A PROBLEM THAT MATH CAN'T SOLVE • NEW TRANSLATION MACHINES

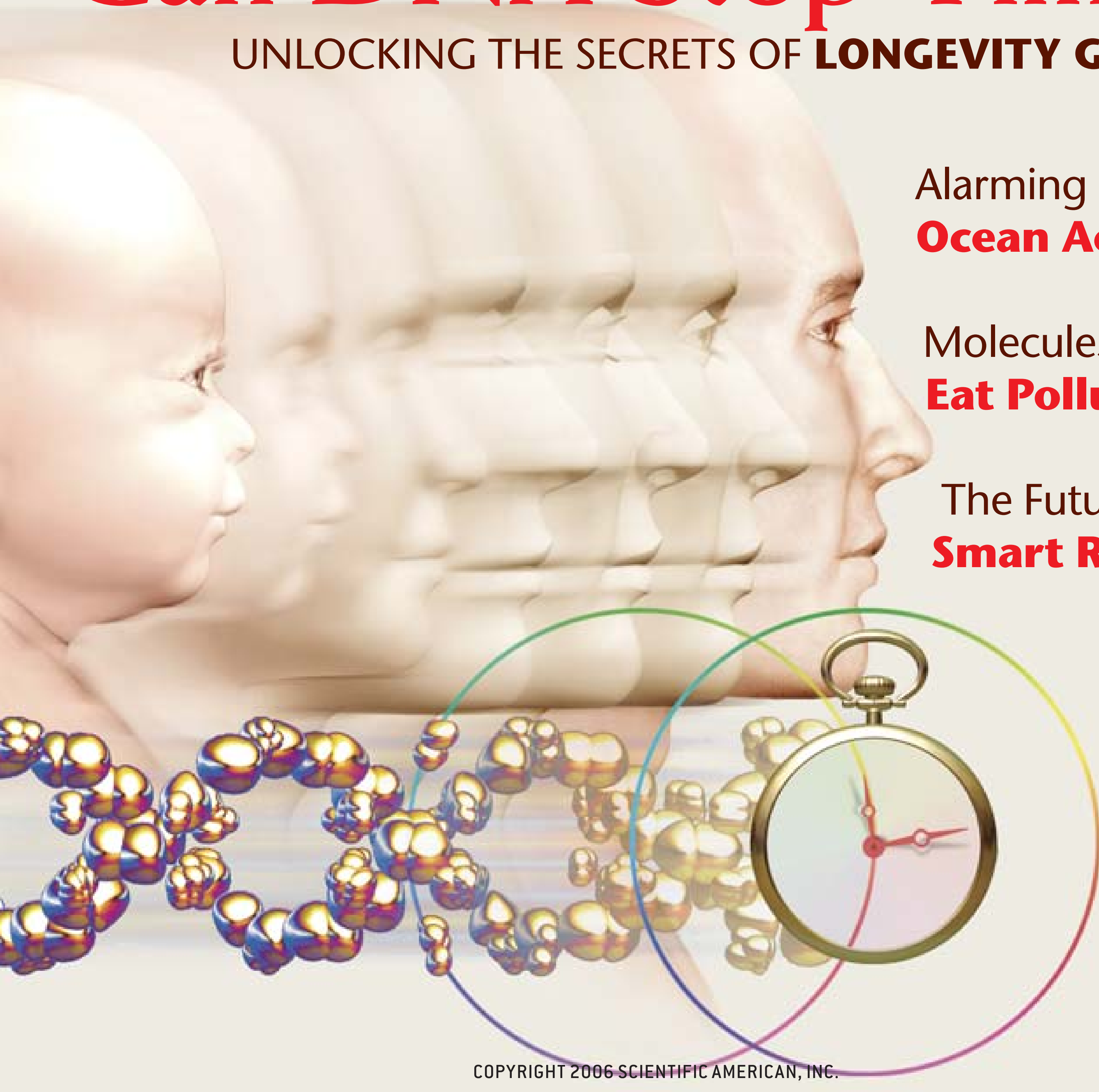
SCIENTIFIC AMERICAN

MARCH 2006
WWW.SCIAM.COM

Cosmic Rays: The Real Threat for Travelers to Other Planets

Can DNA Stop Time?

UNLOCKING THE SECRETS OF **LONGEVITY GENES**



Alarming Rise in **Ocean Acidity**

Molecules That **Eat Pollution**

The Future of **Smart Radio**

march 2006

contents

features

SCIENTIFIC AMERICAN Volume 294 Number 3

74 Where mathematical reason cannot go

SCIENTIFIC AMERICAN Volume 294 Number 3

departments


28


104

36 Kay-Yut Chen, Hewlett-Packard

columns

Cover image by Jean-Francois Podevin; image on preceding page by Kenn Brown; photograph at left by Timothy Archibald.

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 2006 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording for public or private use, or by any information storage or retrieval system, without the prior 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. 40012504. Canadian BN No. 127387652RT; QST No. Q1015332537. Publication Mail Agreement #40012504. Return undeliverable mail to Scientific American, P.O. Box 819, Stn Main, Markham, ON L3P 8A2. Subscription rates: one year $34.97, Canada $49 USD, International $55 USD. 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. **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 248-7684. Send e-mail to sacust@sciam.com** Printed in U.S.A.

Con Men in Lab Coats

Five decades after it was revealed as a forgery, the Piltdown man still haunts paleoanthropology. Now, thanks to the disgraced stem cell researcher Woo Suk Hwang, cell biology has a high-profile scandal of its own to live down. Few recent papers in biology have soared as high in acclaim as Hwang's 2004 and 2005 announcements of cloning human embryonic stem cells—or plummeted as fast into infamy with the discovery that they were rank fakes [see "Down in Flames," by Sara Beardsley, News Scan, on page 20].

WOO SUK HWANG gives stem cell research a bad name.

BAE JAE-MAN *AP Photo/Yonhap*

Embryonic stem cell (ESC) research is no less promising today than it was before Hwang's deceit was revealed; most investigators continue to believe that it will eventually yield revolutionary medical treatments. That no one has yet derived ESCs from cloned human embryos simply means that the science is less advanced than has been supposed over the past two years.

Still, Hwang has badly sullied the reputation of a field that already has more than its share of political and public relations problems. Some longtime opponents of ESC research will undoubtedly argue that Hwang's lies only prove that the investigators cannot be trusted to conduct their work ethically, and the public may believe them. This is one more crime against science for which Hwang should be ashamed. (A minor footnote to this affair is our removal of Hwang from the 2005 Scientific American 50 list; see the retraction on page 16.)

In recent years, fabricated data and other fakery have been uncovered in work on materials, immunology, breast cancer, brain aneurysms, the discovery of new elements and other subjects. As the volume of publication rises, fraud will probably rise with it. Because of the growing financial ties between university researchers and corporations, not to mention the jockeying for leadership among nations in high-stakes areas such as stem cells, some scientists may feel more pressure to deliver results quickly—even if they have to make them up.

These affairs have something in common with the Jayson Blair and Stephen Glass scandals that not long ago rocked mainstream journalism: all these scams exploited the trust that editors extend to submitting authors. The editors and peer reviewers of scientific journals cannot always verify that a submitted paper's results are true and honest; rather their main job is to check whether a paper's methodology is sound, its reasoning cogent and its conclusions noteworthy. Disconfirmation can only follow publication. In that sense, the Hwang case shows how science's self-correcting mechanism is supposed to work.

Yet it is important not to brush off the Hwang case as a fluke without considering its lessons for the future. For instance, Hwang's papers had many co-authors, few of whom seem to have been party to the cover-ups. But what responsibilities should co-authors have for making sure that papers bearing their names are at the least honest?

We should also think hard about whether Hwang's deceit went undetected for months because so many scientists and science journalists wanted to believe that ESC research was progressing rapidly, because that would hasten the arrival of miraculous therapies and other biomedical wonders. Extraordinary results need to be held suspect until confirmed independently. Hwang is guilty of raising false expectations, but too many of us held the ladder for him.

THE EDITORS editors@sciam.com

How to Contact Us

EDITORIAL

For Letters to the Editors:
Letters to the Editors
Scientific American
415 Madison Ave.
New York, NY 10017-1111
or
editors@sciam.com
Please include your name and mailing address, and cite the article and the issue in which it appeared. Letters may be edited for length and clarity. We regret that we cannot answer all correspondence.

For general inquiries:
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8200
fax: 212-755-1976
or
editors@sciam.com

SUBSCRIPTIONS

For new subscriptions, renewals, gifts, payments, and changes of address:
U.S. and Canada
800-333-1199
Outside North America
515-248-7684
or
www.sciam.com
or
Scientific American
Box 3187
Harlan, IA 51537

REPRINTS

To order reprints of articles:
Reprint Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8877
fax: 212-355-0408
reprints@sciam.com

PERMISSIONS

For permission to copy or reuse material from *SA*:
www.sciam.com/permissions
or
212-451-8546 for procedures
or
Permissions Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
Please allow three to six weeks for processing.

ADVERTISING

www.sciam.com has electronic contact information for sales representatives of Scientific American in all regions of the U.S. and in other countries.

New York
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8893
fax: 212-754-1138

Los Angeles
310-234-2699
fax: 310-234-2670

San Francisco
415-403-9030
fax: 415-403-9033

Detroit
Karen Teegarden & Associates
248-642-1773
fax: 248-642-6138

Midwest
Derr Media Group
847-615-1921
fax: 847-735-1457

Southeast
Publicitas North America, Inc.
404-262-9218
fax: 404-262-3746

Southwest
Publicitas North America, Inc.
972-386-6186
fax: 972-233-9819

Direct Response
Special Aditions Advertising, LLC
914-461-3269
fax: 914-461-3433

Australia
IMR Pty Ltd.
Tel: +612-8850-2220
Fax: +612-8850-0454

Belgium
Publicitas Media S.A.
+32-(0)2-639-8420
fax: +32-(0)2-639-8430

Canada
Derr Media Group
847-615-1921
fax: 847-735-1457

France and Switzerland
PEM-PEMA
+33-1-46-37-2117
fax: +33-1-47-38-6329

Germany
Publicitas Germany GmbH
+49-211-862-092-0
fax: +49-211-862-092-21

Hong Kong
Hutton Media Limited
+852-2528-9135
fax: +852-2528-9281

India
Convergence Media
+91-22-2414-4808
fax: +91-22-2414-5594

Japan
Pacific Business, Inc.
+813-3661-6138
fax: +813-3661-6139

Korea
Biscom, Inc.
+822-739-7840
fax: +822-732-3662

Middle East
Peter Smith Media & Marketing
+44-140-484-1321
fax: +44-140-484-1320

Scandinavia and Finland
M&M International Media AB
+46-8-24-5401
fax: +46-8-24-5402

U.K.
The Powers Turner Group
+44-207-592-8331
fax: +44-207-630-9922

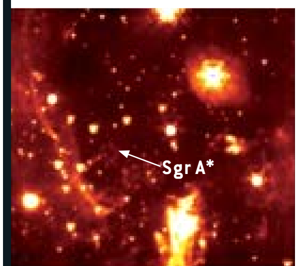
On the Web

WWW.SCIAM.COM

UPDATED EVERY WEEKDAY

Visit **www.sciam.com/ontheweb** to find these recent additions to the site:

Man-Made "Star" Illuminates Milky Way's Mysterious Center



On a clear night, a broad band of stars sweeps across the sky, backed by the faint, milky glow of even more stars. This is the Milky Way, and on a midsummer's eve the very center of it can be picked out within the Sagittarius constellation. But even on the clearest night, the earth's atmosphere obscures the true brilliance of our galaxy, and astronomers have long struggled with images blurred by its mix of gases and turbulence. Now researchers have used a new laser-generated star to obtain the clearest pictures yet of the Milky Way's center.

New Find Pushes Back Date of Mayan Writing

Scientists have identified the oldest known Mayan writing—dating from between 300 and 200 B.C. Previous examples of Mayan script could be confidently dated only to around A.D. 250, leading to speculation that the Mayans had inherited their writing from other, older cultures, such as the Zapotec, despite stylistic differences. The discovery seems to upend that theory, proving that the ancient Mayans were as literate as their descendants.

Science vs. the Death Penalty

The U.S. remains the only developed Western nation to permit executions despite serious flaws in the system. No need for any pacifist proclivity or liberal leaning to see that—just look at the science.

Ask the Experts

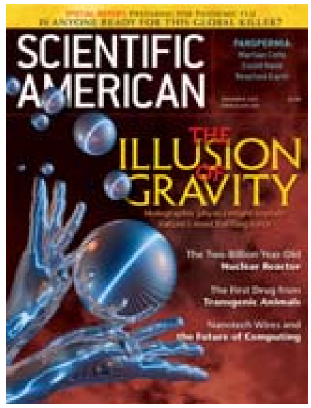
Why does eyesight deteriorate with age?

David Zacks, a retina specialist and assistant professor of ophthalmology and visual sciences at the University of Michigan Kellogg Eye Center, explains.

U.C.L.A. GALACTIC CENTER GROUP/W. M. KECK OBSERVATORY LASER TEAM

Letters

EDITORS@SCIAM.COM



"BE PREPARED," the Boy Scouts' motto, is simple enough for children to remember. Why then can it be so difficult to plan for foreseeable disasters? Two November 2005 articles addressed that question: "Preparing for a Pandemic," by W. Wayt Gibbs and Christine Soares, and "Preparing for the Worst" [SA Perspectives], which also discussed events such as Hurricane Katrina. Reader Colin Buss of Campbell River, B.C., observes, "If global health is at stake, the international community, including the U.S., may be better off not allowing the wealthier nations to dictate health strategy if it is at the expense of the greater good." Kelsey DeForge warned via e-mail not to assume treatment will be free: "The U.S. must assure that those who require vaccinations receive financial help if necessary. Otherwise treatment will be accessible only to those who can afford it, rather than those most at risk."

EXPECTING THE EXPECTED

In "Preparing for the Worst" [SA Perspectives], the editors again do not waste an opportunity to stick a thumb in the Bush administration's eye. And although you are truthful in what you say, what is telling is that which you omit: it was the Louisiana congressional delegation that coveted the funding for levee improvement and then led the charge to divert it to other purposes.

Van Snyder
La Crescenta, Calif.

I was struck by the similarity of the photograph of a Louisiana house surrounded by water that appeared in the *Los Angeles Times* on September 25, 2005, and one in *Scientific American*'s October 2001 issue. The latter was from the article "Drowning New Orleans," by Mark Fischetti, whose introduction states: "A major hurricane could swamp New Orleans under 20 feet of water, killing thousands. . . ."

Our leaders have told us that nobody could have foreseen the devastation wrought by Hurricane Katrina. They have rejected the process of accumulating knowledge through scientific analysis. In these times, when one of the best-established laws of nature is again being put on trial, perhaps your warning was rejected because the word "scientific" is in your title? I suggest that you should substitute the word "fundamental," so that nonscientifically inclined citizens will not be repelled before having a chance to learn from your presentations.

Stuart Spence
Los Angeles

PANDEMIC PREP

Thanks to Christine Soares and W. Wayt Gibbs for their thoughtful overview in "Preparing for a Pandemic." To optimize scarce hospital resources, however, we will need to provide clear direction for safe home care of patients with pandemic influenza.

We do not yet know exactly how H5N1 influenza will spread. We should start, however, to prepare this information for home care now by using our experience with ordinary flu. The public is eager to know: how far in advance of the onset of symptoms someone is contagious; the safe distance from an infected person if one is wearing no protective equipment; the type of masks that are effective in preventing transmission, or if they are entirely worthless; whether eye protection is necessary; how important a factor hand or fomite (contaminated article) transmission is; and how to protect caregivers and identify when a higher level of care is needed. We should begin the research without delay.

Marian McDonald, R.N., C.I.C.
Sebastopol, Calif.

U.S. policy as stated by Soares and Gibbs, which assigns the first shots to key government leaders, medical caregivers, workers in flu vaccine factories, pregnant women, infants, and elderly and ill people who are already in the high-priority group for annual flu vaccines, strikes me as both inconsistent and wrong.

The first three groups make sense if the goal is preventing a breakout pandemic—they are both the most at risk of exposure and the most necessary to prevent further exposure of others. The rest of the priority classes do not fit this test. They may be the groups who will *suffer* most, but they are not most likely to communicate the disease. That dubious distinction goes to school-age children, especially those attending nursery school. They are the proper focus of preventive vaccinations.

