The VASIMR Rocket: Riding Plasma to Mars • AIDS Drugs for Africa

SCIENTIFIC AMERICAN

NOVEMBER 2000 $4.95 www.sciam.com

Late-Breaking News

CLONING ENDANGERED SPECIES

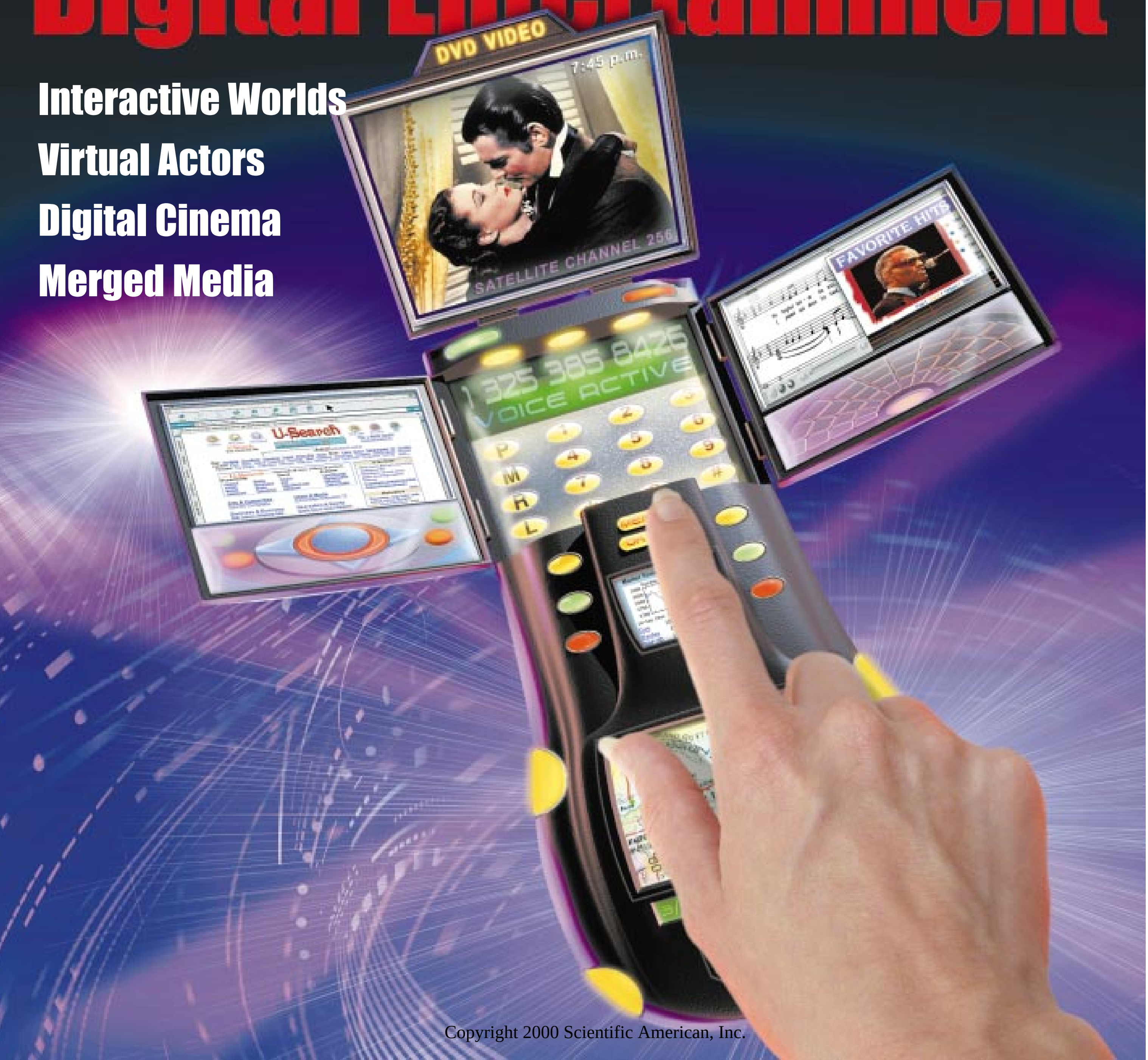
THE FUTURE OF Digital Entertainment

Interactive Worlds

Virtual Actors

Digital Cinema

Merged Media

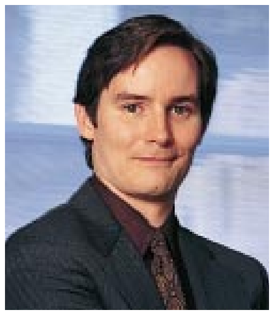


EDITOR_JOHN RENNIE

Cloning and Conservation

Talk of cloning typically inspires speculation and worry about duplicating people. How anthropocentric of us. Other animal species could benefit from cloning technology, too, maybe long before humans do. As the article by Robert P. Lanza, Betsy L. Dresser and Philip Damiani describes, beginning on page 84, it is now possible to clone animals that are on the edge of extinction. Optimists are even hopeful that they might be able to clone some animals that are slightly over that edge, having vanished within recent decades.

The process for multiplying endangered animals—some rare panda, for example—is probably not exactly what might have been envisioned most commonly in science fiction. We can't (yet?) just pluck any cell from our panda and then grow a whole animal from it. Cloning depends on merging DNA from a body cell into an egg cell stripped of its own DNA, then implanting this composite into a female for gestation. On the face of it, that's not necessarily any help, because the females of an endangered species (and their ova) are by definition in short supply. Conventional breeding and artificial insemination would generally still be easier. But that bottleneck can be avoided by borrowing an egg cell and a nurturing womb from a closely related nonthreatened species. Researchers hope soon to be able to point to gaurs born from cows, ocelots born from South American cats called oncillas, and so on. This approach may not work for all species, but it could help pull many back from extinction.



ERICA LANSNER

Paradoxically, cloning could save species while making conservation harder.

So the potential of cloning to preserve species is terrific, and yet it does not solve the endangered species problem. In extreme cases, it could even make matters worse.

How worse? Cloning can be used to help perpetuate an endangered species. But it might also eventually be used, miraculously, to resuscitate a species that survives as no more than a sample of cells frozen in liquid nitrogen. Forgive my paranoia, but I can imagine a future time in which a land-use developer argues that there is no reason to worry about the disappearance of a given species in the wild because we can always resurrect it later through cryogenics and cloning—whereas we need that ranch land now.

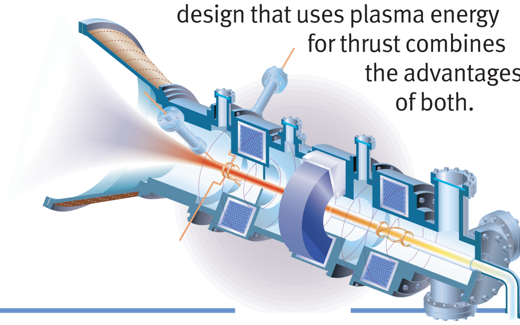
The charismatic pandas, ocelots, tigers and other creatures that decorate ecology posters are most important as bellwethers for their disappearing habitats. Hunting and other human activities may target endangered species in some cases, yet most species face more of a threat from the broad, indiscriminate pressure exerted by the encroachment of our homes, roads, farms, ranches and factories. In saluting the wonderful value of cloning as a conservation tool, let's not forget that real conservation involves preserving the life and lands we might least think to save.

editors@sciam.com

SCIENTIFIC AMERICAN
November 2000 Volume 283 www.sciam.com Number 5

Contents

SPECIAL REPORT



Contents

SCIENTIFIC AMERICAN

November 2000 Volume 283 www.sciam.com Number 5

About the Cover Illustration by Philip Howe and Mac Congrave. Photofest (*Gone with the Wind*). © Derick A. Thomas; Dat's Jazz/Corbis (*Ray Charles*).

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 2000 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of the publisher. Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 242764. Canadian BN No. 127387652RT; QST No. Q1015332537. Subscription rates: one year $34.97, Canada $49, International $55. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. Reprints available: write **Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; (212) 451-8877; fax: (212) 355-0408 or send e-mail to sacust@sciam.com Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.** Printed in U.S.A.

EDITORS@SCIAM.COM

SEARCHING FOR EXTRATERRESTRIAL LIFE

Your article on the Fermi Paradox ["Where Are They?" by Ian Crawford] failed to mention what may be the most important rejoinder to it. In a paper published in *Icarus* in September 1981, William I. Newman and Carl Sagan analyzed how fast a spreading interstellar civilization would expand through our galaxy. They based their work on mathematical models covering everything from the diffusion of molecules in a gas to the observed spread of animal species introduced into virgin territories on the earth.

Newman and Sagan found that how fast the galaxy fills up depends surprisingly little on the speed of interstellar travel. The limiting factor is that there are too many planets to be settled and filled up along the way. A key point is that each step of colonization will not necessarily be directed radially outward from the starting point but instead toward the nearest empty target.

ALAN M. MACROBERT
Bedford, Mass.

Crawford replies:

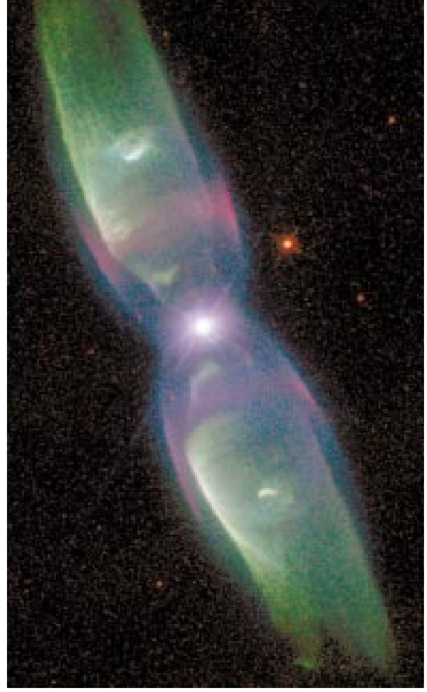
It is true that the colonization timescale is relatively insensitive to the assumed starship speed, depending mainly on the colony consolidation time. But the 1981 Newman and Sagan paper, together with some of their later work and that of Eric M. Jones of Los Alamos National Laboratory, does not support vastly longer colonization times. Indeed, my own estimates were based on a simplified equation provided by Newman and Sagan themselves in a later version of their paper.

The only contentious issue is the population growth rate required for colonies to become established. My range of galactic colonization times of between five million and 50 million years corresponds to annual colony growth rates in the range of 0.2 to 2.5 percent. Newman and Sagan derived longer colonization times mainly because they assumed lower population growth rates, but the justification for this is not clear to me.

George W. Swenson, Jr.'s article "Intragalactically Speaking" leaves one with the sense that multipath interference is an insuperable constraint for SETI. Yet anyone who ever uses a cell phone inside a building knows that multipath can be an advantage. Nathan Cohen of Boston University addressed the multipath problem in SETI with David Charlton of Yale University and published it in 1993.

Called "polychromatic SETI," it is a special version of spread spectrum frequency hopping. Unlike conventional spread spectrum, polychromatic SETI is easily detected. It works by having up to six narrow frequency channels combed over a large frequency range, alternating in groups. This ensures that the multipath actually magnifies, rather than defeats, the signal intensity—at all times.

ROBERT G. HOHLFELD
Research Associate Professor,
Center for Computational Science
Boston University



NASA AND SPACE TELESCOPE SCIENCE INSTITUTE

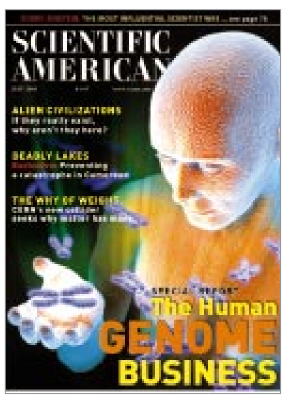
IF INTELLIGENT BEINGS once lived around extinct stars, where are they now?

DARWIN AND DIVINITY

I thoroughly enjoyed Ernst Mayr's article. But I would like to note that a "secular view of life" need not exclude divine action. There is no need to claim that God cannot employ randomness as well as necessity. Those who pooh-pooh Darwin because of their interpretation of the biblical accounts of creation and *On the Origin of Species* are generally unwilling to allow for other interpretations of the same texts. Often writers who wish to defend biological evolution against religious enthusiasts end up shooting down only paper tigers, as well they should. But having done so does not mean that the case against God is closed or that religious thought is invalid. Just as physical science seeks to find and formulate a "unified field theory," theological thought and methods can also strive for the unity of all truth even though there are always unanswered questions—just as there are always unanswered questions in the natural sciences. The dedicated scientist does not walk away from these questions, and the sincere theologian will acknowledge that it is not our business to tell God how to create.

RT. REV. KENNETH C. HEIN
Holy Cross Abbey
Cañon City, Colo.

THE_MAIL



SETI DOES NOT TRANSMIT SIGNALS; it only listens. Readers worried about our being detected (one contingent among the astounding number of respondents to Ian Crawford's article "Where Are They?") should look to the military and the broadcasting networks, not SETI. Besides, the earth is *visible* through any telescope.

Other readers cited historical precursors to Darwin's theory of natural selection ["Darwin's Influence on Modern Thought," by Ernst Mayr], including pre-Socratic philosopher Empedocles and Patrick Matthew, in his 1831 book *On Naval Timber and Arboriculture*. Evan Fales of the University of Iowa writes that "David Hume formulated the key ideas of variation and selection in his *Dialogues Concerning Natural Religion*, crediting Greek philosopher Epicurus with the germ of the hypothesis. It seems very likely that Darwin would have read Hume. That is not to detract from the power and originality of Darwin's insights concerning how strongly the biological evidence supports the hypothesis. But it would be interesting to know whether he got the basic explanatory strategy from Hume." For more on these and other July articles, please read on.

Editors' note:

In his essay, Mayr does point out that "one is certainly still free to believe in God even if one accepts evolution."

BRAVE NEW GENETICS

I find much of the media hype surrounding the Human Genome Project and its commercial applications rather disingenuous. The immediate prediction often reported from this tremendous effort is that the major drug companies will make custom variations of particular drugs based on individual genetic profiles. I wonder, though, if the companies will be loath to develop these optimal formulas once the market potential of the various genetic subgroups is clearly determined. How do companies plan to get multiple versions of a drug through the Food and Drug Administration given that agency's stringent double-blind study requirements? And what will that do to the overall cost of prescription drugs? With health care financing and drug costs for the elderly a major political issue this year, I can't help but think that the industry—and Wall Street—is a little giddy.

KEVIN COLEMAN
Tualatin, Ore.

Scientists need to use the wisdom of Darwin as they seek cures for human diseases. They often ignore the fundamental concepts of evolution, as when they design drugs targeting anxiety, even though under the Stone Age conditions in which humans evolved, built-in anxiety was a reasonable way to ensure that people would react quickly to the slightest rumble, such as the approach of a carnivore. These "bad genes" were used as genetic solutions by our ancestors to survive earlier dangers. Most genetic diseases are not mistakes but are in fact good adaptations, or else evolution would not have selected for them in the first place.

JAMAL I. BITTAR
Toledo, Ohio

Letters to the editors should be sent to editors@sciam.com or to Scientific American, 415 Madison Ave., New York, NY 10017.

ERRATUM

The DNA molecules illustrated on pages 54 and 60 [July] were inadvertently printed as "left-handed" helixes, when in fact they are "right-handed" molecules.

SCIENTIFIC AMERICAN

Sandra Ourusoff
PUBLISHER
saourusoff@sciam.com

NEW YORK ADVERTISING OFFICES
415 Madison Avenue
New York, NY 10017
212-451-8893 fax 212-754-1138

Denise Anderman
Associate Publisher
danderman@sciam.com

David Tirpack
Sales Development Manager
dtirpack@sciam.com

Wanda R. Knox
wknox@sciam.com

Darren Palmieri
dpalmieri@sciam.com

DETROIT
Edward A. Bartley
Midwest Manager
248-353-4411 fax 248-353-4360
ebartley@sciam.com

LOS ANGELES
Lisa K. Carden
West Coast Manager
310-234-2699 fax 310-234-2670
lcarden@sciam.com

SAN FRANCISCO
Debra Silver
San Francisco Manager
415-403-9030 fax 415-403-9033
dsilver@sciam.com

CHICAGO
ROCHA & ZOELLER MEDIA SALES
312-782-8855 fax 312-782-8857
mrrocha@aol.com
kzoeller1@aol.com

DALLAS
THE GRIFFITH GROUP
972-931-9001 fax 972-931-9074
lowcpm@onramp.net

CANADA
FENN COMPANY, INC.
905-833-6200 fax 905-833-2116
dfenn@canadads.com

U.K.
Anthony Turner
Simon Taylor
THE POWERS TURNER GROUP
100 Rochester Row
London SW1P 1JP, England
+44 207 592-8323 fax +44 207 592-8324
aturner@publicitas.com

FRANCE AND SWITZERLAND
Patricia Goupy
33, rue de l'abbé Grégoire
75006 Paris, France
+33-1-4548-7175
pgoupy@compuserve.com

GERMANY
Rupert Tonn
John Orchard
PUBLICITAS GERMANY GMBH
Oederweg 52-54
D-60318 Frankfurt am Main, Germany
+49 69 71 91 49 0 fax +49 69 71 91 49 30
rtonn@publicitas.com
jorchard@publicitas.com

MIDDLE EAST AND INDIA
PETER SMITH MEDIA & MARKETING
+44 140 484-1321 fax +44 140 484-1320

JAPAN
PACIFIC BUSINESS, INC.
+813-3661-6138 fax +813-3661-6139

KOREA
BISCOM, INC.
+822 739-7840 fax +822 732-3662

HONG KONG
HUTTON MEDIA LIMITED
+852 2528 9135 fax +852 2528 9281

OTHER EDITIONS OF SCIENTIFIC AMERICAN

SPEKTRUM DER WISSENSCHAFT
Verlagsgesellschaft mbH
Vangerowstrasse 20
69115 Heidelberg, GERMANY
tel: +49-6221-50460
redaktion@spektrum.com

POUR LA SCIENCE
Éditions Belin
8, rue Férou
75006 Paris, FRANCE
tel: +33-1-55-42-84-00

LE SCIENZE
Piazza della Repubblica, 8
20121 Milano, ITALY
tel: +39-2-29001753
redazione@lescienze.it

INVESTIGACION Y CIENCIA
Prensa Científica, S.A.
Muntaner, 339 pral. 1.ª
08021 Barcelona, SPAIN
tel: +34-93-4143344
precisa@abaforum.es

MAJALLAT AL-OLOOM
Kuwait Foundation for the Advancement of Sciences
P.O. Box 20856
Safat 13069, KUWAIT
tel: +965-2428186

SWIAT NAUKI
Proszynski i Ska S.A.
ul. Garazowa 7
02-651 Warszawa, POLAND
tel: +48-022-607-76-40
swiatnauki@proszynski.com.pl

NIKKEI SCIENCE, INC.
1-9-5 Otemachi, Chiyoda-ku
Tokyo 100-8066, JAPAN
tel: +813-5255-2821

SVIT NAUKY
Lviv State Medical University
69 Pekarska Street
290010, Lviv, UKRAINE
tel: +380-322-755856
zavadka@meduniv.lviv.ua

ΕΛΛΗΝΙΚΗ ΕΚΔΟΣΗ
SCIENTIFIC AMERICAN HELLAS SA
35–37 Sp. Mercouri Street
Gr 116 34 Athens, GREECE
tel: +301-72-94-354
sciam@otenet.gr

KE XUE
Institute of Scientific and Technical Information of China
P.O. Box 2104
Chongqing, Sichuan
PEOPLE'S REPUBLIC OF CHINA
tel: +86-236-3863170

Census 1900

More Americans, More Electricity

NOVEMBER 1950

GENETIC SURPRISE—"Thirty years ago the age-old question of how living things pass on their biological inheritance to their offspring was widely believed to have been solved. Heredity could be traced to invisible factors in the nucleus called genes. In this scheme of things the cytoplasm—the material of the cell outside the nucleus—was just a silent partner. Now claims have been made that the cytoplasm, like the nucleus, houses gene-like factors that take a hand in shaping an organism's heredity. Some biologists have gone so far as to contend that the cytoplasm controls all the basic traits of the organism and the nuclear genes determine only the relatively trivial ones. However, most professional students of heredity reject this extreme view."

FALLOUT SHELTER—"The U.S. has been presented with a 'master plan' for the nation's civil defense, prepared by the Civilian Mobilization Office of the National Security Resources Board and submitted to Congress by President Truman. The plan, however, places 'the primary responsibility for civil defense' on the states and local communities with the philosophy of 'organized self-protection.' In New York City the Sherry-Netherland Hotel has arranged to shelter its guests in its deep cellars, and a projected Madison Avenue skyscraper has included shelter for 4,000 in its plans." [*Editors' note: Not, so far as we are aware, in our building.*]

NOVEMBER 1900

1900 CENSUS—"The population by the Twelfth Census of the United States, of the 45 States and seven Territories, was officially announced by Director Merriam to be 76,295,220, compared with 63,069,756 in 1890; this is a gain of 13,225,464 in ten years, or an increase of 21 per cent. The three most populous cities are: Greater New York (including Brooklyn Borough), 3,437,202; Chicago, 1,698,575; and Philadelphia, 1,293,697."

THE ROOKWOOD POTTERY—"The awards for ceramics at the Paris Exposition have served to awaken fresh interest in a unique institution at Cincinnati, Ohio. The Rookwood pottery has produced not only as artistic ware as ever has been turned out on this side of the Atlantic, but also may be the most thoroughly representative of American ideas and methods in pottery work. Practically no machinery, save the primitive potter's wheel, is used at the Rookwood plant in the actual work of manufacture. From the mixing of the clay to the withdrawal of the completed piece of ware from the kiln, a Rookwood specimen passes through the hands of twenty-one operatives."

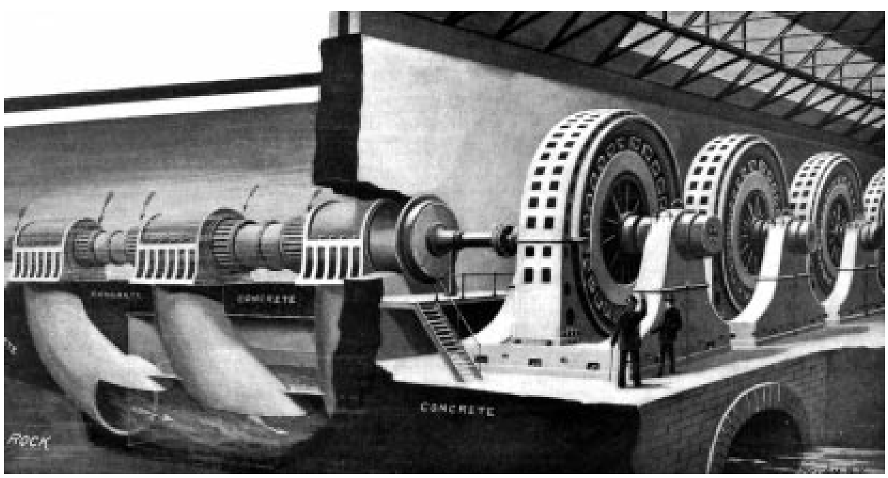
HYDROELECTRIC WONDER—"Nearing completion at Massena, N.Y., near the St. Lawrence River, is one of the latest and largest of the hydraulic electric power plants, which are one of the most significant features in engineering at the close of the nineteenth century. At the Long Sault Rapids the St. Lawrence River is about 42 feet higher than the Grasse River, a tributary stream. Advantage has been taken of this fact, and a canal has been cut across the intervening country. A power plant is now located on the banks of the Grasse River, which is utilized as a tail-race [outflow] for the discharged waters. The present capacity is 75,000 horse power [*see illustration*]."

NOVEMBER 1850

GOITRE AND CRETINISM—"Doctor Grange, a learned Physician of Paris, was commissioned some time ago by the government to pursue, in France and other countries, inquiries into the causes of *goitre* and *cretinism*. He has come absolutely to the conclusion that they are independent of latitude, altitude and climate, and even of circumstances of habitation, poverty, and so forth. Their presence appears to be connected with that of magnesia found in food or drink; their absence often proceeds from the *iodine* which the article contains."

SNAIL MAIL?—"We would not have noticed this story, only we have seen it copied into a number of papers: 'The marvels of the electric telegraph are annihilated, and the means of instantaneous communication between man and man, at any distance whatsoever, has been discovered! The inventors of the alleged marvel have ascertained that certain descriptions of snails possess peculiar properties or sympathies. With the snails placed in boxes, the operator has only to make snail A give a kick (*sic*) and snail B in a corresponding box, which may be in the backwoods of America or the deserts of Africa, repeats the kick. The snails of course must be put in sympathetic communication.' It is a piece of French nonsense."

SCIENTIFIC AMERICAN



TURBINES of the power plant at Massena, N.Y., near the St. Lawrence River, 1900

FROM *SCIENTIFIC AMERICAN*

AVIATION_REMOTE SENSING

Robots in the Sky

Miniature unmanned planes called aerosondes are ready to fly for science

AEROSONDE is transported by a pickup truck, which also serves as the launch vehicle.

Point Barrow, Alaska, may not go down in aviation history alongside Kitty Hawk, N.C., but this past August a group of would-be Wright brothers pushed the frontiers of flight in a new direction. From an airstrip on the shores of the Arctic Ocean, the researchers prepared their aircraft, an orange-crested drone not much bigger than a great blue heron in flight. The plane hurtled down the runway perched on the roof of a pickup truck before soaring skyward as the truck hit freeway speed.

What the 30-pound "aerosonde" lacks in takeoff elegance it promises to make up in its ability to carry a payload of miniature scientific instruments to places where pilots dare not fly. The craft, which can get more than 1,500 miles per gallon, may be able to fill in blind spots in global weather forecasts, monitor hurricanes, and help to decipher the ebb and flow of Arctic sea ice.

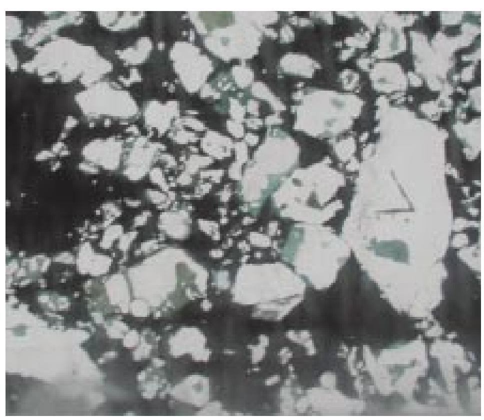
The little robotic plane, which entered the record books by completing the first unmanned flight across the North Atlantic in 1998, is ready to become a scientific workhorse. Funded by the National Science Foundation, the shakedown flights in Barrow, which took place at up to 13,000 feet and away from commercial flight lanes, were the first step in establishing a new aerosonde facility. That site, along with others planned in Australia and Taiwan, could form the nexus of a global aerosonde network over the next few years. "I wanted to gather data in severe weather events, like tropical cyclones," says Greg Holland, CEO of Aerosonde Ltd. in Victoria, Australia. To do that, you need to fly without a pilot in a plane that you can afford to lose, reasoned Holland, who proposed the aerosonde concept in 1992 with Tad McGeer of the Insitu Group in Bingen, Wash.

"We look at these things as disposable aircraft, which is a concept some people have a hard time with," says Juris Vagners of the University of Washington, who leads an engineering group working to give aerosondes the smarts to navigate semiautonomously. With onboard GPS navigation and computing power, aerosondes can follow a programmed flight path and can accept course changes from a ground controller if necessary. Future flights will rely on a satellite link so that the controller can even be a continent away from the plane, Holland remarks.

It was University of Colorado climate researcher Judith A. Curry's search for hard-to-get Arctic climate data that brought the aerosonde to Barrow. There, in April 1999, Curry put a flock of aerosondes outfitted with meteorological instruments to the test. Two that vanished over the Arctic Ocean were victims of ice buildup on the plane. Despite this setback, Curry won NSF support to return to Barrow and establish aerosondes in the Arctic as a routine research tool. The five-year project began this past summer: two weeks of flights to test new instruments, ice detectors and anti-icing coatings. All the aerosondes survived.

For the flights just completed in August, an off-the-shelf digital camera in the belly of the aerosonde collected images of sea ice that Curry's research group will use to study the dynamics of the annual freeze-up process, an overlooked component in global climate models. Surprisingly, the aerosonde did not see more gaps in the ice, as reported this year in other parts of the Arctic. There was actually much *more* sea ice around Barrow than in recent years, Curry says, which demonstrates how hard it is to pick out long-term climate trends from normal year-to-year variations.

But it's the weather data from distant expanses of the Arctic, where there are few reliable atmospheric soundings by balloons or satellites, that has European meteorologists excited. (Weather systems in this part of the world tend to move from west to east, so Arctic conditions affect European weather.) "That's the new ingredient you get from aerosondes: you can move them with the atmosphere," says Tim Palmer, who heads a research group at the European Center for Medium-



ARCTIC VIEW taken by an aerosonde indicates the potential of the unmanned craft to gather sea-ice data over remote areas.

MATTHEW R. ALLEN *University of Colorado (top)*; AEROSONDE "LINCOLN" *(bottom)*

Range Weather Forecasts, based in Reading, England, that is working to improve computer forecasting systems. If a network of aerosondes stood at the ready in key areas of the world, Palmer could run his daily forecasting models, pinpoint data-poor "atmospheric hot spots" that could significantly alter the forecast, and then direct an aerosonde to fly quickly to that area and collect data. Palmer believes that the generally poor European weather forecasts during the summer of 1999 could have been improved if more data from the Arctic had been available. He's testing this hypothesis by rerunning August's forecasts with aerosonde data from Alaska.

Routine science operations won't begin in Barrow until next summer, when Curry and her colleagues return with more miniature instruments, an upgraded aerosonde design and a new catapult device to replace the pickup-truck launch vehicle. She is confident that over the next decade aerosondes will become a standard research platform, especially in remote regions. *—Stephen Cole*

STEPHEN COLE is a science writer and editor based in Washington, D.C.

ENVIRONMENT_RADIATION DANGERS

Radioactive Wrecks

Sunken nuclear subs pose no immediate threat, but they could be long-term ecological time bombs

Many environmental groups became alarmed this past August after the sinking of the Russian submarine *Kursk* in the Barents Sea. Greenpeace International warned that the pristine Arctic waters could become contaminated by radioactive materials leaking from the submarine's two nuclear reactors. Because the vessel lies in relatively shallow water—only 108 meters below the surface—ocean currents could spread the deadly isotopes to the Barents's rich fishing grounds. Greenpeace officials urged world leaders to consider raising the submarine from the seafloor.

RUSSIAN SUBMARINE *KURSK*, shown here a few months before it sank in the Barents Sea, is one of seven nuclear submarines lying on the ocean floor.

AP PHOTO

Nuclear engineers who are familiar with submarine reactors agree that the danger of leakage exists, but in all likelihood the contamination will not occur for a long, long time. Although the explosion that doomed the *Kursk* ripped open the submarine's hull and may even have damaged the thick steel walls surrounding the reactors, the several hundred kilograms of uranium fuel in the reactors have an extra layer of protection. In U.S. submarine reactors, each rod of uranium fuel is encased in a zirconium alloy that is designed to withstand seawater corrosion for several hundred years. Nuclear experts say the fuel rods in Russian reactors have similar casings.

Unless the explosion cracked or smashed some of the *Kursk*'s fuel rods, the highly radioactive by-products of uranium fission will probably not leak out until well into the next millennium. By then, many of the most dangerous isotopes—such as strontium 90 and cesium 137, which have half-lives of about 30 years—will have decayed away. But several longer-lived isotopes could pose a threat when the fuel rod casings finally corrode in 1,000 years or so. Thomas Pigford, professor emeritus of nuclear engineering at the University of California at Berkeley, believes the most hazardous contaminant in the long run may be neptunium 237, which has a half-life of 2.1 million years. "It can get into the food chain if fish or shellfish ingest it," Pigford says, "and if it gets into your body, it can have some very bad effects."

The *Kursk*, however, is just the tip of the radioactive iceberg. Six other nuclear submarines lie on the ocean floor, including two U.S. vessels, the *Thresher* and the *Scorpion*. The U.S. Navy has collected sediments from the areas near its downed submarines and found slightly elevated levels of radioactivity, but the source of the contamination is believed to be the reactors' coolant rather than the fuel rods. Some scientists believe that even when the long-lived isotopes finally leak out, they will settle harmlessly into the mud at the sea bottom. But other researchers caution that neptunium 237 and other fission by-products could spread with the currents under certain conditions.

A more immediate issue is the disposition of dozens of decommissioned Russian submarines carrying spent nuclear fuel. They are rusting away in Russian ports because the government can't afford to dispose of the radioactive waste properly. "That's a more important thing to worry about than the *Kursk*," says Thomas B. Cochran of the Natural Resources Defense Council, an environmental group based in Washington, D.C. "Those reactors are sitting just offshore, in a few feet of water." *—Mark Alpert*

Atlas Shrugged

When it comes to online road maps, why you can't (always) get there from here

The proliferation of online services such as MapQuest and MapBlast would appear to be a godsend to the American man's dream of not having to stop to ask for directions. Today navigational Web sites serve up more than 12 million maps and three million driving directions daily, and monthly traffic is growing at double-digit rates. Wireless personal digital assistants, cell phones and pagers are increasing their geographic reach even further.

But to anyone who downloads directions regularly, there's one slight problem with this geographically enabled infotopia: the directions don't always take you where you want to go. Common complaints include nonlocatable addresses, directions that stop short of their intended destination and the occasional geodisaster that leaves you circling the back streets of terra incognita.

Nick Hopkins, director of software engineering at MapQuest, estimates that as many as one out of every 20 MapQuest driving directions is wrong, a statistic that some experienced users say understates the problem. Unfortunately, MapQuest, which supplies service to many sites, including America Online, its parent company, and Yahoo, is not alone. "There's still a big difference between the state of the art and what people would like to see," admits Scott Young, a senior vice president at Vicinity, which operates MapBlast and provides the navigation engine for Rand McNally's Web site.

To generate directions, the online services first transmit the starting and destination addresses entered into the browser window to an application server that locates these points on a road-network database. This process is called geocoding.

Rather than storing individual street addresses, road databases are organized into road segments—one side of a single block, for instance. Each segment is represented by a string of 256 characters that contains its name and address information, latitude and longitude, and other important attributes such as road class, speed, turn and access restrictions, and links to other connecting segments. A typical U.S. road database contains eight million to 10 million road segments and tens of thousands of "points of interest"—airports, museums, businesses and so forth. One road database occupies several gigabytes of memory.

GEOCODING computes an address's geographic coordinates from a road database network supplied by third-party vendors such as Etak in Menlo Park, Calif. Rather than store every single address in a database, the software can extrapolate. For instance, it can assume that "1736 Eisenhower Street" lies somewhere on the road segment defined by, say, between 1700 and 1800 Eisenhower Street.

Once the addresses have been matched to road segments (149 Main Street would be located midway on the 101–199 Main Street segment), the software calculates an "optimal route" between the segments. Most optimization methods are based on an obscure but powerful piece of graph theory called the Dijkstra algorithm, invented in 1955 by Edsger Wybe Dijkstra, now a computer scientist at the University of Texas at Austin. The algorithm calculates the distance of possible paths between the source and destination node and then selects the shortest one. Imagine an army of rats simultaneously spreading out through a maze in search of the cheese while keeping track of the distance they traverse along the way.

In the case of a road network, each segment is given weights to represent distance, speed limits and other data. Computational speed is also critical, because the software must crunch through hundreds of thousands of road segments for each request while handling dozens of requests each second. As a result, programmers have had to develop shortcuts—for example, choosing paths that favor highways over local streets—to reduce the time it takes to calculate a route to less than 100 milliseconds. Finally, the software translates the resulting set of connecting road segments into a narrative that the user can understand, like "Merge onto Bruckner Expressway in 2.7 miles."

More than half of the bad directions stem from user error, MapQuest's Hopkins says—misspelling a street, for example, or leaving out critical designations such as north or south. Incorrect geocoding accounts for most of the other directional faux pas. Road databases are generally updated four to six times a year and are usually out of date by the time they are published. Although the physical road network changes slowly, attributes such as turn and access restrictions, posted speeds, and street names are effectively in a state of flux, considering the millions of road segments out there. The databases also have inaccuracies, which some estimate to run as high as 30 percent. Sometimes the geocoding process mistakes the destination for one with a similar street name. And there are the usual software bugs.

ETAK, INC.

Fortunately, most navigation services have processes in place for identifying and correcting systematic errors. Each day MapQuest receives several thousand e-mails, which are sorted by an automated system and routed to the quality assurance department. There they are reviewed and forwarded to the software group or the database vendor for correction.

Although the Internet has sped up the error-correction process, Vicinity's Young says that the service's overall 95 percent reliability will not significantly improve until corrections can be made automatically in real time, say, by monitoring the movement of vehicles enabled by GPS (Global Positioning System). These "million ant" correction schemes are still a ways off, he says, but the use of wireless Internet devices with GPS will elevate the service: "If we know where you are and where you're going, we can supply a lot of dynamic information." He predicts that in time, online directions will be as detailed and seamless as those dispensed by the staff at a fine hotel. Of course, even the concierge at the Four Seasons occasionally gets them wrong.

—Michael Menduno

MICHAEL MENDUNO writes about the digital economy from Menlo Park, Calif.

BIOINFORMATICS_STANDARDS

Hooking up Biologists

Consortia are forming to sort out a common cyberlanguage for life science

Imagine that your co-worker in the next cubicle has some information you need for a report that's due soon. She e-mails it to you, but the data are from a spreadsheet program, and all you have is a word processor, so there's no possibility of your cutting and pasting it into your document. Instead you have to print it out and type it in all over again.

That's roughly the situation facing biologists these days. Although databases of biological information abound—especially in this post-genome-sequencing era—many researchers are like sailors thirsting to death surrounded by an ocean: what they need is all around them, but it's not in a form they can readily use.

To solve the problem, various groups made up of academic scientists and researchers from biotechnology and pharmaceutical companies are coming together to try to devise computer standards for bioinformatics so that biologists can more easily share data and make the most of the glut of information resulting from the Human Genome Project. Their goal is to enable an investigator not only to float seamlessly between the enormous databases of DNA sequences and those of the three-dimensional protein structures encoded by that DNA. They also want a scientist to be able to search the databases more efficiently so that, to use an automobile metaphor, if someone typed in "Camaro," the results would include other cars as well because the system would be smart enough to know that a Camaro is another kind of car.

The immediate payoff is expected to be the faster development of new drugs. "Pharmaceutical research is the only industry I know of with declining productivity," says Tim Clark, vice president of informatics for Millennium Pharmaceuticals in Cambridge, Mass. "The R&D effort is at a primitive craft scale, like cottage weavers, although standardization is one of the first problems that got tackled in the Industrial Revolution, with the invention of interchangeable parts."

The issue is what standards to use. In a situation reminiscent of the computer industry in the 1970s, everyone advocates standards, as long as they are his or her own. Formal groups have sprung up worldwide with names like the BioPathways Consortium, the Life Sciences Research Domain Task Force of the Object Management Group, and the Bio-Ontologies Consortium—and each has a different idea of how things should be done.

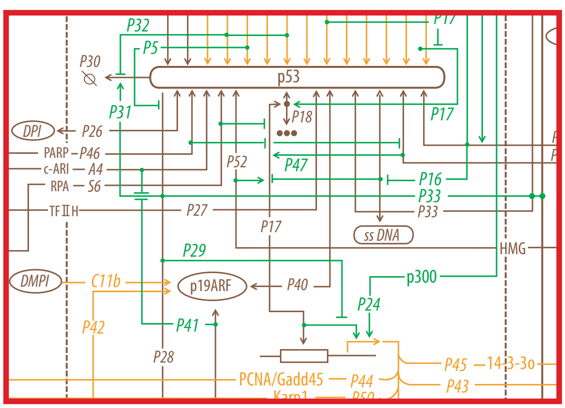
Eric Neumann, a member of both the Bio-Ontologies and BioPathways consortia, is a neuroscientist who is now vice president for life science informatics at the consulting firm 3rd Millennium in Cambridge, Mass. (no relation to Millennium Pharmaceuticals). He says Extensible Markup Language (XML) is shaping up to be the standard computer language for bioinformatics. XML is the successor to Hypertext Markup Language (HTML), the current driver of the World Wide Web [see "XML and the Second-Generation Web," by Jon Bosak and Tim Bray; Scientific American, May 1999].

One of XML's advantages is that it contains tags that identify each kind of information according to its type: "Camaro," for example, would be tagged as a car. Neumann proposes that XML-based languages will "emphasize the Web-like nature of biological information," which stretches from DNA to messenger RNA, proteins, protein-protein interactions, biochemical pathways, cellular function and ultimately the behavior of a whole organism. Current ways of storing and searching such biological information are centered on single genes, according to Neumann, "but the diseases we want to treat involve more than one gene."

Clark says the main problems facing bioinformatics that make standard development necessary are the sheer volume of data, the need for advanced pattern recognition (such as within DNA sequences and protein structural domains), the ability to process signals to eliminate "noise" from data, and something called combinatorial optimization, or finding the best path through a maze of molecular interactions. "You can't build all of it yourself," he contends.

Neumann thinks combinatorial optimization could be the highest hurdle. "Pathways are a lot more complex than [DNA] sequences," he states. "If we don't come up with something, it's going to be a real mess."

—Carol Ezzell



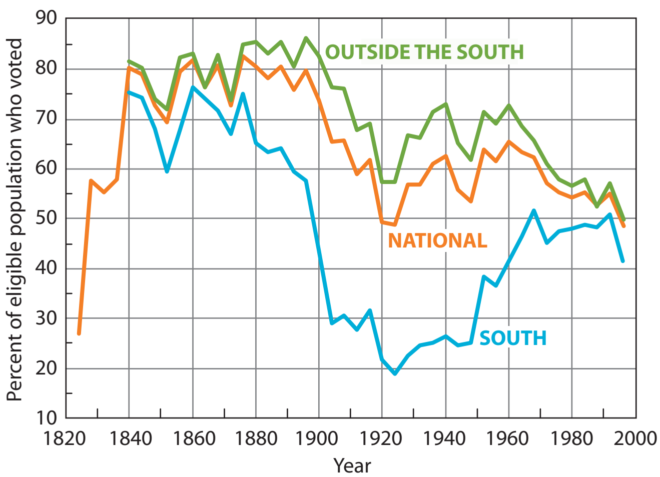
WEB OF INTERACTIONS involving cancer-related gene *p53* is just one that bioinformatics must tackle.

ERIC K. NEUMANN

Voter Turnout

Egalitarian democracy made a spectacular American debut in 1828, when Andrew Jackson won the White House by mobilizing workers, small farmers and frontiersmen in unprecedented numbers. It was the start of a golden age of grass-roots democracy, a time when people—white men, at least—were passionately involved in political discussion. So vigorous was the democratic impulse that it survived the Civil War and Reconstruction, ending only after the presidential election of 1896.

The reasons for the decline in voter turnout since then lie somewhere in the interaction of ordinary people with major economic power groups. The needs and aspirations of farmers and workers, particularly in times of economic or social crisis, came up against the imperatives of the power groups, such as Northern industrialists and the Southern planter merchant class. When these groups were in alliance, popular movements had less opportunity to gain political momentum, but when these groups fell apart, popular movements had a better chance to gain influence.



SOURCE: Adapted with modification from Bill Winders, "The Roller Coaster of Class Conflict: Class Segments, Mass Mobilization, and Voter Turnout in the U.S., 1840–1996"; Social Forces, *March 1999. The "South" includes the 11 states of the Confederacy. The eligible electorate excludes black men before 1870, most women before 1920, 18- to 20-year-olds before 1972, and most aliens throughout.*

Bill Winders of Emory University, who has authored the most systematic analysis of the topic in recent years, identifies four main periods during the past century or so when the major power blocks played a key role in determining turnout. In the first phase, running from 1896 to 1924, Northern industrialists, threatened by populist farmers, striking workers and unruly immigrants, came to an agreement with their erstwhile opponents of the Civil War, the Southern planter merchant class: the industrialists would get a free hand in dealing with unrest in the North, and the Southern planter merchant class would be allowed to reimpose control over the former slaves, using intimidation, the poll tax and literacy tests. Turnout declined heavily through this period in both the North and South.

In the second period—1928 to 1940—militant union activity, protest demonstrations and an unemployment rate that reached 25 percent by 1933 created an unstable situation. Recognizing this, the more realistic industrialists and members of the planter merchant class split with their more conservative colleagues and supported, or at least did not hinder, the efforts of unions and other popular organizations to get out the vote. Each member of the alliance got something: workers got better union and social policies; liberal industrialists got greater regulation of the economy; and Southern landowners got agricultural assistance. Thus was the New Deal coalition born.

In the third phase—1948 to 1968—a similar mobilization of the disenfranchised took place, this time mostly in the South. Northern industrialists broke with the Southern landed class and helped to channel black protest into politics, as for example in the Freedom Summer of 1964, which decisively increased black voter registration throughout much of the South. Furthermore, the courts declared restrictions such as the poll tax and literacy tests illegal, and the federal government passed the Voting Rights Act of 1965 to increase voter registration among blacks. Southern landowners, in an attempt to offset the black vote, made efforts to increase the turnout of lower and working-class whites, thus further boosting voter turnout.

As for the decline in turnout since 1968, Winders notes that the long-standing split in the elite group that led to the popular gains of the 1930s and 1960s no longer exists. The Southern planters were replaced by agribusiness, which has common interests with Northern industry, such as promoting free trade. The decline was reinforced by the social unrest of the late 1960s, which divided the supporters of civil rights.

Furthermore, the economic crisis of the 1970s strengthened the hand of the more conservative industrialists of the North, who allied themselves with industrial segments in the South that had long been antiunion. And then, beginning in the 1970s, a growing stream of special-interest money made politicians far less dependent on mass organizations to get out the vote. Party leaders were reluctant to recruit massive numbers of new members, fearing loss of control over party organization, Winders concludes. As an example, he cites the 1996 primaries, in which Republicans became alarmed about Pat Buchanan's attempt to draw in working-class voters, sensing that they were incompatible with the businessmen who are the backbone of the party.

The U.S. had a 47.2 percent turnout in the 1996 presidential election, well below the 71 percent European average of recent years. The 36.7 percent of those eligible who voted in the 1998 midterm congressional elections was the third-lowest turnout in at least the past 50 years. An important cause involves restrictions on registration and voting, which are greater in the U.S. than in other Western countries. In most jurisdictions, for instance, Americans must reregister when they move, a problem for one sixth of Americans every year; European countries tend to ensure permanent registration. At the most fundamental level, one can argue that U.S. voters don't go to the polls often because the political parties and their allied economic interests have little incentive to promote citizen involvement, while at the same time there is no social or economic crisis strong enough to generate a sense of urgency in the electorate.

—Rodger Doyle (rdoyle2@aol.com)

RODGER DOYLE

TECHNOLOGY_REFRIGERATION

Desert Fridge

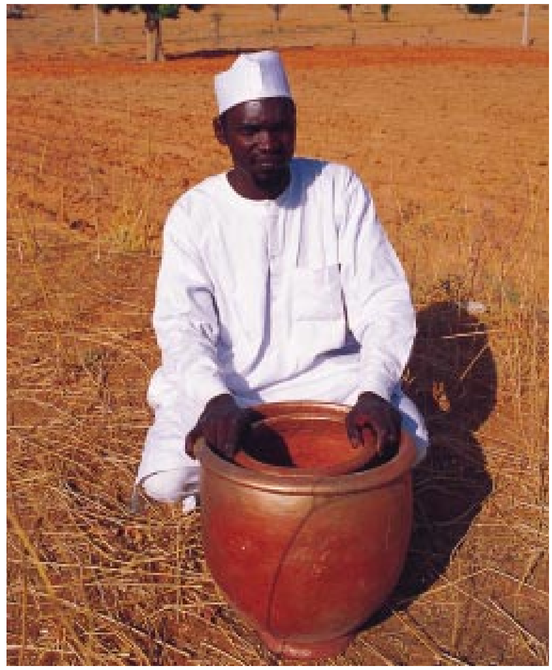
Cooling foods when there's not a socket around

Thanks to the second law of thermodynamics, Mohammed Bah Abba has developed a refrigerator that doesn't need electricity. What's more, it costs 30 cents to make.

The elegant design consists of an earthenware pot nestled inside a larger pot, packed with a layer of damp sand. When the "Pot-in-Pot" system is stored in a very dry, well-ventilated place, the water held in the pots' clay walls and sand evaporates, carrying heat with it. The inner pot therefore cools down—and makes a useful refrigerator in the northern deserts of Nigeria, where Abba lives and works. Abba says his trials showed that tomatoes would last several weeks instead of several days and that African spinach (amaranth), which normally wilts within hours of harvest, can last up to 12 days. (He's never measured, though, just how many degrees cooler the inner pot becomes.)

PHOTOGRAPHS BY TOMAS BERTELSEN

Abba's fridge provides an alternative for desert cultures, which generally dry their foods to preserve them. Drying doesn't diminish protein or calorie content much, notes William R. Leonard, a biological anthropologist at Northwestern University who has worked in the high desert of the Peruvian altiplano. "But things like vitamin C are likely to be in shorter supply" in the dried foods, Leonard says. In addition, some foods, such as spinach and onions, cannot be dried, remarks Abba, a lecturer at Jigawa State Polytechnic in Dutse, Nigeria. The Pot-in-Pot may have great social impact, too: Abba says that young girls who used it would not have to sell their families' freshly picked foods right away and thus would have time to go to school.



POT-IN-POT system developed by Mohammed Bah Abba (*above*) consists of nested clay pots cooled by evaporation from an intervening layer of wet sand (*left*).

For his work, Abba received one of five biennial Rolex Awards for Enterprise on September 27. The others were Elizabeth Nicholls, a Canadian paleontologist who unearthed an ichthyosaur in British Columbia; Maria Eliza Manteca Oñate, an Ecuadorian environmentalist promoting sustainable farming in the Andes; Laurent Pordié, a French ethnopharmacologist who is preserving traditional Tibetan healing methods in northern India; and David Schweidenback, an American recovering used bicycles in the U.S. for shipment to developing countries (see www.rolexawards.com). —*Naomi Lubick*

ROBOTICS_REPLICATION

Dawn of a New Species?

When a future robotic race writes its Book of Genesis, it will surely give a place of honor to Hod Lipson and Jordan B. Pollack. In the August 31 *Nature*, these Brandeis University researchers report that they have designed and built the first robot that can design and build other robots. (In earlier efforts, replicating machines had been simulated only on computers and on special integrated circuits.) The offspring are plastic trusses (like Tinker Toys) propelled by pistons and controlled by simple neural networks. The mother bot is a computer running a genetic algorithm, which draws up plans through trial and error, and a 3-D printer, which can create small plastic sculptures of any shape. The researchers could (almost) leave the system to work at night and come in the next morning to see artificial inchworms crawling around their lab. They still had to strap on motors and connect wires, but—in a reversal of roles—the robot told the humans what to do. The software is available for Windows-based computers (www.demo.cs.brandeis.edu/golem). So will humans soon share the world with cyborgs? If that sounds silly, consider that the researchers felt compelled to say in their paper that "robotic lifeforms" are not dangerous, yet. —*George Musser*

COURTESY OF HOD LIPSON AND JORDAN B. POLLACK

COMPUTER-DESIGNED ROBOT pushes itself along the carpet using the piston at its center (for a video, see http://golem03.cs-i.brandeis.edu/results.html).

