HYPOTHALAMUS	DEVELOPMENT OF FEMALE SEX CHARACTERISTICS	
TESTOSTERONE 19 CARBON ATOMS	TESTIS	SEMINAL VESCICLES, PROSTATE, TESTIS	DEVELOPMENT OF MALE SEX CHARACTERISTICS	
PROGESTERONE 21 CARBON ATOMS	OVARY, PLACENTA	UTERUS, OVIDUCT	MAINTENANCE OF PREGNANCY	
HYDROCORTISONE 21 CARBON ATOMS	CORTEX OF THE ADRENAL GLAND	ALL CELLS	REGULATION OF ENERGY UTILIZATION	
ALDOSTERONE 21 CARBON ATOMS	CORTEX OF THE ADRENAL GLAND	KIDNEY	REGULATION OF ELECTROLYTE BALANCE	
CHOLESTEROL 27 CARBON ATOMS	DIET, LIVER	UNKNOWN	METABOLIC PRECURSOR OF ALL STEROIDS	

STEROID HORMONES are relatively small molecules, with a typical molecular weight of about 300. All of them share a common core structure, consisting of four interconnected rings of carbon atoms. They differ in the pattern of bonding within the rings, and in the side groups attached to them. The structural differences subtly alter the shapes of the hormone molecules, as is shown in the drawings at right; it is principally on the basis of shape that the hormones are recognized by their receptors. Only a few of the

ever, exactly how the hormones activate a gene. Were they inducers, strictly analogous to those discovered by Jacob and Monod in bacteria, or did they function through some more complicated mechanism?

In considering this question it is important to realize that although bacteria and higher organisms share the same genetic code, they differ fundamentally in organization. Bacteria are prokaryotic cells, that is, their genetic material is distributed throughout the cytoplasm. The cells of all the higher plants and animals are eukaryotic: they have a distinct nucleus, which confines the DNA within a specialized membrane. Within the nucleus the DNA is associated in a chemical complex with a number of proteins that are not found in prokaryotes. Moreover, in multicellular organisms the problem of gene regulation is intrinsically more complex than it is in bacteria. In the higher organisms each cell has an identical complement of DNA, but different cells make quite different kinds of protein. A cell in the pancreas, for example, might make insulin, whereas one in the blood-forming tissue might manufacture hemoglobin.



many steroids are shown. Cholesterol is not ordinarily considered a hormone, but it is the precursor from which all other steroids are synthesized in the endocrine glands.

In view of these complications it seems reasonable to suppose the mechanisms of gene regulation in the higher plants and animals differ from those found in bacteria. In the case of the steroid hormones that supposition has been confirmed. Although the steroids do induce gene expression, they do not do so by combining with repressors of the kind described by Jacob and Monod and they do not cause derepression; on the contrary, they become coupled to receptor proteins that have a positive regulatory effect. In eukaryotic cells a gene is thought to be activated when a hormone-and-receptor complex binds to it, rather than when a repressor molecule is removed.

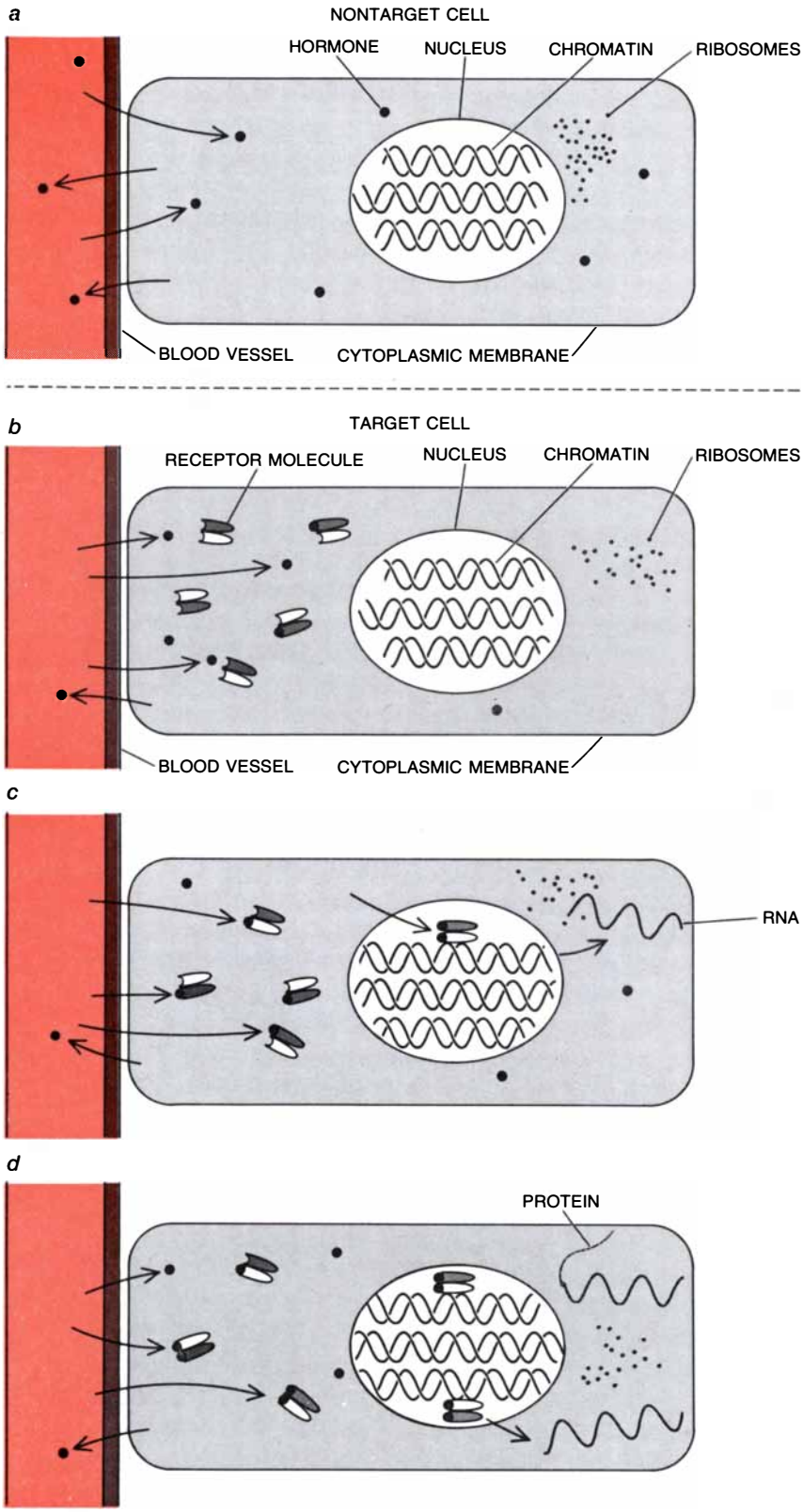
The first step in the investigation of the functioning of the steroid hormones was the demonstration that they have a particular affinity for their target cells. Since the hormones are present in such low concentration, the methods of conventional analytical chemistry are inadequate for the study of hormone physiology; instead it is necessary to employ hormones labeled with radioactive atoms. Highly radioactive steroid hormones were first prepared in the early 1960's by Elwood V. Jensen of the University of Chicago. By injecting female rats with a radioactive preparation of estrogen, Jensen and his colleagues showed that the hormone is retained longer in the uterus than it is in other tissues, such as muscle and blood. The uterus is considered a "target" organ of estrogen, whereas muscle and blood are not. Subsequently I. S. Edelman and his colleagues at the School of Medicine of the University of California at San Francisco demonstrated the retention of a steroid hormone at the cellular level. They employed the technique of autoradiography, in which a thin slice of tissue containing radioactively labeled molecules is coated with a photographic emulsion. The exposure of the emulsion by the radioactive molecules indicated that the hormone appears in the target tissues within minutes after it is administered to the animal. Moreover, it is retained there long after all the radioactive molecules have left the nontarget cells.

In 1968 Walter E. Stumpf, who is now at the University of North Carolina at Chapel Hill, proceeded one step further and showed that within the target cells the hormone tends to accumulate in the nucleus, providing additional evidence that the mechanism of steroid action is a genetic one. Stumpf too employed autoradiography, and he found that the movement of the hormone into the nucleus is extremely rapid: it precedes all other observable changes in the target cell.

The rapid movement of the steroid hormones does not necessarily imply that some mechanism exists to transport them specifically to the target cells; it is sufficient that the molecules be sequestered within those cells. In nontarget cells the hormone passes through the plasma membrane in both directions, and its concentration inside the cell cannot exceed the concentration in the bloodstream. In the target tissues the hormone molecules continue to diffuse into the cells but very few leave them, so that the concentration inside the cells increases [*see illustration on next page*].

These early investigations strongly suggested that some agent in the target cells binds to the steroid hormones and prevents their escape. Indeed, Jensen and his colleagues suggested in 1961 that such a receptor might be involved in the cellular response to steroid hormones. The hypothesis had the further advantage that it could explain how the relatively simple steroid molecules could exert such diverse effects; by forming a complex with a larger "helper" molecule they might become much more versatile. In the subsequent search for such a molecule two important criteria were established. First, any receptor molecule must be present in the target cells of the hormone but absent in all other cells. Second, the receptor molecule should have a high affinity for its particular hormone but low affinity for other steroids with different biological activity.

The identification of steroid receptor molecules was not an easy task, partly because the receptors are present in extremely small quantities (about 10,000



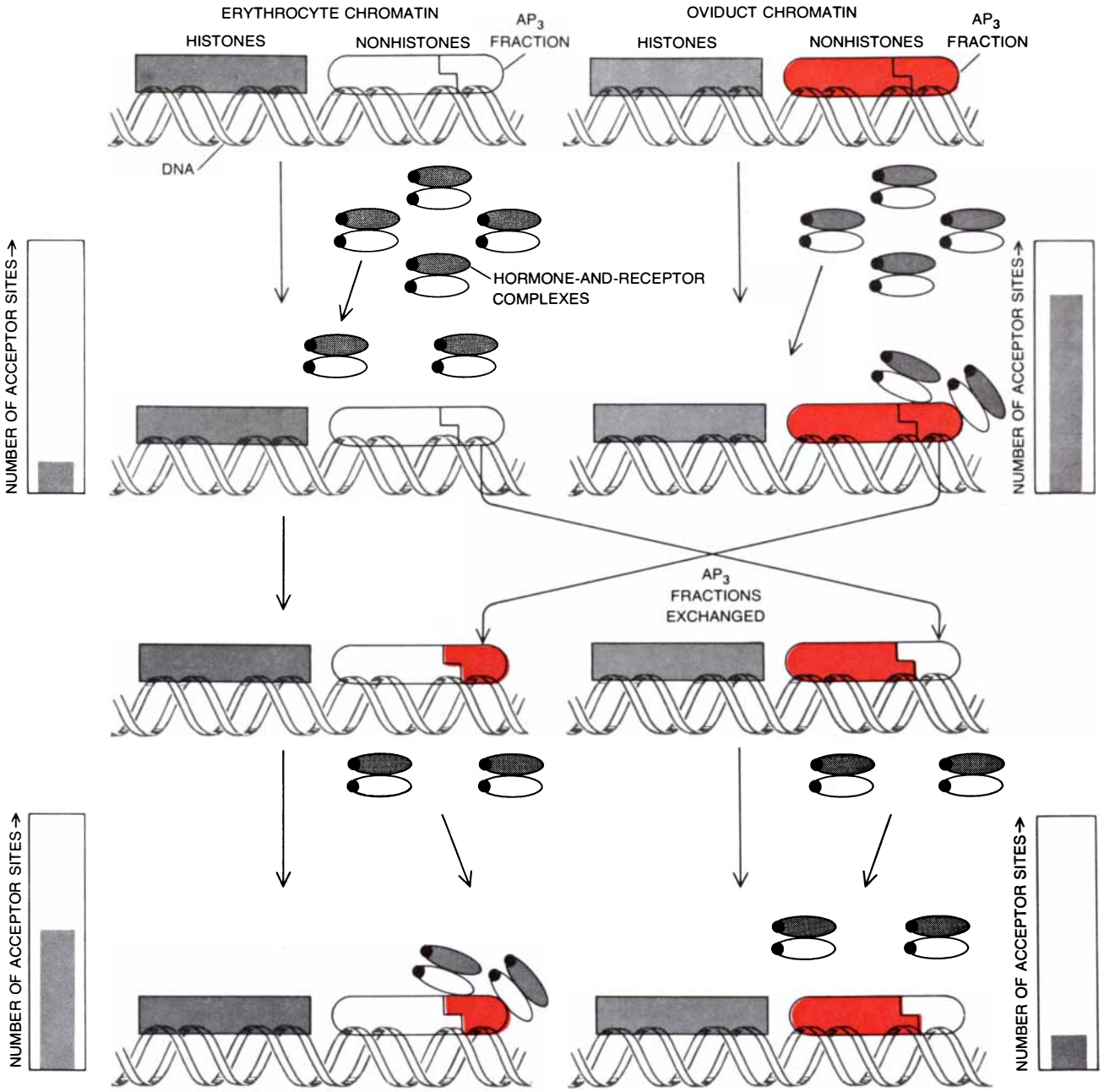
ACTION OF STEROID HORMONES is mediated by receptor molecules found only in cells that respond to the hormone (target cells). In nontarget cells (*a*) the hormone diffuses freely into the cell and out of it, and the concentration remains quite low. In a target cell the hormone is sequestered by receptor molecules (*b*). Each receptor molecule binds two molecules of hormone, forming a complex that enters the nucleus (*c*) and becomes attached to the chromatin, the genetic material. The complex stimulates the transcription of particular genes, so that RNA encoding the information in those genes is synthesized. On the organelles called ribosomes the RNA is translated into proteins (*d*).

molecules per cell). They were ultimately identified in 1966 by David O. Toft and Jack Gorski of the University of Illinois. Toft and Gorski separated hormone-and-receptor complexes from other large molecules by centrifuging an extract from the target cells in a density gradient of sucrose. The method separates molecules on the basis of their molecular weight, yielding fractions containing progressively smaller molecules. By employing a radioactively labeled hormone, the fraction containing the hormone-and-receptor complex can be identified simply as a peak in radioactivity.

Toft and Gorski employed estradiol, a form of estrogen, labeled with tritium, the radioactive isotope of hydrogen. When they tested it in tissue from the rat uterus, they found that the hormone was bound selectively to a protein molecule with a molecular weight of about 200,000. The protein met the criteria established for a hormone receptor: it was present only in the target cells and it bound estradiol very tightly. The protein also displayed an affinity for steroids closely related to estradiol and for certain other nonsteroid substances that have a different structure but a similar molecular shape. Significantly all the substances that were shown to bind to the receptor protein mimic the biological activity of estradiol.

Similar receptor proteins have since been identified in the target tissues of all the known steroid hormones. Tissues that do not respond to a given hormone invariably lack receptors for that hormone. The receptors for estrogen, androgen and the corticosteroids have been characterized in several animals. One of the most difficult of the receptors to identify was the receptor for progesterone. It was eventually found in the chick oviduct, a known progesterone target tissue, by our research group at the Baylor College of Medicine. The progesterone receptor is similar in structure and function to the other steroid receptors.

The role of the receptor proteins was confirmed in the laboratories of Jensen, Gorski and Edelman when it was shown that the receptors not only sequester the hormone molecules in the target cells but also concentrate them in the cell nucleus. We observed the movement into the nucleus by administering a hormone, then after various intervals disrupting the cells and quickly separating the nuclei from the cytoplasm in a centrifuge. We observed a steady decline in the number of hormone-and-receptor complexes in the cytoplasm and a concomi-

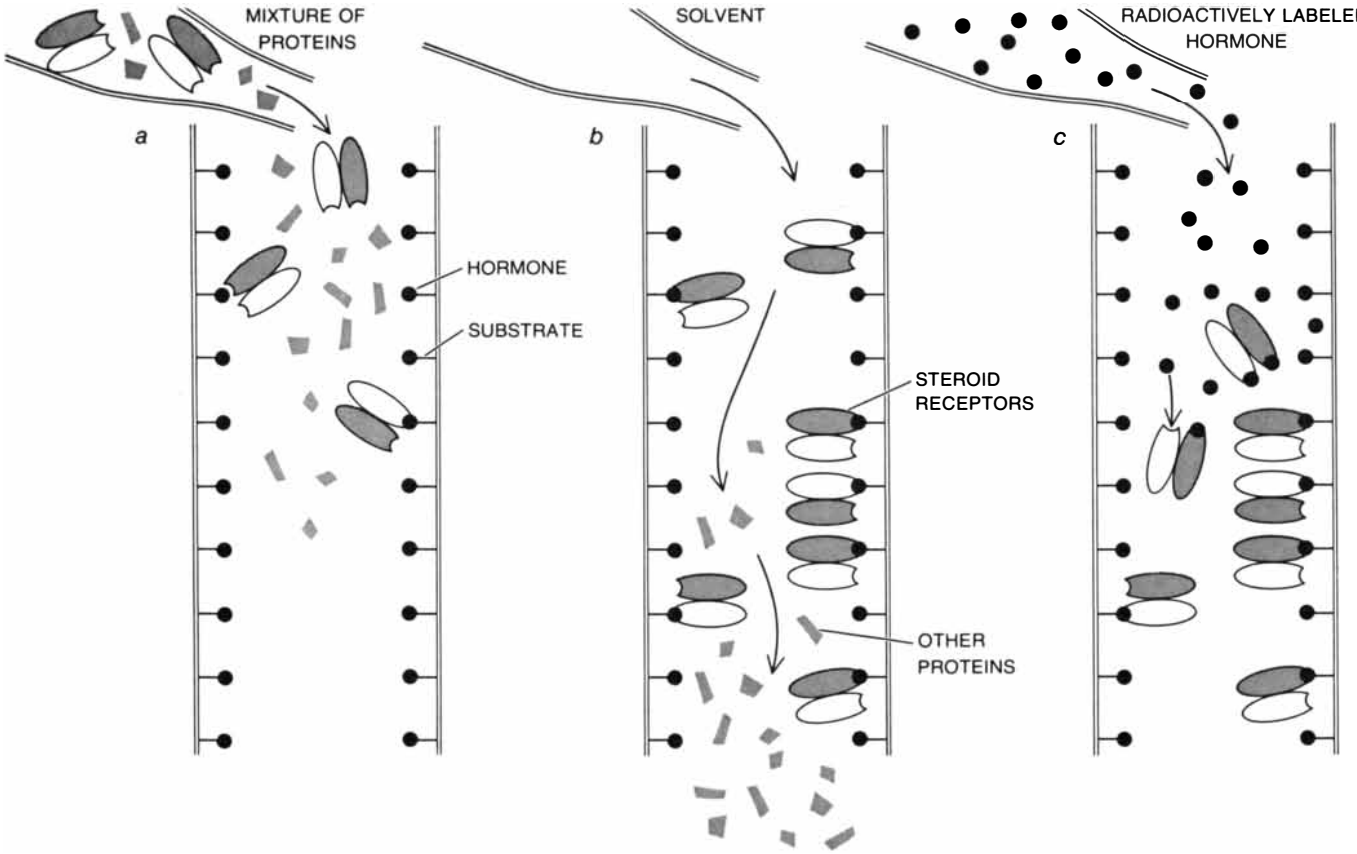


ACCEPTOR SITES for hormone-and-receptor complexes are observed only on the chromatin of target cells. Chromatin consists of DNA and two kinds of protein: the histones, which are much the same in all cells, and the nonhistone chromosomal proteins, which display much more diversity. The ability to bind the complexes resides in a particular fraction of the nonhistone proteins, designated the AP_3 fraction. Ordinarily the complexes bind to chromatin from target tissues, such as the oviduct, but do not bind to nontarget chromatin, such as that from chick erythrocytes. When the AP_3 fractions of the nonhistone proteins from each kind of chromatin are exchanged, the ability to bind hormone-and-receptor complexes migrates with the oviduct AP_3 fraction. The number of acceptor sites found on each kind of chromatin before and after the proteins are exchanged is indicated by the small bar graphs.

tant increase in their concentration in the nucleus. On the other hand, the receptor proteins could not be detected in the nucleus in significant quantity when the hormone was absent. Therefore the presence of the hormone somehow caused the receptors to be retained in the nucleus.

The mechanism by which the hormone-and-receptor complexes were retained in the nucleus was soon revealed. In a series of experiments in our laboratory and in the laboratory of Shutsung Liao at the University of Chicago the complexes were combined in a cell-free system with nuclei isolated from target cells. When the complexes were absorbed into the nuclei from the incubation medium, it was found that they bind directly to the chromosomes. An analysis of the binding reaction suggested that in the nucleus of each target cell there are about 5,000 "acceptor" sites for hormone-and-receptor complexes. Of particular importance was the observation that when nuclei from nontarget cells were employed, the binding of the complexes was greatly diminished.

Although it had been shown that the steroid hormones form complexes with specific receptor proteins in their target cells and that those complexes



AFFINITY CHROMATOGRAPHY, a technique employed in isolating the receptors of steroid hormones, relies on one of the biological properties of the receptors: their affinity for the hormones. Hormone molecules are bound to an inert substrate in a glass column; then an extract from the cytoplasm of target cells is poured into the column (*a*). The receptors are detained by the fixed hormone molecules, whereas all other proteins are washed away (*b*). Finally, the receptors are freed from the substrate by incubating

enter the nucleus and bind to the chromosomes, it remained to be disclosed how the binding of the hormone-and-receptor complexes influences the expression of genetic information. The nature of the interaction must obviously depend not only on the activity of the hormone but also on the architecture of the chromosome.

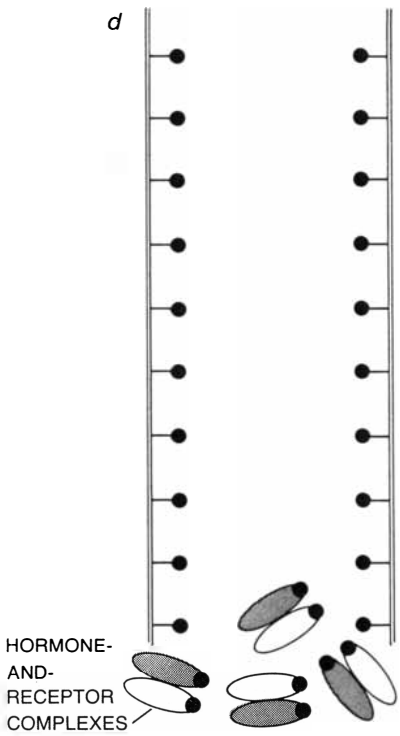
The chromosomes of higher organisms incorporate both DNA and a variety of specialized proteins, organized in the complex substance called chromatin [see "Chromosomal Proteins and Gene Regulation," by Gary S. Stein, Janet Swinehart Stein and Lewis J. Kleinsmith; SCIENTIFIC AMERICAN, February, 1975]. There are two main classes of chromosomal proteins. The histones are highly alkaline proteins and are very similar from one tissue to another; indeed, they are very similar from one organism to another. The remainder, the nonhistone chromosomal proteins, are acidic.

In a series of studies conducted by us in collaboration with Thomas C. Spelsberg, we found that hormone-and-receptor complexes were able to bind directly to chromatin isolated from the nucleus. Moreover, complexes of progesterone and their receptor proteins bound preferentially to chromatin from oviduct nuclei, compared with chromatin from other tissues. Receptor proteins without hormone were incapable of binding to chromatin.

We undertook to determine which elements of the chromatin participated in the binding of the complexes. Gordon M. Tomkins and his colleagues at the University of California at San Francisco had shown that hormone-and-receptor complexes will interact with DNA stripped of all its accompanying proteins. They found, however, that any DNA will suffice; it made no difference whether it was extracted from target cells or from other tissues. This was not a surprising result, since the DNA in all cells of an organism is identical. It seems unlikely for that reason that the DNA itself could regulate gene expression, and Tomkins' findings suggest in particular that the DNA cannot be directly responsible for the tissue specificity of the steroid hormones.

The binding sites were therefore sought among the chromosomal proteins. Because of their comparative uniformity the histones seemed to be unlikely candidates, and indeed Spelsberg found that chromatin from which all the histones had been removed bound the hormone-and-receptor complexes as effectively as intact chromatin. Furthermore, the histones could be removed from the chromatin of target cells and replaced with histones from other tissues or even from other organisms without affecting the binding of the complexes.

The nonhistones are much more heterogeneous than the histones; there may be more than 500 species of nonhistone proteins in a single nucleus, and they vary considerably from one tissue to another. By measuring the binding of receptor molecules to chromatin after removing selected groups of nonhistone proteins, Spelsberg revealed a selective affinity. The removal of two classes of nonhistones had no detectable effect, but when a third class, designated the AP_3 fraction, was extracted, the ability of the residual chromatin to bind hormone-and-receptor complexes declined sharply. When the AP_3 fraction was replaced, binding activity was restored. Finally, in the most conclusive experiment, the AP_3 proteins from target cells were ex-



them with a concentrated solution of radioactively labeled hormone (*c*). The result is a virtually uncontaminated solution of radioactive hormone-and-receptor complexes (*d*).

changed with the same fraction from a nontarget tissue, yielding hybrid chromatins. The chromatin from the nontarget tissue was converted by the addition of target-cell AP_3 fraction into an efficient acceptor for hormone-and-receptor complexes. The target-cell chromatin, on the other hand, lost its ability to bind to receptors when it was treated with the non-target-cell AP_3 fraction [*see illustration on page 37*]. The ability to bind the receptor complexes thus seems to depend on the presence of the AP_3 fraction on the DNA.

The progesterone receptor proteins employed in these experiments were merely crude extracts of hormone target cells from the chick oviduct; the receptor molecules constituted less than one ten-thousandth of the total protein. In order to characterize the receptor molecules more fully, and in order to ensure that the interactions we had observed were not modified by the presence of extraneous substances, we undertook to purify the receptor protein.

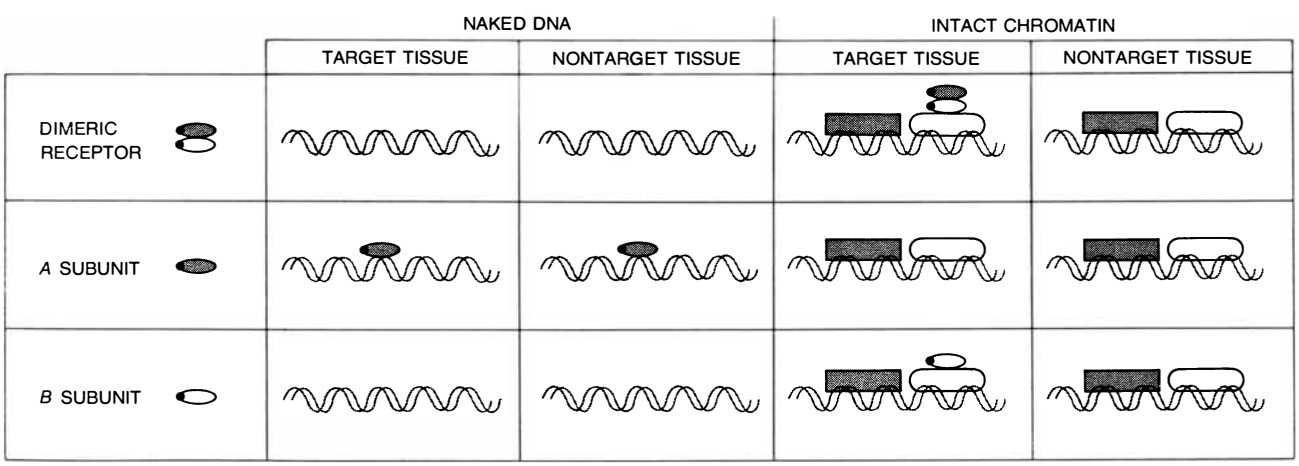
The principal method employed was affinity chromatography, a technique that isolates biological molecules by taking advantage of the specificity inherent in their own functioning. In this case progesterone molecules were chemically bonded to an inert, porous substrate, and the crude solution of proteins was then passed through the substrate. The receptor proteins bound to the hormone, whereas all the contaminants were readily washed away. Finally the receptors were freed from the bound hormone by incubating them with a comparatively concentrated solution of free progesterone. By that method the progesterone receptor was isolated from all other proteins; the purification was accomplished last year in our laboratory.

We found that the receptor molecule is a dimer, that is, it consists of two subunits (labeled *A* and *B*), each an independent strand of amino acids. The subunits are not identical, but both have a molecular weight somewhat greater than 100,000. The subunits are roughly cigar-shaped; when they are bound together as a dimer, they probably lie side by side.

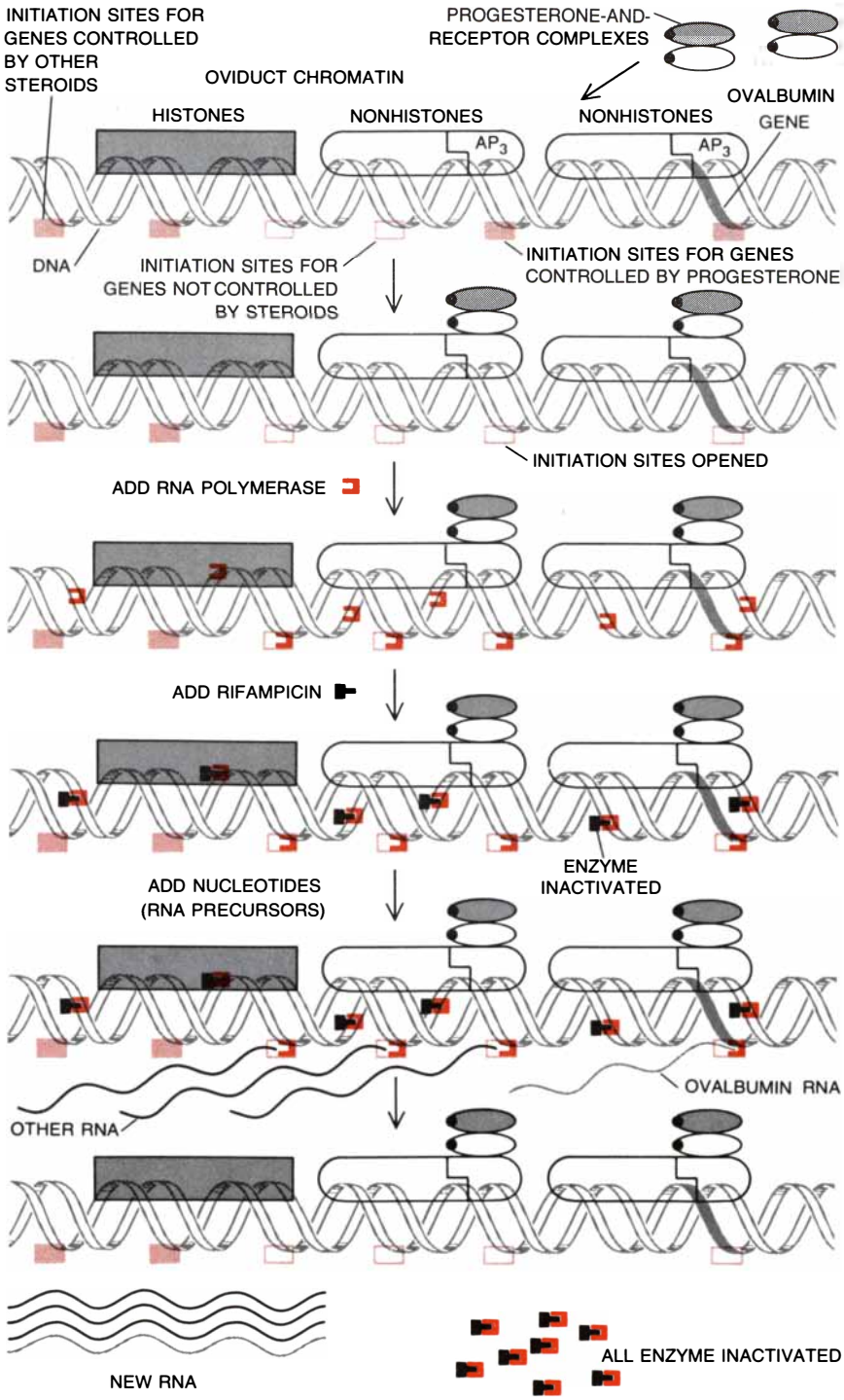
Each subunit has a single binding site for progesterone. No changes in the physical properties of the dissociated subunits have been detected when the hormone is bound to them, but the hormone has an important effect on the intact dimer: after the hormone has been bound, the receptor can be split into its two subunits by exposure to high temperature or to an elevated salt concentration. (In addition, of course, the hormone profoundly alters the biological activity of the molecule.)

When the purified receptors were incubated with oviduct nuclei or chromatin, they bound to acceptor sites on the chromatin just as the crude preparations had. It therefore appeared that no other cytoplasmic proteins were involved in the interaction. We were puzzled, however, to discover that the intact, dimeric receptors had no affinity for naked DNA, even though we had observed binding to DNA in crude preparations.

That paradox was resolved when we examined the behavior of the individual subunits. The *B* protein was able to bind to intact chromatin, and furthermore it showed the same tissue specificity for the AP_3 fraction of the nonhistone proteins. The *A* subunit did not bind to intact chromatin from oviduct tissues or from other tissues. On the other hand, the *A* subunit bound very strongly to naked DNA, whereas the *B* subunit showed no affinity for it. The binding to DNA observed in crude extracts was



RECEPTOR PROTEINS consist of two subunits, labeled *A* and *B*, with different capacities for binding to nuclear material. Although the subunits are not identical, they are about the same size, and each has a binding site for a single hormone molecule. The complete, dimeric receptor binds only to chromatin (made up of both DNA and the chromosomal proteins) from target cells. The *B* subunit displays the same specificity as the dimeric receptor. The *A* subunit does not bind to chromatin but does bind to "naked" DNA. Moreover, it has the same affinity for all DNA, whether it is derived from target cells or from cells of other tissues.



EFFECT OF STEROID HORMONES on chromatin is to increase the number of initiation sites for the transcription of particular genes. The effect was demonstrated with chromatin from the chick oviduct, which is stimulated by progesterone to manufacture RNA coding for numerous proteins, among them the egg-white protein ovalbumin. After the chromatin had been exposed to hormone-and-receptor complexes the enzyme RNA polymerase was added to the solution. A few enzyme molecules occupied stable initiation sites in those regions on the DNA where transcription normally begins, but transcription was prevented by conducting the experiment in a medium that lacked the nucleotides needed for RNA synthesis. Initiation sites for genes controlled by other steroids remained unoccupied. The antibiotic drug rifampicin was then added, inactivating all RNA-polymerase molecules except those at initiation sites. Finally the nucleotide precursors of RNA were added, and at each initiation site a single molecule of RNA was produced. The RNA-polymerase molecules responsible for this transcription, having left their initiation sites, were then also inactivated. Treating the chromatin with hormone-and-receptor complexes increased the production of RNA, and specifically stimulated synthesis of messenger RNA for ovalbumin.

presumably caused by contamination with dissociated *A* subunits.

From that analysis it is possible to construct a hypothetical account of how a steroid hormone is directed to a gene in a target-cell nucleus. Initially the receptor molecule is in the cytoplasm, and it is in dimeric form. After binding two molecules of hormone the structure of the receptor is altered in some unknown way, and it becomes activated. The complex of hormone and receptor then enters the nucleus and binds to the chromatin at an acceptor site defined by the position of some specific protein in the AP_3 fraction of the nonhistone proteins. The binding of the receptor to the acceptor site takes place exclusively through the *B* subunit, since the *A* subunit is incapable of binding to intact chromatin. The dimer may then dissociate, liberating the *A* subunit, the only one that can interact directly with DNA. It is not yet clear whether the *A* protein is capable of recognizing specific nucleotide sequences of DNA, as the regulatory proteins of bacterial cells are. It is possible that the *A* protein merely binds to any nearby sequence of DNA, its location having been determined with sufficient precision by the specificity of the *B* subunit for the acceptor site on the nonhistone proteins.

Perhaps the most intriguing question raised by the steroid hormones pertains to the events after the hormone-and-receptor complex has bound to the chromatin. How is the selected segment of DNA activated by the binding of the receptor? How is the gene "turned on"? In order to answer these questions we must investigate the events that come immediately after the binding of the receptor molecules.

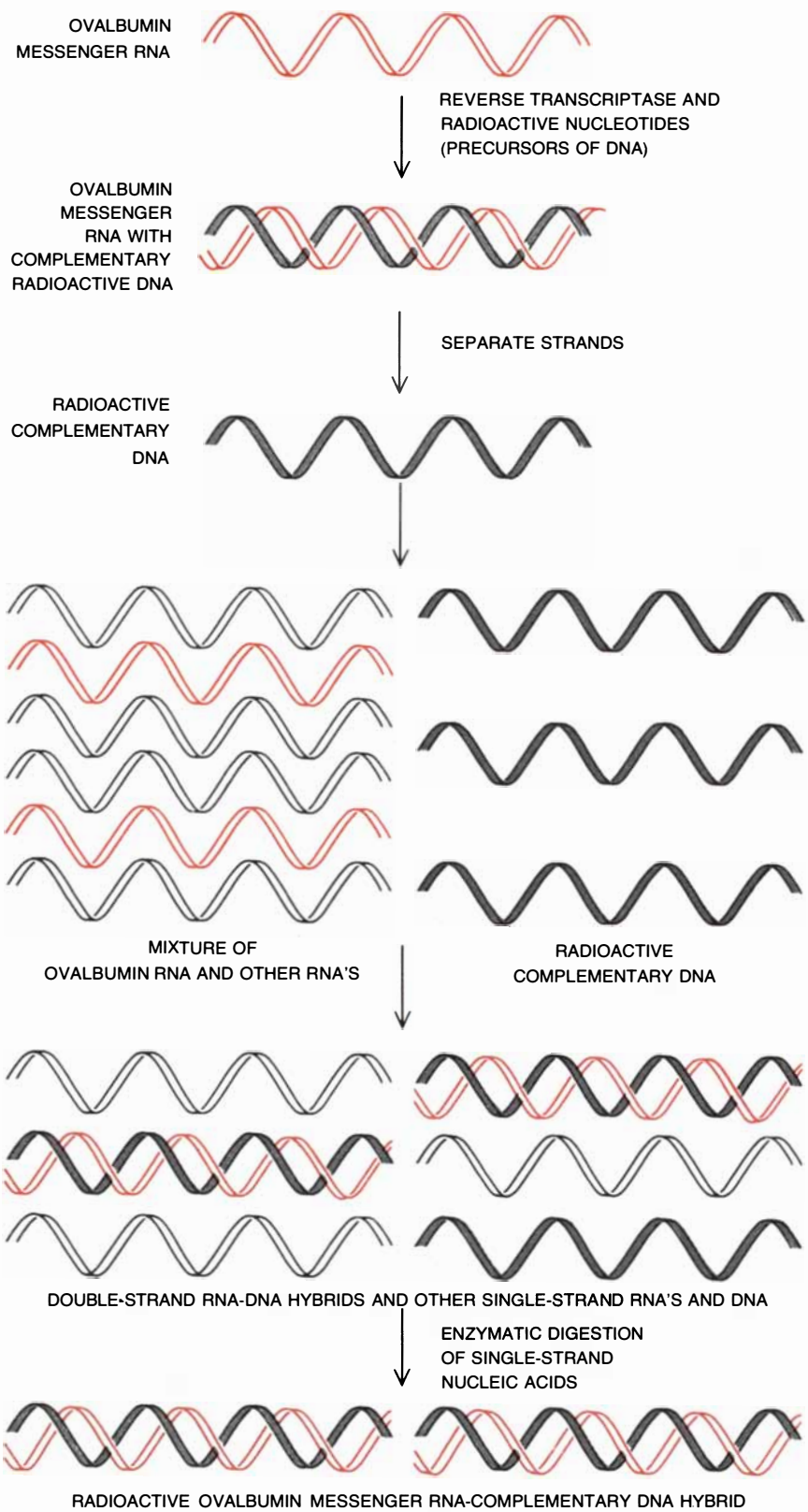
The first step in the sequence of events that leads eventually to protein synthesis is thought to be the binding of the enzyme RNA polymerase to the DNA. In bacteria the enzyme is bound at particular locations, called initiation sites, where the transcription of each gene begins. There is only one initiation site per gene, and hence by measuring the number of sites the number of active genes can be determined. In higher organisms the existence of initiation sites has not been proved, but the behavior of the gene is entirely consistent with the hypothesis that they do exist. We wished to monitor the number of initiation sites as an index of the number of genes expressed following hormone administration.

It was possible to estimate the number of initiation sites in oviduct chromatin by employing a system that produces

one molecule of messenger RNA at each initiation site and then ceases transcription. The system is based on the properties of the antibiotic drug rifampicin, which permanently inactivates molecules of RNA polymerase unless they are occupying an initiation site. In this procedure excess RNA polymerase was added to oviduct chromatin, leading to the binding of enzyme molecules at all available initiation sites; transcription of the genes could not begin, however, because the incubation medium lacked the precursor nucleotides needed for RNA synthesis. Those precursors were then added to the solution, and simultaneously rifampicin was added. As a result all the RNA-polymerase molecules not occupying an initiation site were inactivated; those enzyme molecules at initiation sites each gave rise to a single molecule of messenger RNA, then they too were inactivated, so that all RNA synthesis stopped [*see illustration on opposite page*]. The number of RNA molecules produced is thus a measure of the number of initiation sites on the chromatin.

In our experiment progesterone or estrogen was administered to chicks for varying lengths of time, then the chromatin was isolated from oviduct cells. As in earlier experiments, a radioactively labeled hormone was employed, so that the number of hormone-and-receptor complexes in the cell nuclei could be estimated. The oviduct chromatin was then assayed for initiation sites by the addition of rifampicin. The results strongly supported the role of hormone-and-receptor complexes in gene regulation: the number of nuclear initiation sites per cell rose and fell in conjunction with the number of hormone-and-receptor complexes in the nuclei.

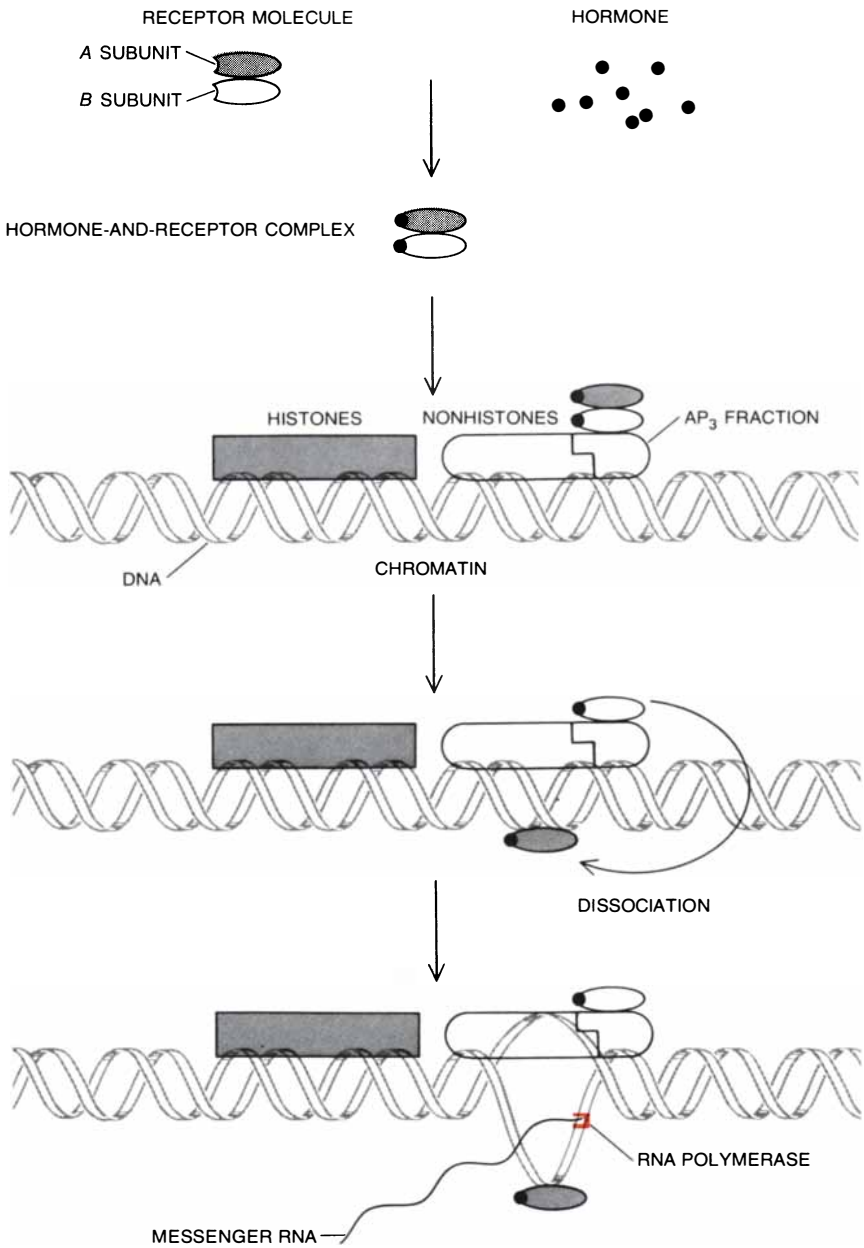
We have now employed the same techniques to demonstrate the hormonal regulation of transcription in an experiment performed entirely outside the living cell. Instead of injecting chicks with progesterone, we purified chromatin from the oviduct of chicks that had never been exposed to the hormone. We then added to the chromatin purified progesterone receptor molecules with bound hormone. RNA synthesis was initiated in the presence of rifampicin, and the number of RNA molecules produced was determined. Once again the number of initiation sites was dependent on the number of hormone-and-receptor complexes present. In this cell-free experiment the number of initiation sites produced was approximately equal to the number of hormone-and-receptor complexes bound to the chromatin. From this fact we concluded that the initiation



IDENTIFICATION OF RNA coding for ovalbumin was accomplished through the technique called RNA-DNA hybridization. The procedure begins with the preparation of a pure specimen of the messenger RNA for ovalbumin. The enzyme reverse transcriptase is then employed to synthesize a strand of radioactive DNA with a sequence of nucleotides exactly complementary to that of the RNA. The complementary DNA is incubated with a mixture of RNA's, and because of its complementary sequence of nucleotides it forms double-strand molecules with ovalbumin messenger RNA but not with any other RNA. All single-strand nucleic acids are then digested, leaving only the double-strand RNA-DNA hybrids. The amount of ovalbumin messenger RNA is calculated from the rate and extent of hybridization.

reaction involves one hormone-and-receptor complex per chromatin acceptor site. The stimulation of transcription is dependent on the presence of the hormone; receptor proteins that lacked hormone had no effect. Furthermore, the extent of stimulation depended on the kind of tissue tested; target-cell chromatin was stimulated more than chromatin from other tissues.

The foregoing experiments demonstrated that hormone-and-receptor complexes can stimulate RNA synthesis. It remained to be shown that the particular genes transcribed were those coding for proteins known to be produced in response to the hormone. Making that determination requires an extraordinarily sensitive method for the detection and identification of specific messenger-RNA molecules. Such a method exists in the procedure called RNA-DNA hybridization. The procedure begins with the preparation of an exceedingly pure sample of the messenger RNA to be detected. We chose to employ the messenger RNA that codes for ovalbumin, an egg-white protein that is completely absent from chicks that have not been treated with estrogen or progesterone but that is produced copiously in the oviduct of chicks that have received either of these hormones. From the purified messenger RNA radioactive single-strand DNA is synthesized with the exactly complementary sequence of nucleotide bases. The synthesis is accomplished with enzymes of viral origin that reverse the normal flow of information in the cell, so that RNA becomes a template for the manufacture of DNA. The complementary strands of DNA are then isolated, and they can be employed as a molecular probe capable of recognizing any molecules identical with the original messenger-RNA template.



MECHANISM OF GENE ACTIVATION by the steroid hormones is thought to involve the separate interactions of both subunits of the hormone receptor molecule with chromatin. After a receptor molecule has bound two molecules of hormone and has entered the nucleus, it binds to the AP_3 fraction of the nonhistone chromosomal proteins. The binding takes place between the nonhistone protein and the *B* subunit of the receptor molecule, which thereby determines what genes will be activated. The subunits then dissociate, and the *A* subunit interacts with the DNA, enabling a molecule of RNA polymerase to occupy an initiation site or the DNA. A segment of DNA is then transcribed, producing a strand of messenger RNA that serves as a template for the construction of a protein.

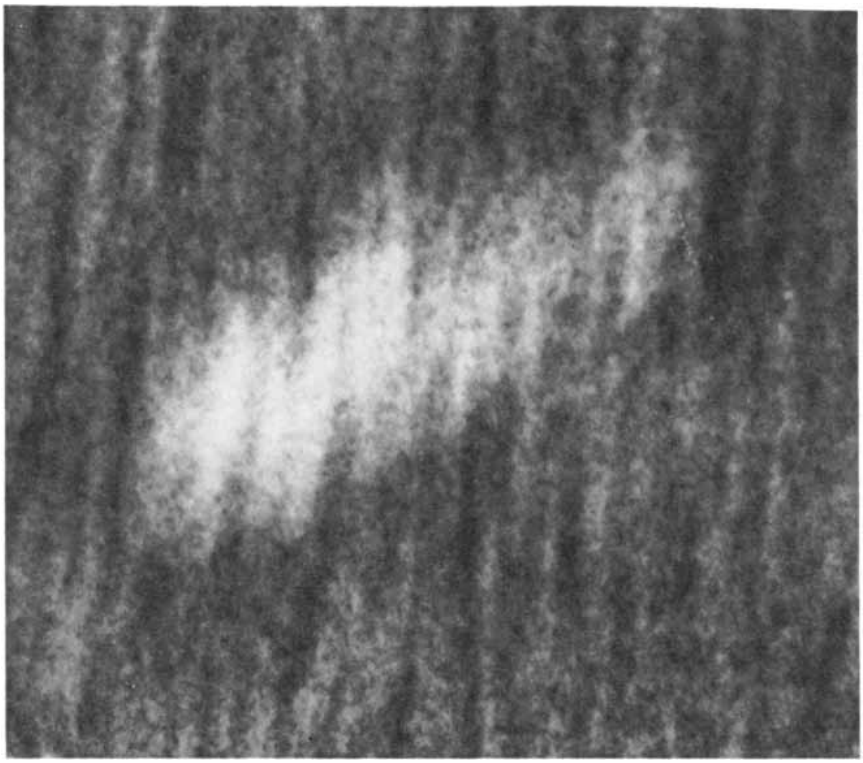
The RNA-DNA hybridization technique was first employed to assay the messenger RNA's produced in the chick oviduct in response to hormone. At various intervals after the administration of progesterone the messenger RNA was isolated from oviducts and incubated with single-strand DNA complementary to the ovalbumin messenger RNA. Because of the complementary sequences all the ovalbumin RNA formed double-strand hybrid molecules with the DNA; the remaining, extraneous RNA's remained as unpaired, single-strand molecules and could therefore be distinguished chemically and eliminated from the preparation. From the number of RNA-DNA hybrid molecules we were able to estimate the number of ovalbumin messenger-RNA molecules present in oviduct cells after hormone injection. The cellular content rose from zero to 10,000 molecules per cell in a matter of hours.

Finally, oviduct chromatin was incubated in a cell-free medium with purified progesterone hormone-and-receptor complexes. RNA was synthesized from these chromatin templates and assayed for ovalbumin messenger RNA by RNA-DNA hybridization. Only the RNA synthesized from chromatin exposed to receptors contained ovalbumin messenger RNA. Thus we concluded that hormone-and-receptor complexes have the capacity to bind to chromatin and to increase both overall gene transcription and the expression of DNA sequences coding for specific messenger-RNA molecules.

At physiological concentrations only the intact, dimeric receptors caused the proliferation of initiation sites and stim-

ulated the synthesis of ovalbumin messenger RNA. The isolated *B* subunit was without effect regardless of the concentration. The *A* subunit did eventually increase the number of initiation sites, but only at concentrations about 50 times higher than the effective concentration of the dimer. That discovery supports the hypothesis that only the *A* protein interacts directly with DNA and stimulates the synthesis of messenger RNA. The *B* subunit apparently serves merely to designate the location of the gene to be expressed, a function accomplished through the binding of the *B* subunit at a position defined by certain of the nonhistone proteins. It is reasonable to suppose the genetic sequences controlled are adjacent to the corresponding acceptor sites on the chromosomal proteins.

The nature of the interaction between the *A* subunit and the DNA remains a topic of active investigation. It is still not known how the hormone-and-receptor complex alters a concealed or inactive initiation site to make it available to RNA polymerase. It is also not yet known how the regulated gene is turned off when enough RNA has been synthesized. The hormone and the receptor protein must eventually dissociate from each other and from the chromatin. The receptor molecule is apparently not destroyed in the process, and it may return to the cytoplasm, form another complex with additional hormone molecules and reenter the nucleus to continue hormonal stimulation. The hormone itself is probably altered in such a way that it becomes inactive, and it must eventually diffuse out of the cell.



SHAPE OF THE RECEPTOR PROTEIN is revealed in an electron micrograph made by P. M. Conn of the Baylor College of Medicine. The protein is the light, oblong form in the center of the image; it is a *B* subunit of a progesterone receptor, with a hormone molecule bound to it. It is shown in its biologically active configuration; when the subunit is denatured, it assumes a globular form. The length of the protein is about 100 angstroms, and the magnification is roughly 700,000 diameters. The image is indistinct primarily because the specimen is destroyed by the intense electron beam as the micrograph is being made.

All the effects of the steroid hormones can now be seen to be accomplished through specific protein receptor molecules; the absence of receptors precludes any response. It seems likely that this relation can be made quantitative: the magnitude of a cell's response appears to be related to the intracellular concentration of receptor proteins. Moreover, the concentration of receptors in a given tissue is not fixed: it can be altered by aging, by changes in the physiological state of development and even by the presence of other hormones. In the chick oviduct estrogen stimulates the production of progesterone receptors, and tissues treated with estrogen are therefore more responsive to progesterone. Conversely, in mammals progesterone depresses the level of estrogen receptors and makes the target tissue less responsive to estrogen. The stability of the hormone-and-receptor complex also influences the magnitude of a target cell's response to a steroid hormone. Cells containing complexes that dissociate rapidly must be bathed continuously in a high concentration of hormone if the response is to be maintained.

The illumination of the molecular biology of the steroid hormones must ultimately reflect some light on clinical endocrinology. For example, the discovery of hormone receptors provides an explanation of an enigmatic syndrome called male pseudohermaphroditism. In that condition the male secondary sex characteristics fail to develop. The cause is not simply a deficiency of androgens, since the target tissues do not respond when additional androgens are administered. It results instead from a defect in the androgen receptor proteins. An entire series of steroid-unresponsiveness syndromes may be discovered, all of them caused by the absence of functioning hormone receptors.

The techniques devised for the investigation of the action of the steroids might also be applied to the identification of certain hormone-dependent cancers. A large proportion of breast cancers, for example, grow in response to estrogens circulating in the bloodstream, and they can be treated by the surgical removal of the estrogen-secreting organs (the ovaries and the adrenal glands). The other breast tumors are unaffected by estrogens. Jensen and his colleagues at the University of Chicago and William L. McGuire and his colleagues at the University of Texas Health Science Center at San Antonio have recently shown that only the hormone-dependent tumors have receptors for estrogen. The distinction can now be employed as a diagnostic tool. Similar methods might be applied to the classification of prostate tumors whose growth is dependent on androgens and to lymphosarcomas dependent on corticosteroids.

A more precise understanding of how steroids bind to their receptor proteins could uncover methods for the sophisticated manipulation of the hormone response. For example, it should be possible to synthesize biologically inactive analogues of the natural hormones that would block the receptor binding sites and thereby suppress the hormonal response. Such "antihormones" could have therapeutic value, and perhaps even social consequences, since they might be safe and efficient contraceptives.

Is Gravity Getting Weaker?

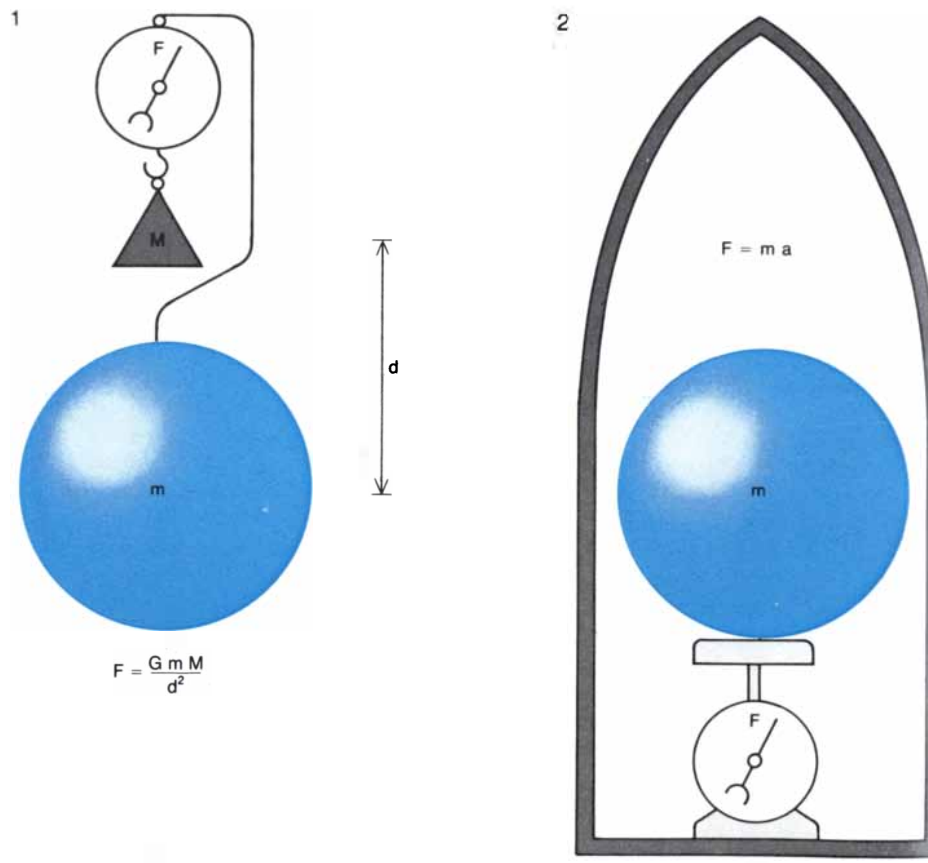
Several theories of gravitation predict that the force of gravity diminishes as the universe expands. Preliminary results of timing eclipses of stars by the moon suggest that it may well be the case

by Thomas C. Van Flandern

According to Newton's law of gravity, every particle of matter in the universe attracts every other particle by a force that is directly proportional to the mass of the particles and inversely proportional to the square of the distance between them. A key term in the equation is the universal gravitational constant G, which defines the exact value of that proportionality. It is generally taken for granted that G is indeed both universal and constant. The universe, however, is expanding. Apart from local irregularities all its constituent bodies—planets, stars, galaxies—are moving away from one another. Is the value of G truly constant and immune to the expansion? Or is the force of gravity between bodies diminishing with the expansion? Some theories of gravity predict that the value of G is decreasing. There now seems to be some observational evidence supporting those predictions. If such evidence is confirmed, the implications for physics will be profound.

Newton's laws assert that there are two relations between a force and a mass. Newton's second law of motion states that the force F exerted on an accelerating body of mass m is equal to the product of the mass times the acceleration a. The equation is written $F = ma$. Newton's law of gravity states that the mass m also has a gravitational field that will attract any other mass M with a force that is the product of the two masses multiplied by some constant—the gravitational constant G—divided by the square of the distance d between the masses. The equation is written $F = GmM/d^2$. One of the assumptions underlying these laws is that the body's inertial mass m in the first equation is the same as its gravitational mass m in the second equation, that is, that the body's resistance to acceleration is directly proportional to its gravitational attraction. That assumption has been verified to high accuracy for small masses but not for large ones.

Two and a half centuries after Newton, Einstein presented his initial description of the nature of space and time: the special theory of relativity. On the basis of the special theory he went on to develop his general theory of relativity, which shows how the behavior of gravity deviates from the behavior predicted by Newton's laws. The deviations are of several types. The first-order predictions of the general theory include three classic tests of the theory. First, light rays passing close to the sun will be bent by the sun's gravitational field twice as much as the amount predicted on the basis of Newtonian theory. Second, the



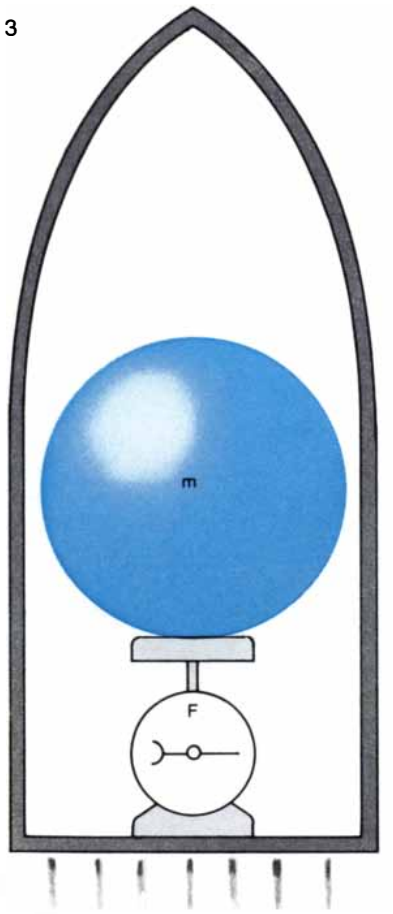
STRENGTH OF A BODY'S GRAVITATIONAL FIELD is assumed to be directly proportional to the body's resistance to acceleration, according to Newton's laws of gravity and of motion. Newton's law of gravity states that any mass m has a gravitational field that attracts any other mass M (*1*) with a force F that is equal to the product of the masses multiplied by

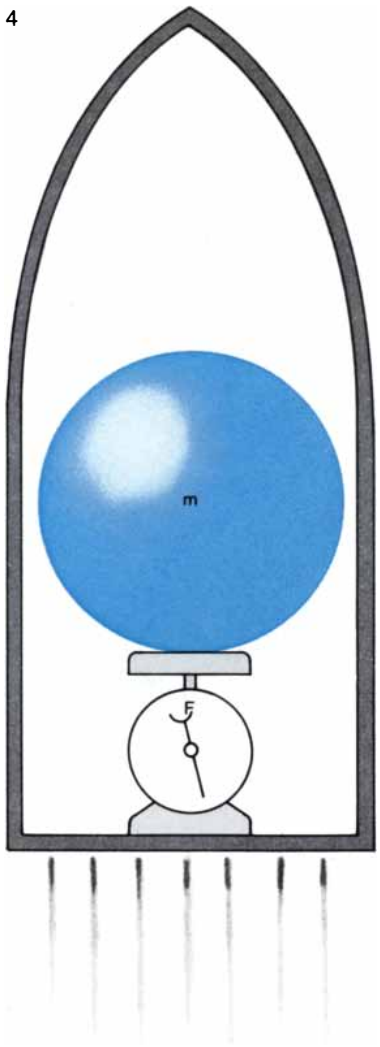
perihelion, or point closest to the sun, of the elliptical orbit of the planet Mercury will advance in the direction of the planet's rotation by 43 seconds of arc per century more than can be accounted for by gravitational perturbations caused by the other planets. Third, clocks will run slower in a stronger gravitational field than they will in a weaker one. The last statement is equivalent to the statement that light "climbing out" of a gravitational field will lose energy, so that its wavelength will be shifted toward the red end of the spectrum. All three of the first-order predictions have been confirmed to an accuracy of 95 percent.

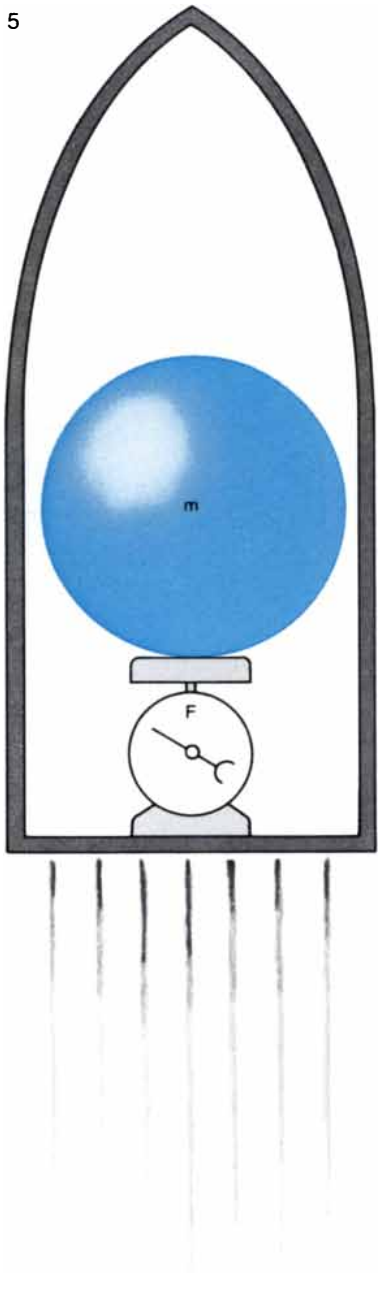
One of the higher-order predictions of the general theory of relativity is that there should be gravitational waves: the gravitational analogue of electromagnetic waves. The theory also predicts that there should be "singularities" in the structure of space-time that lead to the existence of black holes: objects that have collapsed until they are so dense and their gravitational field is so strong that not even light can escape from them. And it predicts that G should be constant.

Although experiments have been devised to test all the higher-order predictions, none of the predictions has yet been either confirmed or ruled out. The sensitivity of the apparatus required for the tests has to be so great that the results so far are at best inconclusive. Although it is tempting to assume that the past successes of the general theory give it a high probability of future successes, it should be remembered that there are newer theories competing with it that make the same predictions for the first-order tests but differ markedly in their predictions for the higher-order tests. Among those theories is one developed by P. A. M. Dirac and another put forward by Fred Hoyle and J. V. Narlikar.

Dirac's theory, which he formulated in 1937, is known as the large-numbers hypothesis. It was the first serious quantitative theory to predict that the strength of gravity decreases with time. Dirac's reasoning was entirely philosophical. Suppose a set of fundamental units for mass, length and time are agreed on. All physical constants can then be expressed in terms of those fundamental units, making it possible to compare constants of different dimensions. Dirac chose the mass of the electron to be the unit of mass, the radius of the electron's orbit in the hydrogen atom to be the unit of length and the time required for a photon traveling at







the gravitational constant G and divided by the square of the distance d between the masses. The formula is $F = GmM / d^2$. Newton's law of motion states that the force F exerted on the body m (*2 through 5*) when it is accelerated is equal to the product of the body's mass times its acceleration a ($F = ma$). If the value of G is decreasing, however, a body's gravitational mass will also be decreasing, that is, its gravitational attraction will weaken, yet its inertial mass (resistance to acceleration) should remain unchanged.