Mark Field
Los Angeles

TINY TECHNOLOGY, BIG CHALLENGE

The technology described in "Crossbar Nanocomputers," by Philip J. Kuekes, Gregory S. Snider and R. Stanley Williams, has a long way to go before it can be used in logic and memory. As in old crossbar telephone offices, a routing path is readily set up through such a matrix, but accessing any bit in such a matrix intended to serve as a memory is defeated by "sneak paths." Structures of this type are not new at a more macro level: 40 years ago researchers knew that blocking diodes were required in a fuse-based read-only memory; the need was rediscovered in "ovonic" memories using amorphous glasses. Logic is difficult even with active elements having only two terminals, as tunnel diodes (also known as Esaki diodes) showed.

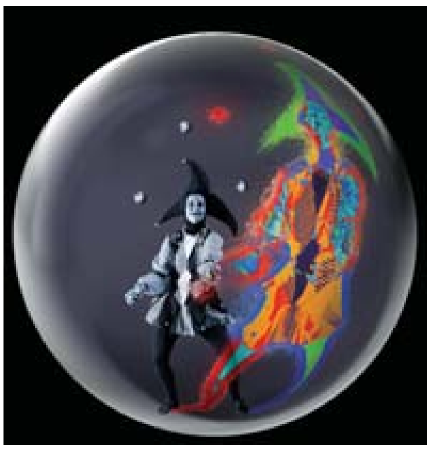
The maturity (read "huge existing investment") of the incumbent technology will make any new processes requiring changes across the whole engineering infrastructure unlikely for many years, if ever.

R. C. Foss
Ottawa, Ontario

KUEKES, SNIDER AND WILLIAMS REPLY: Our article is an extremely abbreviated and fairly nontechnical description of a body of research performed by dozens of scientists that has been ongoing for nearly a decade at Hewlett-Packard Laboratories and elsewhere. We are aware of everything pointed out by Foss but did not have the opportunity to describe all the details in the available space. They can be found, however, in the nearly 100 articles we have published in refereed technical journals and the 50 patents we have been issued in this area.

PLATONIC GRAVITY

When I read in "The Illusion of Gravity," by Juan Maldacena, about the holographic model, I immediately thought: Plato had it exactly wrong. In his allegory of the prisoners in the cave, the two-dimensional shadows on the wall of the cave are mistaken for reality. In contrast, the holographic model says, in a sense, that the "shadows" are the reality. It is the three dimensions we think we see that are the illusion.



PULLING STRINGS: **Could gravity be a holographic illusion?**

Bill Grundmann
Chelmsford, Mass.

MILKING GENETICS

In "The Land of Milk and Money," Gary Stix speculates about the possibility of producing a therapeutic protein, ATryn, in the milk of genetically modified goats. Certainly reducing the cost of medicine is an admirable goal. If such a production process goes forward, however, I hope that very strict quality assessment standards are applied. In particular, ATryn from genetically modified goats should be identical to that derived by traditional methods, rather than merely "substantially equivalent." Genetically engineered foods and other products are considered substantially equivalent even though they contain novel proteins or other substances not found in the original. Allowing such substances in medicines intended for the seriously ill should be considered a grave risk and therefore not acceptable.

Hugh Lehman
Guelph, Ontario

RETRACTION The editors of *Scientific American* have removed Woo Suk Hwang as Research Leader of the Year on the December 2005 Scientific American 50 list. Long after that list went to press, Hwang's seemingly landmark accomplishments in stem cell research were revealed as fakes. Hwang's deceptions misled *Scientific American* along with the rest of the international scientific community. A full statement of retraction is available online at www.sciam.com in the News section for December 15, 2005.

ERRATA In "The Illusion of Gravity," by Juan Maldacena, the description of a DVD having rings of magnetized dots is actually that of a computer hard drive.

In "Case Cracked" [Working Knowledge], it was said that cashews grow inside a peduncle, a false fruit. They grow on the outside.

CLARIFICATION In "Preparing for a Pandemic," by W. Wayt Gibbs and Christine Soares, NexBio's Fludase was described in the box "New Flu Drugs" on page 54 as a drug that blocks the sialic acid receptor. Specifically, it is an enzyme that works by disabling this receptor. The cell-cultured vaccine being developed by Protein Sciences found in the box "New Vaccine Technologies" on page 49 has been qualified by the Food and Drug Administration for accelerated approval, which would make market introduction likely in 2007.

PHOTOGRAPH BY CARY WOLINSKY; PHOTOILLUSTRATION BY BRYAN CHRISTIE DESIGN

50, 100 & 150 Years Ago
FROM SCIENTIFIC AMERICAN

Geology Uproar ■ Big-Gun Battleships ■ Horsepower for Horse Dung

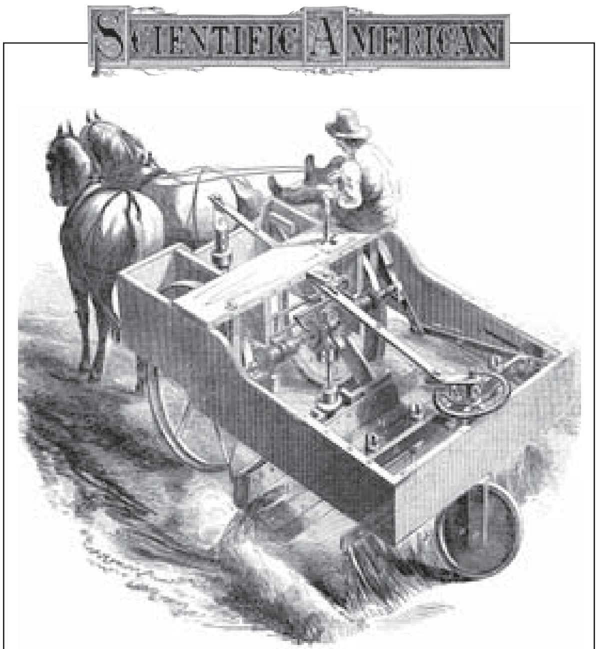
MARCH 1956

COMBAT STRESS—"Military commanders must have a thorough understanding of the effects of combat stress to make the most efficient use of their manpower. With such knowledge they could judge whether a unit was prepared for battle, how long it could fight effectively and how much rest it should be given before being sent into action again. During the Korean war the U.S. military services made a study of combat stress—a first step in a long-range program. One thing is certain: recovery from the acute stress of combat is a matter not of hours but of days. The time-honored rest cure for fighting troops—a hot meal and a good night's sleep—evidently is not sufficient."

EARTH IN UPHEAVAL—"The third in the series of Immanuel Velikovsky's unconventional re-interpretations of earth history has been published, and it is not necessary to read many pages to find that it is as generously packed with nonsense as were its predecessors. After the publication of the first of these books, *Worlds in Collision,* in 1950, emotions have run high; the scientists have said that Velikovsky is crazy; the publishers have said that the scientists are intolerant; Velikovsky implies that he is a martyred genius; the public-at-large is confused. But a scientific controversy implies that scientists argue with scientists, not that scientists argue with editorial writers, literary critics and copy-writers for book-jacket blurbs."

MARCH 1906

THE BATTLESHIP ERA—"The recent launch of the battleship 'Dreadnought' at Portsmouth, England, was an event of more than common significance; for the ship is an entirely new type, and introduces a new era in the art of warship construction. The 'Dreadnought' is the fastest and the most powerfully protected and heavily gunned warship yet constructed. It is in the armament of the 'Dreadnought' that the most striking advance has been made. It was demonstrated in the battles of the Japanese war, that the vital damage was done by shells of 10- and 12-inch caliber, smaller guns proving largely ineffective at the great ranges at which the battles were fought. Therefore the 6-inch guns (the secondary armament) have been abolished altogether."



STREET SWEEPING, by machinery and horsepower, 1856

GREAT ZIMBABWE—"Some puzzling ruins have lately been discovered in Rhodesia. The so-called 'temple' at Zimbabwe (which translates as 'houses of stone') is perhaps the best example of the architecture employed. The objects found in these ruins embrace a large variety, including iron and brass cannon, silver utensils, crockery, beads, glass, etc. These indicate the presence of the Portuguese at some time. Articles of iron and copper are supposed to represent comparatively recent Kaffir occupation; while worked gold in plates, bangles, beads, tacks, etc., are considered to be typical of the ancient builders who, in search of the precious metal, penetrated into what was to them the uttermost part of the world."

MARCH 1856

STREET CLEANING—"Philadelphia, we believe, is the only city in America where street sweeping machines have found a permanent employment. The time is not far distant when hand sweeping in the streets will be wholly superseded by mechanism. Its liberal adoption will contribute greatly to the health and neatness of our towns. The machine herewith illustrated consists of a light three wheeled vehicle. Reciprocating brooms carry dirt up the small inclined leaf, on to the endless revolving belt, which deposits the dirt on the ground in windrows; it is easily shovelled up into dumping carts."

INDUSTRIAL GOLD—"At Coon Hollow in California, and a few other such privileged places, mountains have been leveled with the plain by this process: it consists in conducting the water to the top of a hill composed of auriferous soil; the water is let down through a strong hose of canvas, leather, or caoutchouc; a torrent is shot with the force of gunpowder by the pressure of six or seven atmospheres. The water is directed by one man, and this fluid catapult demolishes and crumbles to pieces avalanches of soil from the hill side. This deluge of mud finds its way through a large sluice, where the gold is found arrested in its course by riffles."

news SCAN

STEM CELLS

Down in Flames

CAN STEM CELL RESEARCH RECOVER FROM WOO SUK HWANG? **BY SARA BEARDSLEY**

In the fledgling world of embryonic stem cells, where Woo Suk Hwang of Seoul National University once took the field by storm, scientists who had faltered behind him are just now realizing why: his data were faked. Hwang's acknowledgment of fabrication and under-the-table dealings late last year eliminated from the record one of the field's most promising therapeutic advances—patient-specific stem cell lines—and left many wondering how powerfully this crisis, one that ethics expert Jonathan Moreno of the University of Virginia calls "the greatest conduct disaster in microbiology," may swing the debate over embryonic stem cells (ESCs).

DISGRACED: Subway passengers in Seoul watch Woo Suk Hwang defend himself on December 16, 2005. (The words on the monitor read, "The sex of the stem cells matched perfectly.") An inquiry later found his data were falsified.

HAN SANG-KYUN *AP Photo/Yonhap*

Starting in 2004, Hwang and his colleagues reported stunning advances in somatic cell nuclear transfer (otherwise known as therapeutic cloning), in which the nuclei of adult cells were put into embryos to create stem cell colonies specific to patients. Following last year's revelations that only one of those successes actually occurred—the cloning of a dog—the stem cell community has been struggling to gauge the impact.

At worst, says Alan Colman of ES Cell International in Singapore, "it may cause a tainting of the whole field" if the public confuses the tiny corner of research spearheaded by the Koreans with mainstream work on embryonic stem cells harvested from preestablished lines. That work, the only kind funded by the U.S. government, is not as scientifically or ethically precarious because it does not demand new human eggs. ESC research opponent Andrew Fergusson, president of the Center for Bioethics and Human Dignity, corroborates that outcome, predicting that this debacle will "make the average American less likely to support stem cell research" when financial investments, ethical tightrope walking and lack of scientific proof are taken into account. Such a reversal would be disastrous for U.S. researchers, who rely on the public's enthusi-

news SCAN

STEM CELLS, ETHICS AND MONEY

The stem cell scandal has also put a lens on the role played by the University of Pittsburgh's Gerald Schatten, Woo Suk Hwang's U.S. collaborator on the experiments published in *Science* in 2005. Schatten was the first to hint publicly at wrongdoing by asking that *Science* remove his name from that paper, a request that struck researchers as odd, similar to "saying you're not the father of your kids," says Douglas Melton of the Harvard Stem Cell Institute.

Schatten may not have been deeply involved in Hwang's work but instead might have had a financial reason to be a senior author. In April 2003 Schatten filed for a U.S. patent on transgenic cloning methods that were, according to the language of the patent application, "directed to methods of using transgenic embryonic cells to treat human diseases." He was thus bound to benefit from Hwang's stem cell publications, states Merrill Goozner of the Center for Science in the Public Interest. That is because those studies "got the whole field moving forward, and this [patent] was tangential to that." Moreover, after the scandal broke, the *Chosun Ilbo,* the most widely circulated paper in Korea, ran reports that Schatten had demanded a 50 percent stake in the international patent filed separately by Hwang—a demand that Hwang refused—a few months before.

In December 2005 the University of Pittsburgh launched an investigation into Schatten's role in the affair, including his commercial interests that had not been disclosed in the published paper. Schatten could not be reached for comment for this article.

asm—translated into private donations and state-sponsored legislation—rather than federal dollars for support.

But at best the scandal has only "set back the clock" on therapeutic cloning, so that "the field is wide open," says Evan Snyder of the Burnham Institute in La Jolla, Calif., who will continue to pursue research similar to that of the Koreans. Tailored stem cell colonies are considered a crucial way "to study pathology in a petri dish, so you can make all kind of advances that are hard to predict otherwise," explains Douglas Melton, co-director of the Harvard Stem Cell Institute; he says that the obstacles to succeeding where Hwang failed are principally technical, not biological, and that the money being poured into the work is still money well spent.

Whether funding sources will agree remains to be seen. Moreno points out that "the effect won't be as great as might have been the case a year ago," because four states—California, Connecticut, Illinois and New Jersey—have already made a financial commitment to embryonic stem cell work. (Still, New Jersey recently tabled its 2005 stem cell measure.) As for institutional backing, opines bioethicist Arthur Caplan of the University of Pennsylvania, "patient advocacy isn't budging"—certainly a positive note for the embattled science.

Beyond the future of research, the bigger issue may be how the scientific community will address the apparent lack of safeguards against misconduct. Defenders note that Hwang's faulty science would have been caught eventually, when the experiments defied replication by other, independent parties. And in terms of ethical responsibility, they point to the National Academy of Sciences's guidelines, published voluntarily seven months before Hwang's fraud, and to the fact that it was caught by other scientists, as evidence of the community's self-correcting nature.

Yet the peer-review process—required to publish papers in scientific journals—is not designed to expose outright wrongdoing, even the staunchest advocates have to admit. They agree that had whistle-blowers not come forward, Hwang's falsified data and unethical means of egg procurement might have gone unnoticed. And this revelation, in turn, has recast the spotlight on missing legislation at home and abroad.

"It's become the Wild West out there, with each state doing what it pleases," says Steven Teitelbaum of Washington University in St. Louis, who has lobbied for changes in the Bush administration policy. "We have nothing that assures the research will be done ethically—laws should be passed on this." Others, including Caplan, believe that international treaties will be necessary to head off concerns over egg sales. One danger is that without oversight, nations may pull away from the international stem cell exchange and cooperative research altogether. "If there are differences in standards, countries could turn isolationist," Colman says.

That slowdown is certain to occur in at least one arena. After seeing how Gerald Schatten of the University of Pittsburgh, a senior co-author who purportedly played a minor role in Hwang's experiments, was carried along in the downward spiral, Moreno says, "people will think twice about collaborating." Potential co-authors of the future may painstakingly assess a project before consenting to give their names—and journals may be pressed to monitor more carefully the contributions of all involved. As for scientific relations with South Korea specifically, Snyder reports that "some of our philanthropic support sent a message: essentially, 'Don't work with the Koreans.' They have no problems with the field, but the Koreans are radioactive now."

Changes there may have to start from the ground up, where the culture is "saturated with distorted patriotism and ultra-nationalism," wrote one Seoul National University professor in a *Korea Herald* editorial. Some even predict a pendulum shift in the way science is conducted. "We'll see very strict regulation set up in Korea," Caplan speculates. "They'll overemphasize high standards."

For all the questions raised by the offenses, many in the business hope that the public can simply home in on the true offender: Hwang himself. Scientists doubt that the man's reputation will ever recover, because "he had every opportunity to come clean, but he went on blaming other people," Colman points out. Ultimately, Moreno asserts, "this is not about profound questions or about ethical line crossing for research. It's about something we can all agree on: we shouldn't lie."