MEDICINE

Universal Soldier

The war against cancer gets a shot in the arm with the promising preliminary results of a universal cancer vaccine. Most potential vaccines are associated with molecules from specific tumors. But researchers at Duke University and Geron Corporation in Menlo Park, Calif., report in the September *Nature Medicine* on an experimental vaccine that depends on a part of the enzyme telomerase that, they note, "is silent in normal tissues but is reactivated in more than 85 percent of cancers." The telomerase vaccine slowed the growth of melanoma, breast and bladder cancers in mice and provoked an immune response in cells derived from human renal and prostate cancers. The search is on for other molecules that, combined with telomerase, would broaden and strengthen the effect. *—Steve Mirsky*

PSYCHOLOGY

Stifled Recall

Emotions may affect how one remembers an event, but so can keeping those emotions in check. A study in the September *Journal of Personality and Social Psychology* found that emotional regulation can take away from finite cognitive resources available to pay attention to an event. Subjects asked to stifle their physical responses to emotions while looking at slides of injured people could not recollect specific details of the images a short time later as well as a control group could. Not all types of emotional regulation affected memory to the same degree, however: subjects asked to view the slides with detachment, as a physician would, effectively precontrolled their emotions and hence had better recall than those who did not have a chance to prepare beforehand (the full text is at www.apa.org/journals/psp/psp793410.html). *—Naomi Lubick*

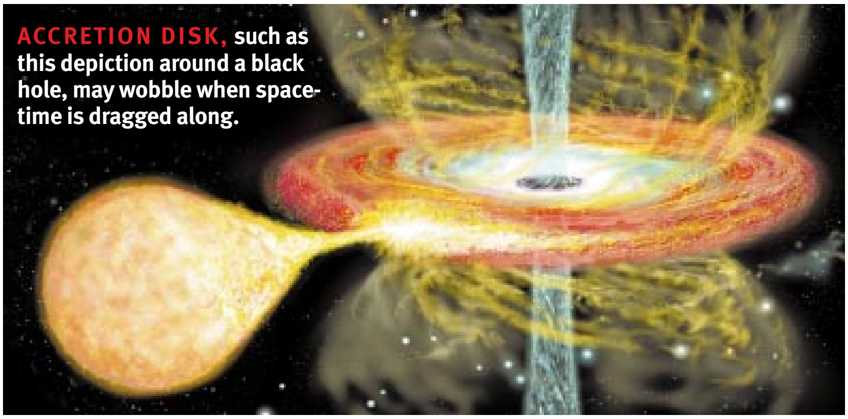
TIM WRIGHT *Corbis*

KEEPING A STIFF UPPER LIP colors memories of emotionally laden images, such as those of an accident.

News Briefs

ASTRONOMY

Broadcasting Space Warp



ACCRETION DISK, such as this depiction around a black hole, may wobble when spacetime is dragged along.

RUSSELL KIGHTLEY MEDIA

Three years ago astronomers reported detecting black holes and neutron stars that were not only sucking in matter but also twisting the very fabric of spacetime around it. The twisting makes the matter precess, or wobble, as it spirals into the dead star. Researchers from the University of Amsterdam report in the August 24 *Astrophysical Journal Letters* that they have detected three neutron stars emitting so-called sideband radiation in the x-rays emitted when material gets drawn in. Such sideband emissions are like the stations carried by AM radio waves. But instead of delivering news, sports and weather, they convey information about the stars and can be used to confirm Einstein's predictions about the dragging of spacetime. For an animation of the precessing matter, see www.physics.uiuc.edu/Research/CTA/news/sidebands/index.html —*Philip Yam*

SLEEP DISORDERS

Narcolepsy and the Lost Peptide

Narcolepsy most likely results from a lack of the neurotransmitter hypocretin, according to two groups, one led by Jerome M. Siegel of the University of California at Los Angeles and the other by Emmanuel Mignot of Stanford University. The studies, appearing in the September issues of *Neuron* and *Nature Medicine,* found that compared with normal human brains, narcoleptic brains had lost practically all their hypocretin peptides, known to regulate both appetite and sleep. Less certain is the underlying cause of the loss of hypocretins; an autoimmune response is a possibility. Therapies repairing the hypocretin system could be a better alternative to the current treatment of stimulants and other drugs. —*N.L.*

DATA POINTS

Medium Rare, Please

MATT COLLINS

Total number of cattle in the U.S., 1999: 99,115,000
Amount of beef produced per cow: 616 pounds

Per capita beef consumption, in pounds, 1999: 64.7
Chicken: 49.2
Pork: 48.8

Total U.S. beef production in 1990: 21.8 billion pounds
Beef production in 1999: 26.4 billion pounds
Retail value: $54.4 billion
Tofu production: 116 million pounds
Retail value: $173 million

Amount of cholesterol in a hamburger: 35 to 50 milligrams
Average daily cholesterol consumption by U.S. males: 337 milligrams
Recommended daily cholesterol intake by the American Heart Association: less than 300 milligrams

SOURCES: U.S. Department of Agriculture; National Cattlemen's Beef Association; Soyfoods Association of North America; Soyatech; American Heart Association. Tofu figures are from 1997, the latest available.

SCIENCE ART

Digital Depictions

A tick and a unicycle race won the 2000 Science & Technology Digital Art Competition, one of the few art contests encouraging submissions through cyberspace. The digital nature of the contest made international accessibility easier—some entries came from England, France and Portugal. The competition, cosponsored by *Scientific American,* was presented by the Arts Alliance Center at Clear Lake in Nassau Bay, Tex. "There are not many venues for artists who specialize in the sciences, and that makes this an important competition," says noted space artist Pat Rawlings, one of the judges. This year's winners, chosen in August from more than 200 entries, can be viewed at www.taaccl.org until the end of November.

—Edward Bell

1st Prize, Science Division

1st Prize, Technology Division

DENA WINKELMAN AND WILLAM MATTHEWS *(left);* RICHARD GREEN *(right)*

GENOMICS

One Transgenic Latte, Coming Up

If Gregor Mendel pondered his pea plant data over coffee, he might have been pleased to learn that the genetic science he founded would one day be on the verge of the perfect decaffeinated brew. Researchers from Japan and Scotland report in the August 31 *Nature* that they have characterized and cloned the gene for caffeine synthase, a key enzyme in the biosynthesis of caffeine. With the gene now known, scientists can set about creating transgenic coffee and tea plants that cannot produce the stimulant. Many consumers eschew decaffeinated beverages because flavor and aromas may be lost in the decaffeination processes. Transgenic decaf, however, would theoretically be otherwise identical to natural brews. *—S.M.*

Decaf coming to coffee beans?

JEREMY HORNER *Corbis*

ARCHAEOLOGIST_OLGA SOFFER

The Caveman's New Clothes

From what they wore to how they hunted: overturning the threadbare reconstructions of Ice Age culture

PARIS—Walking through the human evolution exhibit in the Musée de l'Homme, two things stand out to Olga Soffer: males are depicted to the exclusion of females, and they're wearing the wrong clothes. Only someone who has never sewn before would conclude that this needle could have pierced through hides, she declares, drawing my attention to a delicate sliver of bone in one of the display cases. Rather, the University of Illinois archaeologist asserts, it must have stitched a far finer material—perhaps even something akin to the linen of her blue pinstriped suit.

Such needles—some of which date back more than 25,000 years to the Upper Paleolithic period—vaguely suggest that caveman couture extended beyond the crude animal-skin ensembles envisioned by many of her colleagues. Soffer's efforts are revealing just how sophisticated those first fashionistas were. By scouring the archaeological record for evidence of perishable technologies like weaving, she has uncovered clues to formerly invisible activities of Ice Age men—and women—forcing a reevaluation of the men-in-furs-hunting-megafauna motif that has long dominated reconstructions of prehistoric lifeways. The fabric of their lives, it appears, was much richer than previously thought.

Soffer's passion for fashion predates her interest in the Paleolithic. After graduating from Hunter College with a degree in political science, she entered an executive training program with New York City's Federated Department Stores—owners today of Macy's, Bloomingdale's and others. This led to a 10-year career in fashion promotion, which, she says, suited her just fine early on but grew tiresome as she reached her late 20s. "I started playing hooky," she recalls, chuckling. "I'd go to fashion shows and actually sneak off to the library."

To feed her mind, Soffer decided to take some night school courses in art. In a couple of years she worked her way from Picasso to prehistoric art to prehistoric ways of life and concluded that she "might as well" get her master's in archaeology. Then, after taking a summer off from her job to go to France "to learn digging," Soffer decided to pursue a Ph.D. through the City University of New York while continuing to work halftime in fashion during the first two years.

In 1977 she left for Russia, accompanied by her then husband and her six-year-old daughter, to conduct her dissertation research. Russia was an open niche, she recollects, unlike France, where, she says, there was "a ratio of about two archies for one square meter of territory." Additionally, her Russian parentage meant that she had the advantage of language and cultural sensitivity. There on the central Russian Plain, home of the famous Upper Paleo-

RALF-FINN HESTOFT *SABA*

OLGA SOFFER: FASHION MAVEN TURNED ARCHAEOLOGIST

- ***Born September 9, 1942, in Belgrade, Yugoslavia***
- ***Speaks Russian, Serbo-Croatian, Italian and Czech fluently; converses in a number of others***
- ***Typical airplane reading:*** **Vogue** ***and*** **W**
- ***Favorite designers: Yohji Yamamoto, Geoffrey Beene and Coco Chanel, to name a few***
- ***Received an honorary doctorate from the Russian Academy of Sciences***

lithic mammoth-bone dwellings, she developed her interest in prehistoric subsistence practices. There, too, she began to wonder whether conventional wisdom on the matter was flawed.

"We bring an awful lot of baggage to prehistory," Soffer rues. Take, for example, that perennially popular Ice Age scene, the mammoth hunt. She doesn't buy it. No known living or recent hunter-gatherer groups have ever survived on elephants, she observes. Like elephants, mammoths were dangerous animals, and the close encounters required by hand-held spear hunting would have posed far too many risks. What then of those mammoth-bone assemblages in Russia and elsewhere in eastern and central Europe? The same sites have also yielded the remains of numerous small animals, such as rabbits and marmots. "If they've got all this mammoth meat, why in heaven's name are they hunting *bunnies*?" she demands. A more plausible explanation for most of the mammoth bones is that people collected them off the landscape from animals that died of other causes. She concludes that mammoth and other megafauna hunts were occasional and did not play a central dietary role.

As for bringing down those small animals, Soffer suspects it wasn't with spears. She and James M. Adovasio of Mercyhurst College have identified impressions of netting on fragments of clay from Upper Paleolithic sites in Moravia and Russia that open up an intriguing possibility: net hunting. Ethnographic descriptions of this strategy, Soffer explains, reveal that "you don't need to be a strong, brawny, skilled hunter. You can participate and help with this kind of communal hunt if you're a kid with no experience, if you're a nursing mother, et cetera. It's nonconfrontational" and relatively safe.

Impressions of netting and other perishable materials provide some of the first insight into the lives of prehistoric women, children and the elderly—or, as Soffer describes them, the silent majority. Whereas the activities of prime-age males in hunter-gatherer cultures tend to entail the manipulation of durable materials, those of women, children and the elderly involve more perishables. As a result, the archaeological record has preferentially preserved behavioral remains associated with young men.

Soffer's efforts, however, have demonstrated that it is quite possible to recover evidence of what these other people did. Over the past few years she and her colleagues have identified all sorts of plant fiber artifacts—impressions of cordage, textiles, basketry—from Upper Paleolithic sites across Europe. And research conducted just last year indicates that certain bone and antler objects once thought to be hunting tools actually represent tools used to manufacture these perishable items: net gauges and battens for weaving, for instance.

Although remains of perishables are known from 13,000-year-old Paleoindian sites [see "Who Were the First Americans?" by Sasha Nemecek; Scientific American, September], these Upper Paleolithic materials push back the date for the oldest plant-based technologies by thousands of years. But they're still too advanced to represent the origins of such practices. Indeed, the most basic of these technologies—cordage—probably dates back at least 60,000 years to the first colonizers of Australia, whom many researchers suspect sailed over from Southeast Asia. Considering the limited availability of animal sinew in that region, Soffer says, their rafts would most likely have been lashed together with ropes made of plant fibers.

Most of Soffer's startling observations have been made on archaeological materials that were discovered long ago. Yet until now, no one had noticed them. That's because they weren't looking for it, she asserts. "If you're looking with these questions in mind, stuff that had always been there starts jumping out at you—like the fact that the Venus figurines are dressed. They're *wearing* clothes, for God's sake." Although these voluptuous female statuettes from Upper Paleolithic sites across Europe have been known for decades, most scholars overlooked their apparel. How? "Because an awful lot of people who were studying this stuff were men who looked at the variables that were far more emotionally charged: secondary sex characteristics," Soffer remarks matter-of-factly. "When we started looking at these things as archaeologists, looking at the range of variables and the patterning of those variables—aside from boobs and asses—lo and behold, there's this other stuff."

The other stuff, it appears, includes a stunning array of ritual garb: the famed Venus of Willendorf from Austria wears a woven hat (previously interpreted as a coiled coiffure); the French Venus of Lespugue sports a string skirt; other Venuses model bandeaus, snoods, sashes and belts. Close study of the carvings reveal that all the representations of apparel clearly depict fiber-based items, as opposed to hidewear, further strengthening the case for early textiles.

These Paleolithic representations of women stand in stark contrast to the few known representations of men, none of which show clothing. Whether these mysterious figurines represent sex symbols, fertility goddesses or some other entities, we may never know. Yet to Soffer, the fact that ancient artists took such pains to immortalize their apparel clearly illustrates the importance of these perishable technologies. And if the ethnographic record on perishables is any indication, the manufacturers were probably women. "Women were not just out there to reproduce," Soffer insists. "They were actively involved in production as well, just as women are in any and all societies that we know of."

—*Kate Wong*

VENUS WEAR: Forget those tattered animal-hide getups—Ice Age women had textiles, as seen on Venus figurines that date back as far as 27,000 years ago. A cast of the Venus of Willendorf from Austria (*right*) shows a woven cap; the one from Kostenki in Russia (*left*) displays similar headwear and a twined bandeau.

RALF-FINN HESTOFT *SABA*

As We May Live

Computer scientists build a dream house to test their vision of our future

ATLANTA—To pedestrians walking past in the muggy summer heat, the green house at the corner of 10th and Center streets looks very much like any of the other two-story homes in this quiet neighborhood a block north of the **Georgia Institute of Technology.** Only the loud whir of two commercial-size heat pumps in the side yard hints at the fact that the house is infested with network cables threaded through the floorboards, video cameras staring from the ceiling, sensors tucked into kitchen cabinets, workstations stacked in the basement, and computer scientists bustling from room to room.

Inside the house, some passing student has arranged toy magnetic letters on the refrigerator door to spell out the purpose of this odd combination: "Aware Home of the Futur," a laboratory in the shape of a house where humans can try out living in more intimate contact with computers. There's a piece missing from the message, but the project itself has many gaps to fill. Construction wrapped up only a few months ago, and seven faculty members from Georgia Tech's computer science department are still working with a battalion of students to get the house's sensory systems online.

This house does all the light-switching, stereo-piping tricks of "smart" homes that provide technophiles with electronic convenience, but here that is just a starting point. The goal is to make this place the most ambitious incarnation yet of ideas that have been fermenting in computer research labs for a decade, ever since Mark Weiser launched the first "ubiquitous computing" project at the **Xerox Palo Alto Research Center** (PARC) in the late 1980s. In a seminal 1991 article in *Scientific American,* Weiser predicted that human use of computers would in the early 21st century go through a transition comparable to the shift from shared mainframe machines to personally owned workstations, laptops and handhelds. The third generation of "UCs," he argued, should look like everyday objects—name tags, books, jewelry, appliances, walls—but should be highly interconnected and able to adapt their behavior to different users, locations and situations. In this vision, we will share many kinds of UCs, and the devices will share us.

IT'S AWARE: **a new computer science lab will monitor its live-in test subjects.**

GARY W. MEEK

A decade's work on UbiComp, as it is known in the field, has produced a zoo of ideas and many demos but few real-world tests. **NCR** unveiled a microwave oven that could support e-mail and electronic banking in 1998 and last year demonstrated a trash bin that can use a bar-code scanner on its lid to track the contents of the pantry. Neither has made it beyond prototypes. On a quick stop at the **IBM Almaden Research Center,** Cameron Miner shows me a glass case full of digital jewelry: a tie-bar microphone, earring earphones, a ring with a multicolored LED. "It might flash when you get an incoming call," Miner suggests. But these are mock-ups; they do not actually connect to anything.

No one knows yet what kind of infrastructure is needed to support a UbiComp world, so the designers of 479 10th Street took no chances. Every wall has at least six high-speed jacks to the internal Ethernet network. Cordless devices communicate through a house-wide wireless net. A radio-locating system can pinpoint any tagged object to within 10 feet. The two-gigabit-per-second connection to the university and the Internet is fast enough to transmit several channels of full-screen video and audio. And with some 25 cameras and almost as many microphones trained on the first floor alone, there is plenty of audio and video to go around.

Aaron Bobick, who specializes in computer vision, gives me the grand tour. "Everybody in our department thought building this must be a good thing to do," he says, "although we didn't really have a clear vision of why." The research team eventually decided that those who most need the home of the future are people of the past—not the rich gadget nuts who typically purchase smart homes but rather marginally infirm seniors. "If technology could help you be certain that your parent maintains social contact, takes her medicine, moves around okay, and that means she can stay another 18 months in her own home, then that's a slam-dunk motivator," Bobick says. "When we told that to the people from Intel, they just loved it." **Intel** is now one of the project's corporate sponsors, along with **Motorola Labs, Andersen Consulting** and **Mitsubishi Electric Research Lab**.

Two engineers from **Sprint**, which is interested in the project, arrive on a fact-finding mission and join us as we resume the tour. "On the surface, this could look like *Big Brother* or *The Truman Show*," Bobick concedes, gesturing to the video cameras aimed at us from several directions. Our images pour through wires onto the hard disks of computers in the basement. "But it is important to realize that we want to process video data at the spot where it is collected," he continues. "Then these won't really be video cameras but sensors that simply detect people's location or the direction of their gaze. I want to put cameras in the bathrooms, to make that distinction clear. Suppose your shower could detect melanoma? That's something people are working on." Behind Bobick, Elizabeth D. Mynatt grimaces.

Mynatt, the only woman on the team and the one who suggested the focus on the aged, spends half her time working with caregivers and anthropologists to figure out what problems tend to force seniors from their homes and what annoyances and invasions of privacy they might trade to postpone that. This approach sometimes conflicts with the more typical technocentric style of her colleagues. "I call it the 'boys with toys' phenomenon," she says. "Someone builds a hammer and then looks around for something to bang on."

Mynatt does not want cameras in the bathrooms. She used to work with Mark Weiser at Xerox PARC, and she remembers the lessons of his first experiments with ubiquitous computers. "Xerox tried to make everyone in the building wear these active name badges that we had developed," recalls Dan Russell, who worked in Weiser's group at PARC for several years before moving to IBM Almaden. The idea was to let anyone see where anyone else was at any time. "About half the people said, 'No way.' We also tried to put a Web cam in the coffee room, but again there was a huge backlash." This was at the lab where UbiComp was born.

"Still, I feel uncomfortable about focusing too much on the social implications," says Gregory D. Abowd, co-director of the Aware Home Research Initiative. Abowd is designing software that will automatically construct family albums from the video streams collected by the house—the same streams that Bobick claims he wants to distill at each source. Abowd is also trying to build an intercom system that will allow one person to speak with another simply by saying the person's name. And he enthusiastically describes his idea for a program that would automatically place a phone call to your mother when you talk to her picture—but only after checking with her house to make certain she is awake. "I'm under no illusion about the potential this creates for major privacy problems," he says. "But I'm one of 12 children. I'd rather push the boundary of privacy than cower from it."

Just over Abowd's head, a digital photograph of someone's grandmother sits on the mantle. The photo is bordered by pastel butterflies of various shapes and hues. It is a prototype of a device that one might place on an office desk to keep track of a distant relative living in an "aware" home. Every day the photo would contact the house for a status report from the system that tracks Grandmom's physical movement and social interaction; more activity would add a larger butterfly to the history. The idea, suggests Mynatt, who designed the device, is to find calming technology that helps family members feel close and in control without being invasive.

She describes another active project over lunch: "We know that kitchens are hot spots of activity and that older people suffer some cognitive declines that make it difficult for them to deal with interruptions." So she is designing a reminder program that will use the kitchen cameras and sensors to assemble a running montage of snapshots that can remind people what they were doing just before they were interrupted. She is similarly trying to come up with subtle sounds or images that the house can emit to help inhabitants remember important times of day, such as for appointments or medication. Other researchers want to stick small radio-tracking tags on easily misplaced objects such as keys and remote controls. The list of ideas seems to change weekly, reflecting the enormous uncertainties in the UbiComp field about what society needs and what people will accept.

In a year or so, test subjects will help answer that question as they move into the second story of the house and judge whether all this complex infrastructure and software does in fact simplify and enrich daily life. The project has its skeptics. There is no way to know what Weiser would think, unfortunately, because he died suddenly last year from liver cancer at the age of 46. But his colleague Rich Gold worries that the occupants of a UbiComp house may feel it controls them rather than the other way around. In an essay on "intelligent" houses several years ago, Gold wondered: "How smart does the bed in your house have to be before you are afraid to go to sleep at night?" —*W. Wayt Gibbs*

A Machine for Living In

The four-bedroom, four-bath Broadband Institute Residential Laboratory built by Georgia Tech has more cameras than windows. Amenities include:

- Computers: at least 60
- Video cameras: 25 (first floor only)
- Microphones: at least 1 per room
- Cabinet sensors: 40 (first floor only)
- Televisions (for fun, not research): 60-inch upstairs, 8-by-12-foot projection system in basement
- Network outlets: 48 (at least one per wall)
- Connections per outlet: 2 Ethernet; 2 coaxial; 2 optical fiber
- Internet bandwidth: 2 gigabits per second (via 4 DSL lines and an optical-fiber link)
- Internal wireless network bandwidth: 11 megabits per second
- Construction cost: at least $750,000, not including computer equipment

NETWORK CABLE: about 10 miles' worth in total.

GARY W. MEEK

Wholesale Computation

Companies want to sell your computer's spare processing time. Are there buyers?

The fastest supercomputers in the known universe are virtually free. All you need to beat the performance of a $50-million, massively parallel research machine is a little software and some way to convince 1 percent of the people on the Internet to run it. Unlike a dedicated supercomputer, which generally requires special housing and a staff of attendants to keep it going while it falls rapidly behind the state of the art, the network equivalent increases in power regularly as people upgrade their PCs. And when you're done using the virtual supercomputer, you can stop paying for it. Little wonder, then, that more than a dozen startups should have appeared in the past year, all trying to scoop up spare computing cycles and sell them to the highest bidder.

The best-known example of virtual supercomputing is the volunteer SETI@Home project, a search for radio signals from an extraterrestrial intelligence; it has attracted more than two million participants. Following in the footsteps of code-breaking ventures such as distributed.net, SETI@Home can run as a screensaver; then it is active only when a machine is not doing anything else. Each chunk of radio-telescope data can be processed independently, so machines don't need to communicate with one another, only with a central server. Other embarrassingly parallel problems include DNA pattern matching, Monte Carlo financial modeling, computer-graphics rendering and, appropriately enough, Web site–performance testing. Genome applications alone, says United Devices CEO Ed Hubbard, could soak up all the Net's spare computing power for the next 50 years.

Only two questions stand between the venture capitalists and enormous profits: Can they get millions of users to surrender CPU time to profit-making organizations, and can they sell the resulting power to enough paying customers? Steve Porter of ProcessTree Network has little doubt that his company can retain the 100,000 people currently donating time to nonprofit computations by offering payments of between $100 to $1,000 a year (depending on processor speed and Internet bandwidth). That, he says, will enable him to sell a standard CPU-year (a 400-megahertz Pentium II operating full-time for 365 days) for about $1,500, or less than a fifth the cost of equivalent time on a supercomputer. Nelson Minar of PopularPower expects that even lesser incentives, say between $60 and $200, would still cut individuals' Internet access bills in half—or add up to a tidy sum for schools and libraries. And at Centrata, business development vice president Boris Pevzner says his company intends to bypass individual recruiting entirely

DAVID SUTER

and use its high-powered venture-capital contacts to get computer manufacturers and Internet access providers to build the company's software into their products, where it will operate automatically.

Meanwhile Adam L. Beberg, one of the founders of distributed.net and now an independent software developer, predicts that no one will make money reselling computer power—too many sellers, not enough buyers. Completely open distributed computing has intractable security problems that will prevent firms from putting sensitive code and data out on the Internet for everyone to see. "The only market is behind firewalls," he says.

Andrew Grimshaw of Applied Meta agrees: "Most businesses won't buy consumer-grade [computing] resources from some Linux hacker's dorm room." Beberg and Grimshaw both argue that the real money is to be made with corporate networks, where tens of thousands of well-administered machines sit idle every night. (Applied Meta currently operates for the National Science Foundation a seamless, secure network of more than 4,000 CPUs.)

Proponents downplay such worries, pointing out that encryption, along with the very decentralized nature of the computing, make it unlikely that an adversary will be able to piece together more than a tiny bit of the big picture. Porter says that his company is mostly bidding on projects based on publicly available data and algorithms—it's only the computing power that his clients need. Minar points out that there's just as much need to protect PCs from potentially malicious distributed code. His company places programs in a Java-language "sandbox" that isolates them to prevent unauthorized access to a user's own information.

Moreover, it isn't just cycles that will be for sale. Centrata and Applied Meta, for example, both tout their ability to store information on what looks like one enormous disk. (Redundancy and encryption are just the beginning of the techniques required to make sure that the data are consistently available to the owners and inaccessible to anyone else.) Porter and others are also looking forward to trading in bandwidth: a PC with a megabit-per-second Internet connection, typical of cable modems and DSL connections, could cache data from distant Web sites and serve them to neighboring users, reducing the load on Internet backbones. (Companies such as Akamai are already doing a rapidly growing business in such "edge" caches, but their approach requires dedicated hardware.)

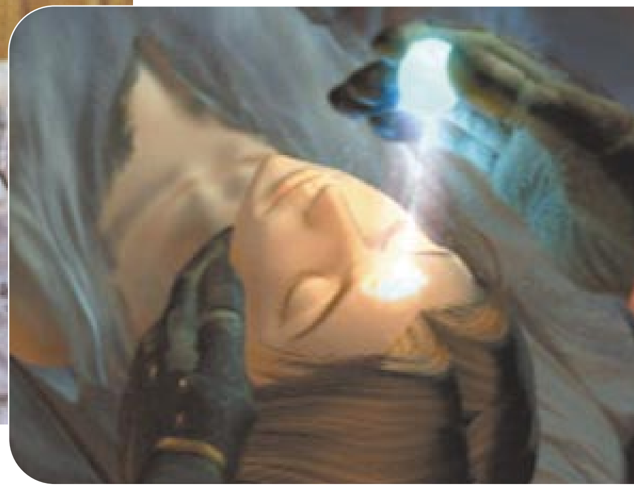
So in a few years, your computer could be surfing the Net looking for the best bids for its spare resources. But will the ready availability of computing power to handle peak processing loads end up curtailing the rapid increases in CPU speed that make distributed computing attractive, or will the ability to solve problems that were utterly unapproachable only a few years ago whet appetites for yet more power? That issue might not even concern the startups. It's possible that widely disseminated distributed-processing software—such as that recently released by Beberg and his friends—will allow buyers and sellers to work directly, leaving the intermediaries hoping to sell your computer power out in the cold. *—Paul Wallich*

the future of digital entertainment

TOM DRAPER DESIGN (*illustration*); SONY PICTURES IMAGEWORKS (*Hollow Man*); ATOMFILMS (*Space Dog*); KATA NATOLA S.I.N./Corbis (*Metallica and guitar detail*); ©2000 WARNER BROS. (*The Perfect Storm*)

Music, movies, television, video games and the World Wide Web are morphing into a single entity. As these previously distinct media switch from essentially analog means of production (like celluloid film) and distribution (like delivery vans) to all-digital ones, their products are converging into one big stream of digital data. Call it d-entertainment. It will come to us on our TV screens, PCs, wristwatches and dashboard displays—anywhere, anytime. And once a few more technical and legal issues are worked out, we'll not only be able to enjoy it, we'll be able to create it and distribute it, too.

TERRY RENNA *AP Photo*



Sony's vision of d-entertainment is built in part on the Glasstron "personal theater" headset and powerful PlayStation 2 graphics such as those in the video game The Legend of Dragoon.

Predicted for years, the convergence of media content, of distribution channels such as cable TV and the Internet, and of PCs, TVs and those wireless personal digital assistants is finally starting to happen. Television shows such as *Who Wants to Be a Millionaire?* and *Monday Night Football* now synchronize their broadcasts with interactive Web sites. Delphi Automotive is taking pre-orders for a service that will bring e-mail, the Internet, digital MP3 music files and other d-entertainment options to our cars.

While transforming our leisure time, the digitization of everything audio and video will also disrupt the entertainment industry's social order. An early sign will be a shakeout in entertainment technology. The TV, VCR and even DVD players could be wiped out by a killer appliance such as the new "personal video recorders" from TiVo and ReplayTV. These magic boxes let us pause and replay live TV and skip through its commercials, as well as search for and store programs on any subject or starring any actor we like.

The shakeout could also be catalyzed by a dark horse such as Sony's PlayStation 2 game machine, released this year, whose microprocessor and graphics capabilities rival those of today's PCs. Sony could take the d-entertainment world by storm if it could sign a deal with a distribution power—say, a cable TV carrier—to complete the chain of content (Sony Pictures, Sony Music, video games), distribution, and platform (PlayStation 2). Indeed, Ken Kutaragi, CEO of Sony Computer Entertainment, which engineered PlayStation 2, says his firm will be the driver for the entire parent company. His next-generation "game" machine, dubbed PS3 by Sony, will offer online shopping and other interactive Internet services. Sony has reportedly signed an agreement to provide online banking through the PlayStation.

Digitizing everything audio and video will disrupt the entertainment industry's social order and force new issues.

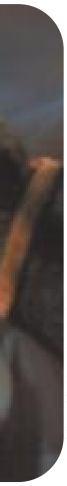
Rival corporate marriages could just as likely change the entertainment world. The proposed merger between media and cable TV giant Time Warner and Internet service provider America Online, awaiting antitrust review, represents the convergence of content and distribution. If the companies could reach a deal with a hardware manufacturer, they, too, could complete a convergence chain.

Broadcast Be Damned

The digital disruption of entertainment's social order will force the industry to confront new issues. For example, record labels are scrambling to find a profitable way to allow music lovers to download tracks online. They may have to forget the $15 CD and offer us a one-time listen of a song for a 10-cent online micropayment, unlimited play for $1, or access to their entire catalogue for $100 a month.

Consumers will clamor for entertainment-on-demand, no longer happy to be passive recipients of what media companies decide to broadcast at a given time. The new technology will let us choose from the world's d-TV, d-music and d-movies [see "Creating Convergence," on page 50], served up on the Internet. By 2020, a more robust, broadband Internet could replace all "broadcast" models—radio, TV, film, newspapers, magazines, books—as the preferred distribution medium for entertainment, predicts Martin Tobias, founder of Loudeye, a Seattle company that encodes and distributes digital media.

Creation of content will be democratized. It used to be that only big Hollywood studios could afford to film and distribute movies or TV shows. No more. Low-cost digital movie cameras and PC


SONY COMPUTER ENTERTAINMENT

video editors allow anyone with an eye to record and edit a movie for just a few thousand dollars, and distribute it through firms such as AtomFilms and IFilm that serve up video over the Web [see "Moviemaking in Transition," on page 61].

Advertising must change if a magic box allows consumers to cut out the commercials. Broadcasters might have to scroll ads along the bottom of the screen during a show to prevent us from stripping them out. Or Coke might have to pay big bucks to get the stars of NBC's *Friends* to wear T-shirts sporting the Coke logo during an episode.

And how will the copyright infringement riot be settled? The trouble besetting the music industry over online swapping of d-music on sites such as MP3 and Napster will play out on a much larger scale once d-TV and d-movies arrive en masse [see "Music Wars," on page 57]. Already, Web sites that enable distribution of d-video, such as Scour.com, are thriving. Yet despite the (mostly young) public's attitude that music and videos should suddenly be free just because they're on the Net, copyright law still dictates that artists, authors and filmmakers control the rights to their creations and deserve to be paid for them. Lawyers may also have to devise new rights and royalty terms for actors who allow believable, computer-generated avatars that look and act like them to be created for d-movies [see "Digital Humans Wait in the Wings," on page 72].

An Entertainment Economy

The emergence of d-entertainment could cause entertaining changes in society, too. Some pundits maintain that the U.S. economy could center on entertainment. Michael J. Wolf, a senior partner at the think tank Booz-Allen & Hamilton, likes to twist Irving Berlin's famous line, saying that in the digital age, "there's no business without show business." As he wrote recently in *Forbes ASAP,* marketers must achieve the same goal as network programmers—"they must now engage, inform, titillate, captivate.... In a word, they must be fun." Witness the phenomenal success of the ice-blue iMac and the Volkswagen Beetle. Hence, Wolf says, the traditional business cycle could evolve into a Hollywood-like entertainment cycle, thriving on hits and dying with flops.

More volatile issues could arise. The new Freenet software program, downloadable from freenet.sourceforge.net, allows PCs on the Net to act as transient nodes that can swap files directly, with no intermediary such as Napster. Whereas Napster swappers can be identified, there's no way to tell who posts or downloads a file using Freenet. Consumers can copy files directly from PC to PC with total anonymity. The implications are far-reaching. Whistle-blowers could post incriminating documents without fear of reprisal, and dissidents in totalitarian states could safely post antigovernment rhetoric. Then again, child pornographers could route their illicit photos and drug dealers could make online trades. Anarchy just got a shot in the arm.

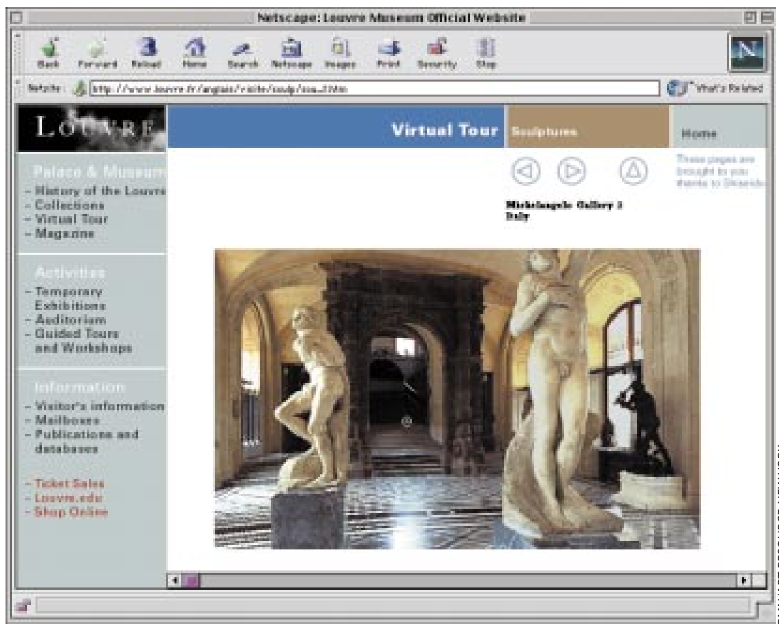
Regardless of whether d-entertainment alters society profoundly, it will change consumer habits. As Robert W. Saint John, founder of the d-video production company Nearly News Productions in San Diego, says, "The whole concept of holding a CD or movie in your hand will disappear." Why plunk down money to acquire one entertainment "thing" at a time when *everything* will be instantly available, updateable, portable and cheap?

The whole concept of holding a CD or movie in your hand will disappear once d-entertainment is widely available.

It's easy to get caught up in the vision of all entertainment going digital. Web surfers can already take virtual tours of world museums. Broadway Digital Entertainment is digitally taping Broadway plays for pay-per-view and streaming on the Internet. But no matter how realistic a virtual-museum tour, walk on the Great Wall or image of a fire-eating street performer may seem, it's not quite the real thing, because we are always to some degree aware that with a single command we can turn the machine off.

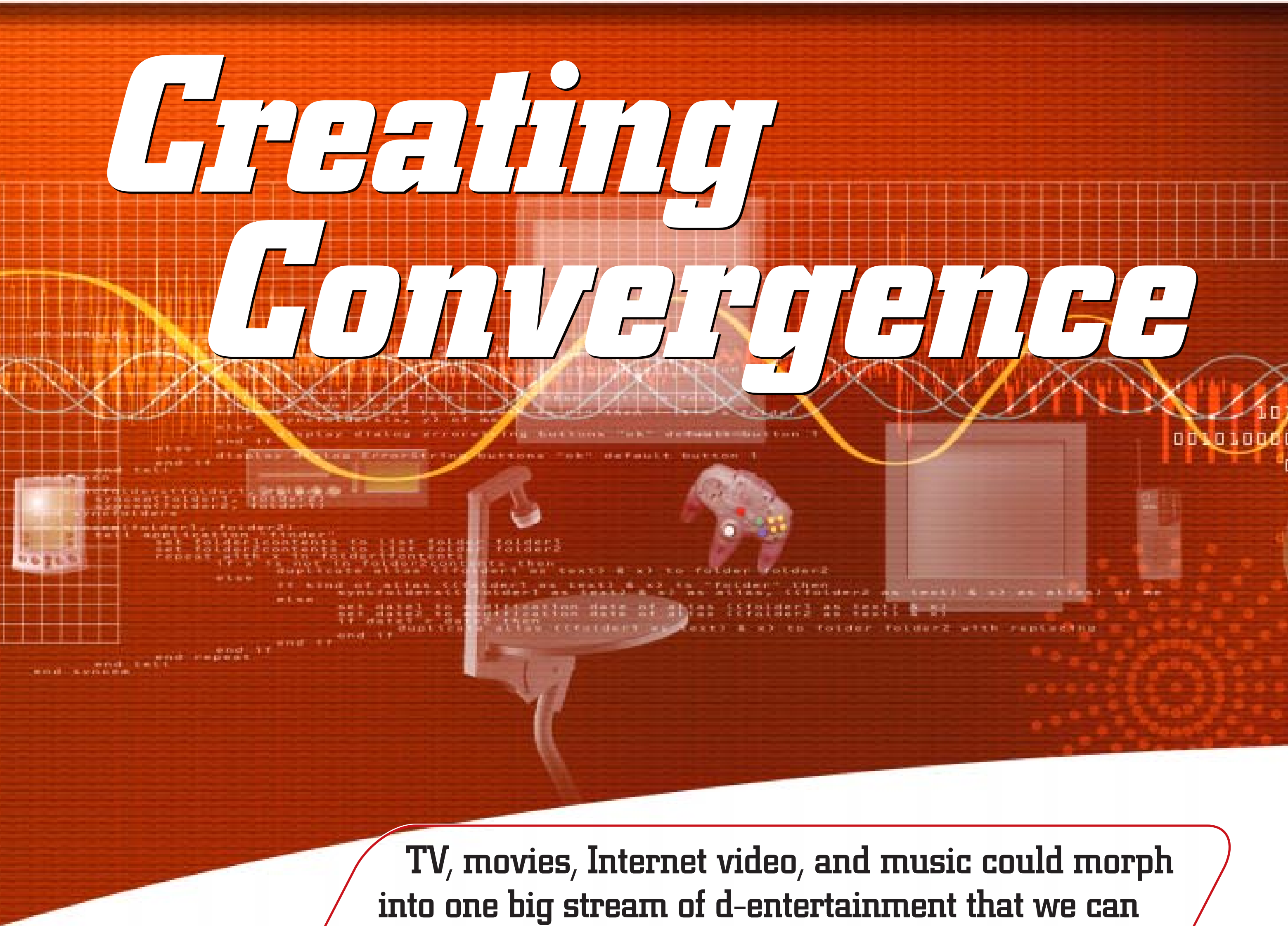
Furthermore, nothing digital can substitute for the neighborhood softball game, the county carnival, the city park. And no matter how "interactive" d-entertainment becomes, it still leaves us pretty much sitting on our butts. Sure, enjoy it. Then grab a loved one and go dance.

—Mark Fischetti, contributing editor



RMN/ART RESOURCE, NEW YORK

Web surfers can take virtual tours of many renowned museums, including the Louvre in Paris; few cultural bastions are beyond digital technology's reach.



Creating Convergence

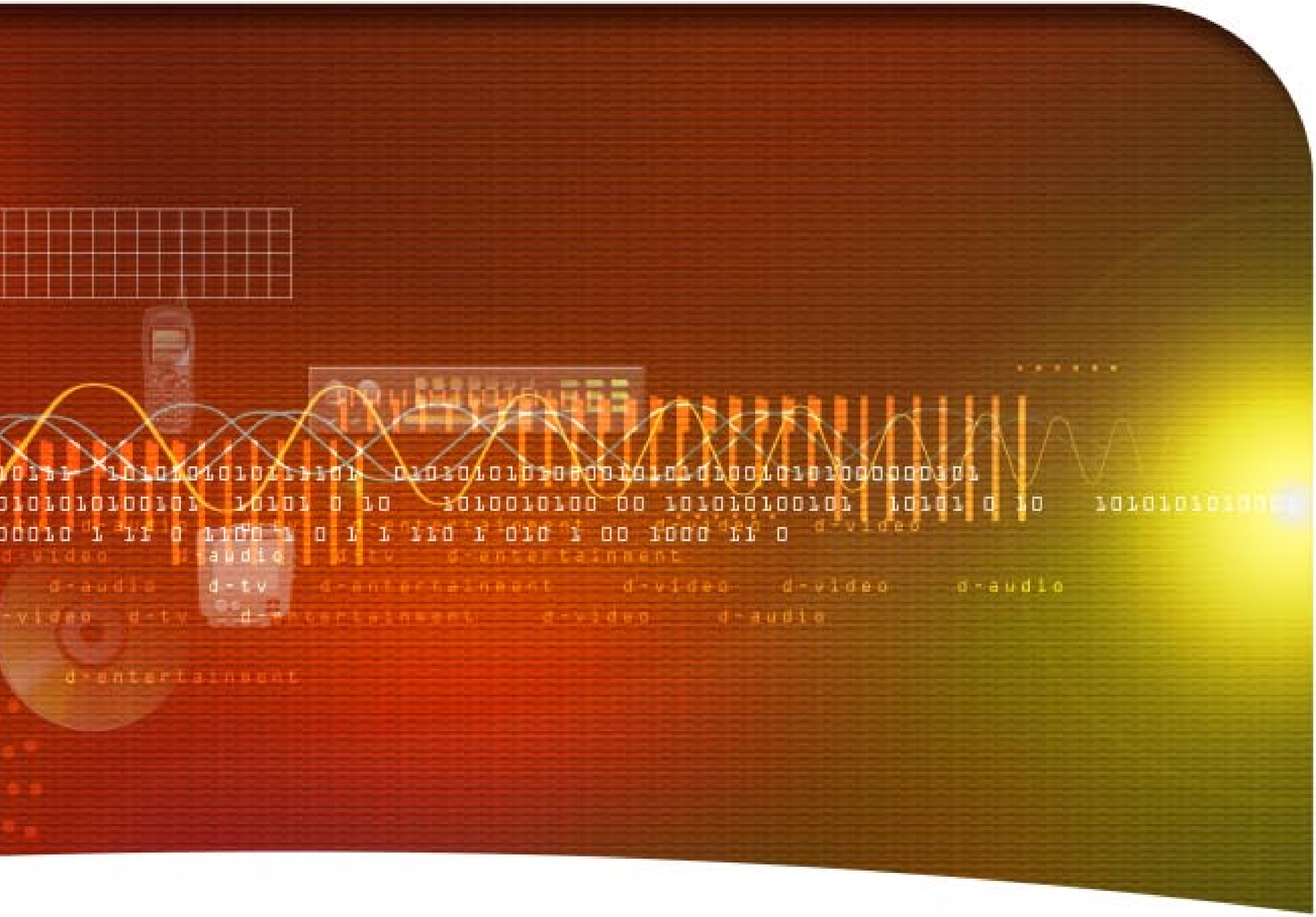
TV, movies, Internet video, and music could morph into one big stream of d-entertainment that we can enjoy on any device, anywhere, anytime. But the devil is in the details by Peter Forman and Robert W. Saint John

The 1939 New York World's Fair featured a formal debut of television broadcast, but the receiver inside the RCA Pavilion was way ahead of its time. The appliance was a combination television-radio-recorder-playback-facsimile-projector set that, in hindsight, suggests that we humans have a fundamental desire to merge all media into one entity. Today this goal has a name: convergence, the union of audio, video and data communications into a single source, received on a single device, delivered by a single connection.

Predicted for decades, convergence is finally emerging, albeit in haphazard fashion. Wireless phones, personal computers and televisions are beginning to take on one another's functions. More important, the patterns by which we are interconnecting these gadgets indicate that we are ready for convergence to sweep us off our feet. Once it does, all forms of digital entertainment will morph into one big stream of bits. We will be able to enjoy movies, TV shows, Internet video, and music on our home theater, computer or wristwatch wherever we are, whenever we want. All that is required is that equipment makers and standards bodies agree on such details as broadband distribution, copyright protection and compatible displays. No small task.

The big convergence is made up of three subsidiary convergences: content (audio, video and data); platforms (PC, TV, Internet appliance, and game machine); and distribution (how the content gets to your platform).

The World Wide Web, spurred on by the "killer



TOM DRAPER DESIGN

app" of e-mail, has greatly accelerated the convergence of entertainment content. And yet the Web's rise has also brought the quick realization that content should be scalable so that it can be delivered to all kinds of platforms, from wireless phones to TVs. This has prompted rethinking, and concern, about who creates and controls the content itself, which depends on how it is packaged and delivered to us. For example, America Online (AOL), no more than an Internet packager and deliverer, is attempting to merge with Time Warner, one of the world's biggest media conglomerates. Unfathomable only a few years ago, the $180-billion combination already seems natural. Unfortunately, Time Warner did little to assure its rivals that the merger would not squelch competition when, in May, it temporarily pulled Disney's ABC network from its cable systems, which go to millions of homes. The move could influence review of the merger by the Federal Communications Commission (FCC) and the Federal Trade Commission. Concerns about distribution have also been raised by Napster, software that allows users to download digital music (the majority of which has not been licensed) directly from other users' computers through the Napster Web site [see "Music Wars," on page 57].

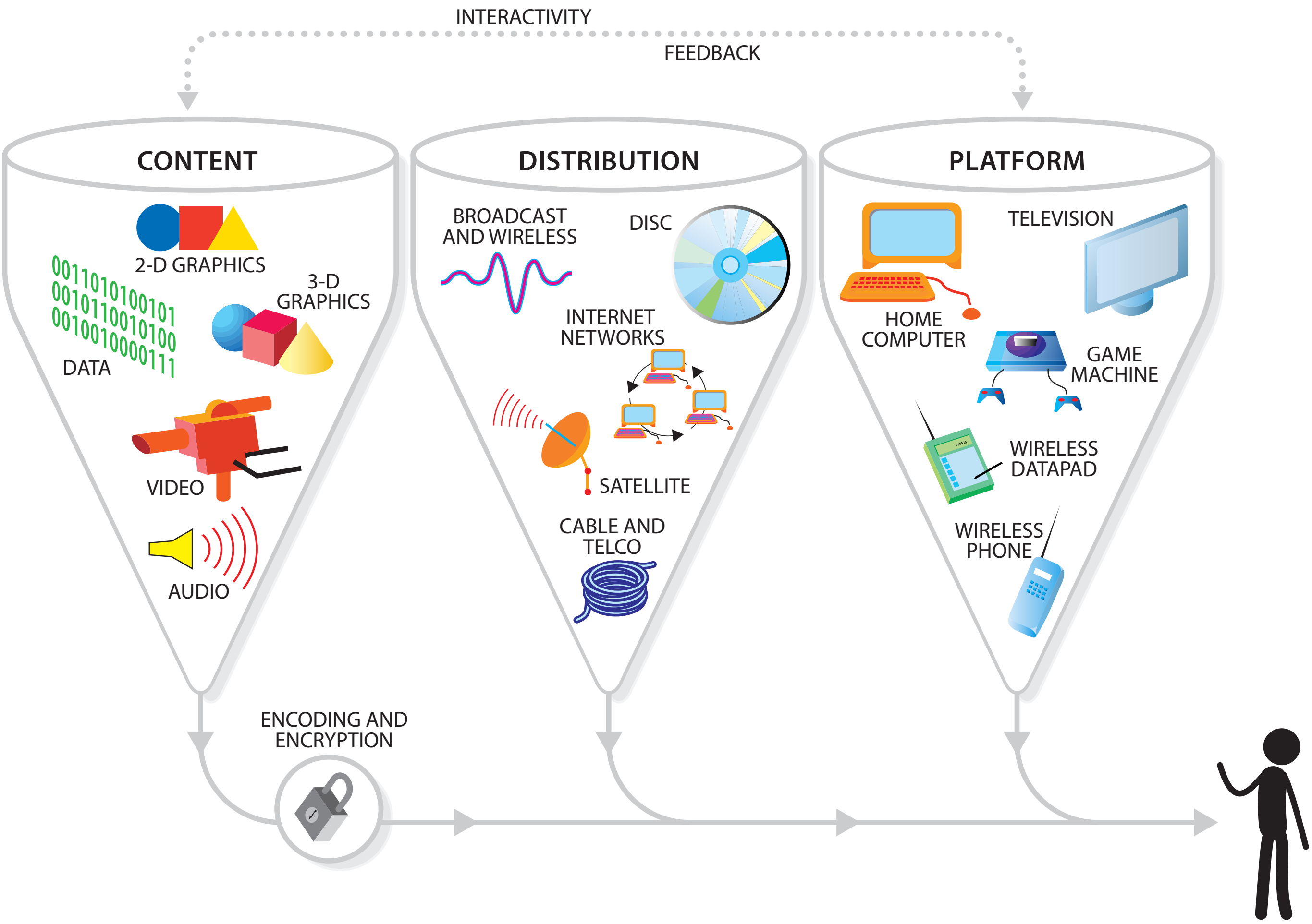
Though not the desired results of convergence, these issues indicate that changes in d-entertainment are so unprecedented that they may require government oversight. What is clear in all the muddy water is that the three elements of convergence are powerfully intertwined. The platforms, however, are where technological choices are most wide open. The number of competing standards and architectures is enormous. How the competition plays out will largely determine how complete, and how soon, convergence will emerge.

Fight or Switch

The earliest example of converging platforms involved crude attempts at interactive television. The most notable trial was Warner Amex's QUBE system tested in Columbus, Ohio, in the 1970s. What it best demonstrated, at colossal expense, was that people did not want to communicate back to the broadcaster.

It took the Internet to "train" people to interact with content. The most recent success story is not quite convergence, but it is a step in the right direction. In March, Disney's ABC Television, in conjunction with its Internet sibling Go.com, launched the "enhanced TV" version of the game show *Who Wants to Be a Millionaire?* By logging on to the *Millionaire* Web site, viewers can play along on their PC while they watch the show on TV. In the first month, 3.5 million visitors accessed the site. This approach is still a "two-screen" experience, however, requiring a TV and PC in the same room. Pundits argue whether the ultimate device will be the PC/TV or the TV/PC. Whichev-

If compatibility standards and data-protection schemes can be worked out, all d-entertainment will converge into a single source that can shine into your life on any screen, wherever you are.