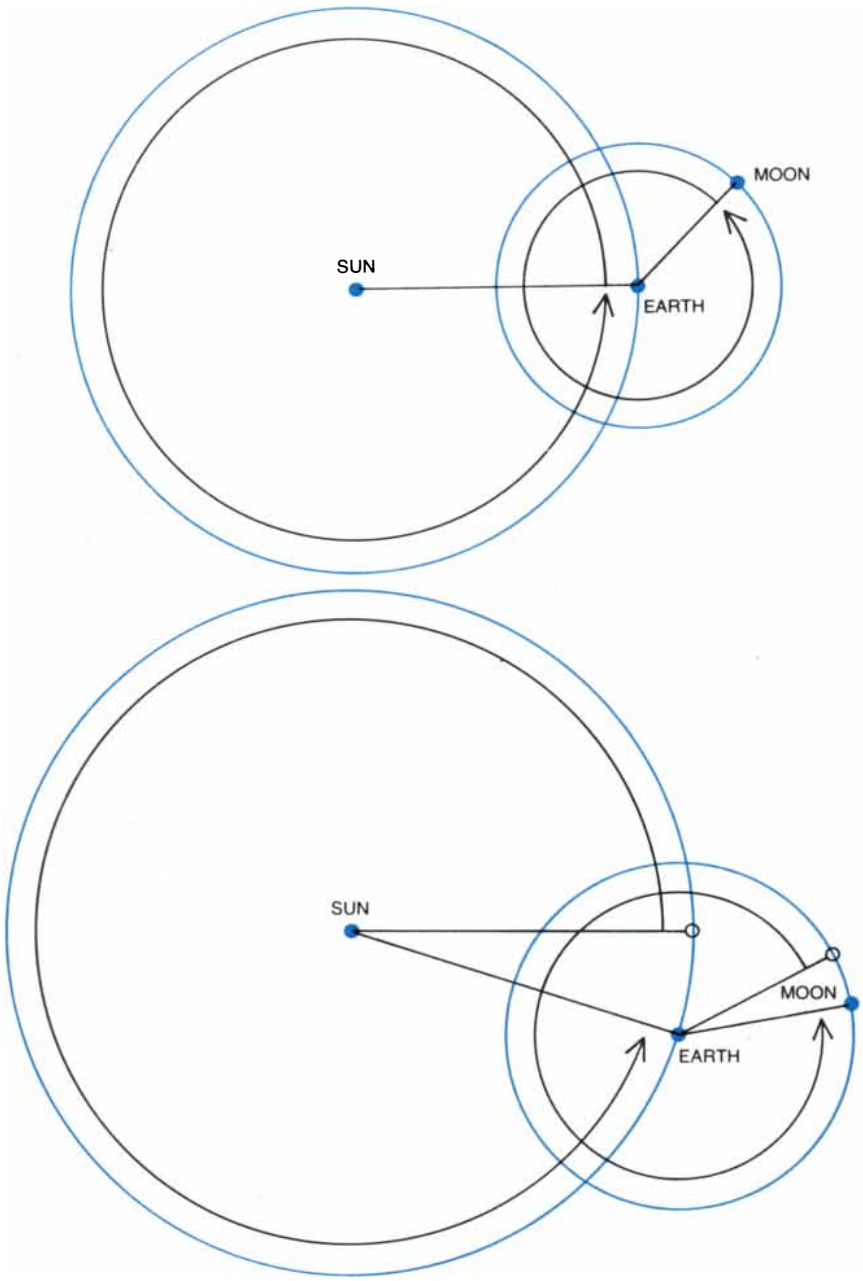
the speed of light to traverse that radius to be the unit of time.

When the basic constants of physics are expressed in the terms of those fundamental units, they turn out to be dimensionless numbers, or ratios. Many are near the value of 1 (10^0) to within a few orders of magnitude. For example, the mass of the proton is 1,836 (about 10^3), the speed of light is 1 and the fine-structure constant of the hydrogen atom is 1/137 (about 10^{-2}). By the same token the electromagnetic force between two charged particles and the strong force that binds protons and neutrons together in the nucleus of the atom are both approximately equal to 1.

Some constants, however, have values that, when expressed in terms of the fundamental units, diverge radically from the value of 1. The coupling constant for the weak interaction between nuclear particles is about 10^{-20}. The gravitational constant G is about 10^{-40}. And the age of the universe is almost 10^{40}.

Thus the value of G in terms of the fundamental units is the reciprocal of the age of the universe. Can it be a coincidence? Or does it reflect some undiscovered relation between the gravitational constant and the age of the universe?

Dirac preferred the latter conclusion, and he pointed out that such a relation implies one of two alternatives. The first is that since the age of the universe changes (increases) with time, the value of G must therefore also change (decrease). The second is that the fundamental units themselves could be changing with time.



ORBITS OF CELESTIAL BODIES WILL SLOWLY EXPAND if the gravitational constant G is decreasing. Therefore the time it takes for the body to complete one revolution in its orbit will lengthen. All orbits, however, will be affected in the same proportion to one another. Thus if one uses the period of the earth's revolution around the sun as a standard of time (ephemeris time) to measure the period of the moon's revolution around the earth, one will not be able to detect any change in the moon's orbital period caused by a decrease in the value of G. If one uses an atomic clock, however, which is a standard of time that is completely independent of the motions of celestial bodies, one could in principle detect a change in the moon's orbital period caused by a decrease in the gravitational constant G.

There is a second philosophical principle that leads to the expectation that the force of gravity is weakening. It is Mach's principle of inertia, conceived by Ernst Mach late in the 19th century. Mach argued on philosophical grounds that the properties of space depend entirely on the matter within that space. One way to visualize Mach's principle is to imagine a universe that is empty except for the earth and a pendulum suspended above the earth's North Pole. Would the pendulum swing in a plane that is fixed in space as the earth rotated under it, as it does in the real world? Or do rotation and direction lose all meaning in an empty universe?

Those questions are equivalent to asking whether or not space has properties independent of the matter within it. Mach argued, again on philosophical grounds, that rotation and centrifugal force could not exist in an empty universe, since rotation is meaningless without any points of reference. If matter is slowly introduced into the distant parts of the empty universe, only then will space begin to take on its familiar inertial properties.

Several theories of gravitation that incorporate Mach's principle predict that as matter recedes because of the expansion of the universe its local influence diminishes, and therefore the gravitational constant G decreases. The exact rate at which G changes depends on the individual theory. All viable theories, however, place the expected rate of change close to the observed rate of the expansion of the universe: $5.6 \pm .7$ parts in 10^{11} per year.

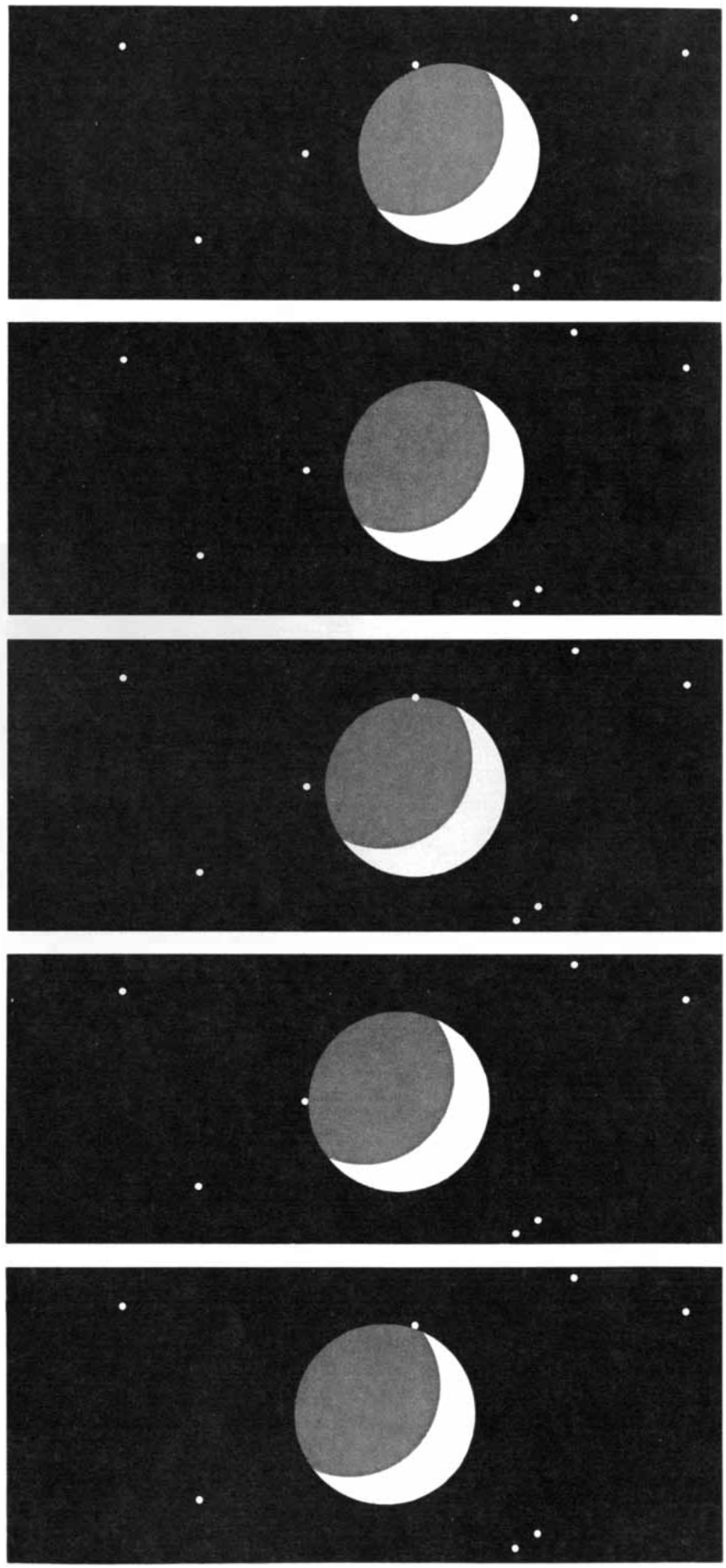
Until recently the measurement of such a small rate of change has not been possible. Now there are several experiments in progress that have the required

precision. One of them, which is based on occultations, or eclipses, of stars by the moon, has already yielded preliminary results.

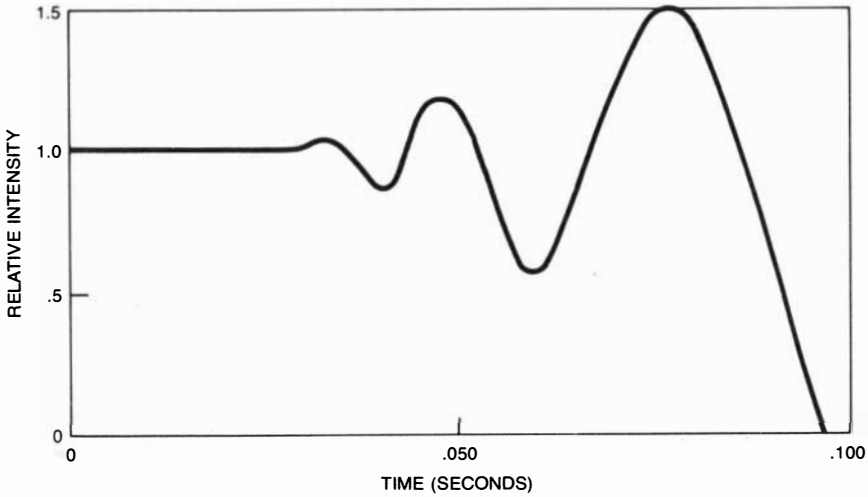
The principle behind the lunar-occultation method is straightforward. During the moon's monthly revolution around the earth it passes in front of numerous stars. Since the moon has no atmosphere, and since the stars are so far away that they are essentially point sources of light, the disappearance of a star at the moon's leading edge and the reappearance of the star at the trailing edge will appear to the eye to be instantaneous. The suddenness of the disappearance and the reappearance can be startling to the inexperienced observer, and it is part of what makes the timing of lunar occultations interesting to amateur astronomers. The occultations that are observed for the purpose of detecting a change in the gravitational constant, however, are monitored with a photoelectric photometer that has a high resolution with respect to time. The photometer yields a detailed light curve of the occultation that is fairly complex [*see top illustration on next page*]; the starlight nonetheless drops from full intensity to near zero within a few thousandths of a second.

That type of occultation is a total occultation, and it is quite common. There is a somewhat rarer type that is known as a grazing occultation, where the path of the star is tangent to the northern or southern limb of the moon. Since the moon's limb is a somewhat rough profile of lunar mountains and craters, it is possible for the star to be occulted many times within a few minutes. The star will cause a typical mountain a mile high on the moon to cast a shadow a mile high on the earth, and the observer must be located within that shadow in order to see the event. For that reason it is usually necessary for observers to travel in groups to the sites where grazing occultations will be visible.

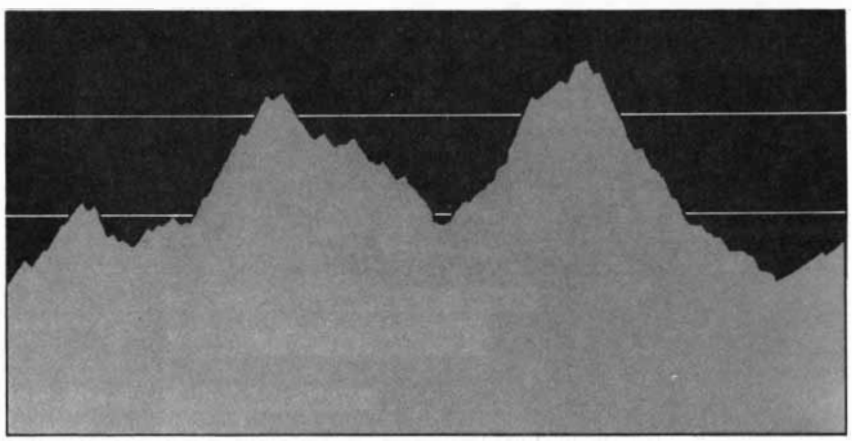
If the strength of gravity is weakening,



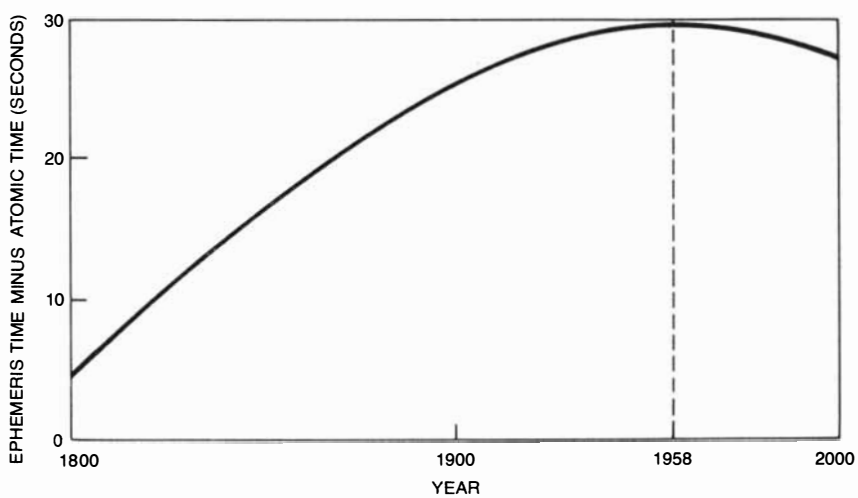
STARS ARE OCCULTED, or eclipsed, when the moon passes in front of them. The moon moves from west to east (*right to left*) in its orbit. The northern limb (*top edge*) of the moon is passing just tangentially to a star, which blinks off and on behind the lunar mountains and valleys in a grazing occultation. Another star is totally occulted at the advancing eastern limb (*left*). If the moon's orbit is expanding and its period of revolution is increasing because of a decrease in *G*, then on each revolution it will occult stars progressively later than predicted, assuming that the occultations are observed with respect to atomic time.



LIGHT OF A STAR DROPS from full intensity to zero in a few thousandths of a second when the star is occulted. When occultation is observed with a photoelectric photometer, starlight is observed to fluctuate in intensity first because of diffraction at the moon's limb.



PROFILE OF THE MOON'S LIMB is jagged with mountains and valleys. Moon is close enough to the earth for its parallax to be large, and observers one mile apart see quite different things during a grazing occultation. White lines represent apparent path of star with respect to moon's limb to two such observers. If observers are stationed 100 feet apart on earth, their observations can establish position of moon to an accuracy of .02 second of arc.



CURVE OF DIFFERENCE between ephemeris time and atomic time accountable to changes in the value of *G* is shown from the data obtained from observations of lunar occultations. Curve is a parabola because the difference is proportional to square of elapsed time.

the moon is very slowly receding from the earth. The moon's orbit will therefore become larger and its orbital period, the time it takes to complete one revolution around the earth, will become longer. On the basis of a theory of the moon's motions that does not incorporate any change in the gravitational constant it can be predicted quite precisely when the moon should occult a particular star. If the gravitational constant is decreasing, it will have the effect of making the actual time the occultation is observed later than the time that was predicted on the basis of the theory.

The amount by which the period is predicted to increase depends on which theory is used to make the prediction. If one accepts the "primitive" theory that *G* is the only fundamental constant that changes, then the proportional rate at which the moon's orbital period changes is equal to minus twice the proportional rate of change of the gravitational constant. If, on the other hand, it is Dirac's fundamental units of mass, length and time that are changing, the proportional rate at which the orbital period changes is equal to only minus the rate of change of the gravitational constant.

If one has a clock that runs at a uniform rate, one can in principle measure the length of the moon's orbital period to determine whether or not it is changing. If the rate of the clock is also affected by a change in the gravitational constant in the same way that the measurement of the moon's orbital period is, however, then one will always find that the orbital period is constant regardless of the actual situation. That was exactly the case until recently, because the revolution of the earth around the sun was used to determine ephemeris time, or astronomical time. If the strength of gravity is weakening, the earth's orbit around the sun will be getting larger just as the moon's orbit around the earth is; therefore the time it takes for the earth to complete one revolution around the sun will also increase, by an amount that is exactly proportional to the increase in the moon's orbital period. Since the relation between the two periods will always be the same, observing occultations by ephemeris time alone will not reveal any change in *G*.

Actually the moon's orbital period does change with respect to ephemeris time, mainly because of the friction of earth tides. The angular momentum lost by the rotating earth as it is slowed by the tides is being transferred to the moon by way of the gravitational coupling between the earth and the moon, causing

QUANTITY AFFECTED	PRIMITIVE THEORY	DIRAC'S LARGE-NUMBERS HYPOTHESIS	
		ADDITIVE CREATION	MULTIPLICATIVE CREATION
G	DECREASES $\propto 1/t$	DECREASES $\propto 1/t$	DECREASES $\propto 1/t$
PERIOD OF A BODY IN AN ORBIT	INCREASES $\propto -2\frac{\dot{G}}{G}$	DECREASES $\propto +\frac{\dot{G}}{G}$	INCREASES $\propto -\frac{\dot{G}}{G}$
MASS OF A BODY	CONSTANT	CONSTANT	INCREASES $\propto t^2$
VELOCITY OF A BODY	CONSTANT	CONSTANT	CONSTANT
RADIUS OF AN ORBIT	INCREASES $\propto t$	DECREASES $\propto 1/t$	INCREASES $\propto t$
RADIUS OF A SOLID BODY	SLOW INCREASE (DEPENDS ON DENSITY)	CONSTANT	INCREASES $\propto t^{2/3}$ (IF DENSITY IS CONSTANT)
RATIO OF ANY ORBITAL RADIUS TO RADIUS OF A SOLID BODY	NET INCREASE	DECREASES $\propto 1/t$	INCREASES $\propto t^{1/3}$
ANGULAR MOMENTUM OF A ROTATING SOLID BODY	SLOW INCREASE	CONSTANT	INCREASES $\propto t^{8/3}$
ANGULAR MOMENTUM OF A BODY IN AN ORBIT	INCREASES $\propto t$	DECREASES $\propto 1/t$	INCREASES $\propto t^3$

PHYSICAL QUANTITIES WILL BE AFFECTED if the value of G is decreasing. The amount by which they are affected depends on which theory holds true. The "primitive" theory assumes that only the value of G will change. P. A. M. Dirac's large-numbers hypothesis predicts that fundamental units of mass, length and time will also change. The letter t indicates the passage of time. The symbol $\propto$ means "in proportion to." The quantity $\dot{G}/G$ indicates the proportional rate of change with time of the gravitational constant G.

the moon to recede from the earth and its orbital period to increase. The most reliable estimate of the proportional amount by which the moon's period is being increased by tidal friction is 15.0 $\pm$ 1.2 parts in 10^{11} per year.

Since 1955, however, a time scale independent of the motions of celestial bodies has been available. It is atomic time, based on the exquisitely regular vibrations of cesium atoms. Observations of the moon's motion with respect to the time measured by atomic clocks should consequently reveal any change in the moon's orbital period caused by a change in the gravitational constant or in Dirac's fundamental atomic units.

Since the positions of the stars and the positions of occultation observers on the earth are usually well known, each timing of an occultation can be reduced to a precise statement about the celestial position of the moon with respect to the center of the earth. The total occultations yield primarily the moon's celestial longitude at any particular time, and the grazing occultations help in determining its celestial latitude. After 20 years of comparing such positions on the basis of atomic time, it has now become possible to determine the change in the moon's orbital period with sufficient accuracy to get a preliminary indication of whether or not the gravitational constant is changing, and if it is, by how much.

The occultations yield the finding that the moon's period is increasing, as a result of all causes, at the rate of 22.2 $\pm$ 3.5 parts in 10^{11} per year. When the rate at which the period is increasing because of tidal friction is subtracted, the difference provides an estimate of the rate at which the gravitational constant is decreasing. Under the "primitive" assumption that only G is changing, the rate at which it is decreasing is 3.6 $\pm$ 1.8 parts in 10^{11} per year. If Dirac's fundamental units are changing, however, that same difference implies that the gravitational constant is decreasing at the rate of 7.2 $\pm$ 3.7 parts in 10^{11} per year.

The agreement of both values with the rate of the expansion of the universe is quite good, although it is obviously aided by the large margins of error. For the present, then, the results of the lunar-occultation observations favor the prediction that the gravitational constant G is decreasing.

It is important to be certain, however, that the observed increases in the moon's orbital period do not have some other cause. Many phenomena that could give rise to such changes have been considered, including the effects of undiscovered planets in the solar system, the "wind" of electrically charged particles streaming from the sun, the radiation pressure of sunlight, the impact of meteorites, the gravitational effects of passing comets and asteroids, electromagnetic forces and a drag force exerted by the interplanetary medium. Making a plausible assumption for the magnitude of each of these effects, the upper limit of what each could exert on the orbital period of the moon is too small to account for the observed change by several orders of magnitude.

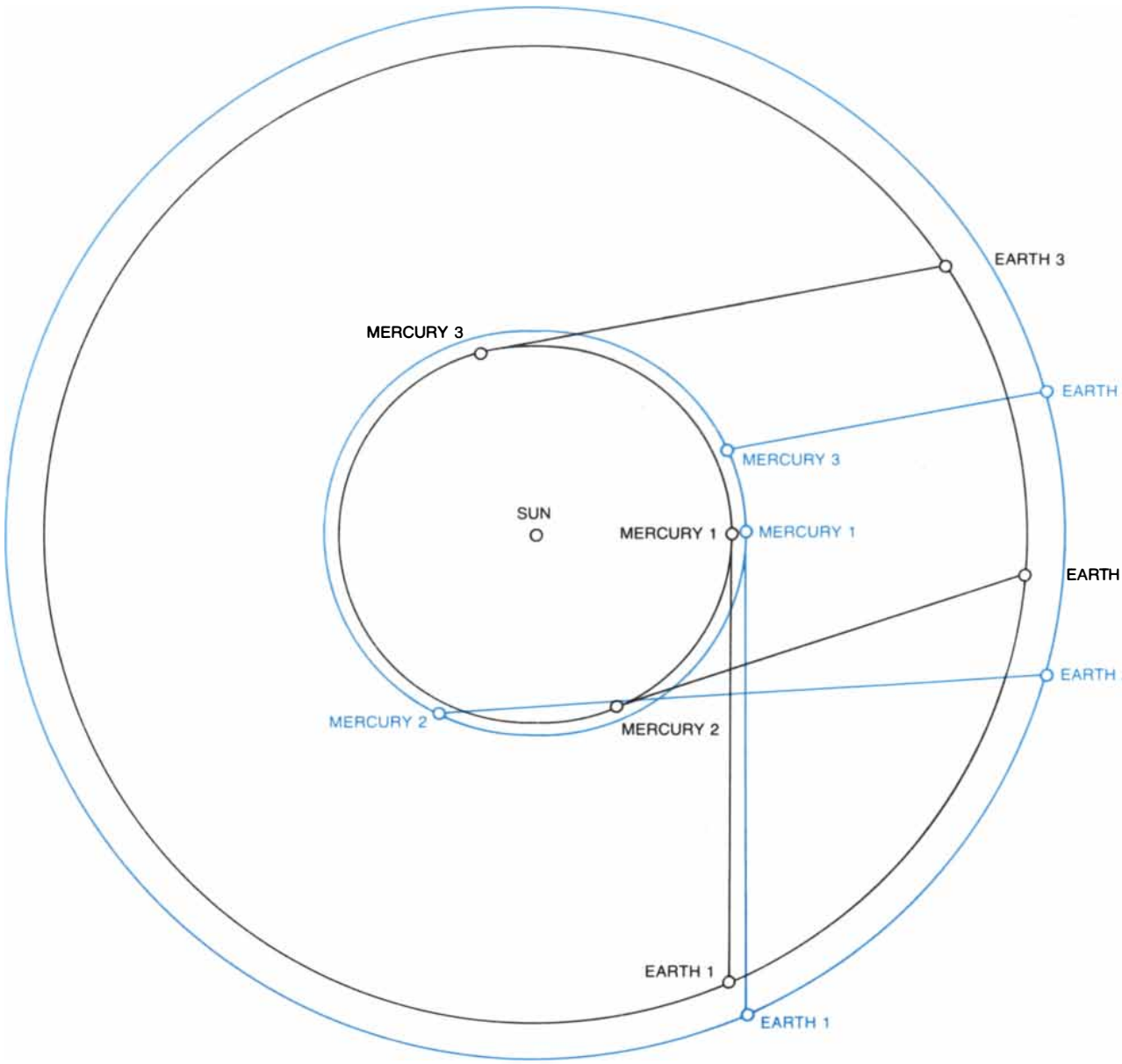
The largest possible alternative effect is one that would obtain if the mass of the sun were decreasing. The sun could lose mass by three processes. First, mass could be lost through the continuous emission of particles, such as those that make up the solar wind. Second, it could be lost through the continuous emission of electromagnetic radiation, which represents a mass loss through Einstein's equation $E = mc^2$, where E is the energy of the radiation, m is the mass lost and c is the velocity of light. Third, mass could be lost in periodic solar eruptions such as prominences or flares. When all the known possibilities are added up, the rate at which the sun could possibly be losing mass is still too small by a factor of 30 to account for the observed increase in the moon's orbital period. We can therefore be reasonably confident that known phenomena other than a change in the gravitational constant have been eliminated as a possible explanation of the observed increase.

It has long been known that the rate of increase in the moon's orbital period caused by tidal friction should be mathematically related to the rate at which the earth's rotation is being slowed by the same cause. That relation has never been satisfactorily observed; the moon seems to be receding at a rate that is greater than can be accounted for by the rate at which the earth's rotation is slowing. The cause of the discrepancy has long been sought. Moreover, the observed rate of the increase in the moon's orbital period seems to be substantially different in modern times from what it was a few thousand years ago. From

analyses of ancient observations of eclipses of the sun several investigators have obtained rates as high as 20 or 25 parts in 10^{11} per year. Now Paul M. Muller of the Jet Propulsion Laboratory of the California Institute of Technology has discovered that if one includes the effect of a known correction to the motion of the plane of the moon's orbit with respect to the plane of the earth's orbit, then the rate of the increase of the moon's orbital period derived from ancient eclipse observations can be brought into agreement with the observed rate derived from telescopic observations over the past three centuries. On this basis if the value of the gravitational constant does decrease, the theoretical relation between the rate at which the moon's period is increasing and the rate at which the earth's rotation is slowing is closely met, resolving the discrepancy.

Some critics of the hypothesis that the gravitational constant is changing have argued that if the earth was closer to the sun in the past than it is now, as the hypothesis requires, the earth should have been a good deal hotter then than it is at present. There is no evidence in the geological record for such a torrid period. Recently, however, Chao-Wen Chin and Richard Stothers of the National Aeronautics and Space Administration's Institute for Space Studies in New York have shown that in one form of Dirac's theory the sun would also have been considerably dimmer in the past, and that the two effects would nearly cancel each other. Thus the objection is not fatal to the hypothesis.

Hoyle has summarized studies of the initial chemical composition of the sun, and he has found that they clearly favor the hypothesis that the gravitational constant is decreasing. A number of planetary astronomers have pointed out that the large rift faults in the crust of the moon and Mars imply that these bodies have expanded somewhat since they were formed, perhaps when their gravity diminished as a result of a decrease in



RADAR-RANGING TO INNER PLANETS is another experiment in progress that will help to determine whether the value of G is changing. If G is constant, then the orbits (*black*) of the planet (in this case Mercury) will remain the same size, and Mercury will reach its greatest elongation (its greatest angular distance from the sun as it is seen from the earth) right on schedule. If G is decreasing, however, then orbits of both Mercury and earth will expand (*color*) and Mercury will reach its greatest elongation late.

G. That expansion could also be one driving mechanism for continental drift on the earth and could help to explain how a continent such as Antarctica can be almost surrounded by a mid-oceanic ridge and yet be apparently drifting away from the ridge at the same time.

In interstellar space there are an excessive number of double-star systems where the two component stars are widely separated from each other compared with the number of systems where the component stars are close together. The theory of the formation of double-star systems indicates that there should be fewer widely separated systems than are actually counted. If the orbits of the stars in such systems have been expanding for billions of years as *G* decreases, that might explain the discrepancy.

The stars in the globular clusters that surround the central disk of our galaxy are ancient; indeed, they appear to be older than the galaxy itself. If they were formed at the same time as or later than the bulk of the galaxy, the inconsistency in the ages is perplexing. On the other hand, if the value of *G* was once larger, stars would have evolved at a higher rate than they are evolving now, and so now they would appear to be older than they actually are. It should also be noted that the apparent expansion of the universe may itself be a result of the weakening of gravity. At the very least one would expect that if gravity were weakening, the expansion of the universe could never be slowed or stopped, and that expectation is in accord with the latest observational results.

In addition to the continuing observations of lunar occultations there are two astronomical experiments currently in progress that will soon be able to measure changes in the gravitational constant. The first is the lunar laser-ranging experiment, in which pulses of laser light are beamed through a telescope on the earth at one of the retroreflectors placed on the moon by the Apollo astronauts, and the reflected pulses are picked up by the same telescope. The time it takes for the light to complete the round trip is a direct measurement of the distance to the moon. The experiment can both measure changes in the moon's orbital period and directly measure the expansion of the moon's orbit. With that information it may soon be possible even to distinguish between the predictions of the various competing theories that assume some change in *G*.

The second astronomical experiment consists of precise radar-ranging measurements of the distance between the earth and the inner planets Mercury and Venus. As with the moon, changes in the orbital period of the planets are being sought. The radar-ranging method has the advantage that, since it does not involve the moon, it is not complicated by tidal friction. It has the disadvantage, however, that the irregular topography of the other planets can confuse the interpretation of the results.

There is also a plan by a group led by Rogers C. Ritter at the University of Virginia to set up a laboratory experiment designed to detect a change in *G*, somewhat along the lines of earlier experiments devised to measure the absolute value of *G*. In the experiment a number of small masses will be attached to a central shaft, and the entire assembly will be surrounded by several large masses. A drag system attached to the shaft will hold the small masses back so that they will not gradually rotate toward the large masses by virtue of their mutual gravitational attraction. The amount of drag exerted on the shaft gives a direct measure of the gravitational force between the masses, and any change in that force should indicate a change in the value of *G*. The hope is that the experiment will reach a precision of a few parts in 10^{12} per year. The laboratory experiment, if it proves feasible, will be independent of celestial bodies and of celestial motions, and it will therefore be very important for checking the interpretation of the results from all the other experiments.

Although it is too early to be certain that the value of *G* is decreasing, the data now available favor that conclusion. If the conclusion turns out to be correct, it will be a strong argument in support of Dirac's large-numbers hypothesis, which predicted the result nearly 40 years ago. What are some of the other implications of the hypothesis?

One implication would have particularly profound consequences for physics. A fundamental number we can add to the list of Dirac's fundamental constants is the total number of nucleons (protons and neutrons) in the universe. That number is estimated to be 10^{80}. By the same reasoning as before, since 10^{80} is the square of the age of the universe, which is 10^{40} in Dirac's fundamental units, one can conclude that according to the large-numbers hypothesis the number of nucleons in the universe must be increasing at a rate that is proportional to the passage of time squared.

Dirac has described two ways in which matter might be created throughout the universe. It might be created additively, that is, new nucleons might be spontaneously appearing everywhere. Alternatively, it might be created multiplicatively, that is, new nucleons might be appearing only where matter already exists, and at a rate that is proportional to the existing mass. Both the results from the lunar occultations and the geological record of the temperatures early in the earth's history favor the possibility that matter is created multiplicatively.

The Dirac theory's implication that matter is being created throughout the universe is one of the principal reasons for the theory's lack of popularity among physicists. Can we take seriously the possibility that new nucleons are continually being created within existing matter? If we can, why is the rate proportional to time squared? And what about the objection that has been raised that crystals, where atoms are arrayed in a fixed lattice, would be destroyed by such a process? I should like to propose an answer to these questions that may be somewhat more palatable than a strict creation theory, and that nonetheless satisfies the large-numbers hypothesis.

Let us imagine that some kind of shield against gravity can exist. Under ordinary circumstances every particle of matter attracts every other particle in accordance with Newton's laws. Suppose, however, it is possible for matter to become so dense that a particle's gravitational field can no longer penetrate it. Clearly ordinary matter is not that dense, but matter in the interior of certain stars might be. If gravitational shielding were possible, then the gravitational and inertial masses of such dense bodies might not be equal, that is, a body's resistance to acceleration would still be proportional to the total number of its nucleons but its gravitational mass would be proportional only to the number of unshielded nucleons.

With that theoretical model in mind, let us face the question presented by the hypothetical creation of matter. If the gravitational constant *G* is decreasing, then massive bodies will expand as the weight of their surface layers diminishes. The expansion will decrease the density of matter throughout the body, allowing additional nucleons to begin contributing to the body's external gravitational field. Hence there will be an apparent increase in the body's gravitational mass, and an increase in the inferred number of nucleons in the body, without any actual creation. Moreover, the increase in mass will be proportional to the square of the time, as required by the large-numbers hypothesis, since the surface area of the shielding layer of sur-

rounding matter increases with time squared. We may also note that there would be no destruction of crystals, since no nucleons are actually being created in the process.

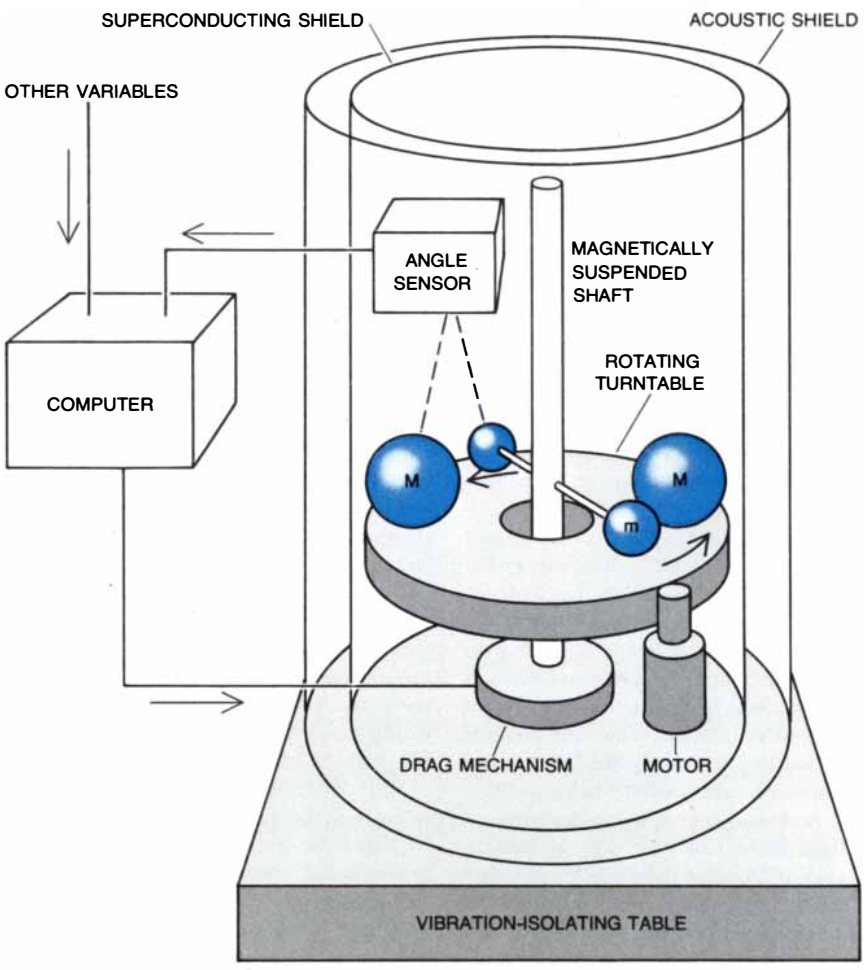
Although this hypothesis is highly speculative, I believe it serves to demonstrate that we cannot ignore Dirac's predictions simply because we cannot imagine a way in which new nucleons could be appearing in the universe. Proponents of the large-numbers hypothesis believe it reflects certain undiscovered laws of nature. In view of the precision of the experiments in progress we can be hopeful that those laws will not remain undiscovered much longer.

Is a decrease in the gravitational constant consistent with the general theory of relativity? As we have seen, the general theory in its present form assumes that G remains constant. Many alternative theories of gravity that are quite similar to the general theory do, however, assume that G decreases. Is there any reason to prefer a description of gravitational fields that is alternative to the description of the general theory?

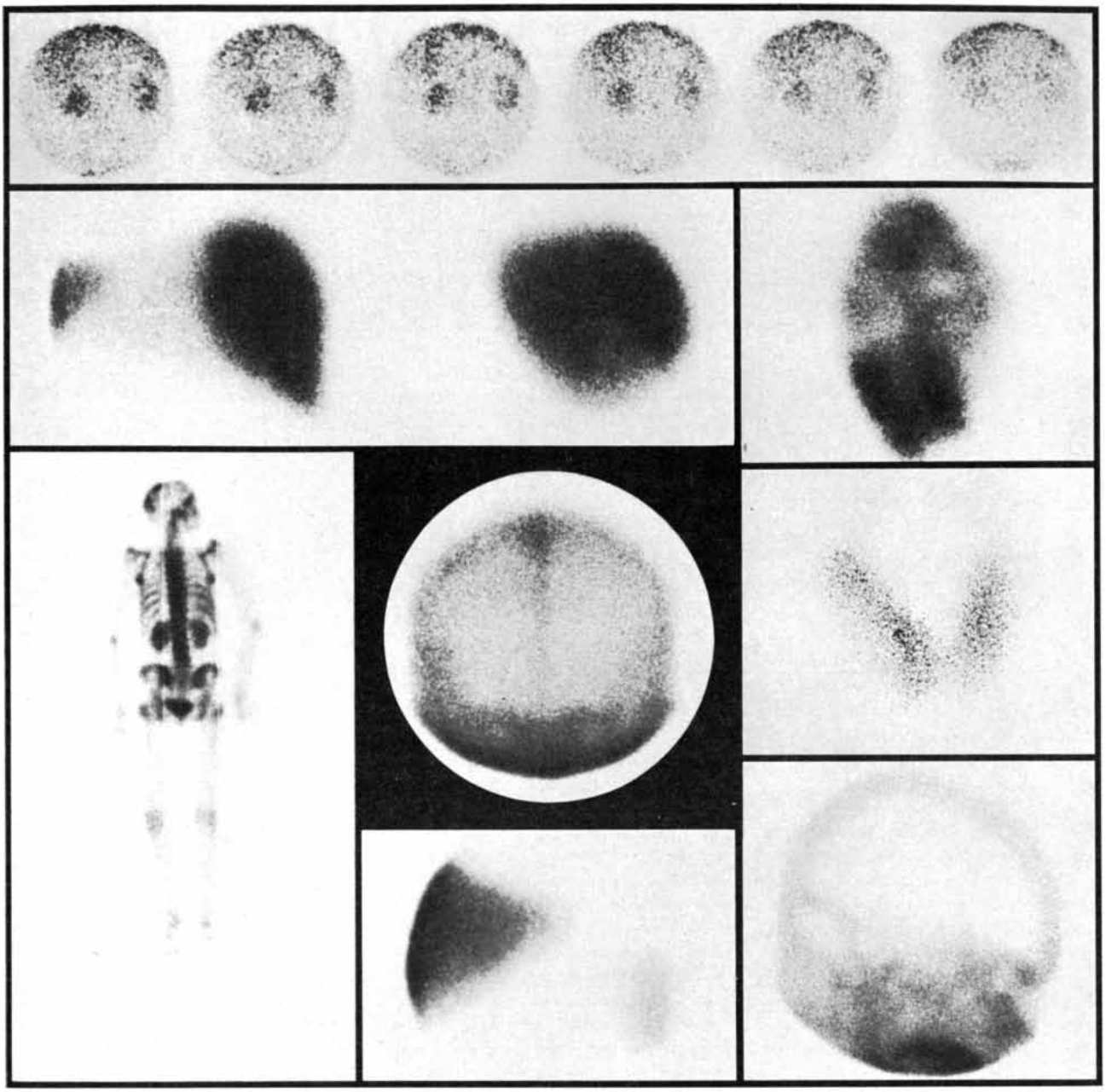
The special theory of relativity has been verified experimentally to high precision. Yet Hüseyin Yilmaz of the Perception Technology Corporation has pointed out that the general theory of relativity and the special theory are inconsistent beyond the first order. Efforts to resolve the inconsistency lead to a slightly different formalism, one that has deep consequences for the higher-order implications of the general theory of relativity. For example, the singularity that appears in Einstein's solution of the field equations disappears in Yilmaz' treatment, making black holes impossible. Donald H. Menzel of the Harvard College Observatory has recently developed a similar treatment. Furthermore, Menzel has examined the physical consequences of the modified field equations, and he finds that the properties of ultramassive stars seem to resemble the properties of quasars (compact and energetic radio sources of an unknown nature) and Seyfert galaxies (compact and highly luminous galaxies). I personally find the possibility that those objects are ultramassive stars more believable than the sparse evidence gathered so far to show that black holes exist.

Another method of preserving the general theory of relativity in its present form while allowing G to decrease is the Milne hypothesis. In the 1940's it was pointed out by E. A. Milne that we cannot readily distinguish between an expanding universe with a decreasing G on the one hand and changes in the atomic structure of matter on the other. If the radii of the orbits of electrons around atomic nuclei are decreasing, and if their periods of motion are diminishing as well, then our measuring rods for distance and time, that is, our atomic clocks, are changing. In such a situation the universe would appear to be expanding and G would appear to be decreasing when actually they were not. Hence the Milne hypothesis removes the conflict with general relativity altogether.

Clearly we have much left to learn about gravity. Although in the general theory of relativity we seem to have an adequate description of weak gravitational fields, there is still doubt that our descriptions of strong gravitational fields are accurate. Moreover, we still do not have any reliable answers to such fundamental questions about gravity as: Why should masses attract each other in the first place? Can gravity be shielded? Is it possible for negative gravity to exist? Observational, experimental and theoretical work on gravitational fields will remain one of the most intriguing, controversial and potentially rewarding areas of physics and astronomy for some time to come. The prospect that within the next few years we shall definitively measure the rate of change of the gravitational constant adds substantially to the excitement.



LABORATORY EXPERIMENT proposed by a group led by Rogers C. Ritter at the University of Virginia will measure changes in G to an accuracy of one part in 10^{12} per year without relying on the motions of celestial objects. The experiment will be isolated from vibrations in the earth and in the atmosphere, and will be cooled close to absolute zero. The small masses (m) are freely suspended and are gravitationally attracted toward the large masses (M). The entire system rotates at a constant velocity of three revolutions per minute to isolate it from extraneous surrounding masses. A drag system holds the small masses back so that they do not slowly rotate into the large masses because of their mutual gravitational attraction. The angle sensor measures the angle between the small masses and the large masses and feeds that information into a computer, which in turn tells the drag system how much force it must exert to keep the angle constant. Magnitude of that force gives a direct measurement of the gravitational attraction between small masses and large masses; any changes in the force will directly reveal any change in value of gravitational constant G.



Radionuclides at work.
Preferences vary on how best to see them.

Röntgen was a physicist, not a physician. He demonstrated the reality of x-rays with a picture of the bones in his wife's hand. Before he did so, physicians had not known how to see the bones underneath the living, intact flesh. Afterwards, though, Röntgen would not have been the best person to ask how your fracture was healing or whether you had a tumor.

So it goes in medical instrumentation. First the physicists fire off a concept, then the engineers rassle it back down to practical hardware, but all it does is make headlines until a substantial number of physicians have learned how to use the new hardware as an extension of the sensoriums they tote around between their personal ears.

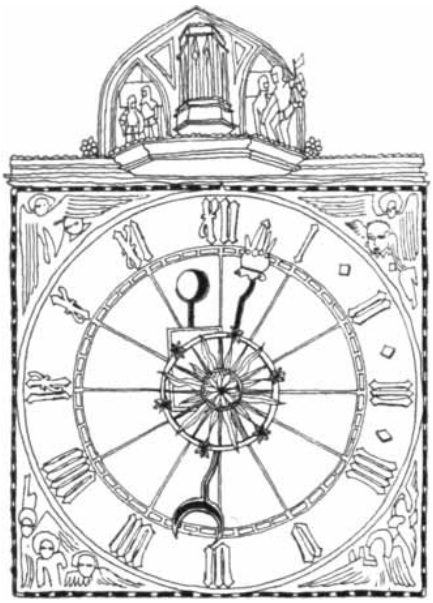
Facilitating in more ways than one the impression of the visual image on the medical mind is basic business with Kodak. This endeavor is in full swing in nuclear medicine, as it has long been in radiology. The term covers the use of radionuclides to delineate things going on within the body.

Thousands of patients each week now undergo such examinations. New radiopharmaceuticals appear. The instrument makers' engineers are in keen competition. We conceive it our obligation to pay close attention to what they are up to. We consult with them and their customers the better to deliver the generated image in its most readily interpretable form to the perceptual apparatus of the practitioners under conditions best suited to the daily tide of patients and patient records.

In addition, we are convincing nuclear medicine departments that we can save them money. Details available from Kodak, Dept. 55W, Rochester, N.Y. 14650.

Kodak

SCIENCE AND THE CITIZEN



Nuclear Test

The utilization of nuclear power will be subjected to an unusual test in June: citizens of California will vote yes or no on an initiative, a popular referendum on a law setting forth the conditions under which nuclear power plants can operate in the state. The official title of the statute is "Land Use and Nuclear Power Liability and Safeguards Act," but the issue is more clearly drawn by the titles proponents and opponents apply to the vote. The proponents call it the Nuclear Safeguards Initiative and say it is an attempt to let the people decide on measures that will ensure the safety of their lives and property. The opponents call it the Nuclear Shutdown Initiative and maintain that it is simply an effort to ban nuclear power plants in the state.

The statute provides for a cutback in nuclear-plant output and a prohibition on new nuclear-plant construction unless specified things happen before specified dates. The first such condition applies to the limit of $560 million set by Federal law on liability for damage caused by a nuclear-plant accident. If the limit is not removed by law or by waiver by a utility within a year after the initiative vote and "full compensation assured" under normal judicial procedures, no new fission-power plants would be allowed in California, and the output of each plant failing to guarantee full compensation would be cut back to 60 percent of full power. Beginning five years after the initiative the power level would be cut back another 10 percent per year until the limit on liability is removed.

The other conditions involve reactor and fuel-disposal safety. Within five years after the initiative the state legislature would have to find, by a two-thirds vote, that "the effectiveness of all safety systems" in California nuclear plants had been "demonstrated by comprehensive testing in actual operation," and that radioactive wastes could be stored or disposed of "with no reasonable chance" of adverse effects. In the absence of such a finding the prohibition of new plants, the cutback to 60 percent and the subsequent 10 percent cutbacks would go into effect. Even before that, by June, 1979, the 60 percent reduction would be required unless the legislature found by a two-thirds vote that it was "reasonable to expect" that the five-year goal would be met. In making its decision the legislature would be guided by an advisory group of at least 15 "concerned citizens" and experts in fields ranging from nuclear engineering to sabotage techniques; no more than a third of the members could be people financially connected with work on "nuclear fission power plants or their components."

Proponents of the initiative, organized as Project Survival and Californians for Nuclear Safeguards and supported by such established environmentalist groups as the Sierra Club and Friends of the Earth, argue that nuclear-reactor safety mechanisms have never been adequately tested, that there have been a large number of reactor accidents, some of them serious, that a major accident could devastate an entire region and that the risks of nuclear power outweigh the benefits, which they maintain have been overstated. The initiative, they say, is aimed at forcing the industry and its Government regulators to improve safety procedures and also to back up their belief in nuclear safety by coming out from behind the current liability limit. The initiative route, they maintain, is the truly democratic way to achieve major reform by popular demand.

Opponents of the initiative, including such largely industry-supported groups as the California Council for Environmental and Economic Balance and Citizens for Jobs and Energy, argue that reactor and fuel safety are reasonably assured, that nuclear power is on balance less harmful than fossil-fuel power and that the need for nuclear power to supplement other sources is clear and increasing. They insist that a yes vote would shut down California's nuclear plants because the conditions are—and indeed were designed to be—impossible to achieve. In particular they attack the two-thirds requirement, which means that 14 state senators could turn off nuclear power in California. Finally, they argue, the sophisticated issues at stake in the controversy cannot be effectively decided by a statewide popular vote influenced by simplistic arguments and catchy phrases such as "Don't Grow Nuclear Plants."

Novel Novas

From the time of Newton until two decades ago the prevailing view of the universe was that it is much like a vast clockwork, with only an occasional nova or supernova punctuating its orderly processes. In the 1960's, with the discovery of quasars, pulsars, celestial X-ray and gamma-ray sources and other unexpectedly energetic phenomena, it became evident that it is a far more turbulent place. Now still another kind of energetic object has been recognized: the X-ray nova.

The X-ray sources that have been studied since the first one was detected in 1962 either have radiated steadily, like the Crab Nebula and other supernova remnants, or have varied regularly, like Her X-1 in the constellation Hercules. Some of the regularly varying sources appear to be members of eclipsing-binary systems; some vary so rapidly (from less than a second up to 10 seconds) that they appear to be pulsars.

The first transient X-ray source to be detected flared up in Centaurus in 1967; it was designated Cen X-2. At a recent meeting of the American Astronomical Society, George W. Clark of the Massachusetts Institute of Technology reported that since then scores of other X-ray novas have been recognized from such X-ray satellites as *Uhuru*, *Copernicus*, *Ariel-5* and SAS-3 (the third Small Astronomy Satellite).

In April of last year a transient X-ray source designated A0535+26 flared up near the position of the Crab Nebula. It emitted hard, high-energy X rays and pulsed regularly with a period of 104 seconds. After fading out in July, it reappeared in August, faded again and

The spear-head, the arrow, the rocket.

From ancient instinct to computer design, the shape that cleaves the air is the wedge.

Now Triumph brings the wedge down to earth in TR7—a fantastic new sports car to steal the American road.

Proven on the Grand Prix race tracks of the world, the edge of the wedge knifes through the air, forcing the front wheels down. Handling is solid and uncannily precise. The air flows over the slippery shape. Drag is minimal. Power enhanced. Miles per gallon increased.

Under its skin, TR7 is a triumph of simplicity. It is utterly dependable, sturdy and sinuous on cantankerous roads and corkscrew curves. It means endless pleasure at moderate price and less maintenance; welcome news today and something no rival sports car can say.

All in all a simply beautiful and beautifully simple machine. Test-drive our words at any Triumph dealer.

For the name of your nearest Triumph dealer call: 800-447-4700. In Illinois call 800-322-4400. British Leyland Motors Inc. Leonia, New Jersey 07605.

BRITISH LEYLAND TRIUMPH

THE SHAPE OF THINGS TO COME TR7

reappeared in early November. Meanwhile an even more spectacular X-ray nova appeared in the constellation Monoceros; on August 13 it reached a maximum output of soft, low-energy X rays that was four times higher than the energy output of the brightest nontransient X-ray source known: Sco X-1 in Scorpius. By early December it was still the brightest X-ray object in the sky.

What is the nature of the new sources? Do they represent a single kind of object or a variety of objects? Kenneth Brecher and Philip Morrison of M.I.T. propose that they should be divided into two classes. The hard-X-ray sources such as A0535+26 may be energized when matter falls onto a rapidly rotating, magnetized, collapsed object similar to the objects that may underlie the nontransient sources Her X-1 and Cen X-3. The soft-X-ray sources, such as the nova in Monoceros, may emit their radiation when gas ejected from the system is heated by shock waves traveling through it. Observations of the Doppler shifts of A0535+26, together with its pattern of disappearance and reappearance, further indicate that it is a member of a binary system in which the two components have an unusually wide separation. Other investigations indicate that the transient X-ray sources are confined to the plane of our galaxy, and that as many as 100 of them may flare up every year.

Disinheritance

The "man-apes" of the genus *Australopithecus* flourished in Africa from at least 1.5 million years ago until perhaps as recently as 750,000 years ago. They therefore comfortably fill the evolutionary gap between the emergence more than 12 million years ago of the first apes with certain manlike features and the appearance some 500,000 years ago of the genus *Homo* (in the form of *Homo erectus*). Evidence of the australopithecines' ability to stand upright and to use (and probably to manufacture) stone tools, which has come to light in recent decades, has strengthened the assumption that they were man's immediate precursor. More than one effort has been made to reclassify at least some specimens of *Australopithecus* as *Homo*, even though few of them have a brain volume greater than 500 cubic centimeters. (Most *Homo erectus* brains were more than 800 c.c.) Now one investigator argues that *Australopithecus* is not even an ancestor of man.

Writing in *Nature*, the British anatomist C. E. Oxnard, who is now at the University of Chicago, reviews the anatomical features in which *Australopithecus* resembles man on the one hand and various apes on the other. Oxnard finds the African hominid unique in some characteristics and in others closer to the orangutan than to man. He further notes that functional analysis of both the pelvis and the talus (an ankle bone) indicates that *Australopithecus* was far from being able to walk upright in the human sense.

Oxnard points to evidence that manlike hominids existed in Africa perhaps millions of years before the appearance of *Australopithecus*. That evidence includes the recent discovery in East Africa by Richard Leakey of a very manlike talus bone that is perhaps 2.6 million years old, the discovery, also by Leakey, of a skull with a brain volume approaching 800 c.c. that is more than two million years old, and the discovery a few years ago in East Africa by Bryan Patterson of Harvard University of a manlike fragment of arm bone that may be more than four million years old. Unless ankle bones, skulls and arm bones evolved through cycles in Africa, being more manlike at early stages, less manlike at intermediate stages and then once again more manlike, Oxnard argues, it is very unlikely that *Australopithecus* occupied a position on the evolutionary line leading to man.

Giant Atoms (as Big as Bacteria)

The biggest atoms are not those of some heavy, transuranic element but quite light atoms in a highly excited state. They can be very large indeed. The diameter of a normal atom (even if it is a heavy one) is about 10^{-8} centimeter; some excited atoms have a diameter of 10^{-5} centimeter and are as large as certain bacteria.

Interest in such inflated atoms has been aroused in part by new methods for creating, manipulating and detecting them. They are under investigation in several laboratories and were the topic of recent meetings in Seattle and Tucson sponsored by the American Physical Society.

The energy state of an atom is denoted by its principal quantum number, designated n, which defines the probability of finding an electron at a particular distance from the nucleus. In the state of lowest energy (the "ground" state) n is 1, and the electrons are effectively confined to a quite small volume; when the atom absorbs energy, n increases in integral steps and electrons are likely to be found at larger distances from the nucleus.

Atoms have now been prepared in the laboratory with n as high as 105, and hydrogen atoms in interstellar space may have even higher values of n. Such atoms are on the very brink of ionization; with only a small additional input of energy they fly apart into an unbound electron and a positive ion. Surprisingly, however, in isolation they are quite stable, in some cases surviving for as long as a millisecond. The stability results from certain dynamical impediments to their decay; transitions to the ground state are inhibited.

James Bayfield and his colleagues at Yale University have created highly excited hydrogen atoms by allowing accelerated protons to capture electrons. In this way neutral atoms are formed in many different energy states, including some with very high values of the principal quantum number n.

Another technique, capable of creating a uniform population of excited atoms, all of which have the same value of n, has been employed by Thomas Gallagher, Stephen Edelstein and Robert Hill of the Stanford Research Institute. They excite atoms of an alkali metal, such as sodium, with light from a tunable dye laser; the wavelength, and hence the energy, of the light emitted by such a laser can be precisely controlled, so that it promotes transitions to particular excited states. The use of sodium atoms instead of hydrogen alters the experiment very little. In highly excited states both kinds of atom can be described as a single electron orbiting a distant, compact center of positive electric charge; in hydrogen that center of charge is a proton, and in sodium it is the rest of the atom.

Sodium atoms excited by laser light are also being studied by Daniel Kleppner, Richard R. Freeman, Theodore W. Ducas, Michael G. Littman and Myron L. Zimmerman of the Massachusetts Institute of Technology. They have combined with the method for creating individual excited states an equally specific technique for detecting them. As n increases and the force binding the electron to the nucleus gets weaker, the electric field required to ionize the atom decreases. By applying a potential of the appropriate intensity Kleppner and his colleagues ionize all atoms in a particular chosen state. Because the ions are electrically charged they can readily be detected and counted. By increasing the potential in discrete steps it is possible to take a census of the population of atoms in each excited state.

Highly excited atoms might be employed in the investigation of a number of processes in atomic physics. For example, they might illuminate the events surrounding the ionization of atoms by an electric field. Collisions of excited atoms are also of interest. It is possible that an atom in the ground state could pass through an excited atom without causing any observable change. Kleppner has proposed that transitions from states with a high value of n could be employed to measure the Rydberg constant, which determines the wavelength of the radiation emitted in transitions in hydrogen atoms.

The giant atoms might also have certain practical applications. The possibility that they could be employed in the separation of isotopes for uranium enrichment is being studied at the Stanford Research Institute. A finely tuned laser would promote atoms of one isotope to a state of moderate excitation, leaving the atoms of other isotopes in the ground state. A second laser would then convert the excited atoms into highly excited ones that could easily be isolated by ionization.

Kleppner has suggested that the atoms could also be employed in a photon detector sensitive in a region of the electromagnetic spectrum extending from the infrared to microwaves with a wavelength of a millimeter. Such long-wavelength radiation is difficult to detect because it has too little energy to alter the state of most atoms or molecules, but it is capable of ionizing highly excited atoms. Such a detector would be particularly useful in astronomy, since the astronomer's customary photon detector—a photographic emulsion—is insensitive to all but the shortest infrared waves.

Diabetes: More than One Disease?

For roughly half a century—ever since the discovery of insulin by Frederick Banting and Charles Best in the early 1920's—diabetes has been widely regarded as a reasonably well-understood, controllable disease: the failure (apparently inherited) of certain cells in the pancreas to secrete enough native insulin, the hormone responsible for mediating the body's normal metabolism of sugar. The telltale chemical symptoms of this debilitating metabolic disorder are well known: an accumulation of sugar (glucose) in the bloodstream, with the excess sugar spilling over into the urine to be excreted. By the comparatively straightforward stratagem of daily insulin injections combined with careful dieting, severe cases of diabetes can be stabilized and the prospect of a quick

death from the disease or its complications can be averted.

Nonetheless, as a recent book on the subject points out, diabetes "remains today one of medicine's most baffling riddles. No one knows what really causes it, how to prevent it, whether it is one disease or two or three, or how to prevent the far-reaching and devastating effects it has on the human body." The quotation is from *Diabetes: New Look at an Old Problem,* by Bertrand E. Lowenstein, M.D., and Paul D. Preger, Jr., Harper & Row, 1976.

The statistics on diabetes are not reassuring. According to the U.S. Public Health Service, the disease is currently a leading cause of blindness, kidney failure, disorders of the nerves and tiny blood vessels, gangrene of the limbs and various sexual malfunctions. In addition diabetes has been associated with heart attack, stroke and other cardiovascular diseases. The upshot is that 50 years after the advent of insulin treatment the mean life expectancy of the diabetic is still 25 percent shorter than that of the nondiabetic. To make matters worse, recent findings suggest that the oral drugs prescribed to control the blood-sugar level accomplish little of lasting value for patients with the forms of diabetes that do not require insulin.

In the face of this generally negative picture, research on the incidence, causes, control and prevention of diabetes is undergoing something of a renaissance. What is more, several promising new approaches to the problem, now in the early investigative stages, may greatly improve the outlook for someone with diabetes. For one thing, as a recent editorial in *British Medical Journal* comments, "it is becoming clearer that diabetes is not a single entity but a number of different diseases in which a raised blood sugar is the common feature." The most widespread form of the disease, called maturity-onset diabetes (because it usually shows up after the age of 40), definitely "appears to have a strong hereditary tendency," the editorial continues. "It has been calculated that 60% of the children of two maturity-onset diabetics will themselves develop mild diabetes by the age of 60. On the other hand, contrary to previous theory, classical juvenile-onset diabetes seems to have a much weaker genetic component. This is heartening for juvenile-onset diabetics who ask about the chances of their children becoming diabetic—they can be reassured that the risk of youth-onset diabetes in their children is probably no more than 1%."

The evidence for a nongenetic origin

of the more acute juvenile-onset, or insulin-requiring, form of diabetes comes from a number of sources. The fact that the incidence of juvenile-onset diabetes has been found to vary with age, season, year and geographic area indicates that environmental factors may be important. Recent investigations in Britain and the U.S. have revived interest in the possibility that certain viruses, such as the mumps virus or the Coxsackie B4 virus, may be a triggering agent in juvenile-onset diabetes. (It has already been shown experimentally that viruses can induce diabetes in certain animals.) In order to explore the possible link between juvenile-onset diabetes and viruses, workers in Britain have been conducting an ongoing series of tests involving identical twins, some of them discordant (only one twin diabetic) and others concordant (both twins diabetic). Although the latest such study reveals "no direct evidence" of a viral cause of juvenile-onset diabetes in the twins, the experimenters report that "the age and time of onset suggested that environmental factors may be important in causing diabetes in the twins."

Referring to the British identical-twin studies, the book by Lowenstein and Preger remarks: "Even though the evidence available is certainly not enough to convict the virus as the villain, it does suggest that investigators should keep on its trail, particularly in view of the fact that if insulin-dependent diabetes is indeed triggered by a virus, then it might be possible one day to produce a vaccine against this type of diabetes." Taking into account not only the viral-etiology studies but also several other currently active lines of research on diabetes (such as the transplantation of part or all of a healthy pancreas or the implantation of some artificial insulin-secreting device), Lowenstein and Preger conclude that "there are glimmers of hope that diabetes may be much more controllable perhaps five years from now."

Rod Cells and Color Vision

When the human eye is fully dark-adapted, vision is provided by the rod cells of the retina, which are roughly 1,000 times as sensitive to light as the cone cells that provide the normal daytime vision. Most people have observed that in dim light, when only the rods are functioning, color vision disappears: all objects are seen in shades of gray. When the light level is high enough, one can see a spectrum of colors from blue to red as the cone cells are stimulated by wavelengths from 400 to 700 nanometers. The curves of photopic, or daylight,

vision have three sensitivity maximums: one for the blue-sensitive cones at 445 nanometers, one for the green-sensitive cones at 535 nanometers and one for the red-sensitive cones at 575 nanometers. The curve for scotopic, or night, vision has a sensitivity maximum at about 500 nanometers. Over the entire spectrum, however, the rods functioning alone do not give rise to any sensation of color.

Anatomical and physiological studies going back many years suggested that the rods might share one or more of the transmission channels by which the cones relay information to the brain. There were also a few experiments indicating that under special conditions a limited sensation of color could result from rods and cones acting together. Then John J. McCann and Jeanne L. Benton of the Vision Research Laboratory of the Polaroid Corporation showed that a multicolored image can be generated by a display in which the rods are stimulated at a wavelength of 550 nanometers and at a light intensity far below the threshold of the cones, while at the same time only one of the three sets of cones is stimulated at a wavelength of 650 nanometers. And when the wavelength stimulating the rods is only a little shorter (500 nanometers), the multicolored image seen is nearly indistinguishable from the image that is generated when the intensity of the light at both wavelengths is high enough to stimulate the cones. In other words, the display evokes essentially the same multicolored image whether the rods are supplying the "short-wave" information or the cones are supplying it. The colors in the display include blue-greens, yellows, oranges, browns and reds, in addition to white, grays and black.

One of the displays used for these experiments consists of what the Polaroid workers call "the Mondrian." This is an assemblage of variously colored papers, cut into rectangles and overlapped so that the papers have surrounds of various hues. The panel is photographed on black-and-white film through a green filter, which yields a short-wave record, and separately through a red filter, which yields a long-wave record. When positive transparencies of the two records are projected in superposition through filters similar to those through which they were taken, one sees a multicolored image remarkably like the original. As Edwin H. Land of Polaroid originally demonstrated some years ago, the superposed images are just as colorful if the long-wave record is projected in red and if the short-wave record is projected simply in black and white, without the

green filter (see "Experiments in Color Vision," by Edwin H. Land; SCIENTIFIC AMERICAN, May, 1959).

In a recent lecture at the Polaroid Vision Research Laboratory, McCann demonstrated to the members of a small audience that they could see a variety of colors in the Mondrian with only their rods and one set of cones. After allowing time for the viewers' eyes to become dark-adapted, McCann first illuminated the Mondrian with a dim light passed through a narrow-band filter with a wavelength of 550 nanometers. When just enough light was used for the patterns to appear distinct, they were seen to be colorless because only the viewers' rods were being activated. Then the light was turned off and the Mondrian was illuminated with a light source of 650 nanometers, again barely above the threshold of visibility. At that wavelength it was apparent that the threshold of the red cones had been reached: the Mondrian was faintly reddish all over, varying only in lightness from patch to patch. When the Mondrian was illuminated with both weak light sources simultaneously, a dim but unmistakably multicolored image could be seen. The light of 550 nanometers falling on the Mondrian was 50 to 100 times weaker than the level needed for cone vision at that wavelength.

In his lecture McCann presented the results of experiments, conducted with Suzanne McKee and Benton, in which the long-wave (red) input was held constant at low levels as the wavelength of the short-wave input was varied between 400 and 600 nanometers. Subjects were asked to adjust the intensity of the short-wave component until the combination produced a variety of well-balanced colors. When the intensity levels are plotted against wavelength, it is clear that the curve matches the sensitivity curve for rod vision and not the curves for any of the three cones.

As in the case of Land's earlier experiments with normal color vision, the rod-cone results have shown that the colors one sees are not necessarily related to the wavelengths of light reaching the eye. Land demonstrated, for example, that the ratio of red and green light reflected from a particular portion of a multicolored chalk drawing can be adjusted so that it exactly matches the ratio from another portion, yet one portion (an awning) will look green and the other (a door) will look red. The same result is obtained, McCann demonstrated, when only the rods respond to the short-wave illuminant and the cones respond to the long-wave one.

FORAGE CROPS

Grasses and legumes that are fed to livestock are the mainstay of U.S. agriculture. Among their virtues is the fact that they can be grown on land that is not suited to growing other crops

by Harlow J. Hodgson

The statistics of agriculture say little about forage (grasses and legumes grown as feed for livestock) because most forage crops are consumed on the farm where they are grown and do not reach the marketplace. Forages are nonetheless the most important crops in the U.S. from several points of view. More land is devoted to them than to all other crops combined. They take up some five times the acreage of all grain crops. The dollar value of forages, expressed in terms of their contribution to human food of animal origin, exceeds the value of any other crop. Indeed, the agricultural system represented by forage crops and the ruminant animals that feed on them (mostly cattle and sheep) can be said to be the backbone of the nation's agricultural economy.