OUTBREAKS

Going to Bat

NATURAL RESERVOIR FOR EMERGING VIRUSES MAY BE BATS

BY CHARLES Q. CHOI

Bats are creatures of the night that are commonly held in fear. At first glance, those fears might seem to have some medical justification. Long known as vectors for rabies, bats may be the origin of some of the most deadly emerging viruses, including SARS, Ebola, Nipah, Hendra and Marburg. Instead of demonizing bats, however, research shows the real culprit behind these outbreaks could be human error.

GOOD WAY TO GET SICK? Live fruit bats hang in a Sumatran food stall. Having the bats in such close proximity to people and live animals for sale in the market could be the reason new viruses have emerged to infect humans.

REUTERS/CORBIS

The Nipah and Hendra viruses were the first emerging diseases linked to bats. Hendra claimed two of its three victims in its first and so far only known appearance in Australia. Meanwhile Nipah has in repeated Southeast Asian outbreaks killed nearly 200 people, and blood tests of wildlife have suggested that the viruses came from the largest bats, flying foxes.

The connection to SARS, or severe acute respiratory syndrome, was less direct. During the outbreak that began in China in 2002, investigators found that civets and two unrelated species harbored the SARS coronavirus, prompting mass culling of the mongooselike civets by the Chinese government. Subsequent research, however, found no widespread SARS infection among wild or farmed civets, indicating that the disease arose in another species and might remain in wait there.

From research with Nipah and Hendra, virologist Linfa Wang of the Australian Animal Health Laboratory knew bats could get chronic infections from the viruses while not getting sick, making them ideal carriers for disease. Bats, civets and a menagerie of other animals were often found caged near one another in live-animal markets in Asia. So Wang hypothesized that bats might harbor SARS as well.

Wang and his colleagues analyzed blood, throat and fecal swabs from 408 wild bats from China. Genetic analysis revealed five bats, which represented three of nine species of horseshoe bats tested, possessed viruses closely related to SARS. They reported last September that the genetic variation within those coronaviruses was far greater than that seen in human or civet SARS. Therefore, bats, probably having lived longer with the diseases, may be the origin of the coronaviruses seen in other species.

Then, in December, researchers connected fruit bats to Ebola, whose origin in the wild had remained unknown since its first recorded appearance 30 years ago. During the Ebola outbreaks in humans, gorillas and chimpanzees between 2001 and 2003 in Gabon and the Republic of the Congo, a team led by virologist Eric M. Leroy of the International Center of Medical Research in Franceville, Gabon, tested some 1,000 animals. Of 679 bats studied, 16 had antibodies against Ebola, and 13 others possessed Ebola gene sequences in their liver and spleen. The sequences demonstrated genetic diversity, "indicating that Ebola probably has spent a long time within bats, suggesting that bats might be the origin," Leroy says. Virologist W. Ian Lipkin of Columbia University notes that scientists suspect that the Marburg virus, a relative of Ebola, also originated in bats.

Leroy vigorously argues that bats should not be culled. Wang agrees, observing that bats play critical ecological roles, such as eating insects and other pests. Besides,

FEAR OF THE NEW

Ebola, SARS and other viruses recently linked to bats "are scary diseases," states senior research scientist Jon Epstein of the Consortium for Conservation Medicine. But he remarks that "far more people die of malaria, cholera and influenza. We need to maintain a sense of perspective with regards to the global burden of infectious diseases." Most emerging diseases transmitted from animals to humans actually come from domesticated animals and carnivores, whereas bats currently account for only 5 percent of such infections, explains the consortium's executive director Peter Daszak. "With bats making up around a fifth of all mammal species, this means they are in fact underrepresented as carriers of emerging diseases," he says.

Wang points out, culling is simply not practical when it comes to bats, which can just fly away. Satellite collars on fruit bats carrying Nipah showed they could fly between Thailand, Sumatra and Malaysia, and the horseshoe bats linked with SARS range across Asia, Europe and Australia.

Preventing future emergences may instead focus on human behavior. Just as SARS is potentially linked to animal markets, so was Nipah linked to pigpens encroaching on bat habitats. And people living in Ebola-endemic areas eat the bats harboring the virus. Knowledge that bats can carry dangerous viruses could work to prevent epidemics, notes Peter Daszak, executive director of the New York City–based Consortium for Conservation Medicine, which studies the connection between emerging diseases and human interactions with the environment. Keeping bats from the wildlife trade might have dramatically cut the risk of SARS emerging, perhaps saving $50 billion worldwide in loss to travel, trade and health care costs "and hundreds of lives," Daszak says.

Charles Q. Choi is a frequent contributor.

PHYSICS

Ion Power

ATOMIC IONS PROVE THEIR QUANTUM VERSATILITY BY GRAHAM P. COLLINS

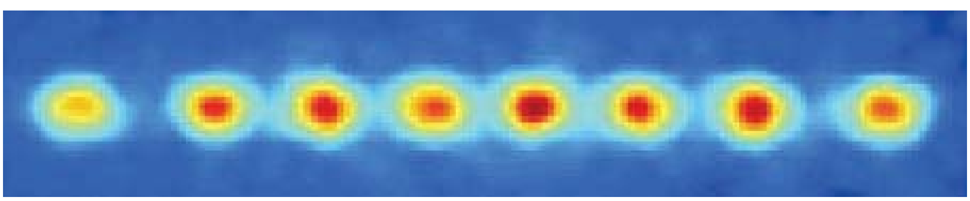
In their quest to build a computer that would take advantage of the weirdness of quantum mechanics, physicists are pursuing a number of disparate technologies, including superconducting devices, photon-based systems, quantum dots, spintronics and nuclear magnetic resonance of molecules. In recent months, however, teams working with trapped atomic ions have demonstrated several landmark feats that the other approaches will be hard-pressed to match.

A quantum computer operates on quantum bits, or qubits, instead of ordinary bits. A qubit can be not just 0 or 1 but also a superposition of the two, in which proportions of zero-ness and one-ness are combined in a single state.

An important class of multiqubit superpositions are entangled states. In these configurations, the state of each qubit is linked in a subtle way to the state of its companions, a linkage that Albert Einstein disparaged as "spooky action at a distance." For example, in a so-called Schrödinger cat state, all the qubits will give the same result—0 or 1—on being measured, even though the choice between 0 and 1 is totally random. (The name comes from the famous thought experiment in which 0 and 1 correspond to the cat being dead or alive and the individual "qubits" are all the particles in the cat's body.)

Cat states are a fundamental building block of techniques for correcting errors in qubits. Such errors inevitably plague all the standard approaches to quantum computation, because states of qubits are exceedingly fragile.

Researchers at the National Institute of Standards and Technology in Boulder, Colo., led by David J. Wineland and Dietrich Leibfried, have now created cat states involving four, five and six beryllium ions.



ENTANGLED: Eight calcium ions held together in a trap are in a special quantum condition known as a W entangled state, in which their properties are subtly correlated. Such states are of use for error-correction schemes in quantum computers. Entangled states become harder to create and maintain as the number of particles increases.

HARTMUT HAEFFNER *University of Innsbruck*

NEED TO KNOW: SCALING UP

Experiments with atomic ions involve custom-built, bulky electromagnetic traps to confine the ions in a vacuum. Though fine for experiments with a small number of ions, they are utterly impractical for the large-scale system that a quantum computer would need to be of any significant use. Now University of Michigan at Ann Arbor researchers Christopher Monroe, Daniel Stick and their co-workers have demonstrated a 100-micron-size ion trap on a semiconductor chip. They used their chip to trap a single cadmium ion and move it to different locations in the trap by applying electrical signals to electrodes. The trap was built using standard lithography techniques, so, Monroe says, it could be scaled up to include hundreds of thousands of electrodes using existing technology.

An electromagnetic trap holds the ions in a row in a vacuum, and lasers manipulate their states. The team estimates that their six-ion cat states last for approximately 150 microseconds.

In Austria, Rainer Blatt and Hartmut Haeffner of the University of Innsbruck and their colleagues relied on a similar technique to produce an entangled state of eight calcium ions. In this experiment a "W state" was created, not a cat state. A W state is in many ways more robust than a cat state. For example, an ion can be lost from a W state and the remaining ions will still be in a W state. Losing an ion from a cat state spoils the entire state.

An important feature of both experiments is that in principle the techniques can incorporate larger numbers of ions. An impediment to scaling up these approaches, however, was that the quality of the entangled state decreased as the number of ions increased. To reduce this error, the scientists might adjust the details of the laser pulses, use different states of the ions to represent 0 and 1, or work with a different ion species altogether.

For a quantum computer to be of use, one must not only create special qubit states but also manipulate them in ways that preserve their quantum characteristics. That is, one must run quantum algorithms on the computer. A group at the University of Michigan at Ann Arbor led by Christopher Monroe and Kathy-Anne Brickman has now demonstrated an algorithm known as Grover's quantum search on a system of two trapped cadmium ions.

The search algorithm rummages through a database with entries in random order. Searching for a particular item would usually demand the examination of every entry. The quantum search algorithm is magically faster because the quantum computer can poll all the database entries at once in a superposition. The speedup becomes more dramatic for larger databases. For example, a million-entry database would take only about 1,000 quantum lookups instead of the full million.

The Ann Arbor experiment operated on the equivalent of a four-entry database, the four entries being represented by two qubits. The researchers say that their system can be scaled up to larger numbers of qubits.

With results coming so thick and fast, it is no wonder that, as Monroe says, "many feel that ion traps are well ahead of other technology in the quest to build a large-scale quantum computer."

INFOTECH

Software Insecurity

OUTSOURCING AND DEFENSE FEAR OF "FOREIGN INFLUENCE" BY DANIEL G. DUPONT

In February 2005 a group of Pentagon industry advisers warned that the "migration of critical microelectronics manufacturing" from the U.S. to other countries compromised national security. To ensure a steady supply of safe microchips, the Defense Science Board—which advises senior defense officials—recommended establishing "trusted foundries" to make critical hardware. But that is only part of the picture. According to the science board, any effort to improve the safety and supply of microchips would be of "limited utility" without a comparable focus on software—especially on what the Pentagon calls "foreign-influenced software."

The Department of Defense once created its own software, but today only the most highly classified code is written in-house, at places such as the secretive National Security Agency. But a good deal of code for some of the military's most sophisticated weapons—fighter aircraft and missile defense systems, for example—is written in other countries.

In 2004 the Government Accountability Office (GAO) found that the military "is experiencing significant and increasing reliance on software and information systems for its weapon capabilities, while at the same time traditional DOD prime contractors are subcontracting more of their software development to lower-tier and sometimes nontraditional defense suppliers." Those suppliers, the GAO added, use "offshore locations and foreign companies" for some software development.

Software developed overseas can be manipulated in several ways, says Nancy Mead, a senior member of the technical staff at the Carnegie Mellon Software Engineering Institute. The code itself can be tampered with and set up to do subsequent damage; it can also be laced with surreptitious "back doors" designed to allow access to a system at a later date. And the possibility exists that software could be copied and sold to adversaries.

"You don't have day-to-day control over what's going on" at some overseas facilities, Mead notes. U.S. companies that look to foreign suppliers must keep an eye on the software-development process as much as possible, she says, because the development phase is the point at which errors or intentional flaws can most easily be prevented. Complex software contains millions of lines of code, and "it becomes more difficult" to spot such flaws later on, Mead explains: "At that point you're just looking for a needle in a haystack."

According to a former Pentagon official who requested anonymity, software written abroad has become the subject of high-level discussions and secret threat assessments within the DOD. The department went back to its science board last October for a look at both why the military has become so dependent on software of "foreign provenance" and what is currently being done to test it. The board will probably finish its analysis sometime this year.

Leading the science board study is Robert Lucky, an esteemed engineer, author and research consultant. Lucky says he was concerned the military might deem too many systems as "mission-critical," meaning that they must have the highest levels of software security. Such classification would make the task of ensuring security that much harder—and more expensive. Lucky and his panel will have to address that question of resources: the Pentagon has asked them to evaluate the investments the DOD could make to increase confidence in military software. Like many choices in life, "it all comes down to money," Lucky remarks. "How much security can you get for how much money?"

Daniel G. Dupont edits InsideDefense.com, an online news service.

JOINT STRIKE FIGHTER, an international project, contains software code written outside the U.S.

LOCKHEED MARTIN

SYSTEMS CHECK

Debugging is notoriously difficult, and researchers are devising new ways to test the complex code used by the military. One such method comes from Ted L. Bennett and Paul W. Wennberg of Triakis Corporation in Redmond, Wash., which makes simulators and provides software validation services. Writing in the *Journal of Defense Software Engineering* last year, they described a research project sponsored by NASA in which they showed how simulations designed to verify system designs can also be used to probe software for faults. Their method allows users to check software for code that is not being utilized. "It either shouldn't be there," Bennett says, "or it should be tested."

Meanwhile a multiyear research project at the Carnegie Mellon Software Engineering Institute focuses on "function extraction"—a means of describing software functionality to determine exactly what it should do and a way to counteract "back doors" or other undesirable software modifications.

LUNAR SCIENCE

Crater Jumper

HOPPING PROBE MAY HUNT FOR ICE ON THE MOON **BY MARK ALPERT**

The entire future of human space exploration may rest on a patch of lunar ice. For the past two years NASA has focused on designing a new crew vehicle and launch system that could return astronauts to the moon by 2018. The agency's ultimate goal is to establish a permanent lunar base and use the program's technology to prepare a human mission to Mars. But the grand plan hinges on a risky prediction: that NASA will find water ice in a permanently shadowed crater basin at one of the moon's poles.

Plentiful ice deposits would be a boon for lunar colonists, who could use the water for life support or convert it to hydrogen and oxygen rocket fuel. And two orbiters sent to the moon in the 1990s, Clementine and Lunar Prospector, found evidence of ice in perpetually shadowed polar areas where consistently frigid temperatures would preserve the water carried to the moon by comet and meteorite impacts. But some scientists have disputed Clementine's radar data, and the anomalous neutron emissions observed by Lunar Prospector could have been caused by atomic hydrogen in the lunar soil instead of ice.

In an attempt to settle the question, NASA plans to launch the Lunar Reconnaissance Orbiter (LRO) in 2008. Traveling in a polar orbit only 50 kilometers above the moon's surface, the one-ton, $400-million probe will train a high-resolution neutron sensor on the suspected ice deposits to determine their locations more precisely. The LRO will also carry a radiometer to measure surface temperatures, an ultraviolet detector to peer into the shadowed crater basins, and a laser altimeter and camera to map the polar regions and to scout possible landing sites.

But because the ice is probably buried and mixed with the lunar dirt, NASA will need to land a probe that could dig up and analyze

LUNAR ORBITER, scheduled for launch in 2008, will scout landing sites at the moon's poles.

soil samples. This mission, scheduled for 2011, is a challenging one given that instruments operating in shadowed areas cannot use solar power. The craft could land at a sunlit site and send a battery-powered rover into a dark crater, but the batteries would quickly die. A radioisotope thermal generator could provide electricity using heat from plutonium decay, but NASA is leaning against this option because it is expensive and controversial.

Another idea under consideration is sending a probe that could hop from place to place on the lunar surface by restarting its landing rockets, which could lift the craft up to 100 meters above its original landing site and move it to another spot in the crater basin to hunt for ice. Investigating more than one site is crucial because the ice may be unevenly distributed. Yet another alternative would be to fire ground-penetrating instruments at several places in the shadowed basin, either from a lander at the crater's rim or from an orbiting craft.

The major pitfall of NASA's strategy is the possibility that its probes will find no ice or discover that the ice is too sparse to be a useful resource. "I'm a little worried that they're counting too much on finding water in a usable form," says Wesley T. Huntress, Jr., a former NASA science chief who now leads the Carnegie Institution's Geophysical Laboratory. If extracting ice from the moon proves infeasible, the space agency may have to choose new landing sites and exploration goals for the human missions. But NASA believes the robotic craft will be worthwhile even if they do not find ice. Says Mark Borkowski, head of the Robotic Lunar Exploration Program: "We can't help but get science from these measurements."

CHRIS MEANEY/GSFC/NASA

THE SURVEYOR THAT HOPPED

NASA has already performed a successful hopping maneuver on the moon. On November 10, 1967, the unmanned Surveyor 6 spacecraft landed in the Sinus Medii ("Bay of the Center") region to take television pictures of the surface and study the lunar soil. A week later mission controllers fired the craft's low-thrust vernier rockets for 2.5 seconds to practice a powered takeoff from the moon. Rising about four meters above the lunar surface, the probe hopped to a new landing spot 2.4 meters away, touched down on its legs and continued operating.