The convergences of digital content, broadband distribution and display platforms create the "big" convergence of d-entertainment and information, with feedback supporting human interactivity.

er hardware leads the way, substantial hurdles must be overcome.

Digital television is evolving along three paths: enhanced resolution, multiplication of channels, and interactive features. Some advances are being manufactured into the television sets themselves, whereas others are being worked into set-top boxes that connect TVs to cable television providers. In 1995 the international Advanced Television Systems Committee (ATSC) issued a set of digital television standards to address the issues involved. But adoption will not come overnight. Digital information necessitates the costly replacement or enhancement of nearly every piece of equipment in the distribution chain, almost all of it analog, from production to broadcast to the television set. In the U.S., the FCC imposed a deadline for all stations to make the changeover to d-TV by 2006, and Congress gave broadcasters free, extra spectrum to make the adjustment. The major networks began various types of digital broadcast in 1998, but conversion has been slow, in part because of political battles. For example, some broadcasters want to lease their extra bandwidth to wireless communications companies, but Congress is telling broadcasters, in effect, "No. You got it for free. Now get moving on digital."

One big technical concern is that the standard mandated for modulation and transmission, known as 8-VSB (vestigial sideband), may be inadequate for reliable reception with antennas; a significant number of people are still not served by cable or satellite. And even with a new digital antenna, a viewer can't easily seek out the best reception by moving it around; a d-TV signal is either received with complete accuracy or not displayed at all. Some manufacturers and broadcasters (who have already spent hundreds of millions of dollars on the transition) are petitioning the industry, the FCC and Congress to change the standard to the arguably more robust European COFDM (coded orthogonal frequency division multiplex), which they say would solve the problem.

Others, unconvinced, are not willing to switch to COFDM equipment when digital transmitters and receivers based on 8-VSB are already in the marketplace. They say a changeover would delay the rollout of d-TV by at least five years. In this environment, consumers are unlikely to replace their analog televisions with digital ones. Most experts expect the entire transition to take 10 to 15 years.

Another issue is what to do with the extra bandwidth, a matter not addressed by the ATSC measure. Some industry leaders favor dedicating it to high-definition television (HDTV), which has greater resolution than standard-definition television (SDTV). Others feel that the extra bandwidth should be used to provide multiple channels of SDTV. Still others want to offer slightly higher resolution for certain types of content along with multicasting and interactivity.

A television that can support all these formats

BRYAN CHRISTIE

will be expensive. Economies of scale will be needed to reduce cost, but until there is enough compelling digital content, consumers will simply defer their purchases. Despite this situation, the d-entertainment industry understands that content really is king. Feverish experimentation is under way to find the successful content formulas that will pull consumers onto a fully digital platform.

The U.S. is behind the rest of the developed world in d-TV. Japan will be converting its analog HDTV service to digital in 2002. The U.K. seems furthest along in Europe, broadcasting widescreen SDTV instead of HDTV. Unfortunately, Japan, Europe and the U.S. are pursuing different TV standards, as they have done for decades. In an ideal world we'd all be using the same standard, but politics and the not-invented-here syndrome seem to disallow it.

Harder to find is true interactive television (ITV). The ATSC and European DVB standards do not yet address ITV, although Europe is implementing some systems. Progress may be sparked by OpenTV, a company in Mountain View, Calif., funded by big players such as AOL, Time Warner and News Corp. It provides the software and middleware (the technical architecture) needed between an ITV broadcaster and viewer and licenses it to cable operators. The software is used in nearly eight million set-top boxes worldwide for cable networks such as the U.K.'s BSkyB and Germany's PrimaCom AG. It is based in part on standard Web languages such as HTML and will soon adopt the newer XML language. The set-top boxes support interactive features such as electronic program guides, e-mail, online shopping, video-on-demand and custom advertising. Meanwhile Britain's Cable & Wireless Communications and others are using the Liberate ITV platform (based on HTML and JavaScript) to deliver interactive services over cable modems, such as grocery delivery and banking services, in addition to telephone, e-mail and digital television.

The National Cable Television Association predicts that the U.S. may catch up by next year, when 75 percent of cable systems will be wired with enough bandwidth for interactive services. Cable is not the only delivery solution for ITV, but it has the greatest penetration in the U.S. television market. At the same time, a cross-industry alliance of computer and broadcast companies called the Advanced Television Enhancement Forum has been formed to try to provide a common development environment for ITV using HTML, XML and JavaScript. More regional efforts include Europe's DVB Multimedia Home Platform, the ATSC's Digital TV Application Software Environment and Japan's ITV standard.

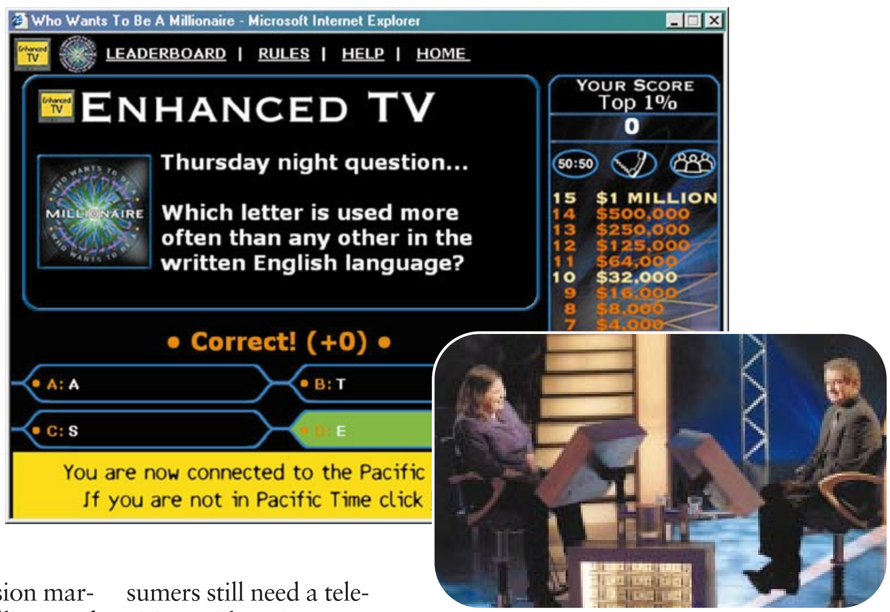
Not everything in d-TV is a potential Tower of Babel. The industry has agreed, worldwide, on the method of compressing video so that it can be transmitted faster and then decompressed so it looks as close to the original as possible. This is achieved through a codec (compressor/decompressor) based on sophisticated algorithms. The standard codec for d-TV is MPEG-2, named after the Motion Picture Experts Group that designed it. MPEG-2, which works with any display platform, has become the pervasive standard for digital television, digital cable, direct broadcast satellite and digital videodisc (DVD).

The MPEG committee has been working on a next-generation codec standard, MPEG-4, which in many ways is defined by convergence. It is the step-up needed to support high-quality, streaming d-video on the Internet, including data that specify interactive elements. It can even stream video at very low bit rates (five kilobits per second and up) that can be handled by mobile wireless networks. MPEG-4, an open standard, will drive interactive d-video for future consumer electronics.

Seeking the Killer App-liance

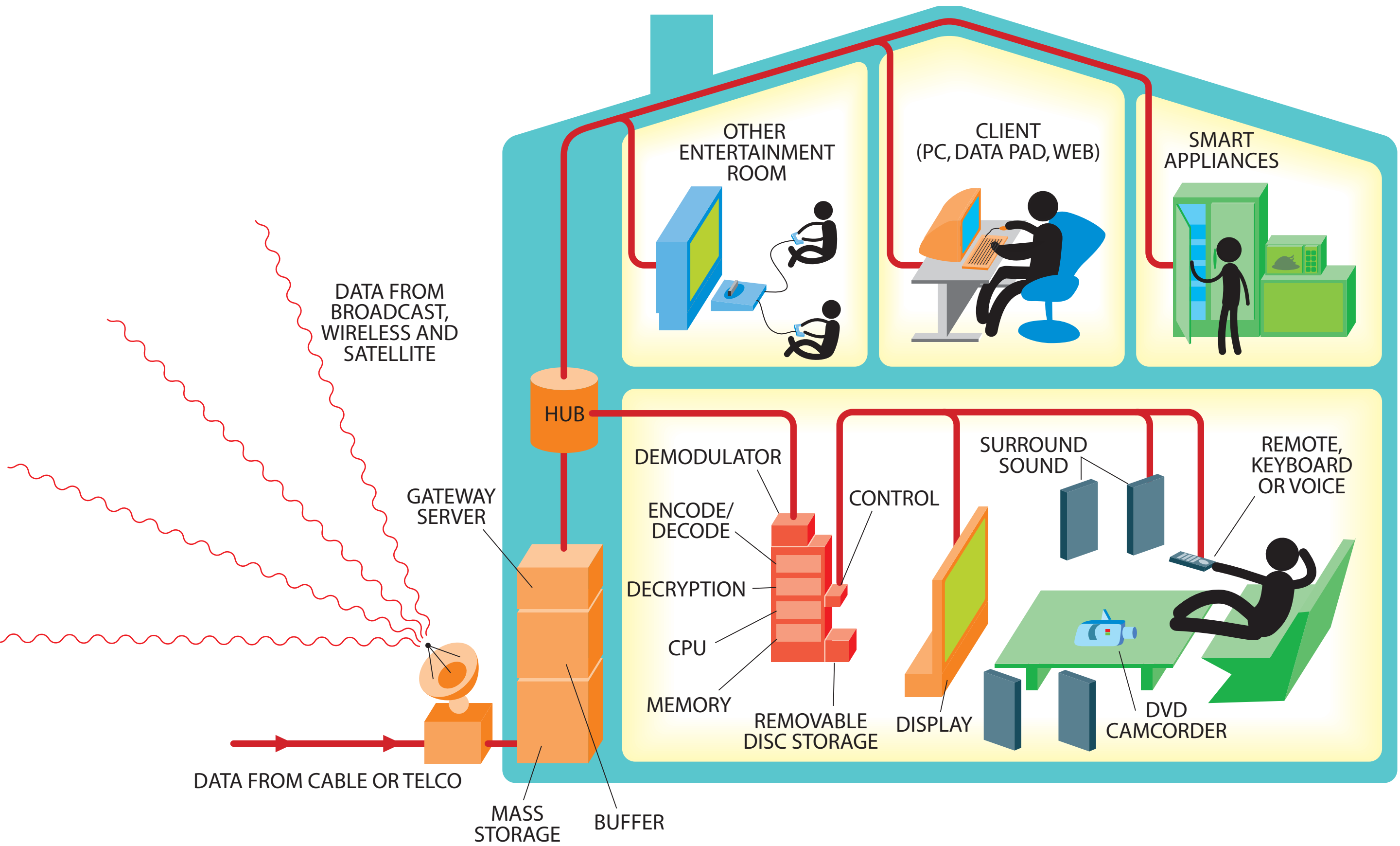
All this technology requires decoding and connection equipment more sophisticated than that found in a standard television. Although some digital broadcast satellite receivers and digital cable TV receivers are already shipping, they are not actually capable of receiving d-TV broadcasts. Consumers still need a television with an integrated d-TV receiver. A "killer" d-TV appliance is needed, and it could spring from any of three current television applications: the DVD player/recorder, so-called personal televisions or game machines.

DVD players constituted one of the most successful product introductions in consumer electronics history. After only two years, more than four million players are in U.S. homes, and the



The *Who Wants to Be a Millionaire?* Web site allows viewers to play along in real time as the hit TV game show is broadcast, a high-profile step toward interactive media.

ABC, INC.



BRYAN CHRISTIE

Homeowners will use a variety of interfaces to enjoy and control an incoming stream of d-TV, d-audio, d-video and data, managed centrally and bused around the house over wires or through the air.

number will easily exceed 10 million by the end of 2000. A DVD can store more than two hours of medium-definition MPEG-2 digital video with multiple tracks of high-quality audio, navigation information and graphics. Some also come with interactive features: the DVD release of *Independence Day* allows viewers to see alternative versions of scenes not shown in the movie and to take part in a space battle game. Players slated for 2001 are certain to displace the VCR, as well as the CD-ROM drives in computers, converging data, audio and video into one source on one medium, at least for the PC.

The personal video recorder, or personal television, is a second promising convergence appliance that emerged in 2000. It is a massive-storage hard drive for TVs made by companies such as TiVo and ReplayTV. A broadcast is cached in MPEG-2 format to the hard disk. The viewer can pause and resume playback while continuing to record the live transmission in the background, allowing him or her to create instant replays—or skip over commercials. An individual can program in viewing preferences, such as "live sports" or "opera," and the device will record such broadcasts in "virtual channels" that can be viewed later. The machine also can scan electronic listings for similar programs and automatically capture them. This kind of "smart" recording, access to enhanced program guides and "live pause" are three must-have functions for future d-entertainment.

The third forerunner of the ultimate d-entertainment platform is the widely popular video-game console. The release of Sega's Dreamcast in 1999, with its 56K modem, marked the debut of a game machine that allowed players to compete with one another over the Internet. Sony raised the stakes this fall with PlayStation 2. It has a DVD drive, Dolby Digital and DTS sound, and a CPU and graphics processor rivaling those of today's PCs. Although it will not initially ship with a modem, it will connect to modems shortly. Even more interesting, Sony recently announced plans to license the technology to other parties, so support could start showing up in TVs, set-top boxes, PCs and even other game machines.

The Entertainment PC

The enticing television technologies could still be thwarted by basic roadblocks. For example, they would require an interface more powerful than the standard TV remote control yet less clumsy than the computer keyboard. None of the technologies provide convenient interactivity. And an all-in-one device may remain forever on the wish list if manufacturer competition forces consumers to erect a growing stack of peripherals next to their television screens. Given these uncertainties, convergence could just as likely be driven by the PC and the Web.

With 1-gigahertz CPUs entering the market, 40-gigabyte hard drives available for $150, powerful graphics processors to handle video manipulation and display, easy on-line interactivity, and rewritable DVDs just around the corner, the PC is ready to take its place as a d-entertainment platform.

The argument over whether the PC or TV will lead to a convergent appliance resembles the "thin" versus "fat" network/client debate of the mid-1990s. The proponents of thin—new-wave companies such as Sun, Oracle and Java—saw a world populated with simple multimedia boxes devoid of hard drives, which would simply download applications (and thus entertainment) from the Net whenever desired. They squared off against the old guard—Microsoft, Intel, Dell, Compaq—which wanted to continue to sell the fat computers stuffed with storage and software. In the d-entertainment world, the TV is a thin client, the entertainment-savvy PC the fat alternative. Which wins will rest on the same unresolved issues: Where are media files stored? Do power and control reside in the hands of the consumer or the network? How much complexity will consumers tolerate in exchange for more features?

There are some good arguments for the fat PC. It easily creates, stores and shares media; the TV does none of these. Economies of scale have led to plummeting prices. And in a world of complicated, competing TV standards, it might make sense to adopt a generic, programmable PC that can support them all, rather than buy and attempt to connect that stack of TV peripherals. Indeed, the entertainment PC could be a cleverly repackaged TiVo Personal TV or PlayStation 2.

It might make sense to adopt a programmable PC that can support competing TV standards, rather than connect a stack of TV peripherals.

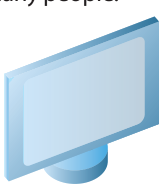
PCs also have the unique ability to produce and receive media streams over the Internet. Streaming audio and video, played "live" as the consumer receives it, has empowered online radio stations and a growing industry of companies such as IFilm and AtomFilms that distribute films by small independent filmmakers. It has also allowed established d-entertainment companies such as MTV and Time Warner (through its subsidiary Entertaindom) to deliver content that has no other venue. Trailers

Winning the Distribution Battle

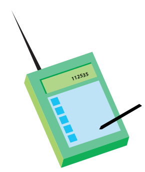
Business models, more than technology, may determine how d-entertainment reaches you. But any long-term winner will have to provide transmission rates upward of 100 megabits per second (Mbps). Here are the options, which would have to improve:

Telephone Network. Ubiquitous, but the standard 56 kilobits per second (kbps) modem can't begin to download d-entertainment fast enough. An "always-on" asymmetric Digital Subscriber Line (DSL) connection provides speeds up to 1.5 Mbps. Very high bit rate DSL promises up to 60 Mbps. DSL cannot handle transmissions farther than three miles from a switching center, however, which leaves out many people.

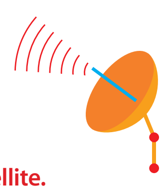
Cable TV. Companies have been laying hybrid fiber-optic/coaxial cable to increase capacity to at least 6 Mbps but have been sidetracked with conventional competition against broadcasters, telephone companies and Internet service providers. The eventual integration (likely through mergers) of broadband, content and a two-way connection will make large cable companies the most formidable d-entertainment distributors. One drag: it will take at least three years for manufacturers to replace the 70 million analog TV converter boxes in homes with simple digital set-top boxes.



Broadcast TV. Stations still depend on the limited 6 MHz/19.39 Mbps of bandwidth for each channel. MPEG compression gives broadcasters the choice to send one high-definition TV signal or four standard signals, or some combination of video, audio and data. But a second connection is required for two-way interaction.



Fixed Wireless. Can provide access at up to 2 Mbps; a good option for rural customers who can't get cable or DSL. The connection is made between radio towers and a rooftop antenna. Wireless is the only way to connect digital PCS phones and handheld computers to the Net.



Satellite. DirecTV provider Hughes Network Systems also offers DirecPC for Internet downloads at up to 400 kbps. Outgoing data must be sent via modem, however, producing noticeable delays. There are a few true, two-way interactive systems, but they are very expensive.

BRYAN CHRISTIE

for such films as *Star Wars: Episode I, The Phantom Menace* and live video streaming directly from CBS's *Big Brother* Web site (24-hour Web-cam surveillance of unrelated inhabitants of a private home) are some of the Internet's biggest hits.

Streaming media will play an important role in d-entertainment, but it won't be a prime mover because of poor resolution; TV-quality video requires much greater bandwidth than most consumers have on their PCs. Furthermore, the Internet's current IP architecture is not robust enough. Media streams suffer under network congestion and weak spots as they pass through servers and routers that have not been scaled for this kind of demand. A few companies like Akamai provide some relief by caching high-demand streams across many dedicated servers closer to end users, but it's not enough to provide the end-to-end quality of service that broadcasters and viewers demand. The continued proliferation of broadband connections, and initiatives to upgrade the Internet such as IP version 6, will help resolve some of these shortcomings.

The roadblock to an entertainment PC could be the PC itself. Even a cheap TV doesn't crash or freeze. The best computers still do.

Another problem is the proprietary architectures and codecs of software such as RealVideo, QuickTime and Windows Media. Hardware manufacturers won't build systems on software that is not open and a certified standard. Software makers will have to comply voluntarily if their products are to work together in convergent fashion.

If these advances occur, then an MPEG-4 stream coming to your PC over a broadband connection would be indistinguishable from cable TV—and you could store or edit those videos as well as distribute your own. The remaining roadblock could be the PC itself. For all its glory, the PC is still unreliable, a function of its complex operating system. Even a cheap TV doesn't crash or freeze. The best computers still do.

Coming Home

By now, we might conclude that there are far too many technologies and trade-offs for convergence to be anything but a consumer nightmare. Keep in mind, however, that the introductions of the personal computer and the Internet were fraught with competing hardware and standards. That didn't stop either from becoming wildly successful.

In the near term, we can probably expect to get the most out of our convergence experience with a mix of separate devices, many connected by "wire" to the Internet but not necessarily to one another. The devices will gradually take on common functions and become more powerful. Various "many-in-one" solutions with similar features will arise, most likely connected throughout the home by a high-bandwidth, wireless local-area network running off a powerful central server tucked away in the closet or basement. The server will maintain an always-on broadband connection to the outside world over fiber-optic lines or satellite links. We will live in a world of many devices, many networks and limitless scalable content, united by invisible connectivity.

Like the platforms, the winning content of convergence will be determined by the economies of scale for the consumer, perhaps bolstered by the mergers and alliances of companies that provide the most compelling packages of content and delivery. More interesting than "how we get there," though, will be what we do with convergence once we have it. History shows that when consumers are given new technological choices, the options that succeed are often different from what anyone imagined or intended. Although it's still far off, the outcome has already been given a name, a term that will probably become the buzzword of the next decade: emergence. SA

COURTESY OF THE AUTHORS

The Authors

PETER FORMAN (*left*) and ROBERT W. SAINT JOHN created convergence software at Ligos Corp. in the late 1990s. Forman is president and CEO of Ligos in San Francisco, which specializes in real-time, software-only video codec technology. He has been president of Intelligence at Large, a developer of Internet telephony and videoconferencing; a vice president of Image G; and founder of New Video Corp., a pioneer in personal computer video and audio codecs for media authoring and training systems. Saint John is an independent video producer and multimedia consultant in San Diego. His company, Nearly News Productions, focuses on video production for videodiscs, the Internet and future d-video platforms. He has been director of technical marketing at Ligos and a digital video editor and has taught courses on video compression and 3-D graphics.

Further Information

VIDEO COMPRESSION. Peter D. Symes. McGraw-Hill, 1998.

THE GUIDE TO DIGITAL TELEVISION. Edited by Michael Silbergleid and Mark J. Pescatore. Miller-Freeman, 1999.

DIGITAL CONVERGENCE. Andy Covell. Aegis Publishing Group, 2000.

SAM'S DIGITAL TELEVISION REPORT (www.teleport.com/~samc/hdtv).

CONSUMER ELECTRONICS ASSOCIATION (www. digitaltelevision.com).

LISA PINES; NEWSMAKERS (*inset*)

Music Wars

Internet distribution of quality d-audio is rapidly being perfected, but the precedent-setting legal battles have just begun

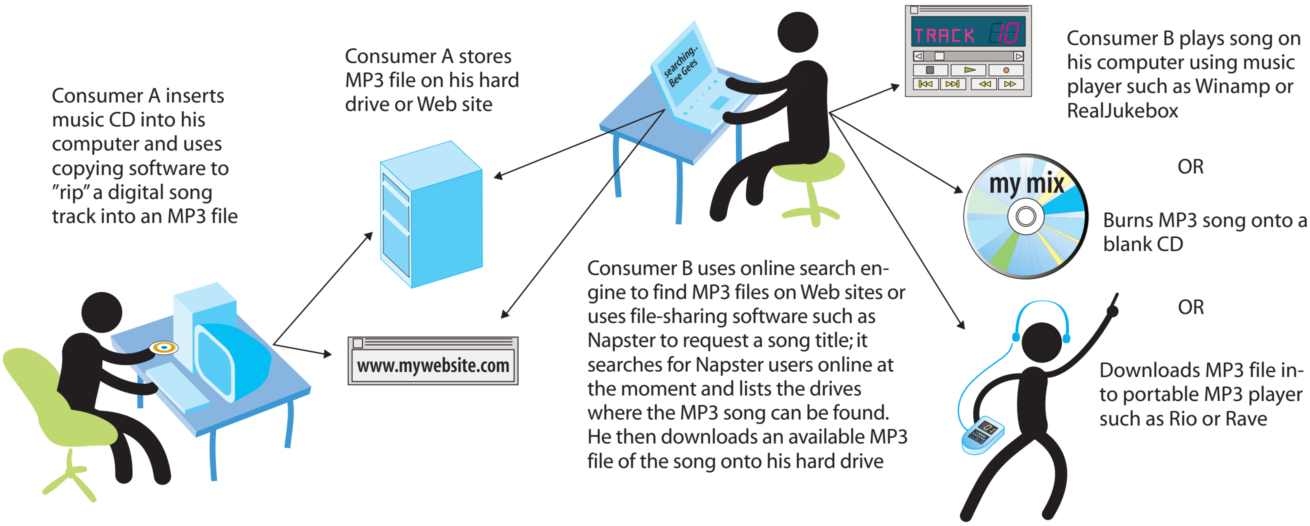
by Ken C. Pohlmann

The free downloading of music files through Web sites such as Napster (*above*) onto computers or portable MP3 players like Rio (*inset*) has touched off a contest of countermeasures to protect artists and record labels.

The technology behind digital music is developing more rapidly than the music business can handle. As headlines constantly remind us, the recording industry is scrambling to cope with new formats and distribution modes that threaten its hegemony in the delivery of recorded music. The rising popularity of the Internet as a conduit for recordings has triggered irreversible changes in the way the public expects to experience music. Ostensibly "futuristic" concepts such as music-on-demand, access to record-label catalogues and the ability to surround oneself with a steady stream of new music for free are already here. Music-as-data is creating a new paradigm for the production and delivery of recordings that has befuddled music's own creators. And the challenges foreshadow those looming even larger on the horizon for d-movies, d-TV and, indeed, all forms of d-entertainment.

The irony is that musicians, their record labels and consumers were all perfectly satisfied with the now ubiquitous compact disc and CD players, both cheaply manufactured, reliable and convenient to use. The interchangeability of discs between computer and stereo systems made the CD family a nearly ideal audio carrier. But technology never rests. The battle over the network distribution of music is driven by how audio is produced and played and by how technology allows copyright protection to be breached.

Digital recordings sample an audio signal and save the amplitude of each sample as a digital "word." A combination of sampling rate and word length determine the final sound quality: the higher the sampling rate, the higher the frequency



Copying songs over the Internet is free and easy with the aid of MP3 and Napster software, available gratis online.

response; the longer the word length, the lower the noise. The industry standard for CD dictates a 44.1-kilohertz (kHz) sampling rate and 16-bit word length. That yields a bit rate of 1.41 million bits per second (Mbps), which delivers adequate fidelity. But it does not allow speedy transmission over the Internet. Depending on Net traffic, a single CD track lasting three minutes might take 90 minutes to download over a typical 56K modem. One solution is to reduce the sampling rate, but fidelity (specifically, frequency response) degrades. A more cunning approach is to reduce the word length. This increases noise; perceptual coding methods, however, enable engineers to shorten the word length considerably with good results. An encoder ignores the parts of a d-music signal that are inaudible and reduces the data needed to convey sounds. Depending on bit rate, the final fidelity can range from nearly indistinguishable from the original to unlistenable.

The family of these perceptual coding algorithms that dominates the audio industry has been devised by the Moving Picture Experts Group (MPEG). One of the MPEG standards defines a "Layer 3" coder, known as MP3, which uses stereo bit rates ranging from 64 to 320 kilobits per second (kbps). A rate of 128 kbps achieves a compression ratio of 11:1 over a CD recording and permits fairly rapid transmission over average Internet connections. A rate below 128 kbps generally introduces audible distortion, whereas rates above 192 kbps should sound identical to an original CD source. Although low rates (such as 64 kbps) might be necessary in today's era of telephone modems, higher rates (192 kbps or better) will become the norm as more broadband connections are established. Whatever the bit rate, MP3 reduces d-music files to more manageable sizes. At 128 kbps, that same three-minute CD track would take about eight minutes to transmit via a 56K modem. And a 20-gigabyte hard drive could hold 300 digital albums, transforming a PC into a digital jukebox.

The spread of MP3 coding has triggered fundamental change in the music industry. A consumer can now transform CDs into MP3 files (a process known as ripping) using programs called MP3 rippers or CD grabbers. The user inserts a CD into his computer and uses an MP3 encoder to condense it into a tidy MP3 file. The individual can post the ripped file on his own computer or Web site, making it publicly available.

Ripping Mad about MP3

That's where the trouble begins. Converting your own CDs to MP3 is not illegal if you use the copy for your own personal use—say, on your MP3 player. But publicly redistributing music without permission from the copyright holder violates copyright law. Organizations such as the Recording Industry Association of America (RIAA) contend that posting MP3 files constitutes copyright infringement, and they have confronted large MP3-swapping Web sites, such as the infamous Napster. Using freely downloadable Napster software, individuals can locate MP3 songs on the hard drives of other people currently logged on to Napster and copy them free—and quickly, with a fast modem connection. Because the copies reside on thousands of Web sites that come and go through connections to Napster, it is difficult to assign blame. So the RIAA and bands such as Metallica have sued Napster for contributory copyright infringement.

Of course, one clever idea begets another. A program called Gnutella allows two users to swap MP3 files directly over the Internet, without having to be routed through a central server such as Napster. Gnutella was developed by Nullsoft (which also created WinAmp, the most popular MP3 player software) and released as a downloadable beta version. Nullsoft is owned by America Online (AOL), which intends to merge with Time Warner, owner of Warner Music—a vocal critic of MP3. Not surprisingly, the Gnutella site was deemed an "unauthorized freelance project" and was shut down by its own administrators. One thousand copies had already been released, however. The genie had left the bottle, and Gnutella clones started popping up everywhere.

The proliferation of bootleg MP3 files and programs like Napster and Gnutella spotlight the efficiency of Internet distribution of d-music and the difficulty in controlling that distribution. Despite the legal war, hardware manufacturers are beginning to embrace the format. New CD players play discs holding MP3 files as well as regular CDs. Meanwhile portable MP3 players such as Rio and Rave have become extremely popular. They have nonvolatile flash memory cards (the same as those used in digital cameras) that store files downloaded from a PC. A player with 64 megabytes of memory holds about 66 minutes of music coded at the quality level of 128 kbps.

Decentralized peer-to-peer networks, such as Freenet, go Gnutella one better. They employ intelligent routing and caching so that a file can move from PC to PC anonymously, making it virtually impossible to identify who posts or downloads a transmitted file. Lawsuits, injunctions and reversals have roiled about Napster and other MP3 trading sites all year. The challenge confronting the recording industry is how to prevent unauthorized copying while still offering the consumer an Internet product he or she is willing to pay for. One popular download site, MP3.com, has settled lawsuits and reached distribution agreements with EMI, Sony Music, Warner Music Group and BMG Entertainment, allowing d-music from these record labels to be stored on the service. But in September it lost a lawsuit to Universal Music, the largest label of all. The recording companies will reportedly share with their artists an undisclosed amount of money from the settlements and awards.

It is to everyone's advantage that issues such as these get worked out, because they will be replayed in larger relief when quality digital video goes online. Gnutella, for example, allows for the swapping of not just audio files but video files. This is a hotly contested area, as lawsuits over the publication of the DVD-Video encryption protocol appear to be heading for the Supreme Court.

Security and the DVD Breach

The music industry has also responded to the popularity of unregulated MP3 files by developing its own Secure Digital Music Initiative (SDMI) specification, to improve on the Serial Copy Management System (SCMS) copy protection used in the current CD format. SCMS is weak: a single bit designates whether a disc can be copied or not. This discourages casual digital piracy, but when a CD is ripped, the copy-prohibit bit is not carried forward. In the SDMI protocol, music data will be encrypted and authenticated, so users will not be able to convert a CD track into an MP3 file that could be posted on the Internet without the decryption key.

Furthermore, SDMI-compliant devices would not play illegally copied SDMI files. The protocol also allows files to be electronically watermarked, so illegally copied files could be traced to their source. D-music files can be encrypted and decrypted without affecting their fidelity. Watermarking, however, embeds a code into the audio signal, and great care must be taken to avoid audibility. This is particularly important because even the record companies envision a near future when customers can log on, listen to music samples, and then purchase and download their selections onto MP3 players or burn them into their own recordable CDs. More than 200 companies from the music content, consumer electronics, information technology and wireless telecommunications sectors have signed on to use SDMI.

SDMI backers hope to avoid the disastrous gaffe that led to the decryption of digital videodiscs—the hot new video format. DVD-Video discs are encrypted with the Content Scrambling System (CSS). A 40-bit key is needed to descramble the video and audio information. Every manufacturer has its own unique key. As a result, every DVD-Video disc has 400 of the 40-bit keys resident on the disc. DVD-Video technology licensees were supposed to encrypt their keys, but one licensee didn't. A Norwegian group called MoRE (Masters of Reverse Engineering), founded by teenagers, reverse-engineered Xing Technologies's DVD-Video player, unlocking its key. Then MoRE was able to work out more than 170 keys by trial and error before giving up out of boredom. Although MoRE has been hit with several lawsuits, no one imagines that the decryption key can once again be made secret. Even if SDMI avoids such obvious missteps, in today's Wild West Internet environment, any scheme of this sort is seen as a provocation to cocky hackers.

The unrestrained availability of MP3 files certainly challenges the status quo. Supporters of Napster, Gnutella, Freenet and their brethren envision an end to the era in which a few large companies dominate music sales. They foresee a democratization of music in which small labels are com-

DENNIS COOK *AP Photo*

Musicians Roger McGuinn of the Byrds (*left*) and Lars Ulrich of Metallica (*center*) express distaste to Hank Barry about his song-swapping Web site, Napster, at a Senate Judiciary Committee hearing held in July concerning copyright violation.

petitive in a secure, downloaded virtual market. Thousands of independent artists could sell their music directly to consumers with no need for record labels for distribution, making a much wider range of recordings commercially viable. Such a renaissance might lower costs for consumers and increase income for artists and their boutique, often self-run, labels.

Another transmission option, real-time streaming, has been less vexing. Data-reduction algorithms reduce file size and bit rate sufficiently to allow the music to be played as fast as it is received. This is how Internet radio stations and Webcasts operate. Listeners log on and download the continually broadcast file into a buffer player, available free online from companies such as RealNetworks. In the best cases, music plays out one end of the buffer as it streams into the other end, although fits and starts are still common today. The low bit rate required for streaming results in low fidelity, so even though these files *could* be recorded onto an individual's hard drive, streaming is not seen as a threat to the record labels. Indeed, streaming has become a valuable tool for labels and independent artists to preview their work to customers.

The music challenges foreshadow those looming even larger for d-movies, d-TV, indeed all of d-entertainment.

Needed: New Business Models

Although MP3 is stealing the headlines, the recording industry is also quietly upgrading its traditional media. The forthcoming DVD-Audio format hopes to piggyback on the spectacular success of DVD-Video. DVD-Audio eschews data reduction in favor of no-compromise, high-fidelity coding, as well as surround sound. A DVD-Audio disc might be coded at a sampling frequency of 192 kHz and a word length of 24 bits, far exceeding the performance of a CD. Whether the average listener can appreciate the improved sound quality—or will pay for it—remains to be seen. New generations of universal players play both DVD-Video and DVD-Audio discs.

A similar super-audio compact disc (SACD) format, introduced by Philips and Sony, also seeks to provide higher performance and surround sound. But its lack of backers and incompatibility with the DVD juggernaut will probably relegate it to a small niche.

DVD-Audio is a technology upgrade made in a tradition of evolution that is comfortable for the hardware and recording companies. But it seems clear that public demand dictates that the industry embrace online d-entertainment. As we advance to broadband and wireless Internet delivery, music's accessibility will only increase, forcing the industry to explore new business models. The choices will probably include purchase, pay-per-listen and monthly subscription. One scenario, proposed by Magex, the digital commerce subsidiary of National Westminster Bank, includes a onetime fee (a "micropayment" of a few cents) for a single play, a larger fee for a set period (say, a 10-hour unit of play) and an even greater fee for unlimited use.

Or perhaps access to d-music will be free, and, as in broadcast television, revenues will come from advertising or corporate sponsorship. Companies might clamor for the glamour. Just as the Hollywood film studios fought videotape and now profit enormously from it, it is conceivable that a shrewd adaptation to new technology might allow for even greater profit.

The distribution technology that will shape our audio future is already in hand. Direct satellite broadcast to cars with small roof-mounted antennas will begin in 2001. The merging of cellular telephones and the Internet, yielding wireless Web access, is profoundly changing telecommunications. And downloading of d-music files is essentially an unstoppable force. The only pitfall to prosperity is the possible lack of cooperation among the many manufacturers and media companies to deliver the future's promise.

If they can deliver, we can envision a day when, as we sleep, our automated agents search Internet catalogues to find music that we might enjoy. As we drive to work, that music seamlessly accompanies us, unless we tune into one of 200 music channels beamed down by satellite. At work, we log on to free Webcast radio stations, streaming music or music videos. In the evening we settle into our home theater and bask in superb DVD-quality sound and pristine HDTV video. Then, late at night, we use a Napster-like program to listen to a live bootleg feed from a rock concert in Tokyo. We revel in that guilty pleasure and resolve that we'll make reparations by paying to download the concert when it's officially posted tomorrow. SA

COURTESY OF THE AUTHOR

The Author

KEN C. POHLMANN is director of the music engineering technology program at the University of Miami and author of *Principles of Digital Audio,* 4th edition (McGraw-Hill, 2000). He is a contributing technical editor to *Sound & Vision* and *Mobile Entertainment*. Freelance writer Wes Phillips contributed to this article.

Further Information

THE MP3 AND INTERNET AUDIO HANDBOOK. Bruce Fries and Martin Fries (contributor). TeamCom, 2000.
DVD DEMYSTIFIED. Jim Taylor. Second edition. McGraw-Hill, 2000 (available October 30).
www.mpeg.org; www.dvdforum.org; www.sdmi.org

Moviemaking in Transition

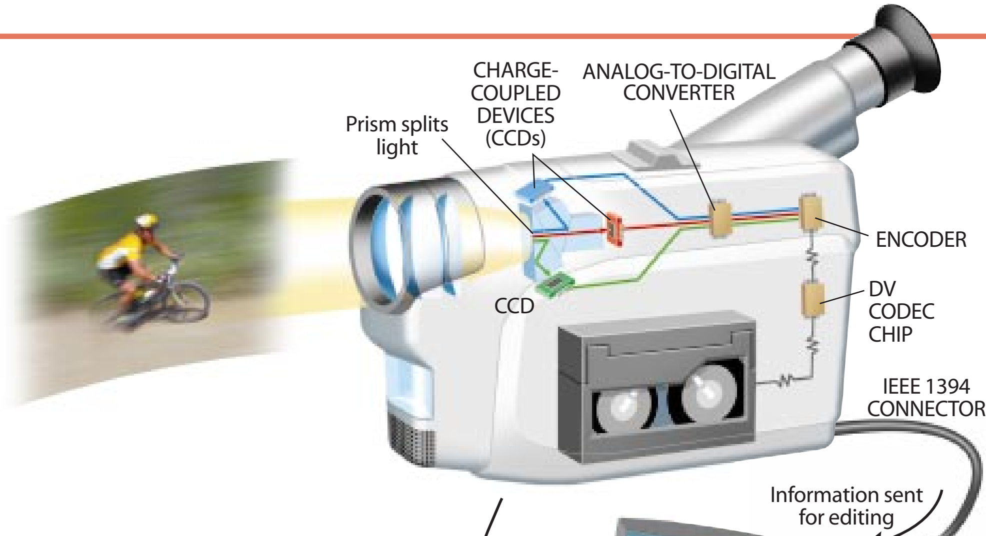
NEXT WAVE FILMS

Digital video cameras and editing equipment are transforming the way movies are made—and even *which* movies get made by Peter Broderick

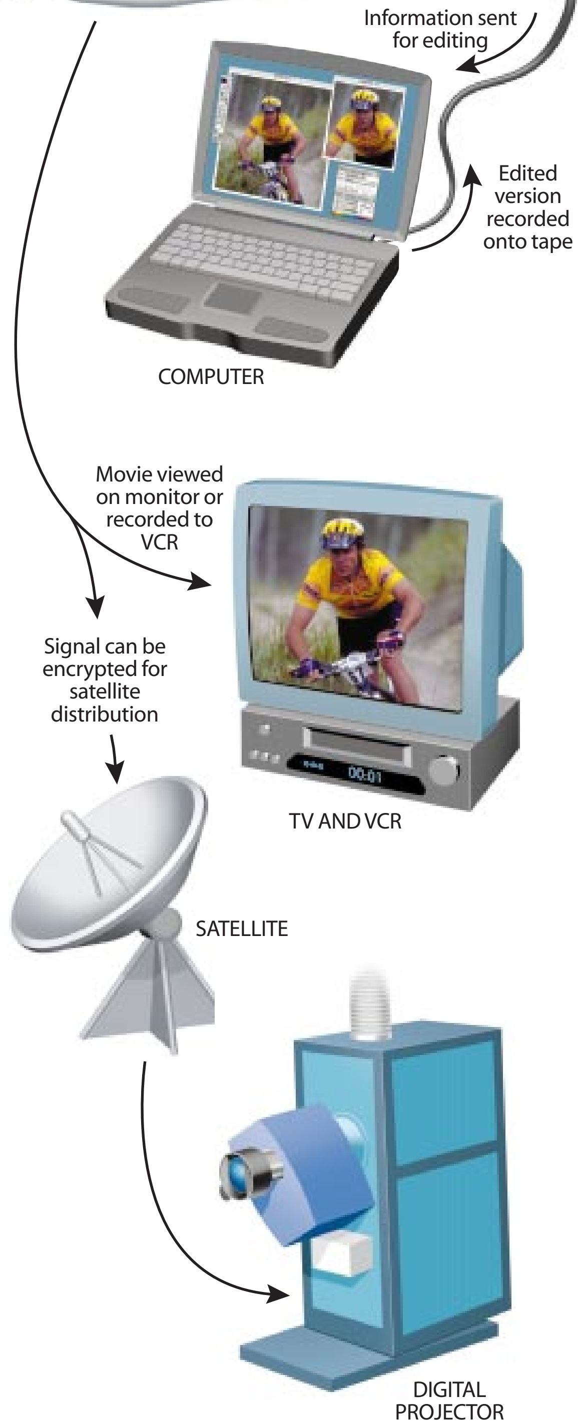
The digital revolution in moviemaking is well under way. New digital tools—from cameras to editing software—are changing not only how movies are made but also which movies are made and who makes them. The technology is in place. The already impressive quality of reasonably priced digital video cameras will continue to improve, as will the power of desktop editing tools. The major hurdle that remains is a traditional distribution system that is unable to handle the new wave of digital moviemaking.

Increasingly, filmmakers shoot documentary and fiction features with digital video cameras, such as that used for *Paper Chasers* (*above*). These cameras improve access to unpredictable situations.

Digital Video Camera



Digital video cameras capture images on silicon chips rather than on the 16- or 35-millimeter film that spools through a traditional film camera (*opposite page*). The data go from the chip to a videotape in the camera. The moviemaker can connect the camera to a computer with an IEEE 1394 cable, move the data to the computer's hard disk and then edit the images without any degradation in quality. For now, the finished movie is usually transferred to film for showing to audiences on a traditional theater projector. But the data will often be transmitted to cinemas via satellite or fiber-optic cables when digital projectors become widespread.

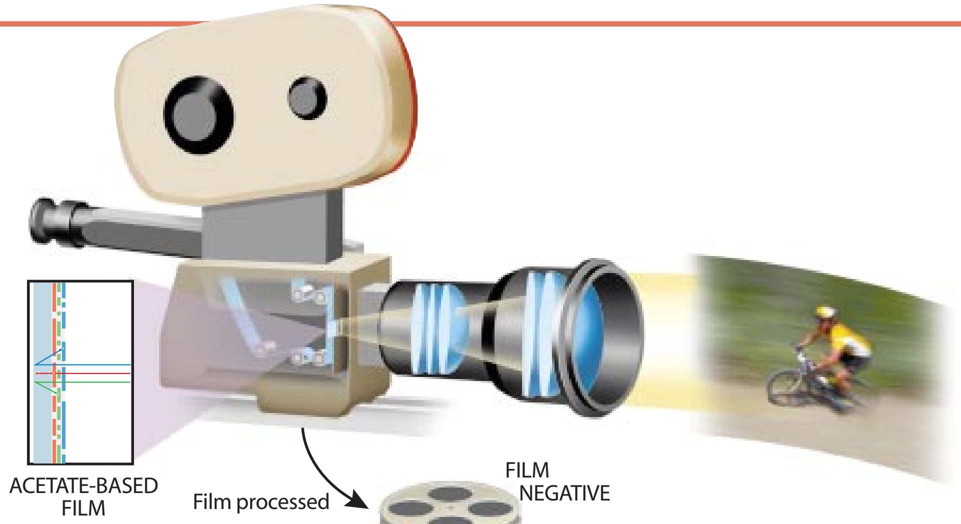


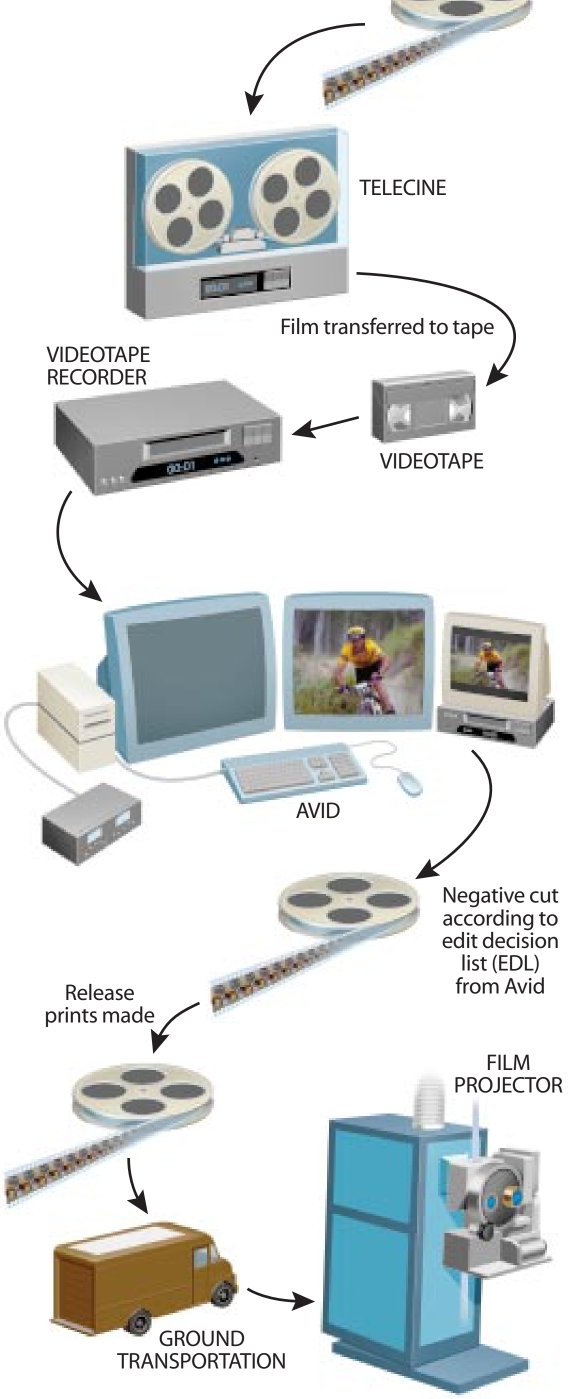
So far the digital revolution has had the greatest impact on the independent production sector. Operating without the infrastructure and inertia of the major Hollywood studios, independent producers have the flexibility to quickly embrace new opportunities and are highly motivated to find ways to reduce production costs.

I have been a part of this development through Next Wave Films, a company of the Independent Film Channel that provides finishing funds and other support for independent filmmakers, whether they are shooting digitally or on film. Between 1998 and 1999 we saw the percentage of digital submissions double, rising to over 34 percent of all requests for finishing funds. During the first six months of 2000, 51 percent of finishing fund submissions were shot on video, and 66 percent of the features added to our database originated on video.

It's not surprising that the earliest adopters of digital production technologies are the filmmakers with the most limited financial resources. No development has lowered the cost of moviemaking as dramatically as digital cameras and postproduction software for editing and special effects. Independent moviemakers can make features they could never afford to shoot on film.

Moviemaking was previously one of the most expensive art forms. Unlike a poet or a painter, a filmmaker needed substantial financial resources and expensive equipment. Now, for the first time, independent filmmakers can afford to own the means of both production and postproduction. When aspiring filmmakers ask how much money they need to make a digital movie, we can now





Film Camera

Traditional movie cameras record 24 images per second on acetate-based film that has three layers of silver halide emulsion sensitized to red, green or blue. After the film is exposed and processed, a telecine machine transfers the negative to videotape. Next, the information (images or sound) on this videotape is input into a computer-based, nonlinear editing system, usually an Avid. The movie is edited digitally on this system, which produces an edit decision list (EDL) indicating precisely which pieces of the original negative will be used and in which order. The negative is cut accordingly. Then release prints are made from the final negative and distributed to movie theaters.

GEORGE RETSECK

say, "Whatever you have is probably enough." Independents with digital cameras and desktop computers have the freedom of novelists to be spontaneous, to improvise, to strike out in new directions and to start over.

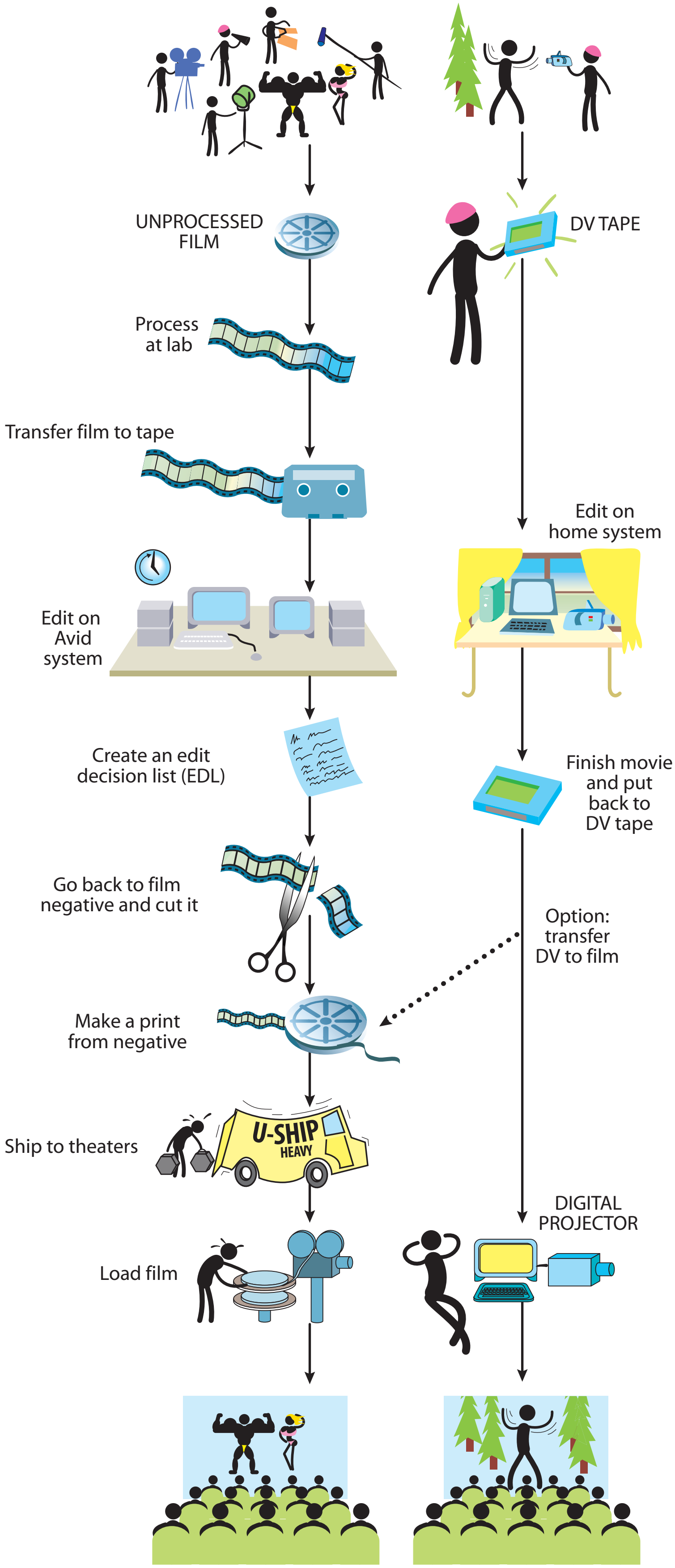
The spectrum of budgets for digital movies is very wide. The upcoming *Star Wars* prequels are being shot with high-definition cameras and are likely to cost more than $100 million. Lars von Trier's latest digital feature, *Dancer in the Dark*, cost about $13 million. Other established directors have made digital features in the $2- to $8-million range, including Mike Figgis (*Time Code*), Spike Lee (*Bamboozled*) and Wayne Wang (*Center of the World*). Many novice filmmakers have directed first features for less than $10,000. Some have even been made for under $1,000. Shot with a consumer digital video camera on a $900 budget, the thriller *The Last Broadcast* is in home video and television distribution in the U.S. and abroad.