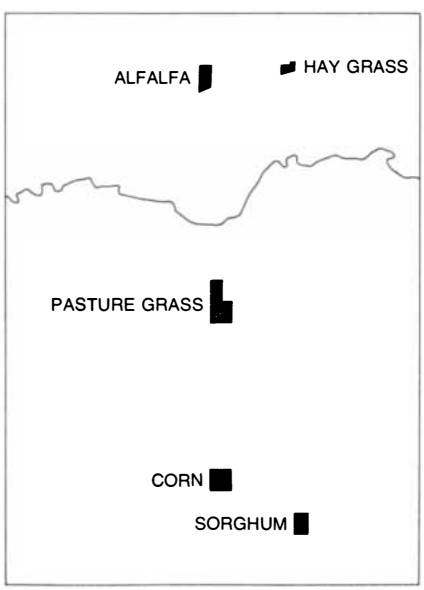
The vitality of the forage-ruminant system is essentially based on the fact that ruminants can digest forage crops. Although the composition of forages varies with the crop, such crops are all characterized by a relatively high content of cellulose, hemicellulose and lignin. It takes the multiple-chamber stomach of ruminants, with its rich flora of microorganisms, to break down those tough fibrous materials and convert them into food. Monogastric animals, such as pigs, chickens and human beings, must rely mostly on more concentrated foods, such as cereals, fruits, nuts and certain vegetables. Ruminants can digest these foods too and may be fed them when they are available at low cost.

A key element in the success of forage-crop agriculture is that much of the land on which forages grow is not suitable for other crops, being too hilly or stony, too wet or dry or otherwise unsuited to intensive cultivation. Large areas of the U.S. fall into such categories, including most of the 245 million hectares of permanent pasture and rangeland. (A hectare, the metric measure of land, is 2.47 acres.) On the other hand, forages are often grown on good land and play an important role in soil management and in providing a maximal economic return on an investment in ruminant livestock. In fact, the area of cropland serving as pasture or for growing forages stored as hay or silage is about equal to the amount planted to food grains and feed grains. Forages grown on those lands are rotated with other crops and therefore must be reestablished every few years.

Much forage is grazed: the animals are turned out into the field or onto the range to feed directly on the growing crop. Large amounts of forage, however, are harvested and processed in some way before being fed to animals. The commonest method of harvesting a forage crop is to make it into hay by letting it dry in the field for a day or two after it has been cut and then store it in bales or as loose hay in stacks or in a barn. Another method is to make green chop, meaning that standing green forage is chopped up and fed immediately to cattle or sheep, which are usually kept confined in a limited area instead of being turned out to pasture.

Forage crops are also made into silage and haylage. Both processes entail chopping the crop finely and storing it in a silo where it ferments. The difference between silage and haylage lies in how long the crop is dried before it is ensiled; silage consists of from 30 to 40 percent dry matter and haylage of from 50 to 60 percent dry matter.

In growing a forage crop the farmer's aim is to maximize the production of vegetative tissues: the leaves and stems. With most other crops management is directed toward the production of fruit or seed. The methods employed by farmers to produce and utilize forages are



FORAGE CROPS COVER A SUBSTANTIAL PART of the South Dakota farmland shown in the photograph on the opposite page. The photograph was made on infrared film from an altitude of 60,000 feet in early August. The fields that are the most intensely green on the ground most strongly reflect infrared radiation and appear in the false-color scheme of the film as red. Fields appearing as blue-green and blue are bare; the darkest ones have been recently plowed and thus have more moisture at the surface than other fields. The photograph was made as part of a test of remote sensing conducted by the Remote Sensing Institute of South Dakota State University on behalf of the National Aeronautics and Space Administration. Remote sensing provides not only a means of identifying a crop but also a means of checking the condition of the crop. To verify the interpretation of the photograph a number of fields were checked on the ground; a few are shown on the map at left.

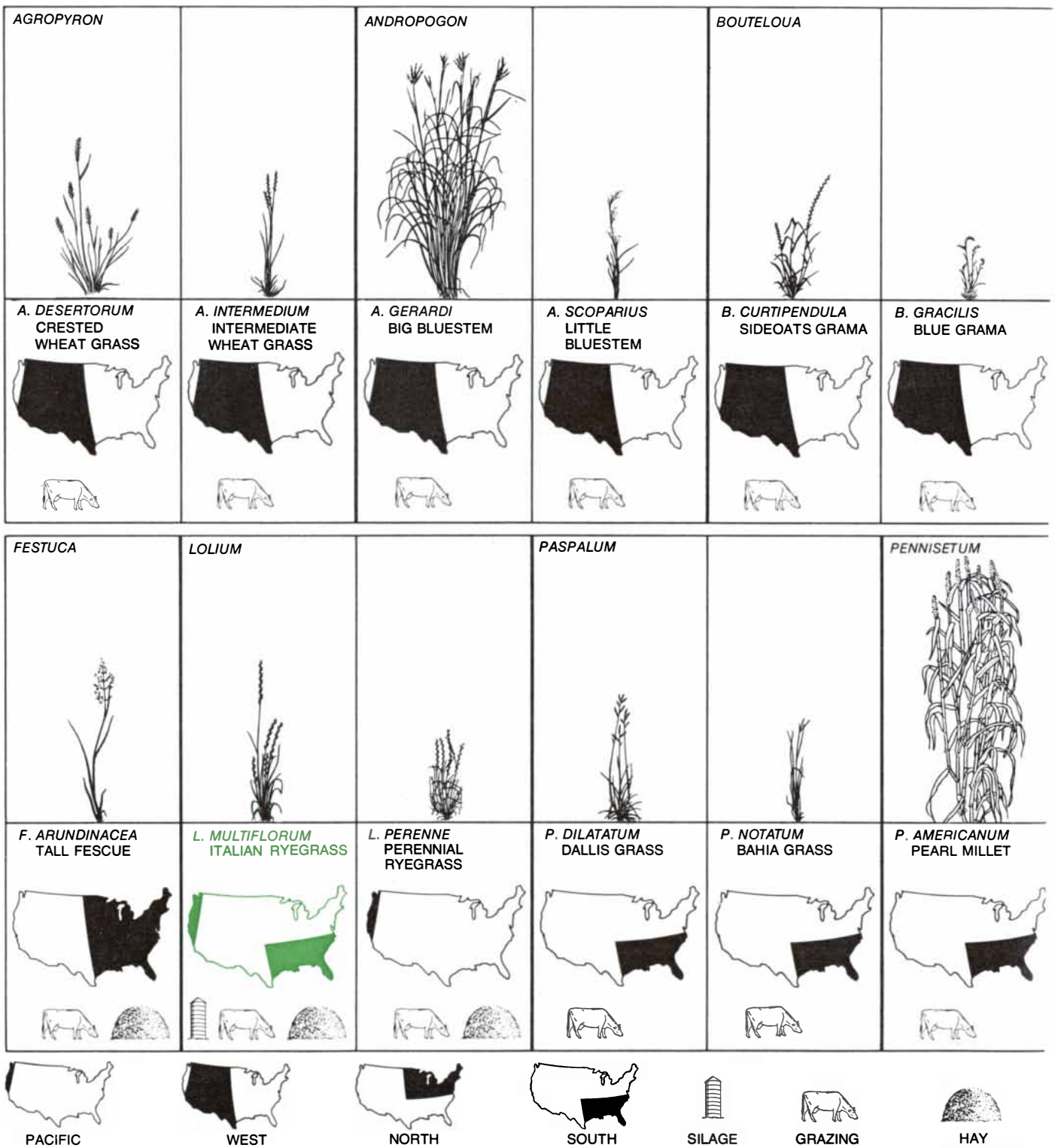
strikingly different from the practices applied to other crops.

The Biology of Forages

Most forage species are members of two large and important plant families: the Gramineae (grasses) and the Leguminoseae (legumes). It is noteworthy that the world's major food crops are also members of those families. Rice, wheat, sugarcane, corn, sorghum, millet, oats, rye and barley are grasses. Soybeans, peas, chickpeas, peanuts and a variety of edible beans are legumes.

About 35 species in some 25 genera of grasses and about 15 species in some 10 genera of legumes are considered important forage crops [*see illustration below and illustration on page 64*]. More than 100 species of grasses and legumes are grown for forage in the U.S. Many of them were introduced from other parts of the world. In addition a number of native species grow on permanent grassland (such as the rangelands of the West) and on permanent pasture in better-watered areas of the country.

The majority of forage species are herbaceous perennials, meaning that

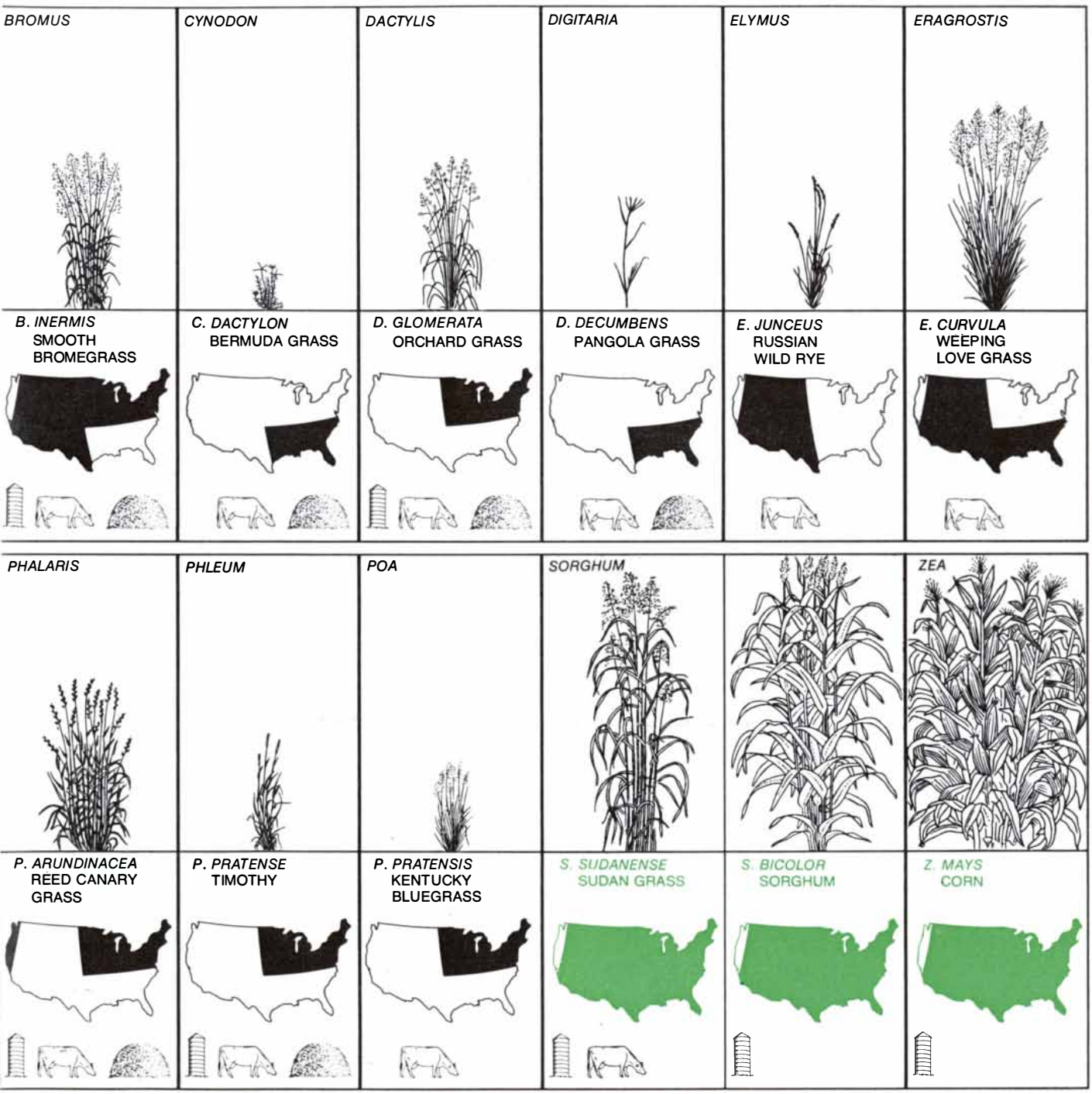


they have little or no woody tissue and do not need to be planted every year. They survive unfavorable conditions, such as freezing, drought and high temperature, by becoming dormant. Active growth resumes from preformed buds when the environmental conditions change for the better.

A few forage crops are annual plants. They include Sudan grass, millet and certain small grains that serve mostly for grazing. They also include sorghum and corn, which as forage crops are mostly made into silage. The acreage planted to such annual forage crops, however, is small compared with that devoted to perennial forages.

Perennial grasses have a wide variety of forms. The fundamental unit of growth is the tiller, or shoot, which consists of a stem with several leaves and with a growing point at its apex. Fibrous roots are at the base of the stem. The grass plant as a whole consists of a number of shoots.

The growing point produces new cells that develop into stem and leaf tissue until a certain stimulus induces the new cells to develop into flower primordia. The stem may remain short and be near



MAJOR GRASSES grown as forage crops in the U.S. are classified according to genus and species, common name (such as crested wheat grass), the region of the country where each grass is mainly grown and the principal use of the crop. Most of the grasses are perennials, meaning that they do not have to be planted every year; a few important ones, however, are annual species (*color*).

the surface of the ground for a considerable time. In this condition the part of the plant that is aboveground consists largely of leaves; the plant is said to be in the vegetative state.

A stem has a number of nodes from which the leaves originate, and internodal spaces called internodes. The stem becomes visible when several of the uppermost internodes elongate as a result of both cell division at the base of the internodes and the elongation of the cells. During the growth process the stem extends to its full height, and the leaves associated with the uppermost nodes are displayed along the stem. By the time the stem has reached its full length (if the plant is allowed to grow that long) an inflorescence will have been produced. If an elongating stem is grazed or harvested and the growing point is removed, further growth must come from the activation and development of buds at the base of the old stem; they form new shoots, which remain associated with the old stem but become largely independent of it.

Some of the Gramineae, such as Bermuda grass and pangola grass, produce stolons, stems that creep along the surface of the ground. Leaves and sometimes new vertical shoots are produced at the nodes of the stolons. Other grasses, among them bromegrass, produce rhizomes, which are stems that grow horizontally below the surface. Rhizomes produce only rudimentary leaves at the nodes until the growing point turns upward and penetrates the surface. Then a shoot appears that is capable of becoming independent of the mother plant. A single plant of a stoloniferous or rhizomatous grass may occupy a large area as a result of such vegetative or asexual reproduction.

The grass inflorescence appears at the end of the stem. It consists of a number of rather inconspicuous small flowers that are arranged in a pattern characteristic of each species. Each flower contains one floret or several florets, each of which is a complete reproduction unit that (in most grasses) contains both the male and the female reproductive organs. When the egg and the pollen cells are mature, the male organs (anthers) are exerted (pushed out of the flower) and the pollen is shed. The pollen is disseminated by the wind, so that the plants are cross-pollinated. Some grasses shed pollen without exertion of the anthers; these species are self-pollinated.

Each floret produces a single seed. In most forage grasses the seeds are quite small; a pound may consist of several hundred thousand seeds. Certain grasses produce such small seeds that several million of them are required to make a pound.

Characteristics of Legumes

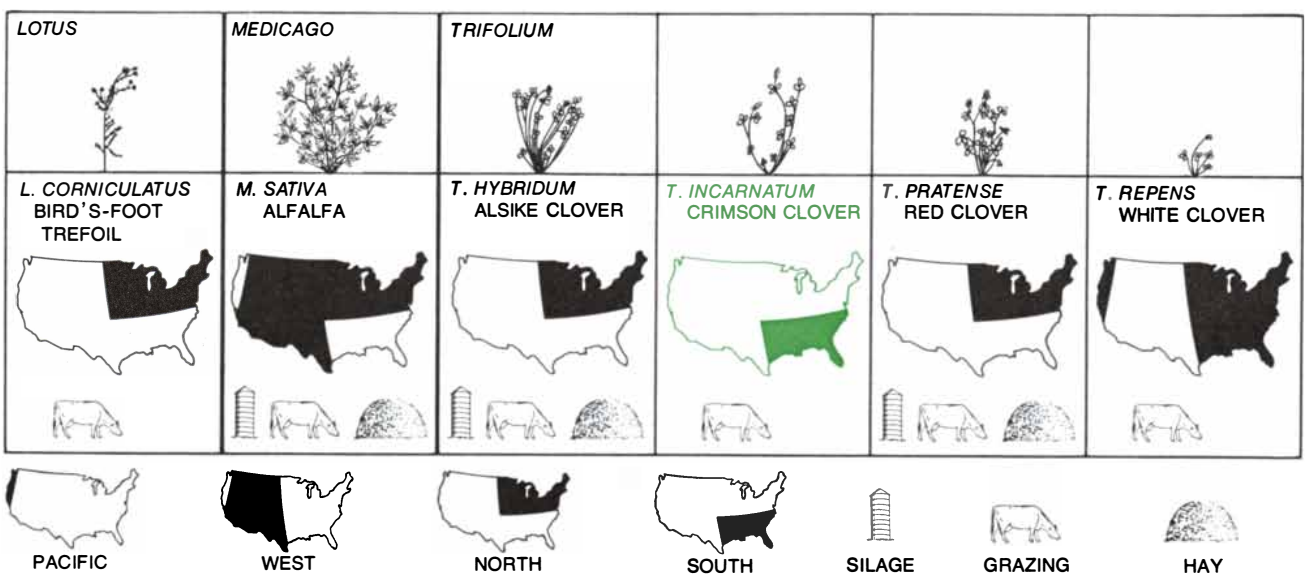
Legumes have somewhat different patterns of growth. A typical legume has a taproot that penetrates deep into the soil. In some species the roots branch moderately. A seedling has a primary root and usually a single stem, which produces buds at the surface of the soil or near it.

The leaves, which in most legumes have a trifoliate pattern, are arranged alternately on the stem. At the end of the stem and of each of its branches is the growing point, encased in immature leaves. When the stem is removed or growth is resumed after dormancy, the buds at the soil surface produce new stems (with branches and leaves), which in turn produce buds at the base. Thus a crown is formed at the soil level. The crown enlarges with age, and as long as it is functional it gives rise to each succeeding crop of stems. A few legumes of importance (white clover, for example) form stolons or rhizomes.

The flowers of legumes develop either from the axil of the leaf (the point where the leaf meets the stem) or from the terminus of the stem. They are usually more conspicuous than the flowers of grasses. The color and size of the flowers of many legumes and the fact that they contain nectar make them attractive to insects, which provide the means of cross-pollination. Legumes vary greatly in the size, shape and color of their flowers. Some legumes, including the soybean and the field pea, are self-pollinating.

The seeds of legumes also vary considerably. Some are tiny and others are quite large. A single flower may produce one seed or several, and the seeds are always borne in a pod.

Perhaps the most singular characteristic of legumes is their ability to estab-



PRINCIPAL LEGUMES grown as forages in the U.S. are depicted according to the scheme employed for grasses in the illustration on the preceding two pages. From an agricultural viewpoint a significant difference between a grass and a legume is that the legume can provide its own nitrogen through its symbiotic relationship with *Rhizobium* bacteria, which convert atmospheric nitrogen into forms the plant can utilize. Legumes in general contain more protein than grass does and are more readily digested by livestock.

lish a symbiotic relationship with *Rhizobium* bacteria, which form nodules on the roots. The plant provides the bacteria with a food supply and the bacteria in turn convert atmospheric nitrogen into forms the plant can utilize. Nitrogen compounds are the essential subunits of protein, and so the relationship with *Rhizobium* accounts for the fact that legumes (both as forage and in seed form) are relatively rich in protein.

In a few instances a similar relationship between a grass and a microorganism has been found. There is keen interest among plant breeders in developing plant-bacteria relationships that will provide nitrogen for grasses. With the cost of manufactured nitrogen fertilizer rising, success in this endeavor would be of enormous agricultural significance.

The most important forage legume is alfalfa, which is grown on about 12 million hectares in the U.S. (In many European countries it is called lucerne.) Alfalfa produces more protein per hectare than any other crop.

ALFALFA IS HARVESTED on a farm in New York State. The cut alfalfa is left in windrows (*left*) to dry for a few hours. When it has dried to about 35 percent dry matter, it is chopped finely and stored in a silo. Alfalfa is the principal forage crop in the U.S.

The Physiology of Forages

The basic physiological processes of forages (photosynthesis, respiration, transpiration and so on) are the same as those in other plants. Perennial forages, however, exhibit certain unique features. In order to survive the winter they require energy reserves both to maintain their living tissues through the period of dormancy and to provide energy for new growth when the period ends. Moreover, plants that are defoliated by grazing or harvesting several times a year, as forages usually are, must maintain energy reserves in a readily accessible form to regenerate stems and leaves. The energy stored for those purposes is carbohydrate, usually in the form of starch or complex sugars.

In regions where the temperature of the air and the soil goes below freezing during the winter living tissues must adapt physiologically in order to withstand the cold. The complete biochemistry of the development of resistance to cold is not yet understood.

Forage grasses and legumes are usually grouped into warm-season and cool-season species according to the type of environment in which they evolved. The members of the two groups exhibit a number of physiological differences, mostly related to temperature. One particularly interesting difference is the process of vernalization required by most cool-season grasses. Only after the apex of the stem has been exposed to low temperature or a short day length (sometimes both) can it be induced to initiate flower primordia. The induction process is photochemical, occurring in the fall when the temperature and the length of the day are decreasing and when the apex of the stem is unelongated and encased within a number of leaf sheaths, which protect it. The actual initiation of flower primordia usually comes in the following spring in response to rising temperature and lengthening days.

Most forage species are polyploid, that is, the number of their chromosomes is some multiple higher than 2 of the basic chromosome number. Usually the polyploid level is such that the plant has four, six, eight or more sets of chromosomes. For this reason and others the genetics of most forage species is quite complicated, and geneticists have not devoted anywhere near as much effort to understanding them as they have to crops such as corn and wheat. As a result relatively little genetic information is available on any important forage species. For some of them scarcely anything is known beyond the chromosome number. Until such information is developed, progress in improving forages is likely to be slow.

Because of the small seeds characteristic of most forage species, establishing a crop by seeding is considerably more difficult than establishing a crop with large seeds. Particular care is required to see that the small seeds are planted shallowly enough for the tiny seedlings to emerge. Close contact between the seed and the soil is necessary to ensure that the germinating seed has a constant and adequate supply of moisture. For these reasons a smooth, firm seedbed is needed for the establishment of a good stand of forage. Even so, many seedlings fail; the seeds must therefore be planted in large numbers in order to establish a satisfactory number of plants. Seed is usually broadcast on the surface and lightly covered or drilled shallowly in rows a foot or less apart.

A special problem arises with legume seeds because of the desirable relationship between the legume plant and the *Rhizobium* bacteria. In order to ensure that the relationship is satisfactorily achieved it is a common practice to inoculate legume seeds with a culture of the appropriate strain of bacteria before planting the seeds. Inoculated seed should be sown and covered as promptly as possible.

A few grasses are propagated vegetatively rather than by seed. The stems of such grasses (examples are Bermuda

grass and pangola grass, which in the U.S. are grown mostly in the Southern states) will produce roots and new shoots from the nodes. A common method of propagation is to collect stolons or rhizomes of the grass and spread them either randomly or evenly over the field. They are then partly covered by disk-harrowing the soil. With favorable moisture conditions the grass becomes established in a full stand quite rapidly.

A common practice is to grow forage as a mixture of grass and legume, sometimes with one species of each and sometimes with more than one. Typical combinations are bromegrass together with alfalfa and timothy together with red clover. One reason is that different species can tolerate different soil conditions, so that a mixture often results in a fuller stand and higher productivity in a given field than a single species would produce. The less ideal the site is, the more advantageous a mixture is likely to be. Furthermore, a mixture is sometimes easier to manage than a grass or a legume alone in a field that is being grazed. If the legume stand is lost, the remaining grass will produce at least a partial crop.

Many cultivars, or varieties, of some forage species have been developed. The cultivars are likely to differ in such important characteristics as winter hardiness, maturity, adaptation to soil and climate, resistance to disease and insects, rate of recovery after grazing or harvest, and yield and quality of forage. It is therefore important for the farmer to make a proper choice of both species and cultivar according to the growing conditions and to the planned use of the forage.

The Need for Nutrients

Although a forage crop needs good nutrition as much as any other crop does, the tendency in the U.S. has been to apply far less fertilizer to forages than to other crops. The reason is that most farmers employ their limited resources of capital to fertilize cash crops. A legume has a high requirement for phosphorus and potassium, although it will usually provide its own nitrogen. It also needs a soil that is only slightly acid or neutral (a *p*H of about 7), which on many fields calls for an application of lime to raise the *p*H.

A grass, on the other hand, has a large requirement for nitrogen and a smaller but still significant requirement for phosphorus and potassium. Grasses grow over a wider range of *p*H than legumes. Both grasses and legumes require a number of minor elements such as boron, magnesium, manganese and copper, but it is not necessary to supply them except in small amounts and on certain soils. The need for doing so increases with the length of time that the field is cropped.

Given a soil of adequate moisture and fertility, a well-managed forage crop can produce as much digestible energy for livestock as any other crop. The independence of legumes (such as alfalfa) from added nitrogen is becoming increasingly a factor in their favor as the price of nitrogen fertilizer increases.

Many of the nation's permanent grasslands have been grazed for decades (some for more than a century) without the addition of fertilizer. Even though animal manure is returned, the result is depleted soil and reduced productivity, often to the point where good forage species have been supplanted by inferior ones. Such lands can be restored if fertilizer and lime are applied as required and superior plant species are reestab-

HAYLAGE is removed from a trench silo for feeding to cows on a dairy farm. Haylage is similar to silage except that it is left to dry longer before it is chopped and stored. This trench silo has a covering of plastic over the top. It is also possible to leave a trench silo uncovered; the top few inches of stored silage or haylage will spoil, but the rest will be adequately preserved as feed.

lished through the procedure called pasture renovation. To rehabilitate land in this way often calls for a certain amount of tilling. Only in recent years has special equipment for accomplishing the task become available. A successfully renovated pasture yields as well as a rotation pasture on good cropland.

Since a perennial forage requires reserves of energy to survive the winter and to recover after grazing and harvesting, it is particularly important for the farmer to regulate the frequency of grazing and harvesting in such a way as to maintain those reserves. If he does not, the plants gradually become weaker and more susceptible to environmental stress. The accumulation of adequate reserves in the fall is particularly important in northern areas to avoid the winter-killing of the plants. Grasses, because of their morphology, are managed much more easily than most legumes.

The production of forage seed is a highly specialized business, requiring machines and managerial techniques different from the ones employed in growing and harvesting a forage crop. Most seed is grown in the drier parts of the country, where the water supply can be controlled by irrigation and where the pollination of legumes by insects is more efficient and more easily controlled. Accordingly most of the seed is produced in the West. A successful variety of forage must therefore produce not only good yields of forage but also profitable yields of seed in a distinctly different environment. These divergent requirements add considerably to the task of the forage breeder.

CORN IS HARVESTED for silage. Corn (*Zea mays*) is a member of the grass family and an annual plant. The harvesting machine being towed by the tractor cuts the corn near the base of the stalk and then chops the entire plant into the fine material that is being blown into the forage wagon. Corn serves both as a forage and as a major feed grain for livestock.

The Utilization of Forages

Although forage crops play an important role in soil conservation and good land use, the primary reason for growing them is to provide feed for ruminant livestock. The driving force in livestock production is the intake of digestible energy. The higher the rate of intake of digestible energy is, the greater the rate of production of meat and other animal products will be.

The nutrient needs of animals are constant from day to day over fairly long periods. Forages, however, do not grow uniformly. Their rate of productivity is markedly influenced by fluctuations in such environmental factors as temperature and moisture. Some forages grow well in the cool temperatures of spring and fall and poorly in the high temperatures of summer; for others the reverse is true.

The most economic utilization of forage, then, is to match the biological capacities of the crop with the nutrient needs of the animals. The successful manager usually grows a number of different forage species and varieties and employs a variety of management practices to maintain an adequate supply of nutritious forage for his livestock throughout the year.

Two factors that govern the intake of digestible energy are the total consumption of forage and the digestibility of the forage. These factors are in turn influenced by, among other things, the species or variety of forage, the maturity of the crop and the palatability of the forage. For most species of forage both digestibility and palatability diminish with the advancing maturity of the crop. On the other hand, the total yield of digestible energy increases with maturity, usually until about the time the plants begin to produce seed. The most practical time to harvest is shortly after flowers first appear, so that a good compromise is reached between the quality of the forage and the total yield of energy. Harvesttime is also influenced, however, by the nutritional requirements of the livestock.

To a large extent ruminant livestock obtain their feed by being put to pasture. Breeding herds of beef cattle and sheep are almost totally maintained on pasture except in severe winter conditions, when grazing is not possible. The farmer can get an adequate gain of weight by beef animals and a reasonably good output of milk out of pasture alone. If he supplements forage with feed concentrates such as grain, the yields will be still higher.

A pasture provides feed at the lowest cost. The animals do the harvesting, so that not much need be spent for machinery and labor. Fencing, however, is a significant cost. Exceptionally high yields of feed can be obtained from pasture if the farmer grows appropriate forages and employs sound management practices.

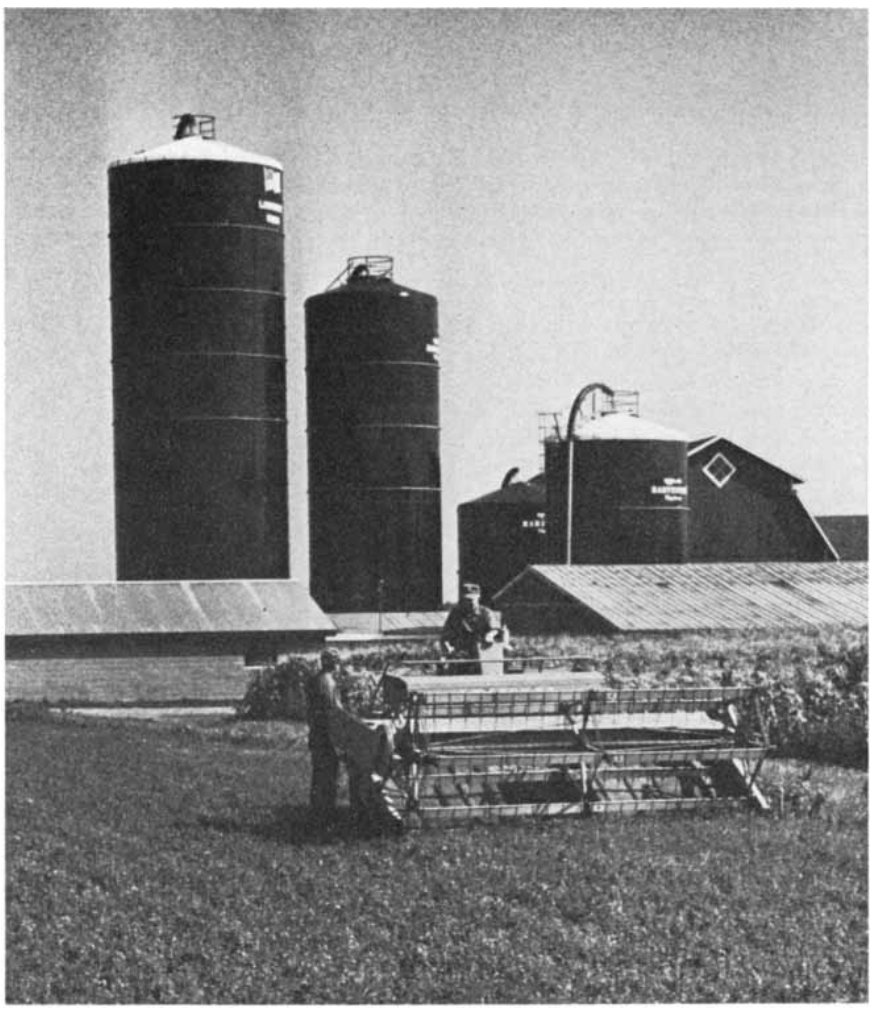
Stored feed comes into play when grazing is not possible. In recent years the practice of feeding large herds exclusively on stored forage has increased. The practice lends itself to mechanization; on some farms machines perform all the operations of harvesting, transporting, processing and storing the forage and of feeding it to livestock. Almost any forage can be stored.

Hay continues to be the most impor-

tant stored forage. The crop is cut and allowed to dry on the ground until the moisture content is from 15 to 20 percent. It can then be baled, stored unchopped or chopped and stored. Most of the hay produced in the U.S. is baled. Recent developments have almost completely mechanized haymaking and greatly reduced labor costs. The principal disadvantage in haymaking is that the loss of nutrients as the crop is cured, handled and stored is high. Losses of from 300 to 500 kilograms of protein per hectare and of similar amounts of carbohydrate are not uncommon. Those losses are physical and biological in nature and usually affect the most digestible portions of the forage.

Haylage and silage are stored in a moister condition than hay is. The interval between mowing and storing is much shorter than it is for hay; for silage it may be only a few hours. The concentration of moisture for ensiling must be no higher than about 65 percent or the preservation of the material will be poor and the palatability will be low. Storing forage as haylage or silage somewhat reduces the amount of labor and time involved in handling the crop, but the loss of nutrients can still be substantial. Greatly improved methods of harvesting and handling forage that is to be stored are needed in order to reduce the loss of nutrients and the amount of energy, labor and time now required.

In ensiling, the crop ferments in the absence of air. When the crop is ensiled, the plant cells continue to respire for a period of time and the aerobic bacteria on the plant increase in number. In a few hours, however, the oxygen is used up. In this anaerobic condition bacteria that produce lactic acid multiply rapidly. They act on the carbohydrates in the stored material to produce a number of acids, which effectively pickle the crop. Within about three weeks the *p*H of the stored material is down to about 4.2, and the crop will then keep for as long as several years if air is excluded.



MODERN SILOS appear behind a field of alfalfa on a dairy farm in New York. They are Harvestore silos, shiny blue structures made of steel plates coated on both sides with cobalt blue glass fused to the steel by heat. The function of a silo, which may also be made of wood or concrete, is to minimize the amount of oxygen that can reach a crop and cause spoilage. When the crop has been chopped finely and put in the silo, it is pickled (and thus preserved) by the action of lactic acid bacteria that thrive in the absence of oxygen.

The Economics of Forages

Since only a small part of the total production of forage enters market channels, statistics on the production and value of forage crops must be derived by indirect methods. A common standard of measurement is what forages contribute as livestock feed in the production of human food of ruminant-animal origin. According to the U.S. Department of Agriculture, about 240 million metric tons of feed units for livestock was supplied from forages in 1974. (A feed unit is the nutritional equivalent of .454 kilogram, or one pound, of corn grain.) All but about 4 percent of it was fed to ruminant animals; the 4 percent went to hogs and poultry. The tonnage of forage feed units was only slightly less than the tonnage of all the grain produced in the U.S. in 1973.

Forages supplied 60 percent of the feed units fed to all livestock in 1974, 82 percent of the units fed to beef cattle, 63 percent of those fed to dairy cattle and 89 percent of those fed to sheep and goats. With increasing amounts of grain going into export channels, the price of grain has risen significantly in the past two years, and so the trend has been to feed less grain and more forage to livestock, particularly beef cattle. Therefore the percent of forage in the feed of beef cattle increased from 73 percent in 1972 to 82 in 1974.

Cash receipts to farmers from beef and dairy cattle and sheep in 1972 were almost $26 billion, which was slightly more than a third of all farm cash receipts. The value of the forage that contributed to the feed of those animals was about $12.5 billion. No other crop approached that value.

Forages and ruminant livestock are important in every part of the U.S. In 1973 beef cattle provided more cash receipts to farmers than any other commodity in 21 states and ranked among the top five income-producing commodities in 47 states. Dairy cattle were the highest producers of income in nine states and among the top five producers in 39 states.

Farmers in the U.S. produced about 233 million metric tons of grain in 1973. About 65 percent of it was fed to livestock (44 percent to ruminants). Excluding wheat, which is the primary grain in the human diet, 75 percent of

WE LOVE SKI LOVERS

the kind who love le super ski

...who love to zip through sweet powder. Ski their hearts out. From early morning to day's end. This is le ski hot. Le ski Québec. In 110 centers. With 130" average snowfall a year. Low cost ski packages. Super *après ski.* Superb French food. Great accommodations. Friendly *hospitalité.* Le super ski is Québec. Not far for some far-out skiing.

come ski our big 4

SKI THE OTHER AMERICA ⚜ QUÉBEC

See your travel agent or write to: **Québec Tourism** dept. B52J2, Québec City, Canada G1R 4Y3. **In U.S.** Dept. B52J2, 17 West 50th Street, New York 10020

Trade your hog

You loved that car of yours when it was new, didn't you?

All bright and shiny and everything?

How could you know the world would change? How could you know it would use too much gas, too much space, too much everything?

How could you know that there would be a VW Rabbit?

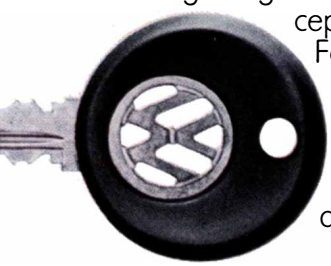
We knew 5 years ago. That's when we started—from absolute scratch—to design the one and only car that honestly makes any sense today.

It's almost as if we re-invented the wheel. We re-thought every detail, every old concept of car design to get to the new concept of the Rabbit.

For example, size:

The Rabbit is a big car. Not to look at. Outside, it parks in almost no space.

But inside, there is more glass area than in a Lincoln Continental Mark IV; as much legroom and headroom as some mid-sized cars; and (with the rear seat down) more luggage space than a Cadillac Fleetwood.



Even the key is padded for safety.

Seat belts put themselves on.

Quite an achievement.

Another example: safety.

The Rabbit helps make you a safer driver.

Front-wheel drive gives you better tracking.

Rack-and-pinion steering gives you better handling.

If one front tire blows, negative steering roll radius helps you to a straight stop.

If one brake circuit fails, a second circuit is still there.

A totally new rear axle makes the car more stable (and safer) on rough roads.

The deluxe Rabbit also has an utterly unique safety belt that puts itself on as you sit down.

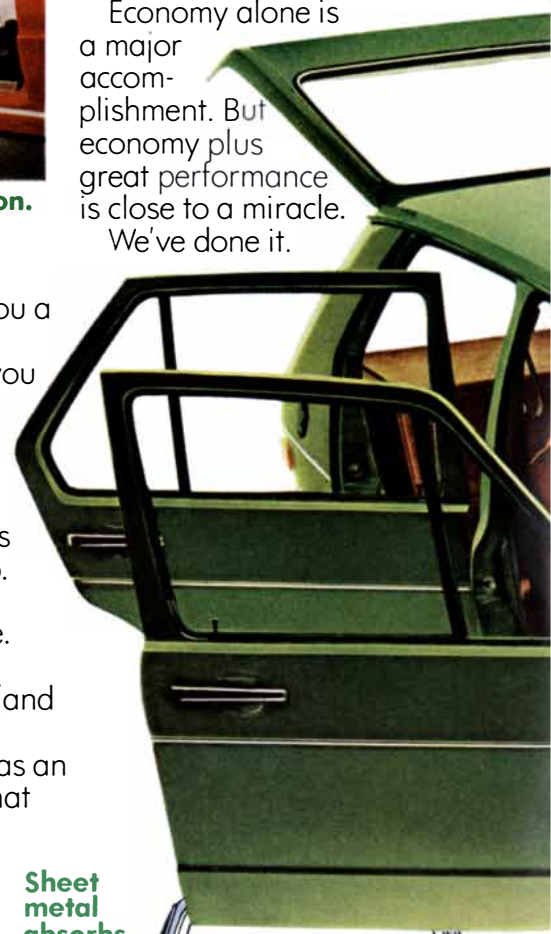
Still another example: economy and performance.

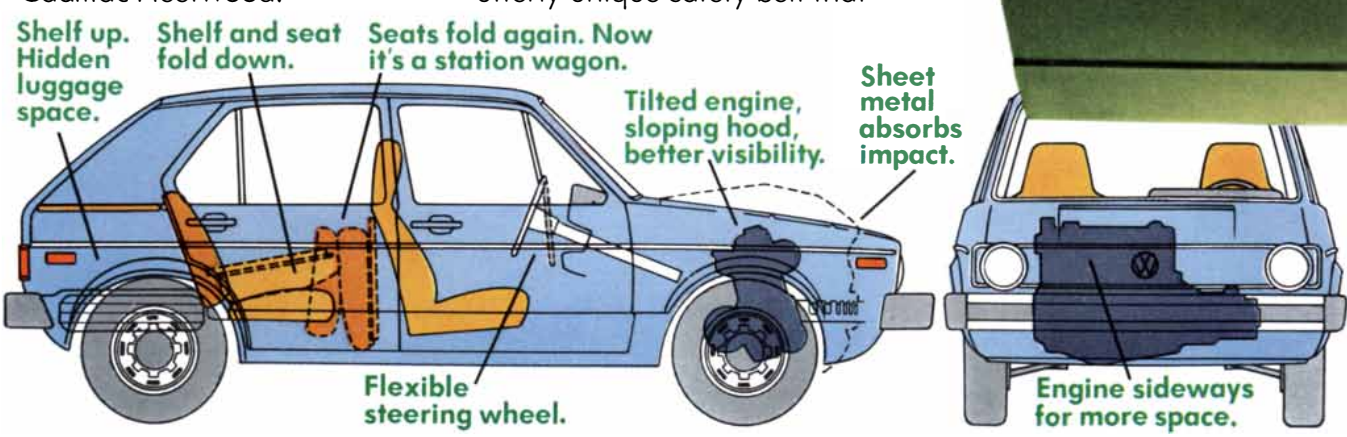
The EPA estimates that the Rabbit with stick shift got 39 mpg on the highway, 25 mpg in the city.

(Actual mileage may vary, depending on type of driving, driving habits, car's condition and optional equipment.)

Economy alone is a major accomplishment. But economy plus great performance is close to a miracle.

We've done it.





for a Rabbit.

0-50 mph in 8.2 seconds. No other car gives you this kind of power with so much economy.

The Rabbit zips from 0 to 50 mph in a mere 8.2 seconds. A $21,000 Maserati doesn't do much better.

No other car gives you this kind of power with so much economy.

There are more than 80,000 VW Rabbits hopping around the U.S. already.

It is the most successful new foreign car introduction in history.

The Amazing Rabbit

Our biggest sale was to Detroit's Big Three–56 Rabbits. We don't think it was a gesture of friendship or because they liked the colors.

They are tearing them to pieces to find out how we did it. We did it the hard way–from scratch.

But it was worth it.

The Rabbit is a happy car, and we're happy with it.

At first, we were worried about the future of the car.

Now we've come up with the car of the future.

Before you buy any computer or peripheral, read about these men.

They've bought hundreds from Digital.

"We depend on over 100 computers so New Yorkers can depend on us."

That's Al Feyler speaking for the Planning and Mechanization Department of the New York Telephone Company, New York City region.

"Our region handles 7 million main stations and some 30 million

"After 4 solid months of competitive testing, we found nothing could compete with Digital."

The business is leasing data communications equipment. The company is Leasco Data Communications Corp., a subsidiary of Reliance Group.

Leasco has put thousands of data communications systems in hundreds of companies all across the country. As part of its agreement, Leasco provides fast maintenance service, no matter where

"We didn't get a computer, we got a company."

Tano Corporation is one of the hundreds of OEM customers that put Digital computers in their products. James Reiss, President of the company, explains the reasons for going with Digital.

"Our company designs complex electronic control systems for oil and gas pipelines, chemical plants, utilities and municipal services. We needed a company that could provide sophisticated hardware and software to match."

At their present rate of growth, Tano expects to double their already significant PDP-11 pur-

Saving managers millions.

Low cost computers that pay for themselves over and over again is what has made Digital one of the two computer companies in the world to install over 50,000 computers.

And what this means to you is that right at the time you need it most you have access to an enormous reservoir of experience and ideas for cutting costs while getting more done.

Digital is the pioneer in minicomputers, the leader in the development of small- to large-scale timesharing systems, and the company that has made many of the major breakthroughs in computerizing techniques.

Because of Digital, interactive computing, computer hierarchies and networks, data communications systems and new uses of computer systems are growing at an unprecedented rate.

And every year Digital places a remarkably comprehensive selection of computer tools within the financial reach of thousands and thousands of new users.

Today anyone who has a good applications idea for a computer can afford one. And just as important, they can get the help they

phone calls daily. Computer power helps us provide our customers with the best possible service."

In only 4 years, Mr. Feyler's department has installed over 104 computers. 76 from Digital Equipment Corporation.

"Our computers perform a wide number of administrative and maintenance operations. Among other things, they report network and switching problems, monitor traffic flow, analyze trunk call data, make equipment assignments, switch messages, and report coin box troubles.

"In short, they help keep one of the world's most complicated communications networks running smoothly."

As far as the future for New York Telephone and New York Telephone customers is concerned, Mr. Feyler explains: "Right now, we're planning a far more complex network by 1985. In fact, we'll be using about 600 minicomputers in all phases of our operation."

Why so many Digital computers? "Cost, architecture and overall reliability were the big factors in our decision." ■

the unit is. Since every service call results in a direct loss to the company, the obvious goal is to minimize service calls. According to Executive Vice President, Carl English, one of the best ways to handle that is to select the most reliable equipment.

"Before we chose a peripheral vendor, we put four competitive teleprinter terminals through a grueling four-month test. Then we had a team of technical experts evaluate the results and the equipment to determine which teleprinter would ultimately be the most reliable. Digital was the unanimous choice."

Since June, Leasco has bought hundreds of DECwriter Teleprinters from Digital. With hundreds more on order.

"Our customers want continuous use of their machines," English says, "and Digital gives it to them." ■

chases in the coming year.

"We depend on timely delivery to keep our production going. That's why we picked Digital. They've given us the true production line, high volume operation we need.

"What's more, they back their hardware and software with good service.

"When we're trying to sell a system to somebody for a remote area, the first question asked is, what if it breaks down. We just give him the telephone number and address of the local Digital field service office. That's a powerful selling point. And one that no other minicomputer company can match.

"After buying 150 Digital computers, I can't say enough about them. Hardware, software, support – they have everything." ■

need from a company with over 4000 field representatives in 25 countries around the world.

To find out how Digital can help you save money, send for your free copy of "50,000 Computers Saving Managers Millions." Fill out the coupon below, or call or write Digital Equipment Corporation, Maynard, Mass. (617) 897-5111, Ext.4089. European headquarters: 81 route de l'Aire, 1211 Geneva 26. Tel: 42 79 50. Digital Equipment of Canada, Ltd.

Please send me your free book "50,000 Computers Saving Managers Millions" to show me how I can get more work done at lower cost.
My computer application is _____.

Name _____ Title _____

Company _____ Address _____

City _____ State _____ Zip _____

digital

50,000 computers saving managers millions.

the grain produced was fed to livestock. Even though the grains other than wheat were available at prices of two or three cents per pound, only small amounts of them became part of the human food supply. Human dietary preferences in the U.S. do not run much to corn, sorghum, barley, oats and rye. As a result such grains have traditionally gone mainly into feeding livestock and into export markets.

Those grains were available in large amounts as livestock feed because of the tremendous agricultural production capacity that developed in the U.S. after World War II. Because no alternative market materialized, the grains (principally corn and sorghum) were available at exceptionally low prices for feeding to livestock. That situation prevailed until two or three years ago, when the demand for exported grain rose rapidly and drove prices up to five to seven cents per pound. In all likelihood the demand for grain for export will continue to rise, so that less grain will be available for ruminant livestock, and the prices will be higher. The dependence of such livestock on forages will increase accordingly. It can be expected that farmers will be under pressure to grow more forage, to produce forage of better quality and to utilize forage crops more efficiently.

Ruminant animals, which contribute beef, dairy products, lamb and wool to the economy, perform a distinctive service by converting plants that people either cannot digest (forages) or do not want (feed grains and by-products) into highly prized products of considerable economic value. Their benefactions extend far beyond the boundaries of the U.S., because much more of the world's land area is in grass than is in crop production. Cellulose is the world's most abundant organic compound. It can only be converted into food on an economic basis by ruminant livestock.

What if, as seems likely, the amount of grain available for feeding livestock continues to decrease? Can supplies of beef and dairy products be maintained? Ruminants, unlike other livestock, can be fed on forage alone if grain is completely unavailable. Supplies of beef and dairy products, however, could not then be kept at the present level unless the amount of digestible energy produced by forages was considerably increased.

Since forages provide more than half of the feed units for all livestock, and since about half of the food nutrients consumed by humans in the U.S. are of animal origin, one can conclude that from 25 to 30 percent of the food supply

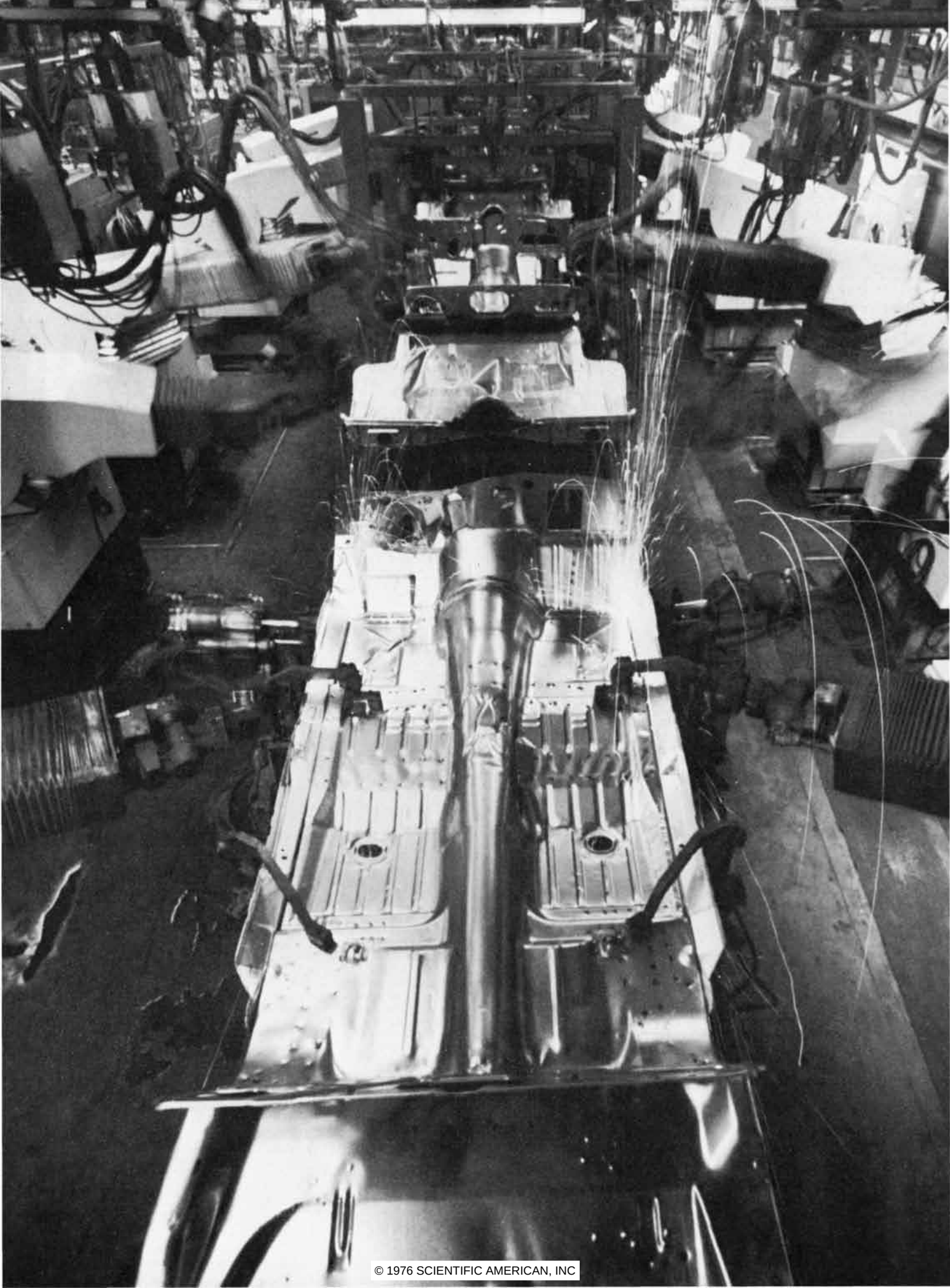
of the typical American is based on forage. Yet forages receive only about 4 percent of the nation's agricultural research funds.

The U.S. now exports more than half of its wheat and soybean crops and increasing amounts of other grains. International agreements now under discussion seem likely to result in the export of still larger quantities of grain, so that decreasing amounts will be available as feed for livestock. The amount of tax-supported research on grains is substantial; indeed, research on some of them has been increased at the same time that research on forages and livestock has been reduced. In effect research on exported commodities amounts to a public funding of research for the food supply of other countries.

Such commodities are important in the nation's agricultural economy and in earning income that helps to balance foreign-trade accounts. They merit the research support they receive. In view of the U.S. population's dependence on foods of animal origin and the increasing dependence of ruminant animals on forages, however, it appears evident that greatly increased support of research on forages and ruminant livestock would be to the direct benefit and interest of most Americans. The nation should be devoting a larger proportion of its resources to research on its own food supply.

It is often suggested that it would be a wise course to eliminate livestock and consume grains directly, thus shortening the food chain. To replace the nutrients of animal origin that the population consumes, however, would require a substantial part of the grain crop. For example, to obtain the minimum daily protein requirement from corn (the principal feed grain) everyone would have to eat about a kilogram (2.2 pounds) of corn per day. The caloric intake would be high. Moreover, the quality of the protein would be poorer than it is now, so that diets would have to be supplemented with minerals and vitamins.

Such diets are not very appealing. Furthermore, the practice would mean that much of the land that is not suited to crop production would not be contributing to the human food supply. There would also be severe and expensive problems of disposing of the by-products resulting from the conversion to human food of the grains that are now fed to livestock. The elimination of livestock in favor of a vegetarian diet appears on close examination to be impractical and unappetizing. It seems unlikely to happen.



ROBOT SYSTEMS

Instructable machines capable of performing simple human tasks are appearing in many factories. Some experimental robots are being taught how to cope with uncertainties in the environment

by James S. Albus and John M. Evans, Jr.

Man's fascination with machines that move under their own power and control is at least as old as recorded history. In Aristotle's Greece plays of several acts are said to have been performed entirely by automatic puppets driven by weights hung on twisted cords. Much later European royalties were enthralled by lifelike automata that could write, draw and play musical instruments. In recent years most of the magical aura surrounding mechanical automata has been dispelled. Today automatic machines and industrial robots are used in factories throughout the world to perform tasks that are too hazardous, too onerous, too boring or simply too uneconomic for human beings to undertake.

The use of automatic-control mechanisms in industry is, of course, not new. It is nearly 200 years since James Watt invented the flyball governor to regulate the speed of his steam engines. That was the first application of the feedback principle to sense and correct deviations from a desired control setting. During the 1930's and 1940's petroleum refineries and petrochemical plants were extensively "robotized" by inserting rather simple analogue control instruments in the feedback loops that regulated the pressure, temperature and flow rates in distillation columns, catalytic crackers and other equipment designed to process continuously flowing materials. During World War II the Servomechanisms Laboratory at the Massachusetts Institute of Technology developed the mathematical basis of the feedback control needed for the high-speed aiming of artillery using radar tracking. Beginning in the late 1940's automatic manufacturing techniques were extended to the metalworking and metal-handling industries with the aid of electronic technology developed during World War II. The trend was accelerated by the development of the digital computer. In the 1950's machine tools controlled by numbers punched on paper tape became available, and in the 1960's the first industrial robots entered the marketplace, controlled by what were essentially special-purpose electronic computers.

It is the application of electronic computers to the control of industrial machines that has transformed robots into a subject of much technological and scientific interest. The insertion of a computer into the feedback loop results in a repertory of control strategies much richer than what can be achieved with the rather simple feedback loops used previously. The capacity for high-speed computation and for decision making enables the computer to change the robot's control programs when it is appropriate and eventually will enable robots to interact with an unstructured environment with some degree of independent judgment.

At present robot technology has two major branches, one technological and the other scientific. In the development of practical industrial robots the primary criteria are reliability and cost-effectiveness. In the scientific study of robot potentialities (often conducted under the heading "artificial intelligence") the emphasis is on exploring fundamental questions of sensory perception, motor control and intelligent behavior.

In factories throughout the world there are several thousand devices loosely called robots. Most of those mechanisms exhibit few of the characteristics the average person would associate with the term robot; they are mostly "pick and place" machines that are capable of only the simplest kinds of motions. They have little or no ability to sense conditions in their environment; when they are switched on, they simply execute a preprogrammed sequence of operations. The limits of motion of each joint of the machine are fixed by mechanical stops, and each detail of movement must be guided by means of an electric or pneumatic impulse originating at a plugboard control panel.

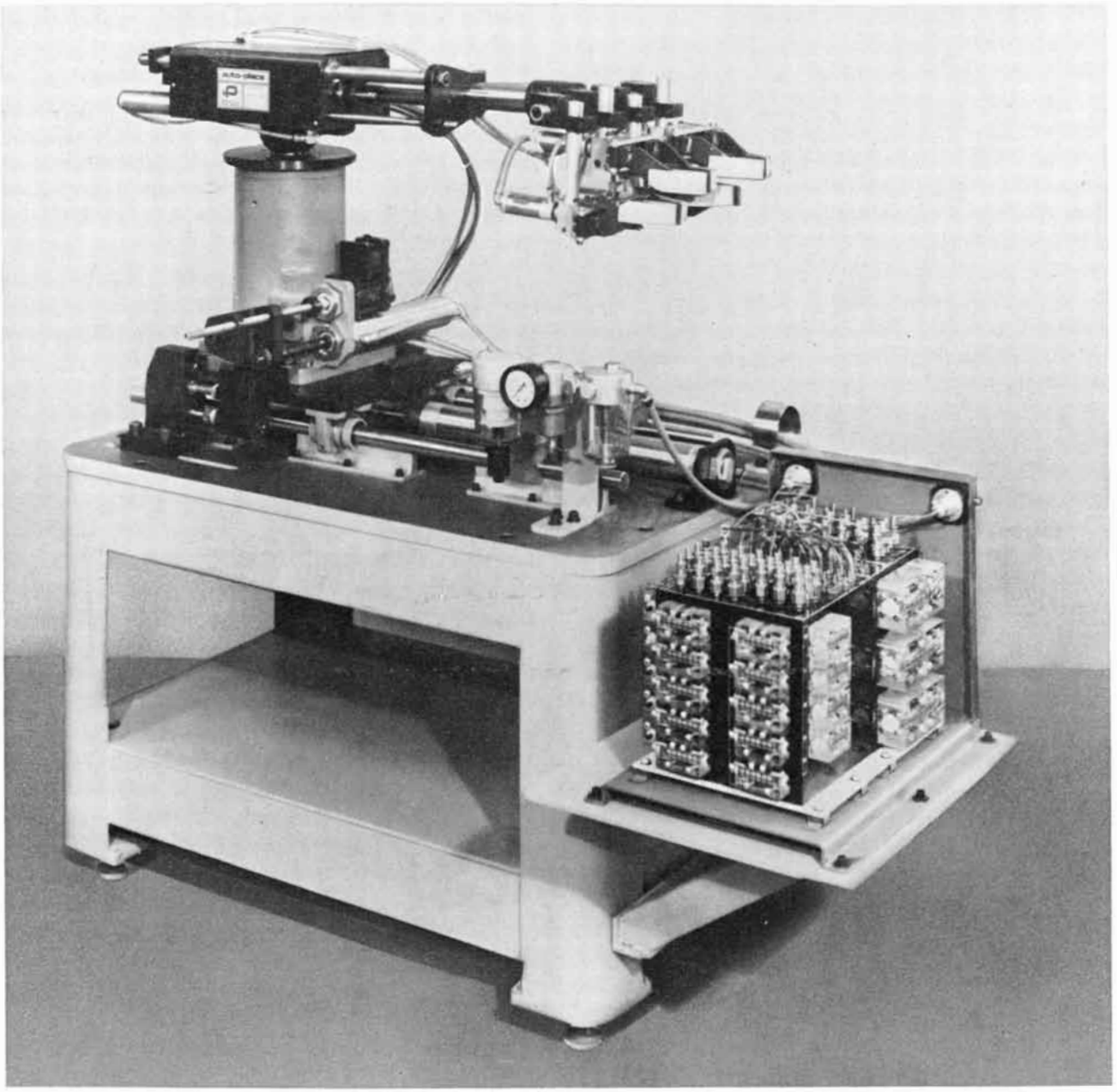
In a typical application each row of the plugboard represents one degree of freedom in a specified robot joint and each column represents one program step. Connections are made in appropriate rows and columns to determine which joints are actuated, and in which direction, at each step. Whenever a new program is needed, the mechanical stops can be repositioned and the programming connections can be relocated to set

SPOT-WELDING ROBOTS on the opposite page are used in assembling the underbodies of Chevrolet Novas, Pontiac Venturas and Buick Skylarks at the General Motors plant in Tarrytown, N.Y. Ten robots known as Unimates (not all visible in this photograph) weld together three underbody components: the dash panel, the floor-pan assembly and the rear-compartment panel. The three components, identical for all three car models, are mounted on an underbody shuttle and carried rear end first past five welding stations in sequence. Each station is manned by two Unimate robots. The underbody shuttle pauses for slightly less than a minute at each station as each robot makes from 17 to 30 welds. The welds are about half an inch in diameter and are spaced about two inches apart. Robots replace press welders, much larger machines with multiple welding guns that had to be extensively modified each time the design of the underbody was changed. Now when the design changes, the robots are simply "retrained," as is shown in bottom illustration on page 79.

up a new series of movements. A more sophisticated level of control can be achieved by adding servomechanisms that can command the position of each degree of freedom to assume any value. The addition of servo control calls for feedback from sensors such as potentiometers, encoders and resolvers, which measure the position of each joint. The measured positions are compared with commanded positions, and any differences are corrected by signals sent to the appropriate joint actuators.

There are several ways to program a servo-controlled manipulator. The simplest conceptually is a potentiometer board, similar to the plugboard system. The positions of the joints at each step are determined by values set on the potentiometers rather than by mechanical stops.

The capability of robots actuated by servomechanisms is greatly enhanced by the addition of an electronic memory and digital control circuitry. The most popular type of servo-controlled robot employed in U.S. industry uses digital servomechanisms and has a plated-wire memory for storing programs. The robot is programmed by leading it through the desired sequence of positions. A human tutor uses a hand-held control box with a rate-control button for each of the robot's joints. Working with these buttons, the tutor guides the robot to the desired position for each program step. By pushing a button the tutor can record the position of each joint. When the program is played back, the control system simply

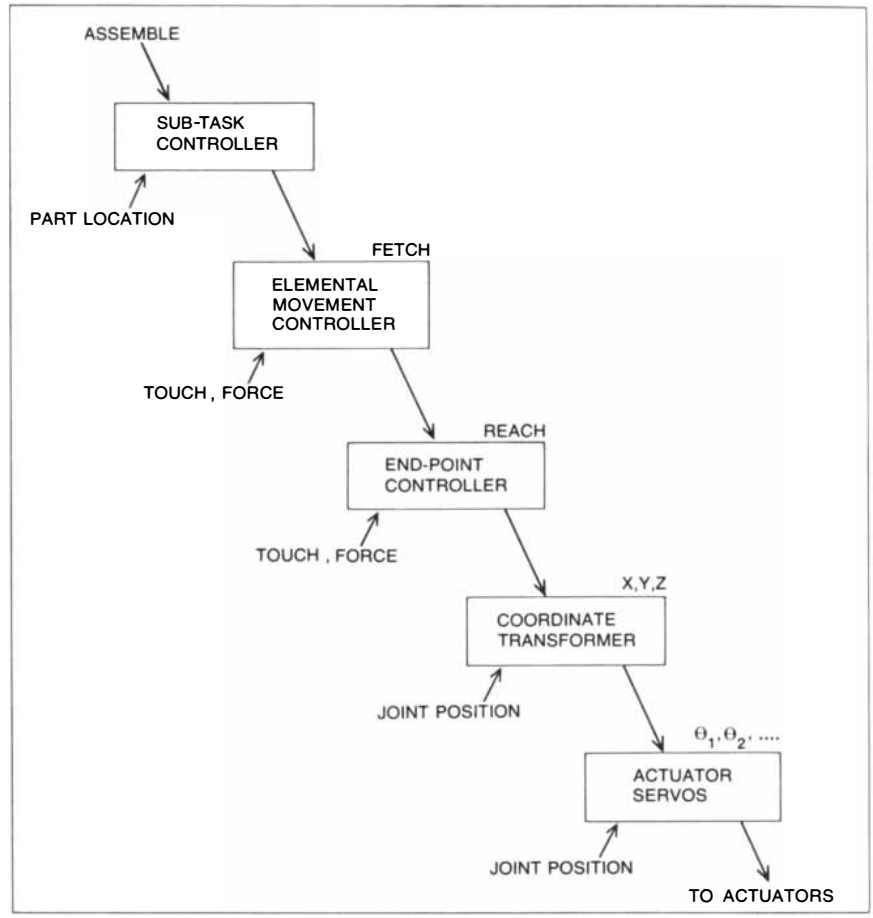


PICK-AND-PLACE MANIPULATORS are the simplest of present-day robots. This model, made by Auto-Place Incorporated of Troy, Mich., is powered by six double-acting air cylinders that enable the robot to lift, reach, rotate and grip objects and also to turn them over. The sequence of operations is programmed by means of a sequencing module and is activated by a power module, both of which are located at lower right. The unit shown can lift objects weighing up to 318 pounds and exert a grip force of 227 pounds.

commands each joint to move to the position recorded for each step. Once the robot goes into operation on the production line it repeats the recorded program over and over again, moving from one step to the next according to a fixed timing cycle, either on completion of the last step or in response to an interlock signal from external machinery.

An electronic memory enables a robot to store several programs and to select one or another depending on different input commands or on feedback from external sensors. For example, robots that spot-weld automobile bodies can be programmed to handle a variety of automobile models intermixed on an assembly line.

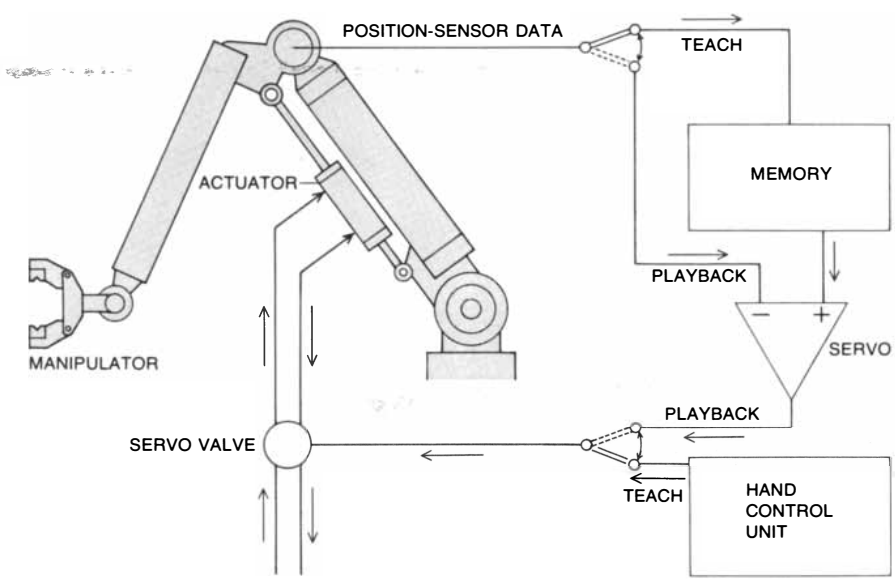
Most robots have memories that hold only a few hundred program steps. The motions involved in each step are large and jerky. If a smooth, continuous motion is required, a magnetic tape capable of storing many thousands of closely spaced steps can be added to the control system. Continuous-trajectory programs are called into play when one wants to have the robot's "hand" move smoothly through space, for example in arc welding or in spraying paint or glass fibers.