ASTROPHYSICS

In the Groove

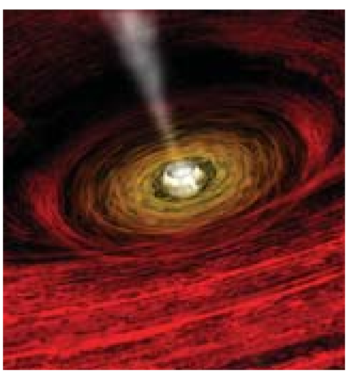
MEASURING A BLACK HOLE'S SPIN BY A CREASE IN SPACETIME BY JR MINKEL

The fortuitous flare-up of a black hole last active in the mid-1990s solidifies an assumption sometimes used to estimate a black hole's spin, one of its two most important properties. (The other is mass.) As matter whips around a black hole, it radiates light perpendicular to its orbital radius, like a lighthouse. The mass and spin of the black hole can furrow a groove in spacetime that makes this orbit wobble in various ways, creating additional fluctuations in its radiation.

In 1996 a black hole system called GRO J1655-40 broadcast a pattern of x-ray oscillations suggestive of such a groove, then clammed up a few months later. In 2005 gas from a companion star got caught in the black hole's drain again, allowing researchers to observe the reenergized system over eight months. Sure enough, they discovered the identical pattern. "Detecting the same frequencies nine years later means what we are looking at here is really a fundamental property," not some gaseous mirage, says Jeroen Homan of the Massachusetts Institute of Technology. His team reported the results on January 9 at the annual meeting of the American Astronomical Society.

The group will now try to determine if factoring in these fundamental frequencies can give them better measurements of the black hole's spin.

JR Minkel is a frequent contributor.



MATTER SPIRALING into a black hole produces beams of light. In the black hole system GRO J1655-40, matter followed the same path as it did nine years previously.

A. HOBERT/CXC/NASA

BIOLOGY

Where the Bacteria Roam

Despite bacteria's ubiquity, their diversity in the world's soils is poorly understood. To get a handle on what makes the organisms thrive, Duke University researchers trekked far and wide to collect a few centimeters of dirt from 98 locations across North and South America, then analyzed each sample for genetic variation. To their surprise, the strongest predictor of high diversity was neutral pH. The acidic soil of the Peruvian Amazon, for example, harbored one half to one third as many species as did the neutral dirt of the arid American Southwest. "There are a lot of variables that didn't turn out to be very important," says co-author Robert Jackson, who adds that a more exhaustive search of different habitats might turn up other stimulators of diversity, such as carbon abundance. The report was published online January 9 by the *Proceedings of the National Academy of Sciences USA*.

—JR Minkel

WHALES

Sense and Sensitivity

The narwhal sports an eight-foot-long spiraled tooth that makes it resemble a unicorn of the sea. Some thought that the whale, typically 13 to 15 feet long, used it to break arctic ice; others theorized that it served as a lance in male jousts. The tooth, in fact, may be a giant sensor for navigating and hunting. Through electron microscopy of two male tusks, researchers from Harvard University, the Smithsonian Institution, and the National Institute of Standards and Technology discovered that a single horn possesses some 10 million nerves running from its surface to its core. Instead of inflicting the narwhals with a massive ice cream headache, this sensitive tooth appears capable of detecting changes in water temperature, pressure and particle gradients linked with salinity and prey. Their findings surfaced last December at the 16th Biennial Conference on the Biology of Marine Mammals in San Diego.

—Charles Q. Choi

NOT FOR JABBING: The narwhal's horn seems to be a sophisticated sensor.

GLENN WILLIAMS

PHYSICS

Nanoparticles Can't Hide

Detecting a virus or any nanosize particle usually means fixing it to a substrate or attaching a fluorescent probe to it. Neither method is practical for detecting particles in real time. Now University of Rochester physicists have assembled a simple system for doing just that. They split a laser beam in two, sending one half to a sample. When the light hit a small particle, it scattered back and recombined with the reserved half of the laser beam, producing a detectable interference pattern only when a moving particle was present. The method works where others do not, the researchers say, because it relies on the light's amplitude rather than intensity. The amplitude is the square root of intensity, so it decays much less than intensity as particles get smaller. The investigators have so far detected single particles as small as seven nanometers across. Peruse their findings in the January 13 *Physical Review Letters*.

—JR Minkel

BRIEF POINTS

- **Buddy system: A dwarf galaxy containing hundreds of thousands of stars seems to be merging with the Milky Way. The new companion galaxy lies 30,000 light-years from Earth, toward the constellation Virgo.**

  Sloan Digital Sky Survey announcement, January 9

- **Patterns of frog extinctions and lethal fungus outbreaks appear to be synchronized. To some scientists, the connection implicates global warming, which stimulates the fungal growth.**

  Nature, January 12

- **Just hand me the leash and go away: nursing home residents felt much less lonely when they spent time with a dog, rather than with a dog and other people.**

  Anthrozoos, March

- **Seeing a person who behaved unfairly get an electric shock triggered the empathy areas in women's brains, but in men the reward centers were activated instead. Men may find pleasure in retribution, although both sexes reported disliking the unfair person.**

  Nature online, January 18

news SCAN

DATA POINTS: PLANETARY STRESS

For its *State of the World 2006* report, the Worldwatch Institute focuses on China and India. The U.S. still consumes the most resources per capita, but if China and India were to catch up, then the resources from a second planet Earth would be needed to sustain the two economies.

Gross domestic product per person in:
China: $4,600
India: $2,500
Europe: $26,900
Japan: $29,400
U.S.: $40,100

Barrels of oil used per person every year in:
China: 1.9
India: 0.9
U.S.: 25.3

Kilograms of grain consumed per person in 2005 in:
China: 292
India: 173
Europe: 561
Japan: 354
U.S.: 918

Ecological footprint* per person in hectares in:
China: 1.6
India: 0.8
Europe: 4.7
Japan: 4.8
U.S.: 9.7

*Ecological footprint refers to areas with significant photosynthetic activity or biomass accumulation.

SOURCE: State of the World 2006

INFECTIOUS DISEASE

Testing Lethality

Plates of sugar could indicate the deadliness of a flu strain. The influenza virus infects the body using hemagglutinin, a viral protein that latches onto sugars containing sialic acids on human cell surfaces. Scripps Research Institute investigators and their colleagues developed arrays containing 200 different carbohydrates and sugary proteins, representing the major types of molecules to which hemagglutinin might attach. They tested eight different flu strains, including the deadly 1918 influenza, and discovered that alterations at as few as two positions on the hemagglutinin protein can transform a bird strain into one that can infect humans. The scientists, who report the work in the February 3 *Journal of Molecular Biology,* say that the arrays could explain why the 1918 flu, which strongly resembles avian viruses, was so devastating. They could also lead to monitors that quickly determine how close another bird strain is to mutating into a pandemic form. *—Charles Q. Choi*

ECOLOGY

Reefer Sanity

PARROTFISH enable coral to thrive by feeding on seaweed, which inhibits coral growth.

Faced with declines in coral reefs from pollution, overuse and climate change, reef managers have begun implementing "marine protected areas"—no-fishing zones. One of the first studies of the ecological effect of such zones shows how they might help. A group led by researchers at the University of Exeter in England studied a sea park in the Bahamas that has not been fished since 1986. An open question was whether the resurgence of predatory grouper fish would kill off helpful parrotfish, which foster reef growth by grazing seaweed. The team found that net grazing actually doubled, coinciding with a fourfold reduction in seaweed proliferation in the protected area compared with fishable areas. The reason: larger parrotfish became more plentiful after fishing ceased. The grouper have trouble feeding on the bigger fish, which in turn graze seaweed more extensively than smaller ones. Dive into the January 6 *Science* for more. *—JR Minkel*

PERCEPTION

What You See Is What You Say

Psychologists argue over whether language influences how people think. It could, however, affect half of what they see. The view from the right eye is processed in the brain's left hemisphere, which also seems to handle language. Investigators at the University of Chicago and the University of California, Berkeley, tested how well the right and left fields of view distinguish between the colors known in English as blue and green. Most of the world's languages actually use a single word for the two, suggesting that for English speakers, language influences the discrimination between blue and green. The researchers found native English speakers were faster at distinguishing bluish squares from greener ones if the differently colored square appeared within the right visual field. This effect vanished if the volunteers had to rehearse simultaneously an eight-digit number, which distracted their verbal working memory. Look for the findings in the January 10 issue of the *Proceedings of the National Academy of Sciences USA*. *—Charles Q. Choi*

SECRET SEA VISIONS/PETER ARNOLD, INC.; ILLUSTRATION BY MATT COLLINS

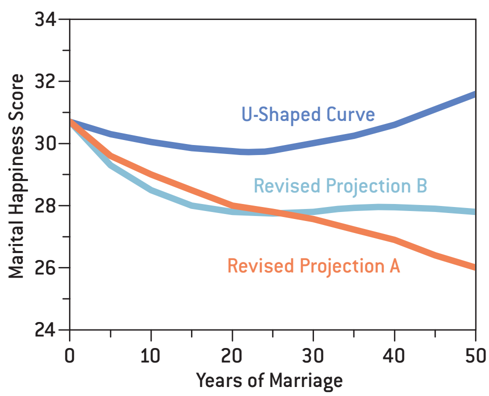
BY THE NUMBERS

The Honeymoon Is Over

AFTER THE VOWS, MARITAL HAPPINESS MAY BE ALL DOWNHILL **BY RODGER DOYLE**

Sociologists have long firmly held that marital bliss is high in the very early years of marriage, declines with the coming of children and rises in later years when children have left home. Happiness thus follows a U-shaped trajectory over the life of a marriage, as shown in the chart.

This belief derives largely from studies that employ a technique—the cross-sectional survey—unsuited to constantly changing phenomena such as marriage because it measures attitudes at only one point in time. A better method is the longitudinal study, which measures attitudes at several points in the life course. Two newer studies employing the more powerful technique now challenge the traditional U-shaped curve of marital happiness.



In the first, a group headed by sociologist Jody VanLaningham, formerly of the University of Nebraska–Lincoln, measured marital happiness using 11 questions, such as "How happy are you with the amount of love and affection you receive?" and "How happy are you with your sexual relationship?" Based on the responses, the researchers created a marital happiness score and calculated a new curve, shown as "Revised Projection A" on the chart. It suggests that happiness declines more or less continuously throughout the lifetimes of married couples. The respondents were interviewed in five waves: during 1980, 1983, 1988, 1992 and 1997. Because 1980 and 1983 were atypical—the divorce rates ran extraordinarily high—the team calculated a second life-course curve based only on the last three waves. This curve—"Revised Projection B"—suggests a leveling off of marital happiness in later years but provides no support for the U-curve theory.

Another longitudinal study, this one by sociologist Debra Umberson of the University of Texas at Austin and her colleagues, confirms the VanLaningham group's finding that marital quality declines over time. In addition, this study measured the independent effect of age—as opposed to duration of marriage—and discovered that the older the spouses, the more likely they are to have a good marriage, perhaps because they are less emotionally reactive in marital conflicts than younger people or because they better appreciate their partner's positive traits. Umberson's team found that although parenthood may have a negative effect early in marriage, it exerted a positive influence in later years after children have left. Childless couples tend to have lower-quality marriages in old age than couples who have children.

Marriage continues to be the choice of an overwhelming majority—94 percent of living Americans have tried it at least once by age 65. Aside from the immediate economic and sexual benefits, marriage confers long-term health advantages, including emotional support and a sense of belonging, which buffer against depression. Indeed, economists David G. Blanchflower of Dartmouth College and Andrew J. Oswald of the University of Warwick in England have calculated the monetary equivalent of marriage's effect on mental well-being. They find that being married is the equivalent of making an extra $100,000 a year. The benefit is greater for men than women, perhaps because they have weaker social support networks before marriage.

Rodger Doyle can be reached at rodgerpdoyle@verizon.net

RODGER DOYLE; SOURCE: VANLANINGHAM ET AL.; THE HAPPINESS SCORE IS AN AVERAGE BASED ON THE RESPONSES TO 11 QUESTIONS

FURTHER READING

Marital Happiness, Marital Duration, and the U-Shaped Curve: Evidence from a Five-Wave Panel Study. Jody VanLaningham, David R. Johnson and Paul Amato in *Social Forces*, Vol. 79, No. 4, pages 1313–1341; June 2001.

How Does Marriage Affect Physical and Psychological Health? A Survey of the Longitudinal Evidence. Chris M. Wilson and Andrew J. Oswald. Warwick Economic Research Papers, University of Warwick, May 2005. Available at www.andrewoswald.com

As Good as it Gets? A Life Course Perspective on Marital Quality. Debra Umberson, Kristi Williams, Daniel A. Powers, Meichu D. Chen and Anna M. Campbell in *Social Forces*, Vol. 84, No. 1, pages 493–511; September 2005.

Cures and Cons

Natural scams "he" doesn't want you to know about By MICHAEL SHERMER

Up to 139 times in one week, Kevin Trudeau pitches late-night viewers about his self-published book, *Natural Cures "They" Don't Want You to Know About,* a rambling farrago of uninformed opinions, conspiracy theories and cheeky jabs at medical, pharmaceutical and governmental authorities ("they"). The book is so risibly ridiculous that even the most desperately ill would not take it seriously—would they?

Apparently they would, to the tune of millions of copies sold, elevating the book to the *New York Times* best-seller list. If readers had purchased Trudeau's *Mega Memory,* perhaps they would have remembered that he spent almost two years in federal prison after pleading guilty to credit-card fraud and that the Federal Trade Commission banned Trudeau "from appearing in, producing, or disseminating future infomercials that advertise any type of product, service, or program to the public, except for truthful infomercials for informational publications. In addition, Trudeau cannot make disease or health benefits claims for any type of product, service, or program in any advertising, including print, radio, Internet, television, and direct mail solicitations, regardless of the format and duration." Trudeau had to pay $500,000 in consumer redress for his bogus infomercials and another $2 million to settle charges against him for claiming that coral calcium cures cancer (it doesn't) and that an analgesic product called Biotape permanently relieves pain (it doesn't).

"Sun block causes cancer."

Amazingly, *Natural Cures* is exempt from this injunction. "Books are fully protected speech. He can author a book and voice his opinions," says Heather Hippsley, assistant director for the division of advertising practices at the FTC who investigated Trudeau's infomercials. "The line is: Informational materials, OK. Products and services, banned."

So Trudeau is free to dole out in print such opinions as these: "Medical science has absolutely, 100 percent, failed in the curing and prevention of illness, sickness, and disease." (Smallpox is not a disease?) "Get all metal out of your dental work." (Won't this help the medical cartel?) "Sun block has been shown to cause cancer." (References?) "Don't drink tap water." (Wrong: studies show it is as safe as bottled water.) "Animals in the wild virtually never get sick." (No need to worry about avian influenza.) "Get 15 colonics in 30 days." (Can I bring a friend?) "Wear white.... The closer you get to white, the more positive energy you bring into your energetic field." (Why is Trudeau wearing all black on the book cover?) "Stop taking nonprescription and prescription drugs." (Including insulin for diabetes?) "This includes vaccines." (Welcome back, polio.) "Have sex." (Without prescription Viagra?)

This 600-page medical advice book contains no index, no bibliography and no references. In their stead are testimonials for the audio edition and a sequel in the works about "weight loss secrets they don't want you to know about."

As for the "natural cures" themselves, some are not cures at all but just obvious healthy lifestyle suggestions: eat less, exercise more, reduce stress. Some of the natural cures are flat-out wrong, such as oral chelation for heart disease, whereas others are laughably ludicrous, such as a magnetic mattress pad and crocodile protein peptide for fibromyalgia. Worst of all are the natural cures that the book directs the reader to Trudeau's Web page to find. When you go there, however, and click on a disease to get the cure, you first have to become a Web site member at $499 lifetime or $9.95 a month. It is a classic con man's combo: bait and switch (the book directs them to the Web page) and double-dipping (sell them the book, then sell them the membership).

Why don't "they" want you to know about these natural cures? "Money and power," Trudeau says. "Most people have no idea just how powerful a motivating force money and power can be." Kevin Trudeau certainly does, and this book is a testimony to that fact.