In general, the lower the budget, the greater the financial benefit of shooting a feature digitally. The cost of film stock and processing makes up a higher percentage of the budget of less expensive movies.

A New Model for Financing

Digital equipment has transformed the economics of low-budget movie production [*see illustration on page 68*]. In the traditional model, a writer creates a script, and then the search begins for third parties willing to provide the money to make it into a movie. After two or three years

SHOOTING FILM VS. SHOOTING DIGITAL VIDEO



BRYAN CHRISTIE

the supplicants usually give up if they haven't raised the money. If they do find financing, they will probably have to trade substantial creative control for it. This often involves giving the financing entity—whether a Hollywood studio or an independent company—final approval of the script, the stars and, ultimately, the film itself.

In the new model of affordable digital moviemaking, filmmakers conduct a no-nonsense resource assessment before writing the script: they determine how much money they have access to, what equipment they can use, and which cast and crew will join the team. The filmmaker then writes a script that can be made into a movie within the framework of available resources. This approach allows filmmakers to devote their time to making movies rather than raising money.

The concept of making films with available resources was central to the ultralow-budget feature movement that began in the early 1990s. Shot on film before the advent of digital video, influential microbudget features maximized the use of available resources: *El Mariachi* (Robert Rodriguez had free access to a dog and a school bus, both of which figure prominently in the film), *Clerks* (Kevin Smith set his film, and shot it, in the convenience store where he worked) and *The Brothers McMullen* (Edward Burns shot his movie primarily in his parents' house, where his mother cooked for the cast and crew). These features inspired filmmakers across the U.S. and overseas to make films for under $100,000 rather than spend years trying to raise millions. The arrival of digital tools accelerated the decline of the budgets of first features by dramatically reducing the amount of cash required. These new tools also provided filmmakers with more creative possibilities.

New Creative Options on the Digital Set

On a film set, the camera is rolling only a small percentage of the time because of the expense of stock and processing and the amount of time required to light and set up each shot. On a digital set, the camera is recording a much greater percentage of the time. Directors often use two cameras, something that is unaffordable on most conventional film shoots. And because digital video production often necessitates a streamlined approach to crew and equipment, the resulting aesthetic choices frequently make lighting simpler and less time-consuming.

On some very low-budget features shot on film, the ratio of footage shot to footage used in the final cut is as low as 3:1. On comparable digital features, the shooting ratio could be as high as 50:1. This lets filmmakers work with actors in ways that would be impossible on film. Directors can shoot rehearsals, capturing inspired moments that would otherwise have been lost. Instead of filming only a few takes of a scene, a director can shoot as many takes as he or she needs to achieve the highest level of performance. On a digital set, the actors usually do not stand around for hours between scenes waiting for the lights to be rigged. Experimentation is now affordable throughout the production process. Digital playback, for example, makes it possible to view what has been shot on the set immediately (rather than waiting one or two days to see film dailies). The director can try many variations on a scene and use video playback to see what is worth pursuing.

Digital cameras also allow filmmakers to take advantage of the real world. Using an inconspicuous digital camera and a small crew, a director can shoot a fictional story in a nonfiction environment. Michael Rehfield, the director and star of *Big Monday,* made his feature on the streets and in the subways of New York. Unlike studio films shot on location in Manhattan using carefully choreographed extras in diligently policed cocoons, *Big Monday* incorporated actual street life in the frame, giving it the authenticity of a documentary.

Director Paul Wagner used a small digital camera to shoot key scenes in Tibet for his feature *Windhorse,* which is extremely critical of the Chinese occupation of that country. The authorities would never have permitted such a movie to be shot there, but they mistook the filmmakers for tourists making a home video.

Digital tools have enabled directors to transform the production process. Instead of having to shoot a movie during a single period, digital moviemakers can shoot and edit, write new scenes, and keep shooting. For the first time, independent filmmakers have "affordable time." This allows the movie to evolve in an organic way—the director can discard the worst material and build on the best. If the project is not working, it can be abandoned at any point and a new production initiated. If directors own the digital camera and are working with a small crew and dedicated actors, they can take advantage of unpredictable factors the way documentary filmmakers do. Last fall I asked a director when he was going to finish shooting his digital feature. He replied, "Tonight, if it rains."

In addition, digital production can eliminate the high cost of creating optical effects on film. The power to create both spectacular and subtle effects on desktop computers is increasing by leaps and bounds. Next Wave Films received an ultralow-budget science-fiction film with amazing desktop effects that would probably have cost at least $1 or $2 million if it had been made with a special-effects house.

Digital filmmakers can also avoid or postpone another major cost of traditional film production—making a print. They can project their movie digitally until they find a distributor willing to finance the transfer to film for distribution in movie theaters. If theatrical distribution is not possible, they would not need to pay for a film transfer, because the video master would already be sufficient for cable and broadcast television, as well as for satellite, home video and Internet distribution.

Independent Enthusiasm, Studio Caution

Although digital moviemaking has been rapidly gaining momentum in the independent sector, change at the studio level has been much slower. This is not surprising, given that studios have an institutional investment in a production process that has been the standard for decades.

Director George Lucas has done more to make the studios seriously consider using digital tools than anyone else in the filmmaking community. His special effects company, Industrial Light & Magic, has pioneered many breakthroughs and continues to push the limits of movie magic. In the spring of 1999 Lucas sped the arrival of digital projection by insisting that *The Phantom Menace* be projected electronically in several theaters [see "Digital Cinema Is for Reel," on page 70].

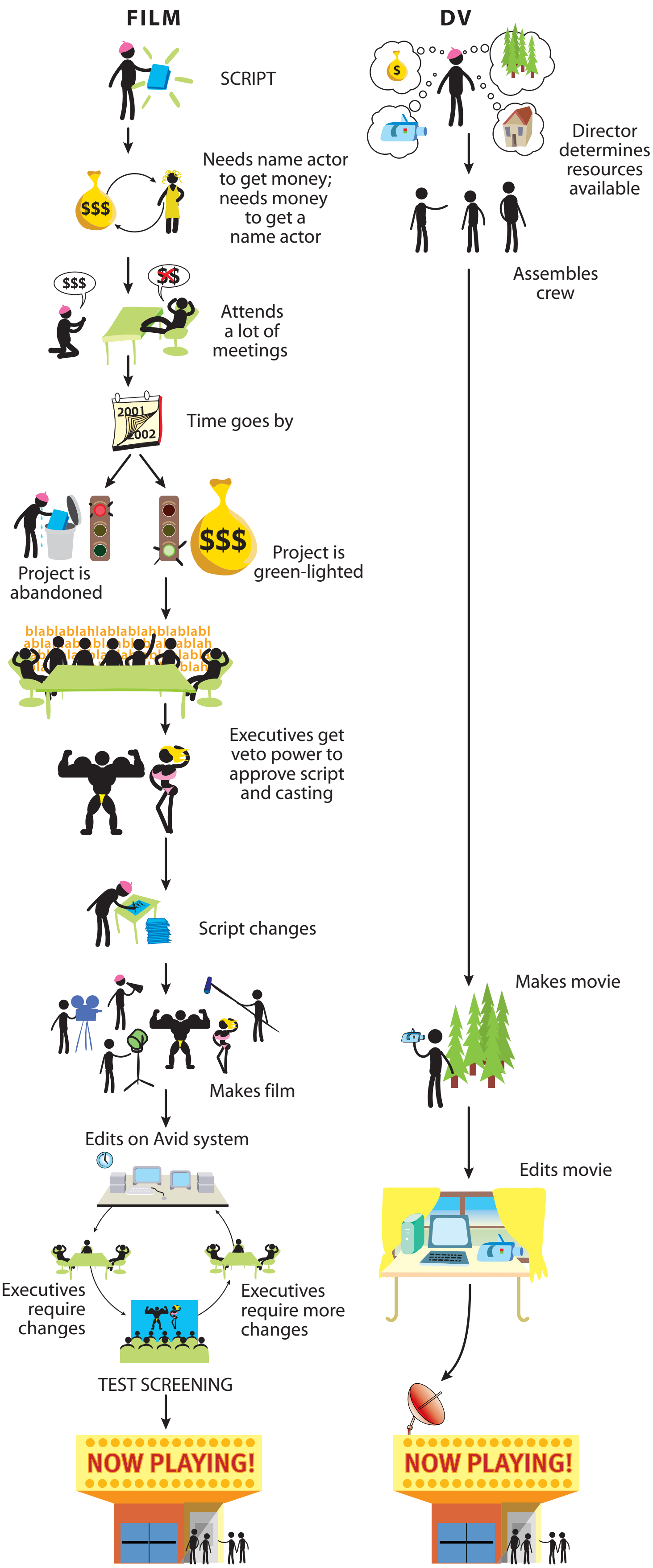
Using an inconspicuous digital camera and a small crew, a director can shoot a fictional story in a nonfiction environment.

He has also been a catalyst in speeding the spread of high-budget digital production. By announcing that the second and third installments of the new "Star Wars" trilogy will be shot with high-definition cameras, he signaled his conviction that $100-million features could be made digitally. His commitment spurred Sony and Panavision to jointly develop new high-definition digital cameras. In addition to having roughly twice the resolution of traditional video, these cameras produce 24 progressively scanned images a second. Ideal for capturing mate-

The lightweight digital video camera allows camera operators to shoot practically anywhere at a moment's notice. At the left, Don Cheadle scores in *Manic,* shot this past June and July.

NEXT WAVE FILMS

A FILMMAKER'S PATH TO FINANCING A FEATURE



The Desktop Studio

It is now possible for all of us to try to become desktop Scorseses. Less than $2,500 will buy a desktop studio that can be used to make numerous movies. The components of the most basic, entry-level digital studio include:

DV camera. High-quality cameras start at under $800 for a 1-chip MiniDV or Digital 8 camcorder. (*Upgrade option:* 3-chip Sony PD-150; $3,700)

Desktop computer. The Apple iMacDV Special Edition comes with a monitor, 128 MB of RAM, a 30-GB hard drive (capable of storing more than two hours of DV material), FireWire ports (which allow your digital video and audio to go from your camcorder directly into your computer with no loss of quality) and free iMovie 2 software (a simple-to-use nonlinear editing program): $1,500. (*Upgrade option*: Apple G4/450 MHz with full-featured Final Cut Pro editing software and Adobe After Effects effects software, including two 19-inch monitors; around $7,000)

Cost of Shooting Film vs. Shooting Video

The differences in cost between shooting a movie on film and shooting on video are significant. The following breakdown compares the average cost of shooting and preparing to edit one hour of 35mm film versus one hour of DV.

Shooting 35mm, editing on video

Film stock (5,400 ft of new Kodak 5274 stock, $.576/ft)	$3,110.40
Processing ($.115/ft)	621.00
Prep for telecine	60.00
Telecine ($200/hr., 5:1 ratio, including tape stock)	1,000.00
Tape stock (one hour BetaSP & 3/4")	73.20
Total cost	$4,864.60

Shooting MiniDV, editing on video

Tape stock (camera original)	$10.00
Tape stock (backup tape)	10.00
Total cost	$20.00

rial designed to be transferred to film at 24 frames per second, these cameras avoid many of the artifacts that can bedevil video-to-film transfers.

Time Code was the first live-action digital feature to be funded and released by a major studio. Its solid visual quality and the favorable critical reception may open the minds of more studio executives to digital production.

But it is much too early to declare film dead. Hundreds of millions of moviegoers around the world have been brought up on celluloid. The recent success of films shot on Imax demonstrates the

BRYAN CHRISTIE

important role celluloid image quality can play. Hollywood does not want to jeopardize its relationship with its global audience. Even if home video is more lucrative, theatrical distribution is the engine that drives the train of ancillary revenues.

The major studios will be very cautious in their use of digital production tools. Until executives are confident that they can shoot digitally and ultimately achieve the equivalent of a film look in theaters (whether via 35-millimeter prints or digital projection), they will move slowly. Some features will be hybrids, mixing film and digital video. Others will use digital video to achieve distinctive looks that do not emulate film in any way.

Can Digital Tools Transform Distribution?

As more filmmakers make digital movies, there will be a growing supply of exceptional features that cannot find theatrical distribution because screens in conventional movie theaters will continue to be devoted to Hollywood product. The number of independent features made annually in the U.S. could easily double in the next few years. This could greatly exacerbate the current crisis in independent feature distribution, doubling the odds against finding theatrical distribution if no more theater screens become available to independents.

Independents now have an unprecedented opportunity to develop new routes to reach audiences instead of relying on a distribution system that serves them poorly and the studios well. This will not be easy, because the well-funded gatekeepers who control access to traditional distribution networks will fight to retain their market dominance. But independents have several things going for them. Public awareness of independent filmmaking has increased, thanks to greater press coverage, high-profile film festivals like Sundance and Toronto, and the expanding reach of cable channels devoted to independent film. The supply of excellent digital features will continue to grow.

And then there is the Internet. It will be key in developing new distribution models and enabling independents to build more direct connections to audiences not satisfied with a steady diet of Hollywood fare. And they will be able for the first time to aggregate audiences across national boundaries. Previously, if distributors in a particular country did not believe a large enough audience existed for a film, no one in that territory had a chance to see it (outside of film festivals).

Anyone with Internet access can now purchase a film online and have it delivered through the mail on videocassette or DVD. Independents from Kevin Smith (*Clerks*, *Chasing Amy*) to the team that made *The Blair Witch Project* have already demonstrated the power of the Internet to build awareness of movies playing in theaters. These techniques will be modified to market digital movies not released theatrically to potential viewers across the country and around the world. Eventually, once bandwidth has increased substantially, films will be delivered online. Viewers with high-speed World Wide Web access and connections linking their computers to their televisions will be able to watch these movies at home with reasonably good image and audio quality.

If independents don't create new distribution mechanisms, they could be marginalized.

Independents will seek to reinvent distribution by using the Web in tandem with other digital routes. Digital projection systems, which have already created opportunities at film festivals, will enable the development of a network of microcinemas. More portable and affordable digital projectors will be used in cafés, museums, community centers and on campuses to show movies on a daily, weekly or monthly basis. Microcinemas can operate with much lower overhead than movie theaters and can cater to regular audiences seeking the best new movies from outside Hollywood. Filmmakers also hope to make greater inroads via digital broadcast satellite and cable.

If independents don't create new distribution mechanisms, they could be marginalized. Although it will be easier for them to make features than in analog times, it will be harder to get those features seen in theaters. But if they can fashion new distribution routes, the diversity and quality of movies available to audiences will soar. The digital revolution in production could usher in a digital renaissance in distribution. SA

COURTESY OF THE AUTHOR

The Author

PETER BRODERICK is president of Next Wave Films, which finances digital features and provides finishing funds to independent filmmakers. He has taught courses on independent production at the University of California, Los Angeles, and given presentations on digital moviemaking at Cannes, Sundance, Toronto and other film festivals. Mark Stolaroff and Tara Veneruso collaborated closely with Broderick on this article.

Further Information

BOOKS

THE IFILM DV FILMMAKER'S HANDBOOK. Maxie Collier. Son #1 Media, 1998. Available at www.dvfilmmaker.com

DIGITAL MOVIEMAKING. Scott Billups. Michael Wiese Productions, 2000. Available at www.mwp.com/pages/booksdigitalmovie.html

MAGAZINES

DV. On the Web at www.dv.com

RES. On the Web at www.resmag.com

Videography. On the Web at www.videography.com

WEB SITES

www.2-pop.com

www.nextwavefilms.com

Digital Cinema Is for Reel

Digital projection works, but it's not at a theater near you—yet

by Peter D. Lubell

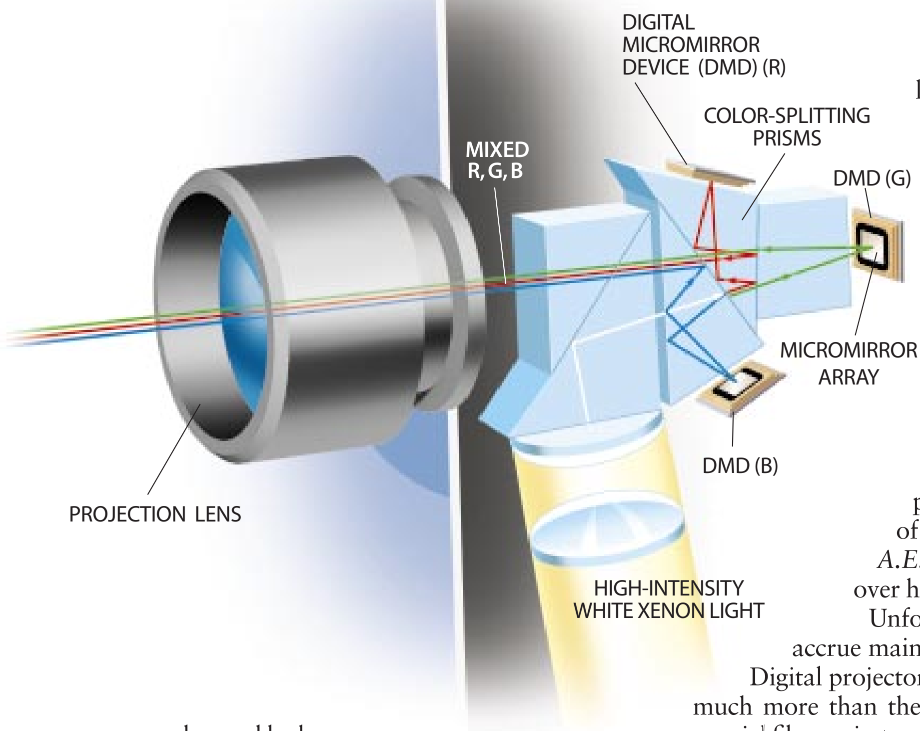
Since his first *Star Wars* film 23 years ago, George Lucas has been a leader in applying technology to the cinema. His most recent movie, *Episode I: The Phantom Menace,* contains almost 2,000 digital-effects shots. Yet Lucas took the digitization of *The Phantom Menace* a step further. During its premiere in the summer of 1999, a few showings were digitally projected. Audiences were amazed at the outstanding audio and the clarity and brightness of the pictures. The d-projectors performed well, but the technology must come down in price before its improved audio and visual presentation reaches a mainstream audience.

For traditionally projected movies, multiple negative copies are made from a positive master print of the edited film. The negatives produce the thousands of positive prints that are shipped to theaters for projection.

Digital cinema also starts with a master print. Each frame of a film is scanned and converted into digital format, and the final edited product is an extremely large data file, about 1,000 gigabytes on average. (If the movie is filmed digitally, the scanning step is unnecessary; see "Moviemaking in Transition," on page 61.) Ideally, the file, once compressed, would be transported electronically to theaters via broadband distribution—either fiber-optic cables or satellite transmission. In the 1999 demonstrations, hard drives were physically delivered to the test theater locations. Later demonstrations of lower-resolution Disney films used digital videodiscs (DVDs) as the portable storage medium.

The 1999 tests used competing projection technologies from Hughes-JVC Technology and the DLP division of Texas Instruments. The DLP system, which had fewer technical problems, now appears to be the front-runner. Disney chose DLP's technology for the digital release of several movies (including *Tarzan, Toy Story 2* and *Mission to Mars*) at 12 North American locations and many others overseas at the end of 1999.

A d-projector uses the trichromatic red-green-blue (RGB) system, first proposed by Thomas Young in 1801, that is the basis for both color photography and television. All color im-



Digital projectors from Texas Instruments feed light to prisms that direct red (R), green (G) and blue (B) wavelengths to the appropriate digital micromirror device (DMD). Meanwhile pixelated information from the movie file is sent frame by frame to the DMDs. A DMD fits in the palm of your hand, and its tiny mirrors, each 16 square microns, turn independently on micromechanical hinges, according to their digital instructions. The reflected components intermingle before reaching the projection lens.

GEORGE RETSECK

ages, as observed by human eyes, are composed of elements of these colors.

The DLP projector employs three digital micromirror devices (DMDs), one each for red, green and blue. Prisms split high-energy white light, obtained from a high-intensity (5,000-watt) white xenon source, into the three colors and direct each to the appropriate DMD. Each DMD is a 1,280-by-1,024 array of digitally controlled mirrors (3,932,160 mirrors in all!). Each mirror reflects varying color intensities as obtained from the data file to project its component of an image.

A d-projector fits within the same envelope as a standard 35-millimeter film projector and can provide the same resolution. The 1999 d-projectors reportedly produced resolutions approaching 2,000 horizontal lines, with contrast ratios (white to black) of 1,000:1, matching contemporary 35-millimeter color prints. These numbers considerably exceed the U.S. performance parameters of high-definition TV. D-cinema has a long way to go, however, to meet the standards of the 70-millimeter prints used for big-budget blockbusters such as *The Phantom Menace* (typically 3,000 lines and ratios of 1,200:1).

But d-cinema holds financial promise as well. Currently so-called saturation distribution throughout the U.S. requires up to 5,000 prints at $2,000 each, for a total of $10 million—not including shipping costs and the logistics of timely delivery. Prints are usually good for only 30 showings before replacement. In contrast, the 19 hard drives, each holding 18 gigabytes, used to store *The Phantom Menace* master print cost about $260,000 and will last indefinitely—and copies can be made cheaply and immediately, given enough available memory.

Digital distribution via satellite or fiber-optic lines would offer shipping advantages, including simultaneous delivery to the theaters. Cisco Systems and 20th Century Fox demonstrated electronic transfer via the Internet this summer at Supercomm, a communications-oriented computer conference held in Atlanta. Downloading the compressed 42-gigabyte file of the animated film *Titan A.E.* took about two hours over high-speed lines.

Unfortunately, these benefits accrue mainly to the movie studios. Digital projectors cost $250,000 each—much more than the $50,000 for a current commercial film projector. Theater operators would be burdened with substantial conversion costs. Even if d-projectors sold for $100,000, retrofitting a 10-screen multiplex would cost from $1 million to $1.5 million, excluding audio adjustments. The studios (and manufacturers) have proposed several economic models for cost sharing, such as no-cost equipment loans and dual-tier (film vs. digital) admission pricing. None have been warmly received by theater owners, who see an expensive technology being forced onto their niche market.

Digital projectors will improve. Whether they are used is not a matter of technical performance but of business practice. Theater operators must find digital distribution and projection attractive both artistically and financially. Until they do, d-cinema is still a few years away. Even George Lucas, who is shooting *Episode II* using Sony digital videotape and editing the film digitally, expects to exhibit his final product in the traditional manner—projected celluloid film. SA

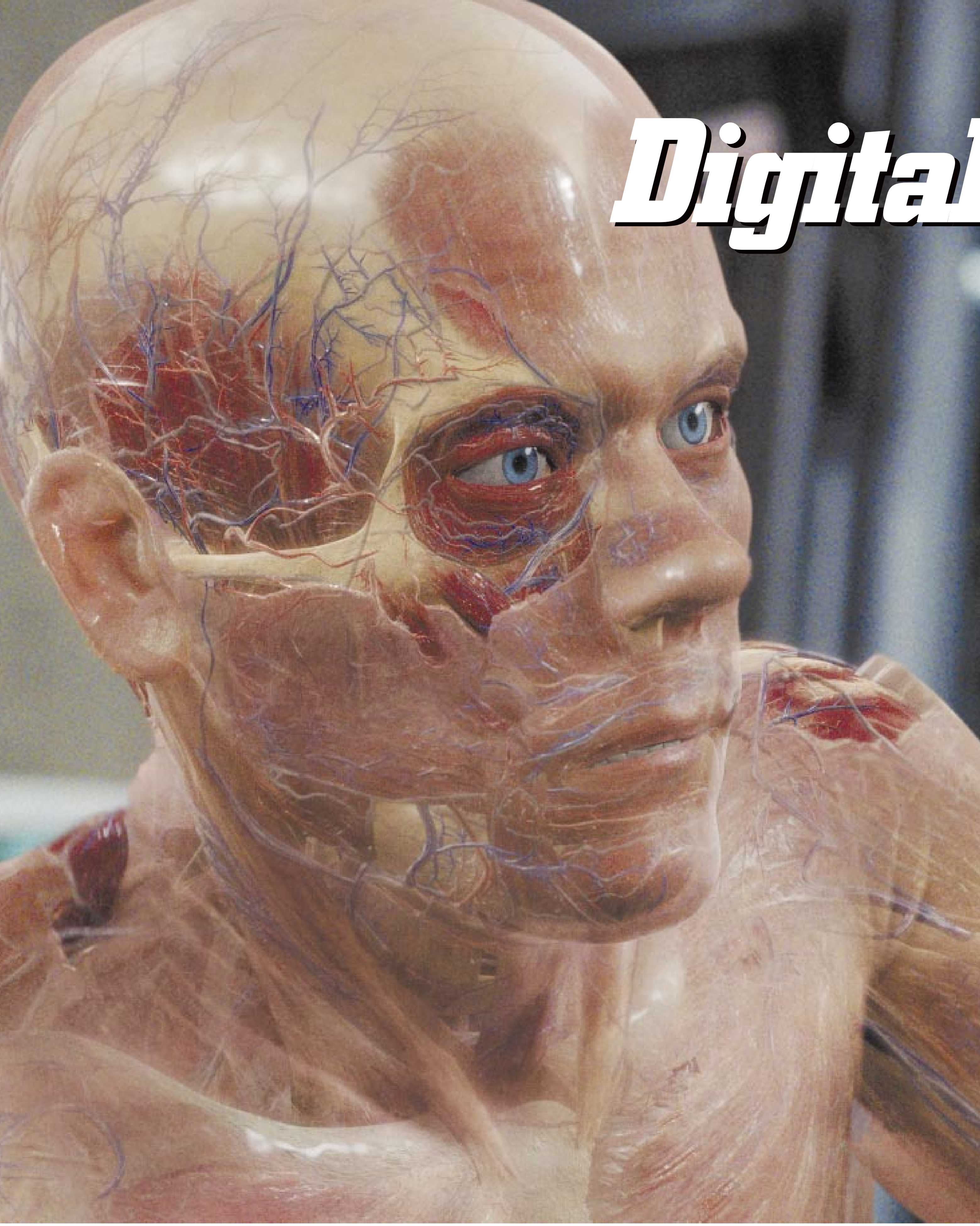
SHERRI A. LIBERMAN

The Author

PETER D. LUBELL is a telecommunications consultant at Communications Strategies in Albertson, N.Y., and adjunct professor at the Polytechnic University of New York. He has also worked for Home Box Office and Rainbow Media.

Further Information

MAKING DIGITAL CINEMA ACTUALLY HAPPEN—WHAT IT TAKES AND WHO'S GOING TO DO IT. Steven A. Morley in *140th Annual SMPTE Technical Conference Proceedings,* Pasadena, Calif., pages 448–462; October 1998.


Digital

Humans Wait in the Wings

Characters, scenes and entire movies have been crafted digitally. But can animators create realistic humans to star in computer-generated films? Actors want to know by Alvy Ray Smith

When movie star Richard Dreyfuss presented a technical Academy Award to me and my colleagues several years ago, he looked at us with a wry but wary smile and said, "We're both indispensable to each other. Don't forget that—the people who made *Toy Story*. We're all going into the 21st century *together* … I hope!" Actors in the audience laughed, some of them nervously.

Dreyfuss's barely disguised fear fed on the growing claim by computer scientists that simulated actors like the characters in *Toy Story* will someday replace real ones as they become ever more humanlike. Zealots would replace Dreyfuss with a simulacrum that walks, talks, gestures, thinks *and* emotes just like the man himself—presumably without hiring him at his substantial rates.

Hollywood can rest assured that actors won't be replaced by computer-generating beings in the expanding world of d-entertainment. At least not soon. But completely computer-generated movies are here and are already altering the actor's role.

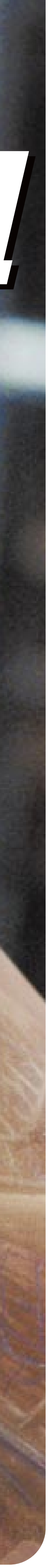
Toy Story, released in 1995, was the first movie to be made entirely by computer. No live-action filming was needed. Three other "cameraless" movies have followed: *Antz, A Bug's Life* and *Toy Story 2.* All were created by Pixar Animation for Disney except *Antz,* made by Pacific Data Images for DreamWorks. Although fascinating, these movies still look like cartoons—wonderful, three-dimensional cartoons that concoct a "real" world, but cartoons just the same. The ultimate challenge is to craft a digital, believable Richard Dreyfuss or Julia Roberts—a d-Dreyfuss or d-Roberts, if you will.

Supposing that we filmmakers can fashion a convincing human character without an actor requires a leap into the controversial computing religions of artificial intelligence, virtual reality and artificial life—the digital "realization" of everything. We are wiser to distinguish between the art of acting and the representation of actors. For the foreseeable future, actors will be needed to infuse characters—whether played by themselves or their computer avatars—with credible actions, expressions, voices and thoughts.

The worlds they act in, however, are becoming increasingly digital, with real-world believability improving fast. The movie *Titanic* used three-dimensional computer graphics to tell a story that could not otherwise be told on the screen. Computers allowed the filmmakers to show scores of people falling off a huge, capsizing *Titanic* in ways that would be too dangerous, contrived or costly to re-create on a real ocean, in a water-tank studio, or with models in miniature. In this year's ocean disaster, *The Perfect Storm,* filmmakers at Industrial Light & Magic engineered ferocious 100-foot storm waves. These couldn't be filmed on real riotous seas or simulated in any other way.

And yet, despite their technical success, whether computer-generated movies succeed at the box office is a different question. The adolescent girls who made *Titanic* so successful weren't flocking to theaters for the graphics. The emotional, rich human characters are what people most love.

Some of today's computer graphics are being



COURTESY OF SONY PICTURES IMAGEWORKS

Sebastian Caine, perhaps the most believable digital human yet, was modeled on actor Kevin Bacon by Sony Pictures Imageworks for *Hollow Man* by anchoring virtual muscle to bone; adding eyes, skin and blood vessels; and controlling fine joint movements.

The ferocious seas engulfing actor Mark Wahlberg in *The Perfect Storm* were simulated by Industrial Light & Magic using computational fluid dynamics for waves and turbulence; rule-based particle dynamics for ripples, foam and spray; and fractal-based shaders that add surface textures and even scatter light inside water drops.

done just for the sake of computer graphics. Consumers apparently like what they see, and the filmmaking business will go where the money is. But as the technology matures, we will get back to the fundamentals of good story and character, with computers used in sophisticated ways audiences won't even be aware of.

How far can computer simulation go? Can we create believable humans on screen? Building a completely simulated person requires solving impressive problems, such as how to convey consciousness, emotion (if indeed the two are separable, as University of Iowa neuroscientist Antonio Damasio wonders), physical nuance and acting, just for starters, not to mention the philosophical question of qualia—how a character knows "blueness." We might someday make an artificial Marilyn Monroe (the usually favored example), because she, as I, is a biological machine that is ultimately explainable in physical terms, at least to a much finer degree than is possible today. But achieving this understanding is still a statement of faith. No one can predict whether the virtual actress will ever exist or if she would be cheaper or more versatile than the real item. We might run into some fundamental logical roadblock, as mathematicians did with Gödel's Incompleteness Theorem.

Kevin Bacon Driving Kevin Bacon

Rather than fantasize about a magical Second Coming called "emergence" or preach about world domination by machines, as some of my colleagues have done, I offer what we might reasonably attain in a current lifetime—an integration of the best of both worlds, human and machine, artistic and technical.

In a conventional movie, cameras capture actors on miles of film, which is edited into a movie reel that is projected at 24 frames per second at your local movie theater. In traditional cel animation, seen in classics from *Fantasia* to *Pinocchio,* animators draw characters with pencil on paper every third frame or so. Artists called inbetweeners fill in the missing frames. Each drawing is traced in ink onto a sheet of clear celluloid, or "cel," and "opaquers" fill in the outlines with colorful, opaque paints. Meanwhile an artist paints a background scene, perhaps elaborately if it will be used for many frames. Character cels and a background are stacked in register, and each combination is exposed on a single frame of film in a conventional movie camera. The camera is advanced one frame, and the whole process repeats thousands of times.

Modern, digital cel animation began in 1990 with a Pixar system used by Disney to make *The Rescuers Down Under* and later stormed theaters with Disney blockbusters such as *The Lion King*. In this technique, the pencil drawings and inbetweening are crafted by artists as before, but computers are used to complete all the other steps. Computer-controlled cameras do the filming. Yet the biggest advantage is digitizing the logistics. This unsung aspect of making animated films—keeping track of the hundreds of thousands of elements in various stages of completion—is a true leap forward, saving great time and effort.

Digital cel animation still has a two-dimensional look. Clay animation is an attempt to conjure a 3-D world and has only recently been used in feature-length films, such as this past summer's hit *Chicken Run*. The elements in each frame are modeled in clay, and the animator must painstakingly move a chicken's leg, say, a fraction of an inch for every frame to make the creature appear to walk.

Claymation will never look lifelike, however, because clay models are clay models. 3-D computer graphics animation (sometimes called simply 3-D animation) is an attempt to fabricate a believable world with no drawings, models or live actors on film. In geometric terms, special software models each character, object and background as well as their shading, lighting, movement and motion blur. Wrapping a picture onto a surface fashions complex, textured objects. All characters, sets, locations and visual effects, plus location sound and sound effects, are manufactured digitally, and the editing and mixing of the components proceed digitally, too. For now, dialogue and music are usually recorded by live people, but these, too, could eventually come from brawny computer chips. Frames are stored in a computer's hard drive and are transferred to conventional film once the movie is done, purely so it can be shown on projectors at movie theaters. As theaters change over to digital projection [see "Digital Cinema Is for Reel," on page 70], they will simply download the digital movie via satellite or run a digital videodisc.

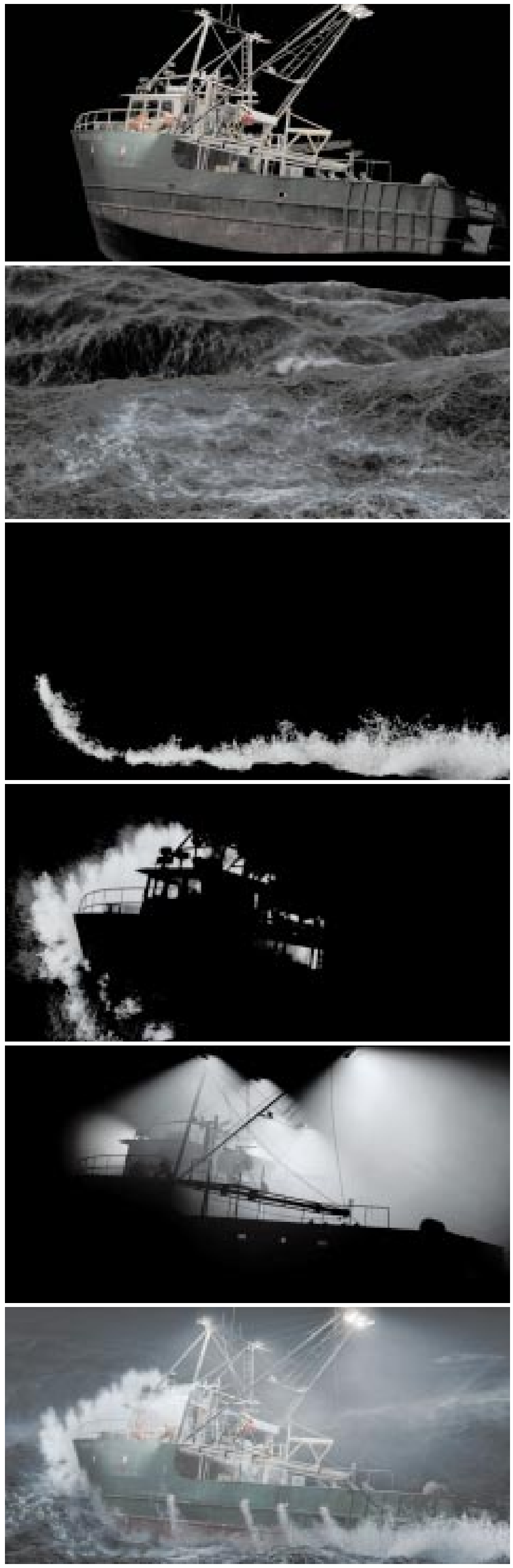
In this computer-generated (cameraless) movie, animators construct a virtual person by building a skeleton, covering it with muscles and skin, and making the character move in nearly lifelike fashion. The most complex attempt at a human yet was seen this past August in *Hollow Man,* which featured a completely digital (but skinless) Kevin Bacon, created by Sony Pictures Imageworks.

Despite the remarkable feat of *Hollow Man,* it may still take another two decades before a plausible Kevin Bacon can be created without filming the real actor as at least a model for his virtual self. To generate some of the movie's effects, Bacon had to wear a green bodysuit so animators could film his movements and then map them onto the computer-generated model. More important, he acted out *Hollow Man*'s character. That hints strongly at the most likely role of actor and animator, of man and machine, in computer-generated movies: real actors driving realistic representations of human beings, perhaps even of themselves.

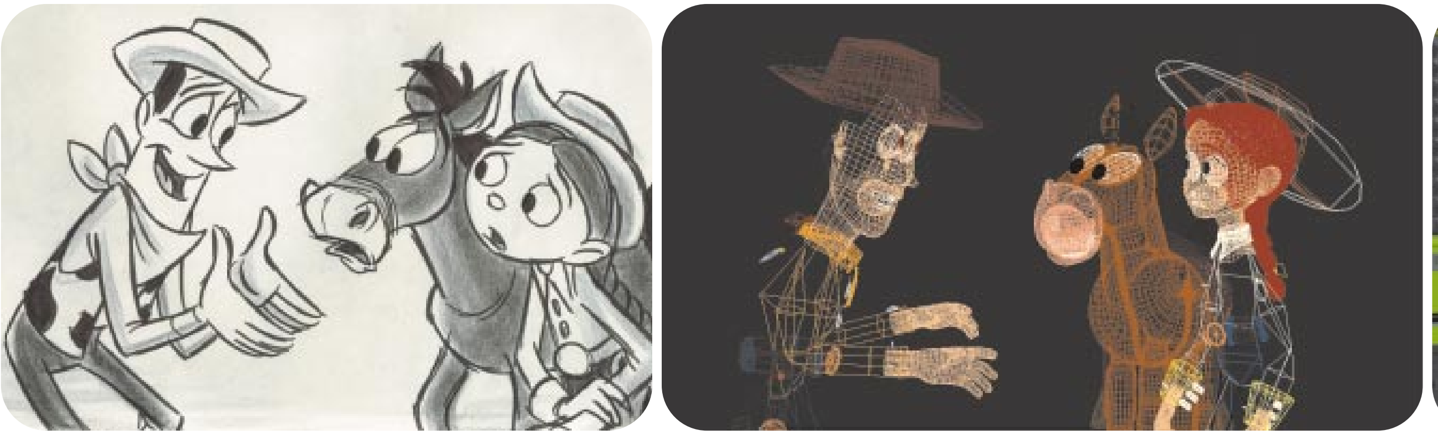
I am sure that the acting abilities of animators and actors can be melded, because we have already done it. Woody, the cowboy in *Toy Story* and *Toy Story 2,* is not just voiced by Tom Hanks. He seems to act like Tom Hanks would in various situations. Hanks would get "into character" when he read Woody's lines: his eyes opened wide at surprise, his shoulders dropped at disappointment. Pixar videotaped Hanks during these sessions, and the animators used his visual cues as inspiration for Woody. To a small degree, Woody became a computer-generated Tom Hanks in chaps and yet still took on his own persona, because the animators infused his character with their own great acting talents.

Here lies a fascinating new avenue for actors. Indeed, when John Lasseter, Pixar's principal artistic director, searches for new animators, he says he doesn't look for drawing ability or knowledge of computer modeling and certainly not for programming experience. He looks for acting skill. In this conception, an animator is just that special kind of talented actor who can make us believe that a collection of colored polygons has heart, gets angry and outfoxes the coyote.

This is not a strange concept. An actor on stage or screen is an animator of his own being. He makes us believe that the body, voice and mind we see are those of an entirely different person—in



A head-on crash (*bottom frame*) was concocted by layering computer-generated images: (*from top*) a virtual *Andrea Gail;* the rendered ocean; a simulated side splash; the bow slam into the oncoming wave; and lighting from rigging lamps.



Scene development in *Toy Story 2* included (*left to right*) the original storyboard sketch, in pencil; a crude computer graphics realization of characters, as wire frames or outlines; a rendering check with colored polygons, typically tiny triangles and quadrilaterals (*not visible*); and the final rendering at movie-screen resolution with full lighting of textures and patterns.

Hanks's case, anyone from *Apollo 13* astronaut Jim Lovell to the maladroit Forrest Gump.

Anyone who has played Doom, Quake, Ultima Online or any other such group online video game has represented himself or herself by a graphical character—an avatar. People readily represent themselves, and many choose an avatar of the opposite sex. Some choose animals or objects. The quality of these impostors suffers greatly because of the bandwidth-starved Internet and microprocessor limitations, but the point is that humans already commonly drive self-manifestations. For the cameraless movie, think of an actor driving a first-rate avatar. Think of Kevin Bacon driving a realistic representation of Kevin Bacon.

Why bother? Because the computer-generated Kevin Bacon can dive from a cliff into the ocean at Acapulco, or off a neuron into his own brain, without getting hurt or shrinking really small.

Reaching 80 Million Polygons

So already there are actors, called animators, driving somewhat credible representations of human beings. What's missing from the fully cameraless movie vision is realism—a significant challenge. Can it be met?

The now famous Moore's Law captures the dynamic of the ongoing digital revolution. Though it is usually applied to the exponential increase in the density of transistors on an integrated circuit, let me express it in an uncommon way: everything good about computers gets 10 times better every five years, or "10x in 5." Circuit designers believe that Moore's Law still has about 10 more years' usefulness—another factor of 100—before a quantum-mechanical wall bars further improvement.

Moore's Law prorated to four years drops the factor 10 to roughly six. Applied to pictures, it seems to tell us that the three million to 17 million polygons used per frame in 1995 to create *Toy Story*–level complexity should have become 18 million to 102 million polygons per frame by the time *Toy Story 2* reached completion in 1999. A polygon is the smallest unit of geometry that is typically manipulated. Digital filmmakers craft every object and character by modeling its surface, gluing together good old Euclidean spheres, cylinders, cones and so on. We often represent a curved surface with a very fine approximation of tiny polygons, so the total number of polygons is used as a good measure of the geometric complexity of an object and hence a scene. When we were at Lucasfilm, my colleagues Loren Carpenter, Ed Catmull and Rob Cook first came up with the notion of 80 million polygons per frame as the reality threshold. The industry has quoted this estimate for years, saying, "Reality begins at 80 million polygons."

Using film's 24 frames per second, the reality threshold is thus 1.9 *billion* polygons per second. The rate of 30 frames per second for television or videotape implies 2.4 billion polygons per second. Today's hottest computers can handle only millions of polygons per second. Generating reality in real time would be marvelous, of course, but the movie business would be happy just to reach reality. To measure how distant real time still is, consider this: *Toy Story* took an average of seven hours of computation for each frame. *Toy Story 2* required several hours per frame, too, and some of its most complex frames took more than 50 hours.

When it was complete, *Toy Story 2* had only doubled its predecessor's complexity, to between four million and 39 million polygons per frame. What *did* increase by a factor of six in those four years was total rendering time per frame, at least for some frames. Pixar's Don Schreiter re-rendered frames from both movies for this article, on the same modern hardware, for a direct comparison. Frames from the newer film took six to 13 times as long to compute. Averages provided by Pixar's Bill Reeves, taken over all frames of both movies, indicate a ratio closer to 5:1.

Even at 80 million polygons per frame, we can't really claim to have replicated reality. True reality might be fractal in complexity: the closer you look, the more an object looks the same—remaining just as complex. But most of us will accept that level of visual complexity as realism. It is a mistake to think that true realism is even our goal. Movies are never real. Dialogue is pieced together, sets feature false fronts, lighting is artificial, and editing plays with time. What "realism" means is a convincing representation of reality. Usage of the representations might be totally surreal. As I have often said, reality is just a convenient measure of complexity.

My colleagues calculated the figure of 80 million polygons by considering what the human eye sees in an image. Assume a scene you are to view is divided into an array of tiny squares, each modeled in the computer by one point of light, called a picture element, or pixel. For reference, the frames in *Toy Story* and *Toy Story 2* had 1.4 million pixels, and *A Bug's Life* had 1.8 million. *Monsters, Inc.* (Pixar's next) will have two million. A pixel represents, with a single color, the average of all the light rays hitting one little square from the projected scene. That bundle of light rays can portray about four levels of surfaces. Each level comprises about eight polygons. Thus, there are 32 polygons for each bundle's pixel, or 32 million polygons for each million pixels of frame resolution. We considered frames with 2.5 million pixels; hence, 80 million polygons per frame.

A New Reality

Our logic has held up empirically—and remarkably, considering the arbitrariness of our argument. But the number of polygons per frame will vary with the final resolution of the images. What is needed is an estimation for achieving maximal reality, from the viewer's perspective, independent of the number of pixels in a frame. According to this perspective (encouraged by my Pixar colleagues Reeves, Catmull, Oren Jacob and Galyn Susman), polygon counting or any geometric measure misses the point. Geometric complexity does not equal realism—it constitutes only an estimated 10 percent of it. So what makes up the rest?

Shape and shade, together, are one big factor. Shape is a 4-D concept that covers the geometry of an object or character plus its motion. The tree-lined street of *Toy Story 2* is much more realistic than the one in *Toy Story* because of subtle movements of the leaves; this contributes more to the believability than does the geometric accuracy of the still objects. Shading colors the objects and includes their illumination as well as their material and textural makeup. Lighting has arisen as a particularly challenging task within shading. Direct computation of the physics is too hard; Mother Nature does it in parallel and real time, but computers cannot.

Richard Dreyfuss should talk with his attorney about protecting his representation. Does he or his avatar command the higher salary?

Overall, the shading portion of the "other 90 percent" of reality should ease greatly when the Moore's Law factor reduces hours of rendering time to seconds. But human viewers add the requirement of accuracy. Look at a brown chair closely; it is not all "brown"—there are blemishes and specks of other colors. Society has a word for an almost-but-not-quite-human: monster. The computer graphics community currently opts for simplifying humans—making them obvious cartoons—rather than risk our perceiving them as monstrous. Or it makes them purposely monstrous, as in *Hollow Man*. Or it goes for animal representation, as with the mouse in *Stuart Little*. Barbara Robertson, a long-time commentator on the computer graphics scene, suggests that the eponymous mouse holds the title for best artificial actor in a live-action film so far, factoring in complexity of representation as well as acting ability. Perhaps the Academy of Motion Pictures Arts and Sciences should designate an award for this new category—for an avatar *and* its actor-animator.

Solving these problems requires new tools that can model intricately detailed representations without becoming hopelessly complex. Physical modeling could help, once the complex physics of form and motion are better translated into manageable computer routines. Instead of an animator having to laboriously animate the action of a pogo stick, for example, he could use a program with a model of how a spring works to compute the frames.

Motion capture will help, too, as it did in *Hollow Man*. Put a human in a bodysuit highlighted with fiduciary marks, film him doing a twist off a diving board, and use the frames to guide the animator. This sounds powerful, but it is not well understood. If you just sample a human's motion point by point, it won't look right. Animators have compensated for this with classic animation tricks of anticipation, squash, stretch and exaggeration. Now we need software that automatically does these things for motion-captured 3-D models.

PIXAR ANIMATION

PIXAR ANIMATION

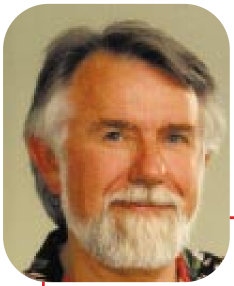
Believability leaped from *Toy Story* (*left*) to *Toy Story 2*. Geometric complexity doubled, and computation time rose by a factor of 10, to improve lighting and texture and to add subtle motions, such as the fluttering of leaves, that greatly enhance the sense of reality.

A technique called image-based rendering will allow us to measure reality itself more accurately. We currently use two different methods to make pictures with computers: geometry (polygons) and sampling (digitizing an object into a set of pixels). The place where these two worlds intersect is image-based rendering. In this approach, a cameraman runs his camera over an object (a vase, say) from all angles. Each frame is digitized. A computer, in a very hairy calculation, figures out by triangulation where each colored pixel "lives" in 3-D space, thus creating a 3-D model of the object or at least of its surface. This advances computer graphics' fundamental imaging element, the 2-D pixel, into a 3-D pixel, often called a voxel.

As the number of controls for a character increases from the hundreds (Woody in *Toy Story*) into the thousands (Al in *Toy Story 2*) to the hundreds of thousands or even millions required for believable humans, we will need powerful tools like these that can be applied with a simple click.

Our society is definitely in the midst of an entertainment revolution. Animation will begin to approach, if not supplant, live action. Is Hollywood ready? Maybe not. When (false) rumors surfaced in August that writer-director Andrew Niccol would cast a computer-generated leading lady opposite Al Pacino for his next film, *Simone*, the Screen Actors Guild (SAG) protested, saying it would cost an actor a job. Although no such character could be lifelike given the current state of technology, the rancor indicates resistance to what will be a fundamental change soon enough. Rather than protesting, SAG should be figuring out how to represent all the creative talent behind a movie, not just the brand-name actors. Animators, by my argument, are actors and should be in SAG.

Actors and avatars are two separate things. Digital filmmakers are replacing representations, not acting. The acting talent will always be integral. Richard Dreyfuss should relax and contemplate a screen body of his that doesn't necessarily age. The cameraless movie needs him. He should start talking with his intellectual-property attorney about protecting his representation. Does he or his avatar command the higher salary? The essential problem becomes interfacing the actor to the realistic model and the actor adapting to that mode of acting. Animator-actors are exploring this area now. It took us 20 years from initial conception to reach the first completely computer-generated movie. Perhaps the next 20 will bring us the completely believable cameraless movie—created together by artists and technicians. SA



LOUIS FABIAN BACHRACH

The Author

ALVY RAY SMITH, cofounder of Pixar Animation, is a digital photographer in Seattle. He has a Ph.D. in computer science, has been director of computer graphics research at Lucasfilm and graphics fellow at Microsoft. Smith has garnered two Academy Awards for technical achievement.

Further Information

THE FEELING OF WHAT HAPPENS: BODY AND EMOTION IN THE MAKING OF CONSCIOUSNESS. Antonio Damasio. Harcourt Brace, 1999.

BUILDING A BETTER MOUSE. Barbara Robertson in *Computer Graphics World*, Vol. 22, No. 12, pages 32–40; December 1999.

VISIBLE DIFFERENCE. Barbara Robertson in *Computer Graphics World*, Vol. 23, No. 7, pages 26–34; July 2000.