HIERARCHY OF CONTROL LEVELS provides an effective way of organizing the flow of instructions and feedback information a robot needs for a given task. At each level in the hierarchy commands from the level immediately above are broken down into simpler commands that are sent to the level directly below. Feedback information, such as force and joint position, can modify commands issued at appropriate levels. At the very lowest level voltage drive signals are sent to motors or valves that operate joints of manipulator.

The addition of a computer with arithmetical and logic capacities to the control system of a robot makes it possible to implement still higher levels of control. With a computer it is possible to program the robot to move its hand along straight lines or in other geometric paths, using an external (rather than an internal) coordinate system. That system has many advantages when the robot is to place objects in regular arrays on a pallet or is to interact with objects traveling on a conveyor belt or an assembly line. The computer makes it possible for the mathematical definition of the desired path in work-space coordinates to be transformed into appropriate drive signals to the servomechanisms for the various joints of the robot's arms.

One computer program for performing such coordinate transformations (known as resolved-motion-rate control) has been developed by Daniel E. Whitney of the Charles Stark Draper Laboratory. Whitney uses a coordinate system fixed with respect to the hand of the robot. Input commands can be given in terms of hand motions, such as "Move y units along the axis of the hand" or "Move x units to the right." A recently developed industrial robot that has a similar capacity for handling work-space coordinates carries out the necessary



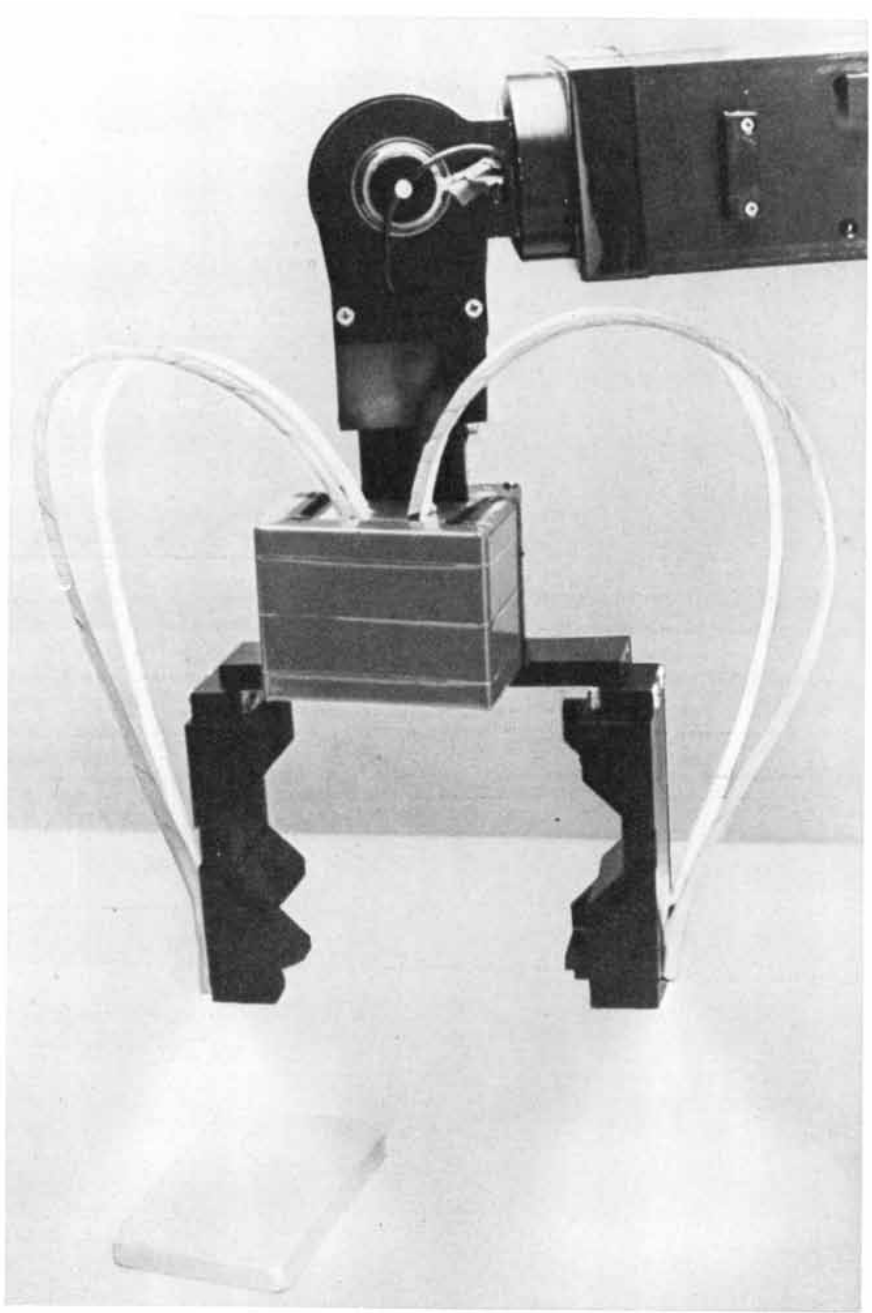
PROGRAMMING AN INDUSTRIAL ROBOT is frequently done by using a hand controller to guide the robot through a sequence of operations. The successive positions of all the robot's joints are stored in a memory. By switching from the "teach" to the "playback" mode stored positions are repeated. Robot is now ready to take its place on production line.

transformations in real time during program execution. The computer control system for this robot also calculates acceleration and deceleration profiles that maximize the speed with which heavy objects can be manipulated.

The great majority of industrial robots have only the most elementary manipulatory capacities. Feedback is limited to information about joint position, combined with a few interlock and timing signals. Hence most robots can function only in environments where the objects to be manipulated are always precisely located in the proper position for the robot to grasp.

For many industrial applications this level of performance is quite adequate for accomplishing a wide variety of tasks. For example, engineers at the Kawasaki laboratories in Japan have shown that robots can put together complex assemblies of motors and gearboxes with no more than high-precision positional feedback, cleverly designed grippers and fixtures for holding parts that yield to a certain extent when the parts are brought together. Other experiments have shown that a small amount of vibration or jiggling, together with properly designed tapers and bevels, can accommodate for slight misalignments and prevent jamming when two pieces with close tolerances are assembled.

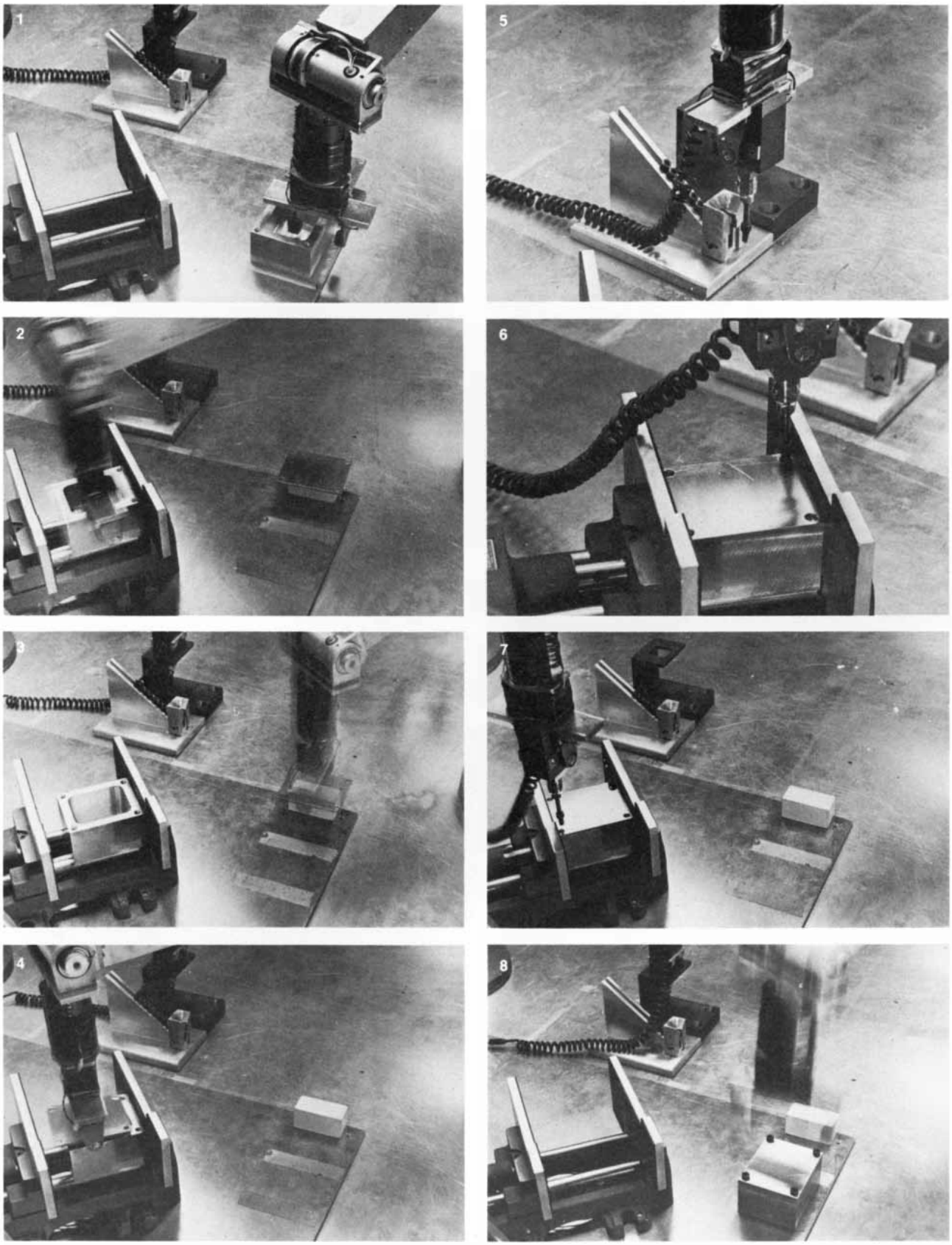
In many other situations it has turned out to be extremely expensive to provide the work environment with the constraints needed for precise robot operation. For example, it is reported that the transfer equipment needed to hold automobile bodies within the positional accuracy required for robot spot welding costs as much as 10 times more than the robot welders themselves. Accordingly a number of government, university and industrial laboratories are studying new feedback methods that could compensate for uncertainties in the location of workpieces.



PROXIMITY SENSOR affixed to a robot's gripper enables a robot to locate and grasp objects that are not precisely positioned. Infrared radiation, produced by light-emitting diodes and carried by fiber-optic bundles, is projected downward in two beams. Radiation reflected from the target enters a parallel set of fiber-optic bundles. Strength of reflected radiation acts as a feedback signal that informs robot how close gripper is to its target. Experimental system in photograph is under development at the National Bureau of Standards.

In order to provide a unifying framework for discussing robot technology the control problem can be viewed as a hierarchy of subproblems. Each level in the control hierarchy accepts commands from the next higher level and responds by issuing ordered sequences of commands to the next lower level, making use of sensory feedback to close control loops where they are appropriate. For example, assume that the goal is to control a manipulator in a complex task such as the assembly of a small gasoline engine. The task can be broken down into sequences of simpler tasks, such as fetch parts *a* and *b*, insert *a* into *b* while nullifying off-axis forces, fasten part *c* to assembly *ab* and so on. Each of these tasks can be broken down further into sequences of elemental movements such as reach, grasp, follow specified trajectory and the like. Those elemental movements can themselves be dissected into sequences of manipulator positions expressed in *x*, *y*, *z* coordinates and gripper orientations. Finally, manipulator positions can be transformed into coordinated sequences of voltages that drive the joint-actuator servomechanisms.

In general there are two types of input to each level in the hierarchy. First there are the input commands from a higher level. At the very top the commands may come from a human operator or from some other external goal-select-

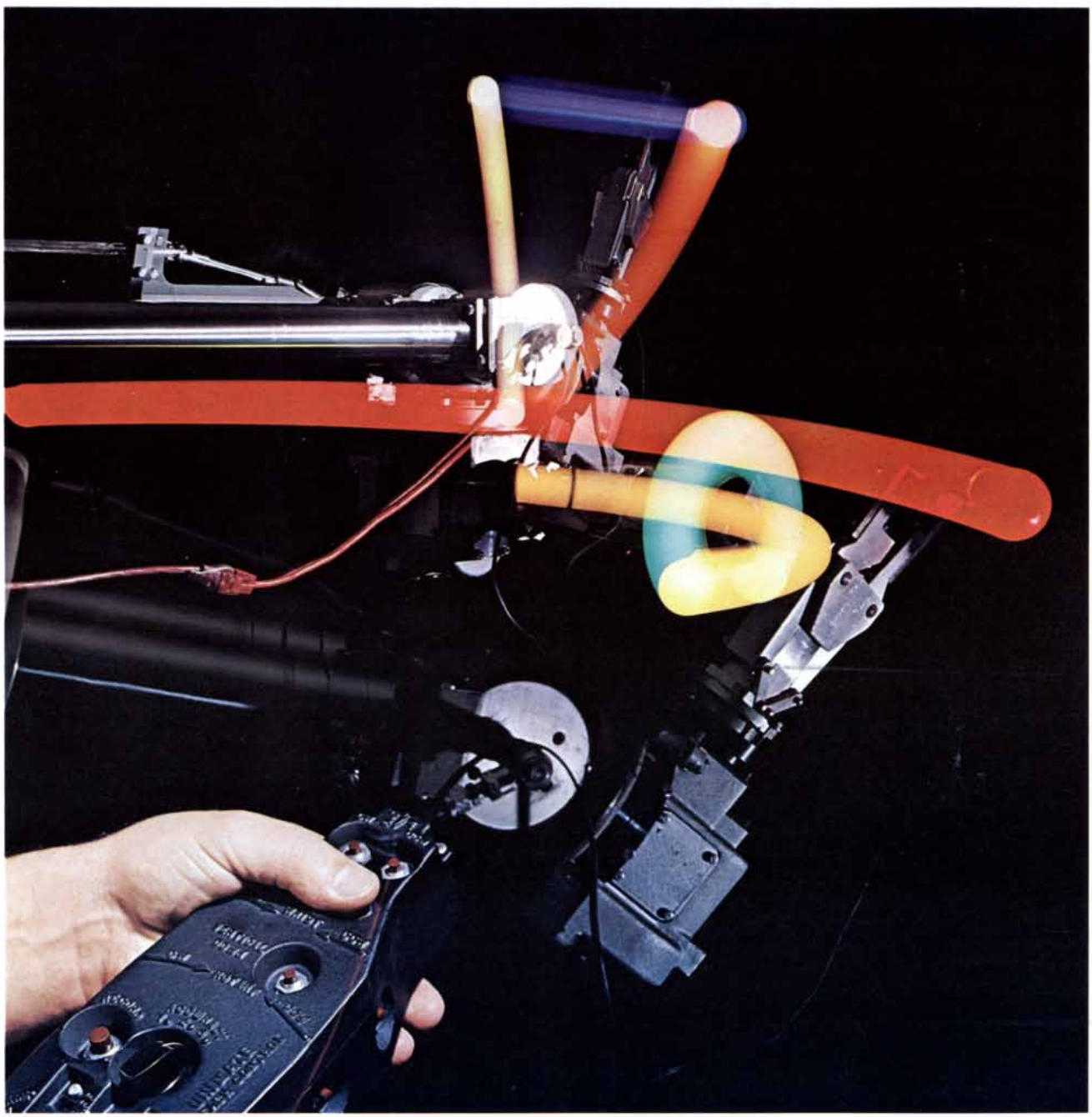


ROBOT ASSEMBLY OPERATIONS are being investigated in the Artificial Intelligence Laboratory at Stanford University. In the demonstration shown in these eight photographs the robot first locates a steel box (*1*) and then places it in a vise (*2*). The robot returns for a cover (*3*) and places it on top of the box (*4*). The vise then closes. The robot now provides itself with a special socket wrench and picks up a bolt from a dispenser (*5*). It locates one of the holes in cover of box and screws in the bolt (*6*). Picking up a second bolt, it repeats the operation (*7*). After all four bolts have been inserted the robot removes box and places it on the table (*8*).

ing mechanism. At all other levels the commands descend as outputs from the next higher level in the hierarchy.

The second class of inputs consists of feedback signals indicating the state of the manipulator and the environment. Such inputs may describe the position and motion of the joints or convey information from sensors of force, touch or proximity on the grippers. In addition feedback data may convey timing signals from external machinery with which the robot must interact. The feedback data may also describe the location of obstacles, the shape and location of workpieces and the motion of reference coordinates or provide other information pertinent to the robot's manipulation task. Particularly at the higher levels of control such input information may require complex processing of optical, acoustical, tactile or other signals of the kind normally perceived by human senses.

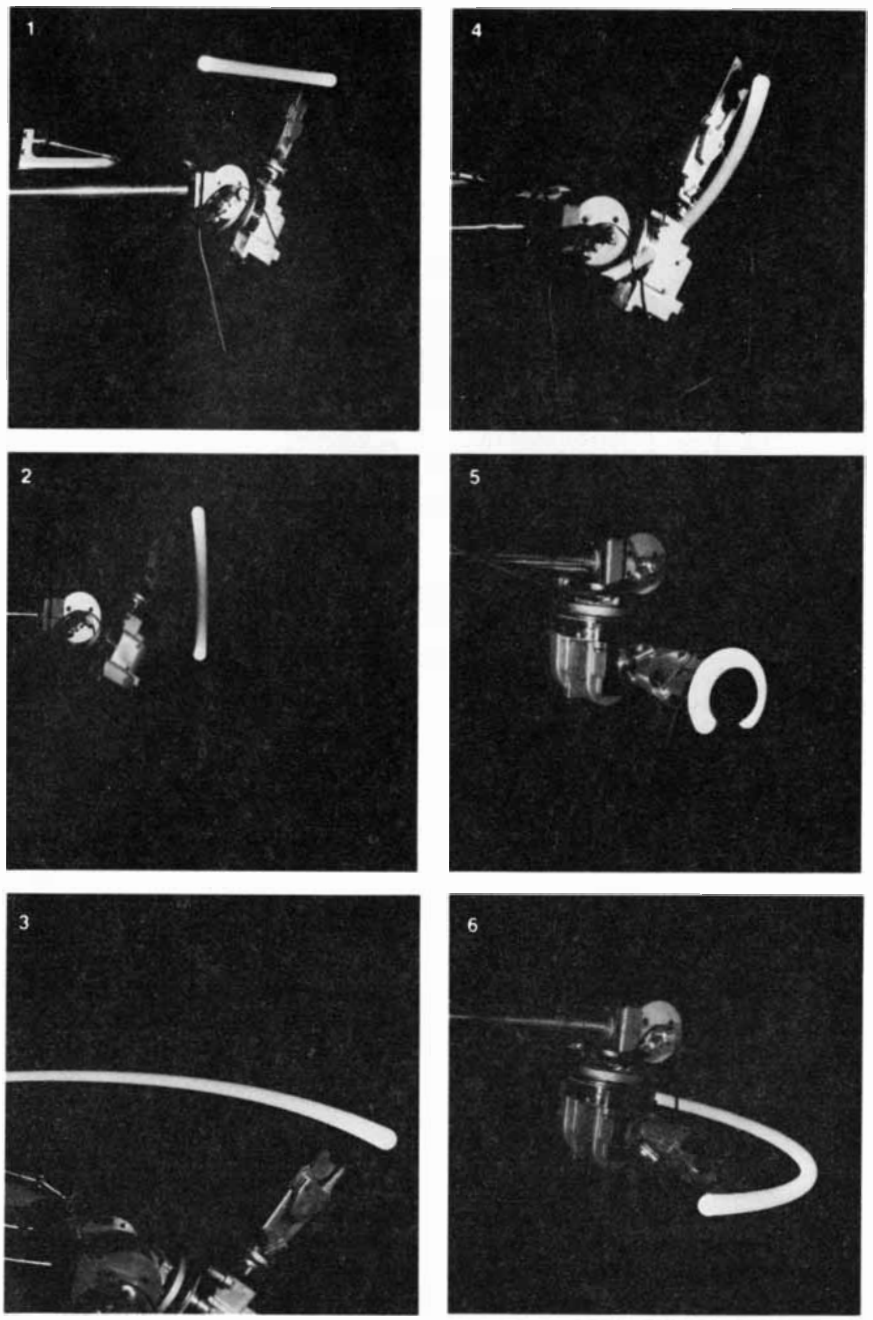
When viewed from a hierarchical perspective, the control of a robot in the performance of a complex task requires that each control level be able to transform commands from a higher level into sequences of commands to the next lower level. The technical difficulty of accomplishing such transformations arises largely from the volume and complexity of the feedback data that must be handled in the computations. The type of feedback in turn depends strongly on the



UNIMATE ROBOT, made by Unimation, Inc., of Danbury, Conn., is capable of executing six basic motions that can be programmed separately by means of a hand controller. In this multiple-exposure photograph the robot's gripper, or hand, is holding lights of different colors to distinguish the six motions. Pistonlike arm is capable of three motions: in-out (*blue*), up-down (*white*) and left-right (*red*). At any position of robot's arm, gripper can execute three more motions: bend (*orange*), swivel (*green*) and yaw (*yellow*).

degree of uncertainty encountered in the environment.

If misalignments are limited to a few millimeters in position or a few degrees in orientation, one can provide the robot with sensors of force, touch or proximity to supply the necessary corrective information. Forces are sometimes measured by detecting motor currents or differential hydraulic pressures in the appropriate joint actuators. In other designs force sensors consist of strain gauges mounted on the robot's "wrist." Alternatively, force sensors can be mounted on the worktable that holds the workpiece as the robot inserts pins and bearings, mounts gaskets and tightens screws. Forces transmitted through the workpiece to the table can be used to control such assembly operations. Most touch sensors are simply microswitches that transmit a signal when two surfaces make contact.



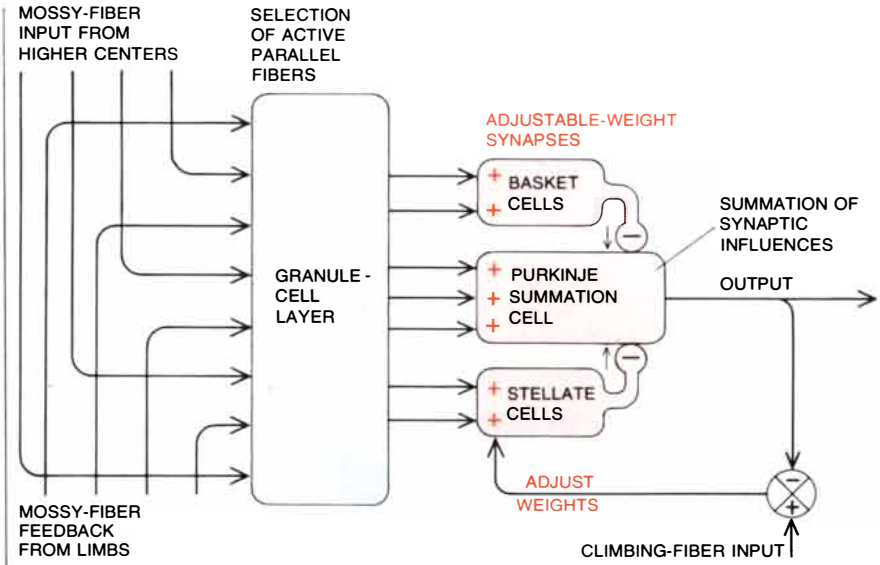
SIX MOTIONS OF UNIMATE ROBOT are illustrated separately. The photographs in the column at left show the three basic arm motions: in-out (*1*), up-down (*2*) and left-right (*3*). In the column at right are the three gripper motions: bend (*4*), swivel (*5*) and yaw (*6*). The pictures do not show the full distance range of the various motions other than swivel.

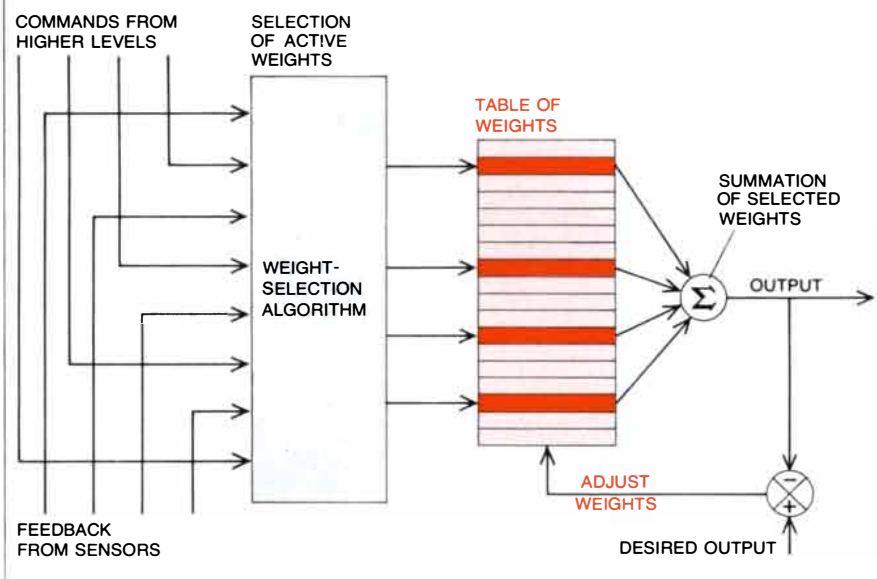
Proximity sensors are usually more complicated. Typical systems employ a laser beam or a beam of infrared radiation and determine proximity by measuring the energy reflected by the target. For distances of more than a meter the range can be computed by measuring the transit time of a laser pulse. A robot equipped with well-designed proximity sensors can search for objects of uncertain position and move toward them at high speed without danger of collision. Since sensors of force, touch and proximity supply information in a coordinate system fixed in the sensory device, the information must enter the control hierarchy at or above the level where commands can be expressed in sensor coordinates.

There are various strategies by which sensory feedback data can be used at different levels in the control hierarchy to compensate for uncertainty in the environment. One method is to determine the exact position of an object and then translate, rotate or stretch previously recorded trajectories to compensate for the measured positional errors. Another method is to introduce the feedback data into the mathematical equations used to compute trajectories for the manipulator end point. The second technique has been used by Whitney and James L. Nevins at the Draper Laboratory to enable a manipulator with force feedback to find surfaces, follow edges and insert pins into holes with close tolerances. Feedback signals can also be harnessed to implement conditional branches in control programs. This capacity has been demonstrated in the Artificial Intelligence Laboratory at Stanford University in the robot assembly of a water pump.

Another technique for incorporating feedback into the control hierarchy is the "cerebellar model articulation controller" (CMAC) being developed in the laboratory of one of us (Albus) at the National Bureau of Standards. CMAC is an adaptive system that uses a distributed storage system to store or learn the correct output for each relevant combination of inputs. Like the cerebellum, after which it is modeled, CMAC can cope with a large number of command and feedback variables and exhibits many of the characteristics of a conditioned reflex. It is potentially useful at various levels in the control hierarchy.

If there are large uncertainties in the robot's world, much more complex feedback information and control techniques are needed. For example, if the robot is to manipulate objects that are





CEREBELLAR MODEL ARTICULATION CONTROLLER (CMAC) undertakes to mimic the information-processing scheme used by the cerebellum, shown schematically in the upper diagram. The cerebellum model was first developed by David Marr, who is now at the Massachusetts Institute of Technology, and was elaborated by one of the authors (Albus). Input commands from higher brain centers and feedback signals from the limbs are carried by mossy fibers to the granule-cell layer, which acts as an address decoder. Each input is an address that is uniquely decoded into a selected set of parallel fibers. Impulses carried by the parallel fibers, after traveling through various synaptic pathways, give rise to an output from Purkinje cells. Each output can be viewed as the contents of a mossy-fiber address. The value of the contents depends on the strength of the adjustable synaptic weights (*indicated by plus signs*). The weights can be adjusted to make the output of Purkinje cells match data carried by the climbing fibers. In this way it is possible for the cerebellum to store a functional relation between input and output. CMAC, the computerized verison of this model, is depicted in the lower diagram. Each input selects a unique set of weights that are added up to obtain an output. The selection process operates in such a way that if two inputs are similar, they select partially overlapping sets of weights. Thus whenever the weights selected by an input address are adjusted so that they change the output, some of the neighboring outputs are also changed because of the overlap in the weights that are selected. The overlap enables CMAC to generalize among similar inputs in much the same way that human beings and other animals generalize among similar experiences. Some examples of the CMAC system's operation are illustrated on page 86.

randomly positioned or heaped together in a pile, simple force, touch and proximity information are usually inadequate. In extreme cases the robot may be required to move at least one obstacle to reach the object it is seeking. It may even be required to select tools with which to achieve its goals. Such maneuvers call for a robot endowed with a capacity for cognitive planning.

Efforts to equip a robot with modest intellectual capacities have occupied workers in the field of artificial intelligence for many years. In most instances the endeavor has required that the upper levels of control incorporate some internal data structure—a "world model," or knowledge frame—that can represent the state of the environment in a meaningful way. The robot must also be endowed with an adequate repertory of sensors and data-processing circuits for analyzing the environment so that the robot can keep its internal knowledge frame up to date.

When a robot is faced with an input command to be executed, it tries to set up a hypothetical desired world model. It then attempts to devise a set of procedures that can convert the existing world model into the desired one. Each step of these internal procedures must correspond to one of the output commands that can be executed by a lower level in the hierarchy. Once such an internal procedure is found it can be converted into the required sequence of output commands.

That technique was embodied in the robot named Shakey, built at the Stanford Research Institute. Shakey could be given a task such as finding a box of a certain size, shape and color and moving it to a designated position. The robot was able to search for the box in various rooms, to cope with obstacles and to plan a suitable course of action. In some situations the performance required the implementation of a preliminary action before the principal goal could be achieved. In one instance Shakey figured out that by moving a ramp a few feet it could climb up on a platform where the box had been placed. It is perhaps needless to add that very substantial problems are involved in endowing a robot with the resources for coping with such tasks.

One promising approach to the problem of analyzing images and scenes automatically has been developed by the Laboratory for Computer Science at M.I.T. Data from a television scanner are processed by computer programs that find edges, define surfaces, analyze shapes and set up data structures to

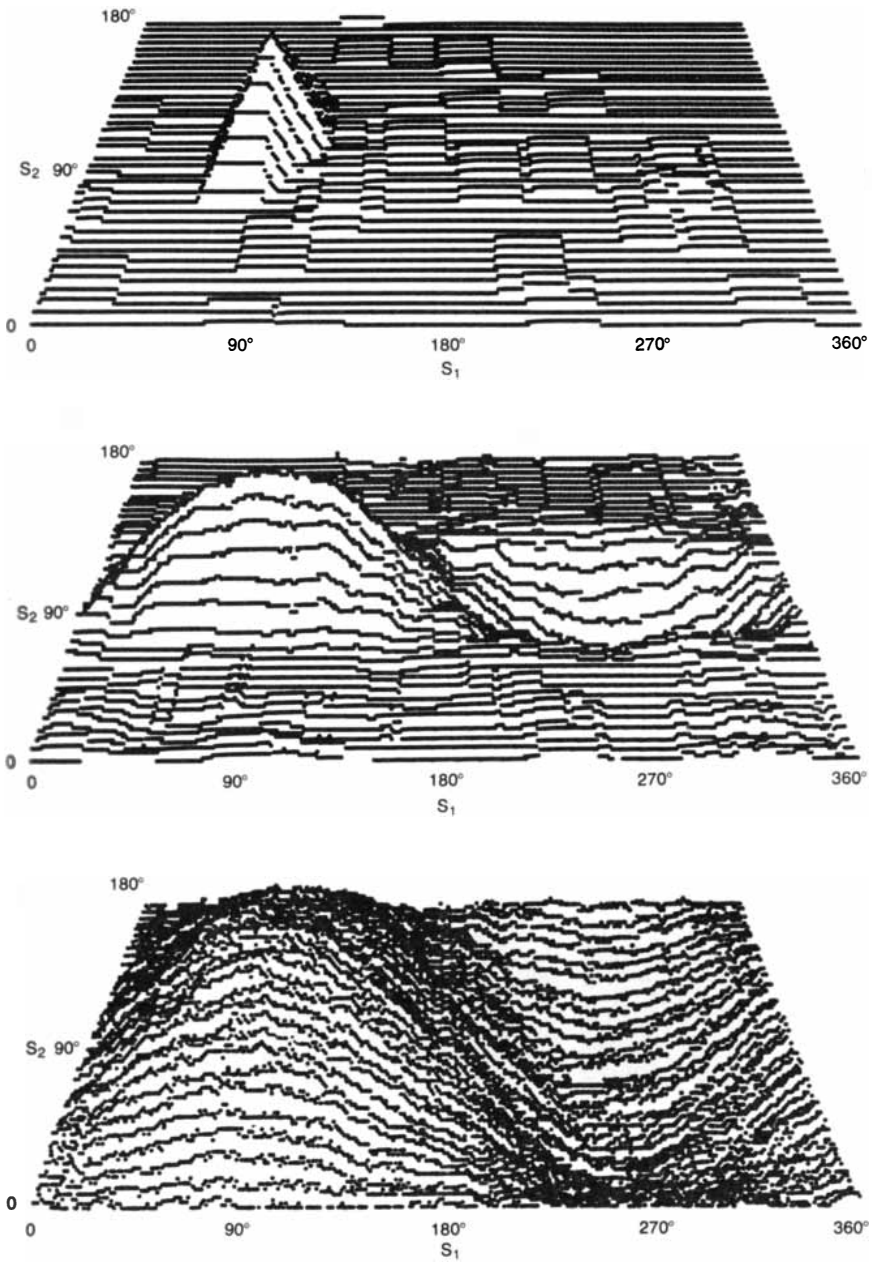
specify how all these primitive qualities fit together into a meaningful description of real objects in the viewed environment. In the most recent work in the Artificial Intelligence Laboratory at M.I.T. the scene-analysis programs are arranged in a nonhierarchical structure so that any subprogram can call any other and exchange information with it. For example, a subprogram for finding edges may reach some tentative conclusion and either report its findings directly to the program that called for it or decide that it needs the assistance of another subprogram. The dialogue includes what might be termed advice, suggestions, remarks, complaints, criticisms, answers, mistakes and conjectures.

Although such investigations have afforded much valuable information about the complex process of visual perception, it does not seem likely that the specific analytical programs developed so far will be embodied in practical robot systems for many years to come. Not only do the programs call for a great deal of computing power but also they require that the objects in the environment have simple geometric shapes and be displayed against an uncluttered background.

A quite different approach to the problem of machine vision has been taken by Thomas O. Binford and his colleagues at the Stanford Artificial Intelligence Laboratory. They illuminate objects in the environment with a narrow beam of laser light that has been spread into a plane by a cylindrical lens. Wherever the plane of light intersects the surface of an object, its contours are revealed by a thin line of light. By moving the plane to a number of different positions and orientations a composite pattern can be built up that indicates the three-dimensional form of the visible surface. The pattern is then analyzed by a computer program that attempts to infer the structure of the entire object from the surfaces that are visible to the scanning system. The method is advantageous in that it avoids many of the complications introduced by shadows and irregular shapes.

At the Jet Propulsion Laboratory of the California Institute of Technology a robot visual system is being developed for locating and picking up rocks from a simulated Martian surface. The system uses a television camera and a scene-analysis program to detect the outline of a rock. A laser range detector is then used to measure the rock's position accurately so that the manipulator can pick it up.

In an effort to devise visual-analysis procedures suitable for industrial applications Charles A. Rosen, David Nitzan and others at the Stanford Research Institute have devised a scanning system consisting of a line of photodiodes mounted above a conveyor belt. As objects pass by in single file—randomly oriented and not necessarily identical—the photodiodes build up an outline image of each. Fairly simple visual-analysis programs determine the outline of the object, measure its position and orientation and conduct a simple inspection for defects. If the object is to be rejected or



CHANGING CONTENTS OF CMAC MEMORY are shown in three computer printouts. They demonstrate how the CMAC memory generalizes among similar inputs. The top pictogram shows how the memory was changed by a single entry, +1, at the input address $S_1 = 90$, $S_2 = 90$. The entry changed the memory not only at the address selected but also at neighboring addresses. As can be seen, the effect on neighboring addresses falls off in proportion to the distance from the selected address. The middle pictogram shows the result of entering a sine wave consisting of only 16 entries spaced along the axis $S_2 = 90$. Again as a result of generalization addresses adjacent to the axis are affected. The bottom pictogram shows the contents of the CMAC memory when the function is the product of two sine waves, one along each axis. As a result of generalization only 175 selected entries were required to store this function over the entire space of 360 × 180, or 64,800, address locations. Illustrations show how any bounded continuous function can be stored in memory, and thus made available as CMAC output, by entering data at relatively few points.

It comes as close to speaking for itself as any car ever built.

Aggressive, unmistakable profile. Strong, mellow, no-nonsense sound. 280-Z, by Datsun.

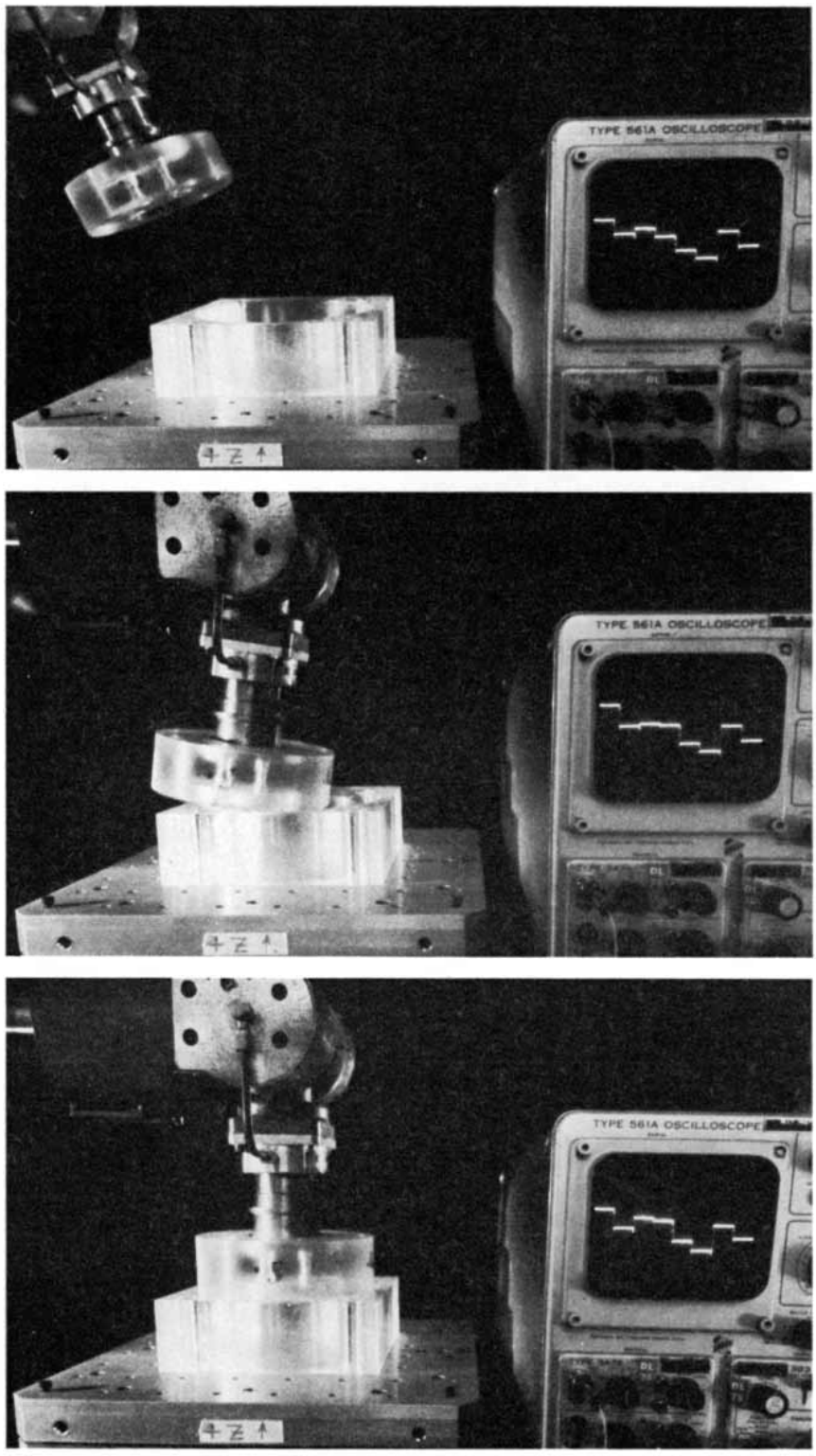
Its particulars bear out its promise. Under the sculpted hood, an efficient 2800cc overhead cam engine with computerized fuel injection. The 6-cylinder plant has an 8.3:1 compression ratio for optimum performance. The transmission is an all-synchro manual 4-speed; an automatic is available.

Perfection. Fuel injection.

Additional GT requisites include rack and pinion steering. Power assisted front disc brakes. 195/70 HR 14 steel belted radials. Reclining bucket seats. AM/FM radio with power antenna and electric rear window defogger standard, of course.

The fuel injected Datsun 280-Z. 2-seater or 4-seater. Commune with one.

The fuel injected
Datsun 280-Z



FORCE-SENSING PLATFORM can be used to guide a robot in an assembly task, such as the insertion of a cylinder into a close-fitting opening. The technique has been developed at the Charles Stark Draper Laboratory, where these pictures were made. The picture at top shows the robot starting its first attempt at an assembly task. The steplike image on the oscilloscope screen represents the forces acting on eight pressure sensors (strain gauges) under the square block on the worktable. Each step in the display corresponds to the force acting on one sensor. The sensors respond to torques as well as to simple directional forces. When the robot tries to fit the cylinder into the block and misses (*middle picture*), the sensors detect exactly where and how the cylinder is pressing against the sides of hole. The oscilloscope display changes accordingly from its neutral, or no-force, configuration. A computer program analyzes the misalignment and calculates a path for the robot arm so that the cylinder enters the hole smoothly on the next try (*bottom photograph*).

worked on, commands are issued to a manipulator that picks up the moving part and presents it in the proper orientation for further processing.

It will undoubtedly be several decades before robots can simulate the elementary processes of intelligent behavior or even the sensory information processing needed for visual comprehension. It will also be some time before robots are capable of inventing strategies to deal with novel situations with anything like the facility of a two-year-old child. Visual comprehension and strategy invention are extremely complex processes, and they may not be understood until the basic functioning of the human brain is far better understood than it is today.

The ability of a robot to perform useful tasks in the highly structured environment of a modern factory, however, is quite another matter. In a factory there is no need for a robot to be creative or to have any significant degree of insight. The primary criteria for an industrial robot are that it be inexpensive, reliable and easily programmed to execute a well-defined sequence of operations. Practical methods are available to enable a robot to compensate for normal dimensional variations and small misalignments of parts. In the constrained environment of industrial manufacturing top-level input commands can be quite stereotyped, similar to the communications between a computer and conventional peripheral devices such as a disk memory. Hence it may be possible to generate high-level command sequences with compilers and other well-known programming techniques. Several programming languages have already been developed for the purpose of programming and operating numerically controlled machine tools. Recently several engineering groups have begun to create similar languages for robot systems.

The developments we have described here suggest that there are no fundamental technical barriers to achieving highly automatic robot factories in the near future. In Japan a major program involving government, industry and the universities is already committed to the development of a prototype unmanned plant incorporating robot technology for the manufacture of machine-tool components. The plant is scheduled to go into operation in the early 1980's, and if it is successful, it will serve as a model for other manufacturing missions. It therefore seems quite likely that highly automatic and even partially self-reproducing robot factories will be in operation in at least some countries within 15 years.

When you sponsor a child...you deserve to know how your money will be spent.

CHRISTIAN CHILDREN'S FUND, INC.
Richmond, Virginia 23261 · Phone (804) 644-4554 · Cable: Childfund

Assisting children around the world

Dear Friend:

We at Christian Children's Fund believe in full financial disclosure, that when you sponsor a child you should be told how your money is spent—as this ad attests. Our annual report, containing a review of CCF programs and Statement of Income and Expense, is sent to each sponsor and to anyone requesting a copy.

We try to keep expenses as low as we can so that as much of your contribution as possible can be used to assist children. About 80% of CCF's total income goes to child care services.

We try to get a sponsor for each child and we ask sponsors to make regular monthly payments. Your contribution is pooled with that from other sponsors who help children in the same programs and is used to support all youngsters, including children in projects who may be temporarily unsponsored.

If you wish to send a special gift to the child you sponsor, the entire amount is forwarded. We ask that such gifts be limited to $10 at one time unless you wish the money to be used for a special need of the child.

Since 1938, CCF has assisted hundreds of thousands of needy children. Yet, today there are nearly 20,000 youngsters in our projects at home and abroad who are unsponsored and urgently need your help.

I know these are hard times for many people. But won't you think about the children?

You can sponsor a child for $15 a month. Please fill out the coupon, and in about two weeks you will receive information about the project and the child who needs your love.

It's infinitely worthwhile.

Sincerely yours,

Verent J. Mills
Executive Director

Venita has suffered a lot. Her mother died, her father was unable to care for her. A children's Home in India, affiliated with CCF, takes care of Venita now. But what about other children who need help?

Write today: Dr. Verent J. Mills
CHRISTIAN CHILDREN'S FUND, Inc.
Box 26511, Richmond, Va. 23261

I want to sponsor a ☐ boy ☐ girl in (Country)_____
☐ Choose any child who needs my help. I will pay $15 a month. I enclose first payment of $_____. Please send me child's name, mailing address and photograph.
I can't sponsor a child now but I do want to give $_____.
☐ Please send me more information.

Name_____
Address_____
City_____
State_____ Zip_____

Member of International Union for Child Welfare, Geneva. Gifts are tax deductible. Canadians: Write 1407 Yonge, Toronto 7.

SS 0520

Christian Children's Fund, Inc. Statement of Income and Expense

For the Fiscal Year Ended June 30, 1975

INCOME:	Percentage of Total Income	Amount
Public Support:		
Sponsorship contributions	80.5%	$23,229,680
Other contributions	9.8	2,815,910
Bequests	1.4	403,971
Total Public Support	91.7%	$26,449,561
Other Support:		
Reimbursements from Christian Children's Fund of Canada	7.2	2,061,463
Other Income	1.1	330,991
TOTAL INCOME	100.0%	$28,842,015

EXPENSES:		
Program services for:		
Family support	60.0%	$17,316,977
Homeless children	9.9	2,844,841
Primary and secondary education	11.2	3,241,216
Total Program Services	81.1%	$23,403,034
Supporting Services:		
Management and general	9.1	2,627,202
Fund raising	11.5	3,319,034
TOTAL EXPENSES	101.7%	$29,349,270
(DEFICIT)—Excess of Expenses over Income	(1.7)%	$ (507,255)*

Christian Children's Fund financial reports are in accordance with the American Institute of Certified Public Accountants industry audit guide, AUDITS OF VOLUNTARY HEALTH AND WELFARE ORGANIZATIONS. Our complete audit report is available upon request.

*Note: The Unrestricted Fund deficit of $274,388 was absorbed by prior years' surplus as was the Restricted Fund deficit of $212,671 and the Land, Building and Equipment Fund deficit of $58,219. The Endowment Fund realized a surplus of $38,023.

The Final Paleolithic Settlements of the European Plain

When the glaciers retreated, the climate of Europe became variable. Tundra gave way to forest and forest to tundra. The technology and social organization of the indigenous hunters changed accordingly

by Romuald Schild

In the period between 14,000 and 10,000 years ago the great plain of northern Europe extended without interruption from the Ural Mountains to Ireland. (The English Channel and the Irish Sea did not yet exist.) The people who lived on the plain were the last hunters and gatherers of the Paleolithic period. They had managed to resettle the plain as the continental glaciers finally retreated at the end of the Ice Age. The story of the Paleolithic hunters' ultimate disappearance is a fascinating chapter of prehistory. In geological terms that chapter extends from the final millenniums of the Pleistocene epoch into the first centuries of the Holocene, the epoch in which we live. In cultural terms it records the end of the Old Stone Age.

On the western and central parts of the European plain there are at least 1,200 living sites representative of the Final Paleolithic period. Their excavation has been reported in an extensive literature, yet the archaeology of the period is little known to English-speaking readers. Almost all the reports have been published in German, Dutch, Polish, Lithuanian or Russian. Those who can surmount the language barrier encounter another obstacle. For more than half a century fieldwork on the European plain has followed an interdisciplinary approach: the specialist who deals with stone tools or animal remains cooperates closely with experts on landforms, soil formation and pollen analysis. Even an archaeologist whose primary interest is in this field finds it difficult to keep abreast of a scholarly literature that is just as likely to be found in journals of geology or pollen studies as in journals of prehistory.

Here I shall take up, among other things, a Final Paleolithic site on the Vistula River not far from Warsaw that has been excavated by my group from the Institute for the History of Material Culture, an institute of the Polish Academy of Sciences. The Calowanie site, as it is called, includes layered deposits left by a succession of hunting groups in the Final Paleolithic period. Analysis of the finds here and at many other sites has enabled us to sketch a tentative picture of those early European societies.

By 9000 B.C. the last major glacial period in Europe, the Würm, had drawn to a close. A shrunken ice cap still occupied the Scottish highlands, and a larger one covered most of Scandinavia. The coastlines of both the Atlantic and the North Sea lay substantially seaward of their present position, and most of what is now the Baltic Sea was a great freshwater lake. Students of early climates have named the period the Late Dryas; dryas is a flowering herb typical of tundra vegetation. The climate of the European plain had become distinctly cooler, ending an earlier warm phase known as the Alleröd interstadial.

The mild-climate vegetation, largely pine and birch forest, was being replaced by a parklike tundra that included scattered shrubs and birches. Reindeer were no longer present in southern France, but herds of them still wandered the great plain to the north, and a few true Ice Age mammals such as the mammoth and the woolly rhinoceros may still have grazed in the northern foothills of the Alps and the Carpathians. The European plain was hospitable to man during the Late Dryas in terms of both animal and plant life; the large number of Final Paleolithic sites in the region is eloquent testimony to it.

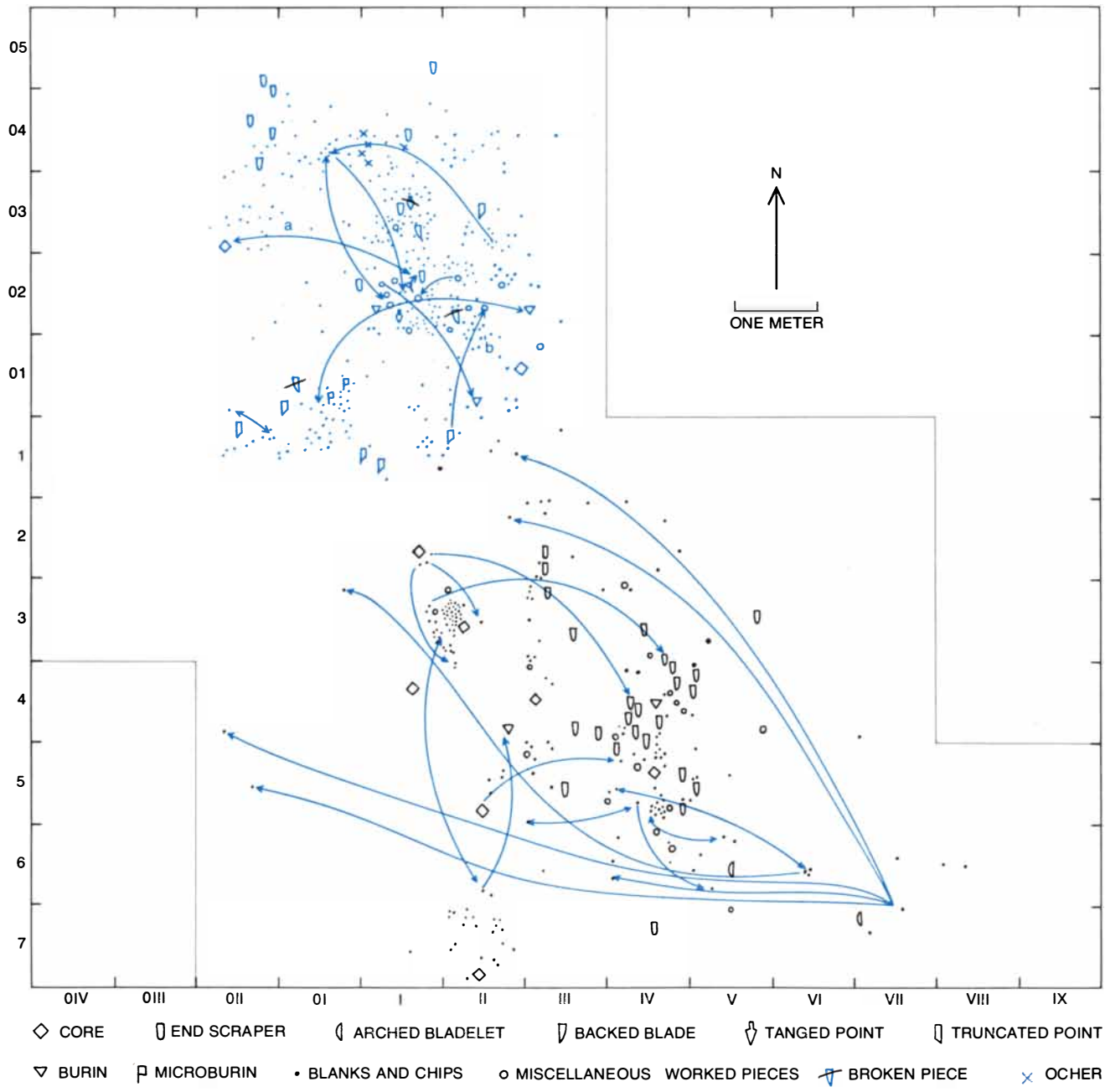
The first reasonably systematic archaeological investigations of the European plain were undertaken in the latter half of the 19th century. Most of the Final Paleolithic sites were found in former sand-dune areas or on the terraces of rivers, and they had usually been exposed to view by wind erosion. The investigators pored over the exposed surfaces, collecting the artifacts (mostly tools made of flint) and other remains that had been uncovered; the work was done by interested amateurs, who were joined by professional workers toward the end of the century. The collections they amassed were huge, but the scientific value of the finds was limited. Not until after World War I was genuinely systematic work undertaken. In Poland it was initiated by Stefan Krukowski, who excavated a number of Final Paleolithic sand-dune and river deposits not far from Warsaw in the vicinity of Swidry Wielkie. Today, thanks to studies of the geological sequences and the pollen deposits at Krukowski's archaeological sites, and more recently to carbon-14 dating, both their chronological order and their absolute ages are known.

Some 10 years before World War II the German archaeologist Alfred Rust began a series of important excavations near Meiendorf, a village not far from Hamburg. A number of Final Paleolithic sites were clustered along a valley that had been carved by glacial runoff and then filled with lake-bottom muds and peat. The several field seasons at Meiendorf brought together scholars representing many disciplines, and their combined results set a standard for all subsequent interdisciplinary campaigns. Rust's sites yielded the first trustworthy

relative dating of Final Paleolithic campsites to be obtained from pollen analysis. Moreover, the many animal bones preserved in the muds provided significant information on hunting preferences in the Final Paleolithic.

Rust's example has been followed ever since. Multidisciplinary excavation at such sites as Bromme in Denmark, Rissen in West Germany and Calowanie and Witów in Poland have provided a firm base for partitioning the Final Paleolithic into several subdivisions. These subdivisions have been established on the grounds of flint-tool typology, of environmental change and of absolute dating. Indeed, the sheer quantity of evidence brought to light by the study of the sites' geomorphology, soils, pollens and animal remains has surpassed all expectations.

The techniques that are primarily archaeological have also been refined. For example, the first study of the horizontal distribution of artifacts at a Final Paleolithic surface site was made in the vicinity of Swidry Wielkie by Ludwik Sawicki early in the 1920's. Ever since then the precise mapping of the horizontal distribution of artifacts has become a standard field requirement. One result is that prehistorians now have at their disposal detailed distribution maps in field reports covering more than 100 sites. The availability of that information makes possible the mounting of intensive investigations of settlement patterns and



SCATTERING OF FLINT TOOLS in one part of Layer III at Calowanie, a Final Paleolithic site near Warsaw, was painstakingly plotted by the excavators to record the horizontal distribution of the artifacts. Such a scattering could once have been mistaken for the camp debris left by a single group; the presence of two separate articulation nets (*arrows*), however, shows that the concentrations were the work of different families. Articulation nets link finished tools to the raw flint they came from. Two articulations in the upper concentration (*color*) serve as examples. The letter *a* marks an arrow connecting an end scraper with the flint core from which it was struck; *b* labels the connection between a truncated point and a piece of worked flint found more than one meter away.

other aspects of the behavior of hunting groups in the Final Paleolithic. Another technical advance has been the integrated study of flint mining, of traffic in this important raw material and of flint-processing "workshops." The pioneer analyses in this field were published by Krukowski in 1919. His work and the work of others in later years has opened a broad field of research touching on economic, social and cultural aspects of prehistoric life.

The most abundant artifacts found at Final Paleolithic sites on the European plain are various kinds of stone tools; the vast majority of them are made from flint. Scholars have spent decades devising various systems for classifying Paleolithic flint implements of all ages, including those of the Final Paleolithic. Among the major tool categories that one finds in the tool assemblages of the Final Paleolithic are flint burins (believed to have been used for working bone and antler), flint scrapers (for working animal hides), large and small flint blades for cutting, perforators for drilling and piercing, and flint points of various kinds, some of them used on projectiles ranging from spears to arrows.

Today, with complex statistical analysis made easier by computers, many prehistorians have come to base their systems of flint-implement classification on analytical techniques that help to evaluate the significance of the presence or absence of certain qualitative and quantitative differences that separate assemblages of artifacts from one another. In Poland we use 50 or more attributes in computing the relations among assemblages. The system was initially established by surveying the best Polish stone-tool assemblages then available, including the rich heritage left us by earlier workers.

To speak rather generally, an assemblage consists of all the stone tools found in an isolated occupational unit at a particular level in a particular site; thus an assemblage is the fundamental unit in any given system of classification. When the attributes of several separate assemblages are "clustered" statistically, the process reveals classificatory units of a higher order. We assign no formal name to these classificatory clusters, although they are sometimes called industries or culture groups.

When we take the next step and statistically group several industries together, we call the result a cycle. An equivalent term in English is culture, and I shall use it here. A culture is traditionally given a formal name (for example the Ahrensburgian culture, which we shall discuss below). Cultures can themselves be statistically grouped into one or another of two higher classificatory units. These higher groupings of course embrace a great many individual tool assemblages, and the more assemblages there are, the fewer detailed resemblances there will be between them. The higher groupings nonetheless share a substantial number of characteristics, not only in terms of stone artifacts but also in terms of the environment of the cultures. The smaller of the two units, a clustering of cultures, is sometimes called a complex. The larger grouping embraces the greatest number of tool assemblages at the bottom of the hierarchy, and the process of amalgamation continues until two or more complexes emerge. When we group the complexes, we call the result a technocomplex, following a usage proposed by David L. Clarke of the University of Cambridge.

How is this classificatory system applied in practice? Consider the Final Paleolithic on the European plain. To begin in about 11,500 B.C., the culture (defined by its tool assemblages) of the people who occupied the European plain from then until about 10,500 B.C. is classified as the Hamburgian. During the main cold phases of the Würm glaciation and up to about 11,500 B.C. the plain had been a practically lifeless arctic wasteland, but at the end of the period a warming trend in the climate allowed a shrubby tundra to replace the icy desert, and herds of reindeer advanced northward with the tundra vegetation.

Hamburgian tool assemblages are known from sites in the Netherlands, in West Germany and in Poland. They include burins, perforators, end scrapers and characteristic shouldered blades, possibly used as projectile points. Many of the blades are truncated, that is, they are blunted obliquely along one edge. The remains of reindeer, which are found in substantial numbers at Meiendorf (Rust's site) and at nearby Stellmoor, show that the men who used the Hamburgian tools engaged in selective late-summer hunting of the migratory reindeer herds.

European archaeologists still debate the relation of Hamburgian culture to higher classificatory groupings. Rust, who has spent most of his career excavating Hamburgian and other Final Paleolithic sites, believes the culture is a unit within a higher grouping known as the Eastern European Upper Paleolithic. Other scholars prefer to associate the Hamburgian with western Europe and specifically with the Magdalenian technocomplex that flourished in what is now northern Spain and southern France and extended into Switzerland, southern Germany and Czechoslovakia. In any event the Hamburgian culture was at best sparsely represented on the plain after 10,400 B.C., and it had disappeared well before the warm Alleröd interstadial began about 9900 B.C.

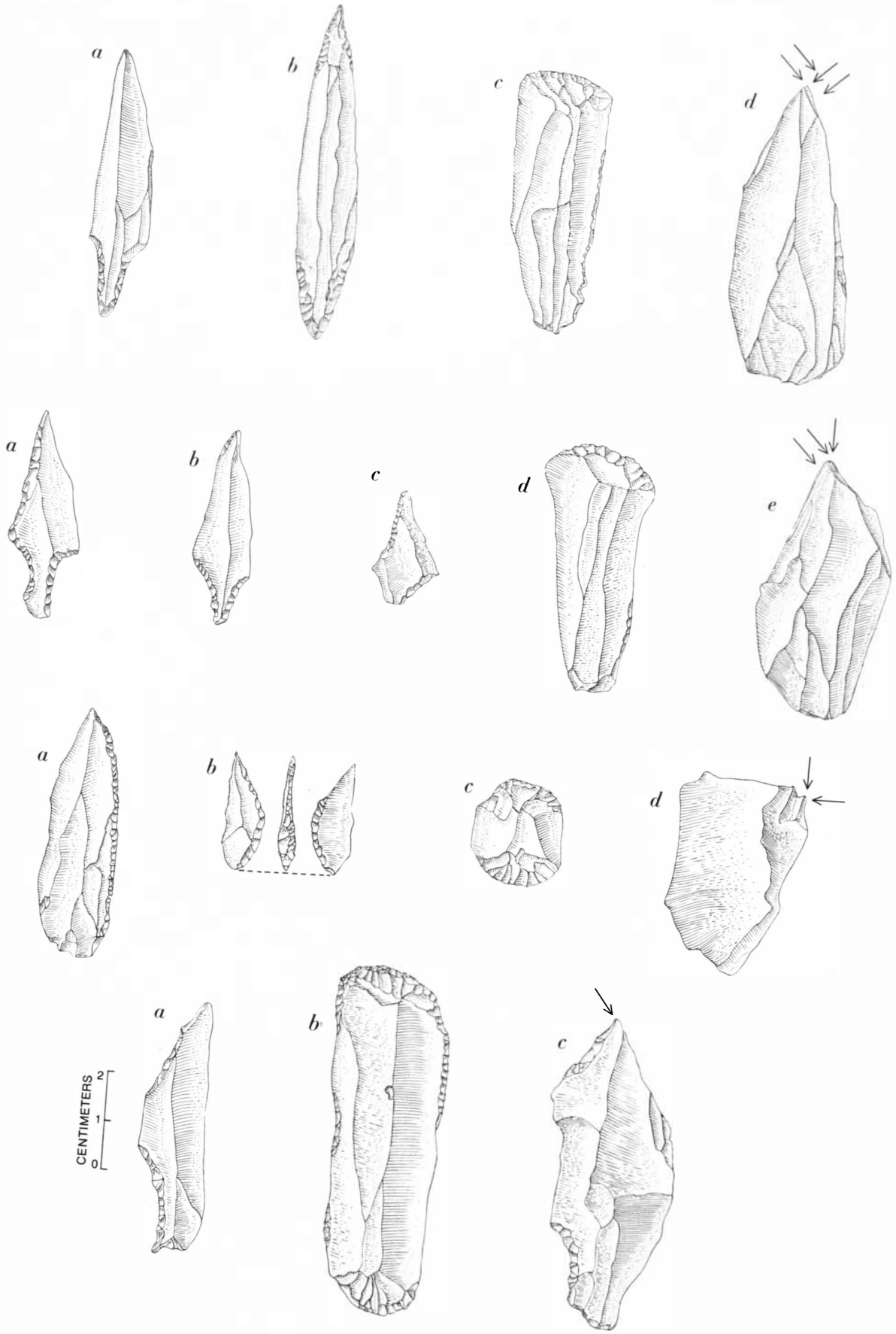
The stone-tool complex that next appeared on the European plain was contemporaneous with the Alleröd interstadial. It is classified as the Arched,

EUROPEAN PLAIN at the end of the Old Stone Age was a wide belt of tundra extending well north and west of today's coastline to include Britain and Ireland. Part of Scotland and most of the Scandinavian peninsula were covered by glacial ice caps. The

Backed Piece Complex. The assemblages characteristic of the complex include an abundance of small flint blades, some two centimeters long, shaped like orange sections or half-moons. The roughly straight side of the bladelet is its cutting edge. The "arched," or rounded, side of the bladelet has been backed, that is, blunted [*see illustration on next page*]. Here I shall call this the arched-bladelet complex. Almost of equal prominence among the artifacts of the complex are small end scrapers, most of them no more than three centimeters on a side, that are often called thumbnail end scrapers. Burins are slightly less common.

Arched-bladelet tool assemblages are found over a wide area of the European plain. In East and West Germany they are said to belong to the Federmesser culture; they are known in the Netherlands and Belgium as the Tjongerian culture and in Poland as the Tarnovian culture. The earliest of the assemblages seem to date to the middle of the Alleröd interstadial. For example, the assemblages at Rissen, a site on the Elbe River in West Germany, have a carbon-14 date of about 9500 B.C. At Calowanie the assemblages appear in Layer III, which has a carbon-14 date of about 9400 B.C. The arched-bladelet complex evidently endured on the European plain for about a millennium, extending from the Alleröd into the Late Dryas period. At another Polish site, Witów, a sandy deposit formed during the dune-building phase of the Late Dryas includes a stratum where arched-bladelet

Baltic was an ice-laden lake that reached far north of its present boundaries, and the cold climate of the Late Dryas period had long since killed the pine forests that once covered the plain. Shown on the map are 20 major sites of the Final Paleolithic. In the cluster near the present mouth of the Elbe River is Meiendorf, one of many sites where stone tools typical of the tundra-adapted Hamburgian culture (11,500–10,500 B.C.) have been found. Of the sites on or near the Vistula River, Calowanie is one of those where tools of two successor groups, a forest-adapted and a tundra-adapted society, have been found. Oronsko, in the foothills of the Holy Cross Mountains, was a major source of the distinctive chocolate-colored flint much favored at that time by the toolmakers of the European plain.



assemblages are found dating to about 8600 B.C. So much for the duration of the arched-bladelet complex in time. As for geographical range, assemblages that include arched bladelets and thumbnail end scrapers characteristic of the complex have been identified in France, Italy, Switzerland, Austria, Czechoslovakia, Hungary and the Ukraine.