There is one lesson that I gleaned from this otherwise feckless author, well expressed in an old Japanese proverb: "*Baka ni tsukeru kusuri wa nai*"—"There is no medicine that cures stupidity." *Domo arigato,* Mr. Trudeau. SA

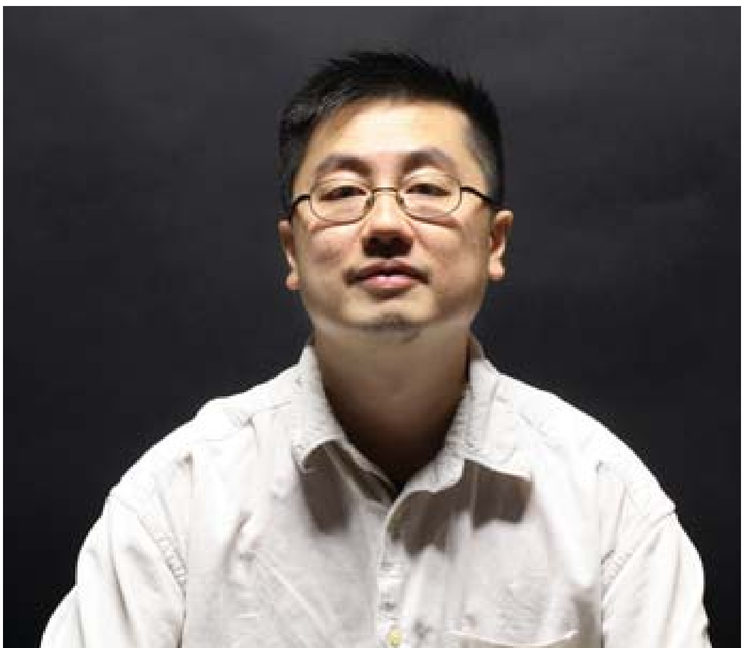
Michael Shermer is publisher of Skeptic *(www.skeptic.com) and author of* Why People Believe Weird Things.

BRAD HINES

Experiments at Work

What's the best way to boost sales or handle competing resellers? By lab-testing business ideas, Kay-Yut Chen gets rid of some of the guesswork By MARINA KRAKOVSKY

Outside Kay-Yut Chen's economics laboratory at Hewlett-Packard in Palo Alto, Calif., the November air is unseasonably warm, even for California, and splashes of yellow and green leaves shimmer against the clear blue sky. But inside, in a windowless, fluorescent-lit room, the 12 visitors participating in today's experiment sit patiently at their randomly assigned computers. When I point to the incongruity, Chen doesn't miss a beat: "That shows one thing—the assumption that people like money is correct."



TIMOTHY ARCHIBALD

KAY-YUT CHEN: HE'S GOT GAME

- Helps corporate decision making at Hewlett-Packard by modeling the economic behavior of players in games involving various payouts.
- Studied physics before turning to economics. "The experimental methods are similar. The only difference is that the 'atoms' I study have minds of their own."
- On how Isaac Asimov's *Foundation* novels continue to inspire: "The protagonists were using mathematics to predict the future of human society in the galactic empire. That is exactly what I am trying to do in my real-life work, albeit on a much smaller scale."

This simple assumption goes to work for Chen whenever the HP principal scientist runs an experiment. The participants—mainly "starving students" from nearby Stanford University—will earn $25 to $75 or more, depending on how well they play today's game. The experiment simulates interactions between sales agents and sales managers. At each period of play, the computer tells sales agents the current market conditions, and based on that information the agents must decide how much effort they will put into making the sale. Although effort incurs costs that take away from the total payout, effort also increases the likelihood of sales success. The agents' total payout is simply the sum of the fixed payment and variable payment for successful sales minus the cost of effort. Managers, in turn, determine the fixed and variable payments they will offer—knowing their own payout will be total sales minus payouts to agents.

Chen explains these rules but says nothing about strategy. He doesn't need to: through a little computer-mediated back-and-forth with their managers, most agents wise up to the fact that this game rewards sales but offers no incentive to tell managers anything. Similar compensation schemes in the real world explain why salespeople tend to sandbag their forecasts, making it hard for their companies to plan ahead.

Chen thinks he has solved the sandbagging problem: have each salesperson choose a personal balance of fixed and variable compensation. For example, the salesperson can choose a high commission percentage with no fixed salary or, at the other extreme, a modest fixed salary and no commission—or some combination in between. Each choice implicitly reveals how much the salesperson plans to sell, much as an insurance subscriber's choice of deductible and premium reveals how sick she is. Based on a truth-telling mechanism from game theory, this design works on paper. But as an experimental economist, Chen will keep testing it empirically, comparing the emerging design with other available models, such as the one he is testing today.

Chen has successfully used that approach to help HP managers design good contracts with retailers and resellers, and he is starting to tackle other thorny problems for his employer: figuring out how to protect HP's bottom line against international currency fluctuations and discovering ways for brick-and-mortar retailers and HP's online store to coexist happily.

The science of experimental economics has boomed with the rise of personal computers—especially after one of its founding fathers, Vernon L. Smith of George Mason University, won the Nobel Prize in economics in 2002 (sharing the honor with Princeton University psychologist Daniel Kahneman). But the field has not caught on in what seems its most logical application: making business decisions within a company. When Chen started his lab in 1994, it became the first economics laboratory inside a corporation, and to this day no other firm maintains a lab like his. "He's basically 'Mr. Experimental Economics in Business,'" says Teck H. Ho, chair of the marketing group at the Haas School of Business at the University of California, Berkeley.

Kay-Yut Chen's experiments showed that a proposed incentive program would backfire. So HP, his employer, scrapped the idea.

Some blame the absence of such labs on the shortsightedness of corporate America. Charles R. Plott, the experimental economist at the California Institute of Technology who pioneered the field, puts it this way: "A lot of us in academia have done business applications, but for someone to take a new science inside a big business that is just rife with competition for funds—and be able to create a science facility that can handle the basic research and the pressures of day to day—that's a monumental task Chen has succeeded in."

Were it not for Plott's influence, Chen would not have become an economist at all. As a Caltech undergraduate, the Hong Kong native majored in physics and participated in Plott's experiments only for the lure of the cash and the challenge of besting other players. But soon the science became the greater draw, and Chen briefly waxes metaphysical describing it: "Doing experiments on people is like playing a game with a higher being. Many people think they have free will, meaning you cannot predict what they will do, but the counterintuition is that people in a lot of cases are predictable." Chen also recognized more opportunity for discoveries in this younger field than in high-energy physics—and Plott easily talked him into switching over.

As he was finishing his Ph.D. at Caltech, HP Labs was trying to invigorate its work in operations research through collaborations with experimental economists. Thanks to professors singing his praises, Chen, now 39, got tapped in 1994 to start the economics lab at HP, where he has been ever since.

HP managers say the benefits of lab-testing business ideas are immeasurable. "That way you're not going to upset the business or your partners or the end users with a bad program," explains Jukka Koskela, an HP sales operations manager who asked Chen to test the idea of rewarding retailers to be number one in sales. When Chen's experiments showed this incentive would push too many retailers to either of two undesirable extremes—giving up at the outset or neglecting other HP goals—the company scrapped the idea. "You could waste millions of dollars implementing a program that isn't good," Koskela says. Chen's tests showed that contracts that reward retailers according to how much business they bring in for HP produce far better results.

In general, a good contract aligns the interests of the principal (such as HP) and the agent (such as a reseller)—but in a business full of interdependent variables, that is much easier said than done. Think of the tangled web of business relationships: office equipment distributor IKON is an HP partner, but it also sells similar products from other firms and competes with other HP partners. In the lab, such mixed motives tempt some players to "game the system," such as the time a student playing a reseller made vastly more than the average payout by hoarding products to lock out other resellers. This incident was an anomaly—and may be even less likely in the real world, where long-term relationships and other variables might prod people to honor a contract's win-win spirit. Still, Chen wants to assure good outcomes even if people are at their worst, so he tweaked the contract to prevent resellers from ordering more inventory than they can really sell.

Such tinkering is anathema to pure theorists, who see experiments as a rickety crutch for those who cannot build good formal models. Chen takes a more pragmatic view, however, and he good-naturedly brushes off criticisms of applied experimental economics. Sure, student subjects aren't exactly like real businesspeople, but if they're bright they're close enough—plus they're affordable. And the experiments are a complement to theoretical reasoning, not its substitute. As Chen sums it up, "I would say this is two parts science, one part art." SA

Marina Krakovsky (www.marinakrakovsky.com) is a writer in the San Francisco Bay Area.

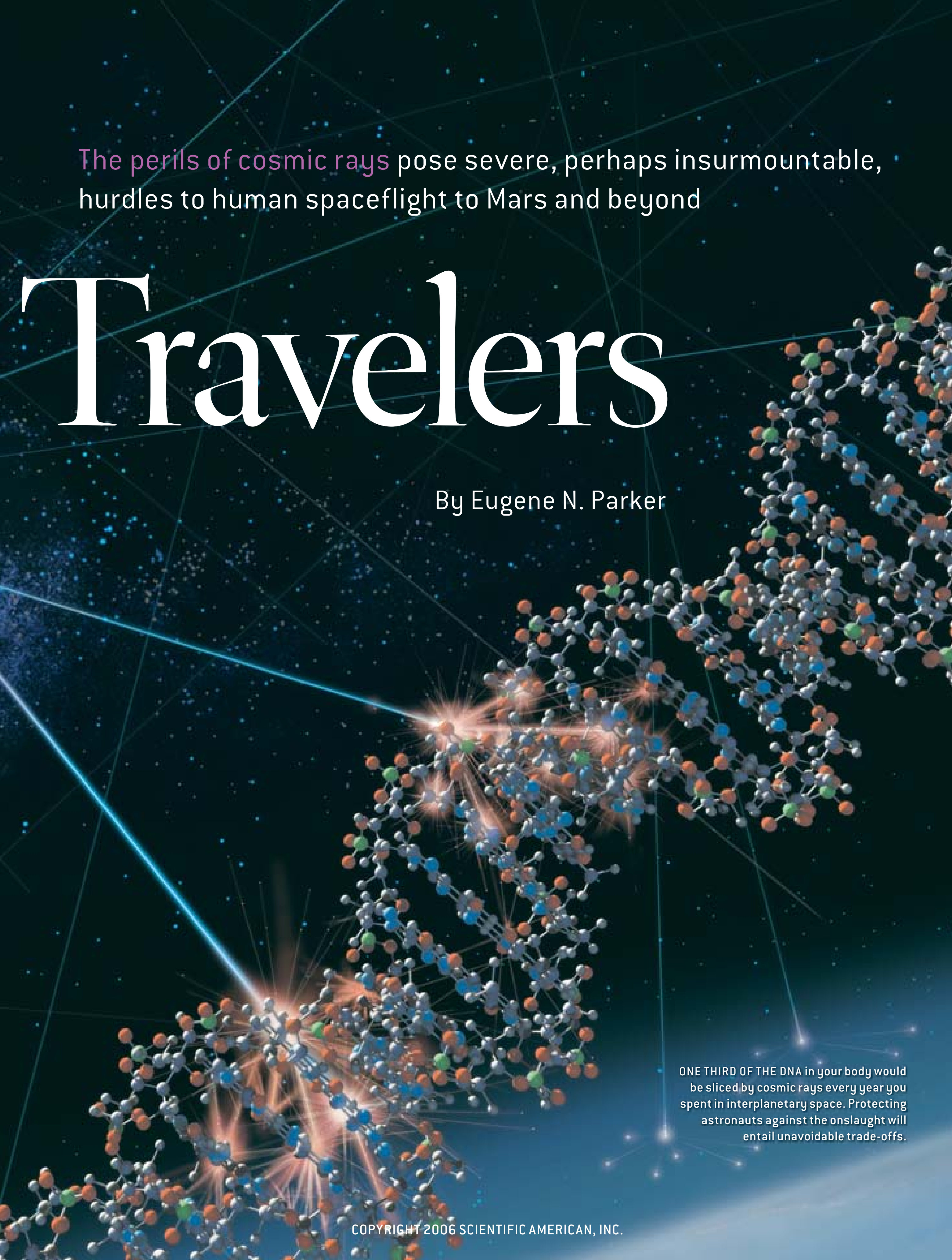
Shielding Space

In science fiction, the worst threats to space travelers are large ones: careening asteroids, ravenous creatures, imperial battle cruisers. In reality, though, the scariest menaces for humans in space are the tiniest: fast-moving elementary particles known as cosmic rays. On a long journey, they would give astronauts a dose of radiation serious enough to cause cancer. Unlike most of the other challenges of venturing into deep space, which engineers should be able to solve given enough time and money, cosmic rays pose irreducible risks, and dealing with them involves fundamental trade-offs. They could be the show-stopper for visiting Mars.

The perils of cosmic rays pose severe, perhaps insurmountable, hurdles to human spaceflight to Mars and beyond

Travelers

By Eugene N. Parker



ONE THIRD OF THE DNA in your body would be sliced by cosmic rays every year you spent in interplanetary space. Protecting astronauts against the onslaught will entail unavoidable trade-offs.

In the laboratory, cosmic rays first presented themselves as a minor annoyance. They were discovered when physicists noticed that electrically charged bodies do not stay that way; their charge slowly leaks away through the air. Something had to be ionizing the air, allowing it to conduct electricity. Many researchers blamed the ambient radioactivity of the soil and rocks underfoot. Austrian physicist Victor Hess settled the issue in 1912, when he went aloft in a balloon and showed that the higher he rose, the faster the charge leaked off his electroscope. So the cause of the ionized air was something mysterious coming in from space—thus the name "cosmic rays."

By 1950 physicists had determined that the term is actually a misnomer. Cosmic rays are not rays but ions—mostly protons, with a few heavier nuclei mixed in—striking the top of the atmosphere at nearly the speed of light. Most come from beyond the solar system, but what catapults them to such a speed remains a question to this day. Experimenters, having once regarded cosmic rays as irksome, embraced them as an observational tool. Variations in cosmic-ray intensity were one of the ways my colleagues and I deduced the existence of the solar wind in the late 1950s.

Contrary to popular belief, it is not Earth's magnetic field that shields people on the ground from the full brunt of these rays but rather the bulk of our atmosphere. Above every square centimeter of surface is a kilogram of air. It takes a vertical column of about 70 grams—about 1/14 the distance through the atmosphere, achieved at an altitude of 20 to 25 kilometers (60,000 to 80,000 feet)—before the average incoming proton hits the nucleus of an atom in the air. The rest of the atmosphere serves to absorb the shrapnel of this initial collision. The impact knocks a proton or neutron or two out of the nucleus and unleashes a shower of high-energy gamma rays and pi meson, or pion, particles. Each gamma ray propagates deeper into the atmosphere and ends up producing an electron and its antimatter counterpart, a positron. These two particles annihilate each other, yielding less energetic gamma rays, and so the cycle continues until the gammas become too weak to create particles.

Meanwhile the pions quickly decay into mu mesons, or muons, which penetrate to the ground. As they pass through our bodies, they produce ions and break chemical bonds but not enough to do us significant harm. The annual cosmic radiation dose of about 0.03 rem (depending on altitude) is equivalent to a couple of chest x-rays.

Outside the atmosphere, the cosmic-ray bombardment is intense. Approximately one proton or heavier nucleus would pass through your fingernail every second, for a total of perhaps 5,000 ions zipping through the body every second, each one leaving a trail of broken chemical bonds and triggering the same cascade that occurs in the atmosphere. The relatively few heavier nuclei among the cosmic rays do as much or more damage than the protons because their ability to break bonds is proportional to the square of their electric charge. An iron nucleus, for example, does 676 times more damage than a proton does. A week or a month of this radiation should not have serious consequences, but a couple of years on a jaunt to Mars is a different story. One estimate from NASA is that about one third of the DNA in an astronaut's body would be cut by cosmic rays every year.

In addition to causing cancer, cosmic rays could lead to cataracts and brain damage.

Shields Up

THE ONLY QUANTITATIVE information available on the biological consequences of energetic radiation comes from the unfortunate individuals who have been exposed to short but intense bursts of gamma rays and fast particles during nuclear explosions and laboratory accidents. They have suffered cell damage and enhanced cancer rates. A Mars traveler would get similar doses, albeit spread out over time. No one knows whether the two situations are really equivalent, but the comparison is worrisome. Natural biological repair mechanisms may or may not be able to keep up with the damage.