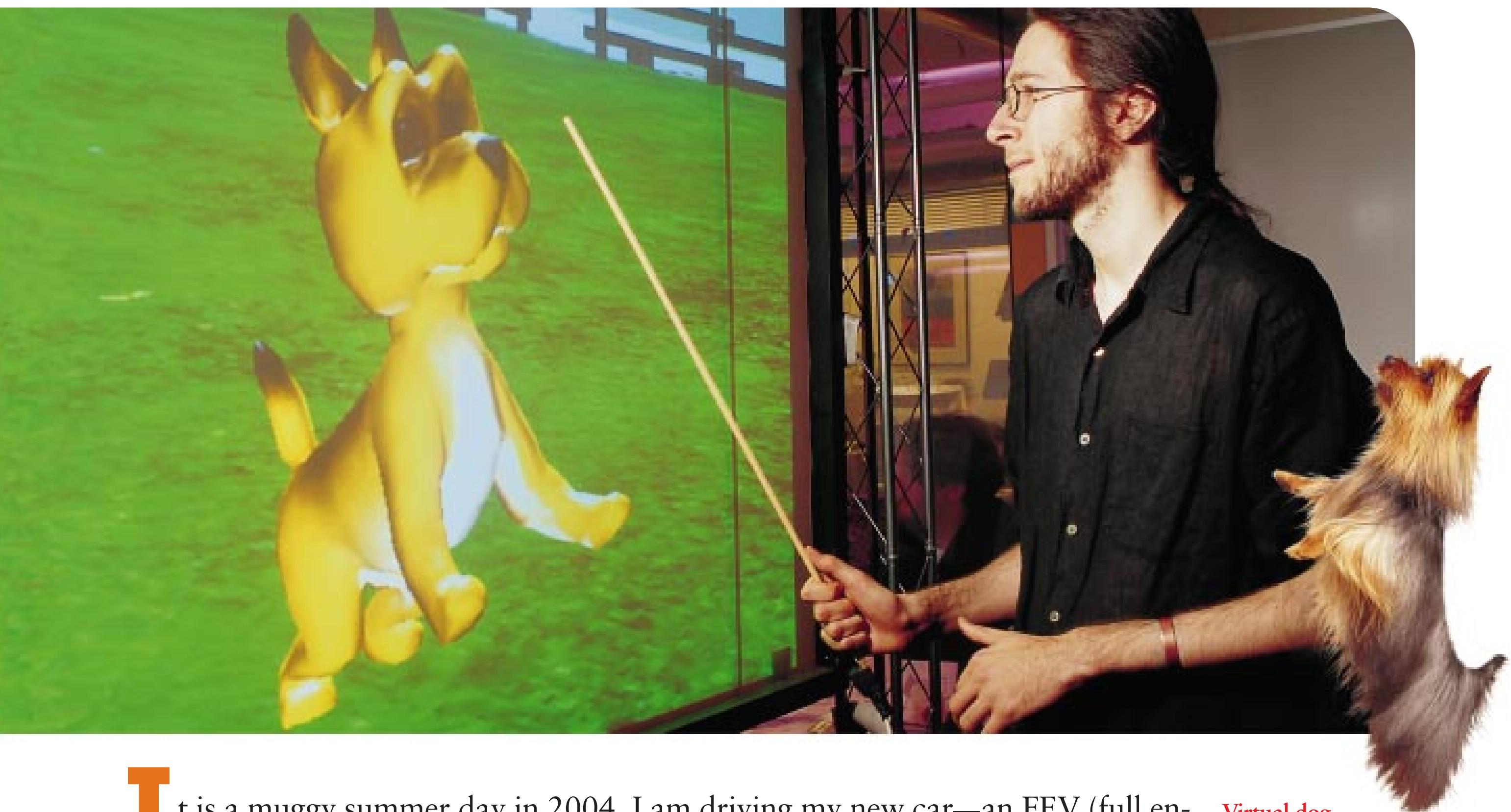
Films notable for their graphics and special effects:

THE RESCUERS DOWN UNDER (Disney, 1990). THE LION KING (Disney, 1994). TOY STORY (Disney, 1995). TITANIC (Fox-Paramount, 1997). A BUG'S LIFE (Disney, 1998). ANTZ (DreamWorks, 1998). STUART LITTLE (Sony, 1999). TOY STORY 2 (Disney, 1999). CHICKEN RUN (DreamWorks, 2000). HOLLOW MAN (Sony, 2000). THE PERFECT STORM (Warner, 2000).

Your Own Virtual Storyworld

True interactive entertainment will arise once engineers and artists create virtual realities that can unfold improvisationally

by Glorianna Davenport



Virtual dog named Duncan learns tricks in much the same way as his real-life counterpart.

WEBB CHAPPELL

It is a muggy summer day in 2004. I am driving my new car—an FEV (full entertainment vehicle)—along the traffic-clogged highway that leads to the kids' summer camp. Glancing into the rearview mirror, I notice a BMW convertible quickly gaining on us. I grumble to myself and pull into the slow lane to let it pass by. For me, driving is the usual tedious task. But for my kids, Jamie and Joy, the trip is an adventure of their own choosing.

From their position in the back seat, the kids see the speeding sports car as a *Tyrannosaurus rex* on the run, legs thrashing and tongue flailing. As the dinosaur roars past the holographic side "window," Joy slaps it with her lollipop, which sticks to the screen. The FEV's onboard computer responds to her assertive gesture by sending a message to the speeder's dashboard display: "Hey, slow down." Suddenly, a police car in hot pursuit appears on the holo-window as a screeching pterodactyl chasing the thunder lizard. The *T. rex* looks over its shoulder, stops and exhales a steamy sigh as the flying reptile perches on its tail. "Got him!" Jamie yells in satisfaction, as lightning and dark clouds invigorate the scene of Jurassic capture.

Does a family car that entertains passengers with

Evolution of Interactive Entertainment

Electronic entertainment has been inviting us to interact since the 1950s, but most forms still lack the capacity for improvisation or serendipity. Merging elements such as "smart" stages and synthetic actors with sophisticated graphics will give rise to virtual storyworlds that go far beyond anything we see today.

1953

Winky Dink and You (CBS)
Early experiment in interactive TV. Viewers place plastic sheet over the monitor, and cartoon characters lead them through drawing exercises.

1972

Pong (Atari)
Passive viewer becomes interactive player who uses a handheld dial to control a graphical Ping-Pong paddle on a TV screen.

1977

Adventure (Will Crowther and Don Woods)
Pioneering example of interactive fiction. Text-based story unfolds on the computer screen as player types commands to initiate the next action in puzzle-solving game.

1983

Dragon's Lair (Don Bluth, Cinematronics)
Laser-disk technology enables the step to the visually stunning illusion of a 3-D world. Player controls a humanlike character rather than an inanimate object.

an interactive fantasy world seem far-fetched? The technology is almost at hand. Global Positioning System (GPS) receivers, which employ a network of satellites to locate precise latitude and longitude, are available in many new vehicles, and automobile manufacturers are scrambling to connect drivers to the Internet. Within the next 10 years ultrathin, holographic monitors could replace windows, and miniature video cameras could track events on the road and inside the car. Special software would then translate the vehicles' relative movements, and the passengers' actions, into digital fantasy—a virtual storyworld. Although it is only one possible future scenario, the FEV heralds a new era of interactive d-entertainment.

Engineers have attempted to create interactive entertainment since the 1950s [*see timeline above*], but most of today's choices still force some kind of undesirable trade-off. If you want a compelling narrative, you must usually accept the passive experience of cinema or television. If you prefer to make decisions, as with computer and video games, you generally sacrifice a good story line. The future of interactive d-entertainment will bring convergence of fixed narrative and personal choice, combined with the computer graphics capabilities it takes to render the story in real time.

Several innovations are key to these imminent advances. Increasing bandwidth is beginning to allow people to choose personalized entertainment from central distribution networks, such as movies-on-demand. Novel input devices, such as touch screens and speech-recognition tools, are changing the way we can communicate with the Internet or other networks. New types of audio and visual displays are enabling ever more realistic sensory environments. And miniaturized wireless technologies are making computing and communications mobile, bringing d-entertainment to your handheld computer, your watch and even your car windows. As media technology moves from the specialized spaces of the theater and living room into the total surround of everyday life, d-entertainment will be available wherever we are, whenever we are receptive and on whatever devices are present.

By its nature, interactive technology will also offer a wealth of choices about how a story unfolds, so no two people's entertainment experience need ever be the same. Writers will not have to script entire tales ahead of time, because the people who enter the story will become the characters whose decisions move the story along. A writer may shape the initial circumstances, but the story will unfold improvisationally. The story environment—and characters within it—will respond to personal messages, news and other forms of information. What this kind of virtual storyworld will require is a database network that is embedded with enough story elements and decision-making algorithms to generate various serendipitous actions with unique content.

Smart Stages and Animated Actors

Another necessary component of the virtual storyworld is an interactive stage where the action can take place. This stage is where content can be found, made and shared as needed; where informational messages can be sent and received; and where virtual characters can engage in their own business or interact with the audience. No such milieu exists, but a new type of cyberspace portal—currently under development at the Massachusetts Institute of Technology Media Laboratory—may be a promising first step. Called Happenstance, it is an animated computer graphics landscape that manages information according to the preferences and situation of a particular user.

As we know from cinema, landscapes provide

1989

Sim City (Will Wright, Maxis)
Introduces interactivity with no preordained outcome. Players try to build successful cities based on choices of utilities, housing, industries and so on.

Back to the Future: The Ride (Universal Studios Florida, Berkshire Ridefilms)
Convergence of sensory perception, other than sight and sound, with a theater experience.

1993

Doom (id Software)
3-D action game popularizes multi-player gaming. People all over the world play one another in real time, via the Internet.

Myst (Cyan)
Inserts puzzle-solving story such as Adventure into a graphically vivid computer world.

1995

The Spot (American Cybercast)
First interactive Web-based soap opera, or Webisodic. Viewers submit ideas about what they think characters should do next.

Cyberdome Theater (Boeing)
Audience members control outcome of theater experience by directing 3-D graphical images, using a five-button keypad in the armrest of each seat.

2000

Big Brother (CBS/Endemol Entertainment)
Audience members affect story line by voting to banish residents of a house where people are living unscripted lives. Viewers can watch an around-the-clock Webcast of the goings-on.

the illusion of continuity in space and time. Graphical representations of weather, plants and other features of the natural environment provide the same illusion in the cybersurround of Happenstance. Your window into this world is your computer monitor, and you use a mouse and keypad to navigate and send commands, but the similarity to current graphical interfaces ends there. This ecological interface translates common computer activities, such as conducting Internet searches, into movement through the landscape.

If you decide, for instance, that you're hungry for Chinese food, you could type a query that gets attached to an icon of a tree seed. You could then plant the seed in the cybergarden of Happenstance to begin a search for nearby restaurants. Today's Internet browsers would list the query results as hyperlinked blocks of text, but inside Happenstance the results appear as leaves sprouting on a tree. Before you can examine the leaves, a volcano begins erupting in the distance—a signal that news relevant to your search is about to arrive. Happenstance is programmed to search for serendipity; the query for Chinese food has also brought you documentary film elements about China. A cinematic editor-in-software directs a graphical sequence that pans away from the tree and zooms in on the cloud of smoke that is emerging from the volcano. The cloud moves overhead, and details from the films rain down into a river that identifies them by their keywords.

Even a stage as smart as a future-generation Happenstance cannot tell a story with all the complexity one would want. For interesting narratives to unwind, characters must be added to the mix. Imagine a fortune-teller sitting beside the tree of Chinese restaurants or a dog fetching various story elements from the river. Some characters could be mapped from nonhuman elements of real life, such as a *T. rex* from a BMW. Human actors—playing the part of characters that they design—could enter the story via their personal computers or other input devices. A third set of actors—semiautonomous, virtual beings that are part of the environment—could also work to move the story forward.

Creating convincing synthetic actors is no easy task. A satisfying interactive narrative requires characters that can convey emotion and improvise actions. They must have minds of their own, and they must be able to surprise us. Such characters contrast sharply with the best of those from the current generation of computer games. These dronelike beings can execute impressive, programmed actions, but they cannot improvise, and they cannot develop relationships with the players.

Media Lab researcher Bruce Blumberg is trying to scale the technological barrier between drones and actors by giving virtual beings the "brainpower" to respond to unscripted situations. He creates computer models of cognitive processes—sensory perceptions, learning, emotions and motor skills—and installs them in animated characters. Most recently Blumberg's team designed Duncan H. Terrier, a virtual sheepdog that they hope will one day emulate his real-world counterpart, a lively Silky Terrier named Sydney [*see illustration on page 79*].

People who enter the story become characters whose decisions move the story along.

Duncan's brain is a catalogue of action "tuples"—probabilistic statements that guide his behavior—that the programmers based on the ethology of how real dogs behave and learn. Each tuple defines a particular action and the conditions under which it should begin and end. Some tuples encode appropriate reactions to emotions, to phys-

COURTESY OF EDWIN BRIT WYCKOFF (*Winky Dink*); COURTESY OF DAVID WINTER (*Pong*); DIGITAL LEISURE, INC. (*Dragon's Lair*); MAXIS, INC. (*Sim City*); CYAN, INC. (*Myst*); SCREEN SHOT FROM AMERICA ONLINE (*Big Brother*) ical needs or to sensory perceptions. The key to making Duncan's responses expressive, Blumberg says, is adding a modifier to each programmed action: "If you see food, then eat it *quickly* until it's gone," or "If you get kicked, then run away *whimpering* until you are a safe distance away." Each tuple also has a value that derives from its consequences. Gobbling a dog biscuit results in a change in hunger, which in turn reduces the value of eating.

Much of Duncan's programming makes him instinctively inclined, as are real dogs, to please his master. Duncan "watches" you from a large video monitor; software translates your presence and your voice into sensory perceptions that Duncan can understand. You stand before him on a surfboardlike device that monitors the direction in which you shift your weight to move you through the virtual Scottish moors where Duncan lives.

Like a real dog, Duncan initially has no idea that performing actions in response to specific verbal commands will lead to good things. Rather he discovers this reality through experience. (Indeed, Duncan is trained using a technique borrowed from real dog training, called clicker training.) Initially, Duncan randomly chooses actions to perform such as "sit" and "shake." When he does something that you want to encourage, you click a button that sends him a signal of praise—a virtual doggie biscuit. Because sitting led to a good thing, its value is increased, making it more likely that Duncan will choose that action in the future. Duncan's statistical memory also keeps track of what seems to be true about the world when he performs the action and tries to find contexts in which he is most likely to get a treat. So if you give Duncan a verbal cue right as he is beginning to sit, and if you preferentially reward him in those cases, he will eventually learn that sitting when you say "sit" is a more reliable strategy for getting a treat than simply doing it when you happen to be around.

A convincing interactive character must convey emotion and improvise actions.

Duncan's 30-odd tuples, which represent both learned and innate behaviors, are always competing. Just when you think that you have Duncan's full attention, he may chase a squirrel that scampers by. The possibilities for surprise behaviors grow when you add sheep and wolves to Duncan's world, each of which are individually programmed. Perhaps ironically, computer power will probably not be the biggest challenge in developing characters more advanced than Duncan. The biggest hurdle will be figuring out how to integrate motivation, emotion and learning abilities in a convincing way. Characters such as Duncan are the first step toward creating synthetic actors who can enliven the interactive, virtual storyworlds of the future.

Getting There

Merging smart, iconic stage sets and synthetic actors within a vast computational network will open the door to interactive d-entertainment scenarios that we cannot imagine now. But as the simple case of Duncan shows, programming all the imagery that could result from the various serendipitous interactions among characters is itself a daunting task. The d-entertainment industry will also have to agree on standards that encourage the building of infrastructure and the mass-production of new devices that can all work together.

Even if engineers and artists can develop the technology, infrastructure and ideas, the economics will still need to be worked out. Who will contribute to these evolving stories, and who will pay for them? One possible strategy would be the use of microtransactions: you pay for the parts of the story you watch, and you are paid for your contributions. Another possibility would be to market a theater experience in which the audience controls the story as a group. In contrast to today's passive cinema experience, a storytelling database could allow a movie experience to be generated at run time. The pieces could be chosen based on the computed sum of the audience members' collective prior theater experience. Stored on a theater key card, this list of prior experiences would ensure that the new theater experience was different from everything any audience member had already seen.

A virtual storyworld where interactive d-entertainment waits for us around each bend of our daily lives might be years off. But projects like Duncan and Happenstance, as well as near-term possibilities such as the FEV, are beginning to take us there. Along the way, we get to enjoy the ride. SA

The Author

GLORIANNA DAVENPORT is director of interactive cinema at the Massachusetts Institute of Technology Media Laboratory, which she co-founded in 1985. Before that time she was a sculptor and a maker of Super 8 films. In 1978 she began teaching filmmaking at M.I.T. in the film school and helped to introduce video to the curriculum. Earlier this year she co-founded Media Lab Europe in Dublin. Davenport and her group currently focus on the creation of media forms for emerging digital technologies.

Further Information

Visit the Massachusetts Institute of Technology Media Laboratory at www.media.mit.edu

Read about the history of Pong at www.pong-story.com/intro.htm

Find information about Myst at http://sirrus.cyan.com/Online/Myst/MystHome

WEBB CHAPPELL

In late November a humble Iowa cow is slated to give birth to the world's first cloned endangered species, a baby bull to be named Noah. Noah is a gaur: a member of a species of large oxlike animals that are now rare in their homelands of India, Indochina and southeast Asia. These one-ton bovines have been hunted for sport for generations. More recently the gaur's habitats of forests, bamboo jungles and grasslands have dwindled to the point that only roughly 36,000 are thought to remain in the wild. The World Conservation Union–IUCN Red Data Book lists the gaur as endangered, and trade in live gaur or gaur products—whether horns, hides or hooves—is banned by the Convention on International Trade in Endangered Species (CITES).

But if all goes as predicted, in a few weeks a spindly-legged little Noah will trot in a new day in the conservation of his kind as well as in the preservation of many other endangered species. Perhaps most important, he will be living, mooing proof that one animal can carry and give birth to the exact genetic duplicate, or clone, of an animal of a different species. And Noah will be just the first creature up the ramp of the ark of endangered species that we and other scientists are currently attempting to clone: plans are under way to clone the African bongo antelope, the Sumatran tiger and that favorite of zoo lovers, the reluctant-to-reproduce giant panda. Cloning could also reincarnate some species that are already extinct—most immediately, perhaps, the bucardo mountain goat of Spain. The last bucardo—a female—died of a smashed skull when a tree fell on it early this year, but Spanish scientists have preserved some of its cells.

Advances in cloning offer a way to preserve and propagate endangered species that reproduce poorly in zoos until their habitats can be restored and they can be reintroduced to the wild. Cloning's main power, however, is that it allows researchers to introduce new genes back into the gene pool of a species that has few remaining animals. Most zoos are not equipped to collect and cryopreserve semen; similarly, eggs are difficult to obtain and are damaged by freezing. But by cloning animals whose body cells have been preserved, scientists can keep the genes of that individual alive, maintaining (and in some instances increasing) the overall genetic diversity of endangered populations of that species.

Nevertheless, some conservation biologists have been slow to recognize the benefits of basic assisted reproduction strategies, such as in vitro fertilization, and have been hesitant to consider cloning. Although we agree that every effort should be made to preserve wild spaces for the incredible diversity of life that inhabits this planet, in some cases either the battle has already been lost or its outcome looks dire. Cloning technology is not a panacea, but it offers the opportunity to save some of the species that contribute to that diversity.

A clone still requires a mother, however, and very few conservationists

SLIM FILMS (*illustration*); LYNDA RICHARDSON *Corbis* (*orangutan*); MARTIN WENDLER *Peter Arnold, Inc.* (*ocelot*); GERRY ELLIS *Minden Pictures* (*panda*); KENNETH W. FINK *Photo Researchers, Inc.* (*bongo*); FRANS LANTING *Minden Pictures* (*cheetah*); ROLAND SEITRE *Peter Arnold, Inc.* (*gaur*); JOHN CANCALOSI *Peter Arnold, Inc.* (*goat*)

Cloning Noah's Ark

Biotechnology might offer the best way to keep some endangered species from disappearing from the planet

by Robert P. Lanza, Betsy L. Dresser and Philip Damiani

would advocate rounding up wild female endangered animals for that purpose or subjecting a precious zoo resident of the same species to the rigors of assisted reproduction and surrogate motherhood. That means that to clone an endangered species, researchers such as ourselves must solve the problem of how to get cells from two different species to yield the clone of one.

A Gaur Is Born

It is a deceptively simple-looking process. A needle jabs through the protective layer, or zona pellucida, surrounding an egg that hours ago resided in a living ovary. In one deft movement, a research assistant uses it to suck out the egg's nucleus—which contains the majority of a cell's genetic material—leaving behind only a sac of gel called cytoplasm. Next he uses a second needle to inject another, whole cell under the egg's outer layer. With the flip of an electric switch, the cloning is complete: the electrical pulse fuses the introduced cell to the egg, and the early embryo begins to divide. In a few days, it will become a mass of cells large enough to implant into the uterus of a surrogate-mother animal previously treated with hormones. In a matter of months, that surrogate mother will give birth to a clone.

In practice, though, this technique—which scientists call nuclear transfer—is not so easy. To create Noah, we at Advanced Cell Technology (ACT) in Worcester, Mass., had to fuse skin cells taken from a male gaur with 692 enucleated cow eggs. As we report in the current issue of the journal *Cloning*, of those 692 cloned early embryos, only 81 grew in the laboratory into blastocysts, balls of 100 or so cells that are sufficiently developed to implant for gestation. We ended up inserting 42 blastocysts into 32 cows, but only eight became pregnant. We removed the fetuses from two of the pregnant cows for scientific analysis; four other animals experienced spontaneous abortions in the second or third month of the usual nine-month pregnancy; and the seventh cow had a very unexpected late-term spontaneous abortion in August.

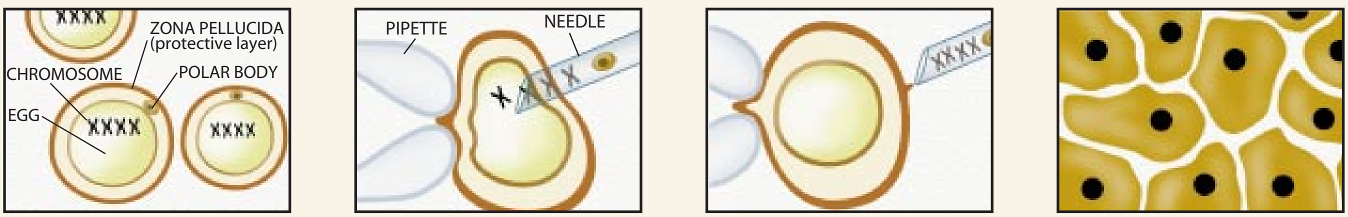
The statistics of the efficiency of cloning reflect the fact that the technology is still as much an art as it is a science—particularly when it involves transplanting an embryo into another species. Scientists, including those of us at ACT, have had the highest success rates cloning domestic cattle implanted into cows of the same species. But even in this instance we have had to work hard to produce just a few animals. For every 100 cow eggs we fuse with adult cattle cells, we can expect only between 15 and 20 to produce blastocysts. And only roughly 10 percent of those—one or two—yield live births.

The numbers reflect difficulties with the nuclear transfer process itself, which we are now working to understand. They are also a function of the vagaries of assisted reproduction technology.

Accordingly, we expect that the first few endangered species to be cloned will be those whose reproduction has already been well studied. Several zoos and conservation societies—including the Audubon Institute Center for Research of Endangered Species (AICRES) in New Orleans, which is led by one of us (Dresser)—have probed the reproductive biology of a range of endangered species, with some notable successes. Last November, for example, Dresser and her colleagues reported the first transplantation of a previously frozen embryo of an endangered animal into another species that resulted in a live birth. In this case, an ordinary house cat gave birth to an African wildcat, a species that has declined in some areas.

So far, beyond the African wildcat and the gaur, we and others have accomplished interspecies embryo transfers in four additional cases: an Indian desert cat into a domestic cat; a bongo antelope into a more common African antelope called an eland; a mouflon sheep into a domestic sheep; and a rare red deer into a common white-tailed deer. All yielded live births. We hope that the studies of felines will pave the way for cloning the cheetah, of which only roughly 12,000 remain in southern Africa. The prolonged courtship behavior of cheetahs requires substantial

THE NUCLEAR TRANSFER (CLONING) PROCESS

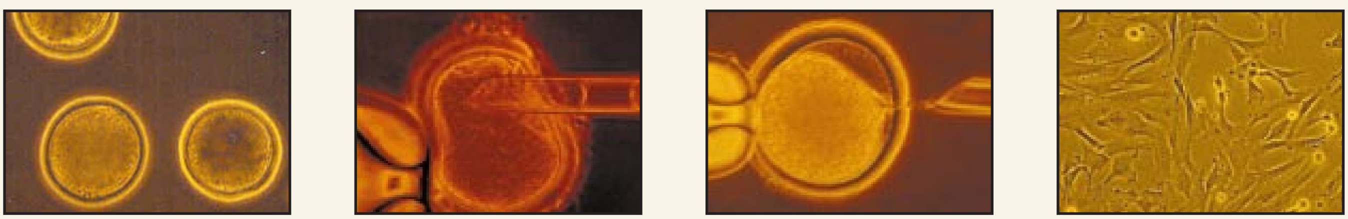


Recipient eggs are coaxed to mature in a culture dish. Each has a remnant egg cell called the polar body.

The polar bodies and chromosomes of each egg are drawn into a needle. A pipette holds the egg still.

Once the chromosomes and polar body are removed, all that remains inside the zona pellucida is cytoplasm.

Skin cells called fibroblasts are isolated from the animal to be cloned and grown in culture dishes.



LAURIE GRACE (*illustrations*); PHILIP DAMIANI (*photographs*) territory, a possible explanation for why the animals have bred so poorly in zoos and yet another reason to fear their extinction as their habitat shrinks.

Panda-monium

One of the most exciting candidates for endangered-species cloning—the giant panda—has not yet been the subject of interspecies transfer experiments, but it has benefited from assisted reproduction technology. Following the well-publicized erotic fumblings of the National Zoo's ill-fated panda pair, the late Ling-Ling and Hsing-Hsing, the San Diego Zoo turned to artificial insemination to make proud parents of its Bai Yun and Shi Shi. Baby Hua Mei was born in August 1999.

Giant pandas are such emblems of endangered species that the World Wildlife Fund (WWF) uses one in its logo. According to a census that is now almost 20 years old, fewer than 1,000 pandas remain in their mountainous habitats of bamboo forest in southwest China. But some biologists think that the population might have rebounded a bit in some areas. The WWF expects to complete a census of China's pandas in mid-2002 to produce a better estimate.

In the meantime, we at ACT are discussing plans with the government of China to clone a giant panda. Chinese scientists have already made strides toward the goal of panda cloning. In August 1999 Dayuan Chen of the institute and his co-workers published a paper in the English-language journal *Science in China* announcing that they had fused panda skeletal muscle, uterus and mammary gland cells with the eggs of a rabbit and then coaxed the cloned cells to develop into blastocysts in the laboratory.

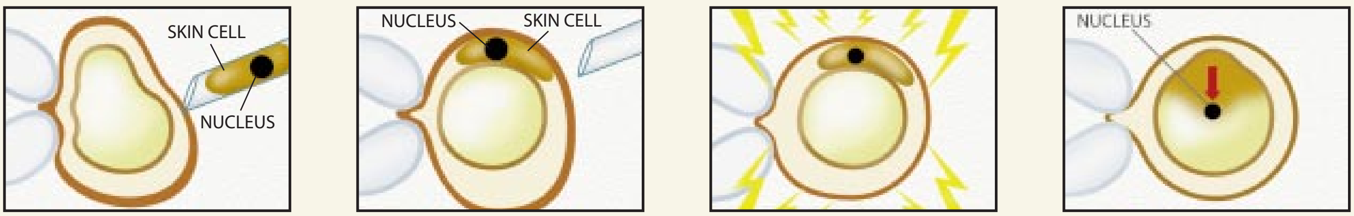
A rabbit, of course, is too small to serve as a surrogate mother for a giant panda. Instead ACT and the Chinese plan to turn to American black bears. As this issue of *Scientific American* goes to press, ACT is finalizing plans to obtain eggs from female black bears killed during this autumn's hunting season in the northeastern U.S. Together with the Chinese, ACT scientists hope to use these eggs and frozen cells from the late Hsing-Hsing or Ling-Ling to generate cloned giant panda embryos that can be implanted into a female black bear now living in a zoo. A research group that includes veterinarians at Bear Country U.S.A. in Rapid City, S.D., has already demonstrated that black bears can give birth to transplanted embryos. They reported the successful birth of a black bear cub from an embryo transferred from one pregnant black bear to another last year in the journal *Theriogenology*.

AICRES scientists hope to take advantage of the success with bongo antelope that one of us (Dresser) had while at the Cincinnati Zoo. In 1984 Dresser and Charles Earle Pope of the University of Alabama at Birmingham (now with AICRES and Louisiana State University) and their colleagues announced the birth of a bongo after moving very early embryos from a pregnant female bongo to an eland surrogate mother.

Most of the mountain subspecies of bongo—a medium-size antelope with vertical white stripes—live in captivity. According to the World Conservation Union–IUCN, the mountain bongo is endangered, with only 50 or so remaining in a small region of Kenya. In contrast, the 1999 Bongo International Studbook lists nearly 550 mountain bongo living in zoos throughout the world. The lowland bongo subspecies is slightly better off: it is listed as "near threatened" and has a population of perhaps several thousand scattered throughout central and western Africa.

A coalition of conservation organizations in the U.S. and Kenya is now planning to send mountain bongo that have been bred in captivity to two sites in Kenya. And in a new approach to reintroducing a species, AICRES is working in Kenya to transfer frozen bongo embryos into eland surrogates. Cloning could support these efforts and possibly yield more bongo for reintroduction.

But what about animals that are already extinct? Chances are slim to nil that scientists will soon be able to clone dinosaurs, à la *Jurassic Park,* or woolly mammoths. The primary problem is the dearth of preserved tissue—and hence DNA. A group of researchers unearthed



An entire skin cell is taken up into the needle, which is again punched through the zona pellucida.

The skin cell is injected underneath the zona pellucida, where it remains separate from the egg cytoplasm.

Each injected egg is exposed to an electric shock that fuses the skin cell with the egg cytoplasm.

The skin cell's nucleus, with its genes, enters the egg cytoplasm. Within a few hours, the fused cell begins to divide.

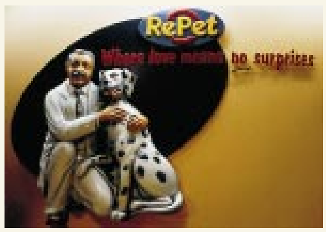
WHAT ABOUT ROVER AND FLUFFY?

The list of domesticated animals that scientists have been able to clone so far includes sheep, cattle, goats and laboratory mice—and now, we expect, the gaur. Compared with that menagerie, you'd think that cloning an ordinary dog or cat would be a snap. Unfortunately, this has not been the case. Both of our research groups have created cloned cat embryos and have implanted them into female cats, but as this article goes to press, neither of our teams has yet obtained a full-term pregnancy. Dogs have presented even more problems.

But we anticipate success soon. At Advanced Cell Technology (ACT), we have undertaken a research program that uses cloning technology to propagate pets as well as service animals such as seeing-eye dogs for the blind, hearing dogs for the deaf, search-and-rescue dogs, and animals used for social therapy. Together with Louisiana State University, the Audubon Institute has teamed up with a company called Lazaron BioTechnologies in Baton Rouge, La., to clone pet dogs and cats.

A surprising number of people are interested in cloning their favorite deceased pet in the hope of getting an animal with similar behavioral characteristics. A good deal of a cat or dog's demeanor is thought to be genetically determined. Although one can argue that there are already plenty of cats and dogs in the world that need homes, people still use traditional breeding methods to try to reproduce a particularly desirable animal. Cloning could offer a more efficient alternative. It could be particularly important in the case of service animals. Currently, for instance, male seeing-eye dogs are neutered at an early age so that they can concentrate better during their expensive and rigorous training. So, unfortunately, even if a dog turns out to be very good at his job, he can't be bred to produce more like him.

Our efforts to clone pets could also pay off for endangered species. We expect to be able to apply the information we obtain from cloning cats and dogs to preserving endangered felines and canines.



SERVICE ANIMALS and pets might soon be cloned. In a new film, *The 6th Day*, grieving pet owners go to a company called RePet to copy their animals.

KLEIN/HOBERT *BIOS Peter Arnold, Inc. (top)*; ROB MCEWAN *Phoenix Pictures (bottom)*

ACT and several other companies now offer pet cloning kits that veterinarians can use to preserve samples from a client's pet for possible future cloning. The kits contain materials for collecting a skin specimen and sending it back to a laboratory. Research assistants there use the tissue to establish a collection of pure, dividing cells called a cell line, which will be the source of donor cells for cloning.

ACT extracts eggs for the cloning procedure from reproductive tracts taken from animals that have been spayed by veterinarians. We remove the ovaries and carefully puncture all visible follicles to release the eggs. Then we collect the eggs and place them in a specialized maturation medium that contains hormones, proteins and nutrients. Once fully matured, the eggs are ready for the nuclear transfer procedure [*see illustration on pages 86 and 87*].

So far our main focus has been the domestic cat, primarily because its reproductive physiology has been well studied, and embryo transfers of early- and late-stage embryos have resulted in the birth of live kittens. Both ACT and the Audubon Institute have been able to establish systems for prompting cat eggs to mature in the lab and have consistently produced cloned embryos that are being transferred to recipients.

But dogs are a different story. The dog's reproductive physiology is unique among mammalian species. Dogs ovulate an immature egg that has a very long maturation time. This means that we need a different maturation system from the one we have used in cats and that we have fewer eggs to work with in the end. So Fluffy will probably have a leg up on Rover when it comes to cloning. *—R.P.L., B.L.D. and P.D.*

what they had hoped would be a well-preserved mammoth last year, but repeated freezing and thawing over the eons had poked holes in the creature's DNA, and molecular biologists have not yet found a feasible way of filling in such genetic gaps.

A similar difficulty has hobbled efforts by Australian scientists to clone a thylacine, or Tasmanian tiger, a wolflike marsupial that died out in the 1930s. Researchers at the Australian Museum in Sydney are attempting to clone cells from a thylacine pup that was preserved in alcohol in 1866, but the DNA is in such poor condition that they say they will have to reconstruct all of the animal's chromosomes.

The recently extinct bucardo may prove a more promising target for resurrection. ACT is arranging a collaboration with Alberto Fernández-Arias and José Folch of the Agricultural Research Service in Zaragoza, Spain. Fernández-Arias froze tissue from the last bucardo. He and Folch had tried for several years to preserve the mountain goat, which in the end was wiped out by poaching, habitat destruction and landslides. Last year they transferred embryos from a subspecies related to the bucardo to a domestic goat, yielding live kids.

But even if interspecies nuclear transfer succeeds for the bucardo, it will yield only a sorority of clones, because we have tissue from just one animal, a female. ACT plans to try to make a male by removing one copy of the X chromosome from one of the female bucardo's cells and using a tiny artificial cell called a microsome to add a Y chromosome from a closely related goat species. The technology has been used by other researchers to manipulate human chromosomes, but it has never before been used for cloning. A nonprofit organization called the Soma Foundation has been established to help fund such efforts.

Why Clone?

Cloning endangered species is controversial, but we assert that it has an important place in plans to manage species that are in danger of extinction. Some researchers have argued against it, maintaining that it would restrict an already dwindling amount of genetic diversity for those species. Not so. We advocate the establishment of a worldwide

Left to right, clockwise: GERRY ELLIS *Minden Pictures* (*cheetah*); CHARLES V. ANGELO *Photo Researchers, Inc.* (*bongo*); FRANS LANTING *Minden Pictures* (*giant panda*); JOHN CANCALOSI *Peter Arnold, Inc.* (*bucardo*); ROLAND SEITRE *Peter Arnold, Inc.* (*gaur*); MARTIN WENDLER *Peter Arnold, Inc.* (*ocelot*)

network of repositories to hold frozen tissue from all the individuals of an endangered species from which it is possible to collect samples. Those cells—like the sperm and eggs now being collected in "frozen zoos" by a variety of zoological parks—could serve as a genetic trust for reconstituting entire populations of a given species. Such an enterprise would be relatively inexpensive: a typical three-foot freezer can hold more than 2,000 samples and uses just a few dollars of electricity per year. Currently only AICRES and the San Diego Zoo's Center for Reproduction of Endangered Species maintain banks of frozen body cells that could be used for cloning.

Other critics claim that the practice could overshadow efforts to preserve habitat. We counter that while habitat preservation is the keystone of species conservation, some countries are too poor or too unstable to support sustainable conservation efforts. What is more, the continued growth of the human species will probably make it impossible to save enough habitat for some other species. Cloning by interspecies nuclear transfer offers the possibility of keeping the genetic stock of those species on hand without maintaining populations in captivity, which is a particularly costly enterprise in the case of large animals.

CLONING CANDIDATES include (*clockwise from upper left*) the cheetah, bongo, giant panda, bucardo, gaur and ocelot.

Another argument against cloning endangered species is that it might siphon donor money away from habitat maintenance. But not all potential donors are willing to support efforts to stem the tide of habitat destruction. We should recognize that some who would otherwise not donate to preserve endangered species at all might want to support cloning or other assisted reproduction technologies. The time to act is now. SA

The Authors

ROBERT P. LANZA, BETSY L. DRESSER and PHILIP DAMIANI share an interest in reproductive biology and animals. Lanza is vice president of medical and scientific development at Advanced Cell Technology (ACT) in Worcester, Mass. He founded the South Meadow Pond and Wildlife Association in Worcester County and is a member of the conservation commission of Clinton Township. Dresser is senior vice president for research at the Audubon Institute and director of the Audubon Institute Center for Research of Endangered Species and the Freeport-McMoRan Audubon Species Survival Center, all in New Orleans. Damiani, a research scientist at ACT, is also a member of the International Embryo Transfer Society's committee on cryopreservation.

Further Information

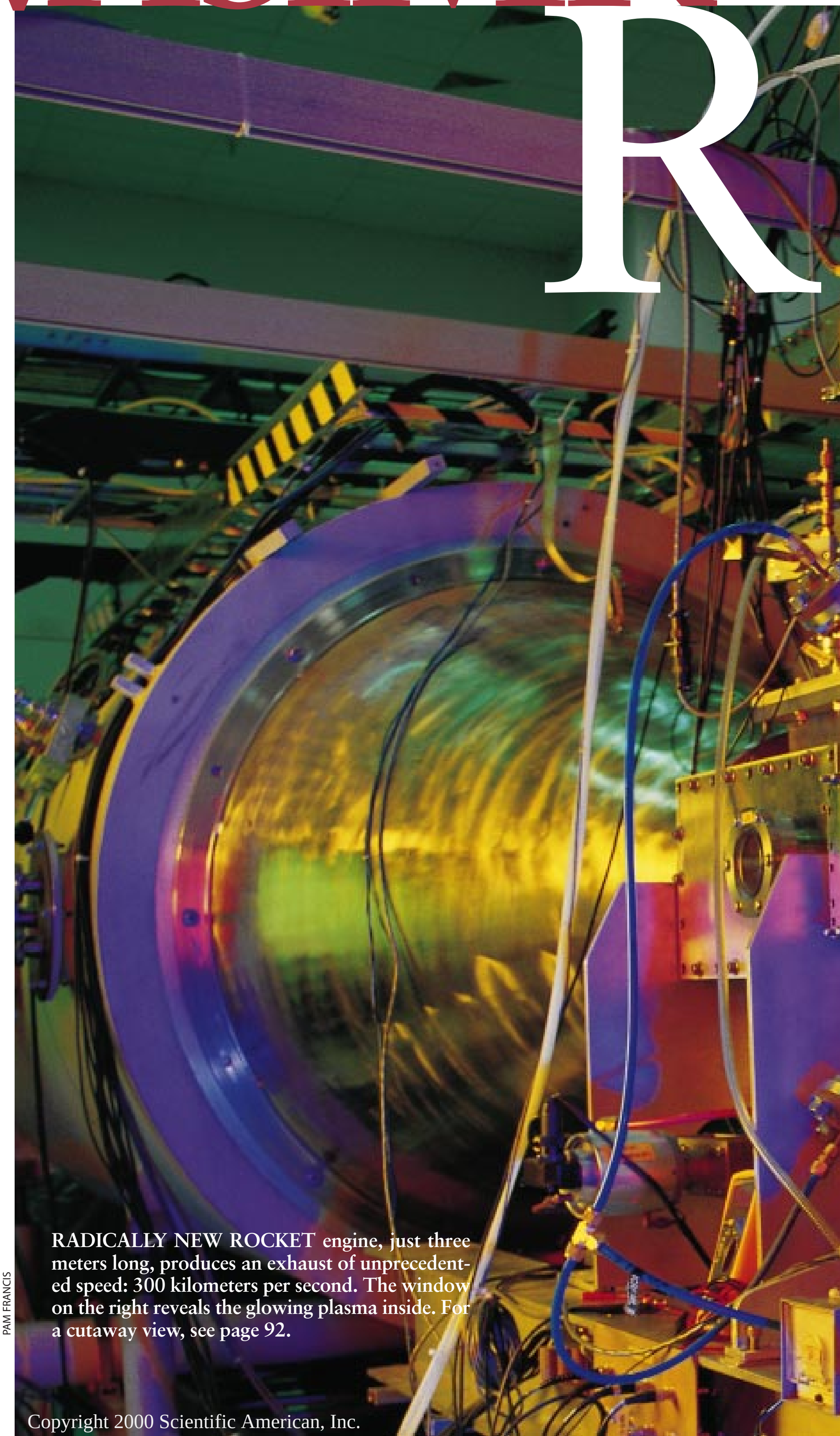
PRESERVATION OF ENDANGERED SPECIES AND POPULATIONS: A ROLE FOR GENOME BANKING, SOMATIC CELL CLONING, AND ANDROGENESIS? Graham E. Corley-Smith and Bruce P. Brandhorst in *Molecular Reproduction and Development,* Vol. 53, No. 3, pages 363–367; July 1999.

BIODIVERSITY HOTSPOTS FOR CONSERVATION PRIORITIES. Norman Myers, Russell A. Mittermeier, Cristina G. Mittermeier, Gustavo A. B. da Fonseca and Jennifer Kent in *Nature,* Vol. 403, No. 6772, pages 853–858; February 24, 2000.

VANISHING BEFORE OUR EYES. E. O. Wilson in *Time* (special report on Earth Day 2000), pages 29–34; April–May 2000.

The VASIMR

by Franklin R. Chang Díaz



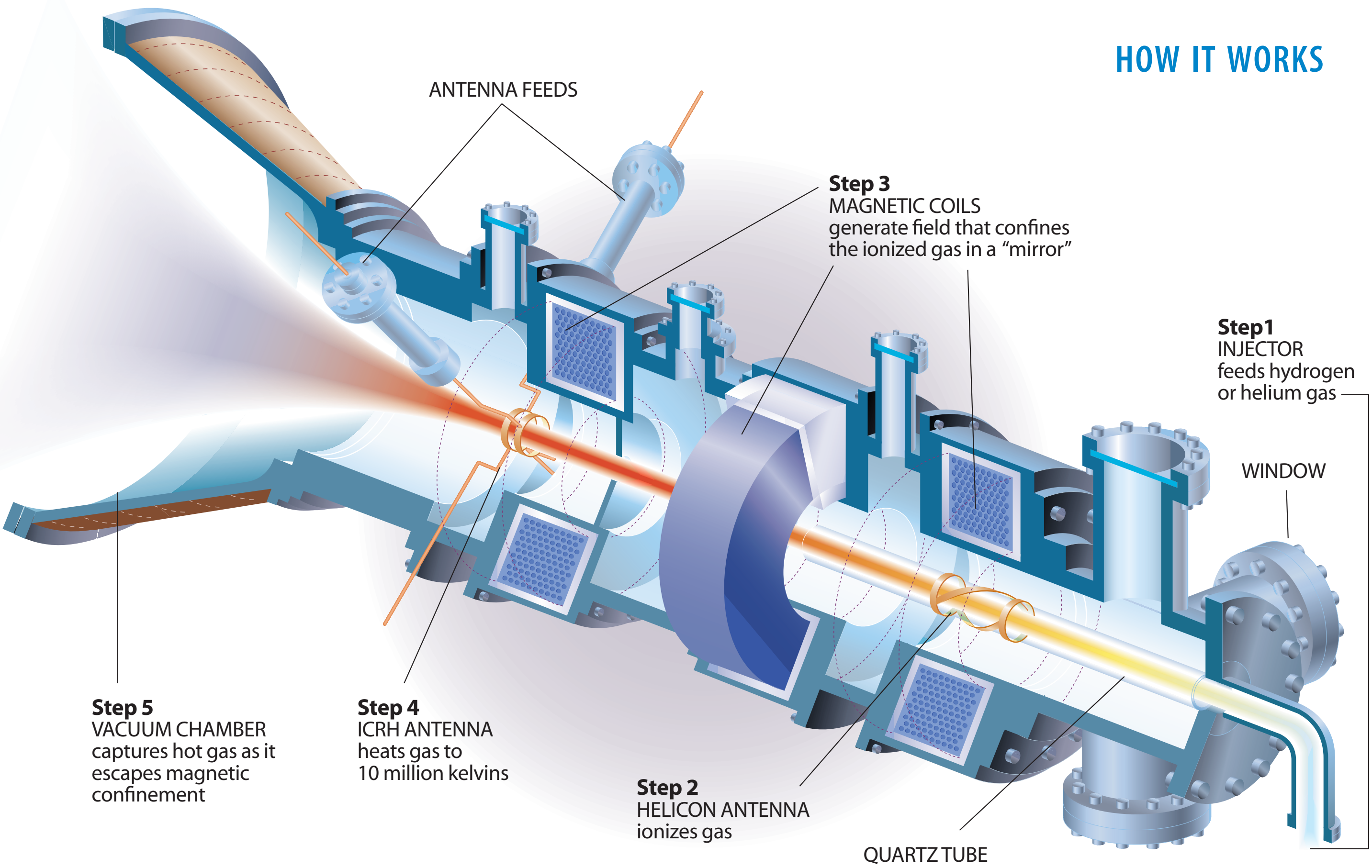
PAM FRANCIS

RADICALLY NEW ROCKET engine, just three meters long, produces an exhaust of unprecedented speed: 300 kilometers per second. The window on the right reveals the glowing plasma inside. For a cutaway view, see page 92.

We dream of going to the stars. As young children growing up in the 1950s, my friends and I were awestruck by the possibility of space travel. As I have learned over the years, our fascination was not unique to my Costa Rican upbringing. Indeed, many of my co-workers today, coming from different parts of the globe, recount similar childhood longings. In the past 50 years, I have had the opportunity to witness the development of the first ships that have transported humans beyond Earth. In the past 20, I have been fortunate to ride on some of these rockets and to get a firsthand glimpse of wonders I could only imagine before. It would seem as if we are destined to burst off our fragile planet and move into the cosmos in a new human odyssey. Such a mighty undertaking will dwarf the westward European expansion of the 16th century.

ocket

There used to be two types of rocket: powerful but fuel-guzzling, or efficient but weak.
Now there is a third option that combines the advantages of both



Yet we lack the ships required to venture far into the vastness of space. With today's chemical rockets, a trip to Mars would take up to 10 months in a vulnerable and limited spacecraft. There would be little room for useful payload. Most of the ship's mass would be taken up by propellant, which would be spent in a few short bursts, leaving the ship to coast for most of the journey [see "How to Go to Mars," by George Musser and Mark Alpert; SCIENTIFIC AMERICAN, March]. If people were to travel to Mars under these conditions, their bodies and minds would suffer considerably. Months of exposure to weightlessness would weaken their muscles and bones, and the persistent radiation of outer space would damage their immune systems.

To be safe, human interplanetary spacecraft must be fast, reliable and able to abort in the event of malfunction. Their propulsion systems must be capable of handling not just the cruise phase of the journey but also the maneuvering near the origin and destination planets. Whereas chemical propulsion can continue to provide excellent surface-to-orbit transportation, new technologies are required to send humans to the planets and ultimately to the stars.

Plasma rockets are one such technology. Utilizing ionized gases accelerated by electric and magnetic fields, they increase performance far beyond the limits of the chemical rocket. My research team has been developing one of these concepts, the Variable Specific Impulse Magnetoplasma Rocket (VASIMR), since the early 1980s. Its genesis dates back to the late 1970s, when I was involved in the study of magnetic ducts and their application to controlled nuclear fusion. In such ducts, a magnetic field insulates a hot plasma from its nearest material surface, letting it reach temperatures of hundreds of millions of kelvins.

I theorized that a duct, properly shaped, could form a magnetic nozzle and convert the plasma energy to rocket thrust. Such a structure functions like a conventional rocket nozzle but can withstand much higher temperatures. Further investigation suggested that the system could also generate a variable exhaust, adaptable to the conditions of flight, just as an automobile transmission matches the power of the engine to the needs of the road. Although the idea of variable exhaust dates to the early rocket pioneers, its implementation in chemical rockets with fixed material nozzles has proved impractical. In VASIMR the concept is finally poised to become a reality.

Newton's Rocket

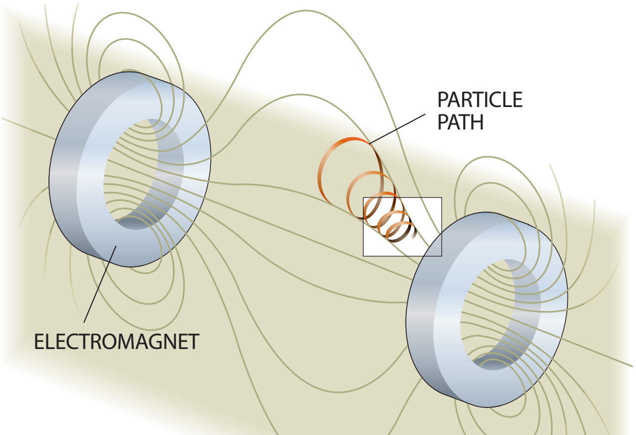
The principle of rocket propulsion stems from Newton's law of action and reaction. A rocket propels itself by expelling material in the direction opposite to its motion. The material is usually a gas heated by a chemical reaction, but the general principle applies equally well to the motion of a simple garden sprinkler.

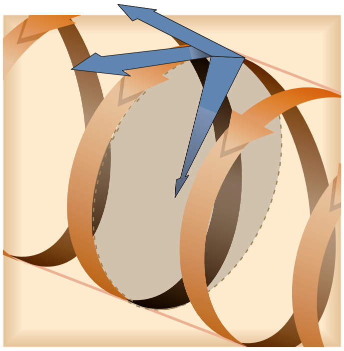
Rocket thrust is measured in newtons and is the product of exhaust velocity (relative to the ship) and the rate of propellant flow. Quite simply, the same thrust is obtained by ejecting either more material at low velocity or less at high velocity. The latter approach saves on fuel but generally entails high exhaust temperatures.

To gauge rocket performance, engineers use the term specific impulse (I_{sp}), which is the exhaust velocity divided by the acceleration of gravity at sea level (9.8 meters per second per second). Although thrust is directly proportional to I_{sp}, the power needed to produce it is proportional to the square of the I_{sp}.