Some students of the arched-bladelet complex would like to associate it with some higher classificatory grouping. For example, Hermann Schwabedissen of the University of Cologne contends that the Federmesser culture represents a wave of colonization from an earlier enclave of the Magdalenian technocomplex located in southern Germany. Those opposed to his view rest their case on paleoenvironmental grounds. At widely separated sites in Europe the appearance of the arched-bladelet complex is clearly contemporaneous with the appearance of temperate-climate boreal forests. During the Bölling interstadial in France, an interval of moderate temperatures that interrupted the cold climate of the Early Dryas beginning about 10,400 B.C., pine forests replaced tundra vegetation in Provence. Arched-bladelet assemblages, representative of the Romanellian culture, also appear in Provence; the carbon-14 dates associated with the earliest of these assemblages are contemporaneous with that period of forestation. Contemporaneous with the forests that arose in the subsequent interstadial, the Alleröd, were the stone-tool industries known in France as Azilian (the artifacts of which are similar in form to those of the arched-bladelet complex), sites in Czechoslovakia such as Kulna (with assemblages that include arched bladelets and thumbnail end scrapers) and Azilian-like sites in the Ukraine such as Borshevo II (upper layer). In the light of this chronology it seems altogether likely that the various arched-bladelet industries that spread across Europe as the pine forest advanced northward somehow reflect a specialized adjustment of stone-tool technology to life in a wooded environment.

Along the northern fringe of the European plain the tundra had now been replaced not by dense forest but by a parklike open forest of birches. In this region one can see a foreshadowing of things to come in the form of tool assemblages that include projectile points, some more than 10 centimeters long, that were chipped so that they have a blunt tang at the end where they were mounted on a shaft. The distinctive assemblages that include tanged points in their inventory have been found in Alleröd-period sites in Sweden, Denmark, West Germany, Poland and most recently in Lithuania. The tanged-point assemblages as a group are known as the Bromme complex, after the well-known Final Paleolithic site in Denmark. Studies of animal remains and pollen at Bromme suggest that the birch-forest landscape of the time included many lakes and ponds. There were reindeer in this subarctic region, but the animal the hunters killed in the largest numbers was the moose.

Shortly before 8800 B.C. an abrupt shift to a colder climate ushered in the Late Dryas period. As tundra conditions returned to the European plain the Alleröd forests literally died out. Fires swept through vast stands of dead pine all across the region; a stratum at Calowanie contains ash and charcoal from one of these conflagrations. The climatic change and the return to a park-tundra environment coincided with a radical cultural change in the hunters' societies on the plain, a change that is also reflected in their stone tools.

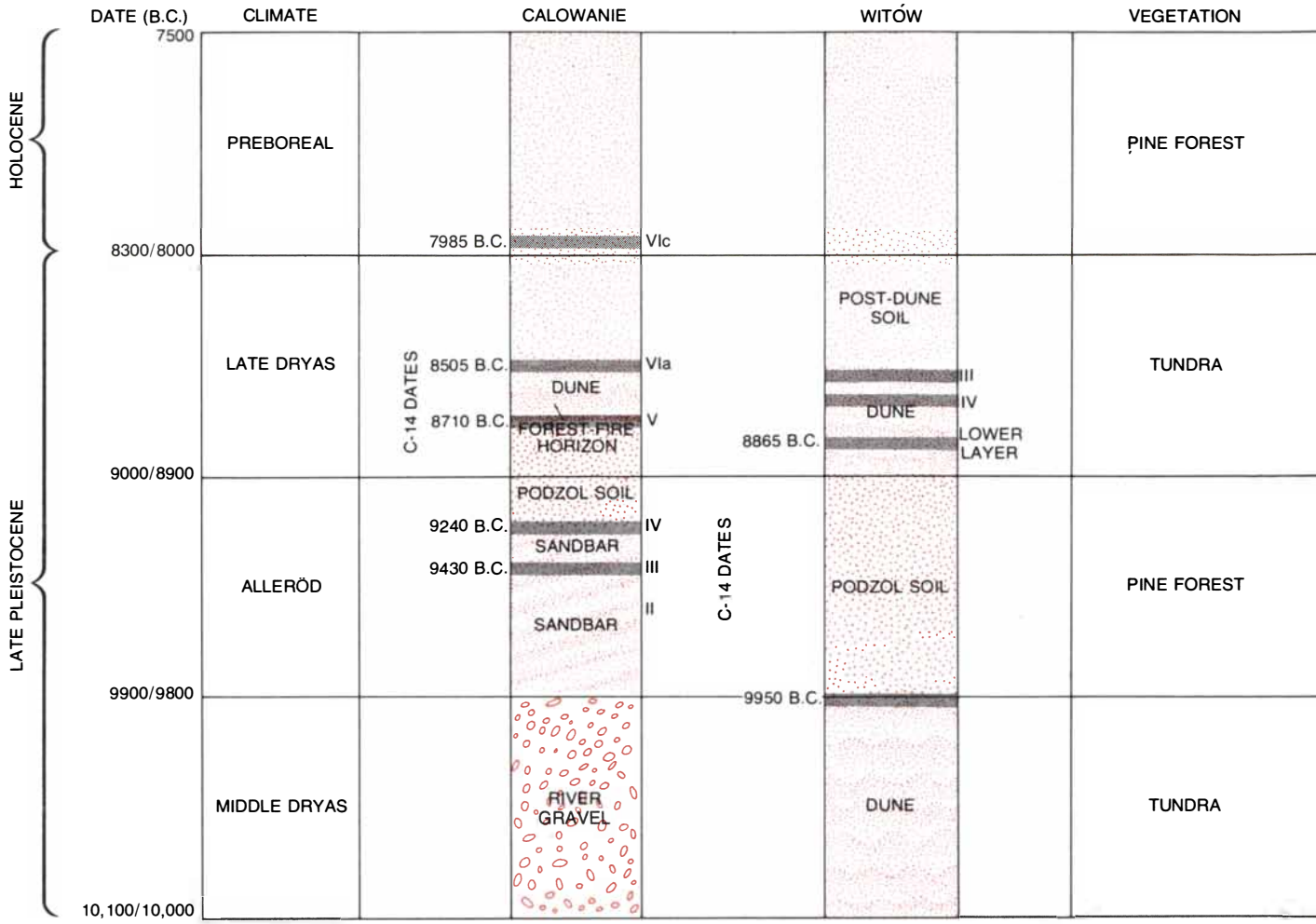
In the western part of the plain (for example in Belgium, in parts of the Netherlands and in West Germany) the tool assemblages are no longer characterized by arched bladelets and thumbnail end scrapers. Instead tanged points are found, together with small truncated microliths and end scrapers that are considerably larger than the thumbnail type. The new tool assemblages show signs of having been derived from the Bromme complex; the culture they represent is known as the Ahrensburgian.

In Poland at the beginning of the Late Dryas arched bladelets and thumbnail end scrapers persist locally in the west; they are found, for example, in two early strata of the Late Dryas at Witów. Slightly to the east of Witów, however, their place is taken at such sites as Calowanie by tool assemblages that include Bromme-like tanged points. Similar tanged points appear at Witów some 200 years later, although arched bladelets continue to be present as well; the combination gives rise to a peculiar classificatory unit. In any event the new assemblages, which include not only tanged points but also large and small end scrapers and numerous burins, form a group known in Poland as the Masovian culture.

An overall chronology for the Final Paleolithic in Poland, which has been established only recently, indicates that the Masovian culture reached its height about 8500 B.C., a time when the dune-building phase of the Late Dryas had ended. By then sites all over Poland, in western Russia and in East Germany west of the Oder River were occupied by hunters who used the Masovian assemblage of stone tools. The Masovian is perhaps the best represented of all Final Paleolithic cultures on the European plain; in Poland alone more than 620 sites have yielded Masovian tool assemblages. At the same time the Masovian culture is basically similar to the Ahrensburgian, and it seems inescapable that both are subdivisions of a single larger unit. We may call that unit the Tanged Point Technocomplex and assign to the earlier Bromme complex the status of an older subdivision within it.

How did the Ahrensburgian and Masovian cultures originate? Some have suggested that the Masovian is associated with a number of tool assemblages that have been found in eastern Byelorussia. The suggestion is that these Russian assemblages are at least as old as the pre-dune-building phase of the Late Dryas, if they are not contemporaneous with the latter part of the Alleröd period. It should be noted, however, that the Russian tool assemblages (collectively known as the Griensk complex) are rather scanty and that their true age is in question. On the other hand, the oldest tool assemblages in Poland that show strong Masovian characteristics (for example the assemblages in Layer V at Calowanie) also include typical Bromme and Ahrensburgian artifacts. These artifacts in turn include

TOOLS OF THE FINAL PALEOLITHIC are arranged in four groups in the illustration on the opposite page. The oldest (*bottom*), typical of the Hamburgian culture, include a blade (*a*) that is partially blunted on one edge by retouching and that may have served as a projectile point, an end scraper (*b*) and a burin (*c*). In the middle are four tools of the arched-bladelet complex: a backed blade (*a*), a burin (*d*) and the two tools most representative of the complex. These are the half-moon-shaped, backed bladelet that gives the complex its name (*b, in three views*) and a small thumbnail end scraper (*c*). Above these are tools of the Ahrensburgian complex: small tanged points (*a, b*), a truncated microlith (*c*) known as a Zonhoven point, an end scraper (*d*) and a burin (*e*). Four tools of the Masovian complex are illustrated at top: a tanged point (*a*), a willow-leaf point (*b*), an end scraper (*c*) and a burin (*d*). The two upper complexes together form the Tanged Point Technocomplex; the makers of these flint tools were the last Old Stone Age peoples of the European plain.



CHRONOLOGY OF FINAL PALEOLITHIC coordinates successive climatic periods between 10,000 and 7500 B.C. with the carbon-14 ages of archaeological deposits both at Calowanie and at Witów, a second Polish site, with changes in the kinds of plants and animals and with the succession of cultural units on the European plain. The deposits found below Layer III at Calowanie relate to unidentified groups that visited the site when the tundra of the cold Middle Dryas period was yielding to pine forest early in the

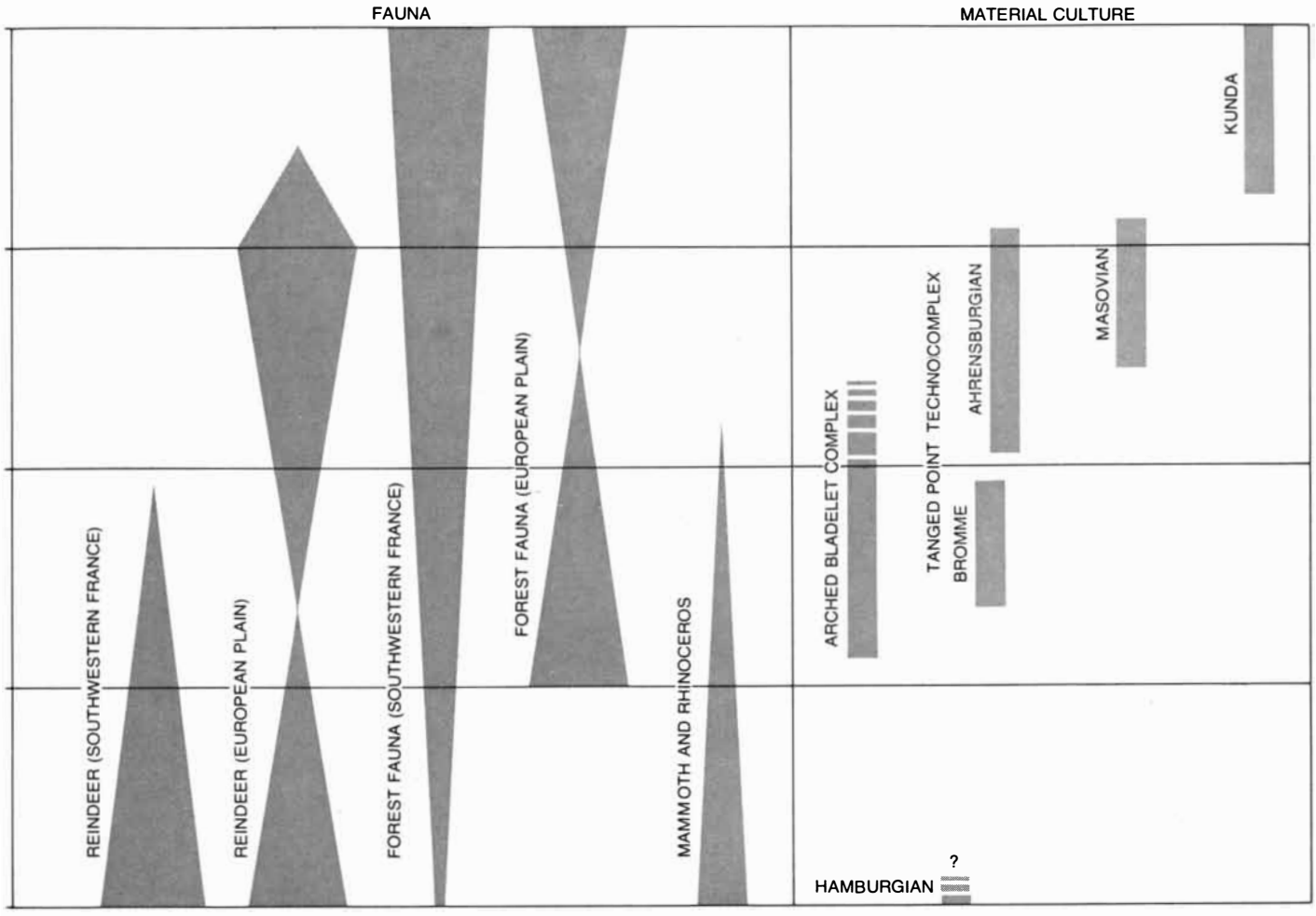
small tanged points and the kind of truncated microlith known as a Zonhoven point. The fact that such artifacts appear together in a number of assemblages that are as old as the beginning of the Late Dryas climatic shift strongly suggests that the Ahrensburgian and Masovian cultures have a common origin.

From this point of view the Tanged Point Technocomplex, most probably in the form of its early subdivision the Bromme complex, is seen as being representative of hunters' communities adapted to the tundra and park-tundra environment that, although it was limited in extent, still existed during the Alleröd interstadial to the north of the pine-and-birch-forest belt. With the arrival of a colder climate in the Late Dryas the tundra quickly reinvaded the south, and the tundra-adapted hunting societies followed it to enter the more southerly parts of the European plain. It was probably after the push southward, perhaps in middle and upper Late Dryas times, that the two components of the technocomplex developed the stylistic differences that enable one to distinguish between the Ahrensburgian and Masovian cultures.

What game animals supported the hunters of the Tanged Point Technocomplex? Ahrensburgian sites have yielded the bulk of the animal remains. In several instances the faunal evidence is strikingly similar to that from Hamburgian sites that are 3,000 years older. The main game animal was the reindeer. The main weapon was the bow and arrow; the characteristic tanged points of the technocomplex have been found more than once still mounted on arrow shafts.

The temporal boundary between the end of the Pleistocene ice age and the beginning of the Holocene was marked by another climatic change. Sometime between 8300 and 8000 B.C. the cold Late Dryas period was superseded by a warmer period known as the Preboreal. Pine and birch forest reinvaded the European plain from the south, and the reindeer herd, following the retreating tundra, moved off to the north and northeast. With the returning forest another community of animals appeared on the plain: red deer, roe deer, moose, wild cattle and wild boar.

The climatic shift that introduced the Holocene epoch coincided with the disappearance of both the Ahrensburgian culture and the Masovian. The last Masovian tool assemblages in central Poland, for example those in Layer VI*c* at Calowanie, are no more recent than about 8000 B.C. In Lithuania and Latvia, for example at such sites as Sindi on the northeastern part of the plain, assemblages of a new kind appear about 7600 B.C. The tools still display some Masovian characteristics, but they have a clear identity of their own as forerunners of the later Kunda complex, which flourished in the Mesolithic period rather than the Paleolithic. With the disap-



following warmer period, the Alleröd. After that at both Polish sites successive deposits contain tools typical of the forest-adapted arched-bladelet complex. Visitors to Calowanie on occasions more than two centuries apart left behind tools of this kind. The visitors who camped there soon after the end of the Alleröd period instead used tools typical of the tundra-adapted Tanged Point Technocomplex. The same was true of subsequent Old Stone Age visitors to the site, including the last groups to camp there some 10,000 years ago.

pearance of the Masovian and the Ahrensburgian, both the Late Pleistocene and the Final Paleolithic on the European plain appear to have drawn to a close.

My review of prehistory on the European plain during the Final Paleolithic is necessarily oversimplified, but it will have provided examples of how the analysis of stone tools, the study of animal remains and the painstaking reconstruction of environmental changes can add to our knowledge of extinct societies. We gain some sense of the ebb and flow of successive technocomplexes and of how hunting societies adapted to the park-tundra environment were also closely confined to it. When the arrival of a cold climate caused the forest to retreat and the reindeer herds to move south, the hunters followed; when a warm period allowed the forest to advance, the reindeer herds followed the retreating tundra northward and the hunters followed the herds again. We also see that, at least in some areas, the transition from one technocomplex to another was not as simple and immediate as the changes in environment were. An example is provided by the 200-year persistence of arched-bladelet assemblages at Witów and their partial coexistence with assemblages including numerous tanged points.

The information I have reviewed so far gives us only a limited understanding of the details of social organization in the Final Paleolithic. Analysis of the structure and behavior of these hunter-gatherer societies calls for exhaustive studies at individual sites. Only in this way shall we come to understand the interaction between society and environment that is the essence of human economic and behavioral adaptation.

Let me illustrate the point by outlining what we have learned from the investigation of successive human occupations at Calowanie. We shall leave aside the two lowest occupation levels, Layer I and Layer II; the assemblages found in them belong to cultures of the early Alleröd period that have not yet been fully identified. Layer III at Calowanie is a level that was occupied by hunters sometime about 9400 B.C., halfway through the Alleröd.

The tool assemblages in Layer III include many thumbnail end scrapers and arched bladelets. Evidently the assemblages belonged to the complex that, as we have seen, represented hunter-gatherer communities adapted to forest life during the Alleröd. We found this to be true also of the tool assemblages in Layer IV, the next higher occupation level. Our first question was: What could be learned about these forest-adapted communities from the horizontal distribution of their artifacts and other debris of occupation?

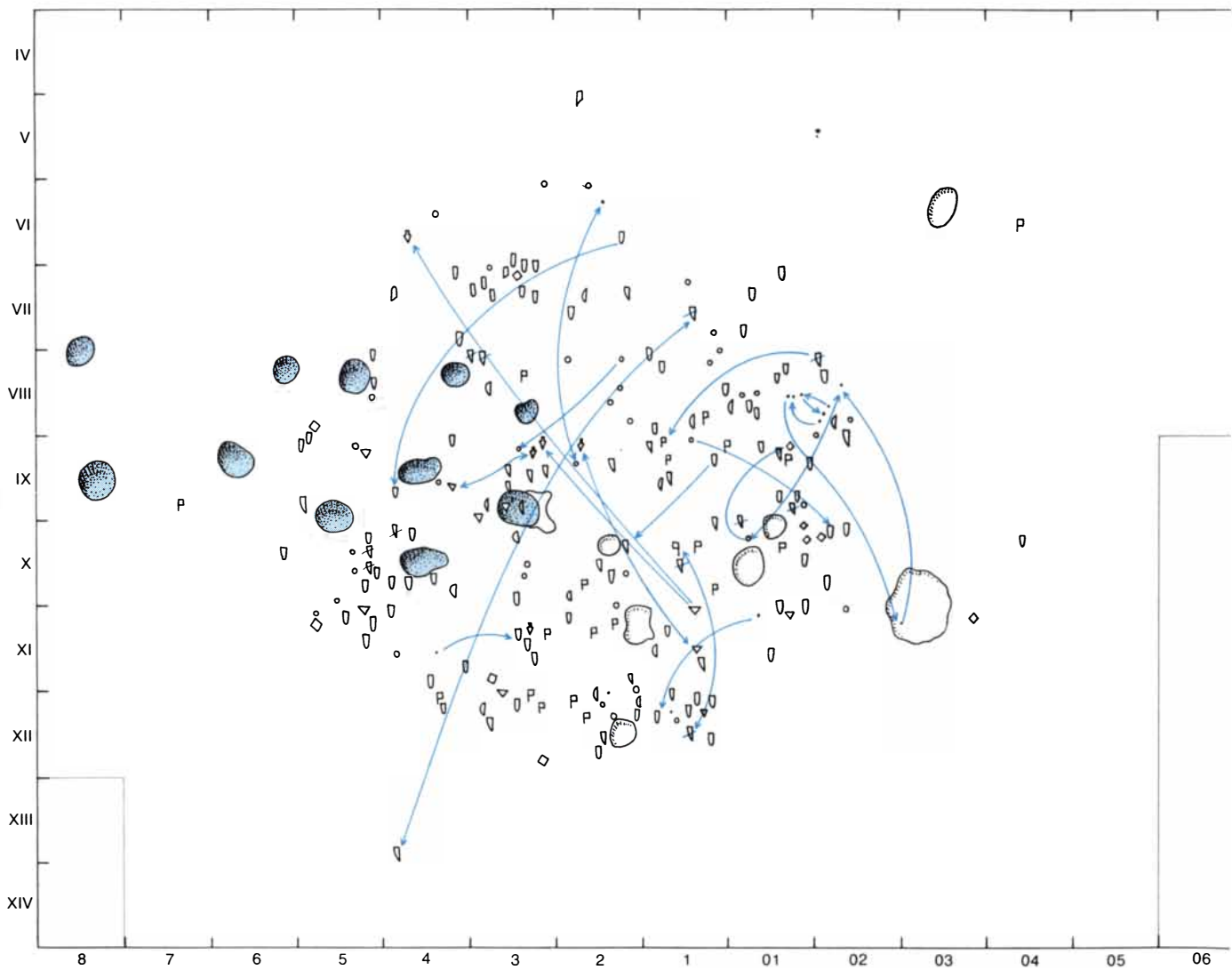
From the time of Sawicki's pioneering studies in the 1920's it has been known that when the locations of the

stone tools at Final Paleolithic sites are mapped horizontally, the tools are often found in arrays that may be as small as four meters in diameter and rarely exceed eight meters. That other pioneer, Krukowski, first named these horizontal arrays *kshemenitsas,* a Polish word meaning flinty places, and the name has persisted. Here, rather than asking the reader to deal repeatedly with this Polish term, I shall call the arrays concentrations.

Layer III at Calowanie contained several of these horizontal concentrations. Two of them might have been mistaken for a single concentration larger than usual if it had not been for the application of a demanding study technique. The technique involves matching stray flint flakes and blades, and even completed tools, with the flint "cores" from which the blanks used for tool manufacture were struck. Such a matching procedure generates what we call articulation nets, which define the boundaries of a single occupation unit, that is, the "household" believed to have been occupied by a single social or task group. For example, two articulation nets recorded in one part of Layer III made it clear that not one but two concentrations were present there.

Why are the single concentrations, when they are undisturbed, so limited in their extent? Where there are signs of dwelling construction, such as postholes or the boundary of a pit house, in association with a concentration of stone artifacts, the size of the shelter is also quite small. The implication is that the group responsible for the concentration was also small in numbers. It may have been a nuclear family, such as a man and a woman and their children, or at most a minimally extended family. Alternatively it may have been a small, temporary task group: a few hunters or a few flintworkers.

Family groups appear to have outnumbered task groups on the European plain during the Final Paleolithic. One is led to that conclusion by the functional diversity of the tool assemblages found in most concentrations. Microscopic studies of the patterns of wear shown by the tools have revealed characteristic traces of use associated with particular kinds of tools. The usual concentration contains tools that, when they



REMAINS OF TWO CAMPS unearthed in Layer IV at Calowanie were mapped in the same detail as the camps illustrated on page 89; here, however, flint blanks and chips that are not part of the articulation nets

are analyzed in this way, indicate that many different tasks were performed at the site. This finding is compatible with family occupation of the camp; when a task group is responsible for a concentration, the tool assemblage is a more specialized one that shows little functional variety.

We deduce from the evidence in this part of Layer III that the two tool concentrations represent either two independent and short-lived occupations or a visit by a band consisting of two families. At the time of occupation the settlement area was a flat and frequently flooded sandbar in the Vistula River that could only have been occupied during the low-water season. This fact, together with the lack of any evidence that permanent dwellings were erected, suggests that the tool concentrations were left at a summer camp. There is no strong reason to either believe or disbelieve that the two camps were occupied simultaneously.

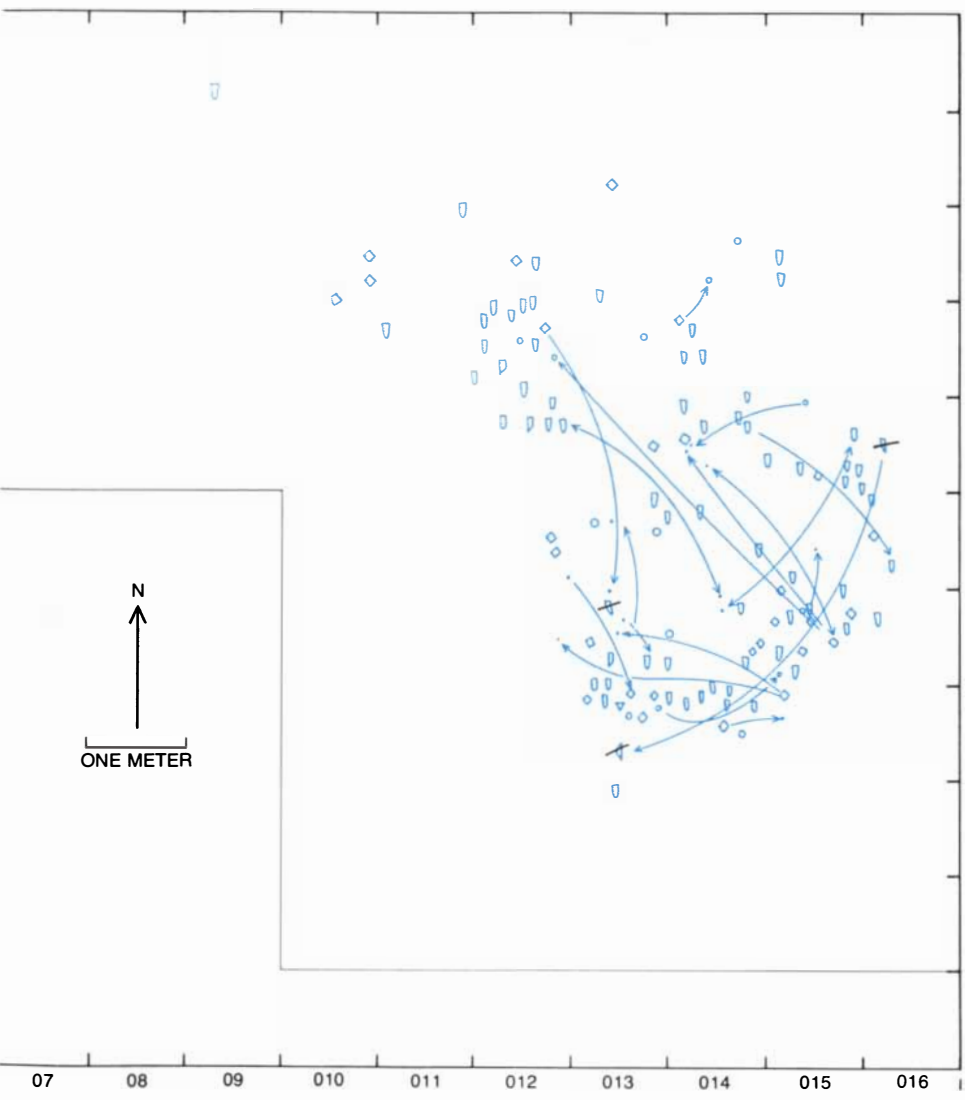
The next occupation level at Calowanie, Layer IV, also held a number of independent tool concentrations. The artifacts continue to be representative of the arched-bladelet complex, although in one adjacent pair of concentrations the tools are representative of two separate industries within the complex. Both the abundance of artifacts in the two concentrations and evidence for a small but sturdy hut in the more westerly of the two suggest that these were long-lived base camps. The tool assemblages are diversified, which suggests that the camps were occupied by nuclear families rather than by specialized task groups. At the same time the number of end scrapers present is somewhat larger than usual. The scrapers show signs of wear, and most of them were evidently retouched more than a few times to give them a fresh working edge. These are signs that the occupants were heavily engaged in the processing of hides. Both soil studies and pollen analysis at Layer IV furnish strong evidence that the sandbar was now covered with trees and brush.

Layer V at Calowanie, lying just above Layer IV, contains thin concentrations of flint artifacts that suggest short-lived camps. Remnants of burned vertical poles at one concentration indicate that a tent or some similar temporary shelter had been raised there. The tool assemblages are no longer representative of the arched-bladelet complex. Instead they contain both Ahrensburgian and Masovian tanged points. These temporary encampments, shelters for tundra-adapted reindeer hunters, were occupied soon after 8800 B.C., as the Alleröd period came to a close and the Late Dryas began. The landscape at Calowanie, dominated by the trunks of dead pines, must have been bleak. The dead trees were later swept by fire, the heat cracking most of the artifacts in the concentrations and the burned timber leaving a layer of pine charcoal.

The succeeding occupation level at Calowanie, the lower portion of Layer VI, overlies a considerable thickness of dune sand that accumulated during the middle centuries of the Late Dryas. The flint concentrations are rich, and the artifacts are typically Masovian. The tool assemblages are quite specialized, with a heavy emphasis on arrowheads and burins. One supposes that the camps were occupied by hunters' task groups.

A span of some five centuries separates the lower and upper portions of Layer VI. During that time, from 8500 to 8000 B.C., a covering of soil developed on top of the Calowanie dune sands. The upper portion of Layer VI contains several flint concentrations. The artifacts are typically Masovian, but the concentrations are rather thin and the tools are few in number. They are the last traces of a Paleolithic hunting society at Calowanie.

The pattern of successive campsites revealed by the excavations at Calowanie is repeated and amplified at scores of other sites on the European plain. Taken together with reconstructions of the changing environment and the changes in technology implicit in tool-



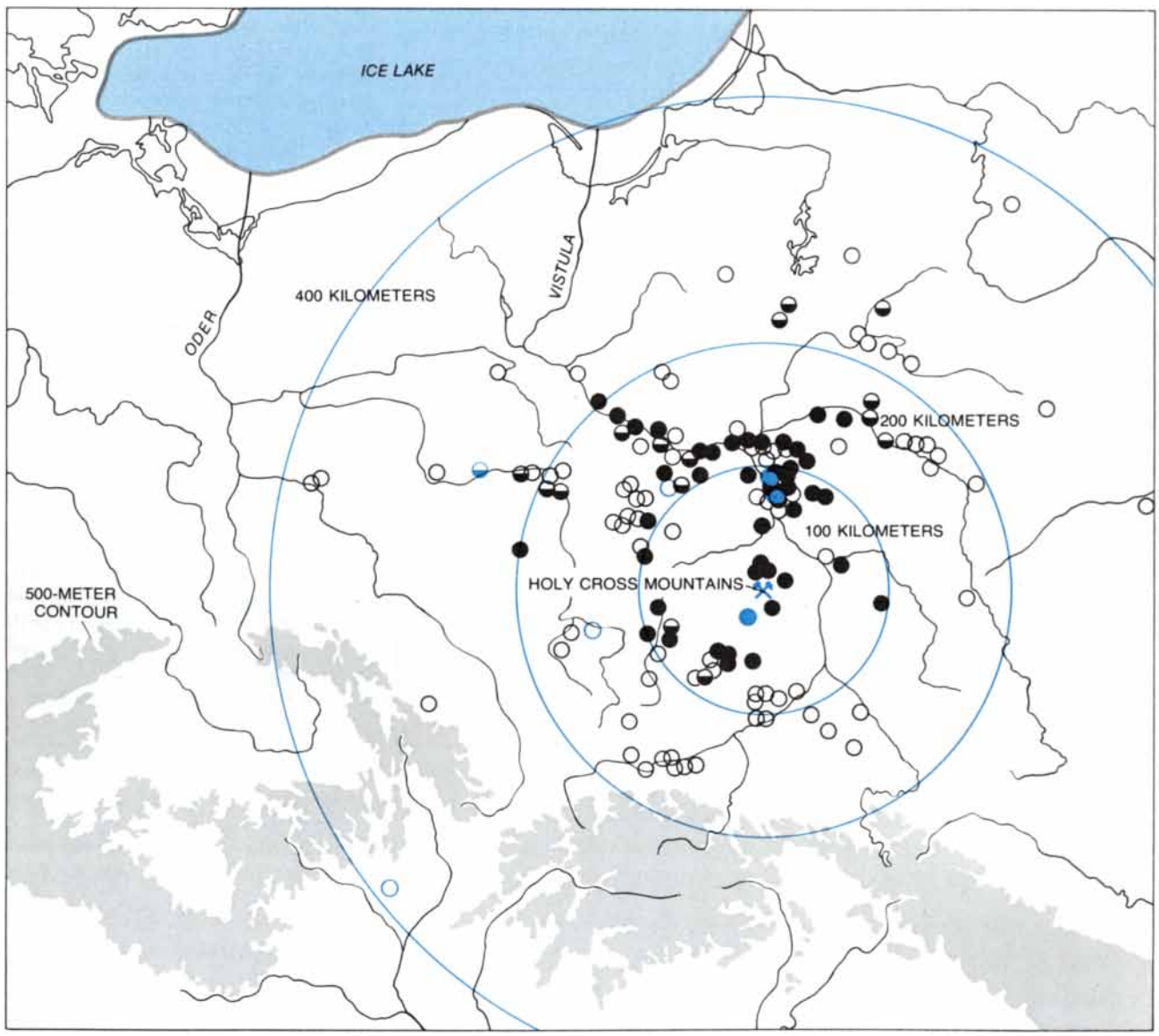
have been omitted from the final plan. The postholes (*color*) in the western encampment suggest that the families that left so many tools behind occupied the camp for some time, perhaps months. Both flint concentrations contain end scrapers in greater than usual numbers, and many of the tools show signs of having been resharpened. The two families may therefore have been heavily engaged in the processing of animal hides during their visit.

assemblage differences, the detailed site analyses yield considerable information about the societies of the Final Paleolithic.

A common technological response to a forest environment is apparent in the long-lived and widespread arched-bladelet complex. The largest population unit appears to have been no greater than two or three nuclear families, and perhaps in some specific instances a single family was the norm. Whatever the size of the units, they occupied base camps for comparatively long periods: perhaps two or more successive seasons. Shorter-lived settlements, probably summer camps, were also occupied, and specialized task groups, characterized by less diverse tool inventories, set up temporary encampments. The existence of these small and seemingly quite flexible social groups is indicative of a successful adaptive response to a dense-forest environment that sheltered a dispersed but nonmigratory animal population.

With the retreat of the forest at the start of the Late Dryas a change appears in the nature of the social group. For example, evidence from the lower occupation level at Witów can be interpreted as indicating the simultaneous presence of four families. Such a social change could easily have been brought about by the return of the tundra environment and, associated with the tundra, a migratory game animal that travels in herds: the reindeer. In any event the artifact concentrations thereafter, with their characteristic tanged points, are representative of a new technocomplex associated with a new society of hunters, well adapted to tundra conditions, that seems to have first developed along the northern fringe of the Alleröd forests.

With a few exceptions at the southern limits of the plain the new tundra hunters almost exclusively favored the rein-



DISTRIBUTION OF TOOLS made of chocolate flint from central Poland (*color*) was more extensive during the Late Dryas period, when tools typical of the Tanged Point Technocomplex were used on the European plain (*black symbols*), than it was during the Alleröd period, when the plain was forested and tools typical of the arched-bladelet complex were used (*colored symbols*). Circles, centered on the Holy Cross Mountains, have radii of 100, 200 and 400 kilometers. Small open circles locate sites where some such tools were present; half-filled circles, those where many were present, and filled circles, those where such tools outnumber others.

deer as a game animal. As a result a new campsite pattern emerges. The reindeer herds were relatively immobile in their winter pasture; in that season the multifamily groups apparently broke up and went into single-family winter quarters. That is at least one way to interpret the evidence from such sites as Rydno and Deimern, with their typical small winter pit houses and diversified assemblages of tools. By the same token, when the reindeer began their migration from winter to summer pasture, multifamily groups would reassemble and follow the herd. That is at least one way to interpret the evidence of short-term encampments in Layer V at Calowanie. Further evidence of the shift from winter social patterns to summer ones in response to reindeer migration may lie in the fact that large numbers of Tanged Point Technocomplex sites are found in the valleys of the European plain that are oriented along a north-south axis. The migratory animals' natural preference for those direct routes in traveling to and from their summer pasture would have made the north-south valleys the most desirable locations for temporary hunting encampments.

To the extent that we can reconstruct the evolution of human societies in the Final Paleolithic, what is evident is a transition from groups rarely larger than one or two nuclear families adapted to living in one territory in a forest environment to groups adapted to an essentially nomadic life, groups forming large multifamily aggregates in the summer and breaking up into nuclear families only in the winter. In this connection it is interesting to note that when early Mesolithic societies appeared in what is now Poland, more than 500 years after the end of the Final Paleolithic, the climate had once more become milder and the European plain was again covered by an Alleröd-like forest. Evidently in a similar response to this similar environment the early Mesolithic hunters and gatherers tended to live in one territory, and their largest encampments seem to have been limited in most cases to one or two nuclear families.

The transition from a comparatively sedentary way of life to a nomadic one in the Final Paleolithic is supported by evidence that is quite independent of either the reconstruction of environments or the behavioral information gained from the analysis of flint concentrations. That independent evidence comes from studies of the mining of flint and the traffic in this essential material during the Final Paleolithic. In Poland such studies have been made fairly simple by an accident of geology.

LEAF-SHAPED POINTS typical of the Masovian complex were used as arrowheads. This cache of arrows is from a grave at a Russian site, Deer Island on Lake Onega. The compound shafts were made partly of bone and partly of wood. Deer Island is an early Holocene site.

The Holy Cross Mountains of central Poland lie roughly in the middle of a line drawn from Warsaw to Cracow. Their northeastern foothills include a zone of Mesozoic limestones and clays that are rich in a unique kind of flint. The flint ranges in color from light brown to dark brown, which explains why the material from the Holy Cross foothills is known as chocolate flint. Seventeen mining complexes of various ages have been identified in the foothills; at one of these, Oronsko, narrow shafts of Final Paleolithic age have been sunk as deep as three meters into the flint-rich clay. Some flint outcrops seem to have been exploited during the Upper Paleolithic, but the mines were worked most intensively during and after the Final Paleolithic.

If one plots on a map the Alleröd-period sites that contain artifacts made from chocolate flint, the sites with the most numerous assemblages of such tools are the three that lie within a 100-kilometer radius of the mining area [*see illustration on opposite page*]. Of the sites outside the 100-kilometer circle, only the one farthest to the northwest yielded more than a few chocolate-flint pieces. This comparatively limited distribution of chocolate flint is believed to indicate that the hunters and gatherers of the Alleröd forests did not move around much. This is not to say that the forest people never left their base camps. It does suggest, however, that the distance they traveled, either in visiting the mines or in flint trafficking, rarely exceeded 100 kilometers.

If one next plots the Late Dryas sites that contain chocolate-flint artifacts, the result is quite different. First, chocolate flint constitutes more than 90 percent of the flint found in many Late Dryas sites within 200 kilometers of the source; the principal concentrations are in sites along the Vistula. Second, although the number of chocolate-flint artifacts found diminishes beyond a 200-kilometer radius, some artifacts made from the material are found almost 400 kilometers from the Holy Cross Mountains. This widespread distribution seems to indicate that among the reindeer hunters of the Late Dryas the way of life was more nomadic and the contact between distant groups was more frequent.

The widespread distribution of the chocolate flint suggests that it was an article of trade in the Final Paleolithic, even though only a few articles that could have been given in exchange for it have been identified. For example, in some of the foothill workshops associated with the Masovian culture, where the flint was processed after it had been mined, certain materials have been found that are alien to the Holy Cross region and must have come from across the Carpathians to the south. Among the materials are the volcanic glass obsidian (probably from sources in the Tokay area of Hungary) and the opaque quartz called jasper (probably from the terraces of the Váh River in Czechoslovakia). These exotic stones are from areas populated by societies with stone-tool assemblages that are entirely different from those of the European plain. It is tempting to perceive these foreign materials as being at least some of the items of exchange.

The independent evidence for a transition in human behavior during the Final Paleolithic provided by the chocolate-flint studies is a further example of how our knowledge of human history depends on the integration of information from a wide range of scientific disciplines. Investigations in this field have led to the practical integration of prehistory, a discipline of the social sciences, with disciplines of the biological sciences and the earth sciences. It is a significant example because it bears on an increasing contemporary need: the integration of all scientific knowledge.

Conversation Pieces

Question of the Century

Is there life on Mars? If there is, we hope this masterpiece of miniaturization will prove it. Built by TRW, it packs into one cubic foot the equivalent of three organic chemistry labs full of equipment that's been designed to determine whether micro-organisms exist in the Martian soil. It also contains the complex electronics needed to gather the data for transmission to Earth. Several weeks after July 4th, 1976, when NASA's Viking (built by Martin Marietta) is scheduled to land on Mars, scientists hope to answer the question that has tantalized men for so long.

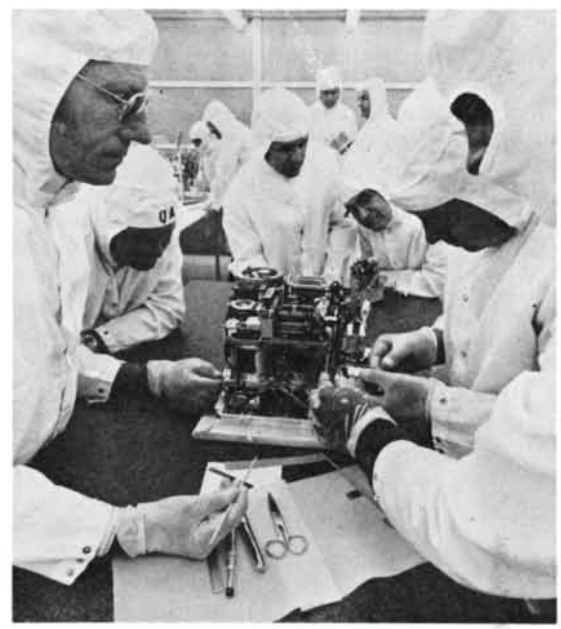
Another TRW-built experiment on Viking will measure the violent winds and cold temperatures of the red planet's tenuous atmosphere. Meanwhile, TRW engineers are working on further experiments designed to probe the hot, dense atmosphere of Venus, later this decade. With Earth's atmosphere "bracketed" by the opposite extremes of these two natural laboratories meteorologists may get new insights into our long-range climatic trends and weather processes.

These experiments are only the latest of many TRW projects designed to help NASA explore the solar system. In 1958, Pioneer 1 was the first spacecraft ever built by a private firm. It showed the shape of the Van Allen belts and measured interplanetary particles and fields. Later Pioneers observed solar disturbances and the solar wind from as far away as the other side of the sun. The most recent Pioneers have crossed the asteroid belt, made close-ups of Jupiter, and sent back more data on the interplanetary medium from record-breaking distances. One is now headed for Saturn and the other will become the first man-made object to leave the solar system.

During the nineteen sixties, TRW built NASA's series of Orbiting Geophysical Observatories, which mapped the Earth's magnetosphere and provided detailed data on phenomena that affect long-distance communications.

We've now started building three High Energy Astronomy Observatories for NASA. Designed to look far beyond our solar system, they'll map sources of X, gamma, and cosmic rays across the entire celestial sphere and then concentrate on the most interesting ones. The results should help answer key questions about quasars, pulsars, black holes, and galactic explosions. They may even throw new light on man's basic theories of energy and matter and on how the universe began.

In skimming the highlights of such projects, it's hard to avoid sounding boastful. But the engineers and scientists who actually do the work have no time for bragging. They're too busy applying the lessons they've learned to new and even more difficult problems. These hard-working people, in fact, are what TRW Systems is all about. If you're interested, we'd like to send you a couple of booklets that give you an overview of our diverse capabilities and more information about the Mars instruments. Just write on your Company Letterhead to:



Building the Viking Lander Biology Instrument involved the use of more than thirty different types of advanced technology.

TRW®
SYSTEMS GROUP

Attention: Marketing Communications, E2/9043
One Space Park Redondo Beach, California 90278

CARBENES

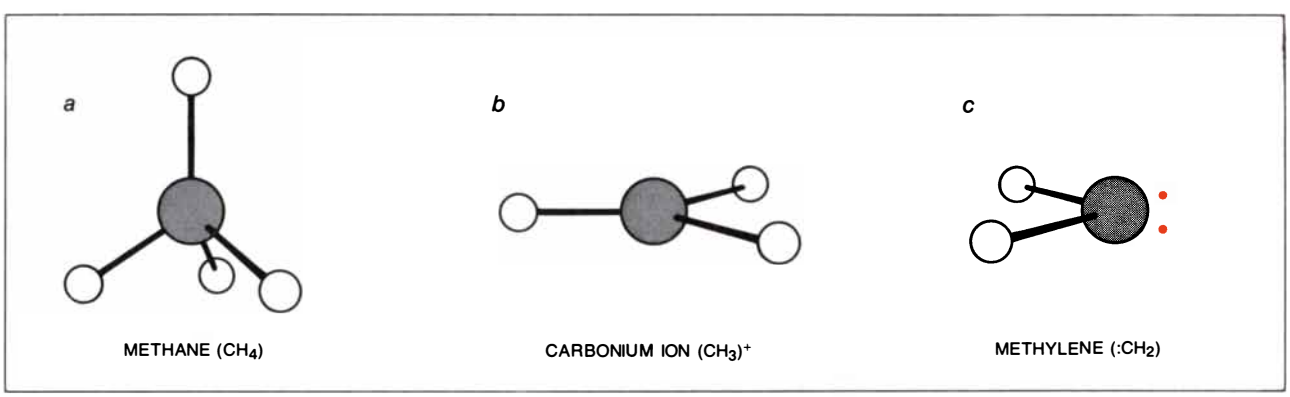
The basic reaction mechanisms of organic chemistry are probed with the aid of these extremely reactive, short-lived molecules, created by depriving a carbon atom of two of its normal bonds

by Maitland Jones, Jr.

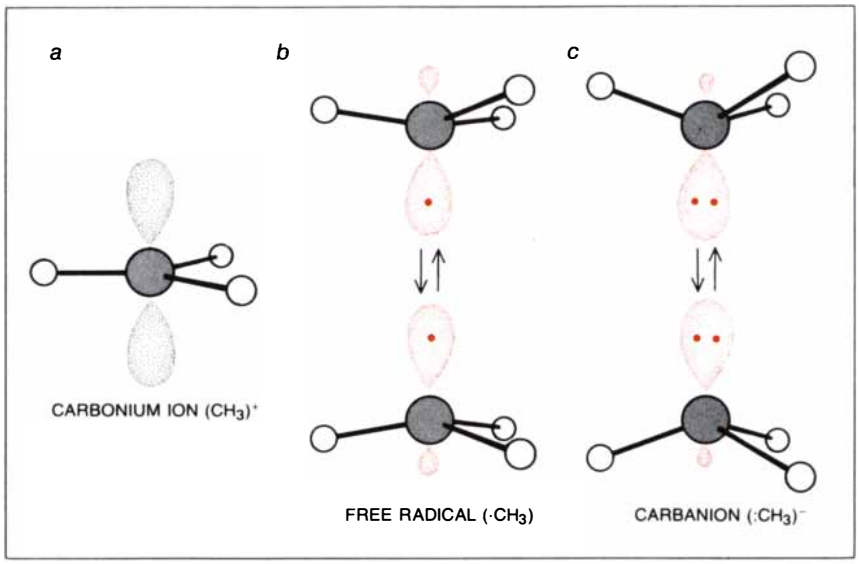
What is most interesting about organic chemistry, it seems to me, is that it involves the chemist directly in the age-old dream of the alchemist: the conversion of one species of matter into another. The investigator can literally watch the changes taking place in the laboratory flask, and to be that close to nature's workings can be immensely satisfying. One is naturally led to wonder how those changes are taking place. Exactly how do molecules come together to form other molecules? Are some molecular shapes better suited than others to engage in chemical reactions? How much energy is required to effect a certain change? Are there intermediate molecular species involved in the conversion from one molecule to another, and if there are, how can one detect and study such short-lived entities?

These days an important subdivision of organic chemistry is concerned with that last question. To be sure, many organic chemists are still busy synthesizing ever more complex molecules, but there is a small group of organic chemists, sometimes called physical organic chemists, which is more concerned with reaction mechanisms. An even smaller group, including my co-workers and me at Princeton University, finds itself focusing on the properties of reactive intermediates: those molecular species that are formed in a chemical reaction only to be consumed almost instantaneously in the creation of new molecules.

Since organic chemistry is basically the chemistry of carbon compounds it is hardly surprising that by far the greatest attention has been paid to reactive intermediates that contain carbon. Normally the carbon atom is tetravalent, that is, it has a tendency to form four covalent, or two-electron, bonds with other atoms, the four bonds generally being directed in space toward the four corners of a tetrahedron. Until recently the commonest reactive intermediates formed in the conversion of one tetravalent-carbon compound into another incorporated trivalent carbon. A typical example is the carbonium ion, in which three atoms placed at the corners of an equilateral triangle surround a central carbon atom bearing a positive electric charge. If an electron is somehow added to a carbonium ion, the positive charge is canceled and a neutral molecule, a free radical, results; in its simplest form this molecule



CARBON ATOM normally forms four covalent, or two-electron, bonds with other atoms, the four bonds generally being directed in space toward the four corners of a tetrahedron (*a*); in this case the four other atoms are all assumed to be hydrogen, and the entire five-atom molecule is called methane. In chemical reactions that involve the conversion of one such tetravalent-carbon compound into another, short-lived intermediate species of carbon that have fewer than four bonds are often formed. For instance, in the carbonium ion (*b*), a typical trivalent-carbon compound, three atoms (again hydrogen in this example) placed at the corners of an equilateral triangle surround a central carbon atom bearing a positive electric charge. By removing two of the atoms normally bonded to the carbon atom one can generate the extremely unstable divalent reactive intermediates known as carbenes; when the two remaining atoms are both hydrogen, the particular carbene shown is called methylene (*c*). In the ball-and-stick diagrams accompanying this article covalent bonds are represented by sticks, carbon atoms by gray balls, hydrogen atoms by somewhat smaller white balls and other atoms (or groups of atoms) by balls with identifying letters inside them; colored dots denote free, or nonbonding, electrons.



CARBONIUM ION CAN BE ALTERED by adding one or two electrons to it, thereby generating two other trivalent carbon species. The region of space in which the added electron (or electrons) is likely to be found is called an orbital and is represented here by the stippled shapes (*a*); occupied orbitals are shown in color. If one electron is added, the positive charge associated with the carbonium ion is neutralized and a free radical results (*b*); this molecule is thought to take the shape of a shallow pyramid that oscillates around an equilibrium point coincident with the carbon atom. The addition of two nonbonding electrons yields the negatively charged carbanion (*c*), which is more pyramidal.

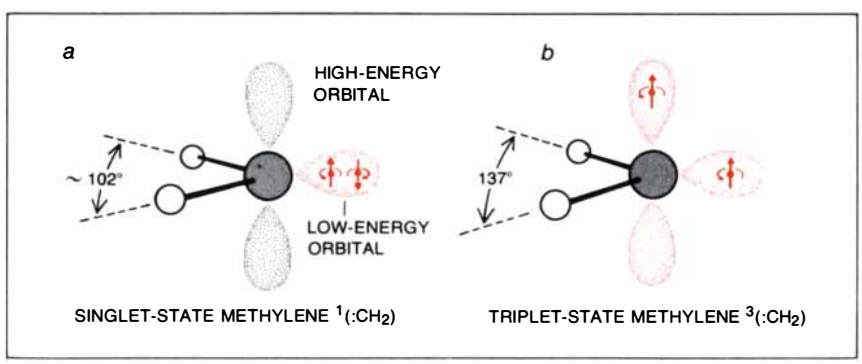
probably takes the shape of a shallow pyramid [*see illustration above*]. The addition of still another electron yields the negatively charged carbanion, which is usually more pyramidal and also oscillates rapidly. The chemistry of these trivalent-carbon compounds has been broadly worked out over the past 40 years or so, although even today it is by no means completely understood.

Interesting and useful as these trivalent species are, one can manipulate the "normal" carbon atom to an even greater degree by contriving to remove two of the atoms ordinarily bonded to it. In doing so one creates the enormously reactive divalent species known as carbenes.

According to legend, it was in the fall of 1951 during a late-night taxi ride in Chicago that three of America's leading organic chemists, William Doering, Saul Winstein and Robert Woodward, jointly conceived the name carbene. Of course, for them to coin the name required that there be something worth naming. Actually the first suggestion that divalent carbon could be a reactive intermediate had been made more than a century earlier, and by the early 1950's the beginnings of more than two decades of intensive work on these extraordinary species had appeared in the chemical literature.

Much of the interest in carbenes derives from their great instability and attendant reactivity. Even carbon-hydrogen bonds, which are among the strongest single bonds to carbon, are ruptured on nearly every encounter with the simplest carbene, methylene (abbreviated $:CH_2$, the two dots before the chemical symbols signifying the two free, or nonbonding, electrons associated with the carbon atom). The reaction of methylene with methane (CH_4) to yield ethane (C_2H_6), for example, releases a considerable amount of heat, a sure indication that the starting material has a much higher energy and hence is much less stable than the product. The existence of such a large energy difference does not by itself mean that a given reaction will be probable or even possible. There may be substantial barriers, represented by energy or entropy requirements, intervening between the high-energy starting material and the comparatively low-energy products. Nonetheless, the picture of a species poised at the top of an energy cliff above its products gives one a sense of the precipitous reactivity of carbenes.

Before looking in detail at the reactions of methylene and other carbenes with various hydrocarbons, a complicating factor must be taken into account. Part of the quantum-mechanical description of the electrons surrounding a carbon atom involves their spin quantum number; the spin of an electron can be assigned a value of either $+1/2$ or $-1/2$, depending (in loose terms) on whether the direction of spin is clockwise or counterclockwise. Accordingly the two nonbonding electrons of a carbene will have a net spin of either 0 (for spins of different sign) or 1 (for spins of the same sign). For spectroscopic reasons having nothing to do with the subject of this article these two conditions of spin are respectively called the singlet state and the triplet state. For any given carbene it follows that one can expect two somewhat different structures. An important question then emerges: Will the chemical reactions of a singlet-state carbene be



CARBENE CAN TAKE TWO FORMS, depending on whether the spins of the two free electrons are in the same direction or in opposite directions. For example, in singlet-state methylene (*a*) the two electrons are paired (that is, spinning in opposite directions), and therefore they are both allowed to occupy the same low-energy in-plane orbital. In triplet-state methylene (*b*), on the other hand, the two electrons spin in the same direction, and as a consequence they are not allowed in the same orbital; instead one electron occupies a higher-energy dumbbell-shaped orbital, which extends above and below the plane of the three atoms. The singlet-state carbene is evidently more bent than the triplet-state one.

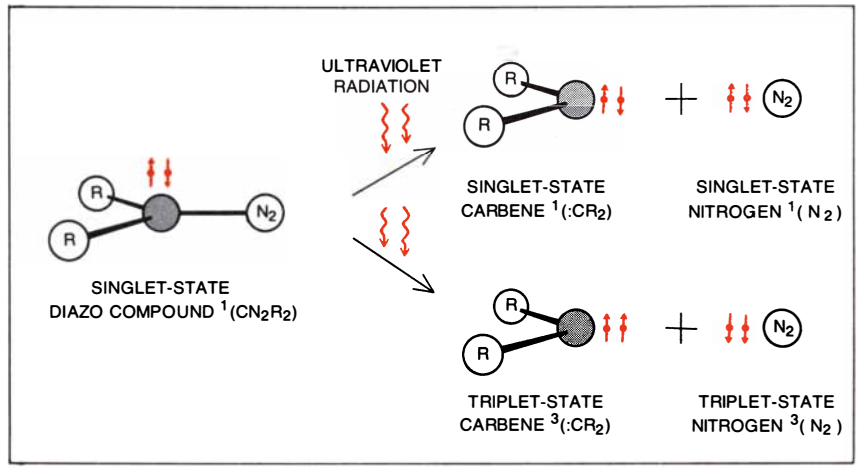
detectably different from those of the corresponding triplet-state carbene?

Methylene is a molecule small enough to be dealt with by the still rather crude theoretical methods available to chemists, and much of what is known about the geometry of singlet- and triplet-state methylene is due to the efforts of theorists. Both molecules are bent, the singlet-state molecule more than the triplet-state one, but the regions of space occupied by the two nonbonding electrons are quite different. These regions of space are called orbitals. In general an electron will occupy the lowest-energy orbital available. The catch is that two electrons with the same spin may not occupy the same orbital. Therefore in the case of triplet-state methylene the two nonbonding electrons are in different orbitals: one is in a comparatively low-energy orbital that lies in the plane of the three atoms, whereas the other occupies a dumbbell-shaped orbital of somewhat higher energy extending above and below the plane of the three atoms [*see bottom illustration on opposite page*]. In the singlet state, where the two electrons have oppositely directed spins, there is no prohibition against double occupancy of an orbital, and here the two electrons are both in the same low-energy in-plane orbital.

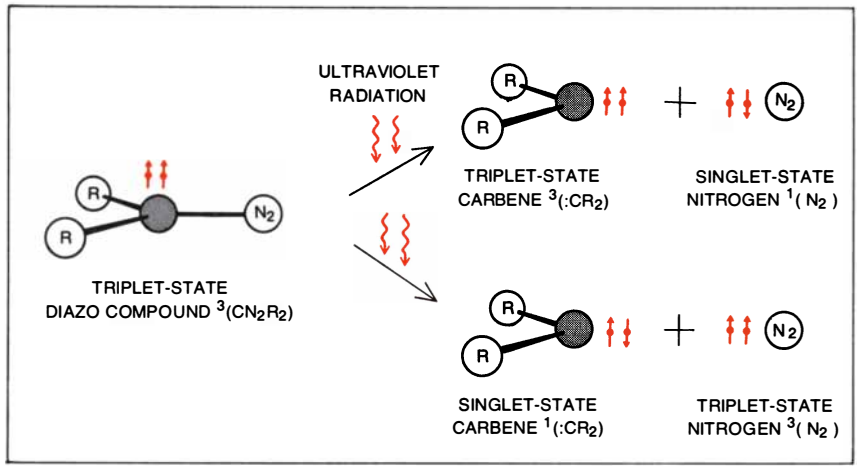
For carbenes where the atoms (or groups of atoms) attached to the divalent carbon atom are more complex than they are in methylene the structures may vary, but the overall features of a sharply bent singlet-state molecule and a more nearly linear triplet-state one persist. The exact magnitude (and even the direction) of the difference in energy between the two spin states has been the subject of a prodigious amount of work and some controversy over the years. It now seems clear that with a few exceptions the triplet states are slightly lower in energy than the singlet states. Thus one of the major questions concerning carbene chemistry reduces to whether or not the subtle change from the singlet state to the triplet state—a change involving only the direction of spin of a single nonbonding electron—would result in two different chemistries.

Let us now look at how carbenes are made in the laboratory. Although there are a number of ways to generate divalent-carbon compounds, we shall concern ourselves here only with the simplest, if most hazardous, method. That approach involves the decomposition by heating or by ultraviolet radiation of the brightly colored nitrogen-containing substances known as diazo compounds. These poisonous, carcinogenic and capriciously explosive compounds give up molecular nitrogen to form carbenes, which must then find some way to normal tetravalency.

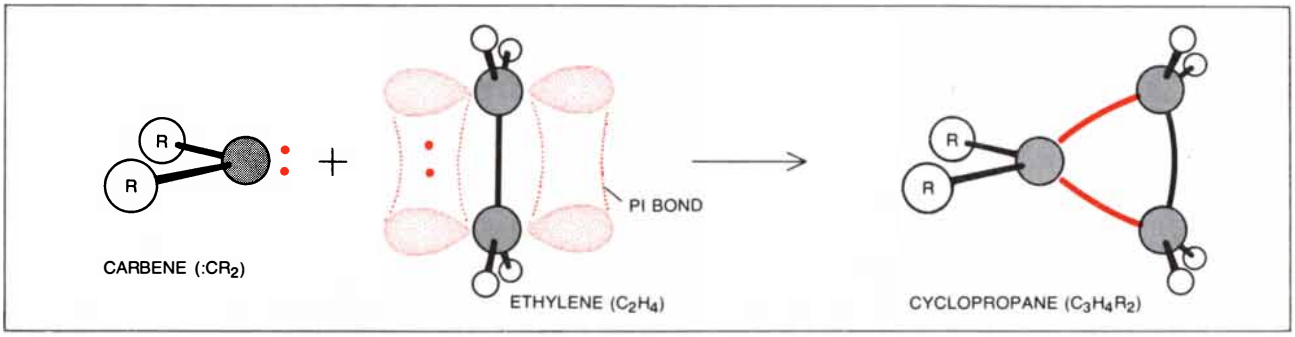
The photochemical decomposition of diazo compounds is a particularly useful means of generating carbenes because it lends itself to the formation of both singlet- and triplet-state carbenes. The direct decomposition of a diazo compound such as diazomethane (H_2CN_2), for example, yields a nitrogen molecule (N_2)



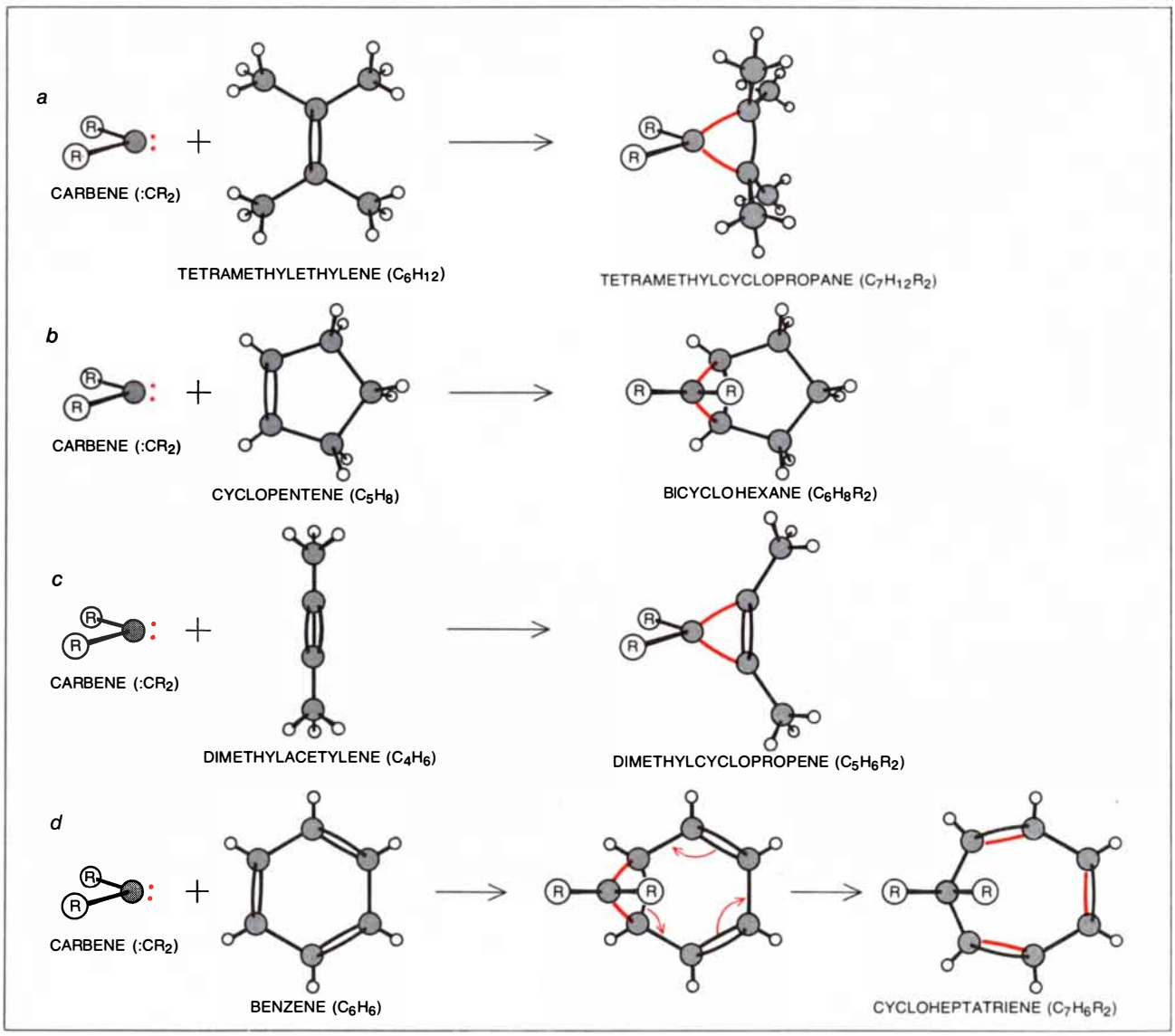
SINGLET-STATE CARBENES are formed by the direct photochemical decomposition of the nitrogen-containing substances known as diazo compounds, which normally exist in the singlet state (that is, with all their electrons forming pairs with oppositely directed spins). Since electron spin must be conserved in the reaction, one must end up with products in which there is no net unpaired spin. This means that when a singlet-state compound such as diazomethane (H_2CN_2) is exposed to ultraviolet radiation to form methylene ($:CH_2$) and molecular nitrogen (N_2), for example, the products must be either singlet-state carbene and singlet-state nitrogen or triplet-state carbene and triplet-state nitrogen. (The symbol *R* is used here to denote any of a variety of atoms or groups of atoms other than hydrogen that can be attached to a divalent carbon atom to form the carbene.) Actually the second route is extremely unlikely, since the energy required to form triplet-state nitrogen molecules is much higher than that required to form singlet-state ones. Hence the direct ultraviolet irradiation of ordinary diazo compounds yields singlet-state carbenes.



TRIPLET-STATE CARBENES are formed indirectly by the decomposition of diazo compounds that have first been converted from the singlet state to the triplet state by the introduction of a special photosensitizing agent. In order for electron spin to be conserved in the ensuing reaction, the products this time must be either triplet-state carbene and singlet-state nitrogen or singlet-state carbene and triplet-state nitrogen. The high energy of the triplet-state nitrogen here strongly favors the former route, and thus the reaction yields mostly the triplet-state carbene. Orbitals have been omitted in both diagrams on this page for the sake of clarity; accordingly the indicated positions of the electron-spin symbols (*colored arrows*) do not necessarily correspond to the actual positions of the electrons.



CARBENE CAN BE ADDED to almost any carbon-carbon multiple bond to form a cyclopropane: a compound that has three carbon atoms in a ring. In this example a carbene is shown reacting with the pi, or "double," part of the double bond in a molecule of ethylene, the simplest multiple-bonded hydrocarbon. The two electrons that constitute the pi bond combine with two nonbonding electrons on the carbene to provide the four electrons needed to make two new carbon-carbon bonds in cyclopropane product.