The implications were recently studied by Wallace Friedberg of the Federal Aviation Administration's Civil Aerospace Medical Institute in Oklahoma City and his colleagues. In a report published last August, they estimated that Mars astronauts would receive a dose of more than 80 rems a year. By comparison, the legal dose limit for nuclear power plant workers in the U.S. is five rems a year. One in 10 male astronauts would eventually die from cancer, and one in six women (because of their greater vulnerability to breast cancer). What is more, the heavy nuclei could cause cataracts and brain damage. (To be sure, these numbers are highly uncertain.)

The constant hailstorm of cosmic rays is not the only ra-

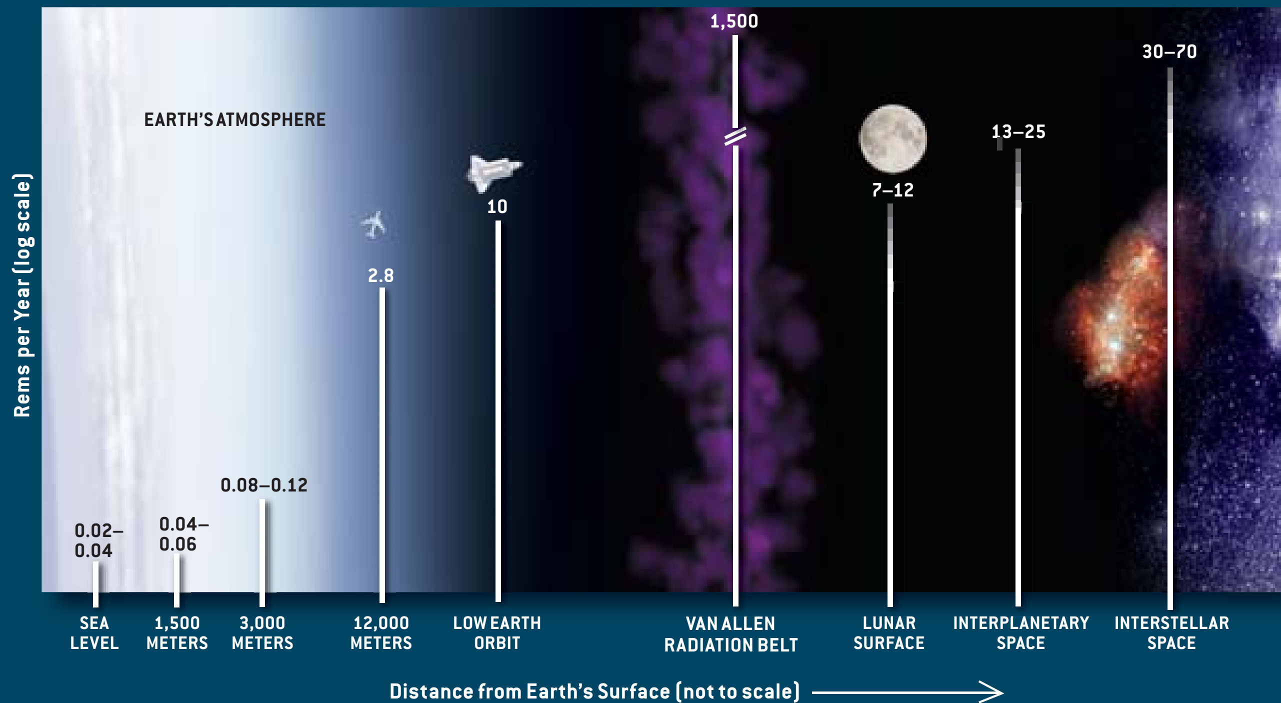
Overview/*Cosmic-Ray Hazard*

- The galaxy is pervaded with fast-moving particles that can rip apart DNA and other molecules. Here at the surface of Earth, we are well protected from this cosmic radiation by the air mass overhead. Astronauts in near-equatorial orbits are shielded by the planet's magnetic field. But those who make long voyages away from Earth will suffer serious health consequences.
- A spherical shell of water or plastic could protect space travelers, but it would take a total mass of at least 400 tons—beyond the capacity of heavy-lift rockets. A superconducting magnet would repel cosmic particles and weigh an estimated nine tons, but that is still too much, and the magnetic field itself would pose health risks. No other proposed scheme is even vaguely realistic.
- Biomedical researchers need to determine more precisely how much long-term exposure to cosmic rays a person can tolerate and whether medicines could stimulate the body's natural repair mechanisms.

PAT RAWLINGS/SAIC (*preceding pages*)

WHAT ARE ASTRONAUTS GETTING THEMSELVES INTO?

It is not quite as bad as venturing inside a nuclear reactor, but traveling through space can still be hazardous to your health. This graph shows one estimate of the yearly radiation dose astronauts would receive from cosmic rays. A rem is a common unit of radiation exposure. Interplanetary astronauts would absorb more radiation in a single year than radiation workers are supposed to receive in a lifetime, and a large number would develop cancer and other illnesses. Solar flares and Earth's Van Allen radiation belts can kill outright but are easier to avoid.



diation threat, of course. The sun, too, can unleash tremendous bursts of protons and heavier nuclei traveling at nearly the speed of light. Such bursts occasionally deliver in excess of a couple of hundred rem over an hour or so—a lethal dose to an unshielded astronaut. The great flare of February 23, 1956, is a notorious example. Whatever measures are taken to ward off cosmic rays should also protect against these solar tempests. Even so, it might be wise to schedule a trip to Mars during the years of minimum solar magnetic activity.

In recognition of the radiation threats, NASA set up the Space Radiation Shielding Program at the Marshall Space Flight Center in Huntsville, Ala., in 2003. The first thought was to protect astronauts by surrounding them with matter, by analogy to Earth's atmosphere. A second proposal was to deflect the cosmic rays magnetically, much as Earth's magnetic field offers some protection for equatorial regions and for the International Space Station. A more recent idea has been to give the spacecraft a positive charge, which would repel the positively charged nuclei.

NASA set up a two-day meeting in August 2004 at the University of Michigan at Ann Arbor to assess where things stood. The conclusion was not hopeful. It was not obvious what the solution to the cosmic-ray problem might be. Nor was it obvious that there is a solution at all.

Force Field

TO MATCH THE PROTECTION offered by Earth's atmosphere takes the same one kilogram of shielding material per square centimeter, although astronauts could comfortably make do with 500 grams, which is equivalent to the air mass above an altitude of 5,500 meters. Any less would begin to be counterproductive, because the shielding material would fail to absorb the shrapnel.

If the material is water, it has to be five meters deep. So a spherical water tank encasing a small capsule would have a mass of about 500 tons. Larger, more comfortable living quar-

THE AUTHOR

EUGENE N. PARKER is the world's leading expert on interplanetary gas and magnetic fields. He is best known for hypothesizing and explaining the solar wind in 1958. Now the stuff of textbooks, the idea of a vigorous outflow of particles from the sun was initially so controversial that the *Astrophysical Journal* nearly rejected his paper. Parker also developed the modern theory of the solar magnetic field, including magnetic reconnection. An emeritus physics professor at the University of Chicago and a member of the National Academy of Sciences, Parker has received numerous prizes, including the U.S. National Medal of Science, the Henry Norris Russell Lectureship of the American Astronomical Society and the Kyoto Prize in Basic Sciences.

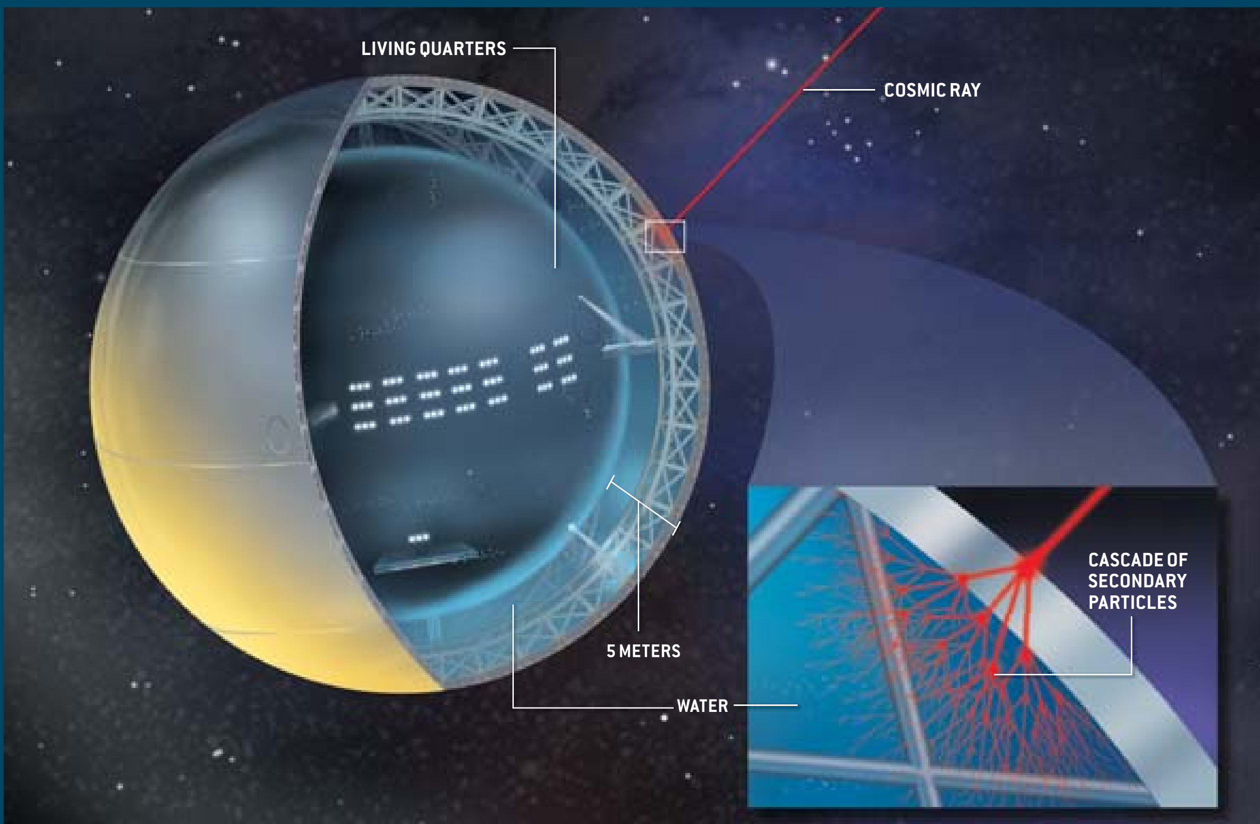
KENT SNODGRASS

PLAN 1: MATERIAL SHIELD

A large mass around the astronauts absorbs incoming radiation and the secondary particles it produces. A spherical shell of water five meters thick provides the same protection that Earth's atmosphere offers at an altitude of 5,500 meters (18,000 feet).

PROS:
- Simple
- Guaranteed to work

CONS:
- Way too heavy



ters would require even more. By comparison, the space shuttle can carry a maximum payload of about 30 tons. Water is commonly proposed because astronauts would need it anyway and because it is rich in hydrogen. Heavier elements make less effective shields because the extra protons and neutrons in their nuclei fall in one another's shadows, limiting their ability to interact with an incoming cosmic ray. To increase the hydrogen content, engineers could use ethylene (C_2H_4), which has the further advantage that it can be polymerized to polyethylene, a solid, thereby avoiding the necessity for a tank to contain it. Even so, the required mass would be at least 400 tons—still not feasible. Pure hydrogen would be lighter but would require a heavy pressurized vessel.

Consider, then, the prospects for magnetic shielding. A charged particle moving across a magnetic field is deflected at right angles to its direction of motion. Depending on the arrangement of field lines, the particle can be sent in almost any direction or even forced to circle endlessly. On approaching the magnetic field of Earth at low latitudes, a charged particle is sent back out into space [*see box on opposite page*] if it is not too energetic. A spacecraft could carry a magnet to do the same.

One big problem, though, is the immense kinetic energy of an individual cosmic-ray proton. Adequate protection for the astronauts means repulsing the very numerous cosmic-ray protons with two billion electron volts (the standard unit of energy used in particle physics). To stop them within the space of a few meters, a shield would have to have a magnetic field of 20 teslas, or about 600,000 times the strength of Earth's field at the equator. So strong a field requires an electromagnet constructed with superconducting wires, akin to those used in particle accelerators. Samuel C. C. Ting of the Massachusetts Institute of Technology headed up a design group that devised such a system with a mass of only nine tons—a big advance over material shielding but still discouragingly heavy to think of carrying all the way to the Martian surface and back.

The magnetic scheme has a number of fine points that should be appreciated. Magnetic fields provide no significant shielding near the magnetic poles, where incoming particles come in parallel to, rather than across, the field. That is why Earth's field provides little protection except for people living

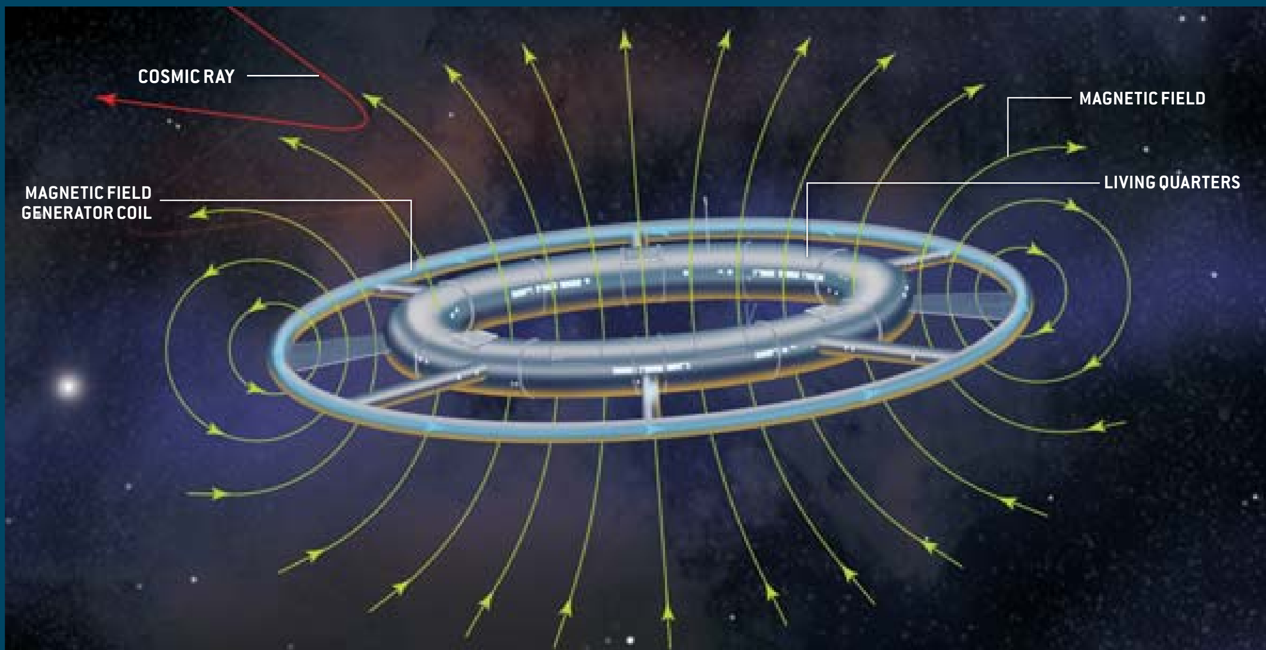
KENT SNODGRASS

PLAN 2: MAGNETIC SHIELD

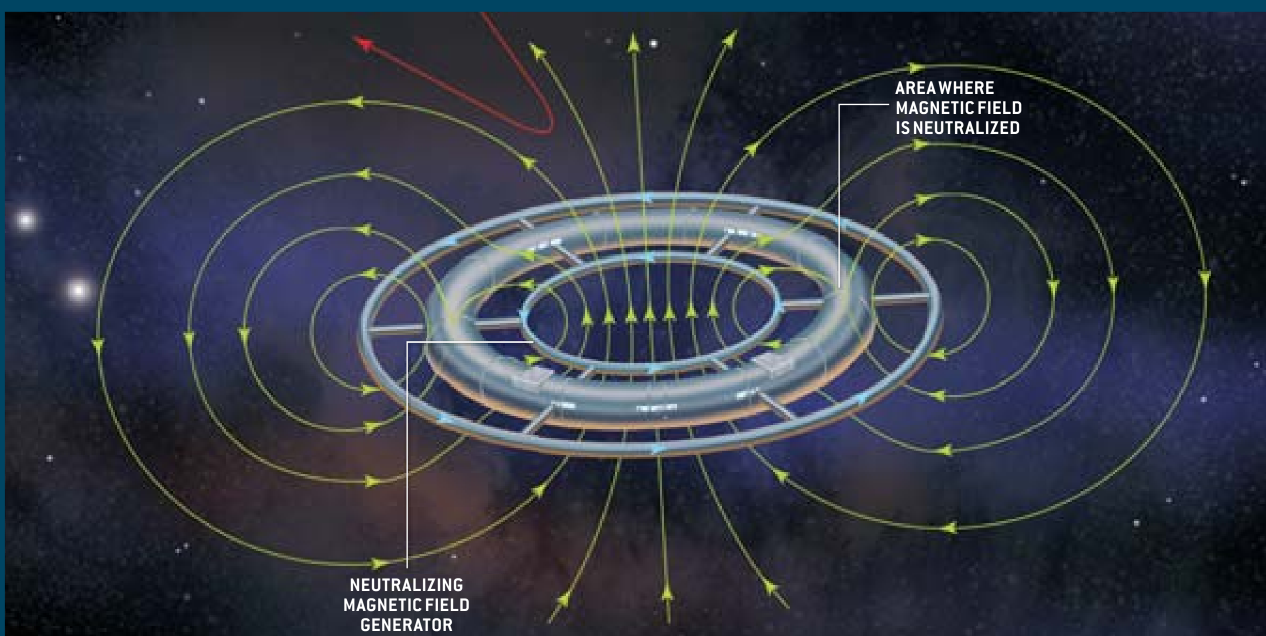
An electromagnet pushes incoming particles back into space. To deflect the bulk of cosmic rays, which have energies of up to two gigaelectron-volts, requires a magnetic field 600,000 times as strong as Earth's equatorial field.