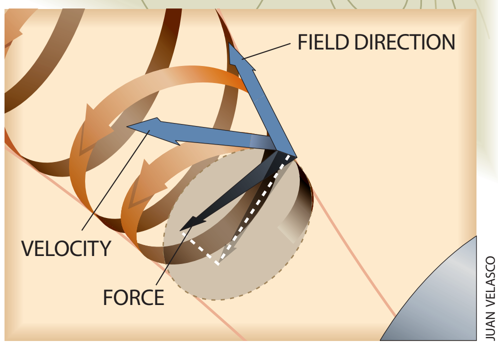
SAMUEL VELASCO

MAGNETIC mirror traps the particles so they can be heated to 10 million kelvins. It consists of two ring electromagnets that set up a bulging magnetic field between them. In VASIMR, a third magnet extends the mirror and provides for a magnetic "nozzle" that pushes particles out.





NEAR THE CENTER of the mirror, the field lines are parallel, so the magnetic force is radial. Particles travel at a constant speed along a helix of nearly constant radius.



NEAR EACH MAGNET, the field lines tilt. A component of the force now pushes the particles away from the magnet. (If the particles are moving toward the magnet, they may be stopped and reversed.) In this region, the helix tightens.

Therefore, the power required for a given thrust increases linearly with I_{sp}. In chemical rockets this power originates in the exothermic reaction of the fuel and oxidizer. In others, it must be imparted to the exhaust by a propellant heater or accelerator. Such systems depend on a power source elsewhere in the ship. Solar panels are generally used; the abundant power requirements of human space exploration, however, will favor nuclear reactors [*see box on page 97*]. This is especially true for missions beyond Mars, where sunlight is relatively feeble.

In our quest for high fuel efficiency, hence high I_{sp}, my research team has moved away from chemical reactions, in which the temperature is only a few thousand kelvins, and entered the realm of plasma physics, in which the temperature is high enough to strip the atoms of some (if not all) of their electrons. The temperature of a plasma starts at about 10,000 kelvins, but present-day laboratory plasmas can be 1,000 times hotter. The plasma is a soup of charged particles: positive ions and negative electrons. At these temperatures the ions, which constitute the bulk of the mass, move at velocities of 300,000 meters per second—60 times faster than the particles in the best chemical rockets.

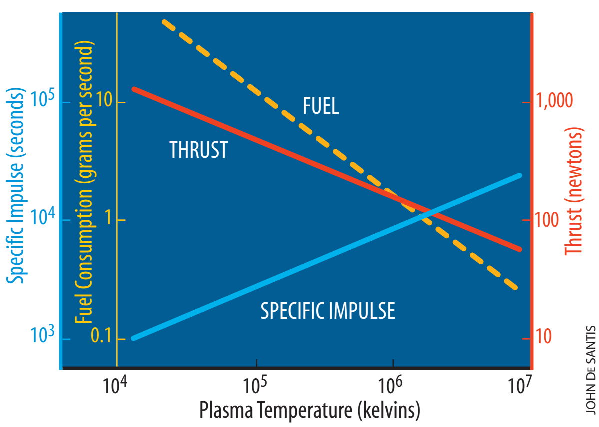
Usually, by design, the power output of the engine is kept at a maximum, so thrust and I_{sp} are inversely related. Increasing one always comes at the expense of the other. Therefore, for the same propellant, a high I_{sp} rocket delivers a greater payload than a low I_{sp} one, but in a longer time. If a rocket could vary thrust and I_{sp}, it could optimize propellant usage and deliver a maximum payload in minimum time. I call this technique constant power throttling (CPT). It is similar to the function of an automobile transmission in climbing a hill or the feathering of a propeller engine in moving through the air.

One way to visualize CPT is by considering the way in which the ship acquires kinetic energy from the exhaust. If this process were totally efficient, the exhaust particles (as viewed by a ground observer) would leave the ship at rest; the ship would be moving at the exhaust speed. All the exhaust energy would have been given to the ship. Thus, for a slow ship, an appropriately slow exhaust better utilizes the power source. As the ship speeds up, a faster (hotter) but leaner exhaust gives better results. Under CPT, the ship starts at high thrust for rapid acceleration. As its speed increases, I_{sp} gradually increases and thrust decreases for greater fuel economy. A car does exactly the same thing when it starts in low gear and steadily shifts up.

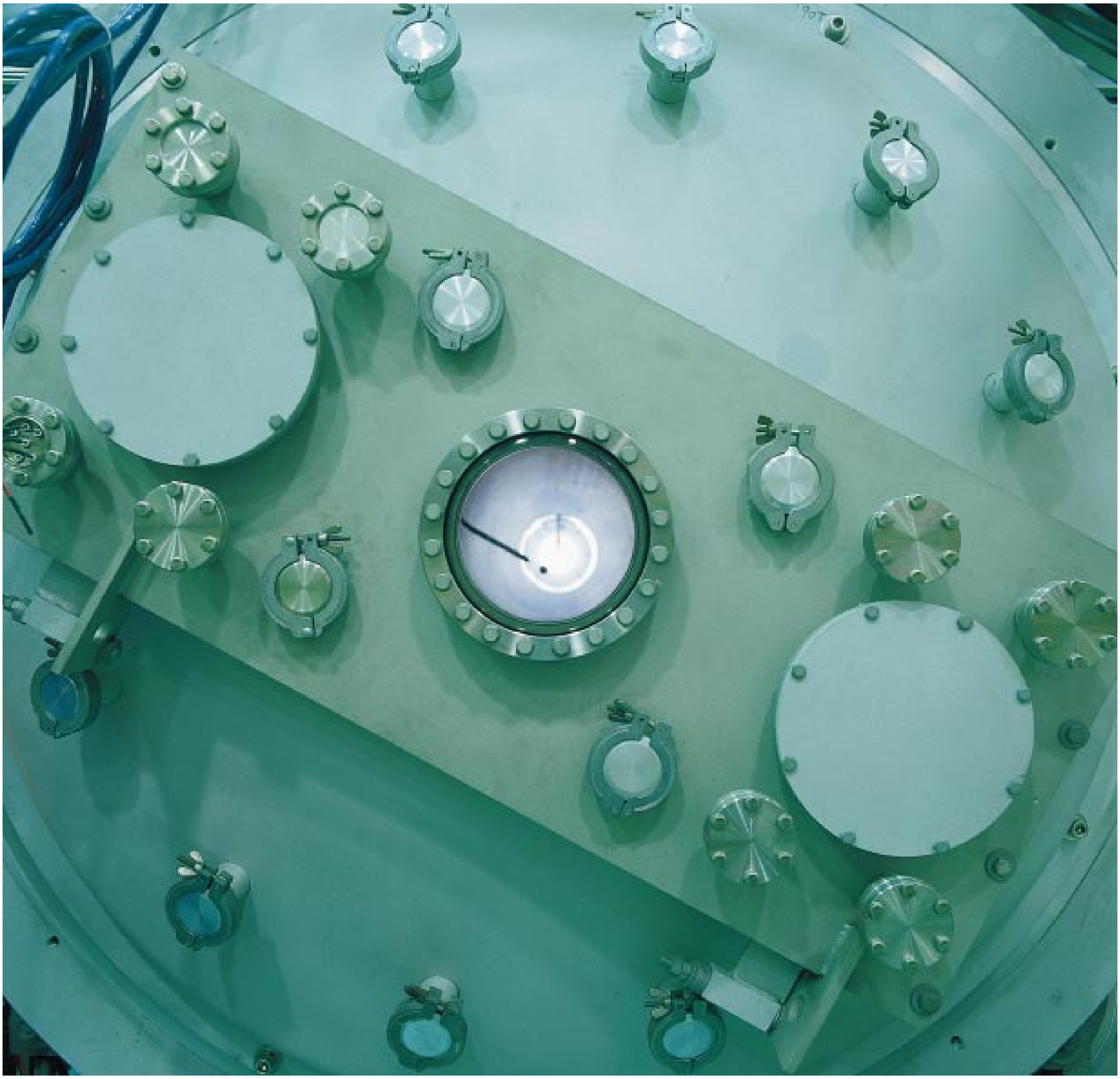
Bouncing Back and Forth

VASIMR embodies a class of magnetic ducts called magnetic mirrors. The simplest magnetic mirror is produced by two ring electromagnets with current flowing in the same direction. The magnetic field is constricted near the rings but bulges out in between them. Charged particles move in a helix along field lines, orbiting around them at a specific radius, the Larmor radius, and at the so-called cyclotron frequency. As one might expect, for a field of a given strength, the heavier particles (the ions) have a lower cyclotron frequency and larger Larmor radius than the light ones (the electrons) do. Also, strong fields lead to a high cyclotron frequency and small Larmor radius. In VASIMR, the ion cyclotron frequency is a few megahertz (MHz), whereas its electron equivalent is in the gigahertz range.

The particles' velocity has two components: one parallel to the field (corresponding to the forward motion along the field line) and the other perpendicular (corresponding to the orbital motion around the line). When a particle approaches a constricted (hence strong-



LIKE A CAR GEARSHIFT, and unlike other rockets, VASIMR can adjust its output. By increasing its temperature, it boosts its specific impulse (*blue*) and reduces fuel consumption (*yellow*)—at the price of less thrust (*red*). The power is a constant 10 megawatts.



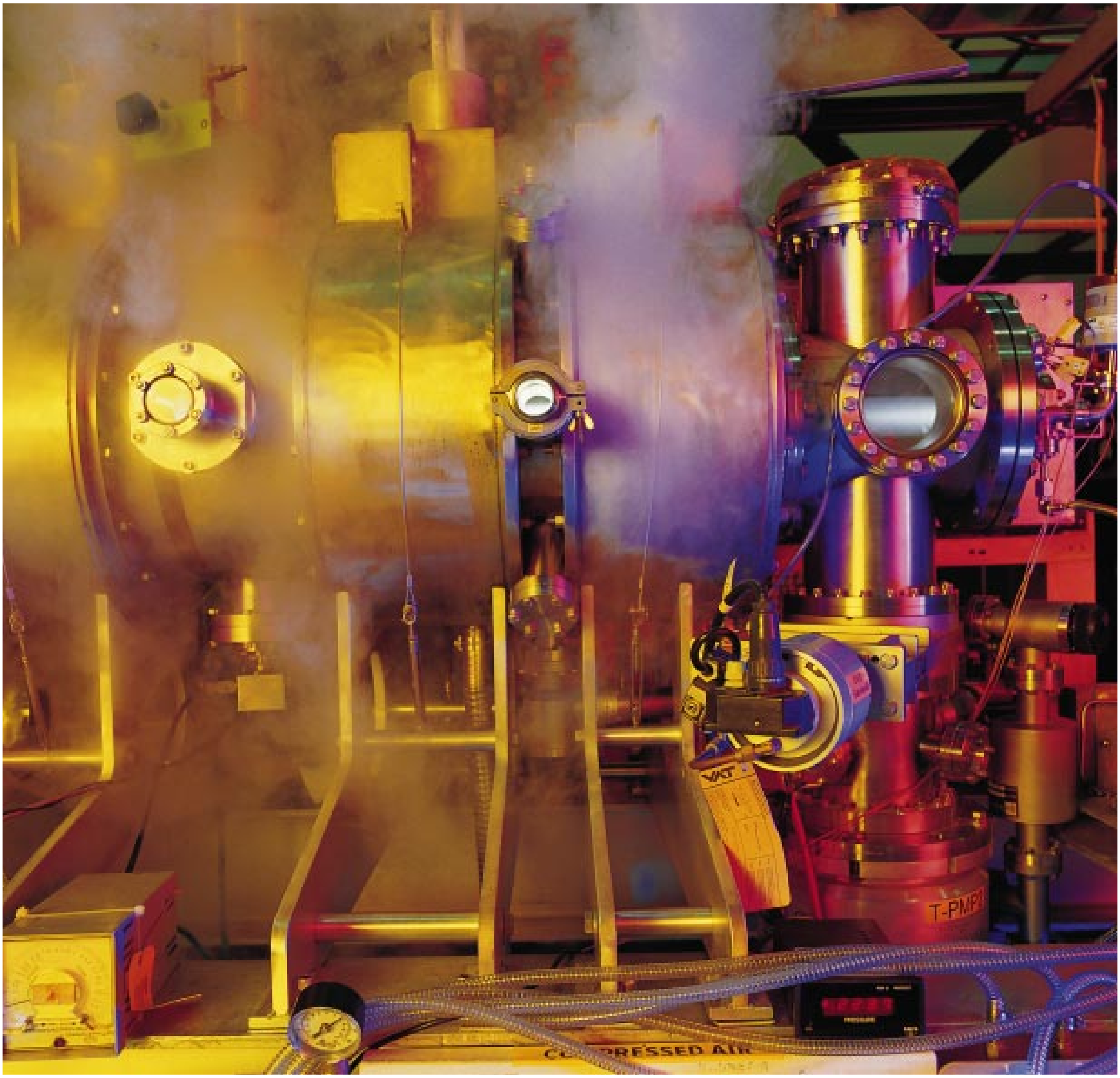
PAM FRANCIS

LOOKING DOWN THE BARREL **of the rocket, you see the plasma coming straight at you. The window is 15 centimeters across.**

er) field, its perpendicular velocity increases, but its parallel one is reduced proportionately to keep the total energy constant. The reason has to do with the direction of the force exerted by the field on the particle. The force is always perpendicular to both the particle's velocity and the field direction. Near the center of the mirror, where the field lines are parallel, the force is radial and so has no effect on the parallel velocity. But as the particle enters the constriction, the force tilts away from the constriction, resulting in an imbalance that decelerates the particle [*see top illustration on preceding page*]. If the particle is exiting the constriction, the field has the opposite effect and the particle accelerates. Because no energy has been added, the acceleration comes at the expense of rotational motion. The magnetic field does no work on the particle; it is simply a vehicle enabling this energy transfer.

These simple arguments hold as long as the field constriction is slow and gradual compared with the particle motion—a condition known as adiabaticity. In their curling motion around the field lines, the particles are guided by them, but like fast-moving vehicles on a slippery highway, they can follow only lines that do not curve sharply.

A magnetic mirror can trap particles if they are sufficiently slow to be reflected at the field constrictions. The particles bounce between them until something disrupts their parallel velocity such that it overcomes the trap or until one of the constrictions is reduced. With enough parallel velocity, the particle will push through and accelerate on the other side. Sudden changes in the velocity of a trapped particle, which may be enough to untrap it, can be brought about by random events such as collisions with other particles, interaction with electromagnetic waves, or plasma instabilities and turbulence. The magnetic field of Earth is a natural mirror. Charged particles from the ionosphere bounce back and forth between the North and South



PAM FRANCIS

BILLOWING CLOUDS OF WATER VAPOR pour off the magnets, which are kept at liquid-nitrogen temperature.

Poles. Some of them penetrate deep into the upper atmosphere, creating the spectacular auroras seen at high latitudes. VASIMR uses three such magnetic structures, linked together: a forward plasma injector, which ionizes the neutral gas; a central power amplifier, which energizes the plasma; and an aft magnetic nozzle, which finally ejects it into space.

Beam in the Power

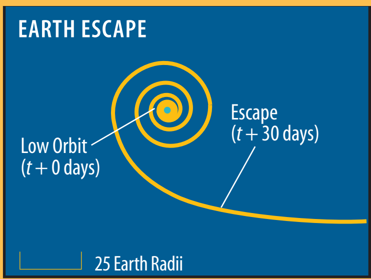
Most plasma rockets require physical electrodes, which erode quickly in the harsh environment. In contrast, VASIMR uses radio antennas. The radio waves heat the plasma just like a microwave oven heats food. Two wave processes come into play. First, neutral gas in the injector stage becomes a dense and comparatively cold (about 60,000 kelvins) plasma through the action of helicon waves. These are electromagnetic oscillations at frequencies of 10 to 50 MHz, which, in a magnetic field, energize free electrons in a gas. The electrons quickly multiply by liberating other electrons from nearby atoms in a cascade of ionization. Although the details are poorly understood, helicons are widely used in semiconductor manufacturing.

Once made, the plasma flows into the central stage, where it is heated by further wave action; the waves of choice here, however, are slightly lower-frequency ion cyclotron oscillations, so named because they resonate with the natural rotational motion of the ions. The wave's electric field is perpendicular to the external magnetic field and rotates at the ion cyclotron frequency. The resonance energizes the perpendicular motion of the particles. This effect, known as ion cyclotron resonance heating (ICRH), is widely used in fusion research. The central stage is ultimately responsible for the high I_{sp} of the rocket.

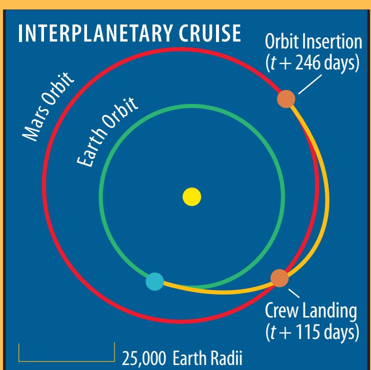
One might wonder how a boost in the ions' perpendicular motion could impart any useful momentum to the rocket exhaust. The answer lies in the

MARS TRAJECTORY

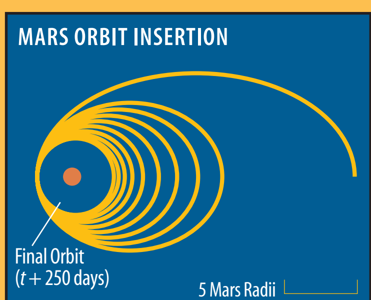
To get to Mars, the ship must first break free of Earth's gravity. Using low gear (maximum thrust and an I_{sp} of 3,000 seconds), VASIMR builds up speed.



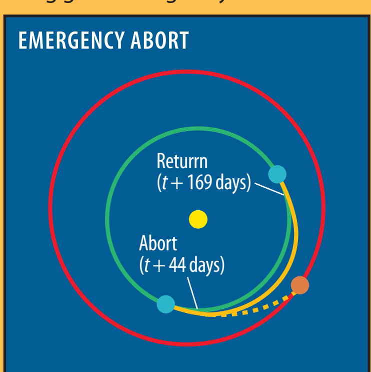
Unlike ordinary rockets VASIMR never coasts. It shifts to higher gears, reaching an I_{sp} of 30,000 seconds 75 days into the flight, and then starts to decelerate.



After flying by Mars to drop off the astronauts, the ship lowers itself into orbit.



VASIMR can turn around if something goes wrong early in the cruise.



JOHN DE SANTIS

physics of the magnetic nozzle, the final stage of VASIMR. The diverging field here transfers energy from the perpendicular motion to the parallel motion, accelerating the ions along the exhaust. Being much more massive, the ions drag the electrons along, so the plasma exits the rocket as a neutral fluid. In VASIMR, this nozzle expansion occurs over a distance of about 50 centimeters.

Once the expansion is complete, the plasma must detach from the rocket. Recent studies by Roald Sagdeev of the University of Maryland and Boris Breizman of the University of Texas highlight the basic physics. The model involves the Alfvén speed, named after Swedish physicist Hannes Alfvén, who first described it. Disturbances in a magnetized plasma propagate along the field at this speed. In a magnetic nozzle the Alfvén speed plays a role similar to that of the sound speed in a conventional nozzle.

The transition from sub-Alfvénic to super-Alfvénic flow delineates a boundary beyond which the flow downstream has no effect upstream—which ensures that the detachment exerts no drag on the rocket. The VASIMR nozzle is designed to expand the plasma past this boundary, to a point where the energy content of the field is small compared with that of the plasma flow. The plasma then breaks free, carrying with it a small amount of the field. A similar behavior is thought to occur in nature when solar flares detach from the magnetic field of the sun. The energy expenditure in the field distortion only minimally taxes the performance of the rocket. Our studies show that plasma detachment occurs one to two meters away from the nozzle throat.

Partitioning the Power

The throttling that makes VASIMR distinctive is done mainly by changing the relative fraction of power going to the helicon and ICRH systems. For high thrust, power is routed predominantly to the helicon, producing more ions at lower velocity. For high I_{sp}, more power is diverted to the ICRH, with concomitant reductions in thrust. We are also studying two other exhaust variation techniques, including a magnetic choke at the nozzle throat for high I_{sp} and a plasma afterburner for high thrust at very low I_{sp}.

A key consideration is the efficiency of the engine over its operating range. Creating a hydrogen plasma costs about 40 electron volts per electron-ion pair. (An electron volt, or eV, is a unit of energy commonly used in particle physics.) This energy expenditure is not available for propulsion; most of it is frozen in the creation of the plasma. The particles' kinetic energy, which is what ultimately generates the thrust, must be added to this initial investment. In the early prototype of VASIMR, at high I_{sp}, the kinetic energy is about 100 eV per ion. Thus, a total energy expenditure of 140 eV yields 100 eV of useful energy, or about 70 percent efficiency. Later VASIMR designs will reach exhaust energies of 800 to 1,000 eV for the same initial investment, leading to greater efficiency. At low I_{sp}, as more plasma is generated for higher thrust, the kinetic energy per particle gets uncomfortably close to the ionization energy, with consequent reductions in efficiency. In the end, however, efficiency must be evaluated in the context of the overall mission. Sometimes brief bursts of high thrust may in fact be the most efficient approach.

The gas ionization involves other inefficiencies. Neutral atoms lingering in this initial plasma cause unwanted power losses if they remain mixed with energetic ions. In an effect known as charge exchange, a cold neutral atom gives an electron to a hot ion. The resulting hot neutral is oblivious to the magnetic field and escapes, depositing its energy on nearby structures. The cold ion left behind is virtually useless.

To avoid this, we are studying a radial-pumping technique, in which the cold neutrals are siphoned out before they wander into the power-amplification stage. They may be reinjected downstream of the nozzle throat, where the ions are already moving in the right direction and charge exchange actually helps the plasma to detach from the rocket. Charge exchange is a serious problem in fusion research today.

Although the helicon is able to ionize nearly any gas, practical considerations favor light elements such as hydrogen and helium. For example, the ICRH process is easiest in light gases, whose cyclotron frequencies at reasonable fields (about one tesla) are compatible with existing high-power radio technology. Fortunately, because hydrogen is the most abundant element in our universe, our ships are likely to find an ample supply of propellant almost everywhere. Another important engineering challenge is the generation of strong magnetic fields. We are investigating new high-tempera-

POWER RICH

In space, power is life. In the spring of 1970 the *Apollo 13* astronauts managed to stay alive by judicious use of their precious battery power. Had their return flight taken longer, they would have met with disaster. The electrical requirements of human spaceflight are set by basic survival needs: rapid transportation and life support. The space shuttle consumes about 15 kilowatts in orbit; the International Space Station, 75 kilowatts. Estimates for a Mars habitat range between 20 and 60 kilowatts—not including propulsion. For a baseline Mars mission, a VASIMR engine would require about 10 megawatts. Higher power means faster transits. A 200-megawatt VASIMR would get to Mars in 39 days.

For human forays into near-Earth space, chemical fuels and solar panels provide sufficient power. But for Mars and beyond, chemical fuels are too bulky and the sun's rays too weak. A 10-megawatt solar array, for example, would be about 68,000 square meters at Mars and 760,000 at Jupiter. Such gigantic panels are impractical; in comparison, the solar panels on the space station are 2,500 square meters. There is only one source up to the task: nuclear.

In the past, nuclear electricity has generally been obtained from "nuclear batteries"—radioisotope thermoelectric generators (RTGs), which rely on the heat generated by the natural radioactive decay of plutonium. Such devices have proved crucial to robotic space missions but are too inefficient for human flight. Far better would be a nuclear reactor, which relies on the fission of uranium in a chain reaction. For each kilogram of fuel, a reactor produces up to 10 million times more power than an RTG does.

To measure the performance of power sources, space engineers use a parameter called alpha, which is the ratio of power plant mass in kilograms to electrical output in kilowatts. Low alphas correspond to high efficiency and high power. Present solar arrays, operating near Earth, have alpha values of 20 to 100. RTGs manage an alpha of 200. But for uranium reactors, alphas can be as low at 0.5.

Reactors are inherently safer than RTGs, because the reactor and fuel can be launched separately and assembled in an orbit far from Earth. Even critics who call for a ban on nuclear power in Earth orbit have acknowledged its importance for deep-space missions [see "Nuclear Power in Space," by Steven Aftergood et al.; SCIENTIFIC AMERICAN, June 1991]. *—F.R.C.D.*

ture superconductors, based on bismuth strontium calcium copper oxide compounds. The magnets will use the cryogenic hydrogen propellant to cool them.

Early VASIMR experiments, which I began in the 1980s with Tien-Fang Yang and others at the Massachusetts Institute of Technology, have led to the present research program. The centerpiece today is the VX-10 prototype at the National Aeronautics and Space Administration Johnson Space Center in Houston. Two smaller experiments at Oak Ridge National Laboratory and the University of Texas support the investigations. We also have partnerships with Rice University, the Princeton Plasma Physics Laboratory, the University of Michigan, the University of Maryland and the University of Houston, as well as with private industry and with NASA centers in Huntsville, Ala.; Cleveland; Greenbelt, Md.; and Norfolk, Va.

We are now formulating plans to test VASIMR in space. A proposed mid-2004 demonstration involves a 10-kilowatt solar-powered spacecraft, which will also study Earth's radiation belts. In another test, a VASIMR engine will try to neutralize the atmospheric drag on the International Space Station. Recent experimental results and rapid progress in the miniaturization of radio-frequency equipment bode well for such space tests.

As a natural progression, a possible VASIMR human Mars mission involves a 12-megawatt system. The ship would climb on a 30-day outward spiral from Earth and cruise through interplanetary space for 85 more days, accelerating much of the way and then decelerating for arrival at Mars. The trip would be twice as fast as one involving chemical rockets. A crew module would detach and land using chemical rockets, while the mother ship would fly by the planet in a fuel-efficient trajectory to rejoin it four months later. To protect the human crew, the Mars vehicle would be provided with a robust abort capability by virtue of its variable exhaust. Its magnetic field and hydrogen propellant would act as a radiation shield.

VASIMR could serve as a precursor to the great dream of those of us in the space program: a fusion rocket. Such a ship would have 10 to 100 gigawatts at its disposal. Although controlled fusion remains elusive, the efforts to achieve it have been relentless and the progress steady. Future generations will use it for rapid access to the planets and beyond. We now find ourselves preparing the groundwork for achieving that vision. SA

The Author

FRANKLIN R. CHANG DÍAZ was scheduled to make his first spaceflight on the ill-starred *Challenger* in January 1986, but at the last minute his team was moved up one flight. Since then, he has flown five missions, highlights of which include the deployment of the Galileo probe to Jupiter, the first two tests of space tethers and the final shuttle-Mir docking. In addition to his astronaut career and his plasma research (in which he has a Ph.D. from the Massachusetts Institute of Technology), he has worked in mental health and drug rehabilitation programs.

Further Information

ELECTRICAL PROPULSION IN SPACE. Gabriel Giannini in *Scientific American,* Vol. 204, No. 3, pages 57–65; March 1961.

ION PROPULSION FOR SPACE FLIGHT. Ernst Stuhlinger. McGraw-Hill, 1964.

THE DEVELOPMENT OF THE VASIMR ENGINE. F. R. Chang Díaz et al. in *Proceedings of the International Conference on Electromagnetics in Advanced Applications*, paper 99-102. Torino, Italy, Sept. 13–17, 1999. For copies of this paper and the next, visit spaceflight.nasa.gov/mars/technology/propulsion/aspl on the World Wide Web.

THE PHYSICS AND ENGINEERING OF THE VASIMR ENGINE. F. R. Chang Díaz, J. P. Squire, R. D. Bengtson, B. N. Breizman, F. W. Baity and M. D. Carter. American Institute of Aeronautics and Astronautics, conference paper 2000-3756; July 17–19, 2000.

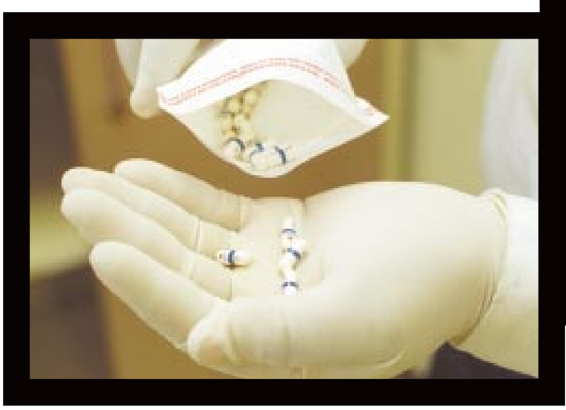
AIDS Drugs for Africa

by Carol Ezzell, *staff writer*
Photographs by Karin Retief/Trace Images

The client, a pretty woman in her 30s with a cap of short, straight hair, takes the small plastic envelope of pills and tucks them into her purse. It is only when she stands up to give her counselor an emotion-filled hug that one can see the reason for her visit to the clinic: a bulge underneath her dress announces her pregnancy, which is now in its eighth month.

Her counselor and the doctors and nurses at this facility in Khayelitsha, a black township outside Cape Town, South Africa, are the only ones who know that the woman, who here will be called Millicent, is infected with the human immunodeficiency virus (HIV). The father of her child is no longer in her life, and she says that her mother—who is dependent on her—would “die” if she knew. Millicent is afraid to tell friends because if her secret leaked out, she could be killed: just days before Millicent’s clinic visit, a 25-year-old woman in Soweto had been found shot through the head with a brown-paper note next to her body bearing the hand-scrawled message “HIV Positive AIDS.” And that was not the first such incident.

Despite the stigma and secrecy, Millicent has hope for the future—or at least her baby’s future—in the form of a bag of white capsules with blue stripes, the antiretroviral drug AZT. Millicent is part of a pilot program administered by the French organization Médecins Sans Frontières (Doctors Without Borders) to use anti-



Most of the 35 million people infected with the AIDS virus live on the African continent, where drugs that fight the virus are rare. Will the world let them die?

COUNSELOR PATRICIA QOLO hugs a client, here called Millicent (*seen from behind, at left*), after discussing how Millicent can protect her unborn baby from contracting human immunodeficiency virus (HIV). Millicent, who is HIV-positive, lives in the township of Khayelitsha (*above*) outside Cape Town, South Africa. She is taking part in a pilot program to receive AZT (*far left*) to prevent mother-to-child transmission of HIV.

retrovirals to prevent HIV-infected mothers from passing on the virus to their unborn babies. Mother-to-child transmission is one of the driving forces behind the AIDS pandemic in southern Africa. According to a report by the South African Department of Health, nine out of 10 children in the world who are infected with HIV live in sub-Saharan Africa.

Antiretroviral drug regimens for preventing HIV-infected mothers from passing on the virus to their infants in the womb or at birth are now routine practice in the U.S. and most developed countries. But the Médecins Sans Frontières clinic is one of only a small number of programs to offer them in all of Africa. Even fewer clinics and hospitals on the continent have enough drugs and procedures in place to provide anti-HIV medicines to the rest of the population of infected men, women who aren't pregnant, and children.

The high cost of the drugs is a critically important factor in preventing AIDS treatments from reaching areas of the world such as Africa that have the greatest need for them. Drug prices overall have been a flash point in the run-up to this month's presidential elections in the U.S. As this issue of *Scientific American* went to press, five global pharmaceutical companies were meeting with the World Health Organization (WHO) in the latest in a series of talks to discuss making their AIDS drugs more affordable for poor nations. As part of the process, the companies—Boehringer Ingelheim, Bristol-Myers Squibb, Glaxo Wellcome, Merck & Company and Hoffmann–La Roche—offered this past May to cut their prices for antiretroviral drugs to Africa by 80 percent, although critics have derided the effort as too little too late.

Pharmaceutical prices aren't the only impediments keeping more Africans from receiving the drugs, many argue. "I wish the problem of AIDS treatment around the world could boil down to high drug prices," says David E. Bloom, professor of economics and demography at the Harvard University School of Public Health. He adds that AIDS "lies at the intersection of some of the most important issues of our day," including poverty, globalization and lack of health care infrastructure.

Beyond drug costs is the testing necessary to measure the success of antiretroviral drugs in individual patients and to monitor the emergence of viral resistance. These tests are expensive, require frequent doctor visits and can be difficult for laboratory technicians to perform. In addition, few leaders of sub-Saharan African countries have supported the use of antiretroviral drugs: indeed, South African President Thabo Mbeki continues to claim that HIV cannot be the sole cause of AIDS. He has also asserted that AZT and other antiretrovirals are too toxic for pregnant African women, even though studies elsewhere in the world have shown that the drugs can cut the odds of mother-to-child transmission by half, with few side effects.

Besides these obstacles, unsettling questions about the ramifications of making antiretroviral therapy widely available could make even a Good Samaritan squirm. Who will care for all the orphans if a country can afford to give antiretroviral drugs only to pregnant women who will later die along with the babies' fathers? And if people in developing countries don't take their medicines consistently and correctly, could they breed super-resistant strains of HIV that could make their way to the developed world?

Living in the Real World

Such questions are merely hypothetical in the shanties of corrugated metal, scrap wood and plastic sheeting that stretch as far as the eye can see from the highway ramp to Khayelitsha. Uncounted tens of thousands live in dire poverty in this apartheid-era relic on the outskirts of Cape Town. More flock here from the rural areas every day to build 8-by-10-foot "informal dwellings" while they search—mostly without success—for work. Instead many get HIV, and only a very lucky few can hope to do anything about it.

The AZT in Millicent's purse was paid for by the South African government, purchased from producer Glaxo Wellcome for the discounted price of nearly three rand ($0.40) per 100-milligram capsule. That's less than the cost of a can of Coke in Cape Town; in New York City, a single AZT pill can go for upward of $5.

Médecins Sans Frontières has so far administered AZT to 1,400 HIV-positive pregnant women in Khayelitsha as part of a pilot program. The first babies were born in March 1999, so the physicians cannot yet determine the program's effectiveness. (The earliest that meaningful antibody tests can be conducted on infants is at 18 months.)

Another drug, nevirapine, may turn out to be just as effective as AZT, and its maker has offered it to Africa and to other poor regions for free. Last July, Boehringer Ingelheim in Ingelheim, Germany, announced it would provide the drug without charge for a period of five years to "developing economies." At the XIII International AIDS Conference, which was held in July in Durban, South Africa, the company reported that nevirapine (sold under the trade name Viramune) reduced the rates of mother-to-child transmission to 14 percent among a group of 652 pregnant women. Ten percent of babies born to a similar number of women taking a combination of AZT and a drug called 3TC, or Epivir, became infected.

The nevirapine regimen requires only a total of three doses of medicine: two to the mother (one during labor and another a day or two after delivery) and one to the newborn. In contrast, AZT must be taken for months. So far, however, South Africa—which has one of the highest proportions of HIV-infected residents in the world—has not taken up Boehringer Ingelheim on its offer, although the provincial government of the Western Cape is expanding the AZT program pioneered by Médecins Sans Frontières to five additional townships.

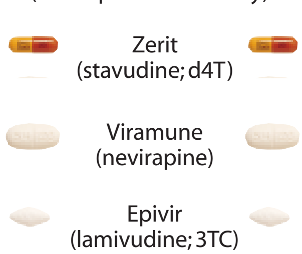
"We don't believe that the only way to manage mother-to-child-transmission is antiretroviral medication," contends South African Health Minister Manto Tshabalala-Msimang. Instead the country's plan for addressing HIV and AIDS through 2005 emphasizes prevention, treatment for opportunistic infections and support for home-based care. "We are beginning to flesh out a program to target households [affected by AIDS] and ask how can we support them," she explains. In August, South Africa and Namibia turned down an offer to 24 sub-Saharan countries from the Export-Import Bank of the United States to provide $1 billion in annual loans for buying anti-AIDS drugs, saying they were burdened with enough debt.

The reluctance of the South African government to use antiretroviral drugs to prevent babies from being born with HIV baffles and saddens many within and outside the country. Even if South Africa chose not to accept Boehringer Ingelheim's donation of nevirapine, it could buy the drug for $4 per pregnant woman. "I not only weep, I feel a sense of injustice and wrong" over the fact that nevirapine is not being given to all pregnant South African women with HIV, Judge Edwin Cameron of South Africa's Constitutional Court—who is himself HIV-positive—proclaimed in a speech at the Durban AIDS conference. The government's stance has also

SWALLOW THIS: THREE DAYS OF ANTIRETROVIRAL DRUGS

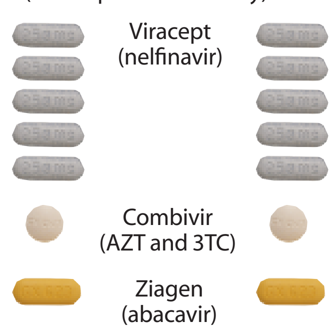
The 16 leading antiretroviral drugs now on the market in the U.S. are used in a dizzying array of combinations, depending on factors such as whether a patient has taken any of them before, the side effects the patient has experienced, and the other conditions he or she has. Multiple-drug combos have been shown to be more effective than one or two drugs used on their own. Below is a selection of three typical drug regimens, showing the number of pills per day a person on each regimen would be required to take. (Nucleoside analogues and nonnucleoside reverse transcriptase inhibitors prevent HIV from copying its genetic material; protease inhibitors block the production of new viral proteins.) Which regimens might best be suited for conditions in Africa—and affordable for use there—is still under debate. —*C.E.*

REGIMEN 1
(three pills twice a day)



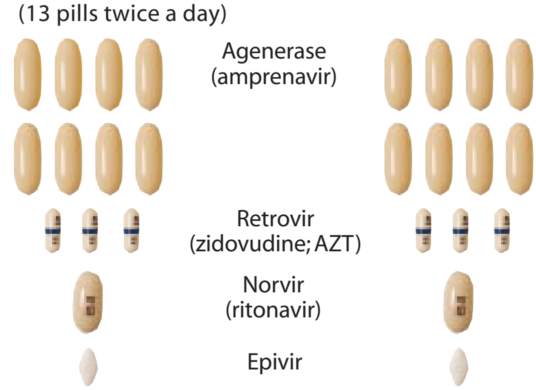
One nonnucleoside reverse transcriptase inhibitor (Viramune) plus two nucleoside analogues.

REGIMEN 2
(seven pills twice a day)



One protease inhibitor (Viracept) plus three nucleoside analogues. Viracept must be taken with a meal or a light snack.

REGIMEN 3
(13 pills twice a day)



Two protease inhibitors (Agenerase and Norvir) plus two nucleoside analogues. Norvir must be refrigerated or kept cool and be taken with a meal. Agenerase should not be taken with meals high in fat.

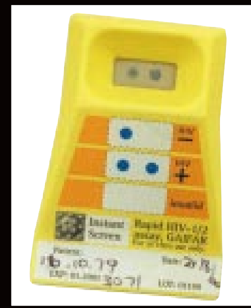
DAN WAGNER; PILLS COURTESY OF IRENE FRANCK *Center for AIDS Research, New York University School of Medicine* prompted legal action from a Durban-based group called Treatment Action Campaign (TAC). The organization has filed a legal proceeding to compel South Africa to provide antiretrovirals to every pregnant woman infected with HIV.

But TAC is against free drugs on the grounds that pharmaceutical company donor programs are not sustainable: they are usually limited to a specific period and sometimes carry restrictions on what the drug can be used for. TAC has exchanged letters with executives at Pfizer, for instance, in attempts to get the company to slash its price for fluconazole, which is popular in the U.S. as a single-dose treatment for vaginal yeast infections. Fluconazole, or Diflucan, is not an antiretroviral drug, but it is the only medicine effective against cryptococcal meningitis, a potentially fatal opportunistic disease that strikes people with AIDS. It is also the drug of choice for treating thrush, yeast infections of the mouth and esophagus that can make it difficult and painful for someone with AIDS to swallow.

A 200-milligram dose of fluconazole can cost up to $20 in the U.S., but generic drug manufacturers in Thailand are now producing it for $0.60 per treatment dose ($0.30 per maintenance dose). Pfizer has offered to provide the drug for free to patients who can't afford it, but only in South Africa and only for cryptococcal meningitis, not thrush. TAC objects to the restrictions; it worries that the company will continue to donate the drug only until its patent expires in a little more than two years and that patients will then be left paying an uncertain price. "What we want is what we asked for—a price reduction" to the Thai price or lower, maintains Zackie Achmat of TAC. Failing that, TAC is petitioning Pfizer to grant a voluntary patent license enabling South Africa to import generic fluconazole.

In a letter to TAC, Pfizer responded that it will ensure that patients in the free-drug program receive the medicine "for as long as they need it" and that the company will "evaluate the success of the program after two years."

NURSE VALERIE SHOSHA draws blood for an HIV test from a pregnant woman in the Khayelitsha Midwife Obstetric Unit near Cape Town (*below*). The woman tested negative and received counseling on how to avoid infection in the future. If a woman tests positive on the initial screening test, the result is confirmed using a second test (*right*), in which two blue dots indicates the presence of antibodies against HIV.



Sticker Shock: It's Contagious

As the debate over fluconazole continues, some countries—principally India, Thailand and Brazil—are producing generic versions of antiretrovirals, including the especially pricey new protease inhibitors, which can cost more than $20 per day for a patient in the U.S. Protease inhibitors block an enzyme that HIV uses to break a large precursor protein into the smaller units it needs to build new copies of itself. The first-line treatment for HIV infection in the U.S. usually includes a protease inhibitor, such as indinavir, plus two drugs that, like AZT, work by shutting down HIV's ability to copy its genetic material. A year's course of treatment with such a combination costs roughly $12,000 in the U.S.

The drug industries of India and Brazil routinely ignore treaties designed to encourage companies from different countries to honor one another's patents. Last May, President Bill Clinton signed an executive order saying that the U.S. government would not interfere if African nations flouted patents held by U.S. pharmaceutical companies by importing generic AIDS drugs made in such countries.

Moreover, the international agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) offers a way for African countries to obtain generic versions of antiretroviral drugs legally. According to TRIPS, which pertains to the 137 members of the World Trade Organization, governments can pass laws on various grounds to override the treaty's patent protections and offer "compulsory licenses" to companies within their borders to enable the firms to make the same products for the domestic market. The companies would have to pay the patent holder royalties, but in most instances they could still sell the products more cheaply.

The pharmaceutical companies are fighting the compulsory licensing idea, which they say will erode the significance of patents, cut into research funding, and possibly lead to poor, and perhaps dangerous, reproductions of their products. Instead the five companies in negotiations with WHO—which together make most of the antiretrovirals licensed for sale in the U.S.—appear to be sticking to their proposal for dropping their prices by 80 percent.

That isn't good enough for David Berman of Médecins Sans Frontières, who would like to see drug companies sell the medicines at just slightly above cost. An 80 percent reduction would bring the cost of a year's worth of a combination of antiretroviral drugs from roughly $12,000 to $2,400, whereas the same medicines in generic form imported from Brazil would come to only $1,000. (To put that into perspective,

Cost Comparison of Selected AIDS Drugs

Drug Name	U.S. Price (wholesale)	South Africa Price	"Best" Price	U.S. Price as Multiple of "Best" Price
Didanosine (ddI; Videx)	$1.80	$0.70	$0.50 (Brazil)	3.6×
Efavirenz (Sustiva)	$4.40	$2.40*	$2.30 (Brazil)	1.9×
Fluconazole (Diflucan)	$12.20	$4.10*	$0.30 (Thailand; maintenance dose)	40.6×
Lamivudine (3TC; Epivir)	$4.50	$1.10	$0.50 (India)	9×
Nevirapine (Viramune)	$4.90	$3.00*	$2.10 (India)	2.3×
Stavudine (d4T; Zerit)	$4.90	$2.50*	$0.30 (Brazil)	16.3×
Zidovudine (AZT; Retrovir)	$1.70	$0.40	$0.20 (Brazil/India)	8.5×
Combivir (lamivudine plus zidovudine)	$9.80	$1.50	$0.70 (Brazil)	14×

*Price the South African government pays to the pharmaceutical company—it is much lower than the price charged to the private sector.

SOURCE: Médecins Sans Frontières/Doctors Without Borders

consider that developing countries generally spend $4 per person each year on health care.) He also doesn't buy the quality argument or the assertion that cheap drug prices for Africa and other developing countries would hinder the search for new pharmaceuticals. "Right now Africa represents only 1 percent of the global pharmaceutical market" because the drugs are priced so high, he contends. "So price reductions here shouldn't affect research."

The 80 percent proposal doesn't impress health minister Tshabalala-Msimang, either. "I have a budget of two billion rand to treat the health issues of the entire country," she says. She calculates that an 80 percent price reduction for a typical course of antiretroviral drugs would work out to 17,000 rand per person per year. Even if she used her budget exclusively to buy the medicines—which would, of course, be impossible to do—she would be able to purchase only enough for roughly 120,000 individuals. According to her department's estimates, South Africa currently has 4.2 million people living with HIV—more than 10 percent of the population.

Harvard's Bloom also calls the behind-closed-doors negotiations of the five drug companies and WHO "shortsighted." As an alternative, he proposes tiered pricing structures according to a country's wealth and the formation of regional buying clubs so nations could buy medicines en bloc and negotiate better prices. Rich countries would end up footing the bill for the R&D, whereas poorer ones would pay only as much as the drugs cost to make.

U.S. Markets for Leading HIV Antiviral Drugs

Reverse Transcriptase Inhibitors (in millions)

Combivir (lamivudine plus zidovudine)	$478.40
Zerit (d4T; stavudine)	$315.90
Epivir (3TC; lamivudine)	$260.20
Sustiva (efavirenz)	$178.60
Viramune (nevirapine)	$108.10
Ziagen (abacavir)	$107.50
Videx (ddI; didanosine)	$78.60
Retrovir (AZT; zidovudine)	$55.00
Hivid (ddC; zalcitabine)	$9.70
Rescriptor (delavirdine)	$7.40
Retrovir IV	$0.70
Total	$1,600.10

Protease Inhibitors (in millions)

Viracept (nelfinavir)	$440
Crixivan (indinavir)	$234
Norvir (ritonavir)	$101
Fortovase (saquinavir soft capsules)	$80
Agenerase (amprenavir)	$48
Invirase (saquinavir hard capsules)	$30
Total	$933

For 12-month period ending February 2000

SOURCE: Chemical Market Reporter, *April 17, 2000*

Beyond Drug Prices

But drug costs aren't the only significant contributors to the bottom line when it comes to AIDS treatment. "If it were simply a matter of making drugs available, people would already have them," asserts Jeffrey Sturchio, executive director of public affairs for Europe, the Middle East and Africa for Merck. Merck sells the protease inhibitor indinavir, which is trade-named Crixivan. "Access to antiretrovirals is just one part of a very complex situation," agrees Stefano Vella, president of the International AIDS Society and chair of the HIV Clinical Research Program at the Italian National Institute of Health in Rome.

Because HIV mutates readily, a given set of antiretroviral drugs doesn't work for an individual forever. Patients must be monitored so they can switch to another drug regimen when the viral strain infecting them becomes resistant. Monitoring involves tests to measure a patient's number of CD4 cells—the immune cells targeted by HIV—as well as assays for viral load, the concentration of HIV in the person's blood.

These tests can be expensive; in India, for example, they run about $1,000 per patient per year. But Praphan Phanuphak of the Thai Red Cross AIDS Research Center in Bangkok claims that viral load monitoring is less important than some have asserted and that CD4 testing can be done more cheaply by diluting the laboratory reagents required.

In the meantime, physicians in the U.S. and Europe are increasingly turning to tests of the genotype, or genetic makeup, and phenotype, or susceptibility to various drugs, of a person's particular strain of HIV to determine which antiretrovirals will work best for that individual. Genotyping involves searching a patient's HIV samples for mutations known to render the virus resistant to certain medications. Phenotyping, which is even more difficult and expensive, requires growing the virus in the laboratory in the presence of a panel of drugs to find out which works best.

Many laboratory technicians in the developing world, however, lack the equipment and training to perform CD4 and viral load tests, much less genotyping and phenotyping assays. This deficiency supports those who argue that Africa is not yet equipped to manage patients on antiretroviral drugs.

"I don't think we can wait for African governments to be able to do viral-load and CD4 tests at each health center," counters Christopher Ouma, who runs a Médecins Sans Frontières clinic in Nairobi, Kenya. Ouma says that if antiretroviral drugs become more affordable for Africa, Asia and other developing regions, the necessary laboratories and staff training will follow. "If drugs are more widely available, we can spur governments to increase infrastructure," he predicts.

In July the Bill and Melinda Gates Foundation announced a $50-million donation to boost the medical infrastructure of Botswana. The grant will be spread over five years and will be coordinated with a matching commitment from Merck in the form of program management and antiretroviral medicines.

Sticking with the Program

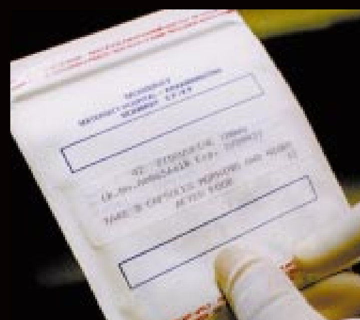
Another hurdle for widespread antiretroviral drug use in Africa is the sheer difficulty of most of the AIDS drug regimens. Under the regimen used to prevent mother-to-child transmission in Khayelitsha, for instance, a woman swallows three AZT pills twice a day during the last two months of pregnancy and then takes the drug every three hours during labor. But to counter her own HIV infection with antiretrovirals, she would need to consume as many as 26 pills every day according to a rigid dosing schedule for years on end [*see illustration on page 100*]. Some medicines must be taken with a meal, some on an empty stomach, some must be kept cool—and missing a single dose can pave the way for the virus to become drug-resistant.

Adhering to such a schedule is hard enough when someone has plenty of food, clean water and electricity for refrigerating the heat-sensitive medications. But between having one's life dictated by a medication regime and the side effects of the drugs, even people in the affluent North cannot or do not adhere to the plan perfectly. People in developing countries will have an even harder time, according to reports from several of the first programs to offer antiretrovirals in such areas.

Papa Salif Sow of the Infectious Diseases Clinic at the University Hospital Center in Dakar, Senegal, reported at this past July's AIDS conference that he was able to get 64 of the 75 patients he was treating with antiretroviral drugs to take 80 percent of their medications, but it required intensive follow-up. And in Brazil, Silvio Aquino of the Municipal Department of Health in Rio de Janeiro said that only 67 percent of patients he followed went for regular CD4 tests, and a mere 40 percent showed up for viral-load testing, meaning they could have undetected viral resistance.

No data on the prevalence of drug resistance in Africa following antiretroviral therapy have been published, according to Christiane Adje of Project Retro-CI in Abidjan, Ivory Coast. But she says that of 68 patients she has treated in her program, which is part of the United Nations AIDS Program's AIDS Drug Access Initiative, 39 have developed genotypic resistance after eight months and 25 have developed phenotypic resistance. (Genotypic and phenotypic resistance can occur in the same patient.) "These findings are very important for Africa," comments John Nkengasong, who is also part of the project. "Before the drugs arrive properly, there is already a lot of resistance circulating."

"We do not have much time to set up a systematic approach" to using antiretrovirals in Africa or patients "will all be resistant," warns Eric Goemaere, director of the Khayelitsha clinic outside Cape Town. He says some of his patients already report having taken a few antiretroviral pills that they have been able to obtain on a sporadic basis on the black market.



BENJAMIN NTWANA, coordinator of the antiviral drug program at the Khayelitsha Midwife Obstetric Unit, talks to a client in the program who gave birth four weeks ago (*left*). The woman, who is HIV-positive, received packets of AZT (*below*) to take during the last few weeks of her pregnancy. The baby will be tested at nine and 18 months to determine whether the drug blocked transmission of the virus.