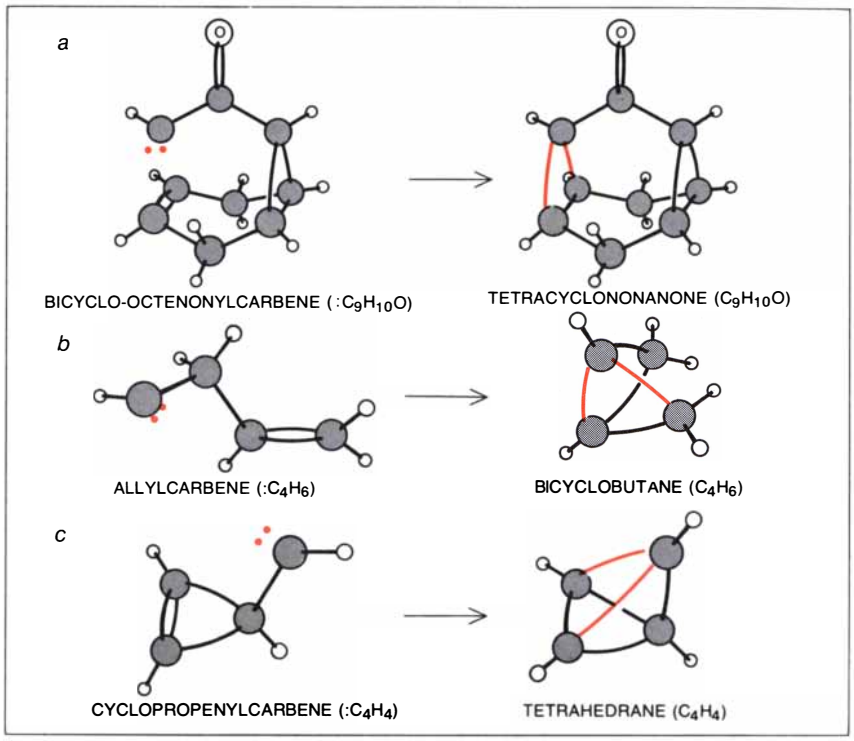
VARIETY OF CYCLOPROPANES that can be synthesized in reactions involving carbenes and multiple-bonded hydrocarbons is represented here by several examples. In *a* a carbene is shown attacking the double bond in tetramethylethylene to form a cyclopropane ring with four methyl (CH_3) groups attached to it. In *b* a carbene reacts with a carbon-carbon double bond in a five-carbon ring to produce a compound with two rings. In *c* a carbene reacts with a carbon-carbon triple bond to yield a very highly strained cyclopropene ring in which one of the original pi bonds remains untouched and reappears in the final product as a double bond. In *d* a carbene forces the normally very stable six-carbon ring of benzene to expand through an unstable cyclopropane intermediate into a seven-carbon ring. The small internal angles of cyclopropanes make them difficult to synthesize in more conventional ways.

and a methylene molecule (:CH_2), both in the singlet state. Spin must be conserved in such reactions. Hence if one starts with a singlet-state diazo compound, in which all the electrons form pairs with oppositely directed spins, one must end up with products in which there is no net unpaired spin. This constraint means that the products must be either two singlet-state molecules or two triplet-state molecules [*see upper illustration on page 103*]. Since the energy required to form triplet-state nitrogen molecules is much higher than that needed to form singlet-state ones, the alternative all-triplet-state route is actually never followed.

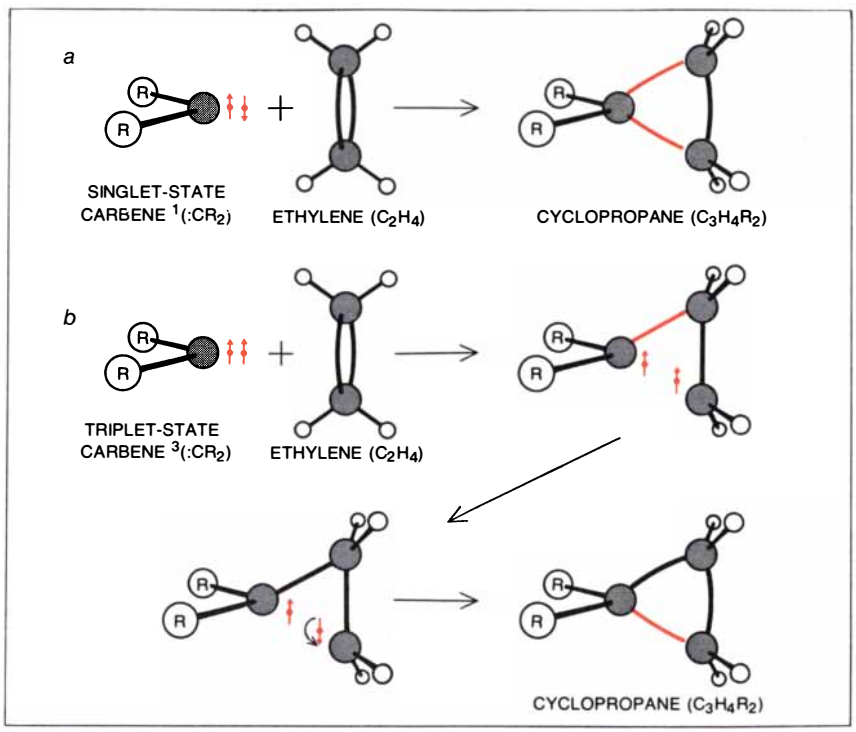
If instead of directly irradiating the diazo compound, however, one first introduces a photosensitizing agent to absorb all the incident radiation, a triplet-state diazo compound can be produced. Again two possibilities for further reaction exist: this time a triplet-state carbene can be formed along with a singlet-state nitrogen molecule, or a singlet-state carbene can be formed with a triplet-state nitrogen molecule [*see lower illustration on page 103*]. Here the high energy of the triplet-state nitrogen works to make the former possibility by far the likelier route.

In short, although quite dangerous compounds must be employed, the availability and versatility of diazo compounds makes them the precursors of choice in the generation of both singlet- and triplet-state carbenes. (There is even an indirect way around the dangers, since the salts that serve as the immediate precursors of diazo compounds can often be used instead. In that case the diazo compound is simply generated where it is to be consumed in the reaction vessel, and it need never be isolated.)

Carbenes have a voracious appetite for electrons. Since they lack two of their normal complement of valence, or bonding, electrons, they react rapidly with even very weak sources of electrons to form tetravalent-carbon compounds in which each carbon atom has a share in eight valence electrons (four of its own and four contributed by the atoms it is bound to). In the laboratory the most common source of electrons to "feed" carbenes is the pi, or "double," part of a carbon-carbon double bond; the product of this reaction is a cyclopropane: a compound with three carbon atoms in a ring [*see top illustration on opposite page*]. The narrow internal angles of those highly strained small-ring compounds make them quite high in energy



INTRAMOLECULAR REACTIONS, in which a carbene adds to a multiple bond within the same molecule, lead to a variety of cage, or multi-ring, compounds. The last example, synthesis of tetrahedrane from cyclopropenylcarbene, has not yet been achieved in laboratory.



DIRECTION OF ELECTRON SPIN affects the mechanism by which carbenes add to double bonds to produce cyclopropanes. In the case of a singlet-state carbene (*a*) the two new bonds are formed directly, since the spins of the four electrons involved in the reaction (two on the carbene and two in the pi part of the carbon-carbon double bond) are properly aligned for the simultaneous formation of two bonds. In the case of a triplet-state carbene (*b*), on the other hand, the alignment of the electron spins does not allow two new bonds to form in a single step. Once the first bond is formed the remaining two electrons have the same spin, and hence they cannot occupy the same orbital to make a covalent bond. One of the electrons must first "flip" its spin before the cyclopropane ring can be closed.

and hence difficult to synthesize in more conventional ways. In a sense the carbene serves as an energy source to make a rather inaccessible class of molecules. Today if an organic chemist is faced with the task of synthesizing a cyclopropane, it is almost certain that his thoughts will first turn to carbenes.

Almost any kind of carbon-carbon multiple bond will react with divalent carbon. For example, on encountering a carbene a simple double-bonded hydrocarbon such as ethylene (C_2H_4) or tetramethylethylene (C_6H_{12}) reacts to yield three-carbon cyclopropane rings; carbon-carbon double bonds in hydrocarbon rings react to yield compounds with two rings; carbon-carbon triple bonds react to give very highly strained double-bonded cyclopropene rings, and the six-carbon ring of benzene is expanded through an intermediate cyclopropane into a seven-carbon ring [*see bottom illustration on page 104*].

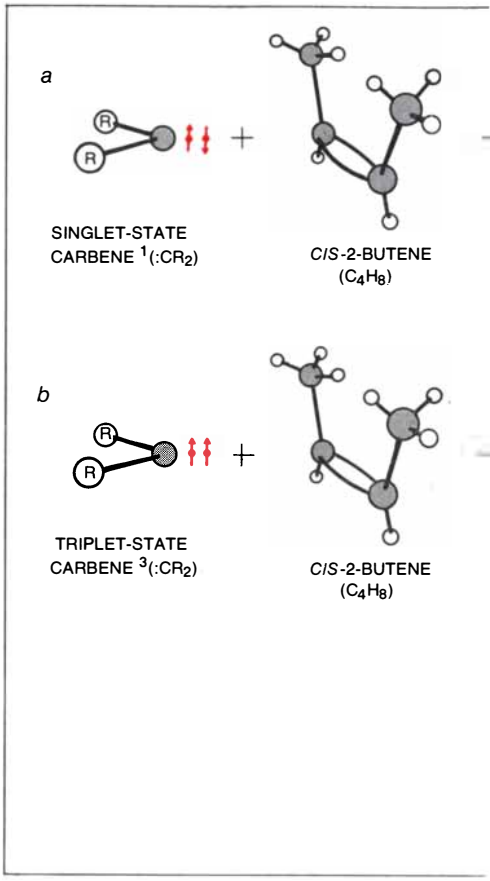
Such intermolecular reactions involving carbenes provide ready access to an immense variety of compounds. Nonetheless, much attention has also been focused lately on the intramolecular version of cyclopropane formation. To produce such a reaction one must manage to put both the carbene and the carbene-accepting multiple bond within the same molecule and in reasonable proximity to each other [*see top illustration on preceding page*]. This internal-addition reaction can lead to highly strained "cage" compounds with many rings. If the carbene and its target double bond are separated by only a single carbon atom, the product is bicyclobutane (C_4H_6), a molecule consisting of two fused three-carbon rings. Even more exotic ring compounds can be produced if the carbene-accepting multiple bond is itself part of a hydrocarbon ring.

An intramolecular reaction involving cyclopropenylcarbene ($:C_4H_4$) would represent the culmination of the cyclopropane ring-building business. If this carbene were to undergo internal addition, the product would be the hypothetical tetrahedral molecule tetrahedrane (C_4H_4). Although the reaction has been tried many times, tetrahedrane has so far eluded isolation. The precise reason for this baffling failure is not known, but preliminary results suggest that tetrahedrane may form momentarily and then react further before it can be detected.

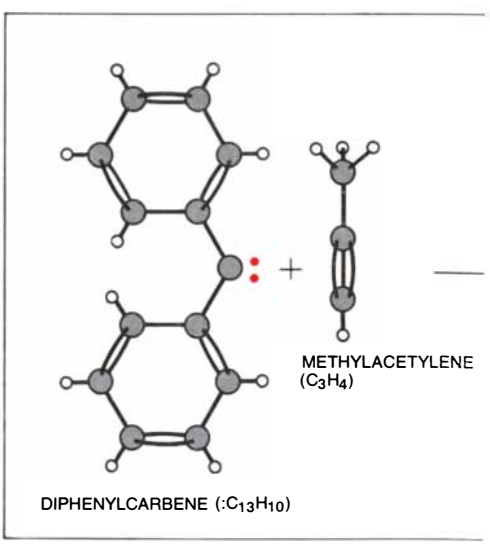
The wealth of experimental data on the cyclopropane-forming process provides surprisingly little information on the mechanism of the reaction. For example, few clues are given on how the atoms approach one another in space or how they twist and turn during the reaction. The question of the reaction mechanism was first attacked by experimenters, but their findings have been greatly augmented in recent years by theorists, who have provided a quite detailed picture of how the reaction proceeds.

The first question to be examined was the one of whether the two new carbon-carbon bonds created in the cyclopropane-forming reaction are formed simultaneously or one at a time. It was recognized quite early by Philip Skell of Pennsylvania State University and by Doering of Yale University that the answer to the question might not be the same for both singlet- and triplet-state carbenes. For example, in the case of a singlet-state carbene the spins of the electrons involved in the reaction (two on the carbene and two in the pi part of the carbon-carbon double bond) are properly aligned for the formation of two bonds, each of which must have two electrons with paired spins [*see bottom illustration on preceding page*]. That is not the case for a triplet-state carbene, however. There a direct reaction to form two new bonds in a single step is not possible, because the alignment of the electron spins does not allow it. An intermediate product that has a net unpaired electron spin must be formed first. Such an intermediate might well revert to cyclopropane if the spin of one electron could be changed. Mechanisms for such a "spin flip" exist in nature, although the time required for the flip is not well known. The lifetime of the presumed intermediate molecule is also of some importance. It does seem that if such an intermediate were to be formed, a cyclopropane would be the inevitable product.

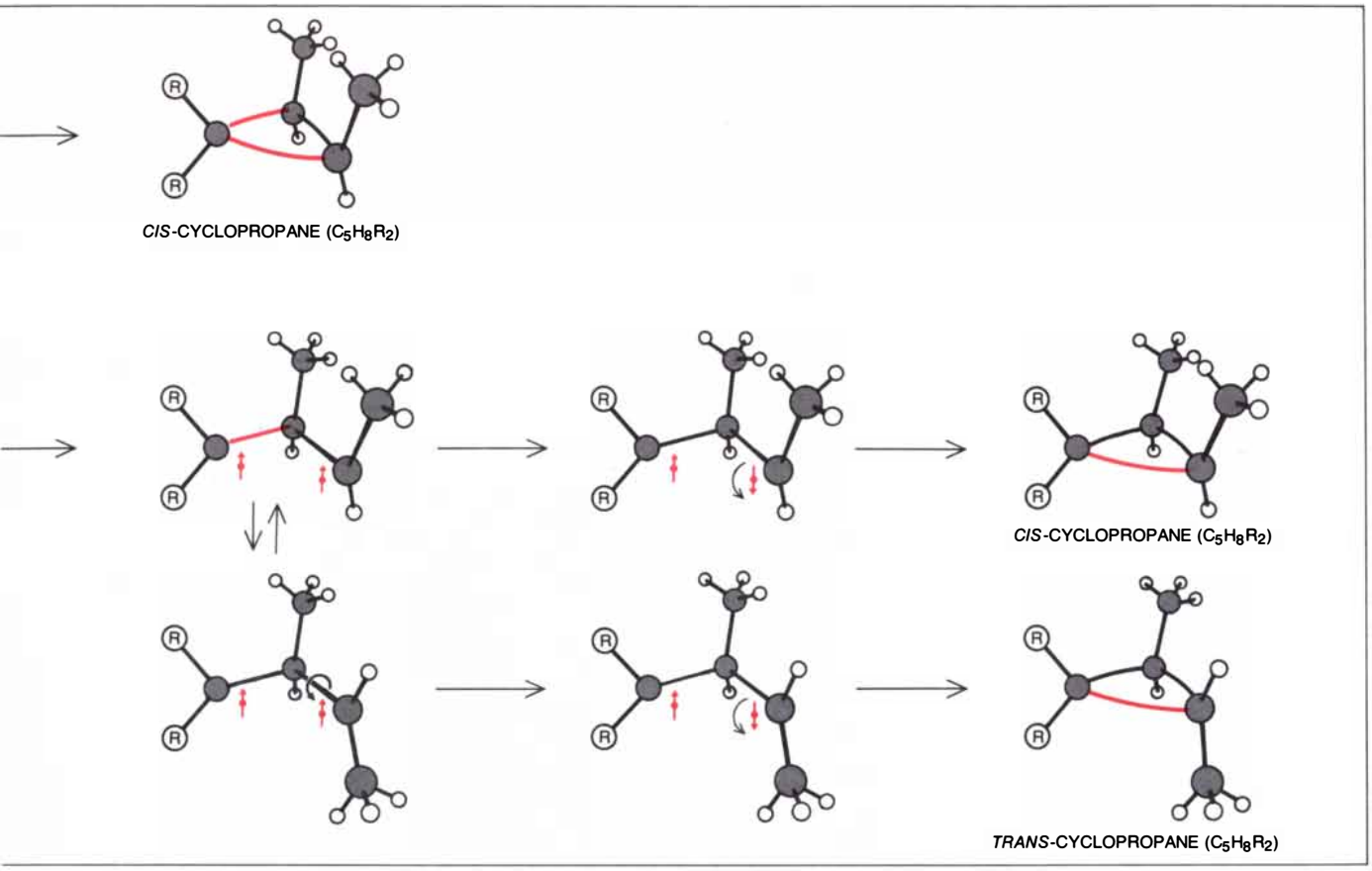
A simple experiment was available to differentiate the two processes: one involving an intermediate and the other not involving one. The experiment rested on the premise that in the intermediate molecule the stereochemistry, or three-dimensional arrangement, of two methyl groups (CH_3) attached to the original double bond in one of the two possible configurations of the 2-butene molecule would be lost because of the rapid rotation known to occur around the carbon-carbon single bond. In the singlet case the simultaneous formation of the two new single bonds would not allow any such rotation and the original stereochemical relations would be preserved. Hence the picture emerged of two slightly different mechanisms: one a singlet



STEREOCHEMISTRY, or three-dimensional arrangement of atoms in a molecule, is either preserved or lost in the cyclopropane-forming reaction, depending on whether the spins of the two nonbonding electrons on the carbene are aligned in opposite directions (*a*) or in the same direction (*b*). Be-

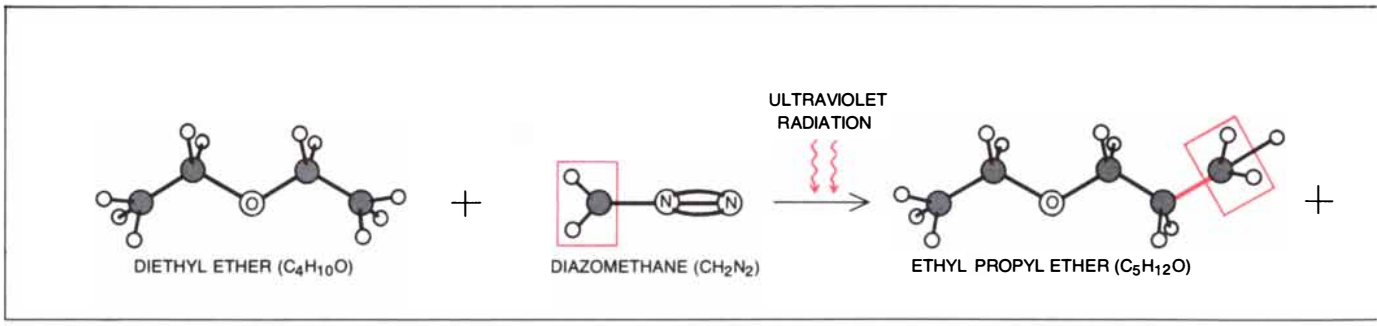


UTILITY OF CARBENES for synthesizing new molecules can be extended beyond the cyclopropane-forming reaction by virtue of the fact that the intermediate formed by the addition of a triplet-state carbene to a mul-



cause the singlet-state carbene can add to the carbon-carbon double bond in a molecule such as 2-butene in a single step to form two new bonds at once, the stereochemical relation of the two methyl (CH_3) groups attached to the double bond is not changed in the reaction. If they start out on the same side (as in the *cis* configuration shown), they will end up on the same side; if they start out on opposite sides (the *trans* configuration), they will end up on opposite sides. In the case of the triplet-state carbene, in contrast, the two new bonds cannot be formed at once, and rotation around the carbon-carbon single bond is possible in the intermediate formed first. By the time an electron has changed its spin, allowing the cyclopropane ring to close, the spatial relation of the methyl groups attached to the original double bond can change. Thus in the latter case the original orientation of the groups is lost.

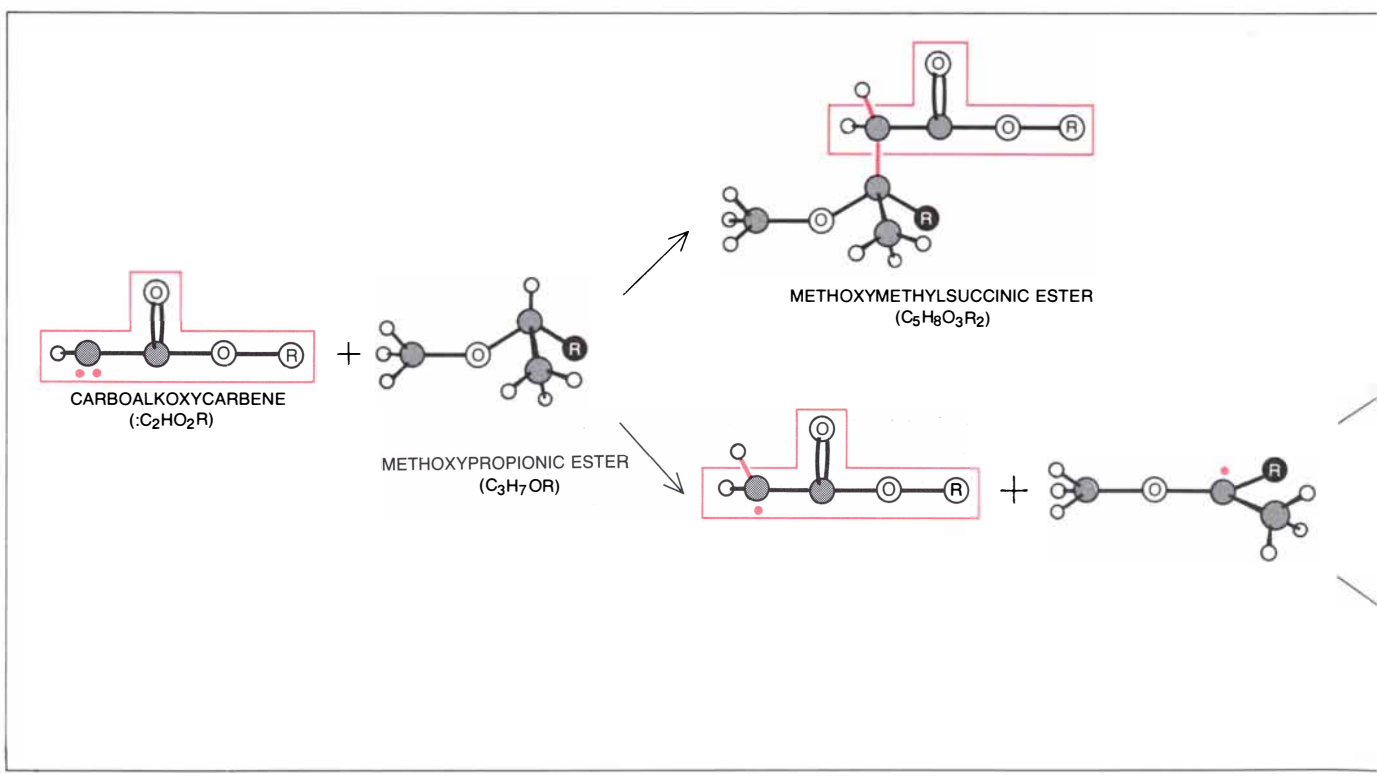
tiple bond need not always close to yield a three-carbon ring. In the addition of triplet-state diphenylcarbene to methylacetylene, for example, the added group instead joins a nearby benzene ring, yielding a molecule with no free electrons, since the two electrons in the double bond initially attacked combine with the two nonbonding electrons in the intermediate to form two new single bonds. The reaction is completed by the migration of a hydrogen atom to give the final product: a member of the class of compounds known collectively as phenylindenes. The addition of singlet-state diphenylcarbene (*not shown*) would give a cyclopropene.



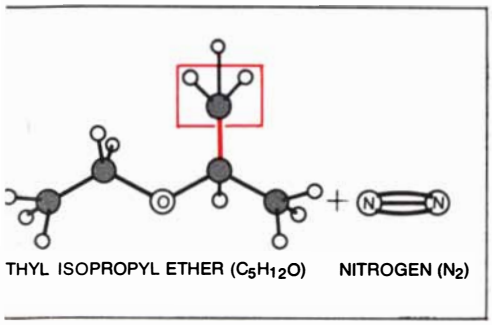
CARBENE CAN BE INSERTED between a carbon atom and a hydrogen atom in a molecule of diethyl ether to generate two isomers, or structurally different versions, of ethyl propyl ether. This reaction, first observed more than 30 years ago, was not recognized until much later as a direct carbene-insertion reaction. Although carbon-hydrogen bonds are among the strongest single bonds to

process where two new chemical bonds are formed directly and the other a triplet reaction proceeding through an intermediate capable of losing the stereochemical relations present in the original molecule [*see top illustration on preceding two pages*]. In other words, a singlet-state carbene can be expected to add to a double bond and retain stereochemistry, whereas a triplet-state carbene should give rise to a mixture of stereochemically differing cyclopropanes. Indeed, in the latter case the final ratio of the two types of cyclopropane should be independent of the stereochemistry of the starting material if the intermediate lives long enough.

Stereospecific carbene reactions (that is, reactions that preserve the stereochemical relations) were commonplace by the time Skell and Doering put forward their concepts, and a small number of nonstereospecific reactions became known shortly thereafter. The assignment of spin state on stereochemical grounds alone became routine. Actually it made little sense to designate one carbene a triplet-state carbene and another a singlet-state carbene without knowing the reactions of both spin states of each carbene. The differences in properties observed might be the result simply of the differences inherent in the properties

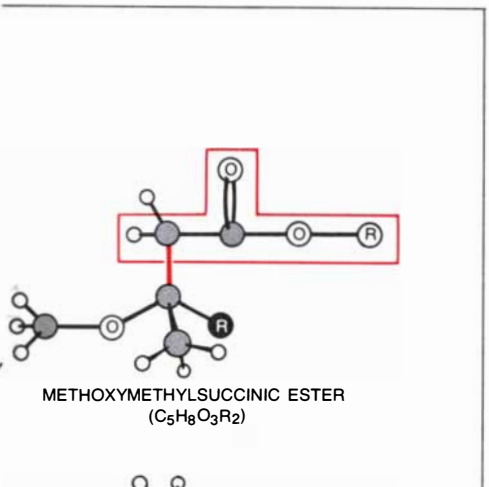


OPTICAL ACTIVITY IS PRESERVED in the direct insertion reaction of a singlet-state carbene such as carboalkoxycarbene into a carbon-hydrogen bond in an optically active molecule such as methoxypropionic ester, which can exist in two forms: one that rotates the plane of plane-polarized light to the left and one that rotates it to the right. In the direct process (*first upper branch*) the particular spatial relation of the groups attached to the key carbon atom is not changed by the reaction, because the two new bonds are formed simultaneously, just as they are in the case of a singlet-state addition reaction. If the process were to involve the removal of the hydrogen atom to produce a pair of intermediate free radicals, however (*first lower branch*), optical activity would be lost



carbon, they are ruptured on nearly every encounter with methylene, shown here (*colored box*) as part of its precursor compound.

of the carbenes and have nothing at all to do with electron spin. Curiously, years passed until the critical experiment was done, but in the end the surmises of Skell and Doering were confirmed. When the stereochemistry of the addition reaction of singlet and triplet states of the same carbene was examined, it turned out that the singlet-state carbene does add specifically and that the triplet-state carbene does not.



owing to the rapid inversion of the intermediate products. Recently the triplet-state counterpart of this reaction (a triplet-state diphenylcarbene insertion) was observed; it was found to follow the latter course, losing the optical activity of the starting material.

Theorists have since provided a much more detailed picture of that reaction. Modern methods of calculation make it possible to explore the potential-energy surfaces for the addition of both singlet- and triplet-state methylene to ethylene, the simplest double-bonded hydrocarbon. The results support the rather meager data supplied by experiment, particularly with regard to the triplet-state addition reaction. What was new was that the singlet-state addition reaction was also found to proceed with the formation of the new bonds one at a time. Thus the singlet- and triplet-state processes are not quite as different as they appeared to be at first.

At Princeton we have tried to exploit the step-by-step addition of triplet-state carbenes in order to expand the utility of carbenes beyond the cyclopropane-forming reaction. To do this we have tried to construct systems where the initially formed intermediate molecule could be trapped internally. A typical example involves the addition of triplet-state diphenylcarbene ($:C_{13}H_{10}$) to methylacetylene (C_3H_4) [*see bottom illustration on pages 106 and 107*]. A singlet-state carbene would simply form a cyclopropene in this reaction, but here the addition of the triplet-state carbene takes place at the less crowded end of the triple bond to yield an intermediate molecule. This species adds internally to one of the benzene rings instead of reacting to form the highly unstable cyclopropene. The reaction terminates with the shift of a single hydrogen atom to regenerate a very stable benzene system. Somewhere in the reaction the spin of an electron must be reversed, but one cannot be certain at exactly what point the reversal occurs. What we have done is to divert the reaction from the formation of a three-carbon ring and develop instead a new approach to the synthesis of a class of molecules known collectively as phenylindenes.

What if a carbene is generated where no addition reaction is possible? Multiple-bonded systems are after all quite open to attack, since the electrons in the pi portion of a double or triple bond are loosely held. Carbenes are so reactive, however, that even the electrons in carbon-hydrogen bonds, which are tightly held indeed, can be attacked. In fact, one of the earliest carbene reactions known was derived from an observation made by Hans Meerwein of the University of Marburg in 1942, namely that ethyl propyl ether and ethyl isopropyl ether (two isomers, or structurally different molecules with the same chemical formula, in this case $C_5H_{12}O$) can be formed by irradiating a solution of diazomethane (CH_2N_2) in diethyl ether ($C_4H_{10}O$). Although it became clear only years later, the two products are created by the interposition of a methylene unit between the carbon and the hydrogen of a particular carbon-hydrogen bond [*see top illustration on these two pages*]. This process, called an insertion reaction, was later found to be general when several carbenes were observed to insert themselves into all kinds of carbon-hydrogen bonds.

In the insertion reaction, as in the addition reaction, the direction of electron spin is vitally important in determining the course of the reaction. Again a simple counting of spins will suffice to show that it is possible (although not necessary) for a singlet-state carbene to form two new bonds at a single stroke, whereas a triplet-state carbene cannot do so and must proceed to the product in a multistep process. As in the case of the addition reaction, the broad features of the mechanism of the insertion reaction were made clear by experiments and later refined by theoretical treatments, which have provided descriptions of the details of the mechanism.

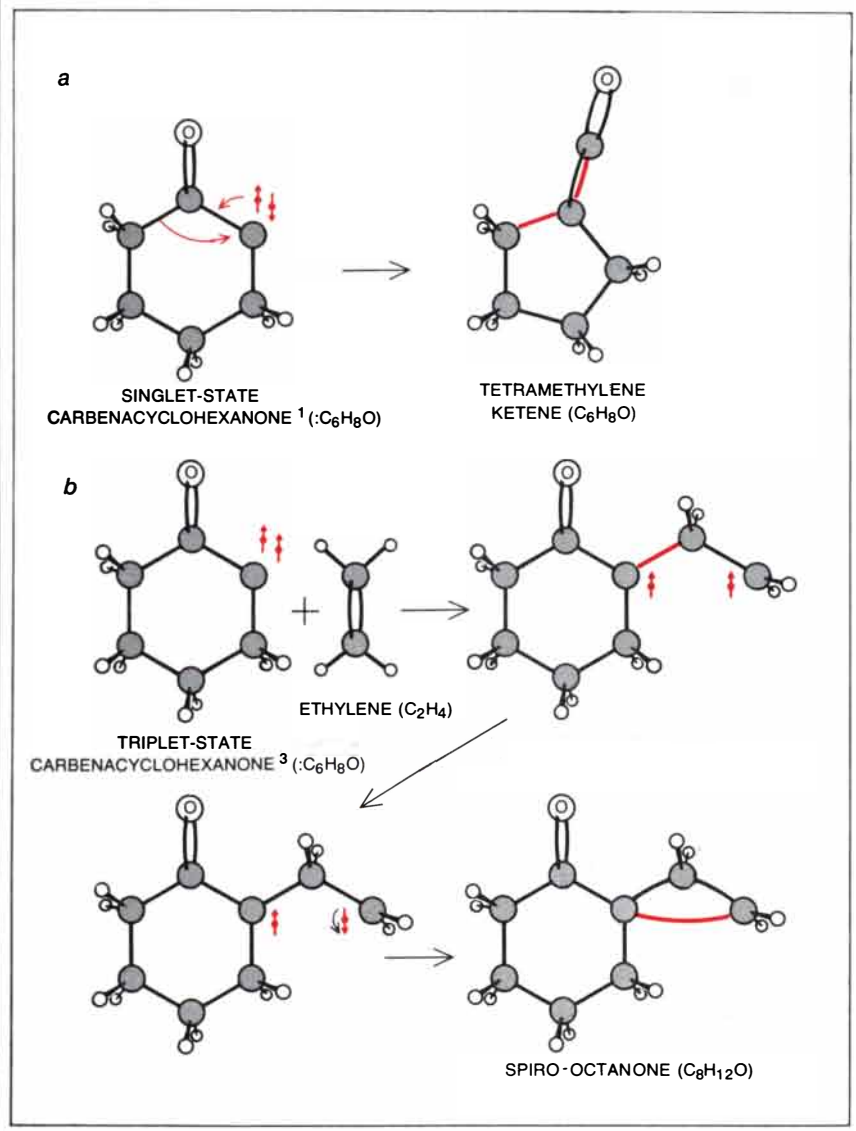
Many organic molecules exist in left-handed and right-handed forms. Such stereoisomers are mirror images of each other and are identical in all chemical and physical properties except for their ability to rotate the plane of plane-polarized light: one isomer will rotate the plane to the left and the other will rotate it to the right. Such molecules are said to be optically active.

The direct insertion of a carbene into a certain carbon-hydrogen bond in an optically active molecule such as methoxypropionic ester would be expected to yield a molecule that would be still optically active. If the mechanism were to involve the removal of the hydrogen atom, however, optical activity would inevitably be lost because of the rapid inversion of the resulting free radicals [*see bottom illustration on these two pages*]. Several singlet-state carbene insertion reactions are known to result in the retention of optical activity. Thus the mechanism of the insertion reaction of singlet-state carbenes involves the formation of the two new bonds with no intervention of discrete free radicals. Recently the triplet-state counterpart of the reaction was also observed. The insertion reaction of triplet-state diphenylcarbene was found to result in the complete loss of optical activity. Thus it is clear that

the triplet-state insertion reaction is not direct but involves a pair of intermediate free radicals.

Again theoretical work has confirmed the mechanism of the triplet-state insertion reaction. The description of the singlet-state insertion reaction, however, is different in detail from the general picture provided by experiment. It is impossible by experiment alone to tell a truly direct process, in which the two new bonds are formed concurrently, from a hydrogen-removal process, which leads to a pair of radicals that cannot escape from each other. Theory draws exactly that picture of the reaction: a pair of radicals is formed, but the members of the pair are close enough at the moment of their creation to be inevitably committed to collapse together to form a single product.

The intramolecular insertion reaction of singlet-state carbenes is so rapid that it completely dominates many chemical reactions. It is difficult to use certain singlet-state carbenes in intermolecular reactions simply because these unstable species react so quickly with nearby internal bonds. Internal reactions require no encounters with external molecules



INTRAMOLECULAR INSERTION REACTION of a singlet-state ketocarbene with an adjacent carbon-carbon single bond results in the extremely fast Wolff rearrangement (*a*), which renders many carbenes useless for synthesis, since reactions with other molecules are too slow to compete. The problem can be avoided by changing the spin of one nonbonding electron to form the triplet state of the carbene (*b*). Now the Wolff rearrangement is not so favorable, and intermolecular reactions (in this case with ethylene) can proceed. Thus triplet-state ketocarbenes can be used to generate conventional cyclopropane products.

and are usually faster than their intermolecular counterparts. Many carbenes have been found to be quite useless for synthesizing new molecules simply because they refuse to participate in intermolecular reactions of any kind. It seemed to us that for a variety of reasons triplet-state carbenes should not be as reactive toward internal bonds as singlet-state carbenes are.

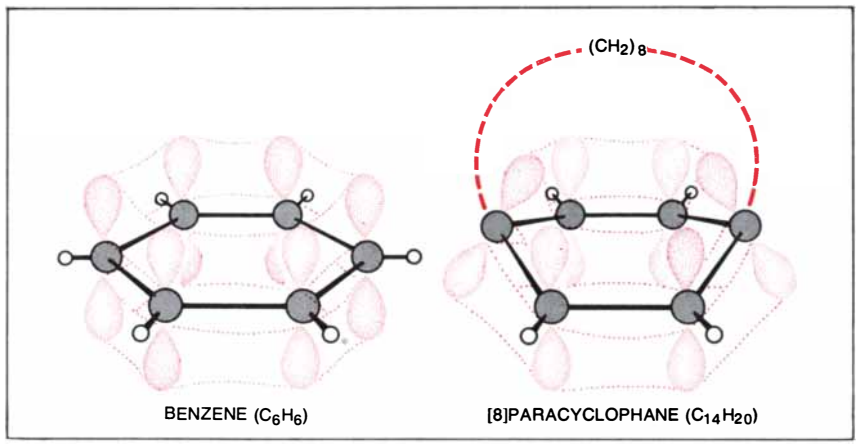
A typical example of the rapid internal reaction of singlet-state carbenes is the insertion of ketocarbenes (carbenes incorporating an oxygen-carbon double bond) into an adjacent carbon-carbon bond to give rise to ketenes [*see illustration on opposite page*]. Even if one generates most ketocarbenes in a carbene-accepting solvent, the intramolecular reaction (called the Wolff rearrangement) is so fast that no useful amounts of addition products are formed. That is not the case with triplet-state carbenes. The generation of triplet-state ketocarbenes with the aid of ultraviolet radiation leads not to ketenes but rather to the conventional cyclopropane addition products. The formation of the triplet state has changed the carbene from one that is prone to unproductive internal reactions into a much more useful species.

To summarize, two kinds of divalent carbon exist. There are singlet-state carbenes, which react rapidly, and without forming detectable intermediates, with all manner of systems incorporating pairs of electrons, and there are triplet-state carbenes, which often lead to the same gross structures as the singlet-state species do but by very different mechanisms. The subtle interplay between theory and experiment has yielded a reasonable understanding of the reactions of carbenes, which should enable future investigators to make greater use of divalent carbon. One can see the beginnings of the new chemistry even today. Not only have carbenes provided an efficient way to synthesize a variety of cyclopropanes and cage compounds but also they are beginning to be useful in making far more exotic molecules. When one calls a molecule exotic, it reflects the feeling that in some sense the molecule is out of the ordinary; it is by studying such molecules that organic chemists seek revelation about chemical bonding.

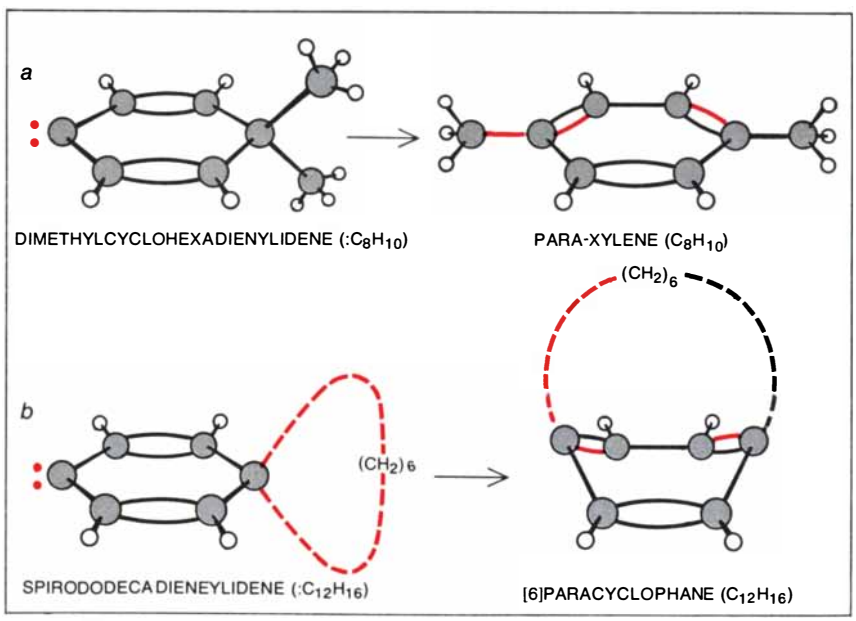
The orbitals that are involved in bonding atoms to each other are reasonably well defined, and to a first approximation the more two orbitals on different atoms overlap, the more strongly will the atoms be bound together. In the case of

a simple double-bonded hydrocarbon such as ethylene, for example, the two dumbbell-shaped orbitals overlap to form the pi part of the double bond. The benzene molecule consists of six carbon atoms in a ring bound together by a particularly stable array of six mutually overlapping dumbbell-shaped orbitals. What will happen if one begins to twist those orbitals with respect to one another? How much twisting and attendant decoupling of electrons will benzene accept before it ceases to be stable? What reactions will a twisted benzene ring enter into? Can one eventually derive a correlation between the angle of twist and the chemical reactions available to a molecule? These are typical of the kinds of question physical organic chemists would like to answer.

In principle the task of synthesizing twisted benzene rings is not a very difficult one. Chemical species known as paracyclophanes had been made in the 1950's and 1960's by several groups, notably those led by Donald J. Cram of the University of California at Los Angeles and by Norman L. Allinger, who was then at Wayne State University. What these chemists and others had managed to do was to bridge two opposite positions of a benzene ring with a chain of carbon atoms [*see upper illustration on this page*]. The benzene ring of the resulting paracyclophane will be forced to bend, provided that the bridge is short enough. There lies the rub, since conventional syntheses had led only to paracyclophanes with bridging chains as short as eight carbon atoms, and they had led to them only after extraordinary efforts. A bridge of eight or more carbons is not enough to warp the benzene ring significantly, and a source of smaller paracyclophanes was sorely needed.



BENZENE MOLECULE owes its great stability to the efficient overlap of six dumbbell-shaped orbitals held parallel in a regular hexagonal ring (*left*). If two opposite positions on a benzene ring were to be bridged with a short chain of carbon atoms (*right*), the orbitals would no longer overlap efficiently and the resulting twisted benzene compound, called a paracyclophane, would be expected to become much more reactive. Conventional methods of synthesis, however, had until recently led only to paracyclophanes with bridging chains as short as eight atoms, not enough to warp the benzene ring significantly.



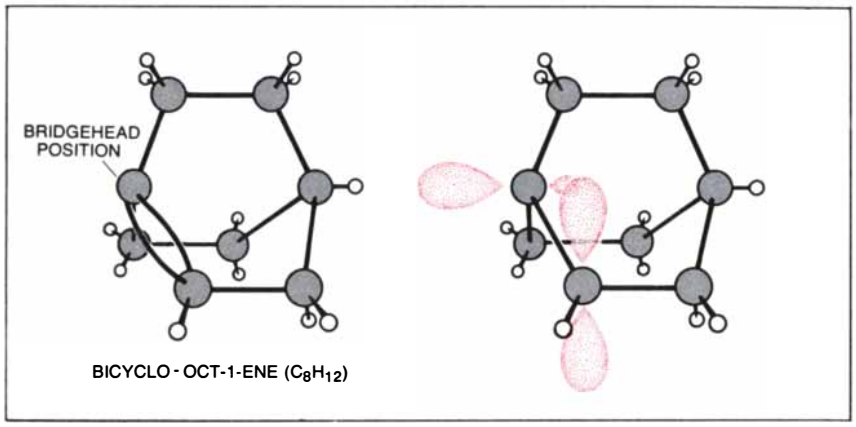
CARBENES PROVIDE a new and convenient route to the synthesis of small, highly twisted paracyclophanes. The discovery that a carbene with the full chemical name of 4,4-dimethylcyclohexa-2,5-dienylidene rearranged itself in the gas phase to form 1,4-dimethylbenzene, or para-xylene (*a*), led by a simple extension of the reaction to the successful synthesis of bent paracyclophanes with bridging chains composed of as few as six carbon atoms (*b*).

At that time our group was engaged in a program of examining the intramolecular reactions of carbenes in the hope of coercing them into ever more difficult reactions. At one point it occurred to us that it would be sensible to isolate the carbene from its chemical environment by generating it in the gas phase, where encounters with other molecules would be minimized. We expected thereby to force the carbene to find some route to normal tetravalency with only its own internal resources. In 1967 Ronald Levin, who was then a graduate student completing the experimental work for his Ph.D., noticed that a carbene with the exotic name of 4,4-dimethylcyclohexa-2,5-dienylidene rearranged itself in the gas phase to produce para-xylene [*see lower illustration on this page*]. What was surprising, and the subject of a subsequent investigation of the mechanism, was the specificity of the reaction. Only one form of the xylene molecule, an isomer with the two methyl groups at opposite positions on the benzene ring, was produced. It was obvious from the first that this reaction provided an opening to the elusive small paracyclophanes. In principle all that was needed was to convert the two adjacent methyl groups into a ring and hope that the ring compound would then follow the same reaction path as the nonring compound. Fortunately the remaining problems were easily solved, and the construction of the precursors for paracyclophanes with six- and seven-carbon bridges was achieved. Even more fortunately, the carbene reaction turned out to proceed as easily in the ring case as in the nonring one, and the bent paracyclophanes were made in that way in only a few days.

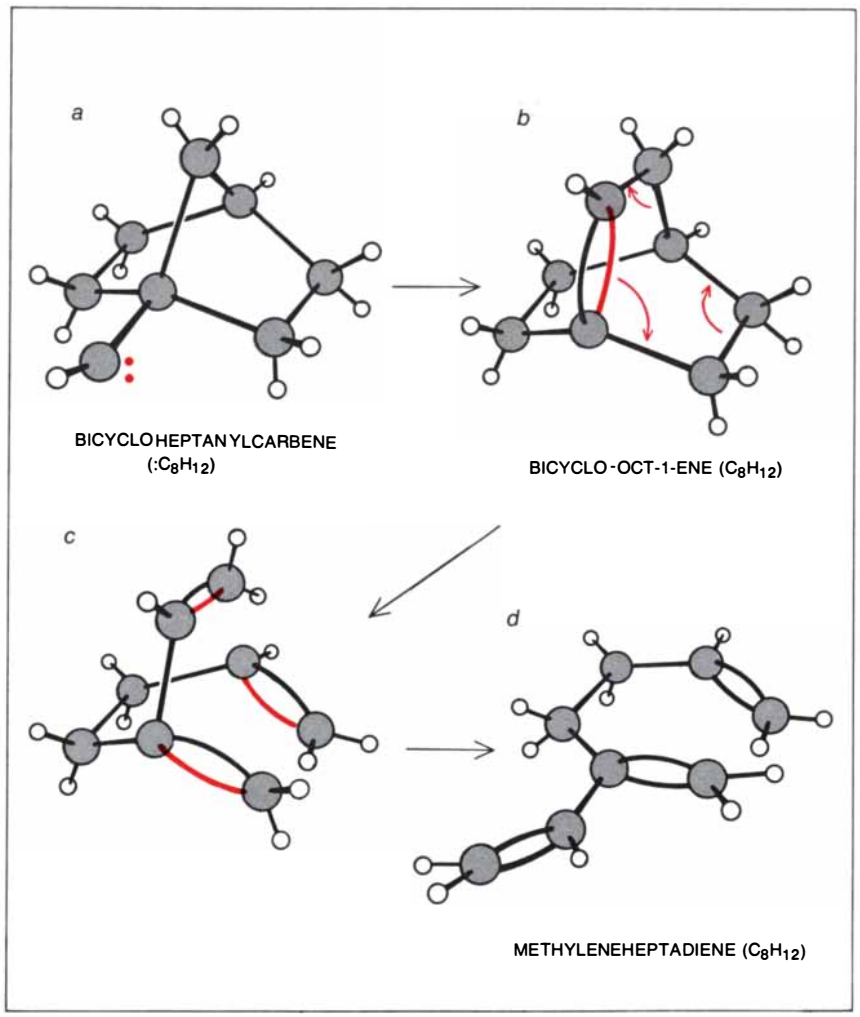
At the same time that our work was proceeding the older methods of synthesis were also brought to bear on the construction of a paracyclophane derivative with a seven-carbon bridge, and success in the effort was announced by Allinger and T. J. Walter of the University of Georgia even before our work was finished. Nonetheless, it is clear that the gas-phase rearrangements of carbenes represent enormously convenient sources of molecules otherwise available, if at all, only through methods requiring months or even years of work.

Carbenes also show great promise as sources of simpler compounds that have twisted double bonds. Once again many chemists had addressed themselves to this problem earlier, but there existed no general approach to the construction of such molecules. The idea was to devise ways of synthesizing cage molecules in which a double bond is present at what is known as the bridgehead position [*see top illustration at right*]. This synthesis is difficult precisely because the rigid cage structure demands that the orbitals making up the double bonds not overlap efficiently. The problem was recognized as early as 1924 by the German chemist Julius Bredt. The difficulty of producing a double bond attached to the bridgehead position is now encompassed in what is known as Bredt's rule. Although many groups of chemists now sporadically violate Bredt's rule, we thought that carbenes might provide a particularly easy route to the synthesis of such species. Our idea was to generate in the gas phase a carbene that would have no recourse but to react to give rise to a compound with a double bond at the bridgehead position [*see bottom illustration at right*]. If the system were properly designed, no option would exist other than for the ring to expand and generate an "anti-Bredt" molecule. Such species are often so unstable that isolation is impossible, but it is clear from the products we have isolated that our method is successful in creating molecules that have strongly twisted double bonds.

Although it would be incorrect to state that all the questions asked at the beginning of this research effort have been answered, we are on the way, and carbenes have been vital in easing our entry into the general problem of how organic molecules react. When we are able to answer the remaining questions, our knowledge of chemical bonding will have advanced much beyond its present state, and we shall be that much better informed about the workings of the chemical world around us.



DIFFICULT PROBLEM OF SYNTHESIS is encountered in attempting to make simple carbon compounds that have twisted double bonds. One approach is to incorporate the double bond at what is known as the bridgehead position in a cage molecule (*left*). The problem arises from the fact that the rigid structure of such a multi-ring molecule makes it difficult for the orbitals making up the pi bond to overlap efficiently (*right*).



CARBENE ANSWER to the problem of synthesizing compounds with twisted carbon-carbon double bonds involves the intramolecular reaction of a carbene attached to the bridgehead position of a cage molecule (*a*). The reaction takes place with an adjacent carbon-carbon single bond to produce the bridgehead double bond in a highly unstable intermediate (*b*). This molecule in turn rapidly decomposes (*c*) to yield an untwisted, and hence lower-energy, final product (*d*). Molecule in *b* is identical with one in top illustration.

The Biological Clock of Insects

With the approach of winter, insects enter into a state of dormancy. This shift in metabolism implies that they have a clock to measure changes in the length of the day. What is the nature of the clock?

by D. S. Saunders

Organisms that have the ability to measure the passage of time are said to possess biological clocks. One kind of clock measures the length of the day or the night, thereby enabling the organism to distinguish between the long days of summer and the short days of winter. The organism may respond to these seasonal cycles with a change in its metabolism, and such a response is called photoperiodic.

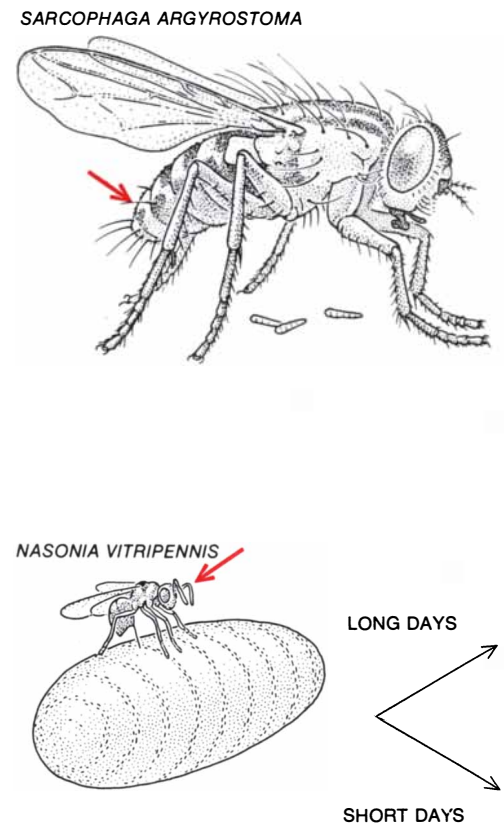
Daily rhythms or oscillations control a wide range of activities in plants and animals, and because such rhythms have a period close to, but not exactly, 24 hours they are called circadian (from the Latin *circa dies*, "about a day"). In insects the usual response to the onset of short days is diapause, a period of dormancy between stages of active growth and development. Changes of as little as 10 to 15 minutes in the length of the day can shift the metabolic regime of a population of insects from growth and development to diapause. There has been much controversy about the nature of the clock in such insects: whether it is comparable to an hourglass or to an oscillator with a period of close to 24 hours. Recent experiments indicate that the distinction between the two mechanisms may not be quite as important as interested biologists once thought.

The selective advantage of diapause in the life cycle of an insect is twofold. It provides a mechanism for survival over periods when food is in short supply or when the climate is unfavorable, and it serves to synchronize the development of the individuals in a population so that all of them feed as larvae or emerge as adults at the appropriate time of year. Reliance on day length for determining the season also confers a selective advantage. Compared with seasonal changes in temperature or humidity, changes in the length of the day are almost free of "noise," or irregularity. Furthermore, response to a certain day length well in advance of winter gives the insect's metabolism time to make the necessary adjustments for entering diapause. In anticipation of diapause the insect lays down reserves of fat, reduces the level of its metabolism and may form special layers of wax in order to resist dehydration.

The minimum requirements for a photoperiodic mechanism are (1) a receptor to detect the presence or absence of daylight, (2) a clock to measure the length of the day or the night and to integrate that information and (3) an effector system to control such metabolic changes as entry into diapause. Most of the early investigations of photoperiodism in insects were concerned with the nature and location of the photoreceptors and the effector system; only recently has progress been made in determining the nature and the properties of the clock mechanism.

Studies of a wide variety of insects have shown that the receptors involved in the photoperiodic response are in the animal's brain. It has also been demonstrated that the insect's eye is not involved in its photoperiodic response. Indeed, the eyes can be surgically removed without impairing the response.

In 1934 V. B. Wigglesworth of the University of Cambridge suggested that insects enter diapause because of a temporary lack of certain hormones, "due sometimes, perhaps, to an inborn rhythm, sometimes, perhaps, to the indirect effects of environmental factors." His hormone-deficiency theory of diapause has since been verified experimentally for a wide variety of insects, and it is clear that the effector system controlling the onset of diapause involves glands in the insect's brain. It appears that when the brain is exposed to short days, the release of the hormones is inhibited, and that when the brain is exposed to long days, their release is promoted.



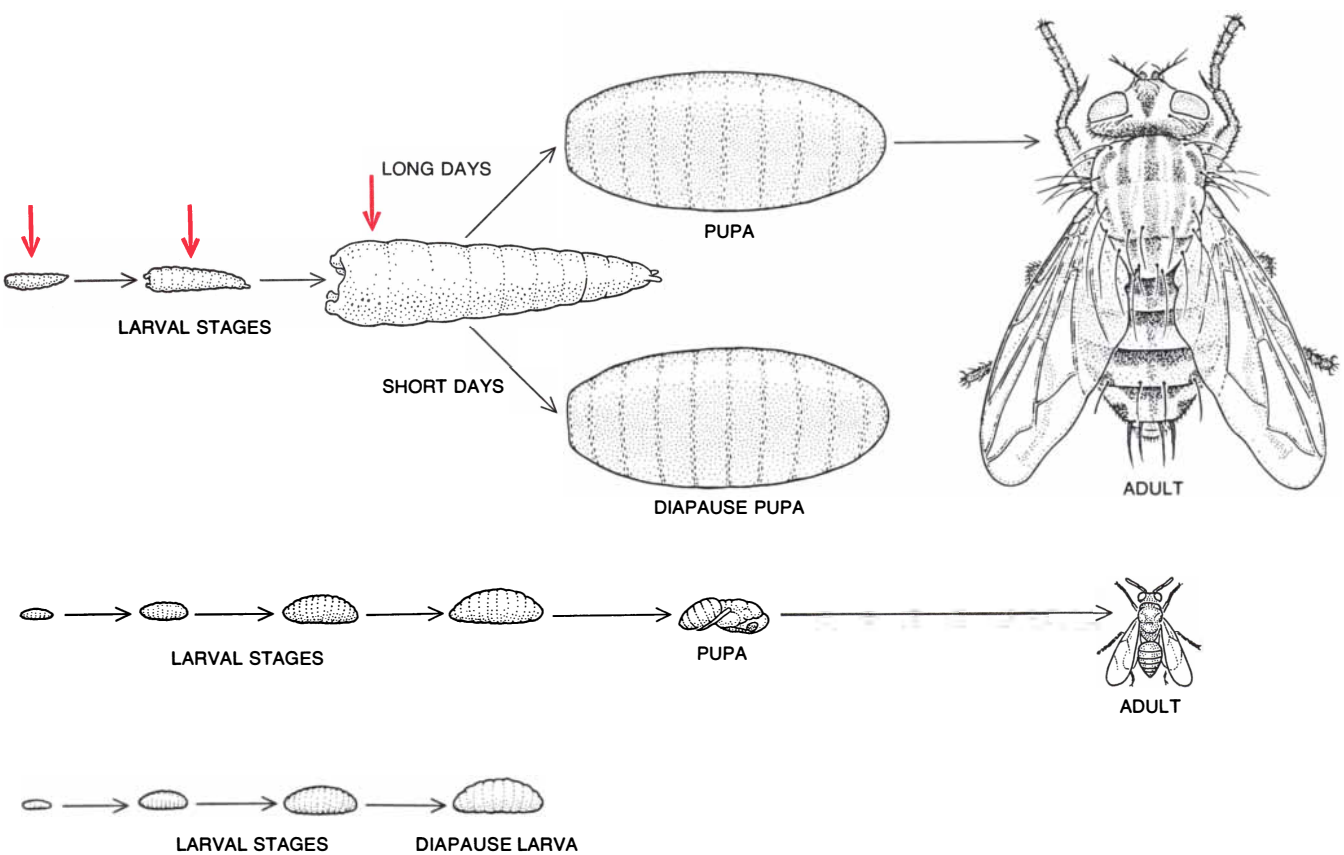
SENSITIVITY TO CHANGES in day length comes at different stages of development in the northern European flesh fly *Sarcophaga argyrostoma* and its parasite *Nasonia vitripennis*, a wasp. In the flesh fly sensitivity to

Since the photoreceptors and the hormone effector system are located in the insect's brain, it follows that the most likely site for the clock is also in the brain. Although there is some experimental evidence supporting this proposition, it is not conclusive. The central question, however, is the nature of the clock. How does it measure the length of the day or the night? How does it integrate successive long days or successive short days? And how does the clock translate the information that signals the effector system to either promote or inhibit the release of hormone by the brain?

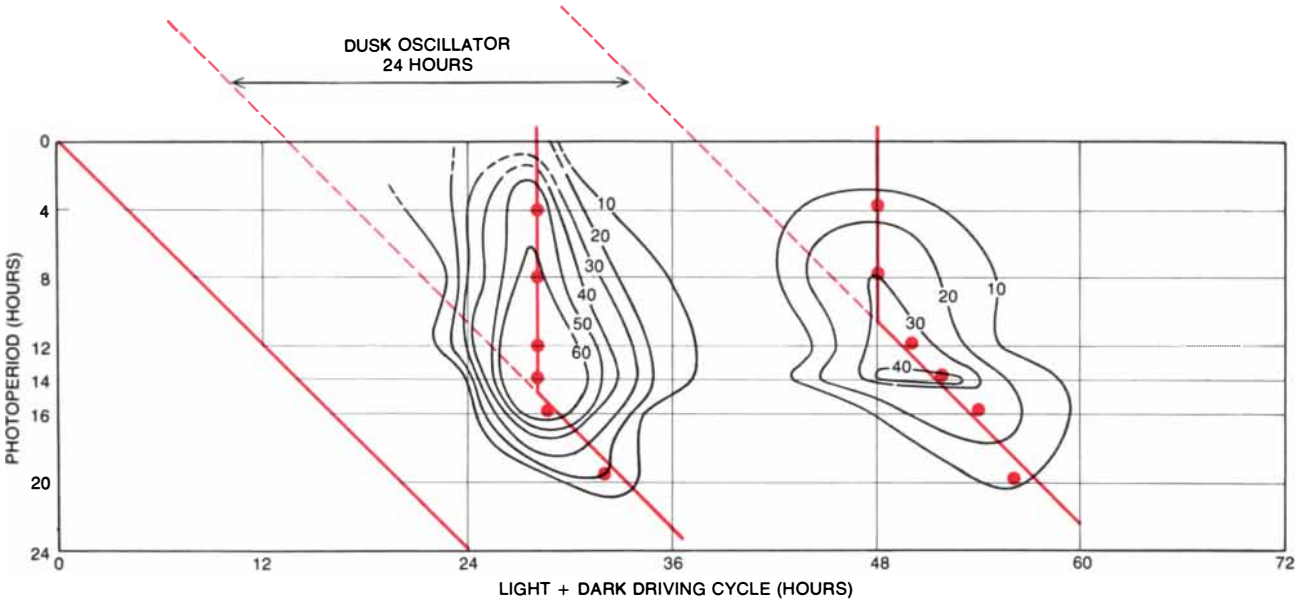
Two models of how insects measure time have been put forward. One proposes that the length of the day or the night is measured by an interval timer, an hourglass type of mechanism. Such a mechanism could be started by dawn and stopped by dusk, or it could be started by dusk and stopped by dawn. The alternative model was suggested by Erwin Bünning of the University of Tübingen in 1936. He proposed that the measurement of the length of the day or the night was accomplished by an endogenous, or built-in, daily rhythm that consisted of two half-cycles, one photophilic ("light-loving") and the other scotophilic ("dark-loving"). He thought that the phase of the endogenous rhythm would be set by dawn. If the daily light period was long, the period of illumination would extend into the scotophilic part of the cycle, and the organism would exhibit its typical long-day responses. If the daily light period was short, the organism would exhibit short-day responses. In effect, Bünning was proposing that organisms measure time with an endogenous circadian oscillator.

At the time Bünning put forward his model there was little experimental evidence to support it. Its great contribution was that it focused attention on the possibility that all types of biological time measurement have a common mechanism. Although Bünning's hypothesis appears to be similar to Wigglesworth's suggestion that diapause in insects could be due to "an inborn rhythm," it is clear that Wigglesworth was thinking more of an inborn annual rhythm than of rhythms with a circadian period.

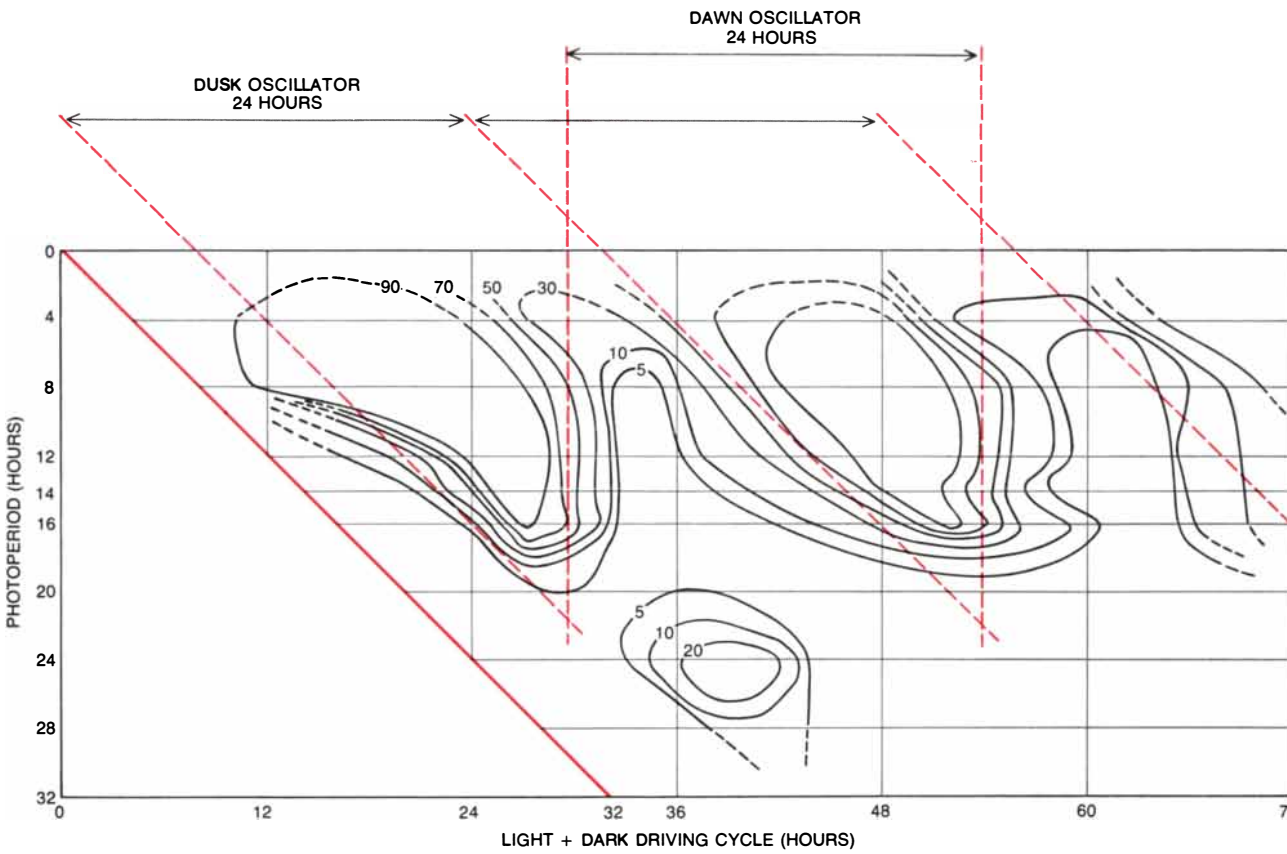
Unfortunately there is still no direct way to investigate the timing mechanisms of insects, and we are limited to treating the mechanism as a "black box." Accordingly most of the research has involved artificially varying the intervals of light and darkness and measuring the response of a population of insects to the various intervals in terms of the proportion of the population entering diapause. One successful technique is to expose the insects to repeated cycles of light and darkness, each consisting of a short period of light, say 12 hours, followed by an extended period of darkness, say 60 hours, with the dark period interrupted by brief pulses of light at a different time in the extended "night." Each experimental group of insects experiences



photoperiod (*colored arrows*) begins when the larvae are at the embryo stage in the maternal uterus and extends through three stages of larval development. Larvae exposed to more than 14 hours of light develop into adults without interruption, whereas larvae exposed to shorter days enter a period of diapause, or dormancy, as pupae. In the parasitic wasp the sensitivity to photoperiod comes while the eggs are in the maternal ovaries. The female wasp lays its eggs in the pupa of the flesh fly. Eggs exposed to more than 15¼ hours of light give rise to larvae that develop without diapause. Eggs exposed to shorter days give rise to larvae that go into diapause.