PROS:
Much lighter than material shield

CONS:
Offers no protection along axis
Strong magnetic field may itself be dangerous



To suppress the field inside the living quarters, the spacecraft designers could add a second, inner electromagnet ring. But the cancellation is only partial and greatly increases the complexity of the system.



KENT SNODGRASS

in equatorial regions. To keep astronauts in the equivalent of an equatorial region, the living quarters of the spacecraft would have to be doughnut-shaped. The astronauts would have to endure a magnetic field of 20 teslas, and no one knows what the biological effects would be. The late John Marshall, a University of Chicago experimental physicist, remarked to me many years ago that when he stuck his head in a 0.5-tesla field in the gap of an old particle-accelerator magnet, any motion of his head produced tiny flashes of light in his eyes and an acid taste in his mouth, presumably caused by electrolysis in his saliva.

My colleague stuck his head into a 0.5-tesla magnetic field. Any motion of his head produced tiny flashes of **light in his eyes and an acid taste in his mouth**.

Given that a strong field can affect body chemistry in this way, researchers need to conduct some laboratory experiments to verify the safety of a 20-tesla shield. If it proves hazardous, engineers may have to cancel out the field within the living quarters using an opposing electromagnet. A secondary magnet clearly makes the system more complicated and more massive.

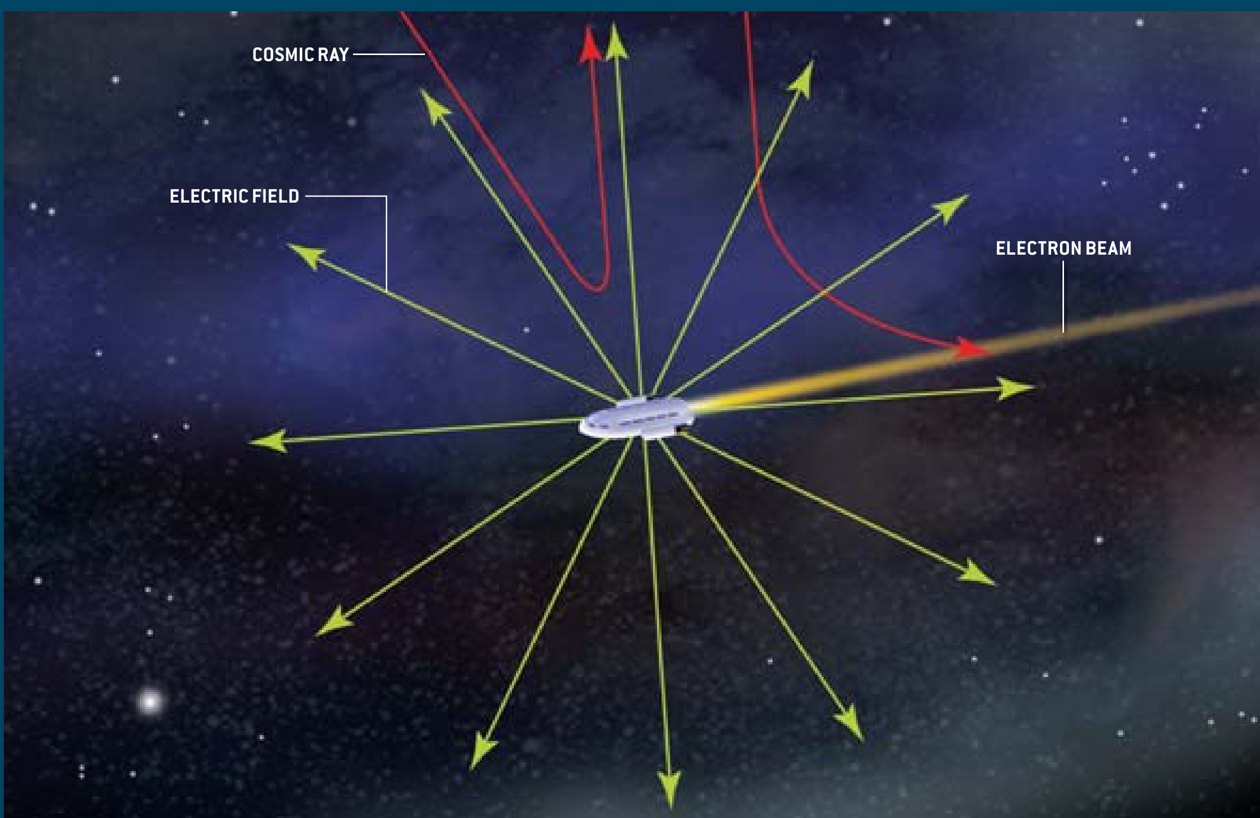
Some researchers have proposed using a field that extends over a distance much larger than a few meters. The field could be pushed out using a plasma, much as the ionized gas of the solar wind carries the solar magnetic field out to great distances from the sun. Advocates claim that such an "inflated" field would not need to be as intense; 1 tesla, or even less, might suffice. Unfortunately, this scheme disregards the fact that plasmas are notoriously unstable. The laboratory effort over the past 50 years to trap plasma in a magnetic field, for the purpose of producing energy from nu-

PLAN 3: ELECTROSTATIC SHIELD

Firing a beam of electrons into space causes a positive charge to build up on the spacecraft. This charge repels cosmic rays. To deflect particles with energies of up to two gigaelectron-volts, the ship would have to be charged to two billion volts.

PROS:
No gaps in coverage
No hazardous magnetic field

CONS:
Creates nasty influx of negatively charged particles
Requires gargantuan electric current



KENT SNODGRASS

clear fusion, has shown the remarkable ability of a plasma to wiggle free of any attempt to control it. Even if the plasma could be harnessed to inflate a magnetic field, it would serve only to weaken, rather than enhance, the shield. The field lines would be pushed out radially and spread around a larger circumference, so that an incoming proton would have to cross fewer field lines. The shield strength would fall, just as it does in the midlatitudes and polar regions of Earth.

No Charge

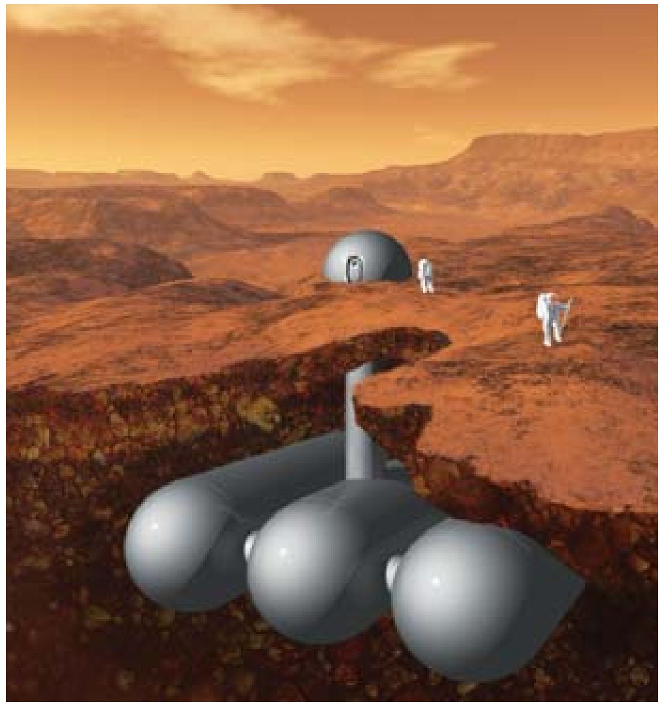
STRIKING OUT in another direction, other researchers have proposed to charge the spacecraft electrically. If the outer walls had a voltage of two billion volts relative to the surrounding space, they would repel all the cosmic-ray protons with energies up to two billion electron volts. A similar scheme has been proposed for a moon base.

The proposers seemed to be unaware that space is not empty. In the vicinity of Earth, the solar wind fills space with about five ions and five electrons per cubic centimeter. These electrons, being negatively charged, would be powerfully attracted by a positively charged spacecraft. Because the electric field would extend out to where its potential energy fell below the thermal energy of the electrons—a distance of tens of thousands of kilometers outward from the spacecraft—it would pull in electrons from an immense volume. They would hit the walls with an energy of two billion electron volts and behave just like cosmic rays, each having as much energy as the protons the system repels. Therefore, the natural cosmic-ray flux would be replaced with a vastly more intense artificial one. The electrons would produce gamma rays on impact with the spacecraft, and the intensity of that bombardment would be staggering, dwarfing the original problem.

That is not all. Simple estimates of the power requirements to maintain the charge of the spacecraft are mind-boggling. One ampere of current at two billion volts amounts to 2,000 megawatts—the output of a good-size electric power plant. Rough estimates suggest the current would exceed 10 million amperes. The proposers have not spelled out how they hope to charge the spacecraft to two billion volts in the first place. Curiously, like the idea of inflating magnetic fields, the notion of charging the spacecraft to shield the astronauts has received substantial attention and funding without a clear explanation of how it might work.

Others have proposed more prosaic options. Larger rockets or advanced propulsion technologies could speed the journey and lessen astronauts' exposure time. But the optimum travel time to Mars is more or less a fixed fraction of the orbital period of the planets, and to trim it by much would take a good deal more fuel (and hence money). On Mars itself, the problem does not go away. The atmosphere is scrawny, a mere 10 grams per square centimeter. Burying the base under hundreds of tons of soil would provide protection but require heavy machinery.

At the present time, then, the proposals for protecting astronauts from cosmic rays give little encouragement. But on the bright side, researchers are only beginning to explore the biomedical side of the problem. Natural healing processes in the cell may be able to handle radiation doses that accumulate over an extended period, and some people's bodies may be better at it than others'. If so, the present estimates of the cancer incidence, all based on short, intense bursts of radiation, may overestimate the danger.



RON MILLER

MARS'S PITIFUL ATMOSPHERE is scant protection against cosmic rays. Astronauts will have to bury their base under tons of dirt and limit their exposure outside. The prospect for permanent settlement hinges on whether biomedical researchers can develop antiradiation medicine.

In 2003 NASA set up the National Space Radiation Laboratory at Brookhaven National Laboratory to determine the molecular pathways of cell damage, with the hope of finding drugs to reduce or repair it. The lab is investigating precisely how radiation batters DNA and what types of injury do not readily heal. So far the only known chemicals that improve the resistance of laboratory rats to radiation damage are themselves toxic.

It would be too bad if the romance of human space travel ended ignominiously with cosmic rays making it infeasible. Capable people might be willing to go to the moon or Mars just for the adventure, come what may. Even so, the radiation hazard would take the luster off the idea of human space travel, let alone full-scale colonization. SA

MORE TO EXPLORE

Shielding Space Explorers from Cosmic Rays. Eugene Parker in *Space Weather*, Vol. 3, No. 8, Article no. S08004; August 18, 2005.

Presentations from the 2004 NASA workshop on radiation shields are available at **aoss.engin.umich.edu/Radiation**

NASA's own Web site on space radiation is at **www.radiationshielding.nasa.gov**

Unlocking the Secrets of Longevity Genes

A handful of genes that control the body's defenses during hard times can also dramatically improve health and prolong life in diverse organisms. Understanding how they work may reveal the keys to extending human life span while banishing diseases of old age

By David A. Sinclair and Lenny Guarente

TAPPING THE POWER of longevity genes could change the arc of a typical human lifetime: instead of vitality and growth giving way to the decline of old age, a person might be able to retain the youthfulness he feels at 50 when he is 70, 90 or well past 100.

You can assume quite a bit

about the state of a used car just from its mileage and model year. The wear and tear of heavy driving and the passage of time will have taken an inevitable toll. The same appears to be true of aging in people, but the analogy is flawed because of a crucial difference between inanimate machines and living creatures: deterioration is not inexorable in biological systems, which can respond to their environments and use their own energy to defend and repair themselves.

At one time, scientists believed aging to be not just deterioration but an active continuation of an organism's genetically programmed development. Once an individual achieved maturity, "aging genes" began to direct its progress toward the grave. This idea has been discredited, and conventional wisdom now holds that aging really is just wearing out over time because the body's normal maintenance and repair mechanisms simply wane. Evolutionary natural selection, the logic goes, has no reason to keep them working once an organism has passed its reproductive age.

Yet we and other researchers have found that a family of genes involved in an organism's ability to withstand a stressful environment, such as excessive heat or scarcity of food or water, have the power to keep its natural defense and repair activities going strong regardless of age. By optimizing the body's functioning for survival, these genes maximize the individual's chances of getting through the crisis. And if they remain activated long enough, they can also dramatically enhance the organism's health and extend its life span. In essence, they represent the opposite of aging genes—longevity genes.

We began investigating this idea nearly 15 years ago by imagining that evolution would have favored a universal regulatory system to coordinate this well-known response to environmental stress. If we could identify the gene or genes that serve as its master controllers and thereby act as master regulators of an organism's life span, these natural defense mechanisms might be turned into weapons against the diseases and decline that are now apparently synonymous with human aging.

Many recently discovered genes, known by such cryptic names as *daf-2, pit-1, amp-1, clk-1* and *p66Shc,* have been found to affect stress resistance and life span in laboratory organisms, suggesting that they could be part of a fundamental mechanism for surviving adversity [*see table on page 54*]. But our own two laboratories have focused on a gene called *SIR2,* variants of which are present in all organisms studied so far, from yeast to humans. Extra copies of the gene increase longevity in creatures as diverse as yeast, roundworms and fruit flies, and we are working to determine whether it does the same for larger animals, such as mice.

Animals that remain on this diet not only live longer but are far healthier during their prolonged lives.

As one of the first longevity genes to have been identified, *SIR2* is the best characterized, so we will focus here on its workings. They illustrate how a genetically regulated survival mechanism can extend life and improve health, and growing evidence suggests that *SIR2* may be *the* key regulator of that mechanism.

Silence Is Golden

WE FIRST DISCOVERED that *SIR2* is a longevity gene by asking what causes individual baker's yeast cells to grow old and whether a single gene might control aging in this simple organism. The notion that an understanding of yeast life span would tell us anything about human aging was deemed preposterous by many. Aging in yeast is measured by counting how many times mother cells divide to produce daughters before dying. A typical yeast cell's life span is about 20 divisions.