Goemaere and his colleagues are currently putting together a network of physicians in his area who are now prescribing antiretrovirals and are working to ensure that the doctors have greater supplies of the drugs. He hopes that as people find out about the treatment program—which will start in 2001—more will come forward to be tested for HIV and that the denial and stigma surrounding being HIV positive will diminish. "AIDS did not 'exist' in Khayelitsha last year," Goemaere likes to say. "Now it does." SA

Further Information

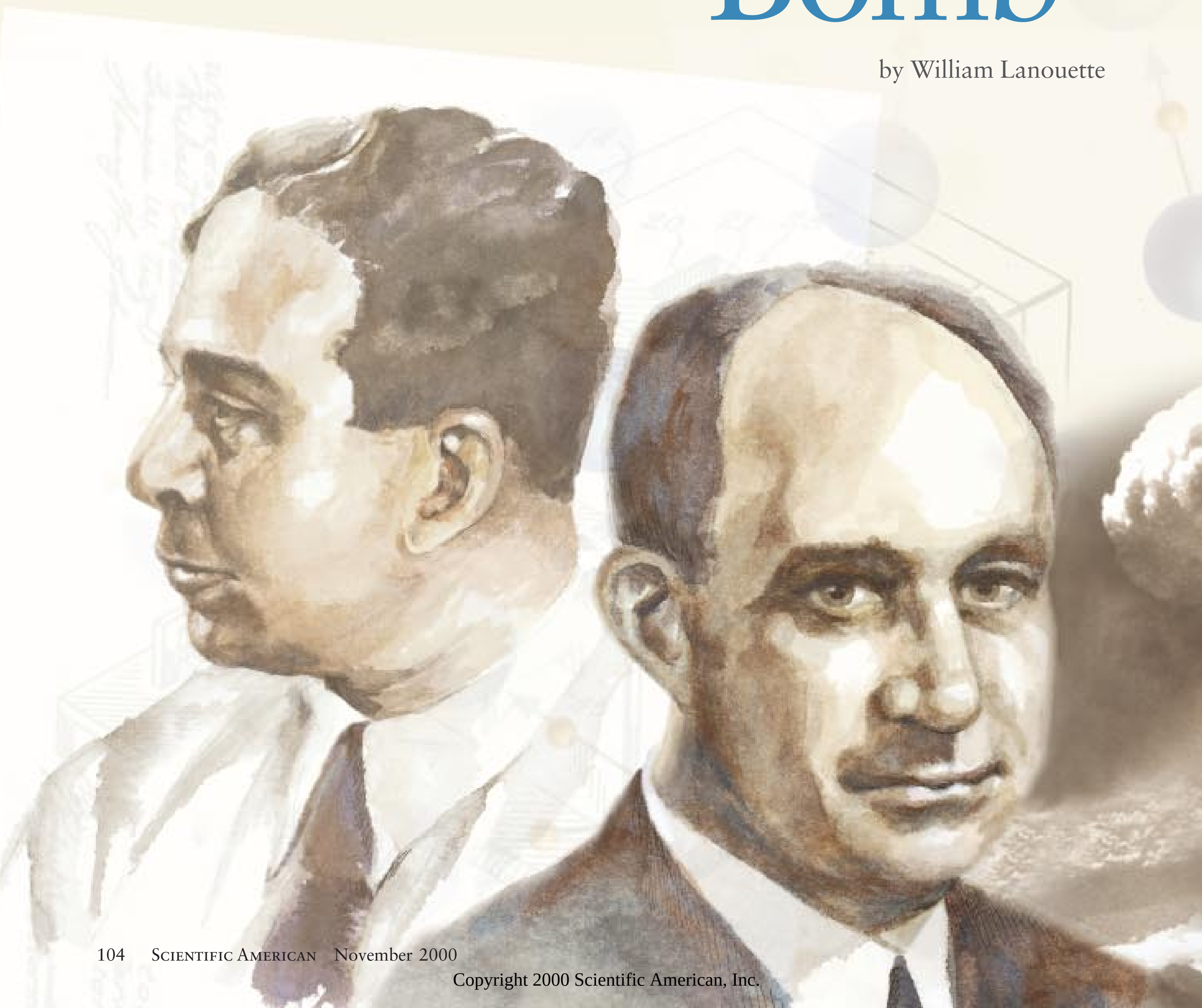
BEYOND OUR MEANS? THE COST OF TREATING HIV/AIDS IN THE DEVELOPING WORLD. Panos Institute, London, 2000.

SOMETHING TO BE DONE: TREATING HIV/AIDS. David E. Bloom and River Path Associates in *Science*, Vol. 288, No. 5474, pages 2171–2173; June 23, 2000.

Like a story by Victor Hugo as told to Neil Simon, the events leading up to the first controlled nuclear chain reaction involved accidental encounters among larger-than-life figures, especially two who did not exactly get along—but had to

The Odd Couple and the Bomb

by William Lanouette



On the eve of World War II, European physicists Enrico Fermi and Leo Szilard both moved into the King's Crown Hotel, near Columbia University in New York City. Although they had previously exchanged letters, they met by chance at the hotel in January 1939. The encounter led to one of the more colorful—and contentious—partnerships in the history of science.

Each man was a refugee from European fascism, and each possessed essential pieces to the puzzle that would ultimately release the energy of the atom. They quickly realized, however, that a joint effort would require them to overcome deep differences in their worldviews, work styles and basic personalities. Had Fermi and Szilard failed to persevere in their often uncomfortable collaboration, the world's first controlled nuclear chain reaction would not have been developed by 1942, and the Manhattan Project would not have built the first atomic bombs by 1945. As Szilard later reflected, "If the nation owes us gratitude—and it may not—it does so for having stuck it out together as long as it was necessary."

Crossed Paths

The 38-year-old Enrico Fermi had just arrived in New York from Rome. The trip included a stop in Stockholm to receive the 1938 Nobel Prize in Physics, for work in which he had bombarded the element uranium with neutrons, which created new transuranic (heavier-than-uranium) elements. Fearing new racial laws in fascist Italy, Fermi and his Jewish wife decided against returning home. Instead he accepted one of four American offers and took a job at Columbia.

Leo Szilard, a 40-year-old Hungarian Jew, came to New York by a more circuitous route. He left his native Budapest in 1919 for Berlin, where he studied and worked with Albert Einstein. Initially, the two shared some ideas and several patents for an electromagnetic refrigerator pump [see "The Einstein-Szilard Refrigerators," by Gene Dannen; Scientific American, January 1997]; two decades later their relationship would take on vast historical significance.

When Adolf Hitler took power in 1933, the wary Szilard fled to London. That same year, he conceived the idea for a nuclear "chain reaction" that, according to his 1934 patent application, might produce "electrical energy" and possibly "an explosion." Such chain reactions would eventually take place in nuclear power plants and in nuclear weapons. First, however, an element that could foster a chain reaction would have to be discovered. After four years of failed experiments at the University of Oxford and then at the universities of Rochester and Illinois in the U.S., Szilard, too, came to Columbia.

Fermi was a rigorous academic whose life centered on a brilliant physics career; he had little interest in politics. A homebody, he soon moved his family from the King's Crown to a house in suburban New Jersey. He awoke at 5:30 each morning and spent the two hours before breakfast polishing his theories and planning the day's experiments. Rare among 20th-century scientists, Fermi was a gifted theoretical physicist who also enjoyed working with his hands. When not lecturing, he toiled in the laboratory with his dedicated assistants, making and manipulating equipment.

An unemployed "guest scholar" with no classes or lab of his own, the bachelor Szilard rarely taught, published infrequently and dabbled in economics and biology. He lived in hotels and faculty clubs and enjoyed soaking for hours in the bathtub to dream up fresh ideas. (One later inspiration was that the National Science Foundation should pay second-rate

CONTENTIOUS COLLABORATION between Leo Szilard (*opposite page, left*) and Enrico Fermi (*right*) had a multiplicity of consequences, involving Albert Einstein, President Franklin D. Roosevelt and the realization of the atom bomb.


Albert Einstein
Old Grove Rd.
Nassau Point
Peconic, Long Island

August 2nd, 1939

F.D. Roosevelt,
President of the United States,
White House
Washington, D.C.

Sir:

Some recent work by E.Fermi and L. Szilard, which has been
municated to me in manuscript, leads me to expect that the elemen
ium may be turned into a new and important source of energy in th
future. Certain aspects of the situation which has arisen

ILLUSTRATION BY JENNIFER JOHANSEN; AIP EMILIO SEGRÈ VISUAL ARCHIVES (*Szilard and Fermi*); CORBIS (*mushroom cloud*); COURTESY OF ARGONNE NATIONAL LABORATORY (*letter from Einstein to F.D.R.*)

scientists *not* to conduct research.) Szilard read newspapers avidly, speculated constantly about financial, political and military affairs, and always kept two bags packed for hasty escapes from any new eruptions of fascism.

A late sleeper, he often appeared at Columbia only in time for lunch, after which he would drop in on colleagues, posing insightful questions and suggesting experiments *they* should try. "You have too many ideas," future physics Nobel laureate Isidor Isaac Rabi finally said to him. "Please go away."

The late Massachusetts Institute of Technology physicist Bernard Feld worked with Fermi and Szilard as the latter's research assistant at Columbia. He summed up the two men: "Fermi would not go from point A to point B until he knew all that he could about A and had reasonable assurances about B. Szilard would jump from point A to point D, then wonder why you were wasting your time with B and C."

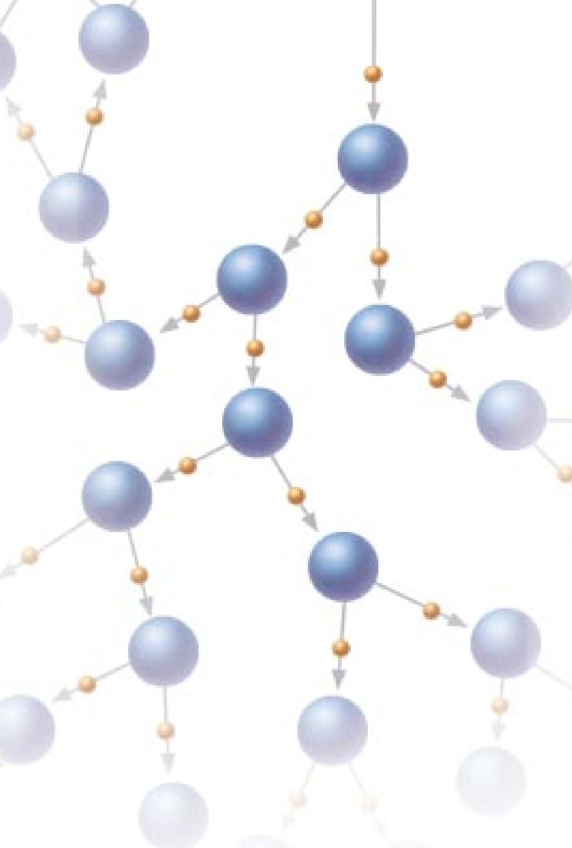
Within days of the chance meeting between Fermi and Szilard at the King's Crown Hotel, Danish physicist Niels Bohr landed in New York with important word from Europe: physicist Lise Meitner, a Jew who had fled from Germany to Stockholm, had determined that Berlin chemists Otto Hahn and Fritz Strassmann had caused uranium to undergo "fission" via neutron bombardment. They had split the atom. (In 1966 the three would win the Enrico Fermi Award for this work.) Bohr's report helped Fermi come to a more complete understanding of his own 1934 uranium experiments; in addition to creating transuranic elements, he had unknowingly split atoms.

To Szilard, the news was more ominous. He realized that uranium was the element that could fuel the chain reaction described in his 1934 patent application. Betting on his political insight, he had assigned that patent to the British Admiralty in secret, lest he alert German scientists to the possibility of atomic explosives. The discovery of fission confirmed Szilard's fears that an atom bomb could soon be a decisive reality.

The notion of the nuclear chain reaction had first come to Szilard while he was standing on a London street corner in 1933. The neutron had been discovered only the previous year, and physicists now thought of the atom as resembling a solar system, with negatively charged electrons orbiting a nucleus of positively charged protons and neutral neutrons. Having no charge, a neutron hurled at an atom might stealthily penetrate the nucleus without being repelled. Szilard imagined that if a neutron hit a nucleus and split the atom, the breakup might release the binding energy that holds the atom together. Some of that atom's neutrons might in turn be released, which could hit and split other atoms. If more than one neutron was released from each split atom, the process could exponentially expand, with millions of atoms splitting in a fraction of a second and freeing vast amounts of energy. (Szilard would later learn that Bohr's news enabled Fermi likewise to envision a chain reaction, although he considered one extremely unlikely.)

While Szilard was filing his patent in 1934, Fermi was in Rome, becoming the world's expert on neutron bombardment of atoms. He found that by passing the neutrons through paraffin wax he could slow them down, increasing the chance that they would be absorbed by the target nucleus. His work with uranium was puzzling. Sometimes the nucleus absorbed neutrons. (Because atomic identity is governed by the number of protons, the neutron absorption produced only heavier variants, or isotopes, of uranium.) But sometimes neutron bombardment created entirely new elements. German chemist Ida Noddack, following Fermi's experiments in journal reports, suggested a chemical analysis of the new species to see if they were the fragments of split atoms. But Fermi, concentrating on the physics of bombardment and absorption, did not pursue the implications of those new elements. Had he done so, he might have recognized nuclear fission years before Meitner.

At Columbia in the spring of 1939, Fermi and Szilard each tried experiments aimed at a better understanding of fission. Szilard offered Canadian physicist Walter Zinn a radium-beryllium neutron source he had just ordered from England. With it, Zinn and Szilard showed that more than two neutrons escaped during fission. Fermi and his assistant Herbert Anderson tried a similar experiment using a more powerful radon-beryllium source, with inconclusive results. Szilard guessed that the source was too strong, enabling some neutrons to pass right through the nucleus and making it hard to know if they were counting neutrons from fission events or merely the original neutrons. Szilard loaned Fermi his English



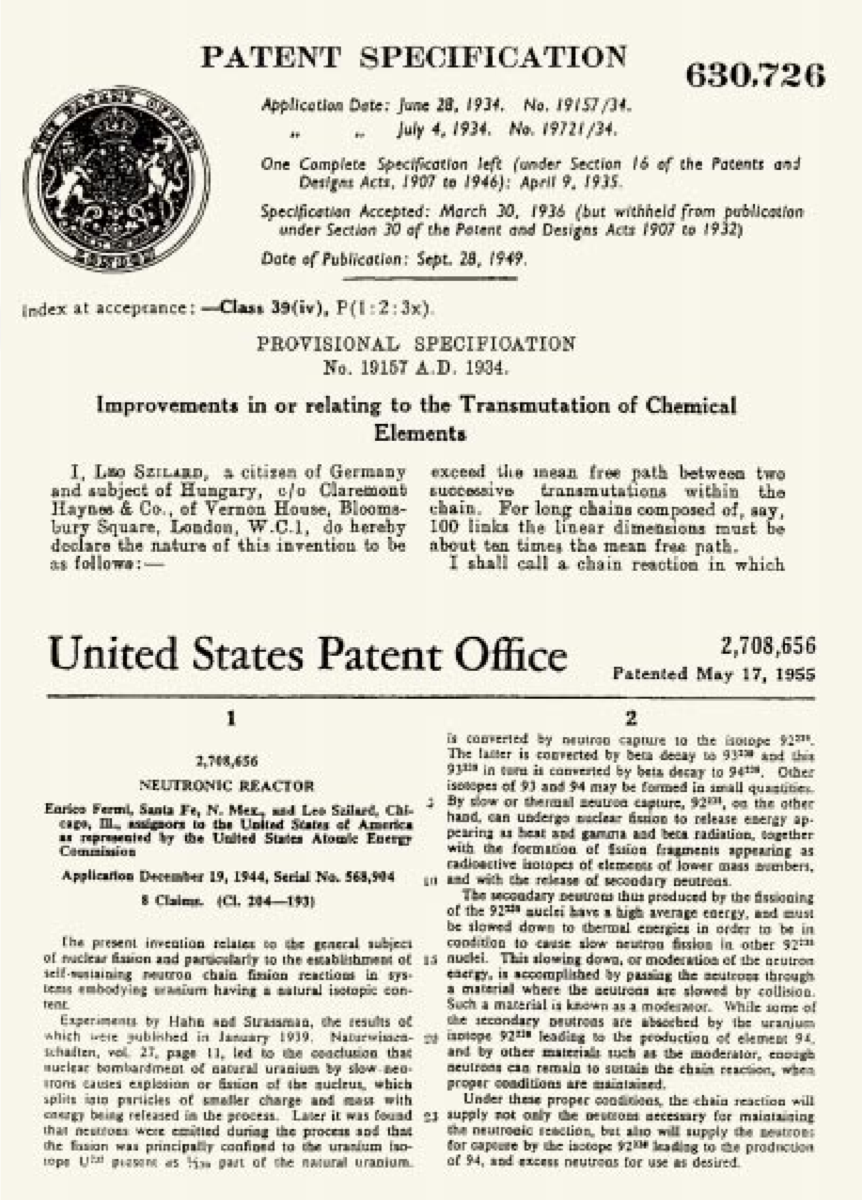
JENNIFER JOHANSEN

neutron source, which gave much clearer results.

The two men then attempted to work together—with a resounding clash of individual styles. Szilard shunned manual labor in favor of brainstorming, but Fermi expected all his team members to participate in hands-on experiments. Although the men respected the other's abilities, they bristled in the other's company. Recognizing their mutual need, however, they reached out to Columbia's physics department chairman, George Pegram, who agreed to coordinate their separate work. Pegram's shuttle diplomacy harnessed Fermi's precision and Szilard's prescience. With Anderson, the combative colleagues succeeded in determining that by using slow neutrons "a nuclear chain reaction could be maintained."

Building the Chain

Although collisions between Fermi and Szilard were all too common, collisions between neutrons and nuclei were at first too rare. Passing the neutrons through so-called moderators, such as Fermi's paraffin, helped to slow them, making their collision with an atom's nucleus more likely. By 1939 physicists also knew that "heavy water" was an efficient moderator. Ordinary, or "light water," consists of two hydrogen atoms and an oxygen atom, the familiar H_2O. In heavy water, two heavy iso-



PATENT SPECIFICATION **630,726**

Application Date: June 28, 1934. No. 19157/34.
" " July 4, 1934. No. 19721/34.

One Complete Specification left (under Section 16 of the Patents and Designs Acts, 1907 to 1946): April 9, 1935.

Specification Accepted: March 30, 1936 (but withheld from publication under Section 30 of the Patent and Designs Acts 1907 to 1932)

Date of Publication: Sept. 28, 1949.

Index at acceptance:—**Class 39(iv)**, P(1:2:3x).

PROVISIONAL SPECIFICATION
No. 19157 A.D. 1934.

Improvements in or relating to the Transmutation of Chemical Elements

I, Leo Szilard, a citizen of Germany and subject of Hungary, c/o Claremont Haynes & Co., of Vernon House, Bloomsbury Square, London, W.C.1, do hereby declare the nature of this invention to be as follows:—

exceed the mean free path between two successive transmutations within the chain. For long chains composed of, say, 100 links the linear dimensions must be about ten times the mean free path.

I shall call a chain reaction in which

United States Patent Office 2,708,656 Patented May 17, 1955

1

2,708,656

NEUTRONIC REACTOR

Enrico Fermi, Santa Fe, N. Mex., and Leo Szilard, Chicago, Ill., assignors to the United States of America as represented by the United States Atomic Energy Commission

Application December 19, 1944, Serial No. 568,904

8 Claims. (Cl. 204—193)

The present invention relates to the general subject of nuclear fission and particularly to the establishment of self-sustaining neutron chain fission reactions in systems embodying uranium having a natural isotopic content.

Experiments by Hahn and Strassman, the results of which were published in January 1939, Naturwissenschaften, vol. 27, page 11, led to the conclusion that nuclear bombardment of natural uranium by slow neutrons causes explosion or fission of the nucleus, which splits into particles of smaller charge and mass with energy being released in the process. Later it was found that neutrons were emitted during the process and that the fission was principally confined to the uranium isotope U^{235} present as 1/139 part of the natural uranium.

2

is converted by neutron capture to the isotope 92^{239}. The latter is converted by beta decay to 93^{239} and this 93^{239} in turn is converted by beta decay to 94^{239}. Other isotopes of 93 and 94 may be formed in small quantities. By slow or thermal neutron capture, 92^{235}, on the other hand, can undergo nuclear fission to release energy appearing as heat and gamma and beta radiation, together with the formation of fission fragments appearing as radioactive isotopes of elements of lower mass numbers, and with the release of secondary neutrons.

The secondary neutrons thus produced by the fissioning of the 92^{235} nuclei have a high average energy, and must be slowed down to thermal energies in order to be in condition to cause slow neutron fission in other 92^{235} nuclei. This slowing down, or moderation of the neutron energy, is accomplished by passing the neutrons through a material where the neutrons are slowed by collision. Such a material is known as a moderator. While some of the secondary neutrons are absorbed by the uranium isotope 92^{238} leading to the production of element 94, and by other materials such as the moderator, enough neutrons can remain to sustain the chain reaction, when proper conditions are maintained.

Under these proper conditions, the chain reaction will supply not only the neutrons necessary for maintaining the neutronic reaction, but also will supply the neutrons for capture by the isotope 92^{238} leading to the production of 94, and excess neutrons for use as desired.

PATENT awarded to Szilard in England for the chain reaction idea was assigned to the British Admiralty and remained secret until after the war. A U.S. patent for the actual reactor was awarded jointly to Fermi and Szilard.

topes of hydrogen, called deuterium, unite with the oxygen. (Heavy water is still used as an effective moderator for natural uranium fuel in today's nuclear reactors, whereas light water is used for enriched uranium fuel.) But heavy water was expensive and scarce. The large-scale experiments that Szilard had in mind would require a more common and affordable moderator. He would discover one that his German counterparts had overlooked.

As Szilard had feared, German atom-bomb research was well under way by the spring of 1939. Both German and American physicists also recognized that graphite—the soft form of carbon that is used as pencil lead—could be a moderator. But German scientists gave up on it because it absorbed too many neutrons; they instead concentrated on heavy water, always in short supply. Szilard, who often personally took trains to Boston or Buffalo to procure raw materials for Fermi's experiments, realized that commercial graphite also contained small amounts of boron—a voracious absorber of neutrons. He ordered custom-made, boron-free graphite, which eventually led to one of the most caustic Fermi/Szilard confrontations.

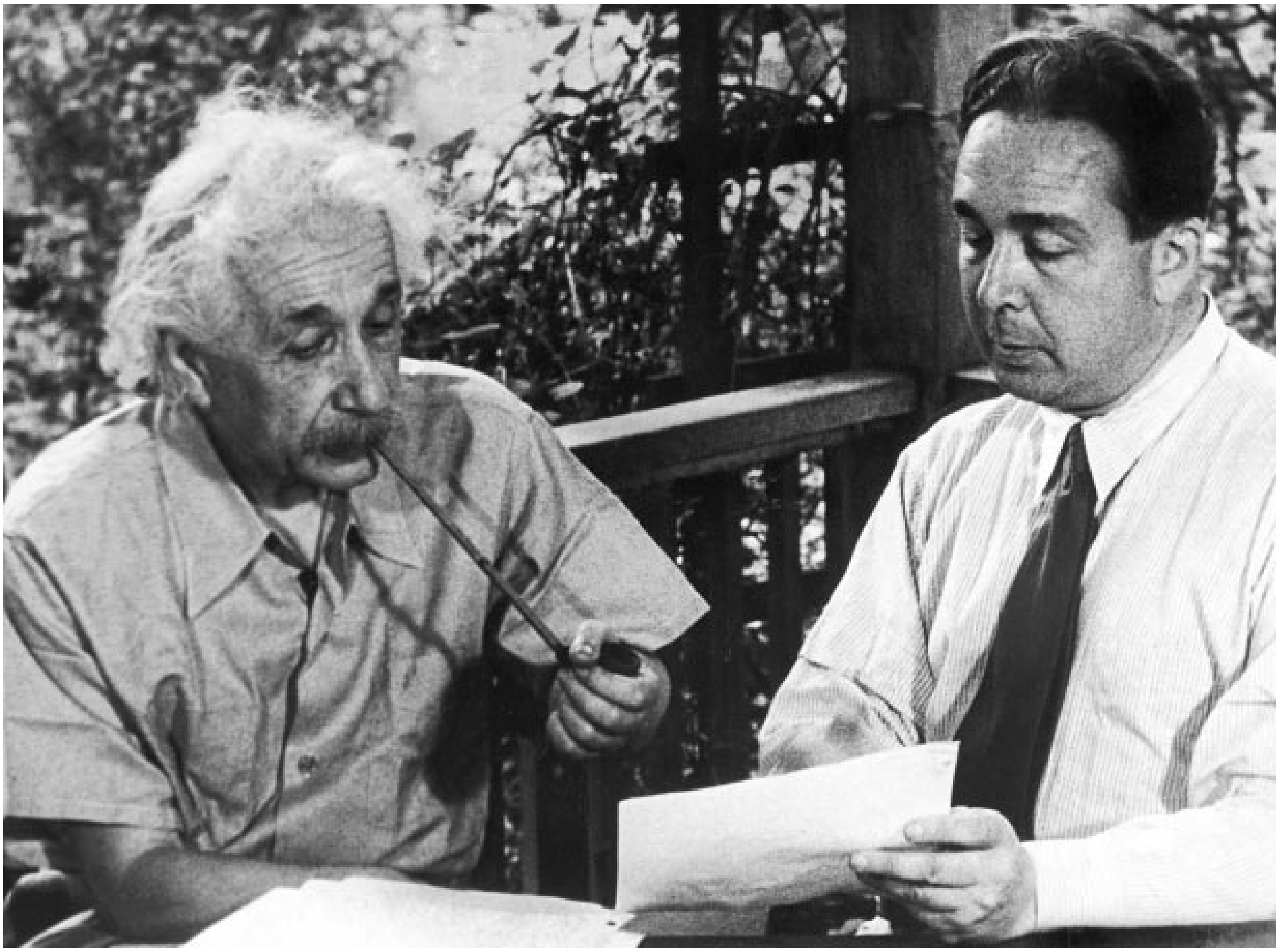
Anderson measured neutron absorption in the pure graphite and found that it would indeed make a good moderator. Szilard recommended that the test results remain secret. Fermi, ever the professional scientist, objected to the breach of the long-standing academic tradition of peer-reviewed journal publication. "Fermi really lost his temper," Szilard would later recall. "He really thought this was absurd." Pegram once again interceded, however, and Fermi reluctantly agreed to self-censorship under these special circumstances.

With the graphite moderator, Fermi thought there might now be at least a ray of hope for a self-sustaining chain reaction. On the question of how realistic that hope was, Fermi and Szilard had also shown distinctly different modes of thinking. Szilard fretted that the Germans were ahead in a nuclear arms race; in the American vernacular that Fermi enjoyed trying out, he reacted to Szilard's speculation with "Nuts!" Fermi thought that any atom bombs were perhaps 25 to 50 years away and told colleagues that actually creating the self-sustaining chain reaction was "a remote possibility" with perhaps a 10 percent chance.

"Ten percent is *not* a remote possibility if it means that we may die of it," Isidor Rabi replied. Szilard noted how differently he and Fermi interpreted the same information. "We both wanted to be conservative," Szilard later recalled, "but Fermi thought that the conservative thing was to play down the possibility that this may happen, and I thought the conservative thing was to assume that it would happen and take the necessary precautions."

These precautions included Szilard borrowing $2,000 to support Fermi's research. Nevertheless, in the summer of 1939 Fermi showed his relative lack of concern over the implications of nuclear research by leaving for the University of Michigan to study cosmic rays. The world's first successful design for a nuclear reactor was thus created neither in a lab nor a library but in letters.

Szilard, typically, urged starting "large scale" experiments "right away." Fermi, typically, remained skeptical. Szilard proposed stacking alternating layers of graphite and uranium in a lattice, the geometry of which would define neutron scattering and subsequent fission events. Fermi countered with a homogeneous design in which the uranium and graphite would be mixed like gravel. The suggestion angered Szilard, who concluded that Fermi preferred it only because it was an easier configuration about which to make calculations. Fer-



"MARCH OF TIME"/TIME PIX

EINSTEIN AND SZILARD confer over the letter that would convince the U.S. to develop nuclear weapons. This photograph of the physicists is actually a re-creation of their meeting, taken in 1946 for the "March of Time" film *Atomic Power.*

mi responded that further reflection had convinced him of Szilard's lattice idea. Once sold, Fermi applied his substantial ingenuity to determining the lattice's physical properties and coordinating the personnel necessary to make a reactor.

Friends in High Places

Szilard recognized that despite his and Fermi's brainpower, they would still need help from important allies for their collaboration to succeed. They would get it from an unlikely trio: Franklin D. Roosevelt, J. Edgar Hoover and Albert Einstein.

During the summer, Szilard learned that Germany was restricting uranium supplies. He assumed that this indicated fission research and wanted to alert the federal government. With the instincts of a public relations expert, he turned to his mentor and friend Einstein, who was living at a summer cottage on Long Island, about 70 miles east of New York City. Szilard told the renowned physicist about the chain reaction. "I haven't thought of that at all," Einstein replied, seeing at last a mechanism that might make real the mass-energy conversion of his famous equation.

Szilard made two visits to Einstein, the second to discuss a letter for him to sign. "Szilard could do anything, except he could not drive a car," recalls his second-trip chauffeur, a fellow Hungarian refugee scientist. "And I could drive a car. And, therefore, I drove Szilard to the summer place.... Einstein was a democrat in that he invited not only Szilard for a cup of coffee but also his driver." Edward Teller was thus present when Einstein, wearing an old robe and slippers, read and agreed to sign the now well known letter to President Roosevelt. The letter, dated August 2, 1939, began, "Some recent work by E. Fermi and L. Szilard...." It proceeded to warn of German atomic weapons research and urged the U.S. to do its own.

Szilard passed the letter to investment banker Alexander Sachs, who was a New Deal adviser and had access to the president. World War II began on September 1, and in October, when Roosevelt finally received the letter, he agreed that some action was needed "to see that the Nazis don't blow us up." To that end, he created a federal Uranium Committee, with Szilard and other émigré scientists as members. Within weeks they had gained a commitment of $6,000 for research at Columbia.

After the war, Einstein said he had "really only acted as a mailbox" for Szilard. In 1940, however, Einstein was once again forced to play a decisive role when the U.S. Army almost denied Fermi and Szilard security clearance. Investigators, basing their conclusions on information from "highly reliable sources," came to the paradoxical conclusions that Fermi, a refugee from fascism, was "undoubtedly a Fascist" and that Szilard, in terror of the Nazis, was "very pro-German." Perhaps Szilard's cries that Germany could win the war accounted for the latter misinterpretation. (The report also spelled Szilard's name in two different ways, both of which were wrong.) The army decided of each man that "employment of this person on secret work is not recommended," despite the fact that the only secret work in question in the U.S. at the time was taking place in the minds of Fermi and Szilard.

Had the army been heeded, of course, funds would have run out, and all the embryonic federal atomic research by Fermi and Szilard would have ceased. This mistake was averted when the Federal Bureau of Investigation, under pressure from the White House, was ordered to "verify their loyalty to the United States." FBI director J. Edgar Hoover sent agents to interview Einstein (whose pacifist views would later cause his own loyalty to be questioned). With Einstein's good word, federal money flowed in to Columbia in November 1940, although suspicions of Fermi and Szilard would abate only years after they became U.S. citizens.

Funding in place, Fermi's team now worked systematically to construct "piles" (Szilard's lattice) of uranium and graphite, to test for the ratio and geometry that would optimize a chain reaction. The day before the Japanese attack on Pearl Harbor, President Roosevelt approved an all-out federal commitment to research the A-bomb. In the spring of 1942 Fermi, Szilard and the rest of the Columbia team moved to the University of Chicago, where they established a top-secret "metallurgical laboratory" for chain-reaction research. The army's Manhattan Project took over control of the effort in June. Ironically, at this same moment in history, Germany scaled down its own A-bomb work, convinced that the undertaking was impractical for the current war.

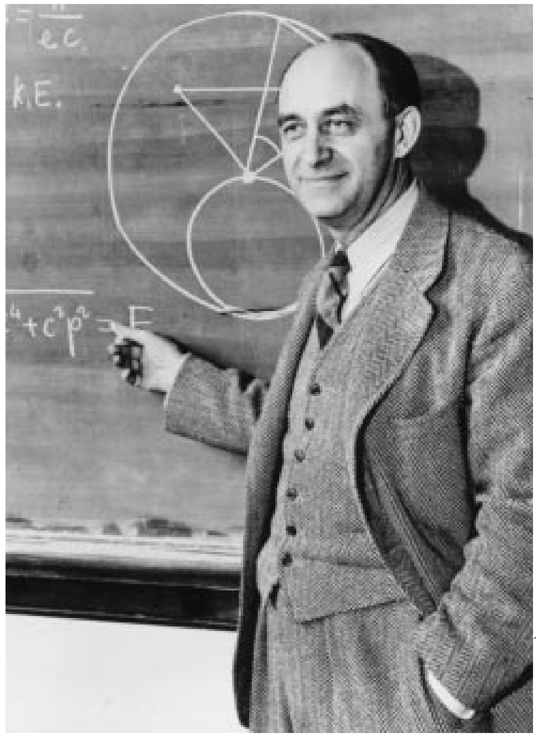
In the fall, a pile was constructed, with uranium spheres embedded in graphite blocks. On December 2, 1942, in a squash court under Stagg Field, the university's football stadium, Fermi directed the experiment that initiated the world's first controlled, self-sustaining nuclear chain reaction. After the historic experiment, Fermi and Szilard found themselves alone with their reactor. They shook hands, Szilard remembered, "and I said I thought this day would go down as a black day in the history of mankind."

Later Conflicts and Harmony

Near the war's end in 1945, Fermi and Szilard differed once again. Szilard had hastened the A-bomb's development as a weapon of defense against Germany. With Hitler's defeat, Szilard argued that the bomb should not be used offensively against Japan but instead be demonstrated to encourage surrender. Fermi, as scientific adviser to the administration's high-level committee on options for bomb use, argued that a demonstration would be impractical. The administration agreed, with the subsequent August devastation of the cities of Hiroshima and Nagasaki.

After the war, Fermi favored continuing army control of atomic research, while Szilard successfully lobbied Congress for a new, civilian Atomic Energy Commission. The two men found common ground in opposition to Szilard's old friend Teller in 1950, when both objected to U.S. development of the hydrogen bomb. Fermi called the H-bomb "a weapon which in practical effect is almost one of genocide."

A joint patent for the Fermi-Szilard "neutronic reactor" was first published in 1955, a year after Fermi's death. Szilard pursued molecular biology and nuclear arms control until his death in 1964. Fermi summed up Szilard by calling him "extremely brilliant" but someone who "seems to enjoy startling people." Szilard reflected on Fermi by writing, "I liked him best on the rare occasions when he got mad (except of course when he got mad at me)." SA



AIP EMILIO SEGRÈ VISUAL ARCHIVES

NOBEL LAUREATE FERMI was both a brilliant theoretician and a gifted experimentalist. Few 20th-century physicists could make such a claim.

The Author

WILLIAM LANOUETTE received a doctorate in politics from the London School of Economics in 1973. His thesis, comparing the use and abuse of scientific information by U.S. and U.K. legislators and government officials, prepared him well for his current work as an energy/science policy analyst at the U.S. General Accounting Office. He has written about atomic energy and science policy for more than 30 years, in such publications as the *Atlantic Monthly,* the *Bulletin of the Atomic Scientists* and the *Economist.* The author of a biography of Leo Szilard, Lanouette has lectured widely about the politics and personalities of the Manhattan Project. He is an avid oarsman, and his next book will be about the lucrative rise and scandalous end of professional rowing in 19th-century America. Lanouette thanks Nina Byers, professor of physics at the University of California, Los Angeles, and independent scholar Gene Dannen for helpful additions to this article.

Further Information

ENRICO FERMI: COLLECTED PAPERS. Vols. 1 and 2. University of Chicago Press, 1962 and 1965.
COLLECTED WORKS OF LEO SZILARD. Vols. 1, 2 and 3. MIT Press, 1972, 1978 and 1987.
GENIUS IN THE SHADOWS: A BIOGRAPHY OF LEO SZILARD, THE MAN BEHIND THE BOMB. William Lanouette (with Bela Silard). University of Chicago Press, 1994.
ENRICO FERMI, PHYSICIST. Reprint. Emilio Segrè. University of Chicago Press, 1995.
An Enrico Fermi Web site can be found at www.time.com/time/time100/scientist/profile/fermi.html
A Leo Szilard Web site is at www.dannen.com/szilard.html

DIAGNOSIS_AT HOME

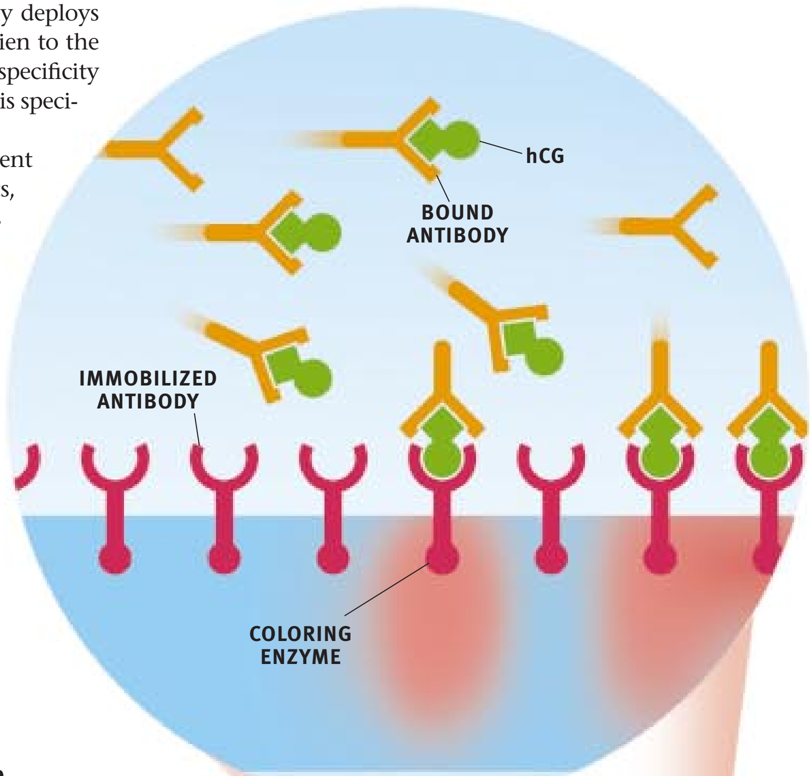
Pregnancy Tests

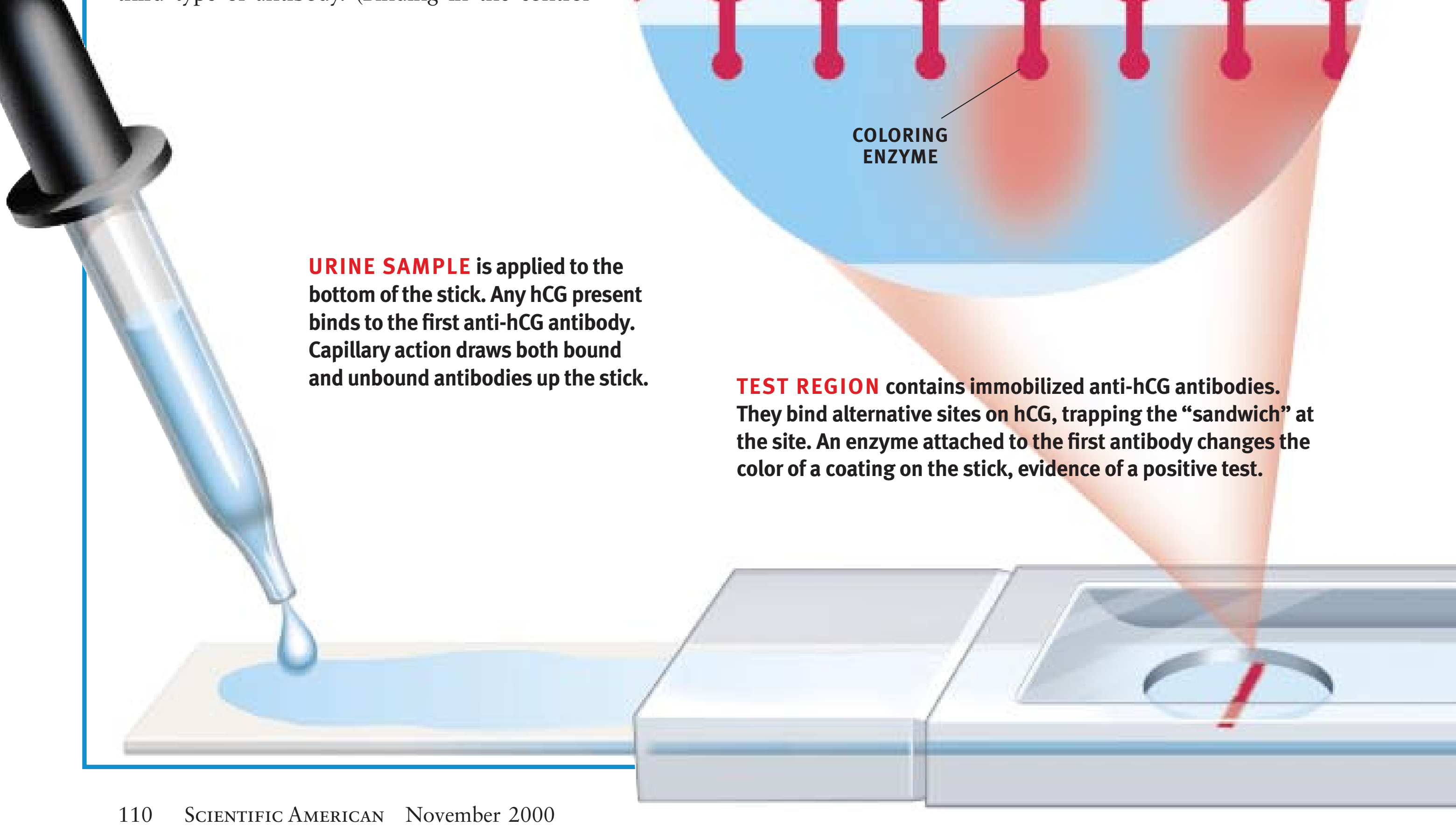
At-home medical diagnosis dates back to the early 18th century and the invention of the enclosed thermometer. Since then, human ingenuity has devised ways to detect an assortment of medical conditions. The market for such devices has flourished since 1977, when the Food and Drug Administration approved the home pregnancy test.

The pregnancy test checks a woman's urine for human chorionic gonadotropin (hCG), a hormone produced by the placenta after egg fertilization. Its concentration in the blood doubles every two to three days, peaking around the eighth week of pregnancy. The technology for the test depends on antibodies—Y-shaped proteins that our immune system normally deploys against invading viruses, bacteria or anything else alien to the body. Every antibody binds tightly and with high specificity only to certain foreign molecules, or antigens, and this specificity is what guides the immune defenses.

Pregnancy tests use a combination of three different types of antibodies, two from mice and one from goats, the first two of which bind to hCG. A urine sample is applied to the bottom of a test stick. The first antibodies flow up the stick as soon as they come in contact with the urine sample, attaching to any hCG that is present. The second antibodies are moored to the strip's test area, and they immobilize the hCG (along with the bound antibodies). Any of the first antibodies that lack hCG continue to move into the control region, where they are grabbed by the third type of antibody. (Binding in the control area verifies that the first antibody moved up the stick properly.)

The molecules of hCG and the antibodies are minute and colorless and so would naturally be hard to detect. But the first anti-hCG antibodies also carry an enzyme that converts a colorless substance plated onto the test and control regions into a colored one. Depending on where the first antibody ends up, the person conducting the test will see a colored line in just the control region (meaning no pregnancy) or a colored line in both the control and the test regions (pregnancy). —*Rebecca Lipsitz, staff writer*

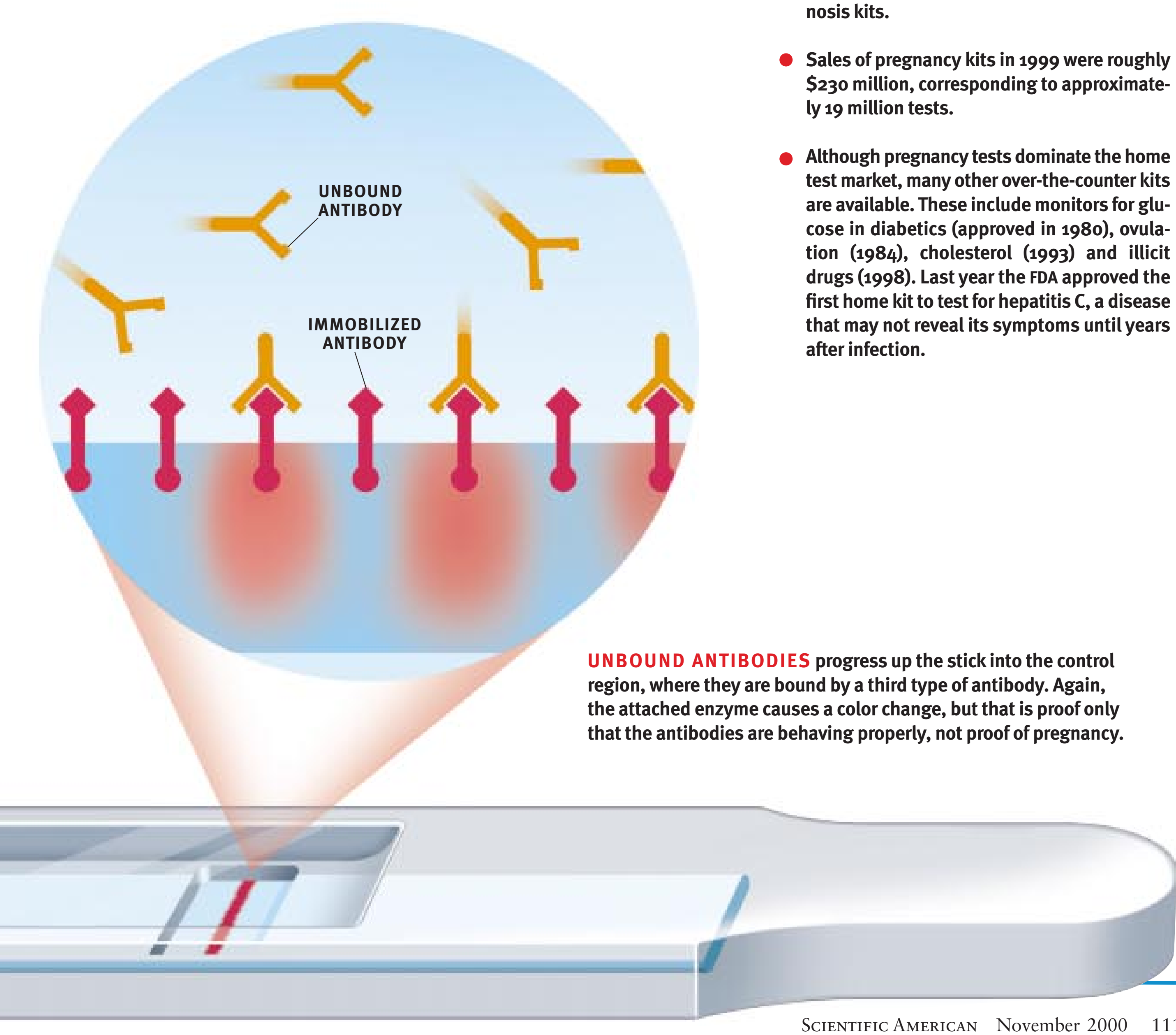




URINE SAMPLE is applied to the bottom of the stick. Any hCG present binds to the first anti-hCG antibody. Capillary action draws both bound and unbound antibodies up the stick.

TEST REGION contains immobilized anti-hCG antibodies. They bind alternative sites on hCG, trapping the "sandwich" at the site. An enzyme attached to the first antibody changes the color of a coating on the stick, evidence of a positive test.

ANTIBODIES, protective proteins made by the immune system, are similar in overall shape but contain slight variations that enable them to recognize and bind to different specific antigens (foreign substances). Antibodies from mice are used in pregnancy kits.



UNBOUND ANTIBODIES progress up the stick into the control region, where they are bound by a third type of antibody. Again, the attached enzyme causes a color change, but that is proof only that the antibodies are behaving properly, not proof of pregnancy.

DID YOU KNOW . . .

- One of the characteristics that distinguishes different pregnancy tests is how soon after egg fertilization they can detect hCG. This depends directly on the minimum threshold levels of hCG needed for a positive test. Many manufacturers claim their tests can detect hCG at a concentration of 25 milli-international-units per milliliter, which appears approximately 10 days after egg fertilization. False negatives often result when hCG levels are below the detection limits of the test. False positives can sometimes occur for women who have received hCG as part of an infertility treatment program.

- Since 1977 the FDA has approved more than 100 different pregnancy tests. The confidentiality offered by these tests has made them the most popular of all types of home diagnosis kits.

- Sales of pregnancy kits in 1999 were roughly $230 million, corresponding to approximately 19 million tests.

- Although pregnancy tests dominate the home test market, many other over-the-counter kits are available. These include monitors for glucose in diabetics (approved in 1980), ovulation (1984), cholesterol (1993) and illicit drugs (1998). Last year the FDA approved the first home kit to test for hepatitis C, a disease that may not reveal its symptoms until years after infection.

ILLUSTRATIONS BY GEORGE RETSECK AND BRYAN CHRISTIE

ARTIFICIAL_LIFE

Boids of a Feather Flock Together

Shawn Carlson explains how to simulate simple organisms on your computer

Scientists sometimes struggle to understand why certain animals act as they do, especially social animals. A school of fish or a flock of birds, for example, behaves in many ways like a single creature. Yet exactly how the individuals organize themselves into a "superorganism" is still very much a mystery.

But believe it or not, these days insights into such self-organizing communities seem to come more often from computer hackers than from field biologists. Many programmers are creating on their desktops virtual environments populated with simulated animals. The nature of these artificial life-forms (or "a-life," for short) usually hinges on a special data string, which is analogous to the DNA blueprint of a living organism. This digital code defines how an a-organism interacts with its cybersurroundings and determines the likelihood that the simulated creature will reproduce.

To mimic real DNA, the cybercode is programmed to experience random mutations, which can alter the fitness of the artificial animal. So by tracking many generations of these byte-size beasties, you can in several minutes watch their digital DNA evolve in ways that might take nature millions of years to accomplish with a real genetic code. With much of this software available online (consult the premier a-life Web site, www.alife.org, or see www.aridolan.com/ad/adb/adib.html for an index of sites where you can download Java scripts), any interested amateur can now plumb the depths of evolution, at least in these virtual worlds.

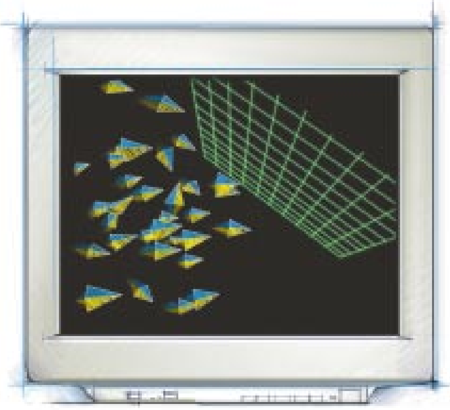
If you take up this challenge, you'll be joining the ranks of people such as Craig Reynolds, now with Sony Computer Entertainment, who in 1986 developed an impressive model of flocking birds. Reynolds speculated that each bird in a flock acts on a simple set of directives. So he programmed his a-life creations, which he whimsically dubbed "boids," to follow just three rules. First, don't get too close to anything, including other boids. Second, try to match your velocity to that of the other boids around you. And third, always move toward the center of the pack of nearby boids.