INDUCTION OF PUPAL DIAPAUSE in populations of the fly *S. argyrostoma* that were exposed to various cycles of light and dark is shown in the form of contour plots. The contours define the driving cycles of light and dark that induce diapause in the fly at the levels of 10, 20, 30, 40, 50 and 60 percent. The solid colored circles mark peaks of diapause induction. When the photoperiod is less than 12 hours, the diapause peaks come at the same time. When the photoperiod is more than 12 hours, the peaks form a slope that parallels the "light off" slope of the photoperiod, indicating that the oscillator now takes its principal time cue from beginning of dusk.



DIAPAUSE CONTOURS for the wasp *N. vitripennis* reveal three high-diapause plateaus at roughly 24-hour intervals. Diapause in the wasp is induced by exposing the females to short photoperiods while the eggs are still in the maternal ovaries. The contours define the driving cycles of light and dark that induce between 5 and 90 percent diapause in the wasp larvae. Each plateau has an ascending slope on its left side that is parallel to the "light off" slope of the photoperiod and a descending slope on its right side that is parallel to the vertical "light on" slope of the photoperiod. The slopes of the diapause plateaus are interpreted as being manifestations of two independent oscillators, one entrained, or set, by dusk and the other entrained by dawn. Each oscillator has a period of 24 hours.

the interrupting light pulse at a different time in the night. If the photoperiodic clock has something to do with a circadian rhythm, the diapause response should peak roughly at 24-hour intervals. If, on the other hand, the clock is like an hourglass, no such periodic effect should be observed.

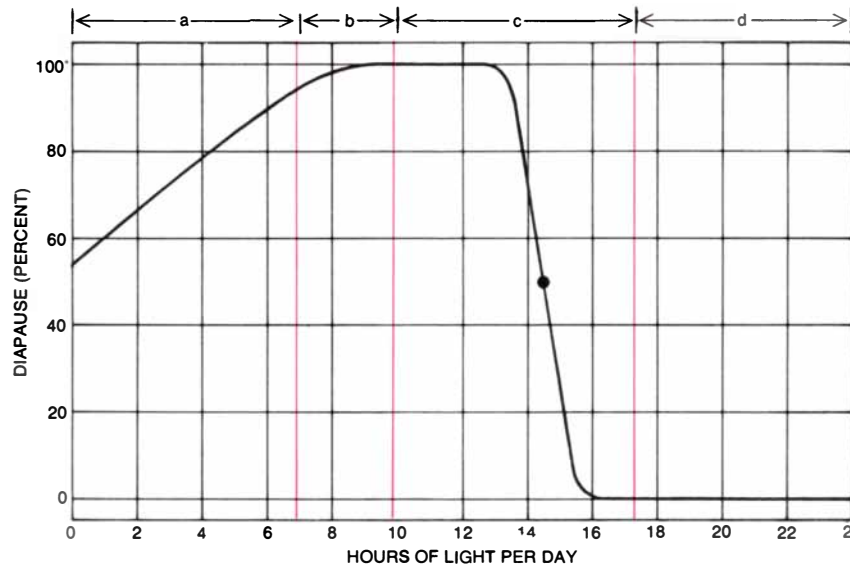
A second technique is to alternate a fixed period of light with a fixed period of darkness and to expose different populations of the insect to different combinations of light and darkness. If the driving cycle (the light period plus the period of darkness) is close to 24 hours or to multiples of 24 hours, one would expect the diapause response to peak roughly at 24-hour intervals if the clock is a function of the circadian rhythm. This approach has been called the resonance technique because it shows that the circadian oscillator functions most effectively when it is driven close to its natural period, that is, in resonance with the normal 24-hour cycle of night and day.

These two techniques have been successfully used to demonstrate the circadian nature of the photoperiodic clock in plants and birds. Until recently the results with insects were largely inconclusive. The first demonstration that photoperiodic time measurement in insects is a circadian phenomenon has come from my experiments with the large flesh fly *Sarcophaga argyrostoma* and its parasite *Nasonia vitripennis*, a wasp.

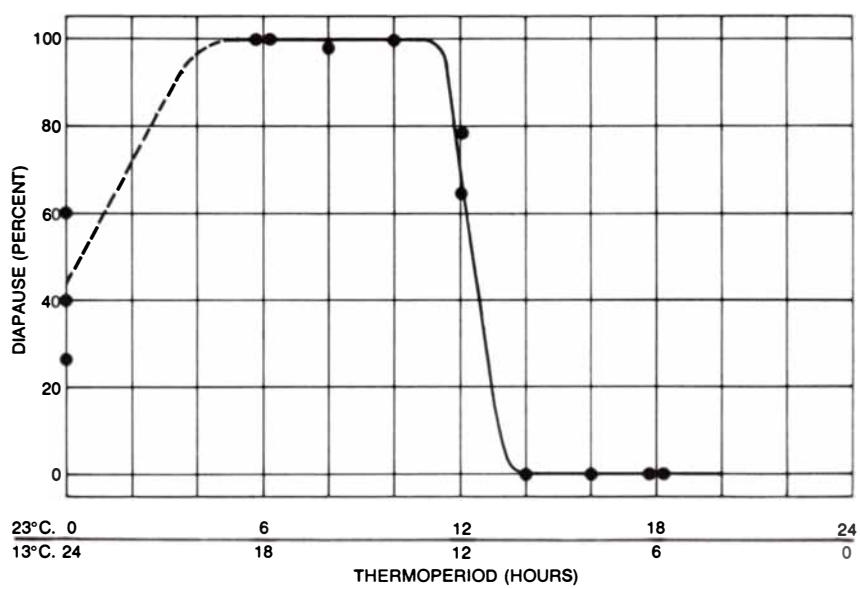
S. argyrostoma is a fly common in northern Europe. The larvae of the fly, which are deposited on carrion, are sensitive to photoperiodic influences at all stages of their development, from the time they are embryos within the maternal uterus until they metamorphose into pupae. If the larvae are artificially exposed to short days (less than 14 hours of light), the pupae enter diapause rather than metamorphosing into adult flies. If the larvae are exposed to more than 14 hours of light, the pupae metamorphose into adult flies without interruption.

N. vitripennis is a small wasp that drills through the outer pupal case of *Sarcophaga* and other flesh-eating flies. It lays its eggs on the pupa, and when the wasp larvae hatch out, they feed on the soft pupal tissues. They pupate within the host pupa and emerge as adults after biting a hole through its shell. The wasp larvae do not exhibit any photoperiodic response. The wasp's sensitive stage is the time when its eggs are still in the maternal ovaries.

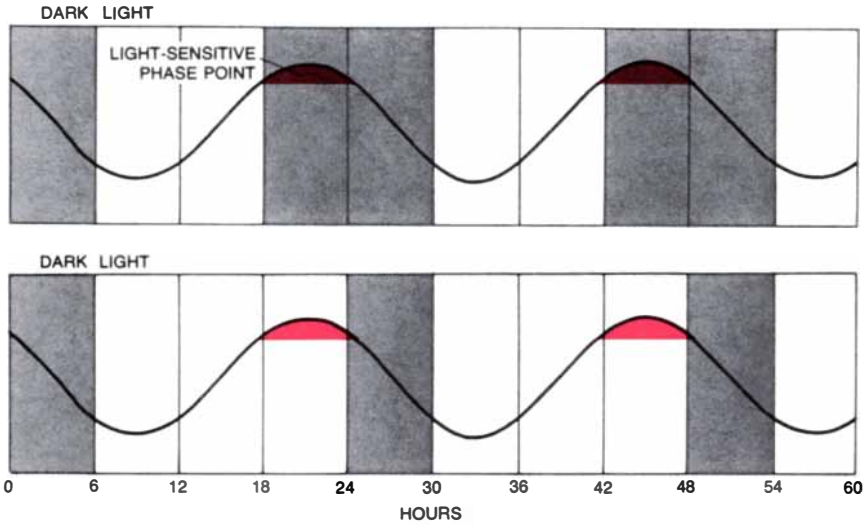
If the female wasp is exposed to long days (more than 15¼ hours of light), the eggs will give rise to larvae that do not enter diapause. If the female wasp is exposed to shorter days, the larvae emerging from those eggs will go into diapause in the last stage of their development. I initially demonstrated that the photoperiodic response of this wasp was a circadian phenomenon by means of experiments where the female wasps were exposed to repeated cycles of a short period of light followed by a long period of darkness that was interrupted by brief pulses of light. Maximum diapause was observed when the brief interrupting pulses came at 19, 43 and 67 hours after the onset of the main photoperiod of the



PHOTOPERIODIC RESPONSE CURVE for the fly *S. argyrostoma* shows the proportion of a population of larvae that produce diapausing pupae at various day lengths. The critical day length for the induction of diapause is 14½ hours. The photoperiods *a* and *d* do not occur naturally in the fly's normal environment. The photoperiods *b* occur during the winter, when the insects are in diapause, and photoperiods *c* occur during spring, summer and fall.



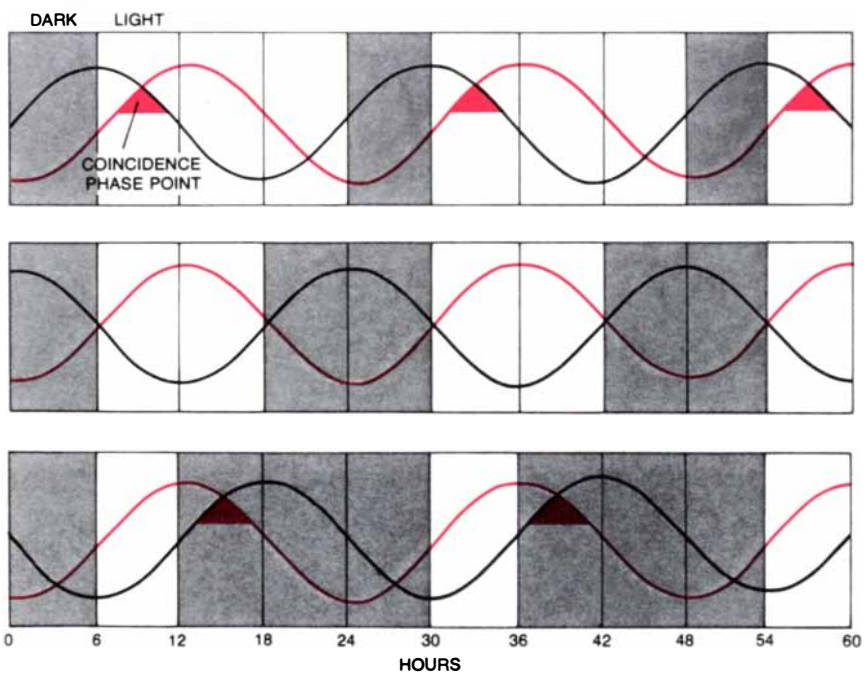
DAILY TEMPERATURE CYCLES can control diapause induction in *N. vitripennis* when the wasp is reared in complete darkness. Female wasps raised in a cycle that has more than 11 hours at 13 degrees Celsius give rise to progeny that go into diapause. The results show that the wasp's oscillators can be entrained by temperature cycles as well as by light cycles.



EXTERNAL-COINCIDENCE MODEL for the measurement of time by the photoperiodic clock, as proposed by Colin S. Pittendrigh of Stanford University, consists of a circadian oscillator with a light-sensitive phase point (*colored area*). When the day length is short, say 12 hours, the light-sensitive phase point is not illuminated and diapause occurs (*top*). When the days are long, the phase point is illuminated and diapause does not occur (*bottom*).

cycle (that is, at 24-hour intervals), and minimum diapause was observed when the interrupting pulses fell at intervening times.

In 1971 I spent a year at Stanford University working in the laboratory of Colin S. Pittendrigh. Pittendrigh has made intensive studies of the relation between photoperiodism and circadian rhythms, particularly in the fruit fly and the pink bollworm moth. He has extended Bünning's hypothesis with the proposal that the endogenous clock consists of two independent oscillators, one entrained, or set, by dawn and the other entrained by dusk. As the photoperiod changes, the phase relations of the two oscillations also change. In this internal-coincidence model, as it is now called, light has only one role: entrainment. Long days would shift one set of oscillations until it was in step with the other, and the combined effect of the two sets of oscillations would result in development without diapause, perhaps because any in-phase temporal relation between two cellular components would, for example, result in the synthesis of a third substance that is responsible for initiating development. Short days, on the other hand, would put the two sets of oscillations out of phase, and in the absence of the substance that promotes development, diapause would occur [*see bottom illustration on this page*]. A similar model was independently proposed by the Russian entomologist V. P. Tyshchenko in 1966.



INTERNAL-COINCIDENCE MODEL, also proposed by Pittendrigh, consists of two circadian oscillators, the phase of one set by dawn (*colored curve*) and the phase of the other set by dusk (*black curve*). When the days are long, the two oscillations are partially in phase (*top*) and development proceeds without diapause. When the days are short, the oscillations are out of phase and diapause occurs (*middle*). Very short days (*bottom*) may produce some coincidence of the two oscillations, which would result in development without diapause.

In Pittendrigh's laboratory I conducted resonance experiments with both the fly *Sarcophaga argyrostoma* and its parasitic wasp *Nasonia vitripennis*. The results demonstrated that the photoperiodic response of both insects was circadian but that the details for the two species were quite dissimilar. In experimental populations of the parasitic wasp the results indicated the presence of independent "dawn" and "dusk" oscillations. In the flesh fly, however, the photoperiodic clock seemed to consist of a single oscillation that was entrained to the entire photoperiod when the photoperiod was less than 12 hours, but once the photoperiod exceeded 12 hours the oscillation obtained its principal time cue from dusk. The two species therefore showed the properties of the two current theoretical models for oscillatory photoperiodic clocks: respectively internal coincidence and external coincidence. The parasitic wasp displayed the properties of Pittendrigh's model with two oscillators, and the flesh fly displayed the properties of Bünning's model with a single oscillator.

If the photoperiodic clock of the parasitic wasp is of the internal-coincidence type, a simple test is available. It is known that a variety of circadian rhythms can be entrained by temperature. Since light has only the role of entrainment in this model, it should be possible to simulate the effects of photoperiodic entrainment with periods of higher and lower temperature in the complete absence of light.

Female wasps raised from the egg

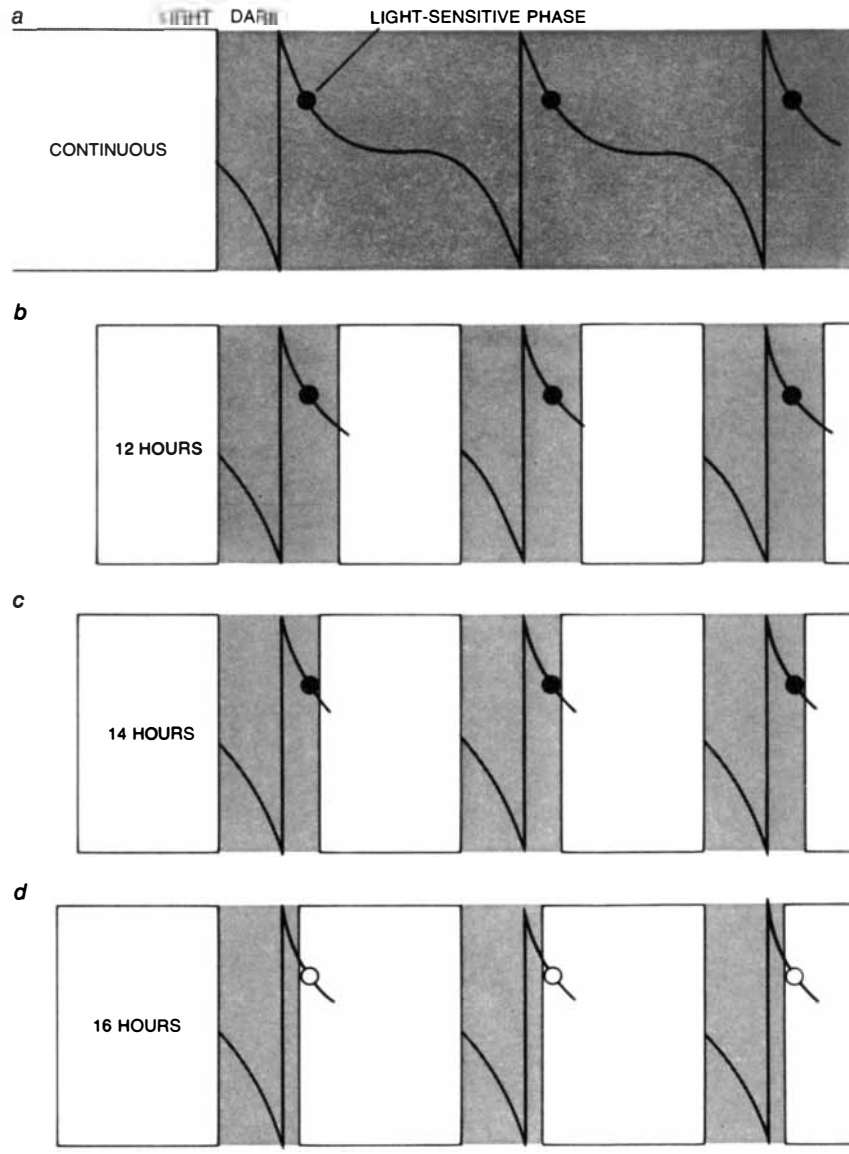
stage in total darkness were exposed to various daily temperature cycles consisting of a period at 23 degrees Celsius followed by a period at 13 degrees C. Wasps raised in a cycle that had more than 13 hours at the higher temperature gave rise to progeny that did not go into diapause, and wasps raised in a cycle that had more than 11 hours at the lower temperature gave rise to progeny that did go into diapause [*see bottom illustration on page 117*]. The results of the experiment not only show that temperature cycles can be substituted for light cycles in this wasp but also rule out for the wasp any model of a clock in which light induces, or starts, diapause or development by photobiological means other than mere entrainment.

The clock of the flesh fly, on the other hand, does suggest such an inductive property for environmental light. Pittendrigh has proposed an external-coincidence model in which light plays a dual role: entrainment and induction. That is, light can both entrain the oscillation and control induction. This model consists of a circadian oscillator that is entrained by light in such a way that a photoinducible phase of the oscillator is not illuminated in short days but is illuminated in long days. Illumination of the photoinducible phase results in long-day effects, or uninterrupted development, but when the photoinducible phase falls in the dark period, diapause supervenes.

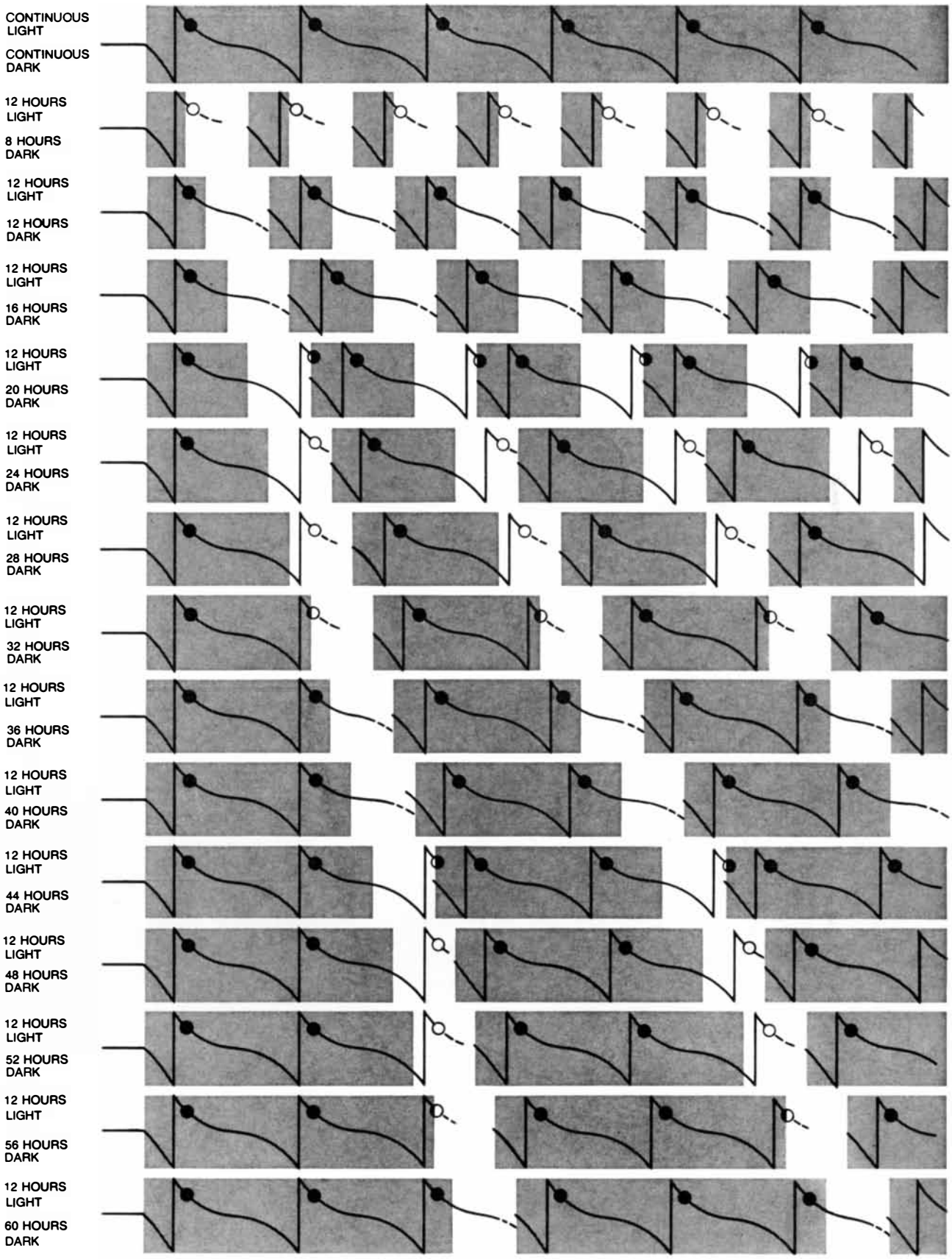
Evidence supporting the external-coincidence model comes from Pittendrigh's intensive study of the emergence of fruit flies of the species *Drosophila pseudoobscura* from their pupae. The circadian rhythm of adult emergence is damped out by periods of light greater than 12 hours but resumes when the pupae are transferred to darkness. In a natural 24-hour light cycle with 12 hours or more of light the oscillation restarts each dusk and measures the length of the night as if it were an hourglass. Only when the night is extended to more than 24 hours is the oscillatory nature of the system revealed.

The resonance experiments with the fly *Sarcophaga argyrostoma* produced similar results and demonstrated the hourglass nature of the night-length measurement. Thus photoperiods greater than 12 hours had a damping effect on photoperiodic oscillation in *S. argyrostoma*. It therefore seems to be a reasonable working hypothesis that the clock of the fly is of the external-coincidence type.

Additional evidence for the existence of a dark-period hourglass in insects comes from studies of the aphid *Megoura viciae* by Anthony D. Lees of the Imperial College of Science and Technology. Lees found that the aphid's clock is set in motion at dusk and measures night length as an hourglass does. Moreover, the aphid, like the flesh fly, needs a photoperiod of a minimum length before the clock can act as an hourglass. When the aphid is subjected to extended nights, however, it does not display oscillatory properties. This suggests that the aphid's clock is an oscillator executing only a single cycle before it is extinguished. If that is the case, the dividing line between oscillators and hour-



RHYTHM OF PUPAL EMERGENCE in the fruit fly *Drosophila pseudoobscura* shows the characteristics of the external-coincidence model for the photoperiodic clock. When pupae are transferred from continuous light to continuous darkness (*a*), the rhythm with which adult flies emerge from the pupa always starts at a fixed phase, called circadian time 12 because it is equivalent to the phase of the oscillation at the end of a 12-hour light period (*b*). The light-sensitive phase first occurs about nine hours later at circadian time 21, and in constant darkness it appears at about 24-hour intervals thereafter. With cycles of 12 hours of light and 12 of darkness (*b*) and 14 hours of light and 10 hours of darkness (*c*), the light-sensitive phase (*solid circles*) falls in the dark period, and since it is not illuminated, diapause occurs. When the cycle is 16 hours of light and eight hours of darkness (*d*), the light-sensitive phase is illuminated (*open circles*) and diapause does not occur. Since the rhythm always starts at dusk at the same circadian time, as the length of the light period increases dawn moves backward in relation to the oscillation. Since the oscillation is reset by each photoperiod, it measures the length of the night as though it were an hourglass.



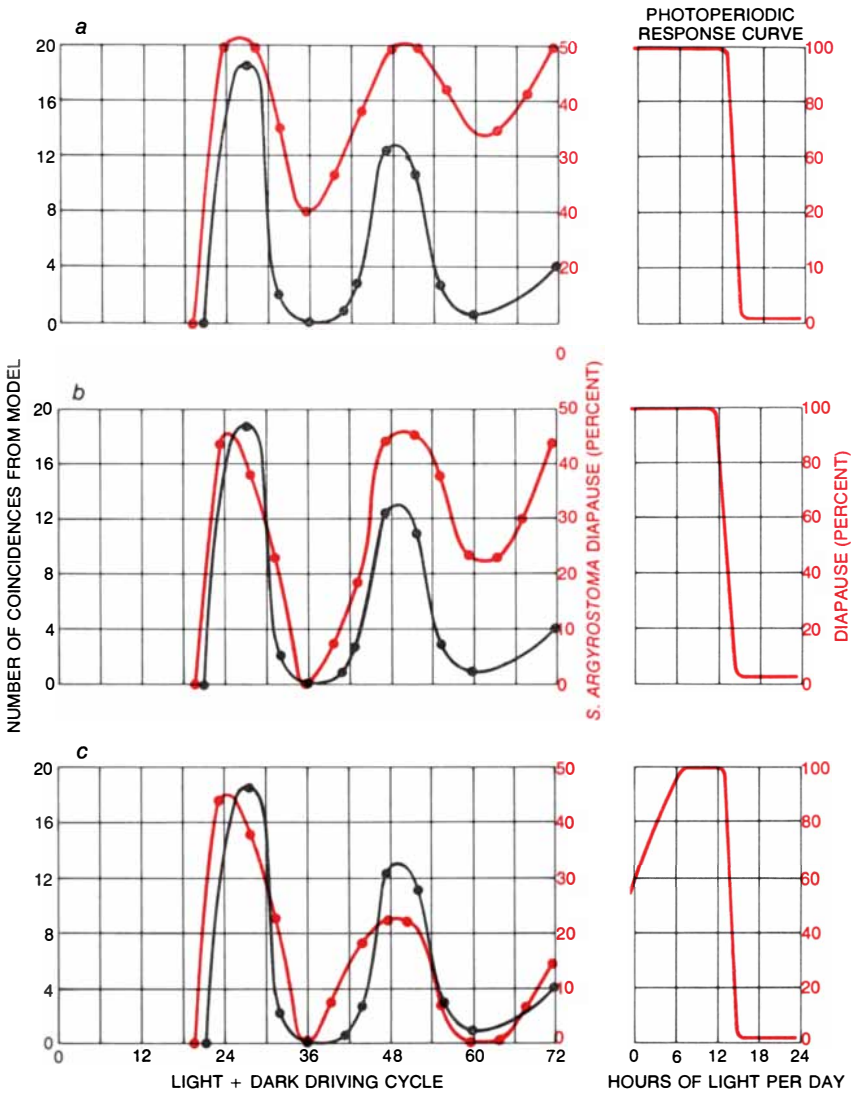
THEORETICAL APPLICATION of the external-coincidence model to resonance experiments involves using a 12-hour photoperiod and varying the periods of darkness. When the light-sensitive phase of the circadian oscillator falls in a dark period, diapause occurs (*solid circles*). When the light-sensitive phase is completely illuminated (*open circles*), diapause does not occur. Partial illumination of the sensitive phase results in a lower incidence of diapause. In all cases the oscillations take their time cue from onset of dusk.

glasses becomes very fine, perhaps even academic.

Bünning's single-oscillator model, the external-coincidence model and the two-oscillator internal-coincidence model have one thing in common: long days have a positive effect in that they give rise to development or prevent the insect from going into diapause, whereas short days are considered to be essentially neutral. All insects add up successive light periods before they display photoperiodic induction. It is therefore probable that long days lead to the production of a chemical substance, and the summation of long days is represented by the accumulation of that substance to a level where it stimulates the release of a brain hormone that initiates the chain of events in egg development, molting or metamorphosis.

In the flesh fly *Sarcophaga argyrostoma* and its parasitic wasp *Nasonia vitripennis,* however, the summation of short days has been shown to be virtually independent of temperature, and this temperature-compensated mechanism should probably be regarded as an integral part of the clock. The rate of development of the fly larvae or of the wasp eggs, however, is dependent on the temperature, like most other physiological processes. Therefore an interaction between on the one hand the temperature-dependent period when the insects are sensitive to photoperiod and on the other the temperature-insensitive number of short-day cycles required for the induction of diapause means that female wasps or fly larvae kept at higher temperatures show a lower diapause response than those kept at a lower temperature. The reason is that they are unable to summate a sufficient number of short-day cycles before the wasp eggs are deposited or the fly larvae pupate and the sensitive period comes to an end.

The conceptual difficulties with regard to whether long days have a positive inductive influence and short days are neutral are well illustrated in these responses. For example, the two insects clearly summate short days as well as long ones, so that short days are not neutral in the same sense that constant darkness is. In the flesh fly, however, these difficulties can be surmounted if we incorporate an additional component in the external-coincidence model of the insect's clock. That component measures the duration of the day. It has hourglass properties in that it requires at least six hours of light to register a short day and does not reset itself in extended periods of darkness.



THREE EXTERNAL-COINCIDENCE MODELS are compared with the results of diapause experiments with the fly *S. argyrostoma*. The black curves show the number of inductive coincidences between the photoperiod and the light-sensitive phase in the external-coincidence model that is depicted on the opposite page. In the unmodified external-coincidence model (*a*) the data from the experiments with the fly (*colored curve*) do not correspond to the curve obtained from the model. Furthermore, the photoperiodic response curve (*right*) does not show the decrease in diapause at very short day lengths that was found experimentally. In the second version of the external-coincidence model (*b*) it is assumed that the diapause-promoting substance is synthesized when the photoperiod does not coincide with the light-sensitive phase and that the substance is destroyed when there is coincidence. The resulting curve (*color*) corresponds more closely to the theoretical model, but the photoperiodic response (*right*) lacks the characteristic drop at very short day lengths. The third version (*c*) incorporates a second component, an hourglass, that measures the length of the day. The two curves now correspond closely, and the photoperiodic response curve also matches the results that have been obtained experimentally for *S. argyrostoma*.

There are strong similarities between the two-component-clock model for the parasitic wasp and the actual properties shown by the photoperiodic clock of the flesh fly. The two-component clock accounts for the summation of short days as well as for the summation of long days, and it explains the decrease in diapause associated with very short periods of light. The principal value of such a model lies in its predictive powers. If the model is valid, the experiments it suggests should bring us an important step closer to understanding the fundamental nature of biological timekeeping. Even as matters stand, it seems clear that the biological clock of an insect such as the flesh fly is not simply an hourglass or an oscillator. It is some subtle combination of both.

MATHEMATICAL GAMES

Some elegant brick-packing problems, and a new order-7 perfect magic cube

by Martin Gardner

"Pack my box with five dozen liquor jugs."

—ANONYMOUS PANGRAM

A pangram is an attempt to pack as many different letters as possible into the shortest intelligible sentence. It is considered not cricket to use names and initials, such as "Schwartz" and "X. Q. Zym," or strange words, such as "pyrzqxgl," which in *The Magic of Oz* enables you to change instantly into any kind of animal you like if you know how to pronounce it correctly.

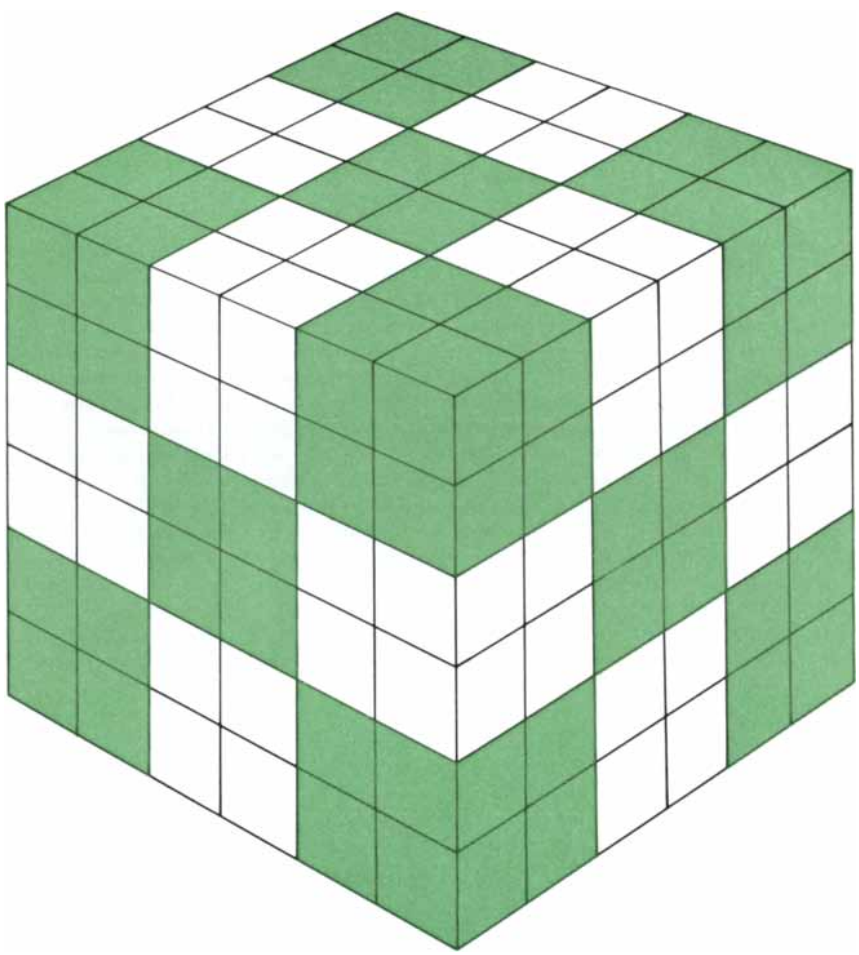
Many ultimate pangrams of 26 different letters have been constructed by word players, but they tend to be inelegantly obscure, for example "Vext cwm fly zing jabs kurd qoph," by Dmitri Borgmann. It means that an annoyed fly in a Welsh mountain hollow, humming shrilly, pokes at the 19th letter of the Hebrew alphabet drawn by a Kurd. Cryptographers find a perfect pangram amusing because it can be written in a simple letter-substitution cipher as ABCDE FGHIJ KLMNO PQRST UVWXY Z.

The difference between creating pangrams and working on packing problems in combinatorial geometry is not as great as one might suppose. The restraints in the former are the formation rules of English spelling and grammar and those in the latter are the rules of mathematics. At least two eminent mathematicians, Augustus De Morgan and Claude E. Shannon, are on record as having spent considerable time composing pangrams, and I know of many lesser mathematicians who have tried their hand at it.

In mathematics a packing problem in general is one in which a given set of mathematical objects are to be packed as efficiently as possible into a given space according to given rules. Computer scientists, for example, are concerned with finding fast algorithms for packing sets of numbers into "bins," with the sum of the numbers in each bin not to exceed a specified limit. Such algorithms are needed for the efficient storage and retrieval of information. Geometrically the task can be viewed as a problem in one-dimensional packing: packing rods of varying lengths inside long pipes into which the rods fit snugly.

In a complex industrial society all kinds of problems arise involving the packing of three-dimensional objects into a specified area: the storing of objects in a warehouse; the packing of supplies into ships, planes and freight cars; the packing of objects in cartons for distribution to stores, and so on. Perhaps it is the increasing need for packing algorithms that has stimulated some mathematicians in recent years to spend more time on such problems.

This month we consider only the simplest kind of solid packing: the packing of "bricks" (rectangular parallelepipeds) into a "box" (also a rectangular parallelepiped). To simplify still more, we assume that all three dimensions of both bricks and box are integral, and that the volume of the box exactly equals the total volume of the bricks to be put inside. As David A. Klarner says in his article "Brick-packing Puzzles" (*Journal of Recreational Mathematics*, Vol. 6, No. 2, 1973), many people are surprised to learn that even with these strong simplifications there are problems that are both elegant and challenging. By elegant Klarner means the following. If the bricks will pack the box perfectly, the problem is elegant if finding a way to pack it seems simple but is actually difficult. And if the bricks will not pack the box, the problem is elegant if there is a



Color scheme for an order-6 cube-packing problem

simple but subtle way to prove impossibility. Klarner, now at the State University of New York at Binghamton, is one of the pioneers of brick-packing theory. It is to him that I am indebted for most of what follows.

About 1960 the Dutch mathematician Nicolaas G. de Bruijn was struck by the fact that his son, age 7, was unable to fill a $6 \times 6 \times 6$ box with 27 bricks, each $1 \times 2 \times 4$. The two volumes are the same and the packing seems easy, but one always ends up with at least one hole that the last brick will not fill. Studying the matter led de Bruijn to interesting results. They were first published as problems in a Hungarian journal, then later summarized by de Bruijn in his paper "Filling Boxes with Bricks" (*American Mathematical Monthly,* January, 1969).

De Bruijn calls a brick harmonic if its three measurements are integral and can be ordered so that each length is a multiple of the preceding one. In algebraic terms a harmonic brick has the form $a \times ab \times abc$, where the letters are positive integers. The $1 \times 2 \times 4$ brick is, of course, harmonic. It is called the canonical brick because it not only is the simplest harmonic brick of three distinct measurements but also approximates the shape of ordinary bricks used in masonry [*see top illustration on this page*].

De Bruijn was able to prove that a collection of identical harmonic bricks, each $a \times ab \times abc$, will perfectly pack a box if and only if the box is $ax \times aby \times abcz$. "Perfectly pack" means to fill completely; it is the same as saying that the brick will tile the box. De Bruijn showed that perfect packing is possible only if the box's dimensions are multiples of the brick's dimensions. To put it another way, if the bricks pack the box at all, there will be a way to do it trivially. That means they will pack when all are identically oriented. (Of course, they may also pack in nontrivial ways.) If the bricks are not harmonic, there are boxes they will fill only in a nontrivial way. For example, five nonharmonic bricks of $1 \times 2 \times 3$ will pack a $1 \times 5 \times 6$ box, but they cannot pack the box if they are all parallel.

De Bruijn's results generalize to hyperbricks in all higher Euclidean spaces, and they also hold for 2-space "bricks" (rectangles). The canonical plane brick—the 1×2 domino—will pack a rectangle only if one side is even, and then, of course, a trivial packing is possible.

Let us return to the task that puzzled de Bruijn's son. Since 6 is not a multiple of 4, we know from de Bruijn's work that the canonical brick will not pack an order-6 cubical box. Is there a simple impossibility proof?

There is, and it is a generalization of the solution to the old brainteaser about an order-8 checkerboard that has had two diagonally opposite corner squares removed. Can the board be covered with 31 dominoes? The fact that it cannot is evident once you realize that the two missing squares are the same color. The board therefore contains 32 squares of one color and 30 of the other. Since a domino must cover two squares of opposite color, after 30 dominoes are put down there will always be two uncovered squares of the same color that cannot be covered by the last domino.

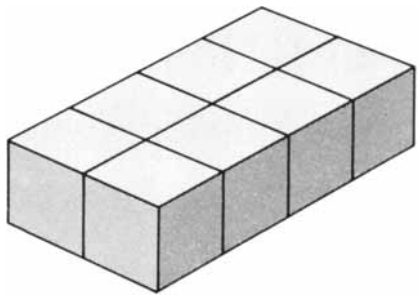
To apply the same kind of parity check to the cube-packing problem, imagine that the order-6 cube is divided into 27 order-2 cubes and that the order-2 cubes are colored as is shown in the illustration on the opposite page. No matter how a canonical brick is oriented inside such a cube, it will fill four colored cells and four white cells. The cube, however, has eight more colored cells than white cells. Therefore after 26 bricks are placed eight cells of the same color will remain. Clearly the last brick cannot fill them.

Will 125 canonical bricks pack a $10 \times 10 \times 10$ box? They will not, and the same impossibility proof applies. Indeed, the proof applies to any cube with a side that is even and not a multiple of 4. Will 250 bricks of $1 \times 1 \times 4$ pack the order-10 cube? As the impossibility proof shows, the answer is no. Sometimes more than two colors are needed for elegant impossibility proofs, but it is surprising how much can be done with just two colors.

Here is a delightful tiling problem from Klarner that is easily proved impossible by two-coloring, although not a checkerboard coloring. You have a 25×25 square you want to tile (no overlap, no vacancies) with a mixture of 2×2 squares and 3×3 squares. I shall give the elegant impossibility proof next month.

One of Klarner's theorems introduces the concept of cleavability. If a rectangle can be tiled with identical rectangles, there is always a way to tile it so that it can be cut into two smaller rectangles, each of which can also be tiled. Such a rectangle is said to be cleavable.

Does this unusual theorem have a 3-space analogue? That is, if a box can be fully packed with identical bricks, can it always be cut into two smaller boxes, each packable with the same bricks? The answer, Klarner found, surprisingly is no.

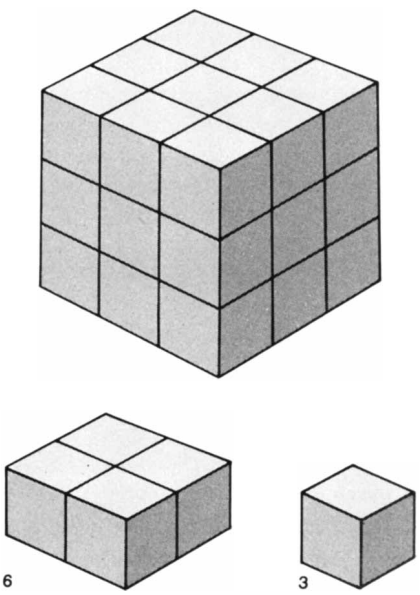


The canonical brick

The smallest example (discovered by David Singmaster) is the $5 \times 5 \times 12$ box. It can be packed with $1 \times 3 \times 4$ bricks, but not in a way that is cleavable.

Is there an infinite number of noncleavable boxes a given brick will pack? Again the answer surprisingly is no. Klarner was the first to show that in any Euclidean space an infinite set of boxes packable by a given brick has a finite subset of packable boxes that can be used for packing all the others. In addition he showed that for every brick there is a finite set of packable but noncleavable boxes. In 1971 two Hungarian mathematicians, G. Katona and D. Szász, refined Klarner's findings by giving a constructive proof that included specific numerical bounds.

If instead of identical bricks we allow a mixture of different bricks (as in Klarner's tiling problem), many beautiful new problems arise. One of the simplest is a $3 \times 3 \times 3$ box puzzle that Klarner says first appeared in a Dutch book in 1970. We want to pack it with six $1 \times 2 \times 2$ bricks and three unit cubes. It



The $3 \times 3 \times 3$ packing puzzle

Your kind of Books...

And you can examine them free for 15 days.

COMPUTER POWER AND HUMAN REASON
From Judgment to Calculation

Joseph Weizenbaum

We do not know how to make a "wise" computer, and yet, as this book brilliantly documents, scientists, politicians, and ordinary citizens are today relinquishing to computers the responsibility for making genuinely human choices—choices that demand true wisdom. *Computer Power and Human Reason* is an account of the impact of technology on man's self-image and of his abandonment of judgment for calculation. It is a distinguished scientist's searching probe of the limits of computer power and of scientific rationality itself. It is above all an eloquent and moving defense of the sanctity of the human spirit. Joseph Weizenbaum is professor of computer science at MIT.

1976, 300pp., 13 illus., $9.95

A Book-of-the-Month Club Alternate Selection

FIELD PHOTOGRAPHY
Beginning and Advanced Techniques

Alfred A. Blaker

The former Director of the University of California Scientific Photographic Laboratory shares his many tricks of the trade in this clear, comprehensive, and beautifully illustrated book. *Field Photography* includes more than 70 original photographs (18 in full color), 85 drawings, and a separate, portable field supplement of technical data. This is sure to become the standard, indispensable reference for both professional and amateur nature photographers.

1976, 445 pp., 156 illus., $19.95

GEOLOGY ILLUSTRATED

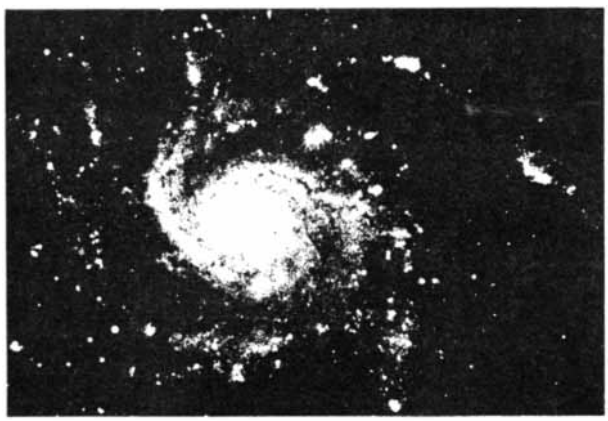
John S. Shelton

Drawings by Hal Shelton

Geologist/pilot Dr. John Shelton has been taking aerial photographs since 1939, and in this, his master work, he assembles nearly 400 of these photographs to produce a breathtaking view of the natural features of our home planet.

"Worth buying for the photos and book design alone, but you'll probably find yourself becoming interested in geology regardless of your original intentions."
—The Whole Earth Catalog

1966, 434 pp., 382 illus., $14.50



Hale Observatories

HIGHLIGHTS IN ASTRONOMY

Fred Hoyle

As his millions of readers will attest, Fred Hoyle has a flair for sharing the very latest astronomical discoveries with the general reader. Hoyle's latest book, *Highlights in Astronomy*, is a lavishly illustrated tour of the universe. The 149 illustrations include nearly 100 photographs, some of them never before published. This is quite simply the best popular description of what modern astronomy is all about.

1975, 179 pp., 149 illus., softcover, $5.95

HELPLESSNESS
On Depression, Development, and Death

Martin E. P. Seligman

"A place to look for the problem when that hideous vague sense that 'something's wrong' begins to grow."
—CoEvolution Quarterly (of the *Whole Earth Catalog*)

Helplessness is Dr. Martin Seligman's new theory of anxiety and depression, the result of a ten-year search for the causes and cures of these widespread disorders. Dr. Seligman's research suggests that depressed individuals can *learn* to exert greater control on their environments and thereby lessen their feelings of depression. Of special interest is Seligman's suggestion that early childhood experiences with control may effectively immunize a child against depression for life.

1975, 250 pp., 8 illus., softcover, $4.95

SCIENTIFIC AMERICAN Books

Clearly written and beautifully illustrated books by the most outstanding scientists of our day.

THE ADVISORS
Oppenheimer, Teller, and the Superbomb

Herbert F. York

Among President Truman's scientific advisors, Edward Teller and J. Robert Oppenheimer represented opposing factions in the top-secret debate over whether or not to build "the super"—the hydrogen bomb. *The Advisors* is an absorbing and long overdue look at the technical, strategic, and human content of that great debate. Drawing on his own participation in Project Superbomb, on interviews with other participants, and on several recently declassified documents, Herbert York tells how history has shown that Oppenheimer's position—in opposition to the bomb—was correct.

1976, 175 pp., illustrated, $6.95

A GUIDE TO PSYCHOLOGISTS AND THEIR CONCEPTS

Vernon J. Nordby and Calvin S. Hall

Here is a fascinating look at the lives and works of the men and women who helped shape modern psychology. Turn to any page and you'll find biographies, discussions of the major ideas, lists of the most important published works, and an artist's original portrait of each psychologist.

1974, 187 pp., 42 illus., softcover, $3.95

ARCHAEOLOGICAL ATLAS OF THE WORLD

David and Ruth Whitehouse

With 103 maps drawn by John Woodcock and Shalom Schotten

An essential new reference for everyone with an interest in archaeology, this handy volume contains 103 maps pinpointing more than 5000 archaeological sites throughout the world.

1975, 272 pp., 109 illus., softcover, $8.95

A HISTORY OF AMERICAN ARCHAEOLOGY

Gordon R. Willey and Jeremy A. Sabloff

A History of American Archaeology traces the trends in archaeological investigations of the New World from 1492 to the present. The book contains more than 100 illustrations, including some excellent historical photographs from the nineteenth and early twentieth centuries, and a bibliography of more than 850 references.

" . . . The first book-length history of American archaeology by professional archaeologists. . . . An invaluable source book."—*Science*

1974, 252 pp., 124 illus., softcover, $5.50

THE SOLAR SYSTEM

In less than twenty years of exploring the solar system we have learned more about the sun, the planets, and interplanetary space than in all the preceding centuries of earth-bound observation. This book is an exciting compendium of our new knowledge.

1975, 145 pp., 96 illus., softcover, $4.50

ENERGY AND POWER

The Environmental Monthly calls this "just possibly the most important book published in recent years for the concerned layman. . . . " *Energy and Power* gives expert coverage to the scientific, technological, social, and environmental aspects of energy use.

1971, 144 pp., 116 illus., softcover, $4.50

THE HUMAN POPULATION

"It is indeed refreshing to read a book dealing with population which is not only written by experts, but also written in readable and intelligible language for the benefit not only of colleagues but also laymen."—*Choice*

1974, 147 pp., 98 illus., softcover, $4.50

LIFE AND DEATH AND MEDICINE

Life and Death and Medicine provides a solid framework for understanding the complex and pressing health care crisis confronting our society. The book discusses the unique biological needs and frailties of human beings, how medical science has responded to these needs and frailties, and the economic factors influencing the quality of health care.

1973, 147 pp., 93 illus., softcover, $4.50

THE BIOSPHERE

For perhaps two billion years, the air and water of our planet have remained virtually the same, maintained by all living things and by the interrelated cycles of nitrogen, carbon, water, minerals, food, and energy. *The Biosphere* discusses the dynamic equilibrium between the living and non-living parts of the earth and the necessity for maintaining this equilibrium.

1970, 134 pp., 101 illus., softcover, $4.50

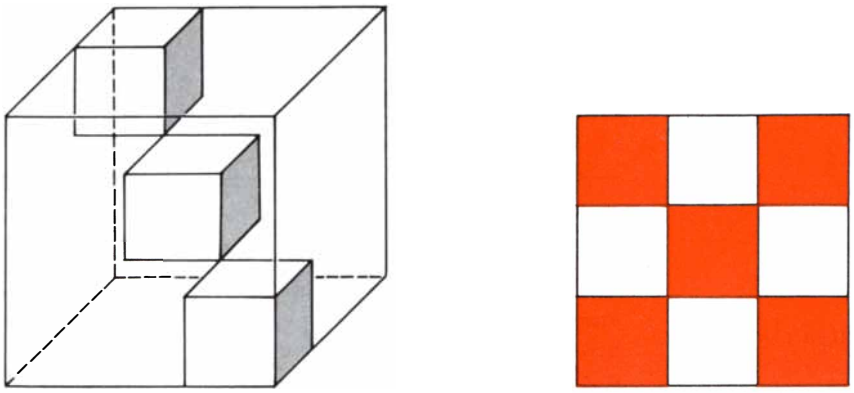
THE OCEAN

Here is a comprehensive, authoritative review of our modern scientific knowledge of the seas, with special reference to our new understanding of how they were created and our increasing utilization of their resources.

1969, 140 pp., 103 illus., softcover, $4.50

To order any of these books, use the adjacent card. If card is missing, send order with payment to:

W. H. FREEMAN AND COMPANY
660 Market Street, San Francisco, California 94104
58 Kings Road, Reading, England RG1 3AA



Key to the 3 × 3 × 3 packing puzzle

looks ridiculously easy, yet many find it irritatingly difficult. The reader is urged to construct a set, either by making the pieces out of wood or by gluing cubes together. Plastic cubes that snap together can be obtained from Creative Publications, Box 10328, Palo Alto, Calif. 94303. The company sells mathematics teaching materials, and it will send a catalogue on request.

The packing has a unique solution (not counting rotations and reflections) requiring that the three unit cubes be along a space diagonal [*see illustration above*]. To prove it, first consider any 3 × 3 cross section. If it is checkerboard-colored, five cells are of one color and four are of the other. No matter how a 1 × 2 × 2 brick is placed, it will occupy 0, 2 or 4 cells of each cross section, with the colors of the occupied cells evenly divided. As a result each of the nine sections must have one and only one cell occupied by a unit cube. (A section cannot contain all three cubes because that would force some sections to be without cubes.) In addition in every section the unit cube must be at the center or at one of the corners. The only way to meet these requirements is to place the cubes along a space diagonal. Placing the six bricks then follows automatically.

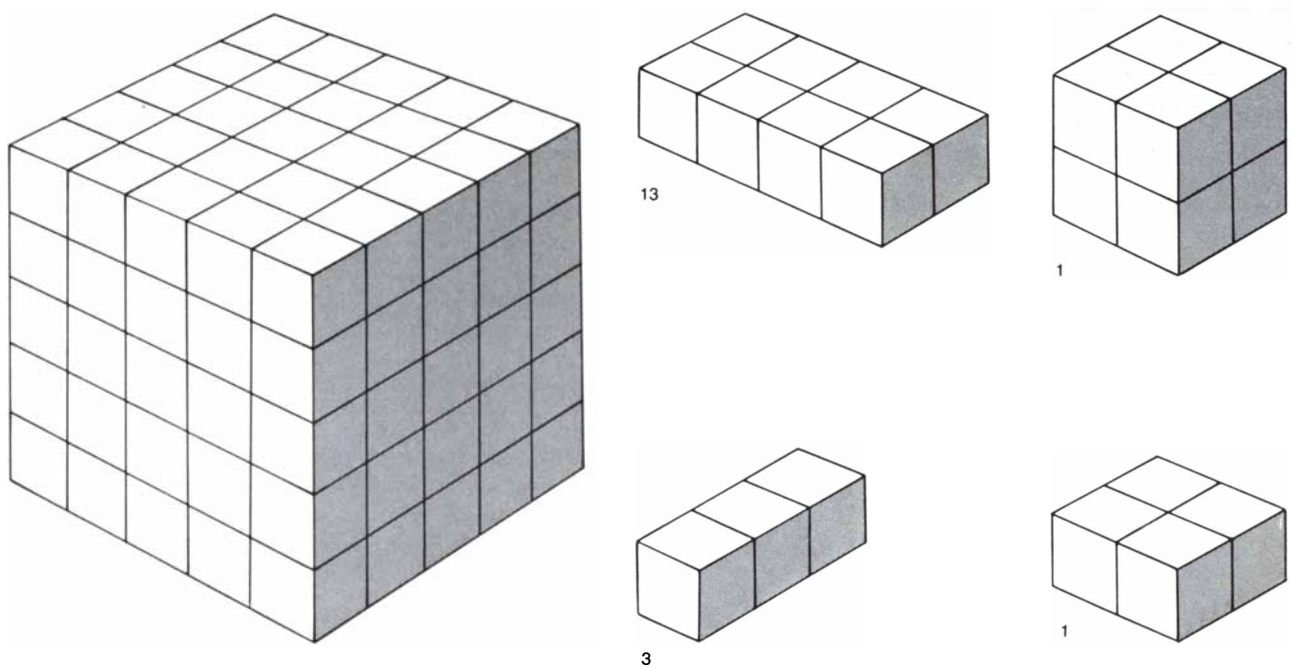
John Horton Conway of the University of Cambridge, who is well known to regular readers of this department, set himself the task a few years ago of designing a more difficult cube-packing puzzle. Conway's cube, although it appears to be almost as easy as the 3 × 3 × 3 puzzle, is so hard that some people cannot solve it until they are told its whimsical secret.

Conway's cube has many variants. The one he likes best requires 18 harmonic bricks [*see illustration below*]. The task is to pack them into a 5 × 5 × 5 box or, what is the same thing, to build an order-5 cube. It is fiendishly difficult if one tries to solve it by trial and error. Next month I shall give the key, and a proof of the key's necessity.

Returning to the packing of identical bricks, we can ask the following general question. If a box is not perfectly packable with a set of bricks, what is the maximum number of bricks that will go into it? Even when the bricks are harmonic, it is an extremely difficult problem, although a good start has been made in a paper published in 1974 by Richard A. Brualdi and Thomas H. Foregger.

The authors define a "representing set," abbreviated *R*, as a set of cells in the box such that no matter where a brick is placed it will occupy at least one cell of *R*. When *R* is made as small as possible, it is said to have minimum cardinality. The authors show that the maximum number of identical bricks (not necessarily harmonic) that will go into a box is equal to or less than the minimum cardinality of *R*.

In 2 space the maximum number of harmonic bricks (dominoes) invariably equals the minimum cardinality of *R*, but that does not hold for higher spaces. Consider canonical bricks and cubical boxes. The order-4 cube is the smallest that is packable, and of course the pack-



The 18 bricks of John Horton Conway's cube-packing puzzle

ing is trivial. The order-6 cube, as we have seen, is not packable, but all bricks but one are easily put inside.

What about the order-5 cube? It has 125 cells. That is not a multiple of 8, so that it is not fully packable by canonical bricks. Will 15 such bricks (total volume 120) go inside? To put it another way, can you build an order-5 cube with 15 canonical bricks and five unit cubes? The minimum cardinality of R is 15, but try as you will it is impossible to get 15 bricks inside. After 14 bricks have been placed the 13 remaining cells will never accommodate the last brick. Brualdi and Foregger have a complicated coloring proof of the impossibility, but in meditating about it one afternoon I had a happy inspiration. It led to the following *reductio ad absurdum* proof.

Assume that 15 bricks will pack. The total surface area of the order-5 cube is $6 \times 25 = 150$. Since each face has an odd number of squares, one cell on each face must be filled by a unit cube. (No more than one unit cube can occupy a cross section because there are 15 such sections and only five unit cubes.) This leaves a surface of $150 - 6 = 144$ to be packed by faces of the canonical bricks. Now each brick must have one and only one of its 1×2 ends on the surface. This leaves a surface of $144 - (15 \times 2) = 114$ to be packed by 1×4 and 2×4 rectangles. But 114 is not evenly divisible by 4. Therefore the original assumption is false.

A similar proof is obtained by considering the cube's three planar midsections. Each plane is a 5×5 matrix, making 75 cells in all. One cell in each 5×5 square must be intersected by a unit cube. (This could be done with a single unit cube at the center, or two or three unit cubes suitably placed.) Each of the 15 canonical bricks must intersect two cells, or 30 cells in all. Taking 33 cells from 75 leaves 42 to be intersected in sets of 4 or 8, and since 42 is not a multiple of 4, impossibility follows.

The next cube of interest is the order 7. It is easy to put 41 canonical bricks inside it, but will it hold 42, leaving seven holes to be filled by unit cubes? Surprisingly the answer is not known. Foregger posed this as an unsolved problem (E2524) in the March 1975 issue of *American Mathematical Monthly*. Although every canonical brick must intersect a pair of cells in the planar midsections, after the necessary subtractions are made the remaining cells are a multiple of 4. Therefore the previous impossibility proof does not apply.

Brualdi and Foregger found many special cases in which the maximum number of harmonic bricks equals the minimal cardinality of R. For example, if the smallest face of the brick packs each face of the box, there is equality. For canonical bricks there is equality if one of the box's dimensions is even. The general problem, however, is far from solved.

I can think of no better way to conclude than by quoting the final sentence of Klarner's article: "A word of warning in closing! Engaging in experiments with little wooden blocks is fraught with the danger that friends, family and colleagues will assume you are entering your second childhood, and that you should be put away. A good defense is to have a few copies of Conway's puzzle on hand to divert their attention while you make a getaway."

Shortly after last month's report on magic squares and cubes was written, Richard Schroeppel and Ernst G. Straus each independently constructed perfect magic cubes of order 7. It is not known if 7 is the lowest possible order. The existence of perfect magic cubes of order 5 and order 6 remains an open question.

Victor Meally called my attention to the construction of a perfect magic cube of order 8 in the late 1930's by J. Barkley Rosser and Robert J. Walker. It is pandiagonal in the sense that if any of its three sets of parallel sections are cyclically permuted, the cube remains perfectly magic. The cube was not published, but its construction is given in a manuscript deposited in a library at Cornell University. In the manuscript the authors show that perfect pandiagonal cubes exist for all orders that are multiples of 8 and all odd orders higher than 8. A brief summary of the construction technique for the order-8 cube is given in the last two editions of W. W. Rouse Ball's *Mathematical Recreations and Essays,* revised by H. S. M. Coxeter.

Rosser and Walker were not, however, the first to make a perfect pandiagonal cube of order 8. That was done by C. Planck and reported in his *Theory of Path Nasiks,* privately published in Rugby, England, in 1905. Planck showed that in k space, where k is 2 or higher, the smallest perfect pandiagonal "cube" is of order 2^k, and the smallest one that is also associative (symmetrically opposite pairs of numbers have the same sum) is $2^k + 1$. We saw last month how that is true of magic squares. In 3 space the smallest perfect pandiagonal cube is of order 8, and the smallest one that is also associative is of order 9. The perfect order-8 cube given last month is associative but not pandiagonal.

THE AMATEUR SCIENTIST

Thin-layer chromatography: a convenient way of separating things that are very much alike

Conducted by C. L. Stong

Thin-layer chromatography offers the most versatile means of separating a mixture into its component substances. The separation is accomplished on a thin layer of adsorbent material wetted with solvent. The technique is quite simple. The specimen is applied as a spot to the edge of the layer. That edge is then wet with the solvent, and the components of the mixture are separated as the solvent carries them across the surface.

The technique is within the reach of most amateurs in terms of cost, but it has a number of limitations. Roger Baker, Jr. (Box 7854, University Station, Austin, Tex. 78712), has therefore developed a new form of the technique, which he calls ultrathin-layer chromatography. The thinness of the layer allows the separation procedure to be reduced in scale so that smaller samples can be separated. Baker describes the technique and its uses as follows.

"The thin layers most commonly used in thin-layer chromatography are applied to a substrate such as glass in the form of a slurry of adsorbent and liquid (usually water) in much the same way that butter is spread on bread. For mechanical reasons, such as minute irregularities in the substrate, it is difficult to apply uniform layers less than about .1 millimeter thick by spreading a slurry.

"The method of making thin layers described here is based on an old technique for dusting resin powder onto copper plates for photomechanical printing. The basis of the procedure is to shake or agitate a powder of the adsorbent material, a silica gel, inside a box so that some of the powder is left suspended in the air inside the box [*see illustration on page 130*]. A certain period of time is allowed for the larger particles of the powder to settle to the bottom of the box. Then the substrates are introduced through a previously sealed door near the bottom of the box. An additional period of time is allowed for the finest dust particles to settle uniformly onto the substrates, which are then removed. The dusted substrate is chilled so that a film of water condenses on the surface, drawing the particles into close contact with the substrate and causing the layer to adhere when it is dry.

"The technique was developed with microscope slides (75 by 25 millimeters in size) as substrates. After a slide is coated the adsorbent layer of silica gel is wiped from the edges, leaving a central strip. This strip can be scribed into narrower strips, so that several dozen individual samples can be resolved on one slide simultaneously.

"After the samples are applied to the adsorbent layer near one end the slide is placed adsorbent side down over a shallow cavity covered on the bottom with blotting paper that has been saturated with solvent. Solvent vapor is induced to condense as a liquid at the end of the adsorbent near the sample by cooling the top surface of the glass slightly. The chromatogram is developed as the solvent is drawn by capillary force toward the dry end.

"Improvements in both sensitivity and resolving power can be achieved by marshaling the sample into a line before beginning the development. That can be done by applying a warm zone through the glass at the opposite side of the sample from the cool zone. Since the position of the solvent front is held stationary by evaporation caused by the warmth, the condensed solvent flows toward the solvent front, leaving the sample constituents deposited along a thin line at the front. When the warm zone is removed, development proceeds with the components separating into narrow bands rather than the overlapping round spots that would otherwise result. This step of sample concentration is called longitudinal concentration.

"A little-known variation of thin-layer chromatography, called vapor-impregnation-gradient chromatography, is in some ways similar to the solvent-gradient elution employed in column chromatography. This gradient effect is easily achieved with my apparatus. The gradient technique tends to compress the sample components into narrow bands as development proceeds, so that longitudinal concentration of the sample before it is developed is not necessary.

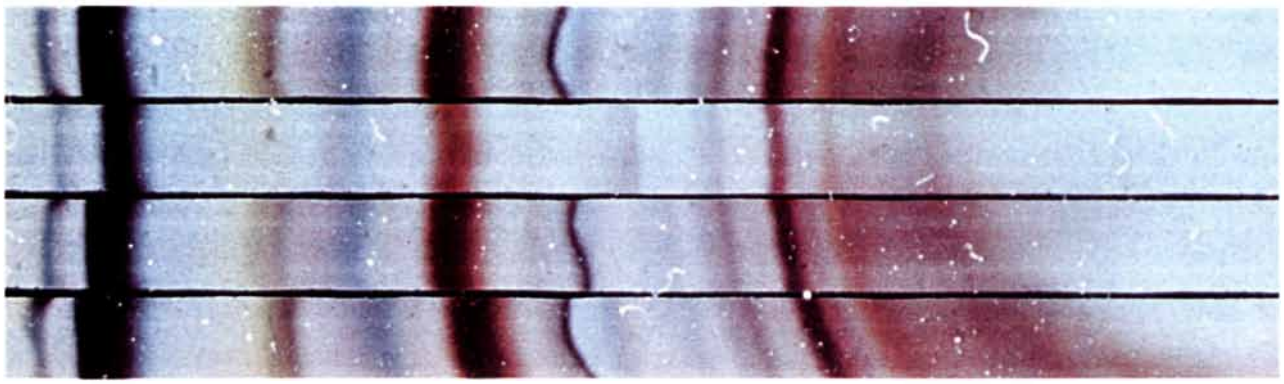
"The gradient technique requires that the developing chamber be lined with two sections of blotting paper soaked

C. L. Stong

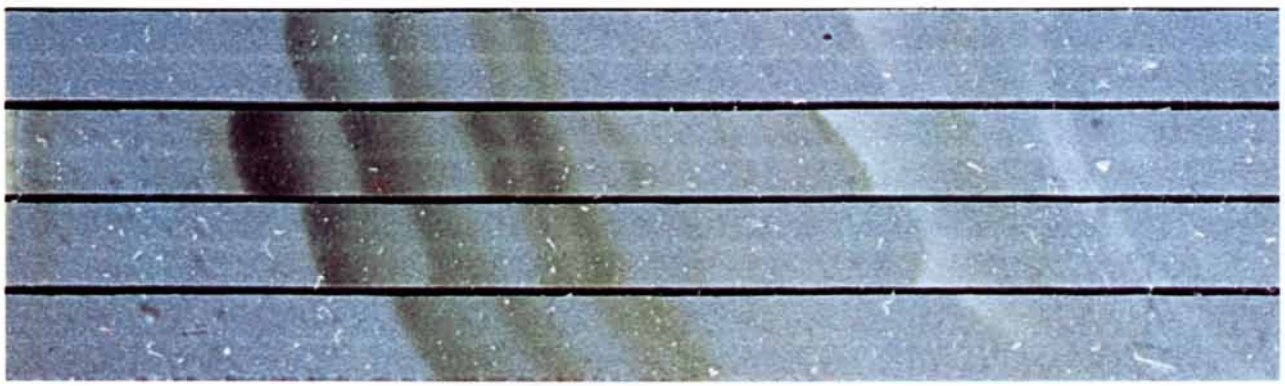
C. L. Stong, who had conducted "The Amateur Scientist" since 1957, died of lung cancer on December 9. His age was 73.

Stong, who was known to one and all as "Red," was born in Iowa, studied electrical engineering at the University of Minnesota, was a barnstorming pilot in the era of the Curtiss "Jenny" and worked for the Western Electric Company in various capacities for 36 years until he retired in 1962 to devote his full time to "The Amateur Scientist."