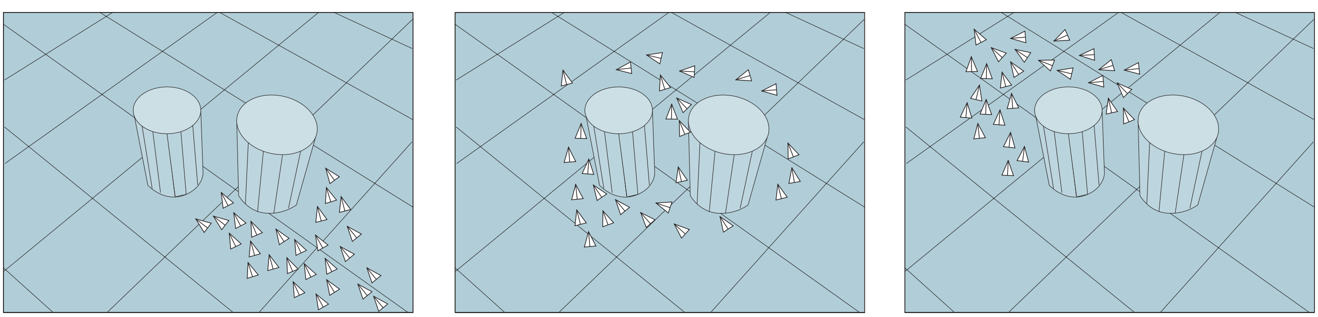
The results of his simulation are remarkable. (Check out www.red3d.com/cwr/boids/ for an eye-popping animation.) No matter how the boids are initially scattered, they quickly form a flock. When the group encounters an obstacle in cyberspace, it splits into two groups and reassembles on the far side.

Reynolds's boids seem to support the fascinating theory of emergent behavior, which describes how complex social interactions can arise when individuals obey a few rudimentary but very special rules. Reynolds's code contains no reference to flocking, much less any instructions for how a flock should navigate obstacles. And whereas most programs can't deal with situations that the programmer does not anticipate, in this simulation the boids execute a surprising array of sensible responses to unforeseen challenges. Reynolds remains a leading pioneer on the virtual frontier: his applications of computer animation to motion pictures won him an Academy Award in 1998.

Ariel Dolan, a computer programmer in Ramat-Gan, Israel, has added a delightful twist to Reynolds's creation by developing carnivorous boids, dubbed "floys" (pronounced "flow-eez"). Your gateway to Dolan's a-life aviary is at www.aridolan.com/eFloys.html, where you will find Java scripts that produce simulations that are truly a treat to watch.

Dolan's floys follow just two rules: stick close to your fellows but not too close, and when you spot an intruder, move toward it and attack. These instructions also generate a flock. But the assembly benefits individuals in an unexpected way. Avoiding too much togetherness





CYBERBIRDS, dubbed "boids," can be created on a computer (*left*) using software developed by Craig Reynolds. Although each virtual creature follows only a few simple rules, when a group of them encounters an obstacle (*below left*), the individuals split up (*center*) and later rejoin (*right*), mimicking a flock of real birds.

ALL ILLUSTRATIONS BY DANIELS & DANIELS

causes the floys to move in an ensemble that spreads out over a substantial region. They thus patrol airspace that extends far beyond the sensory range of any one floy. Whenever an outsider approaches, the nearest floy moves toward it. And because its compatriots are all programmed to remain together, the nearest neighbors follow along, as do their wing men, and so on.

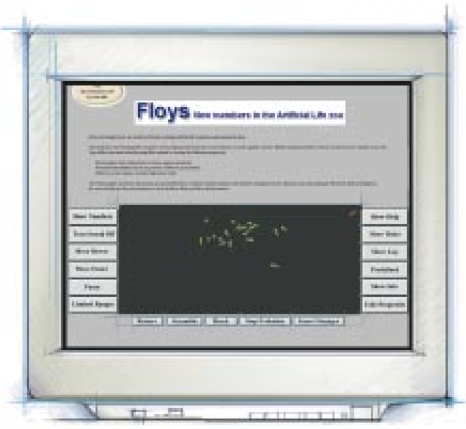
The result is quite striking. Challenged by an enemy, the entire flock turns quickly and pursues the intruder, including those floys that were initially too far away to sense the presence of the interloper. The swarm soon engulfs the invader, which ends up fighting for its a-life. This scene is not unlike what would happen to a hapless water buffalo that wandered into a lake infested with piranha.

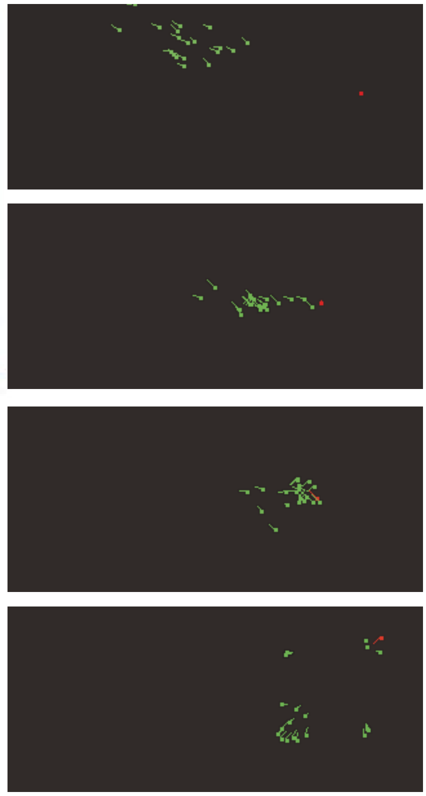
Like feasting piranha, each floy is rewarded for every successful "bite" that it takes. In Dolan's boid-eat-boid world, each morsel delivers one unit of energy from prey to predator. Fast floys have more opportunity to gorge, but they burn up energy rapidly while flying. Slow floys use less energy, but they tend to reach intruders last, and so they collect less sustenance.

Dolan's code generates mutations in the instructions for speed and energy consumption and in other parameters as well, including how closely individuals approach one another, how fast they accelerate and even the probability that they will disobey the rules. Dolan's program also ensures that the floys with the most energy are the most likely to reproduce and pass on their traits. So as the cybergenerations pass, the population becomes ever better adapted to live within its virtual realm.

Armed with Dolan's code and a little imagination, anyone with modest programming skills can perform all kinds of original investigations. I am now trying to understand why high levels of aggression survive in a population when this tendency so often seems self-destructive. My new version sets two bands of floys against each other. A floy still loses energy when it gets bitten by one from the enemy camp. But because warriors don't gain strength when they strike a foe, successful floy fighters in my simulation don't gain energy when they bite an opponent. The result is a melee in which the winner is determined entirely by numbers, energy reserves and the rules of probability.

I plan to modify the code to allow for floys with three distinct levels of inborn hostility. When a floy receives one cybergene for aggression from each parent, it will be especially combative and will fight to the death with all foreigners. A floy that inherits only one gene for aggression will battle until its energy reserves get low and then retreat. And a floy with no genes for aggression will withdraw from all fights.





JAVA APPLET allows interested Web surfers to simulate carnivorous boids, called "floys" (*above*). A sequence of snapshots (*top to bottom, right*) shows what happens when a band of floys (*green*) detects a lone intruder (*red*) and hunts it down as the besieged interloper tries to flee.

Here's what I think might happen. When the population encounters an enemy group of overwhelming strength, only some of the band will do battle, while the aggressionless individuals will retreat. When energy reserves eventually run low, the modestly aggressive floys will join their nonaggressive comrades. But the highly aggressive floys will continue to engage the enemy, thereby protecting the retreating floys from attack. Although these steadfast warriors will ultimately be destroyed, most of the original population will survive—not bad, considering that a uniformly aggressive population would be killed off completely. And because more hyperaggressive individuals will appear in the next generation when the surviving floys with single genes for aggression mate, this strategy may be continued indefinitely. So some individuals in a warring population might always have destructively aggressive tendencies—not for their own protection but to ensure the survival of their pacifist brethren. Dolan's floys allow me to explore this notion.

Artificial life-forms can also be used to examine problems that have nothing to do with biology. For example, some investigators have used computer simulations of this kind to probe the mysteries of traffic flow (see, for instance, www.theo2.physik.uni-stuttgart.de/treiber/MicroApplet/). My good friend Greg Schmidt believes that Dolan's floys may, in fact, need little modification to model the way people drive on California highways. As with my version, the actions of each floy could be determined in part by an aggression parameter, which would make some floys more likely to speed, swerve in front of others or drive on the shoulder during traffic snarls. Such a model, borrowed from a simulation of birds, could offer important insights into traffic management. At the very least, it might one day explain why drivers prone to rush-hour road rage always seem as common as crows. SA

For more information about this and other projects, direct your browser to the Web page of the Society for Amateur Scientists, www.sas.org, and click on the "Forum" button. You may write the society at its new address, 5600 Post Rd. #114-341, East Greenwich, RI 02818, or call toll-free: 877-527-0382. To purchase Scientific American's *new CD-ROM containing all installments of this department published in the 20th century (more than 800 articles), visit www.tinkersguild.com or dial toll-free: 877-503-0148.*

BIOMATHEMATICS_PATTERNS

Spiral Slime

Ian Stewart finds mathematics in creatures great and small

This past summer I attended a conference in Portugal on the mathematics of pattern formation, and one of the lectures reminded me of my second favorite animal. My first favorite animal is the tiger, partly because I like its dramatically striped fur. I like my second favorite because of its patterns, too, but this creature isn't as elegant as the tiger. It is the slime mold, or more specifically, a species of cellular slime mold known as *Dictyostelium discoideum*.

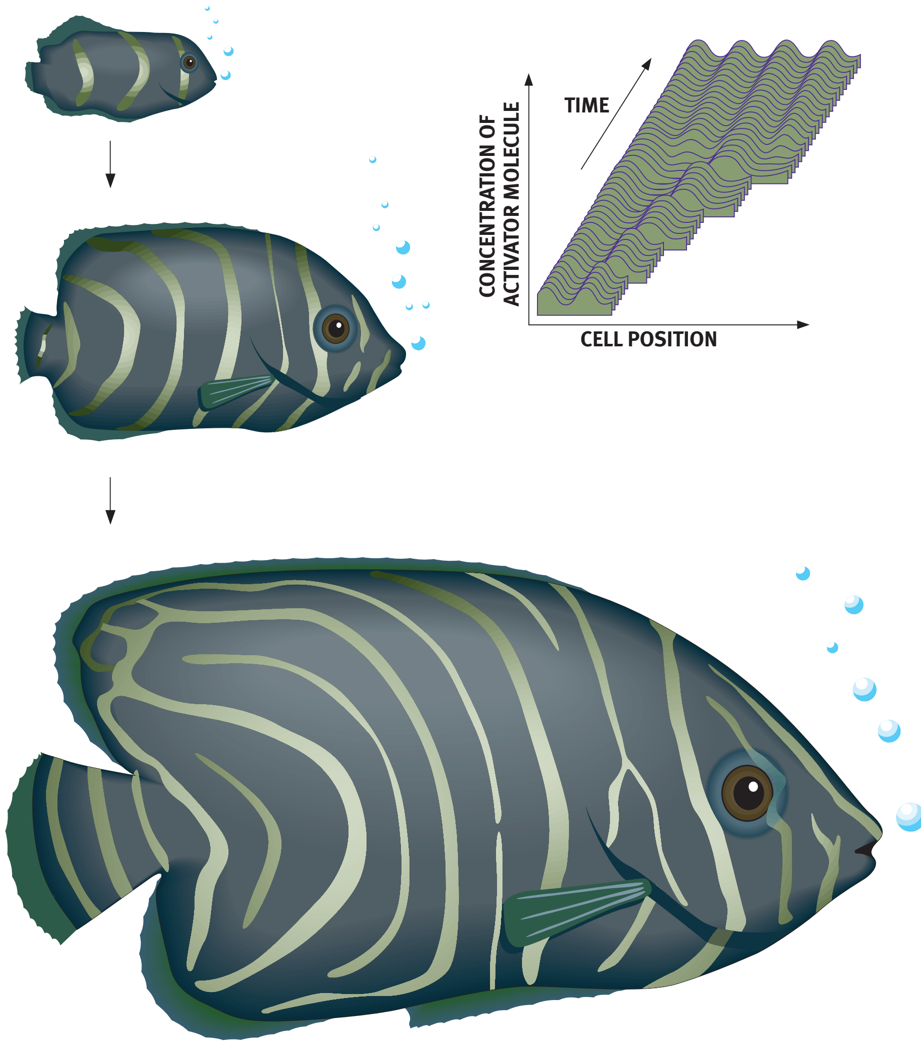
Biologists find the slime mold fascinating because it lies on the borderline between single-celled protozoa and multicellular organisms. The slime mold also illustrates a biological truth that the explorers of the human genome should take to heart: it's not just your genes that matter but what you do with them. Despite its lowly position on the tree of life, *Dictyostelium* manages to create astonishingly beautiful spiral patterns [*see illustration on page 118*]. To what extent are these patterns encoded in the slime mold's genes? Is there, in fact, a gene for spirals?

To answer that question, we need to know how the slime mold makes its spirals. The pattern is actually the result of collective activity. The life cycle of *Dictyostelium* begins with a microscopic spore wafting along on the winds. If the spore happens to land on a nice, moist resting spot, it germinates into a single-celled amoeba and starts hunting for food (mostly bacteria). When the amoeba gets big enough, it reproduces by splitting in two. Pretty soon there are lots of amoebas.

The artistry appears when the food runs low. The amoebas clump together, and as they make their way toward the center of the clump they sometimes form an elegant spiral. The crowd of amoebas gradually becomes more dense and the spiral more tightly wound. At some point it breaks up into "streaming patterns" that look like roots or branches extending from the center. The streams thicken, and as more and more amoebas try to get to the same place, they pile up in a heap known as a slug (not to be confused with the mollusk of the same name).

The slug is a colony of amoebas, but it moves as if it were a single organism. Once it finds a dry place, it attaches itself firmly to the ground and puts up a long stalk. At the top of the stalk is a round blob called a fruiting body. The amoebas in the fruiting body turn into spores and blow away on the wind, thus continuing the cycle.

Thomas Höfer, a biophysicist at Humboldt University, Berlin, has discovered a simple system of mathematical equations that reproduces both the slime mold's spirals and its streaming patterns. Cornelis J. Weijer of the University of Dundee has shown that very similar equations can model the movement of the slug. The main factors that determine the patterns are the density of the amoeba population, the rate at which the amoebas produce a chemical known as cyclic AMP and the



CHANGEABLE ANGELFISH develops new patterns of stripes as it matures. In a computer simulation of the process (*upper right*), the wave peaks correspond to stripes.

ALL ILLUSTRATIONS BY BRYAN CHRISTIE

ARTFUL AMOEBAS in a slime mold colony form spirals, as shown in this artist's rendering.

sensitivity of individual amoebas to this chemical. Roughly speaking, each amoeba "shouts" its presence to its neighbors by sending out cyclic AMP. The amoebas then head in the direction from which the shouts are loudest. The spiral pattern is a mathematical consequence of this process. It forms when the amoebas at the center of the clump are rotating as they send out waves of cyclic AMP.

It therefore seems that most of the slime mold's genes simply tell it how to be an amoeba. The genes tell the cells how to send out chemical signals, how to sense them and how to respond to them—but the spiral patterns they produce are not specified in the genes. Instead the patterns emerge from the mathematical rules that the amoebas are obeying. Mathematics may define the life cycle of the slime mold as much as genetics does.

The equations that lead to this far-reaching (and controversial) conclusion are modifications of equations devised nearly 50 years ago by English mathematician Alan Turing, who is best known as one of the founders of computer science. Turing was also interested in morphogenesis—the formation and differentiation of biological tissues and organs. In 1952 he postulated that ordered patterns in living creatures don't need an ordered precursor. He argued that the patterns could arise from chemical substances called morphogens that react with one another as they diffuse through tissue.

When Turing first published his ideas, they were purely theoretical, but a striking example of "Turing patterns" soon appeared: the Belousov-Zhabotinsky (BZ) chemical reaction. Russian scientist B. P. Belousov and later his compatriot A. M. Zhabotinsky discovered that mixing just a few ordinary chemicals—including sodium bromate, sulfuric acid and malonic acid—in a petri dish will produce concentric rings and spirals very similar to those made by the slime mold. Similar reactions can produce stripes, spots, dappling and many other patterns that are common in the animal kingdom.

Nevertheless, Turing's ideas were rejected by biologists. A major problem with his thesis was that the patterns that appear spontaneously in the BZ reaction are not fixed—they move across the petri dish. The same is true for all the other Turing patterns observed by chemists. In contrast, the patterns in most living creatures are fixed. We don't see zebras with moving stripes or leopards with moving spots. Turing had shown theoretically that his equations can produce both stationary and moving patterns, but laboratory experiments seemed to create only moving ones. Later on, the chemists discovered why: if you carry out the reactions in a gel rather than a liquid, the patterns become stationary. Living organisms resemble gels more than they do liquids. But by the time this distinction became clear, biologists had lost interest in the debate.

Mathematicians, though, continued to ponder Turing's ideas. Although his equations were far too simple to model real biological phenomena, they did produce the same kinds of patterns typically seen in animals. If pigments are deposited according to the peaks and troughs of parallel waves, you get stripes. More complex waves produce spots. The challenge for mathematicians was to flesh out Turing's scheme using theoretical models that more closely simulate the workings of biology.

In 1995 two Japanese scientists found the first convincing evidence of Turing patterns in living things. Shigeru Kondo and Rihito Asai of Kyoto University observed growing angelfish over several months and noticed a gradual rearrangement of their stripes. In mammals, the skin patterns simply enlarge as the creature grows, but in maturing angelfish new stripes are constantly forming as the older stripes split in two. What is more, the changes can be predicted by mathematical equations very similar to Turing's. A computer simulation of molecular interactions in a one-dimensional array of cells yielded a wave pattern that closely matched the rearrangement of angelfish stripes [*see illustration on page 116*].

The movement of the stripes is rather slow, which is why we don't generally notice it. But, as Galileo said, "It moves all the same." Mathematics changes the appearance of the angelfish just as it influences the life cycle of the slime mold—and perhaps our own as well. SA

READER_FEEDBACK

In a recent column on tilings ["Rep-Tiling the Plane," May], I said that nobody has classified all possible rep-tiles (tiles that can be assembled to form larger replicas of themselves). Aaron Meyerowitz of Florida Atlantic University sent me an e-mail pointing out that this is almost certainly a hard problem in the sense of algorithmic complexity—that is, how long a computer solution to the problem would run.

Suppose we restrict our attention to polyominoes, which are polygons made by joining squares edge to edge. It is known that the problem of determining whether a given set of polyominoes can tile a rectangle is intractable. Any algorithm used to solve the problem would take an absurd amount of time to yield an answer. The same is also likely to be true for the problem of determining whether a single polyomino can tile a rectangle. Furthermore, we know that any polyomino that tiles a rectangle is a rep-tile: you can always fit the rectangles together to make a square and then fit copies of this square together to duplicate the shape of the original polyomino. So it seems intuitive that classifying the polyomino rep-tiles alone is an impossible task. *—I.S.*

Books

REVIEWER_CLAIRE PANOSIAN DUNAVAN

Politics and Plagues

Laurie Garrett sounds an alarm for the disastrous state of global public health

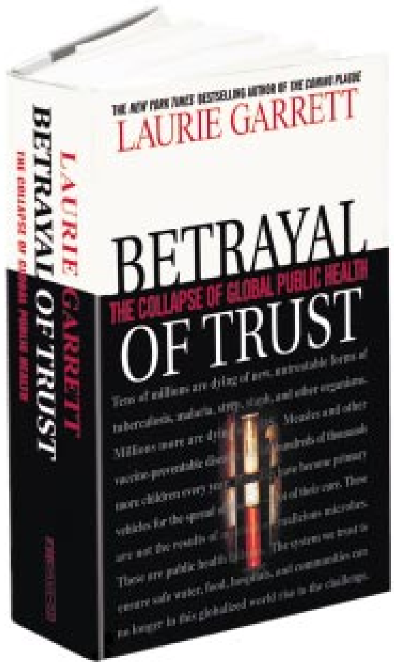
"At the dawn of the twentieth century the Western world fused the ideas of civic duty and public health. Conquering disease was viewed as a collective enterprise for the common good.... Where did we go wrong?" asks *Newsday's* medical writer and future-shock-meister Laurie Garrett in the epilogue to her new book *Betrayal of Trust: The Collapse of Global Public Health.*

Six years after her best-seller *The Coming Plague* fueled nightmares of microbial Armageddon, the journalist-turned-prophet is now on to politics and plagues. Her thesis? With globalization, no person or corner of the planet is safe from antibiotic-resistant superbugs, epidemics or biowar. Why then, Garrett asks, have recent investments in disease control declined, while relative spending on curative medicine—goods and services beyond the reach of most global villagers—skyrocketed? And will this brazen self-interest by the world's wealthy someday rain doom on us all? Needless to say, for many disease sufferers in *Betrayal of Trust,* it already has.

It's hard to know what to credit—Garrett's in-your-face storytelling, her avalanche of historical and contemporary health data, or simple force of will—but her proposition is convincing. By the final pages of *Betrayal of Trust,* we also believe.

The road to faith, however, is long—more than 700 pages of text and notes from five years of stalking disease and bureaucratic decay in India, Africa, the former Soviet Union and our own United States. Interspersed are epidemiologic statistics that sometimes raise more questions than they answer. But, in the end, even these support the cause. The mere fact that field data on, say, multidrug-resistant tuberculosis or HIV-tainted blood transfusions can be fragmented, anecdotal or downright flawed argues in itself for better epidemiologic intelligence and global monitoring.

In 1994 a flurry of plague on the west coast of India sounded an early warning. Over a single week, the quasi-outbreak produced a panicked exodus of half a million people from the economic boomtown of Surat. Headlining her chapter "Filth and Decay: Pneumonic plague hits India and the world ill responds," Garrett rebukes the World Health Organization and Indian health authorities for a palsied response to the potential global wildfire. After all, just one spew of bacteria-laden phlegm on an international flight could have lit the blaze. ("Well, maybe," a seasoned plague pro—of whom there are precious few these days—might reply. Among killer respiratory pathogens, smallpox beats *Yersinia pestis* transmission hands down any day.)



Betrayal of Trust: The Collapse of Global Public Health
by Laurie Garrett
Hyperion, New York, 2000 ($30)

In any case, while international airlines doused their planes with pesticide, Sudan jailed all travelers from India for six days, China barred Indians period, and WHO waffled, how big—really—was the outbreak of plague in Surat? Several thousand cases or fewer than a hundred? The truth is, we'll never know. India's government, no doubt embarrassed by its own crippled public health machinery and mishandling of the crisis, never released its cache of biologic specimens for outside review.

In contrast, global teamwork helped to rein in Zaire's 1995 epidemic of Ebola virus hemorrhagic fever. Once the world finally learned, that is, of desperate patients, relatives and hospital workers perishing from the blood-borne *landa-landa* (evil spirit) in the benighted town of Kikwit. Garrett's original Ebola coverage, complete with sights, smells and anguished cries in the night ("*Afwaka!* Someone has died! Someone has died! He was my husband! He was my husband!"), won her a Pulitzer Prize. Her recap in *Betrayal of Trust* is a cautionary tale of political corruption, human misery and the sometimes hidden dangers of sick bays. Ironically, with its utter lack of everyday gear like sterile needles and latex gloves, Kikwit's own general hospital was the single biggest boon to Ebola's spread there.

But the former Soviet Union's disastrous state of health and the post–cold war defections (who knows where: Iraq? Libya? Sudan?) of key military biologists are the twin thunderbolts most likely to jolt Garrett's readers from lingering complacency. First she surveys the carnage of an unraveled Soviet economy and medical system: a twofold excess in death versus birth rates, rampant alcoholism and heroin addiction, scores of sickly and abandoned children, widespread environmental pollution, crackpot science (a still pervasive legacy of Lysenko, the Stalinist era's hero of anti-intellectualism), mistrust of vaccination, and decrepit hospital gulags where ordinary infected people go without treatment while, elsewhere, power bosses hoard lifesaving drugs as they once did caviar and cognac.

Three hundred pages later, we join Garrett on the first visit by any U.S. journalist to VECTOR, the former U.S.S.R.'s once premier virus weapons lab outside Novosibirsk, Siberia. In 1997 it is a ghostly complex with broken windows, surrounded by Russian soldiers with threadbare uniforms and pathetic little ken of the lethal contents of Building Number 1—in Garrett's own words: "row upon row of industrial freezers [housing] Ebola, Lassa, smallpox, monkeypox, tick-borne encephalitis, killer influenza strains, Marburg, HIV, hepatitis A, B, C, and E, Japanese encephalitis, and dozens of other human killer viruses. And there were dozens of different strains of smallpox viruses—140 of them were natural, wild strains. Some were handcrafted by the bioengineers of VECTOR, giving them greater powers of infectivity, virulence, transmissibility."

Some experts take issue with Garrett's (and others') Jeremiah-like vision of biowar. In fact, the debate is now routine fare in major medical meetings. Nevertheless, after reading her eyewitness account and recalling recent history, it's hard to deny her claim that biologic weaponry vies today with nuclear arms–making as the world's most potent intellectual property.

As for American public health and medical science, Garrett names heroes and villains and openly states her mistrust of the new frontier: molecular and genomic medicine. Its illusory promise of health for all, she argues, plays directly into pharmaceutical profit mongering and ignores historical fact. After all, haven't nutrition, housing, sewage and water systems, epidemic control, immunization, education and literacy, and prenatal, maternal and child health programs—in short, public health—saved more lives than any high-tech medical intervention yet employed or imagined?

Haven't nutrition, housing, sewage systems, and immunization saved more lives than any high-tech medical intervention yet employed?

This seeming antiscience stand—along with Garrett's impassioned, attention-getting style—could raise the hackles of many a dedicated doctor or research scientist. In response, how easy it would be to quibble with her occasional factual error or to bloodlessly intellectualize her case: yes, of course medicine needs public health, and vice versa, and so on.

Don't go there.

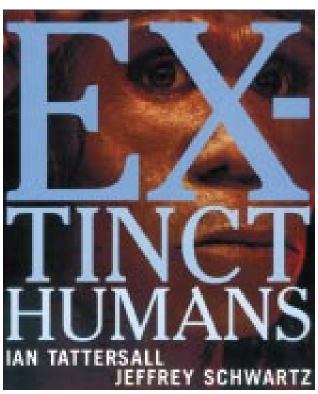
Someday Garrett may be thankful for medicine's imperfect craft. But right now she deserves thanks. While hardworking professionals in public health silently battle on, she gives voice to their stifled moral outrage over lost lessons and opportunities. At its core, *Betrayal of Trust* is a call for international activism long overdue. May its readers include policymakers, finance ministers and global philanthropists. With Garrett's glaring challenge before them, they, too—like the frontline heroes she honors—hold the power to save lives. SA

CLAIRE PANOSIAN DUNAVAN is a medical journalist and professor of medicine, infectious diseases and international health at the University of California, Los Angeles. Her articles and columns appear regularly in Discover *and the* Los Angeles Times.

THE EDITORS_RECOMMEND

IAN TATTERSALL AND JEFFREY H. SCHWARTZ'S *Extinct Humans,* with principal photography by Jeffrey H. Schwartz. A Peter N. Nevraumont Book, Westview Press, Boulder, Colo., 2000 ($50).

Tattersall (curator of anthropology at the American Museum of Natural History) and Schwartz (professor of physical anthropology at the University of Pittsburgh) have traveled around the globe to examine essentially every known hominid fossil. Their study of the anatomy has led them to the following conclusion. The pattern of evolution in our own species is no different from that of the rest of the earth's fauna: "repeated evolutionary experimentation, diversification and, ultimately, extinction." This reasoning may seem only commonsensical to those unfamiliar with the more usual picture that paleoanthropologists sketch of a rather linear development—"a single-minded struggle," as the authors put it, "from bestial benightedness to uplifted enlightenment."



They develop their theme with great style (and great photographs) and conclude by suggesting what accounts for *H. sapiens'* being the lone hominid on the earth today. We won't spoil the fascinating read by divulging what they (very convincingly) propose. The book is an intellectual adventure that would be well worth undertaking for this intriguing denouement alone, but there are in addition a wealth of informative stops en route.

MARTIN DAVIS'S *The Universal Computer: The Road from Leibniz to Turing.* W. W. Norton & Company, New York, 2000 ($26.95).

"As computers have evolved from the room-filling behemoths that were the computers of the 1950s to the small powerful machines of today that perform a bewildering variety of tasks, their underlying logic has remained the same," Davis says. "These logical concepts have developed out of the work of a number of gifted thinkers over a period of centuries. In this book I tell the story of the lives of these people and explain some of their thought." Davis, professor emeritus of mathematics at New York University, has devoted his career to "this relationship between the abstract logical concepts underlying modern computers and their physical realization."

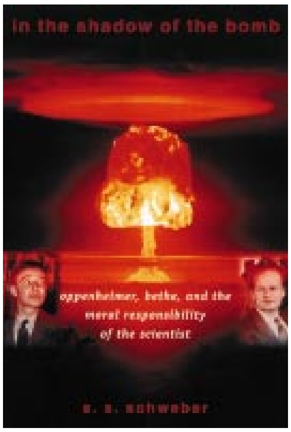
His tale encompasses seven mathematicians who contributed to that relationship: Gottfried Leibniz, George Boole, Gottlob Frege, Georg Cantor, David Hilbert, Kurt Gödel and Alan Turing. Leibniz, one reads, dreamed of "machines capable of carrying out calculations." Boole put forward an algebra of logic. And on to Turing, who envisioned a "universal machine" that could play games like chess, be induced to learn much as a child does and ultimately "could be made to exhibit behavior one would be led to call intelligent." Davis believes that the story he tells "underscores the power of ideas and the futility of predicting where they will lead."

HOWARD C. HUGHES'S *Sensory Exotica: A World beyond Human Experience.* MIT Press, Cambridge, Mass., 1999 ($26.95).

Can a dog sense in advance that its owner is about to have an epileptic seizure? A dog described in a recent news report does that, evidently by detecting certain chemicals associated with the onset of a seizure. It is an example of a sensory capability beyond the human range. Many animals can sense things that people are unaware of or sense weakly. Such animals are the subject of the story recounted by Hughes, who is a professor of psychology at Dartmouth College. He describes sonar in bats and dolphins, biological compasses (based on the sun or stars or geomagnetism) in birds and insects, electricity sensing in fish, and pheromones (chemical signals) in insects and apparently in people. And he takes pains to pin down the mechanism of the sensory capability in each case. "We don't yet have all the answers," he says, "but at least we are learning how to ask the right questions."

S. S. SCHWEBER'S *In the Shadow of the Bomb: Oppenheimer, Bethe, and the Moral Responsibility of the Scientist.* Princeton University Press, Princeton, N.J., 2000 ($24.95).

Many a scientist has to think today about what Schweber calls "the danger of the knowledge of certain technologies." It is an issue that came into sharp focus with the development of the atomic bomb. Schweber treats it by tracing the careers of physicists J. Robert Oppenheimer and Hans Bethe, with emphasis on their answers to the question "What is the role of the scientist in a democracy?" In so doing he sets the stage for an inquiry into other profound and troubling questions: "What did it in fact mean for scientists to address problems affecting all of humankind? What moral and political responsibilities did it entail, particularly during the beginning of the Cold War and in the McCarthy era? And how did scientists respond to these demands?" Schweber is a physicist and a historian of science, serving as professor of physics and professor of the history of ideas at Brandeis University. He brings both talents compellingly to his discussion.



GERALD M. EDELMAN AND GIULIO TONONI'S *A Universe of Consciousness: How Matter Becomes Imagination.* Basic Books, New York, 2000 ($27.50).

A woman senses that a room is light or dark and is aware that she has done so. A photocell senses the same thing without awareness. The difference is consciousness—something everyone recognizes but no one can fully explain. Edelman (director of the Neurosciences Institute in San Diego) and Tononi (a senior fellow there) propose what they call the dynamic core hypothesis to explain the neural basis of conscious experience. "This hypothesis states that the activity of a group of neurons can contribute directly to conscious experience if it is part of a functional cluster, characterized by strong mutual interactions among a set of neuronal groups over a period of hundreds of milliseconds." They call such a cluster the dynamic core because of "its ever-changing composition yet ongoing integration."

In telling their tale, the authors describe brain structure and function, review earlier efforts to explain consciousness and come to a discussion of higher-order consciousness—the kind that humans have. "Our position has been that higher-order consciousness, which includes the ability to be conscious of being conscious, is dependent on the emergence of semantic capabilities and, ultimately, of language. Concomitant with these traits is the emergence of a true self, born of social interactions, along with concepts of the past and future."

PETER TYSON'S *The Eighth Continent: Life, Death, and Discovery in the Lost World of Madagascar.* William Morrow, New York, 2000 ($27.50).

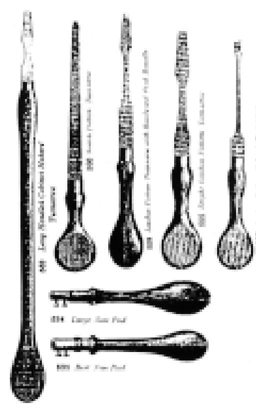
Madagascar is the world's fourth-largest island (after Greenland, New Guinea and Borneo) and the largest oceanic one. Isolated in the Indian Ocean, 250 miles off the coast of Africa, it is of singular interest because of the many animals and plants that are found only there. Science writer Tyson went there four times between 1993 and 1997. On each trip he traveled about with a scientist—a herpetologist, a paleoecologist, an archaeologist and a primatologist. His book describes the island largely through their research. It is a technique that produces a sharp picture of an intriguing place. Tyson also presents a brief history of the island, beginning with the odd fact that the first human inhabitants apparently arrived no more than 2,000 years ago, and ponders what can be done about the ecological devastation being wrought by the impoverished population there today.

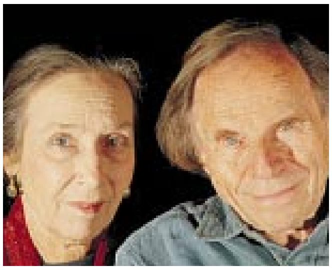

FROM *THE EIGHTH CONTINENT*

WITOLD RYBCZYNSKI'S *One Good Turn: A Natural History of the Screwdriver and the Screw.* Scribner, New York, 2000 ($22).

Rybczynski, a professor of urbanism at the University of Pennsylvania, has written best-selling books on such subjects as domestic comfort, building his own home and Frederick Law Olmsted. He felt a bit let down when the Sunday magazine of the *New York Times* asked him to write an article about "the best tool" of the second millennium. But he is good with tools and interested in them, and so he took on the assignment.

Many tools, he soon found, predate the second millennium. Consulting William Louis Goodman's *History of Woodworking Tools,* published in 1964, he read somewhat disbelievingly that the screwdriver did not appear until the 19th century. That set him off on a search for earlier references to this "laughably simple tool." The result is this splendid account of a number of tools, of the evolution of the screw and finally of his discovery that the "turnscrew" is indeed much older than Goodman thought. His search led him eventually to the 15th-century *Medieval Housebook,* where he found a drawing of a screw-turning lathe with a puzzling tool resembling a chisel lying on a workbench. "One day, while I am puzzling over the drawing again, I realize that the blunt end [of the tool] is exactly the same size as the slot in the head of the cutter. Of course. It's not a chisel, it's used to adjust the cutter. It's a screwdriver. Eureka! I've found it. The first screwdriver."


FROM *ONE GOOD TURN*



Anniversary of Needles

Philip and Phylis Morrison celebrate 400 years of magnetic understanding

The magnetic compass sailed westward for centuries out of China, until even Atlantic mariners used it widely. In the round year of 1600, a prominent physician of Elizabethan London, William Gilbert, authored—in Latin—the first classic of that springtime century of science, under the grand title *About Magnets and Magnetic Bodies, and that Great Magnet, the earth,… & demonstrated by experiments*. That English rendering is still in print (Dover paperback). No wonder Galileo prized Gilbert's amazing work, 10 years earlier than Galileo's own celebrated Latin classic on the telescopic sky.

Natural magnets, called lodestones, are samples of magnetite, the complex black oxide of iron, and were then chief donors of magnetization. Any piece of smelted iron could be magnetized by stroking it with a good-size lodestone. (The old word "lode" occurs still in the poet's word "lodestar" and carries the sense of leading one on.) Gilbert tells just how to proceed even without a lodestone: ordinary iron metal, "if it be drawn out long" and held in the north-south line, can magnetize compass needles quite well.

A magnetic needle aligns itself very roughly north and south along the meridians wherever it is taken (no iron ships then). No one knew how the needle attended to one suspiciously cosmic direction everywhere it went, until Gilbert demonstrated the circumstances of that amazing power. Stars aside, this magnetism is terrestrial.

Gilbert explains it through action. Make a "terrella" by rounding and polishing a small piece of lodestone. Little iron wire bits—"one barley-corn in length"—will align nicely about it, their stance depending on their location. Around the equator of the terrella they lie flat on the surface but turn to point to the poles. At each magnetic pole they stand upright, pointing into the ground. At a place like London they find the north-south plane and seek to point downward, but not vertically as at a pole. This "dip" of real compass needles toward the downward direction, once mounted to swing freely in the north-south plane, had been first displayed at about the time of Gilbert's earliest researches.

If the body of the earth holds a huge lodestone, its two magnetic poles lying a little displaced from the poles of earth rotation, the magnetic technology that puzzled all before Gilbert can be rationalized. Worldwide observations of local misalignment between the magnetic and the celestial pointers fit this strangely simple picture quite well, and their dip angles, too. Today a grade school science class can have a small bar magnet and a hand-size split globe of the earth made of foam plastic. Push the magnet into the foam at about the right latitude, and many small bits of soft iron will align themselves on the spherical surface just as Gilbert drew them. Staples are bits of iron wire easily made by working your desktop stapler without any papers to link.

At the southernmost tip of Africa, well south of the Cape of Good Hope, Portuguese navigators noted around 1500 that the magnetic needle pointed closely in the same direction as the noon shadow of the sundial gnomon (or of the mainmast), the true north and south by sun and stars. They named that austere and windswept strand the Cape of the Needles, where magnetic needle and earth axis concurred. That agreement no longer works there, for the great magnet of the earth drifts irregularly up to 10 degrees a century.

Is the earth really a great lodestone? Certainly not. Neither magnetite nor iron metal remains strongly magnetic at the temperature of the earth's molten outer core. But long-term geomagnetic change provides strong and simple evidence for our present understanding. Compass needles are not ruled by any permanent magnet but by a great electromagnet, the earth. We find an ordered succession of magnetized rocks on the ocean floors that record 100 and more full reversals of the present direction of the magnetic poles during the past couple of 100 million years of seafloor spreading. The earth's magnetic state fades sharply to zero at highly irregular intervals on a million-year timescale. The deep electric currents switch off or nearly off, to reappear generally aligned near the rotation axis, sometimes with reversed magnetic polarity, sometimes repeating the previous state. The core region is a dynamo, where complex electric currents wind within a thick conducting fluid shell to generate the surface fields.

Theorists assure us that magnetic forces between interacting currents can amplify an initial seed field, taking energy both

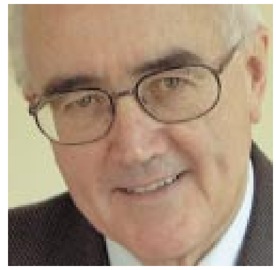
DUSAN PETRICIC

from the heat they tap and from the kinetic energy stored in the rotation of the earth. The seed might be the magnetized plasma repeatedly expelled our way from the active sun. During the past five years, three-dimensional simulations of the real fluid global dynamo have become impressive, capable of mapping plausibly tangled

Continued on page 127

Survivals

James Burke considers the primitive, the antiquarian, the mathematical, the electrical—and draws some dramatic conclusions


DAVE PAGE

I accidentally spilled some salt at dinner the other night and was absentmindedly tossing a pinch over my left shoulder when I recalled that Victorian adventurer Edward Burnett Tylor had first identified that kind of thing: the now meaningless modern survivals of once meaningful ancient practices. Tylor ("father of cultural anthropology") was the guy who linked Little Red Riding Hood to solar myths and kicked off the argument (still raging) about whether early technology diffused or was independently invented by different cultures around the planet.

He began this work after visiting Havana in 1856, where he met Henry Christy, traveling son of a banker, loaded and ethnology-minded. The pair went on to Mexico, where Tylor made notes for his first anthropological book, *Anahuac,* and Christy filled his suitcase with all that paraphernalia you bring back from holiday: masks, dolls, wicker baskets, and so on. Christy did this magpie trick anywhere he went, including Perigord, France, where he later bankrolled and accompanied Édouard Lartet ("father of paleontology") on explorations in Upper Paleolithic caves that revealed some of the first examples of cave art. A few of which Christy hacked off and took back to England to add to his massive collection of prehistoric loot, which in 1865 he would leave to the British Museum, in the tender hands of keeper of ethnography Augustus Wollaston Franks. The collector's collector.

Apart from piling up stuff for himself (Italian majolica, Japanese and Chinese porcelain, Japanese sword guards, medieval brass-rubbings, various ancient drinking vessels and finger rings), Franks was a whiz at separating the rich and aged from their hoards of amassed bric-a-brac for the benefit of the museum. One such was the Fountaine family treasure, established 100-odd years earlier by Sir Andrew of that ilk, tutor to the English royal family, successor to Newton at the Mint, pal of Gottfried Leibniz and antiquarian extraordinaire. In 1714 Sir A. spent a few years in Italy, buying up serious quantities of the past (coins, paintings, books and ceramics) and becoming bosom buddies with Cosimo III de' Medici, second-last grand duke of Tuscany, while that place and family were going to rapid wrack and ruin. I mean, as recently as Cosimo's dad, Ferdinand II, the resident in-house Medici intellectuals had been none other than Galileo and his sidekick Evangelista Torricelli, who only discovered the vacuum.

Torricelli is less well known for his work in early gobbledygook.

Torricelli is less well known for his work in early gobbledygook, a.k.a. the geometry of indivisibles. If you read this column regularly, you'll know this is one of my many blind spots, so the best I can do is say that Torricelli made discovering the volume of barrels easier. I think.

Well, there was infinitely more to it than that, as was evidenced when Torricelli's notes on infinitesimals blew away one John Wallis, English mathematician and no slouch. Knew Greek, French and Latin by the age of 15, became the government's crack cryptologist, introduced the infinity symbol ∞, helped to found the Royal Society with his pal Robert Boyle and took Torricelli's work to the stage that it was ready for Newton to turn it into calculus. All that, as well as churning out the homilies and tracts required when you were also a working clergyman. Wallis's puritanical sermons were edited and published in 1791 by Charles de Coetlogon, Calvinist preacher, whose pièces de résistance were his blurbs, that is, prefaces to other people's weighty tomes, the 18th-century equivalent of "Read it! A real page-turner!"

In the 1770s de Coetlogon was assistant chaplain at Lock Hospital in London to the Rev. Martin Madan, reformed sinner. Back in 1748 Madan had belonged to a drinking club for young tearaways, attended one of John Wesley's sermons to get material for a satirical after-dinner routine, was converted on the spot and took holy orders. My guess is you know Madan only because he edited the present version of Wesley's brother Charles's "Hark! The Herald Angels Sing." In 1780 Madan really blotted his copybook with a carefully reasoned work on how polygamy was okay by the Bible and would solve the problems of adultery, seduction and prostitution. Lead balloon.

Meanwhile the umpteenth-upgrade edition of Wesley's personal health care handbook, *Primitive Physic,* was hitting the stores with an extra chapter on the amazing curative properties of the new electricity. Very much in fashion at the time, thanks to quack promoters like fave-rave "Dr." James Graham and his "London Temple of Health and of Hymen," where hypochondriacs waited in line to sit on an "electrico-magnetic throne" and receive shock treatment for everything from acne to lack of issue. Once Graham shared the lecture stage with a large woman named Ann Siddons,

PATRICIA J. WYNNE

who had a squint, who later tried to poison herself very publicly in Westminster Abbey and whose sister paid her an annuity to stay more than 150 miles away.

The sister was Sarah Siddons, tragedy-queen megastar of the 18th-century theater, who would play every major female Shakespearean role, give lessons in speech to the royal princesses and spend over 50 years thesping to SRO crowds, many among which were so impressed at her performances that they fainted.

In her 1801 *King John,* the part of teenage Prince Arthur was played by a 14-year-old newcomer, Edmund Kean, who took elocution lessons from his ventriloquist uncle and who stage-managed a career every bit as boffo as Siddons's, making so much money that it was piled on his apartment floor. Kean believed his own hype, became the melodrama villain of all time, was elected an honorary Huron chief and died of alcoholic excess. His son Charlie, trading on the family name, was a moderately successful theater manager. In 1858 he hired an 11-year-old girl to play Prince Arthur. She went on to become Henry Irving's leading lady and to gain worldwide fame as Ellen Terry, with more years on the Shakespearean stage even than Siddons.

In 1905, late in her career, Terry played the lead in *Alice Sit-by-the-Fire,* written for her by an up-and-coming playwright, J. M. Barrie, who had just produced what his friends told him would be a disaster that nobody would ever buy tickets for: a modest little play for children entitled *Peter Pan*. Barrie authored a list of plays as long as your arm, most of them now forgotten, including his only real flop: *Jane Annie,* written in collaboration with a guy who three years earlier had been a practicing doctor and one year later was to become a name in every household as the creator of the most famous detective of all time. Arthur Conan Doyle's Sherlock Holmes turned into such a cult that to this day thousands fetch up at his "home" (now a museum) on Baker Street, London, where he "lived" with his medical straight man, Dr. Watson.

Later in life Doyle dreamed up another winner: Professor Challenger, explorer-anthropologist hero of *The Lost World,* who finds a mysterious, isolated plateau in the South American jungle that is inhabited by living prehistoric animals and half-ape, half-human "missing links."

Archaic survivals suggested by the work of anthropologist E. B. Tylor. Pass the salt, please. SA

Wonders, *continued from page 125*

worldwide voyages of the modeled poles.

It is worth recalling the elaborate wire coils with iron cores that compose the power generators now turning to feed your reading lamp. That functionally similar dynamos can grow of themselves at cosmic scale is a wonder. The power demand of the global electromagnet is not of cosmic magnitude but on the scale of the biggest power generators we ourselves have built. The earth's magnetism has been supplied over time by using under a billionth of the earth's stored energy of spin.

The first experiment Gilbert describes in *De Magnete* is surprisingly not about magnetism at all: "Make yourself a rotating needle, of any sort of metal, . . . pretty light, poised on a sharp point, after the manner of a magnetic pointer." No magnetism, no north or south, this needle signals not a magnetic attraction but an electrical one as it turns to follow the position of a rubbed piece of golden amber. Gilbert extends his fresh and skeptical thinking to fit the facts: a second distinct material attraction was already known to Greek and Roman authors but never so carefully examined. A spray of water will shield this needle from amber, but a coat of olive oil will not. Gilbert lists many substances that attract this needle, not only amber and jet, and scoffs at wilder claims, mainly magical or animate explanations, as "disgracefully inaccurate."

His mechanics and his chemistry are still very much derived from Aristotle and the Greek elements and humors, so his long, philosophical arguments do not seem conclusive to us. But he is a diagnostician of talent, who tells us what he has tested and compared during 18 years of work. His electrics are only a brief digression, but they point right at electric charge and its forces, now linked into electromagnetism. This early touch of generality, a kind of unification based not on words alone but on shared action, illuminates Gilbert's work. Four centuries later his simple model still lies closer to our developing science than to technical innovation or to learned speculation alone. SA

DAN WAGNER

Smart Thinking

Readers of this magazine are obviously highly intelligent, but what other clues make us assume someone is brainy, wonders **Steve Mirsky**

Here's how you know people think you're smart: they call you "Einstein." Here's how you know people think you're really smart: you're the guy whose name, Einstein, is what smart people get called. The really smart guy's name may also be used sarcastically—ya follow me, Einstein? (Quick digression: Was this habit popular throughout history? Centuries ago did anyone ever say, "You forgot to muck the stalls again, Newton." Or, "Hey, Pythagoras, your sandal's untied about halfway up your calf.")

Finally, here's how you know people think you're really, really smart: in a study designed to examine perceptions of intelligence, undergraduates consistently and often by a wide margin picked Albert Einstein in response to the instruction, "Think of an ideal example of an intelligent person." The research, performed at the University of British Columbia and recently published in the *Canadian Journal of Behavioural Science*, was designed to get a better grasp on judgments of intelligence. The authors point out that such perceptions and conceptions "influence attitudes and behaviour in everyday social interactions, voting preferences ... and can affect personnel decisions." Ask any kid whose teachers already taught that kid's older sibling.

The study examined the responses of hundreds of sophomores and juniors over a two-decade stretch, with data gathered in 1982, 1984, 1989, 1993 and 1997. Even though each student could name only one person, Einstein always topped the list and never had less than a 16 percentage point lead over his closest rival—with the exception of 1982, when he edged Pierre Trudeau by a single point. (For you younger readers, Trudeau was the dashing, erudite Canadian prime minister who presided over Canada's final, full independence from Great Britain. His wife Margaret, apparently much more of an Anglophile than Pierre, eventually left him and wound up dating half the Rolling Stones.)

The Trudeau votes are part of a fascinating trend. Both the Canadian and British prime ministers—whoever they happened to be at the time of a particular polling—do well in the voting. And whoever is the president of the U.S. does even better. This prejudice has led to oddball assessments of brainpower in which Ronald Reagan beat Leonardo da Vinci (last superpower over *Last Supper*), George Bush topped Stephen Hawking (read my lips over try to read my books) and Bill Clinton outpolled Isaac Newton (castigation over gravitation).

The results reveal the importance of three factors in assumptions about intelligence: familiarity, likability and power. Who could be more familiar than the prime minister or the president, both of whom probably appear on Canadian television on an almost nightly basis? And if those people didn't have a talent for being liked, chances are they wouldn't be in office in the first place.

Both familiarity and likability probably take a back seat to power, however, which explains why the Canadian P.M.'s usually finished behind the U.S. presidents. "Why aren't Canadians biased toward their own leaders?" the authors ask. "The paradox, we suspect, is based on the fact that Presidents ... are viewed as more powerful than our Prime Ministers—even with respect to Canadian affairs. Indeed, there is previous evidence that leader power supersedes leader likability in determining attributions of intelligence."

So caveat emptor, Cicero. Next time you're under the impression that somebody is smart, make sure it's not just charm or nuclear stockpiles that has you making that assumption. For example, other frequent high finishers include Steven Spielberg, Oprah Winfrey and Madonna. Now, these people are no doubt no dopes. But the fact that they appear in the top 15 in multiple years, whereas John Bardeen is nowhere to be found, is a strong indicator that in conceptions of intelligence, fame and power trump (oh yeah, the Donald made the lists, too) two quiet Nobel Prizes in Physics. That's right, Bardeen won two Nobels, which is one more than Einstein got. And probably at least one more than you've got, Einstein. SA

SIDNEY HARRIS