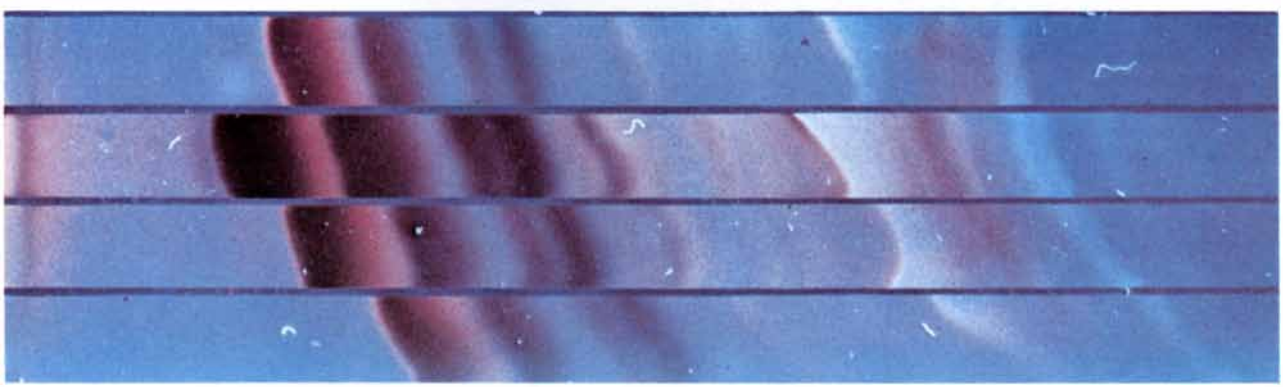
Stong was the amateur scientist personified. He delighted in all things pertaining to applied mathematics, to physics and to electrical, optical and mechanical devices. For the last he had a rare practical bent, a measure of which is the fact that he built and tested virtually everything described in "The Amateur Scientist," whether or not he was the original builder. One could be sure that something described in the department would work, because Stong had made it work himself. This installment of "The Amateur Scientist" is his last, and the department will be discontinued until further notice.



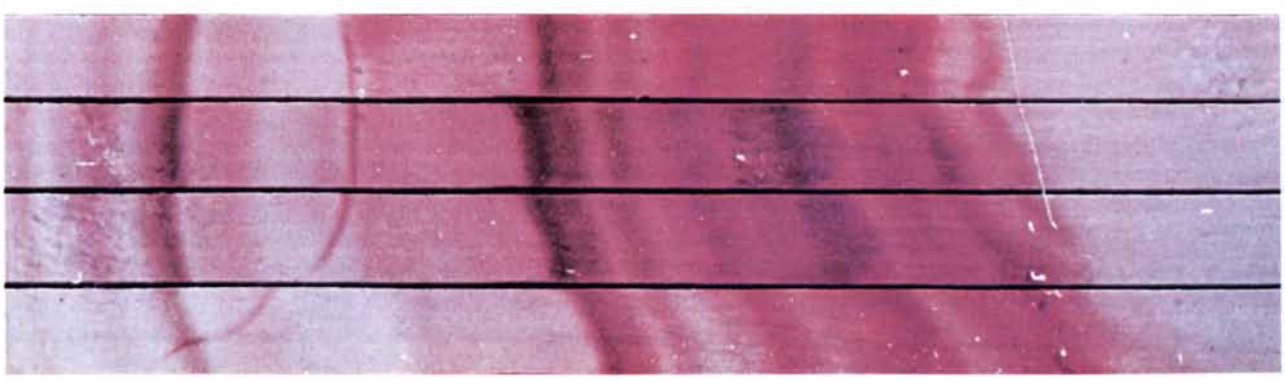
Components of ink separated on a thin-layer chromatogram. (White specks are artifacts.)



Components of grass-leaf extract in normal light



Components of the same extract in ultraviolet



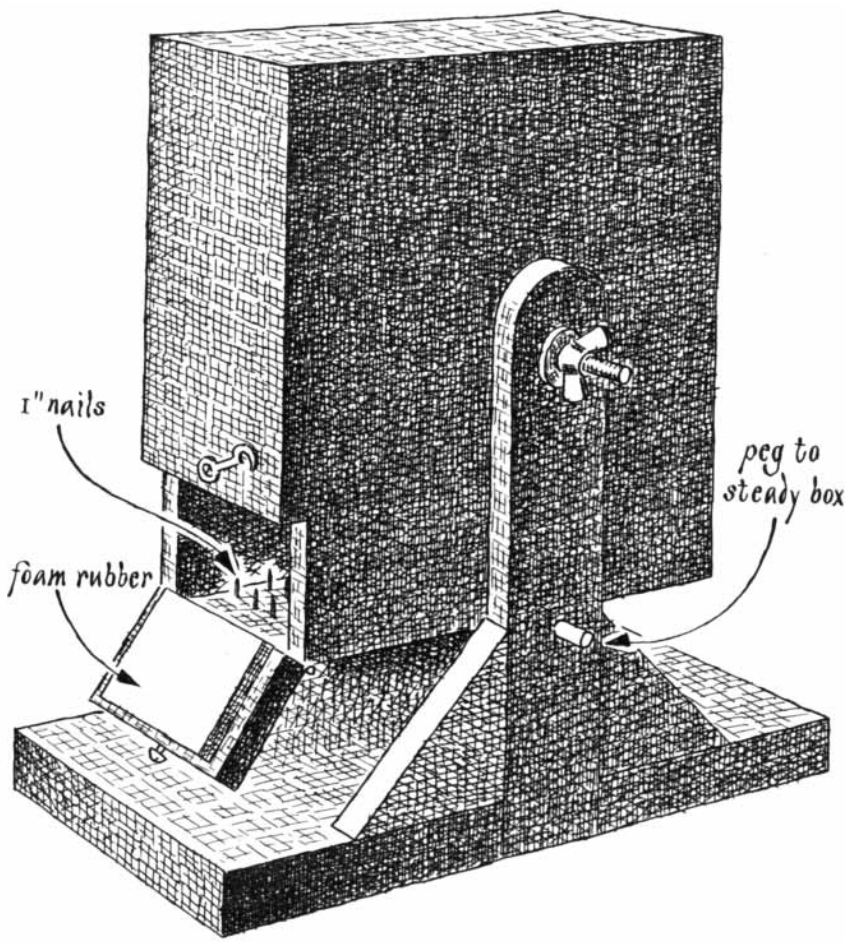
Components of grass-leaf extract stained with gentian violet

with different solvents. The two sections should not touch. At one end of the chamber, below the sample, install a short double-layer section of blotting paper soaked with a relatively strong polar solvent such as acetone. The rest of the length of the chamber is lined on the bottom with a longer strip of blotting paper soaked with a relatively weak nonpolar solvent such as naphtha (lighter fluid). Various solvent combinations are satisfactory, but they should be miscible in all proportions and chosen so that the components to be resolved migrate rapidly in the first solvent and slowly in the second.

"The sample is carried down the length of the chamber by the stronger solvent, which passes over the area of the chamber lined with the blotting paper soaked with the weaker solvent. An exchange of solvent vapors occurs, causing the migration to slow with distance. Certain components of the sample are able to migrate farther than others under these conditions, so that the components are separated into bands along the length of the adsorbent strip. In addition to the sharp separations yielded by the gradient technique a rather wide variety of substances can be separated with a simple solvent combination such as acetone and naphtha. Thus the search for solvents that yield satisfactory resolution is simplified.

"Whether the sample is resolved in the ordinary way with a single solvent or with a gradient, it is useful to apply the warm zone to the far end of the adsorbent strip in order to continuously remove the solvent by evaporation as it reaches the end of the chamber. In this way the development can be continued as long as it is wished, so that even the more slowly moving components will be moved far enough to be well separated.

"After the sample has been resolved chromatographically it is possible to concentrate the bands of components sideways with a procedure resembling the one used in longitudinal concentration. Hence the components are deposited along a line running the length of the adsorbent strip. The procedure is called lateral concentration; it yields a greatly increased sensitivity of detection



Box for dusting silica-gel layer onto glass slides

at the expense of a loss in resolution.

"An interesting method of detecting colorless materials that are not soluble in water is to stain them with a dye as they remain adsorbed on the silica gel. After the silica gel has been treated with a strong dye solution the excess dye is washed off, leaving the separated components visible as more highly stained areas on a light background. The method is sensitive, nondestructive and fairly general. It also gives a permanent record of the position and the relative amounts of sample components.

"The details of construction of the dusting box can be varied considerably without much affecting the results. I did my first experiments with a tin can. The settling periods involved should be increased in proportion to the height of the box.

"Satisfactory results can be obtained with a plywood box about 50 centimeters in height. The amount of silica-gel dust retained in suspension in the air, and accordingly the thickness of the layer produced during one coating operation, depend on factors such as the height of the box and the particle size of the powder. About a gram of adsorbent per 10 square centimeters of settling area is a suitable amount. The powder can be agitated by inverting the box end over end about a dozen times. The box should be tapped immediately after agitation to knock down any loose particles.

"After the initial settling period the cleaned microscope slides are introduced into the box on an aluminum plate. My plate holds 12 slides. The opening through which they are introduced should be sealed during agitation with a strip of foam rubber on a hinged door. Some kind of standoff, such as nails driven through the bottom of the box, is needed to support the aluminum plate above the bulk of the powder. With a dusting box 50 centimeters high, suitable results are obtained if the dust is allowed to settle for about four minutes before the substrates are introduced. The substrates should then remain inside the box for about 10 minutes to collect the adsorbent dust.

"The best procedure for dusting on the adsorbent layers has not been determined with certainty. On the basis of work that has been done with pure silica gel prepared for thin-layer chromatography, however, it seems best to apply the desired layer over several stages. Presumably other finely powdered adsorbents could be applied in much the same way.

"After an initial layer of adsorbent is

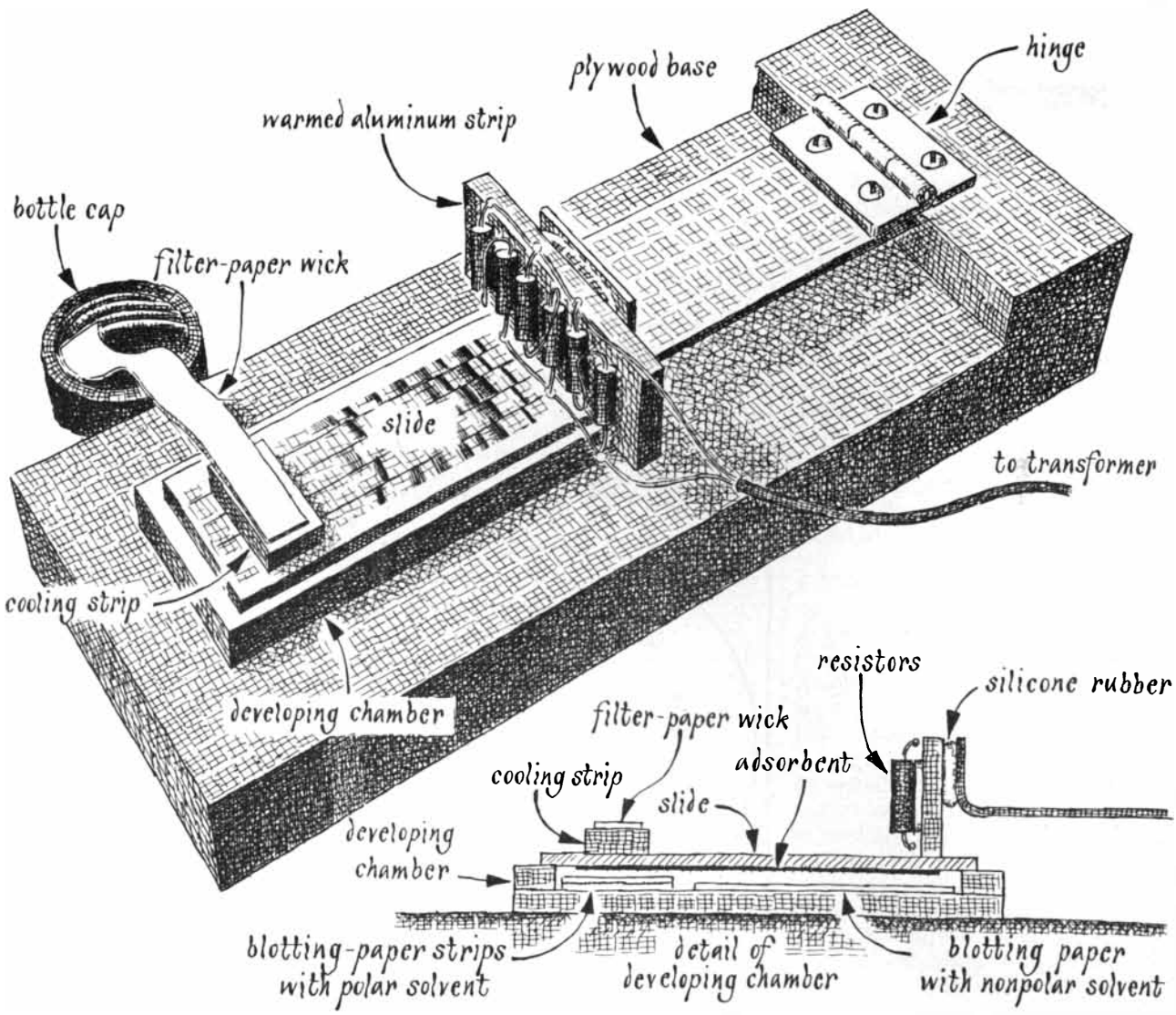
dusted onto the slides the aluminum plate on which they rest is chilled by placing it on a layer of crushed ice. The temperature of the glass is thereby lowered below the dew point of the air, causing a film of water to condense on the adsorbent layer. When the layer is wet and shiny, the excess water is wiped off the sides and bottom of the plate and the plate is dusted and wet by condensation once or twice more to thicken the final layer so that it is free of pinholes.

"The greatest uniformity is achieved by shifting the positions of the substrates between dusting steps. It appears that the layers should have a certain minimum thickness in relation to the size of the particles in order to minimize the effect of point-to-point variations in the capillary channels between the particles. The silica-gel layers that appeared to be the most useful for chromatographic purposes were from 15 to 50 micrometers thick. The thickness was measured with a micrometer both before and after wiping the adsorbent from a portion of the glass. A densitometer can be employed to compare the relative thicknesses of several layers.

"Since the slides are covered with adsorbent over their entire top surface at the time they are prepared, the adsorbent is wiped off each border for about five millimeters inward, leaving a central strip of adsorbent measuring about 65 by 15 millimeters. That can be done by sliding the end of a matchstick along the border, using your thumbnail as a guide. I made a tool for scribing lines about three millimeters apart along the length of the adsorbent strip by soldering a row of ordinary pins to a brass strip with the points extending like the teeth of a comb. A pin at the edge extends a little farther so that it rides along the edge of the slide and guides the rest of the pins. Any excess dust is blown off. The slides are stored in a microscopist's slide holder.

"The developing chamber can be machined from a solid metal such as aluminum. My chamber was constructed from two pieces of aluminum an eighth of an inch thick measuring 79 by 28 millimeters. A rectangular hole 70 by 18 millimeters was cut in the center of one of them with a jeweler's saw. The two pieces were cleaned well and cemented in close contact with epoxy glue to form the chamber [*see illustration below*]. The top surface of the chamber, on which the slide rests, was ground flat with fine abrasive paper and finished by rubbing it against a flat piece of glass covered with a paste of kitchen cleanser and water.

"The sample solution should be applied to the adsorbent layer about 15 millimeters from one end. I make the



Developing chamber designed by Roger Baker, Jr.

application with a fine glass capillary drawn over a flame from larger tubing and having a square-cut or broken end. It can be difficult to apply the sample to the adsorbent layer from such a hand-held capillary without disturbing the layer, particularly when the sample is to be applied to the narrow scribed strips. A small plywood jig makes the job easy. The jig consists of a microscope slide cemented to a piece of plywood [*see bottom illustration on this page*]. Another slide prepared with adsorbent can be made to adhere to the top of this first one with a small drop of naphtha.

"A small plywood bridge is made to straddle the slides and extend upward at a right angle to the surfaces. A glass or metal tube is fastened to the bridge with a rubber band. The tube acts as a guide for the capillary pipette, which is slid through until it touches the adsorbent surface at the desired place. By holding the jig so that the slide is nearly vertical and the capillary nearly horizontal and then tilting the two slowly, you can delicately adjust the height of the liquid column in the capillary, and therefore the rate of flow of the sample into the adsorbent, to match the rate of evaporation of the solvent from the sample spot.

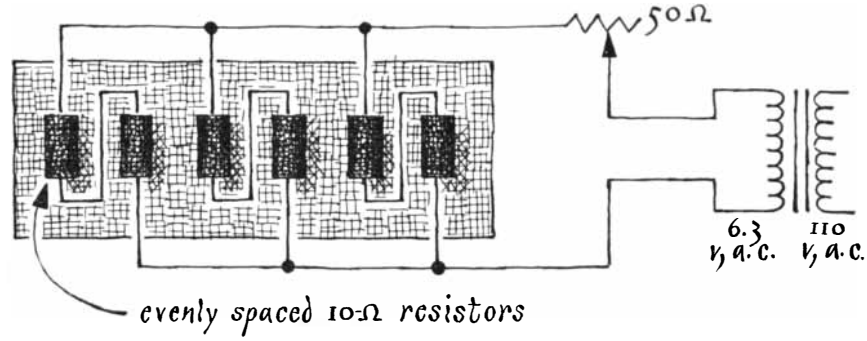
"The glass is cooled to initiate condensation of the solvent by placing the adsorbent strip face down on the developing chamber and covering the glass above the end of the adsorbent strip (just beyond the sample spot) with a small, flat piece of aluminum about 15 by 10 millimeters. This piece of aluminum is cooled slightly by covering it with a strip of filter paper about a centimeter wide that extends over to a bottle cap filled with water alongside the chamber. The strip acts as a wick, drawing fresh water from the bottle cap and cooling the aluminum by evaporation.

"The time required for development will vary according to how rapidly the components migrate, how quickly the solvent is condensed and whether or not the experimenter is working with the gradient technique. It may be sufficient to simply let the solvent travel down the length of the strip, which might take 15 minutes. Often it is desirable to apply the warmth at the end of the strip, so that the solvent is continuously removed as it reaches the end and development is continued for a longer time. Since the migration rate of the components is restrained by the conditions of gradient development, an hour or more is often necessary for this type of development.

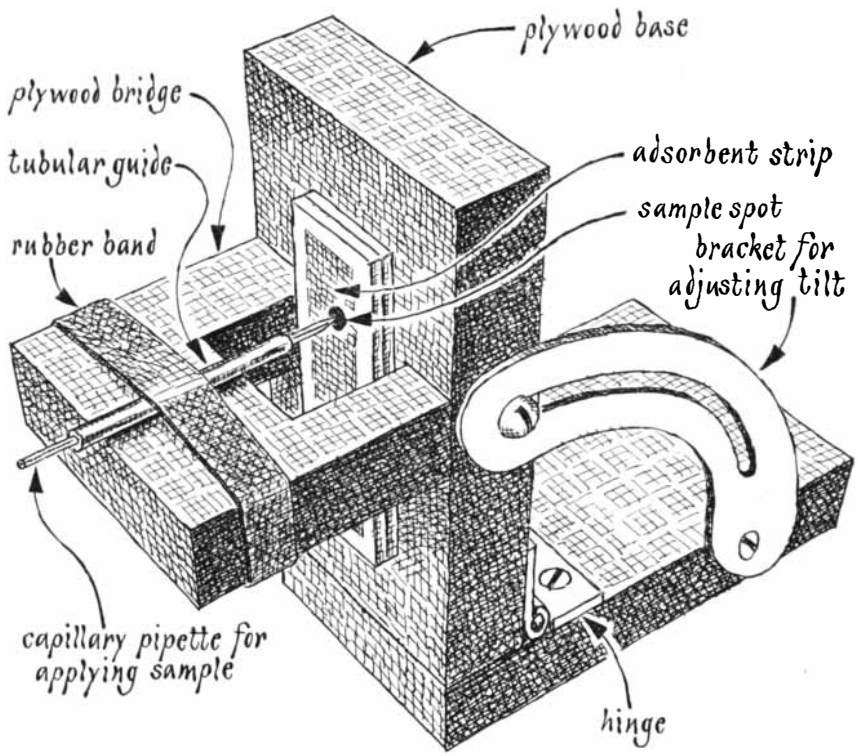
"It is convenient to mount the warmed aluminum strip at the end of a hinge so that it is free to swing down and rest gently on the glass. The strip is about 70 millimeters long, 20 millimeters wide and 1/16-inch thick. It carries six .5-watt, 10-ohm resistors attached with silicone rubber, parallel to one another and evenly spaced [*see illustration above*]. They are wired together in three groups of two. The warmth of the strip is controlled by wiring all three of the groups in series with a one-ampere, 50-ohm rheostat (Ohmite). Everything is then connected to the low-voltage output of a 6.3-volt filament transformer (Radio Shack).

"My method of detecting the location of colorless components of the sample is to stain them with a water-soluble dye such as rhodamine B. The stain can be applied by dipping the slide into a strong water solution of dye. The stain is allowed to dry on the slide. Then the slide is rinsed with a dilute solution of acetic acid in water, which will remove most of the dye from the background areas of silica gel and cause the areas that have adsorbed material from the sample to appear darker.

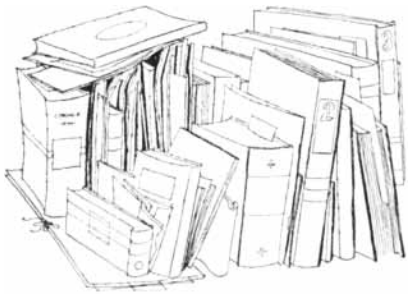
"Ordinary thin-film chromatography of black ink from a ball-point pen, with benzene as the solvent, is capable of resolving about 14 component pigments. The gradient technique, with acetone and naphtha as the solvents, can resolve the same ink into more than 20 pigments. The same solvent gradient will serve to separate acetone extracts of plant leaves into various yellow and gray pigments and as many as 10 distinct blue or green pigments."



Circuitry for warming aluminum strip



Jig for applying a sample to a slide



BOOKS

Uri Geller: international Pied Piper of the credulous, and other matters

by Philip Morrison

THE MAGIC OF URI GELLER, by The Amazing Randi. Ballantine Books ($1.75). SUPERMINDS, by John Taylor. The Viking Press, Inc. ($10.95). MY STORY, by Uri Geller. Praeger Publishers ($8.95). URI: A JOURNAL OF THE MYSTERY OF URI GELLER, by Andrija Puharich. Bantam Books, Inc. ($1.95).

Raffiniert ist der Herr Gott, aber Boshaft ist Er nicht. Einstein said it: The Creator is subtle, but he is not malicious. That is the faith by which the physicist lives. There is order deep in his world, which sufficiently cunning experiment will disclose. But that is no stance in which he can safely buy a used car, appraise the operative statements of a White House press officer or bet against an artist with the cards. Not Einstein, but Machiavelli and W. C. Fields must be his guides then. For things will not be as they seem, plausible assumptions of logical independence ("How could he know?") will usually fail, and systematic mischief will rule the little world. Human beings—unlike electrons—can understand the intent of an observer. In particular, conjurers know the questions likely to be in the minds of the audience and contrive to capitalize on them for their own purposes. The hand is not often quicker than the eye, but the left hand you are coaxed into not watching can work wonders.

These four books span an international misdirection industry that is now a couple of years old in the U.S. It is chiefly a construct of the media, out of celebrity and talk shows on television in the U.S., over the BBC and half a dozen other national networks, and articles in such publications as the *Psychic News, Der Spiegel* and *Paris-Match*. The furor centers on a quick-thinking, handsome, charming, bold but disarmingly informal former Israeli paratrooper named Uri Geller. There are a dozen claimants to this or that occult power in the public eye today, but Geller earns the attention of a general scientific reader because in large measure he offers not mere video wonders but credentials issued by physicists. He has been studied by the Electronics and Bioengineering Laboratory of the Stanford Research Institute (often confused with the great university it neighbors) in work published at length in *Nature*. A few well-known professors of physics, mainly in London, have similarly published their experiences with Geller, which they view as being at least enigmatic and at most revolutionary.

The main Geller effects form a short list: reproducing concealed drawings; bending spoons, keys and the like; starting stopped watches; deflecting magnetic compasses, and recording high count rates on Geiger counters. It is evident that in themselves these effects are ordinary: anyone can achieve those results sans the occult. The issue turns on how he does them: on control. Geller and his believers say he does these things by extraphysical means. Critics observe that almost all the effects belong to the familiar repertoire of stage magicians. The physicists—with no record at all of previous successes in this domain—assert that their experimental skills rule out the paltry tricks of the prestidigitator.

Enters now the Amazing Randi, an elfin gentleman with a white-fringed chin and a flashing wit, a magician of long experience and patent ingenuity. He pursues Geller and finds simply a fraud, "a clever magician, nothing more—and certainly nothing less.... Uri's pattern of deception was unmistakable. Clever, yes. Psychic, no."

That ESP of Geller under rigid experimental control? First of all the think tank's shielded room is rendered suspect by their unwillingness to let Randi examine it. The room, which was built for another experiment, has a passage for cables, a passage stuffed with gauze. Geller's one perfect telepathic hit on a sketch of a bunch of grapes, correct to the very number of grapes in the bunch, was too good. For that particular one, Randi feels, "Shipi made a small drawing of the target—an exact drawing—and pushed it through the hole to Geller." (Shipi is Geller's closest associate, a quiet young man, unobtrusive but almost always around.) The psychologists at SRI—not the physicists—ran a double-blind test, genuinely controlled. Geller did not attempt a specific guess on any of their 100 sealed envelopes. Charles Rebert of the life-sciences division at SRI suspects even the one hit Geller obtained once conditions were somewhat loosened in response to the pleas of the physicists: "I'm convinced that he [Geller] swapped envelopes on us.... He flunked our test."

The watch-starting? It works often on studio watches, and it is widely contagious. In Britain, the U.S., West Germany, Japan—in many places—derelict watches long stopped responded nationwide, and happy viewers telephoned in to report the miracles. But in New Zealand: "We sent out a few students to fix broken watches, and out of 16 watches tried, 14 of them started, and seven have kept going for at least four to five days." Stopped watches are often jammed; winding, shaking and holding them enough to warm the oil will get them temporarily under way. Uri takes the credit, as he did in his famous stopping of the funicular near Munich. A reporter suggested that. They tried it out, Geller in the moving car, concentrating. Sure enough, "the cable car stopped right in midair.... The main switch had flipped off for no explainable reason." So Geller writes in his own story. Did someone flip that switch? We cannot say, but it was certainly no bent-metal miracle.

So the chronicle goes. When it seems that controls are working, Geller often fails. When he succeeds, there has been uncertainty in the controls, often total collapse. Sometimes he suddenly carries a key off to a nearby water tap; he has found that running water helps keys to bend. Perhaps, but it is surely true that more torque can be placed on a key when its end is inserted into the opening of the faucet. Try it yourself.

How does Geller himself explain his successes? We have his autobiography. It seems that from an early age, since a preschool epiphany in the garden, he has

had these strange powers. Since his hypnosis in 1971 by his first American sponsor, the psychic expert Andrija Puharich, he has received mysterious self-erasing messages on tape cassettes; they foretell his future and explain his past. (Puharich's book, probably the most implausible deadpan chronicle to see print in a long time, confirms Geller fully and has identified his sources. Some live on the spacecraft *Spectra,* "fifty-three thousand sixty-nine light ages away.") Geller writes: "The tapes bring with them a tremendous message to the world, even though they do sound like science fiction." One such pregnant message, self-activated and self-erased, is: "The ultimate powers, whether on the particle level or the cosmic level, are in rotation and drawing off the gravitational power from the center of the system. There are special rays... where the skin of the envelope of the cosmic rays is utilized for power." Automatic taping seems no more instructive, although it is less objective, than automatic writing was in the old days.

Then there is John Taylor, a reputable London theoretical physicist with an easy pen and a theatrical bent. His last book was on black holes for the layman. (It impressed this reviewer as being wildly uncritical and shamelessly tendentious, with three final chapters on spiritualism, on Christianity and on dialectical materialism, all as they are affected by black holes. The too stern Hume might have wanted to burn such a work, but one has to admit that these are not demonstrable errors but legitimate moral opinions.) The current book goes much further. It centers on the theory and practice of metal-bending. No references are made to Geller's life or Puharich's witness; instead Taylor, hopelessly misdirected by what he has "seen," undertakes to study metal-bending empirically. Measure the bent rod with care.... His main discovery is a group of child metal-benders, who will do the Geller bit willingly and repeatedly for Taylor. To be sure, they are 11-year-old innocents, and so he lets them go out of his sight to bend their samples, even go home over the weekend. Sometimes he comes along, but (as Randi quotes Taylor) "this feature of bending not happening when the object is being watched—'the shyness effect'—is very common." Common indeed! Lately a group from the University of Bath has published experiments with little volunteer metal-benders, drawn like Taylor's from the great British pool recruited by Geller over television. These small adepts, mainly innocent young girls, went to work bending metal. The observer did not see anyone cheating. But the skeptics from Bath had taken one precaution: the room where the experiment was being conducted was being carefully watched by several other observers through one-way mirrors from outside. Result: Five of the six gifted youngsters cheated visibly. They hastily bent metal by hand, tabletop and chair edge, presumably whenever they thought the overt observer was distracted. (The sixth subject bent no metal at all.)

Taylor dimly grasps the idea of experimental control. He sealed some metal samples in plastic tubes. There one is, before and after wonder-bending *in absentia,* a full-page color spread. Alas, Randi visited Taylor and popped the rubber stopper out of such a tube almost inadvertently, without disturbing the sealing wax (which turns out to hold a screwhead and not the stopper itself); "it was a very poor piece of preparation." Randi's drawing and the photograph in Taylor's book form a compelling sequence demonstrating that professors expert in the theory of a complex variable should leave sealing wax to magicians and private detectives.

It has all happened before; it will all happen again. We can laugh at the comedy and the incongruity; we can weep a little at the pomposity, the credulousness, the air of callous exploitation and the crash of reputations. Randi cites a wonderful parallel or two. The two great psychophysicists Weber, then advanced in age, and Fechner, partly blind, were taken in by Henry Slade, a medium who caused knots to appear in a string, the ends of which had been tied in a loop and sealed with wax, after many days of storage. According to the psychologists' astrophysicist colleague (the theorizing Taylor of that day), one Johann K. F. Zöllner, the medium Slade simply had access to the fourth spatial dimension! The magicians of the time knew better, as they do today. A new loop had been substituted. Randi draws a systematic methodological parallel between the aspects of the Geller case and the events of a century ago. His chapter is a telling guide to recognition of the central phenomena.

Not everything is made clear. Randi not only reads sealed envelopes and bends keys unnoticed but also can drive down the street while he is blindfolded with pizza dough and three layers of tested blindfolds. He will not explain that at all. His work is not unique, but magicians by long tradition will not expose their contrivances; they hold that the appeal of their clever art depends on guile. Indeed, Randi has already gone too far for many of his confreres, who see in Geller a new approach to the public. Not human ingenuity but the hint of mysterious power is what draws audiences today. A magician can do better with a bumbling, nervous performance. ("The mentalist who is too perfect loses credibility.... Stress, as does Uri Geller, that what you are attempting to do is little understood and doesn't always work on command.... They want you to succeed." The elegantly flawless card trickster is being supplanted by a performer who simulates an earnest but uncertain parapsychological striver. The article cited (from an American magicians' periodical) goes on more cynically: "Find out what their budget will allow them to pay. Ask at least that much!... Later, concentrate on the best-paying audiences."

The Geller wave seems now to ebb. The contrast between the galactic claims and the limited, banal performance in the end drains interest. Another will arise, and another, just as waves at the shore. Perhaps we shall learn at least this: The physicist is no person to ferret out the truth once *Boshaft* enters. Credentials ought to be more specific; a plasma physicist does not pronounce on the tracks of the magnetic monopole, at least until he has etched many a film of plastic. Clumsy-fisted theorists who trust the innocence of children ought not to set themselves up as knowing investigators. They will generate only more thick documents of human credulousness. Finally, there is no canon of logic or inference from common sense or use of past experience that can speak more strongly to the mind than the simple will to believe.

THE CAMEL AND THE WHEEL, by Richard W. Bulliet. Harvard University Press ($16). THE WHEELWRIGHT'S SHOP, by George Sturt. Cambridge University Press ($13.95). The labyrinth of the *casbah* is a commonplace thrill to travelers struck by the intricately winding streets, steep and narrow alleys, quaint stairs and culs-de-sac that nucleate most old cities from Fez to Kabul. Moreover, until recently the building and upkeep of surfaced roads outside the city was not usually taken as a serious task by the indigenous governments of North Africa and the Middle East. Used to the rectilinear city plans of the Roman world, convinced that road building was a chief task of the state, scholars have sought deep within Islam some inner reason for this fancied deficiency of public order.

The witty, reflective argument of *The*

Camel and the Wheel explains the facts by a surprising proposition. The economics of bulk transport in a dry climate favors the pack camel over the wheeled vehicle drawn by ox or horse. Heavy oxcarts hauled the city's goods from remotest antiquity until about the fifth century, after which the camel took over until the days of Otto and Benz. The level, straight streets of the Frankish world gave the carters access for heavy loads to all points in the city, once the streets were made wide enough to admit one rigid axle, or even two passing. But if the surefooted camel alone serves your town, you have no need for squared corners, wide streets or gentle grades. Local considerations, from wind protection to the direction of shade, from high habitation density to easy defense, take over; narrow, curving ways, little street markets and steep, changing grades were all incorporated efficiently into the medieval cities of Islam. It was no philosophical mystery; economic rationality determined. After all, Mecca itself, holiest of Islamic cities, is built with straight streets —a design older than camel transport.

The support of this proposition is manifold, drawing from art and archaeology, old texts and new, camel husbandry, anthropological studies, linguistic forms. The doubter must look at the photograph of a line of pack camels, haughty and graceful under heavy loads, swinging past a bemired horse cart on a bad Chinese road. The changeover was made wherever camels could be effectively bred to yield their competitive edge over wheeled transport, a matter of fully 20 percent. The camel is faster and stronger, its pack-saddle is cheaper than a wagon, its driver is more easily shared among half a dozen animals. Domestication of the camel had taken place very much earlier, probably in southern Arabia and the eastern horn of Africa. The ancient overland trade in "frankincense and myrrh" brought the nomad's camel to the desert borders of the settled lands of Syria and Mesopotamia by the seventh century B.C., but nearly 1,000 years passed before the creaky oxcarts of the ancient Middle East gave way to the silent caravan. It was a social transformation that did it: the sway of "the camel-borne warrior mounted upon a North Arabian saddle." The Arab nomads who bred camels and understood them became the cultured and powerful sovereigns of all Islam. Once the camel was available its productivity was telling, and along the narrow ways of the camel a wheelless society grew. Camel-mounted field artillery, fired from the back of the couched animal, existed in the Persian army well into the 19th century, and that same army, without wheels or roads for loads exceeding the camel's standard quarter ton, "carried unformed metal on camelback to the scene of battle and there cast it into seige guns." What counts to a prosperous and well-governed wheelless state in an arid region is not road surfacing; it is bridges, caravanserais and police protection for the traveler. (Compare, *passim*, *The Thousand and One Nights*.)

The book excels in its remarkable levelheadedness and its critical and candid writing. The esoteric topic has often attracted—as the bibliographical essay and the notes here fully demonstrate—not mere enthusiasts but true zealots. Professor Bulliet, a Berkeley Arabist, escapes from the trap and recognizes fully the danger of one-sided explanations and easy correlations. His critical stance is of interest to historians of technology and to students of social change in general. The nature of the camel lies at the base of the argument, of course. The reader will learn that camels store not water but heat; they tolerate both remarkable increases of body temperature and severe water loss, their desert endurance coming from both traits; no sweat for camels!

The era of the pack camel is near its end, done in by that other energy gift of the Arabian Gulf. But camel efficiency, evolved by nature and by human insight in that climatic region, may yet reveal the camel in the role of meat producer, "and the disdainful expression on its face will ever call to mind the past era when the camel was superior to man's proud invention, the wheel."

Precisely how the wheel could be the everyday load-bearer of an agricultural economy in the south of England is the burden of George Sturt's classic and beautiful evocation (first published 50 years ago) of the craftsmen he knew in Surrey in the 1880's, plying a trade in blacksmith's iron and in the elm, ash, bone-hard beech and stout oak of the local woodlands. That trade—making wagons, carts, plows and barrows—was itself the adaptive growth of six centuries or more, a matured folk craft whose hardworking, low-paid, unlettered but skilled and subtle journeymen and masters knew their way from collective experience and tradition alone, without so much as a single calculation. Rules and gauges were used, mostly undivided, but no computations entered, not even in the pricing of the goods, which often went for "less than the mere iron and timber. ... They never knew.... These matters ...were settled by guess work, not by calculation." Over the unchanging generations errors would even out and a reasonable charge emerged, but change came very hard. The keen drawshave deftly broke every corner of each squared timber and reduced each shaft and "pillar" until it looked right. "A man skillful with the draw-shave enjoyed this work, lingering over it like an artist." It was not decoration, however; its vital object was to lighten the horse's load, and the corner-shaving removed an eighth of the dead weight of the wagon timbers. Every longitudinal piece of a wagon was slightly curved, following ancient wood patterns long stored in the shop. The timber was not steam-bent; rather it was chosen for its natural curves. ("How accurately woodland nature seemed to know the shape of moving horses.")

Across this detailed account of the demanding love of hand for work Sturt paints the picture of social relations in such a shop, where the working man found deep pride, "that satisfaction which of old... streamed into their muscles all day long from close contact with iron, timber, clay, wind and wave, horse-strength."

BUTTERFLIES, by Thomas C. Emmel. Alfred A. Knopf, Publisher ($35). THE DICTIONARY OF BUTTERFLIES AND MOTHS IN COLOR, by Allan Watson and Paul E. S. Whalley, with an introduction by W. Donald Duckworth. McGraw-Hill Book Company ($29.95 until May 31, $39.95 thereafter). These large, costly and colorful volumes both treat the insect group Lepidoptera with deliberate and successful visual emphasis, but in two distinct styles, which one might call hot and cool. Emmel is a zoologist from the University of Florida; on the dust-cover of his book one sky blue butterfly against a green leaf is shown as big as your two hands held out flat. The creature is a tiny European butterfly, however. The book opens with a full-page photograph, bled to the edge, of elegant spherical eggs of one species magnified a couple of dozen times, followed by similar striking enlargements of the caterpillar, the pupa and the imago. More than 300 color photographs embellish the book, almost all showing living butterflies, with some wider shots to evoke rain forest or mountain meadow. Caterpillars and chrysalids are included, and a few extraordinary micrographs in color exhibit individual wing scales at high magnification, the examples glowing like feather cloaks for some forest king. The text, studded with maps and drawings, develops the elements of butterfly life

cycle and behavior and describes the main families. With a few pages for each habitat, the text goes on to give an overall account of the world's butterflies of rain forest and desert, temperate forest and grassland, mountains and tundra, oceanic islands, all the regions where these beauties are found. Within these pages most of the pictures appear, some as big spreads but most at modest size, six or eight to the page. The crux of the text is in its last chapter, on collecting. Here the engaged author evokes the bittersweet delight of butterfly collectors over the centuries, famous Churchill, gifted Nabokov and Will C. Minor, a sheepherder from Fruita, Colo. "Minor was once offered $25 each for pairs of perfect specimens, but on the day he and another collector went up to Black Ridge to get freshly emerged butterflies, they could only stand helplessly by and watch the valuable butterflies made worthless as pugnacious male pairs battled furiously above the edge of a cliff." This book glows with warmth; it explains as it captures wonder and beauty; indeed, it celebrates butterflies with the almost hyperbolic skill of the photographers worldwide and of the Milanese color printers.

The second book, by two British Museum experts, with a Smithsonian curator writing the brief introduction, is a cool one, equally devoted and as colorful. The book is in fact a visual dictionary. About half of its pages display (the color printers this time are Dutch) some 1,000 forms of winged moth and butterfly. No eggs, no caterpillars, no long shots at all; the aim is identity. Many of the photographs are made from life, but a great number are made from the "flawless set specimens" arrayed on pins in the British Museum.

The rest of the volume is given over to an alphabetical list of some 2,000 entries, each of which hangs on the master peg of the proper binomial identification paragraph or two of information about the insect listed. (Cross-references are given for hundreds of common names.) Once again the delight of the collector is well served. The general text occupies only a few pages, but the list and plates of the British book sweep a much larger net across the fluttering world. There are some 165,000 known species of the Lepidoptera; only about an eighth of them are the generally showier, benign and day-flying butterflies. The rest are moths, large and small, which receive only the sketchiest of allusions from Emmel but are fully sampled in the other volume.

The fundamental problem for any describer of the scalewings is that extraordinary plentitude of forms. Between these two sizable books there is really not much overlap of species. Only one publication "attempts to cover the world fauna of both butterflies and moths": the 16 volumes (with more to come from Stuttgart) of A. Seitz's *The Macrolepidoptera of the World,* a set now costing $2,200. The general reader will gain much from either of the present books, however. One is more like an instructive album of the theater, the other more like a well-labeled jewel chest.

Enough of books; a little of the animals themselves, and of what they have meant to their students and admirers. Their very variety is an inexhaustible lesson in adaptation; even the specific food plants are still unknown for most forms. The energy source for most butterfly species in adulthood is the nectar of a blossom, for which payment is usually made by pollination. Many longer-lived forms need protein as well, and some are lured by rotting fruit, bird droppings or urine-soaked ground. The long-domesticated forms are moths: the unremarkable-looking silkworm proper and the tussore moths, the chief variety of which is a five-inch orange beauty. (Silkworms will feed on lettuce or the osage-orange leaves as well as on mulberry.) We look away from the corn borer, the gypsy moth and the Mexican jumping-bean moth. The prince of collectors must have been Carl Linnaeus himself. He named one North American species of brimstone butterfly *Gonepteryx eclipsis.* Alas, the type insects are still preserved among his specimens, and they prove to be familiar European forms with new wing patterns beautifully painted in by hand. He even cited as a reference a book where exactly his hand-painted forms are carefully represented. Who fooled whom? We do not know.

MISSILES AND ROCKETS, by Kenneth Gatland. Illustrated by the John Wood Studio, principal artist William Hobson. Macmillan Publishing Co., Inc. ($6.95). From a map of the "Peenemünde Rocket Establishment (circa 1943)" to a colorful airbrush rendering of the external tank of the Space Shuttle Orbiter, the 80 color plates of this pocket-size hardback exhibit the look of the world's ballistic missiles, chiefly military, of course. Here they are, from the V-2 to the shuttle, with some cutaway and exploded views, with their engines, carriers, silos, auxiliary structures such as telemetry dishes and blockhouses and with a number of useful maps and tables. The text is brief and knowing, an eyewitness account of visits by the author to sites worldwide, and a capsule history (taking a warlike view of political developments that is implausibly hard-boiled). The work is scrupulously international, arising as it does out of the London world of aerospace journalism. The Guiana Space Centre of the French near Kourou is mapped along with Tyuratam and Vandenberg, although the Tanegashima Space Center of Japan is only described, and the Chinese launch center is merely located on the map of China.

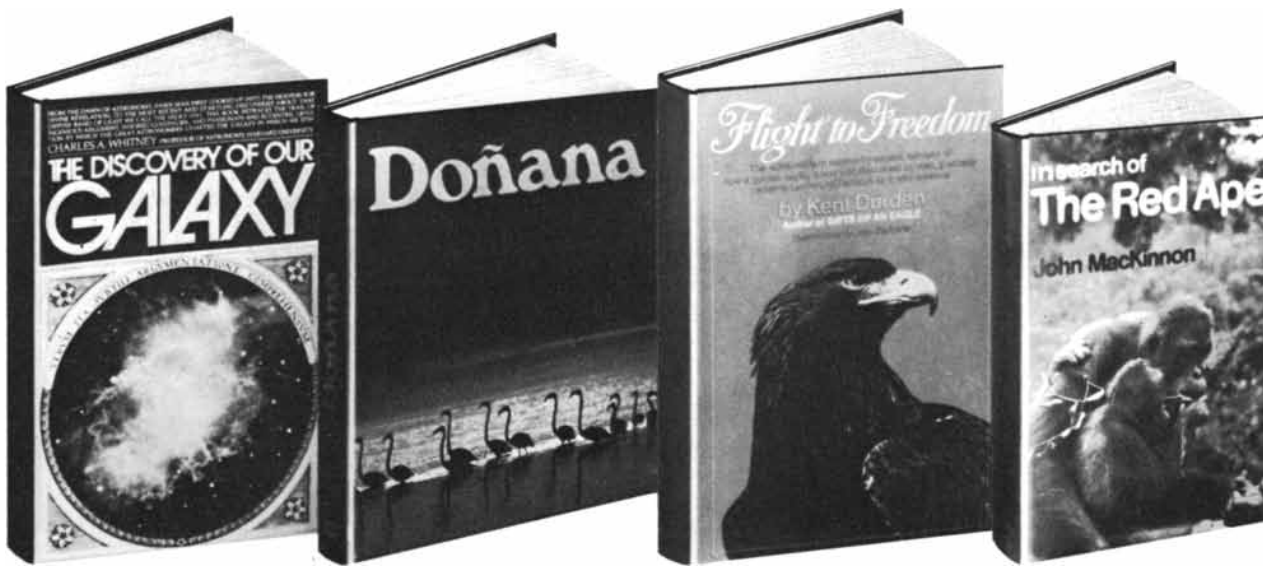
A two-page color spread, in the familiar style of a carefully rendered, detailed, retouched photograph, shows the cosmodrome at Tyuratam: the rocket on the pad, the Soyuz payload high in the nose, the big hinged support arms drawn away, the fuel tank cars in place, all making ready the manned launch. The escape system and the flight sequence follow. There is no doubt that this presentation of Russian practice makes more real to any reader that powerful technology we too often associate with enigmatic news reports and an occasional picture of the parade in Red Square. It is interesting to study the portraits and tables that recount the uprating of the Russian ICBM of 1957 from Sputnik to Voskhod and Soyuz. The Type *A* launcher grew more and more powerful upper stages, but it retained the four strap-on boosters and the central sustainer, each of these mounting four thrust chambers employing liquid oxygen and kerosene. The general conception of this very reliable vehicle, like its total thrust, reminds one of a more powerful and more versatile Atlas. *Salyut* or the new Venus landers use a newer, larger launch vehicle, painted here provisionally "in the absence of official Soviet information" much more like the U.S. Saturn *IB* or the newly boosted versions of Titan.

The missile submarines are here with their solid-fuel rockets, similar enough whether they are French, British, U.S. or Russian. Convergence is at work in this domain of evolution as well as among marsupial predators. The two superpowers quantitatively dwarf the other two by about an order of magnitude in boats and missiles. The Russian Delta-class nuclear submarine "seems to have become operational in January, 1973," with a missile range of 4,000 nautical miles, not enough to reach the U.S. from the Barents Sea. Thus have we justified our new Tridents and the long-range *C-4* missile, its prime contractor Lockheed. The arms race is given terrifying and fascinating concreteness in this small book of reference.

Get Away From It All...

as a member of

Natural Science Book Club



Investigate a pulsar. ***Wander in Europe's greatest wildlife preserve.*** ***Fly with an eagle.*** ***Live with an orangutan.***

(Publishers' Prices shown)

84285. TIDELINE. *Ernest Braun.* Beautifully photographed and written 9-by-12-inch volume on that long, thin strip of wilderness where land, air, and sea meet. Contains almost one hundred full-color photographs vividly depicting one day's change along the tideline. **$16.50**

54240. HUNTING FOR DINOSAURS. *Zofia Kielan-Jaworowska.* The three-year effort to find and decipher the meaning of the ancient bones that litter the Gobi desert. Crammed with fascinating descriptions of dinosaurs and their evolution, and the team's work and discoveries. **$12.00**

64875. NUBIAN RESCUE. *Rex Keating.* Dramatic story of the race to save hundreds of Egyptian monuments from the rising waters of the Assuan Dam. Fully illustrated. **$12.50**

58900. LOOKING AT ANIMALS. *Hugh B. Cott.* Containing over 150 pen-and-ink drawings and photographs (many in color), this lavish volume explores through word and picture the world's richest wildlife area—East Africa. **$15.00**

46500. THE EVOLUTION OF RADIO ASTRONOMY. *J.S. Hey.* One of the pioneers of radio astronomy offers an absorbing history of this powerful technique's remarkable development," . . . an important work in the history of astronomy." M.I. Meeks, in *Sky and Telescope.* **$10.00**

78820. THE SOCIAL BEHAVIOR OF THE BEES. *Charles D. Michener.* First definitive work on these most amazing insects. A publishing classic. *Counts as 2 of your 3 books* **$25.00**

34400. ARCHAEOLOGY BY EXPERIMENT. *John Coles.* Spotlights archaeological experiments from the science's primitive beginnings 150 years ago to the highly sophisticated simulated Iron Age settlements that have been established today. **$7.95**

42325. THE DISCOVERY OF OUR GALAXY. *Charles A. Whitney.* A distinguished Harvard astronomer recreates the lives and discoveries of great astronomers from the sixth century B.C. to black holes, pulsars, supernovae, and Seyfert galaxies. **$10.00**

42680. DOÑANA: Spain's Wildlife Wilderness. *Juan Antonio Fernandez.* Sumptuously illustrated portrait of one of the world's most important wildlife sanctuaries. Stunning 9-by-11-inch volume contains over 240 pages of full-color photographs. *Counts as 2 of your 3 books.* **$29.95**

84210. THIS LIVING REEF. *Douglas Faulkner.* Lavishly illustrated word-and-picture masterpiece on one of the South Pacific's most beautiful island chains. Outsized volume contains 120 pages of glorious six-color pictures—all with detailed fascinating commentaries. *Counts as 2 of your 3 books.* **$25.00**

88740. ZOO. *Bernard Livingston.* Delightful wide-ranging account of how zoos began, how they are run today, and where they are heading in the future. "A must for zoo buffs and animal lovers." —*Roger Caras.* **$8.95**

48675. FLIGHT TO FREEDOM. *Kent Durden.* "Like the eagles he loves, Durden is a vanishing species. He has expertise, he sees what he observes, and puts both into words that make unforgettable reading. In this book he even surpasses his first classic, *Gifts of an Eagle.*"—*George S. Small.* **$5.95**

36090. THE BEST OF BABCOCK. *Havilah Babcock.* A delightful collection of the matchless outdoor pieces by the late humorist, naturalist, and distinguished teacher, Havilah Babcock. A classic. **$8.95.**

37195. BY THE EVIDENCE. *L.S.B. Leakey.* Fascinating autobiography of one of the century's greatest paleontologist's life and work from 1926 to the early 1950s. **$9.95**

34720. ATLAS OF ANCIENT ARCHAEOLOGY. *Edited by Jacquetta Hawkes.* With the aid of over 350 maps, site plans, and drawings, two dozen internationally renowned archaeologists highlight 170 major sites throughout the world. **$19.50**

55120. IN SEARCH OF THE RED APE. *John MacKinnon.* A former field researcher for Jane Goodall tells of his exciting three-year study of the orangutan in the wild jungles of Borneo and Sumatra. **$7.95**

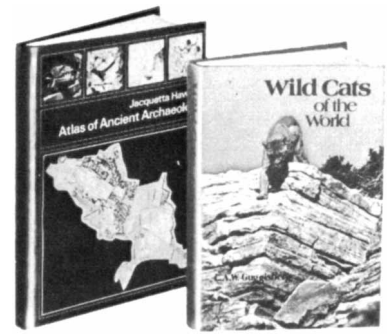
87195. WILD CATS OF THE WORLD. *C.A.W. Guggisberg.* One of the world's leading authorities on wild cats offers definitive natural histories of each of the 37 species of wild cats. **$15.95**

Take any 3 books

(values to $51.95)

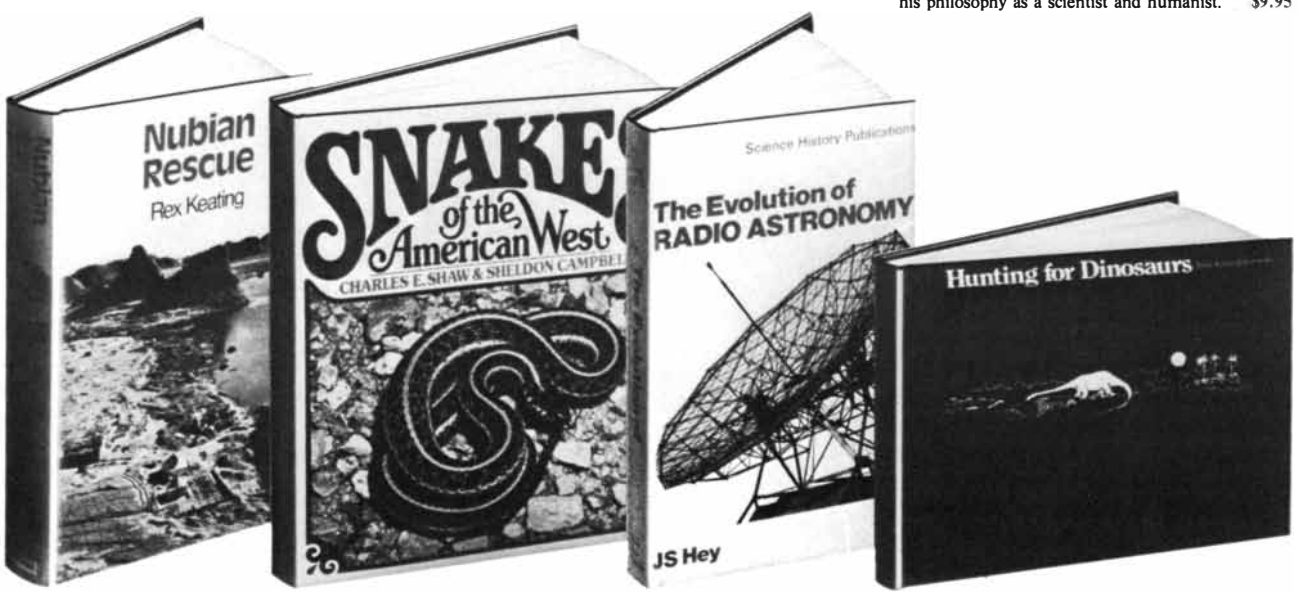
for only 99¢ each

if you will join now for a trial period and take three more books —at member discount prices— within the next 12 months



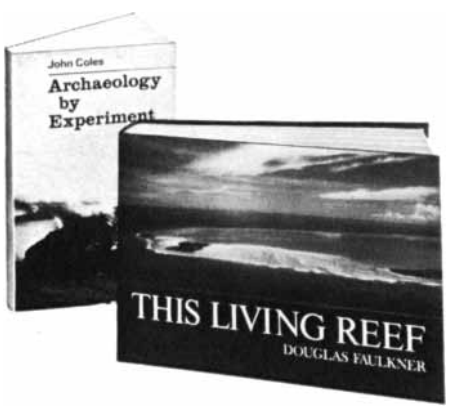
78680. SNAKES OF THE AMERICAN WEST. *Charles F. Shaw and Sheldon Campbell.* Two skilled herpetologists explore virtually everything that is known about the snakes of the American west, their habits and habitats, their appearances and characteristics. **$12.50**

32618. ALL THE STRANGE HOURS. *Loren Eiseley.* Autobiography of one of America's outstanding naturalists and scientists reveals the stunning Eiseley mind focusing on his formative, adventurous youth, his creative period as a writer and thinker, and his philosophy as a scientist and humanist. **$9.95**



Rescue an ancient monument. ***Make friends with a rattler.*** ***Pioneer in a new way to look at the stars.*** ***Catch a dinosaur.***

IMPORTANT GUARANTEE! Examine your 3 introductory books for 10 days. If not delighted, simply return them and your membership will be canceled. You pay nothing, owe nothing!



85820. VENOMOUS ANIMALS OF THE WORLD. *Roger Caras.* One of America's pre-eminent nature writers explores the entire scope of this remarkable group of animals—from the common honeybee and ant to such exotic, and often deadly, creatures as the black widow spider. *Counts as 2 of your 3 books.* **$25.00**

How the Club Operates • The Book Club **News**, describing the coming Main Selection and Alternate Selections, will be sent to you 15 times a year at 3-4 week intervals. • If you wish to purchase the Main Selection, do nothing and it will be shipped to you automatically. • If you prefer one of the Alternates, or no book at all, simply indicate your decision on the reply form always enclosed with the **News** and mail it so that we receive it by the date specified. • The **News** is mailed in time to allow you at least 10 days to decide if you want the coming Main Selection. If, because of late mail delivery of the **News**, you should ever receive a Main Selection without having had the 10 day consideration period, that Selection may be returned at Club expense. • Continued membership past the trial period brings you the added benefits of our Bonus Book Plan—an important way to save even more, at least 70% off publishers' prices.

Natural Science Book Club
Riverside, New Jersey 08075 **4-5AE**

Please accept my application for membership and send me the three volumes indicated, billing me only 99¢ each. I agree to purchase at least three additional Selections or Alternates during the first year I am a member, under the club plan described in this ad. Savings range up to 30% and occasionally even more. My membership is cancelable any time after I buy these three books. A shipping and handling charge is added to all shipments. **Send no money. Members are billed when books arrive.**

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

A few expensive books (noted in book descriptions) count as 2 choices.

Name _____

Address _____

City _____ State _____ Zip _____

(Offer good in Continental U.S. and Canada only. Prices slightly higher in Canada.)

BIBLIOGRAPHY

Readers interested in further reading on the subjects covered by articles in this issue may find the lists below helpful.

THE ETHICS OF EXPERIMENTATION WITH HUMAN SUBJECTS

EXPERIMENTATION WITH HUMAN SUBJECTS. Edited by Paul A. Freund. George Braziller, Inc., 1970.

EXPERIMENTATION WITH HUMAN BEINGS: THE AUTHORITY OF INVESTIGATOR, SUBJECT, PROFESSIONS, AND STATE IN THE HUMAN EXPERIMENTATION PROCESS. Jay Katz. Russell Sage Foundation, 1972.

RESEARCH ON HUMAN SUBJECTS: PROBLEMS OF SOCIAL CONTROL IN MEDICAL EXPERIMENTATION. Bernard Barber, John J. Lally, Julia Loughlin Makarushka and Daniel Sullivan. Russell Sage Foundation, 1973.

HUMAN SUBJECTS IN MEDICAL EXPERIMENTATION. Bradford H. Gray. Wiley-Interscience, 1975.

THE RECEPTORS OF STEROID HORMONES

RECEPTORS FOR REPRODUCTIVE HORMONES. Edited by Bert W. O'Malley and Anthony R. Means. Plenum Press, 1973.

STEROID-CELL INTERACTION. R. J. B. King and W. I. P. Mainwaring. University Park Press, 1974.

THE ROLE OF ESTROPHILIN IN ESTROGEN ACTION. Elwood V. Jensen, Suresh Mohla, Thomas A. Gorell and Eugene R. De Sombre in *Vitamins and Hormones*, Vol. 32, pages 89–127; 1974.

FEMALE STEROID HORMONES AND TARGET CELL NUCLEI. Bert W. O'Malley and Anthony R. Means in *Science*, Vol. 183, No. 4125, pages 610–620; February 15, 1974.

IS GRAVITY GETTING WEAKER?

SECULAR ACCELERATIONS OF THE EARTH AND MOON. Robert R. Newton in *Science*, Vol. 166, No. 3907, pages 825–831; November 14, 1969.

COSMOLOGICAL MODELS IN A CONFORMALLY INVARIANT GRAVITATIONAL THEORY—II: A NEW MODEL. F. Hoyle and J. V. Narlikar in *Monthly Notices of the Royal Astronomical Society*, Vol. 155, No. 3, pages 323–335; 1972.

NEW APPROACH TO RELATIVITY AND GRAVITATION. Hüseyin Yilmaz in *Annals of Physics*, Vol. 81, No. 1, pages 179–200; November, 1973.

A DETERMINATION OF THE RATE OF CHANGE OF *G*. Thomas C. Van Flandern in *Monthly Notices of the Royal Astronomical Society*, Vol. 170, No. 2, pages 333–342; February, 1975.

FORAGE CROPS

FORAGES: THE SCIENCE OF GRASSLAND AGRICULTURE. Edited by Maurice E. Heath, Darrel S. Metcalfe and Robert F. Barnes. Iowa State University Press, 1973.

FORAGE MANAGEMENT IN THE NORTH. Dale Smith. Kendall/Hunt Publishing Company, 1975.

UTILIZING PLANT AND ANIMAL RESOURCES IN PRODUCING HUMAN FOOD. W. F. Wedin, H. J. Hodgson and N. L. Jacobson in *Journal of Animal Science*, Vol. 41, No. 2, pages 667–686; August, 1975.

ROBOT SYSTEMS

EXPLORATORY RESEARCH IN INDUSTRIAL MODULAR ASSEMBLY: REPORT R-800. James L. Nevins, Daniel E. Whitney et al. The Charles Stark Draper Laboratory, Inc., 1974.

PROCEEDINGS OF THE FIFTH INTERNATIONAL SYMPOSIUM ON INDUSTRIAL ROBOTS: SEPTEMBER 22–24, 1975. Society of Manufacturing Engineers, 1975.

A NEW APPROACH TO MANIPULATOR CONTROL: THE CEREBELLAR MODEL ARTICULATION CONTROLLER (CMAC). J. S. Albus in *Transactions of the American Society of Mechanical Engineers: Series G*, Vol. 97, No. 3, pages 220–227; September, 1975.

THE FINAL PALEOLITHIC SETTLEMENTS OF THE EUROPEAN PLAIN

DIE JUNGPALÄOLITISCHEN ZELTANLAGEN VON AHRENSBURG. Alfred Rust. Karl Wachholtz Verlag, 1958.

DIE STIELSPITZEN-GRUPPEN IM NÖRDLICHEN MITTELEUROPA. Wolfgang Taute. Böhlau-Verlag, 1968.

ESQUISSE DE LA PALÉOGÉOGRAPHIE DE LA PLAINE DE L'EUROPE SEPTENTRIONALE DANS LE TARDIGLACIAIRE. Romuald Schild in *Przeglad Archeologiczny*, Vol. 21, pages 9–63; 1973.

CARBENES

CARBENES, NITRENES AND ARYNES. T. L. Gilchrist and C. W. Rees. Thomas Nelson and Sons, 1969.

CARBENE CHEMISTRY. Wolfgang Kirmse. Academic Press, Inc., 1971.

CARBENES: VOL. I, VOL. II. Edited by Maitland Jones, Jr., and Robert A. Moss. Wiley-Interscience, 1973, 1975.

GAS-PHASE CARBENE REACTION. Maitland Jones, Jr., in *Accounts of Chemical Research*, Vol. 7, No. 12, pages 415–421; December, 1974.

THE BIOLOGICAL CLOCK OF INSECTS

PHOTOPERIODISM AND SEASONAL DEVELOPMENT OF INSECTS. A. S. Danilevskii. Oliver & Boyd, 1965.

THE PHYSIOLOGICAL CLOCK: ENDOGENOUS DIURNAL RHYTHMS AND BIOLOGICAL CHRONOMETRY. Erwin Bünning. Springer-Verlag New York, Inc., 1973.

THE PHOTOPERIODIC CLOCK IN THE FLESH-FLY, *SARCOPHAGA ARGYROSTOMA*. D. S. Saunders in *Journal of Insect Physiology*, Vol. 19, No. 10, pages 1941–1954; October, 1973.

CIRCADIAN OSCILLATIONS IN CELLS AND THE CIRCADIAN ORGANIZATION OF MULTICELLULAR SYSTEMS. Colin S. Pittendrigh in *The Neurosciences: Third Study Program*, edited by Francis O. Schmitt and Frederic G. Worden. MIT Press, 1974.

EVIDENCE FOR 'DAWN' AND 'DUSK' OSCILLATORS IN THE *NASONIA* PHOTOPERIODIC CLOCK. D. S. Saunders in *Journal of Insect Physiology*, Vol. 20, No. 1, pages 77–88; January, 1974.

INSECT CLOCKS. D. S. Saunders. Pergamon Press, in press.

MATHEMATICAL GAMES

FILLING BOXES WITH BRICKS. N. G. de Bruijn in *The American Mathematical Monthly*, Vol. 76, No. 1, pages 37–40; January, 1969.

MATCHING PROBLEMS. G. Katona and D. Szász in *Journal of Combinatorial Theory: Series B*, Vol. 10, No. 1, pages 60–92; February, 1971.

BRICK-PACKING PUZZLES. David A. Klarner in *Journal of Recreational Mathematics*, Vol. 6, No. 2, pages 112–117; Spring, 1973.

PACKING BOXES WITH HARMONIC BRICKS. Richard A. Brualdi and Thomas H. Foregger in *Journal of Combinatorial Theory: Series B*, Vol. 17, No. 2, pages 81–114; October, 1974.

THE AMATEUR SCIENTIST

THIN-LAYER CHROMATOGRAPHY. Kurt Randerath, translated by D. D. Libman. Academic Press, Inc., 1966.

"Unbelievable."

The Aspen sedan has the look of a fine European road car.

The Aspen sedan was designed with as much attention to its outside appearance as to its inside engineering. The shape of its classic lines and the high sweep of its glass give a luxurious look to this finely engineered automobile.

The Aspen sedan has more total room than a Cadillac Seville.

There's plenty of room to seat six people quite comfortably. And that's more than enough to make an Aspen sedan the envy of some very expensive cars.

The Aspen sedan is priced lower than Ford Granada or Mercury Monarch.

Based on manufacturer's suggested retail price for Aspen sedan, excluding state and local taxes, destination charge, and optional equipment.

The Aspen sedan has a ride that rivals that of a full-sized car.

The unique Isolated Transverse Suspension is rubber isolated to reduce noise as well as vibration transmitted to the passenger compartment. This gives Aspen the comfortable ride you usually find in a much bigger car.

The Aspen sedan got up to 27 MPG on the highway according to EPA estimated mileage results.

According to EPA estimated mileage results, the Aspen sedan and coupe got 27 MPG on the highway and 18 city. The wagon got 30 MPG highway and 18 city. All were equipped with a 225 Six and manual transmission. (Your actual mileage depends on driving habits, condition of your car, and options. In California, see your Dealer for mileage results.)

The Aspen sedan offers the important convenience features of a luxury car.

A long list of options includes everything from power seats and windows to electric door locks and automatic speed control. So you can enjoy the same conveniences as you would in a big luxury car.

Here's "The Clincher."

"For the first 12 months of use, any Chrysler Corporation Dealer will fix, without charge for parts or labor, any part of our 1976 passenger cars we supply (except tires) which proves defective in normal use, regardless of mileage." The owner is responsible for maintenance service such as changing filters and wiper blades.

Dodge

ASPEN

The new Dodge Aspen. For a small car at a small price, it's unbelievable.

THE SILVER MARTINI.
FOR PEOPLE WHO WANT A SILVER LINING WITHOUT THE CLOUD.

SMIRNOFF SILVER. NINETY POINT FOUR PROOF. SMIRNOFF LEAVES YOU BREATHLESS.®

SMIRNOFF SILVER® VODKA, 90.4 PROOF. DISTILLED FROM GRAIN. STE. PIERRE SMIRNOFF FILS. (DIV. OF HEUBLEIN, INC.) HARTFORD, CONN.