One of us (Guarente) began by screening yeast colonies for unusually long-lived cells in the hope of finding genes responsible for their longevity. This screen yielded a single mutation in a gene called *SIR4,* which encodes part of a complex of proteins containing the Sir2 enzyme. The mutation in *SIR4* caused the Sir2 protein to gather at the most highly repetitive region of the yeast genome, a stretch containing the genes that encode the protein factories of the cell, known as ribosomal DNA (rDNA). More than 100 of these rDNA repeats exist in the average yeast cell's genome, and they are difficult to maintain in a stable state. Repetitive sequences are prone to "recombining" with one another, a process that in humans can lead to numerous illnesses, such as cancer and Huntington's disease. Our yeast findings suggested that aging in mother cells was caused by some form of rDNA instability that was mitigated by the Sir proteins.

In fact, we found a surprising kind of rDNA instability. After dividing several times, yeast mother cells spin off extra copies of the rDNA as circular rings that pop out of the genome. These extrachromosomal rDNA circles (ERCs) are copied along with the mother cell's chromosomes prior to cell division but remain in the mother cell's nucleus afterward. Thus, a mother cell accumulates an ever increasing number of circles that eventually spell her doom, possibly because copying the ERCs consumes so many resources that she can no longer manage to replicate her own genome.

When an extra copy of the *SIR2* gene was added to the yeast cell, however, formation of the rDNA circles was repressed and the cell's life span was extended by 30 percent. That finding explained how *SIR2* could act as a longevity gene in yeast, but amazingly, we soon discovered that extra copies of the *SIR2* gene also extended the life span of roundworms by as much as 50 percent. We were surprised not only by this commonality in organisms separated by a vast evolutionary distance but by the fact that the adult worm body contains only nondividing cells—thus, the replicative aging mechanism in yeast could not apply to worms. We wanted to know exactly what the *SIR2* gene does.

As we soon discovered, the gene encodes an enzyme with a completely novel activity. Cellular DNA is wrapped around a complex of packaging proteins called histones. These bear chemical tags, such as acetyl groups, that determine how snugly the histones package DNA. Removing acetyl groups from histones tightens the packaging further and renders the DNA inaccessible to the enzymes responsible for popping the rDNA circles out of the chromosome. This deacetylated form of DNA is said to be silent because any genes in these regions of the genome are rendered inaccessible to being activated.

Sir proteins were already known to be involved in gene silencing—indeed, SIR stands for silent information regulator. Sir2 is one of several enzymes that remove acetyl tags from the histones, but we discovered that it is unique in that its enzymatic activity absolutely requires a ubiquitous small molecule called NAD, which has long been known as a conduit of many metabolic reactions in cells. This

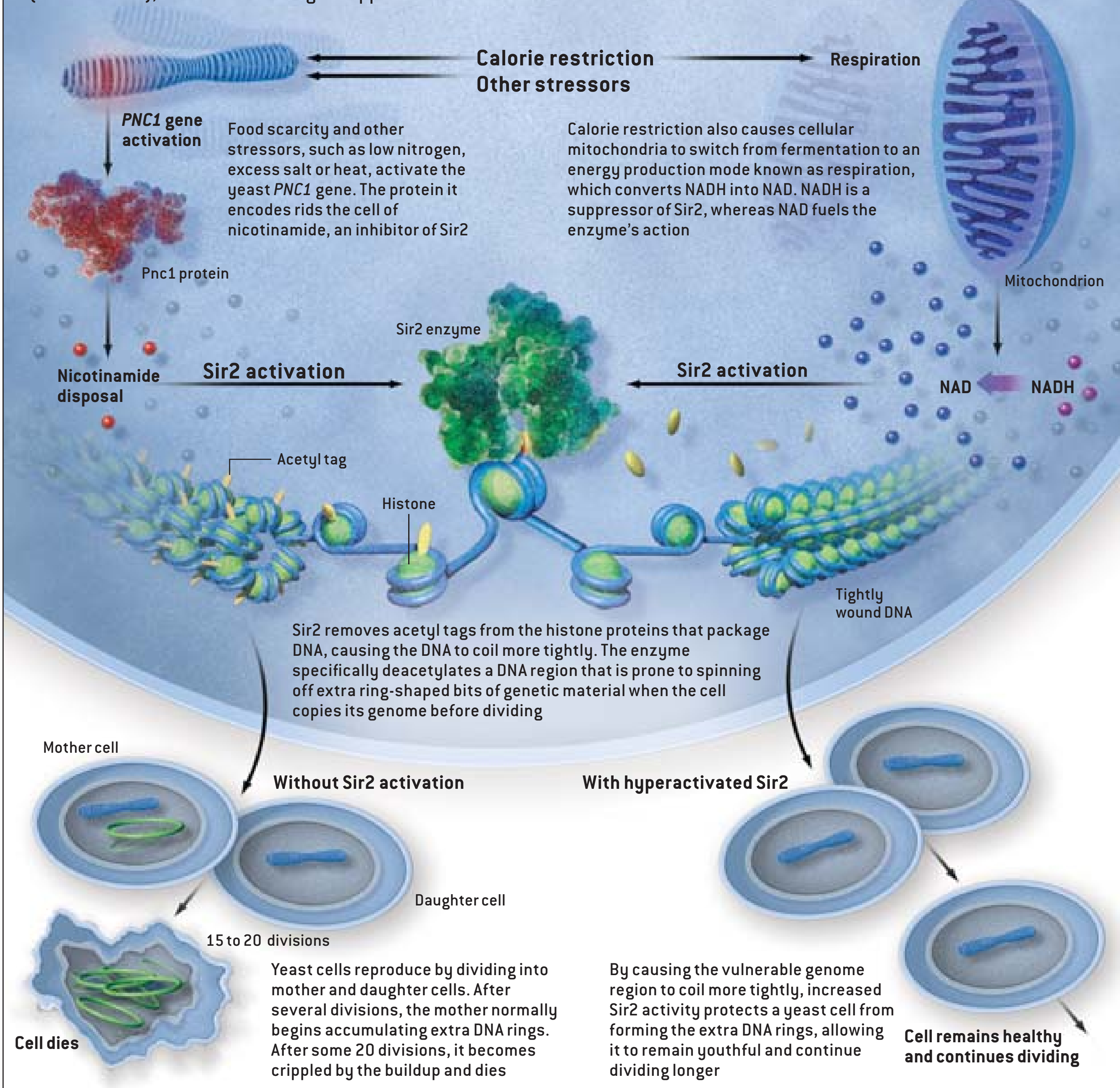
Overview/*Lifting Limits on Life Span*

- Genes that control an organism's ability to withstand adversity cause changes throughout the body that render it temporarily supercharged for survival.
- Activated over the long term, this stress response prolongs life span and forestalls disease in a wide range of organisms.
- Sirtuins are a family of genes that may be master regulators of this survival mechanism.
- Understanding how they produce their health and longevity-enhancing effects could lead to disease treatments and ultimately longer, disease-free human life spans.

PRECEDING PAGES: CARY WOLINSKY (*photographs*); JEN CHRISTIANSEN (*photocomposition*)

SIR2 AND STRESS IN YEAST

Moderate stress extends the life span of yeast about 30 percent by stimulating increased activity of the Sir2 enzyme. Stressors can boost Sir2 activity via two distinct pathways (*shown below*), each one resulting in suppression of a Sir2 inhibitor. The hyperactivated Sir2, in turn, represses a form of genome instability that normally would contribute to the yeast cell's death after about 20 cycles of cell division.



TAMI TOLPA

association between Sir2 and NAD was exciting because it linked Sir2 activity to metabolism and thus potentially to the relation between diet and aging observed in calorie restriction.

The Calorie Connection

RESTRICTING AN ANIMAL'S calorie intake is the most famous intervention known to extend life span. Discovered more than 70 years ago, it is still the only one absolutely proven to work. The restricted regime typically involves reducing an individual's food consumption by 30 to 40 percent compared with what is considered normal for its species. Animals ranging from rats and mice to dogs and possibly primates that remain on this diet not only live longer but are far healthier during their prolonged lives. Most diseases, including cancer, diabetes and even neurodegenerative illnesses, are forestalled. The organism seems to be supercharged for survival. The only apparent trade-off in some creatures is a loss of fertility.

Understanding the mechanisms by

which calorie restriction works and developing medicines that reproduce its health benefits have been tantalizing goals for decades [see "The Serious Search for an Antiaging Pill," by Mark A. Lane, Donald K. Ingram and George S. Roth; SCIENTIFIC AMERICAN: THE SCIENCE OF STAYING YOUNG, 2004]. The phenomenon was long attributed to a simple slowing down of metabolism—cells' production of energy from fuel molecules—and therefore reduction of its toxic by-products in response to less food.

But this view now appears to be incorrect. Calorie restriction does not slow metabolism in mammals, and in yeast and worms, metabolism is both sped up and altered by the diet. We believe, therefore, that calorie restriction is a biological stressor like natural food scarcity that induces a defensive response to boost the organism's chances of survival. In mammals, its effects include changes in cellular defenses, repair, energy production and activation of programmed cell death known as apoptosis. We were eager to know what part Sir2 might play in such changes, so we looked first at its role during calorie restriction in simple organisms.

In yeast, we have found that restricting food availability affects two pathways that increase Sir2 enzymatic activity in the cells. On one hand, calorie restriction turns on a gene called *PNC1*, which produces an enzyme that rids cells of nicotinamide, a small molecule similar to vitamin B_3 that normally represses Sir2. Consistent with the idea that calorie restriction is a stressor that activates a survival response, *PNC1* is also stimulated by other mild stressors known to extend yeast life span, such as increased temperature or excessive amounts of salt.

A second pathway induced in yeast by restricted calories is respiration, a mode of energy production that creates NAD as a by-product while lowering levels of its counterpart, NADH. It turns out that not only does NAD activate Sir2, but

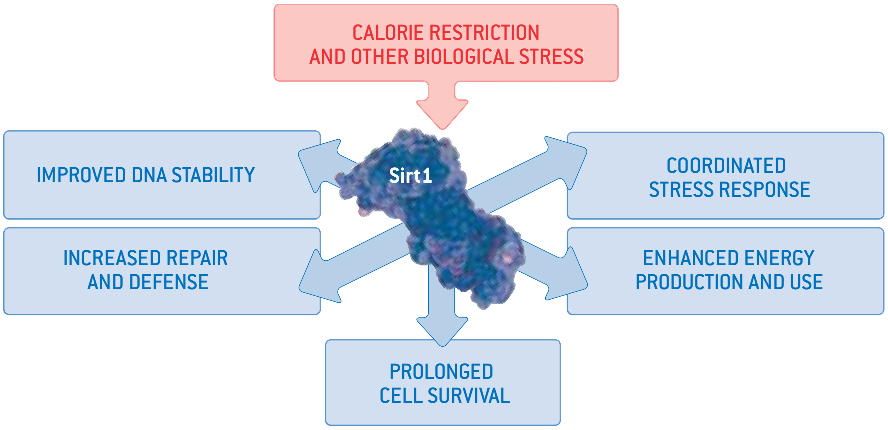
GENETIC PATHWAYS THAT EXTEND LIFE SPAN

Scientists studying longevity have identified an assortment of genes that can influence life span in different organisms. Like *SIR2* and its gene relatives (the Sirtuins), some of them promote longer life when more copies of the gene are present or activity by the protein it encodes is increased. Many of the genes and their proteins have a negative effect on life span, however, so reducing their activity enhances longevity.

In worms, for example, the gene that encodes cellular receptors for insulin and insulinlike growth factor 1 (IGF-1) is called *daf-2*. Suppressing the *daf-2* gene's activity in adult worms interferes with signaling via insulin and IGF-1 and extends the organisms' lives by as much as 100 percent. Suppression of several other growth-related genes or intervening in the pathways of molecular activity they trigger has also been found to promote longevity.

Several of the genes listed below or their proteins have been shown to regulate or be regulated by Sirtuins during calorie restriction, suggesting that they could be part of a master regulatory network for aging. The authors speculate that *SIR2* and its relatives may orchestrate this network.

GENE OR PATHWAY (HUMAN EQUIVALENT)	ORGANISM/LIFE SPAN EXTENSION	MORE OR LESS IS BETTER	MAJOR PROCESSES INFLUENCED	POSSIBLE SIDE EFFECTS OF MANIPULATION
SIR2 (*SIRT1*)	Yeast, worm, fly/ 30 percent	More	Cell survival, metabolism and stress responses	None known
TOR (*TOR*)	Yeast, fly, worm/ 30 to 250 percent	Less	Cell growth and nutrient sensing	Increased infections, cancer
Daf/FoxO proteins (Insulin, IGF-1)	Worm, fly, mouse/ 100 percent	Less	Growth and glucose metabolism	Dwarfism, sterility, cognitive decline, tissue degeneration
Clock genes (*CoQ* genes)	Worm/30 percent	Less	Co-enzyme Q synthesis	None known
Amp-1 (*AMPK*)	Worm/10 percent	More	Metabolism and stress responses	None known
Growth hormone (Growth hormone)	Mouse, rat/ 7 to 150 percent	Less	Body size regulation	Dwarfism
P66Shc (*P66Shc*)	Mouse/27 percent	Less	Free-radical production	None known
Catalase (*CAT*)	Mouse/15 percent	More	Detoxification of hydrogen peroxide	None known
Prop1, *pit1* (*Pou1F1*)	Mouse/42 percent	Less	Pituitary activity	Dwarfism, sterility, hypothyroidism
Klotho (*Klotho*)	Mouse/ 18 to 31 percent	More	Insulin, IGF-1 and vitamin D regulation	Insulin resistance
Methuselah (*CD97*)	Fly/35 percent	Less	Stress resistance and nerve cell communication	None known



ORCHESTRATOR OF BENEFICIAL CHANGE, the Sirt1 enzyme appears to be responsible for the health and longevity-enhancing effects of calorie restriction in mammals. Food scarcity and other biological stressors trigger increased activity by Sirt1, which in turn alters activities within cells. By boosting manufacture of certain signaling molecules, such as insulin, Sirt1 may also coordinate the stress response throughout the body. The enzyme produces its effects by modifying other proteins [*see illustration on next page*].

NADH is an inhibitor of the enzyme, so altering the cell's NAD/NADH ratio profoundly influences Sir2 activity.

Having seen how life-extending biological stress increases Sir2 activity, the question became, Is Sir2 necessary to produce the longevity? The answer appears to be a resounding "yes." One way to test whether Sir2 is essential to this process is to remove its gene and determine whether the effect remains. In organisms as complex as fruit flies, calorie restriction does require *SIR2* to extend life span. And because the body of an adult fruit fly contains numerous tissues that are analogous to mammalian organs, we suspect that calorie restriction in mammals is also likely to require *SIR2*.

Yet if humans are ever to reap the health benefits of calorie restriction, radical dieting is not a reasonable option. Drugs that can modulate the activity of Sir2 and its siblings (collectively referred to as Sirtuins) in a similar manner will be needed. Just such a Sirtuin-activating compound, or STAC, called resveratrol has proven particularly interesting. Resveratrol is a small molecule present in red wine and manufactured by a variety of plants when they are stressed. At least 18 other compounds produced by plants in response to stress have also been found to modulate Sirtuins, suggesting that the plants may use such molecules to control their own Sir2 enzymes.

Feeding resveratrol to yeast, worms or flies or placing them on a calorie-restricted diet extends their life spans about 30 percent, but only if they possess the *SIR2* gene. Moreover, a fly that overproduces Sir2 has an increased life span that cannot be further extended by resveratrol or calorie restriction. The simplest interpretation is that calorie restriction and resveratrol each prolong the lives of fruit flies by activating Sir2.

Resveratrol-fed flies not only live longer, despite eating as much as they want, but they do not suffer from the reduced fertility often caused by calorie restriction. This is welcome news for those of us hoping to treat human diseases with molecules that target Sir2 enzymes. But first we want a better understanding of the role of Sir2 in mammals.

Leader of the Band

THE MAMMALIAN VERSION of the yeast *SIR2* gene is known as *SIRT1* ("*SIR2* homolog 1"). It encodes a protein, Sirt1, that has the same enzymatic activity as Sir2 but that also deacetylates a wider variety of proteins both inside the cell nucleus and out in the cellular cytoplasm. Several of these proteins targeted by Sirt1 have been identified and are known to control critical processes, including apoptosis, cell defenses and metabolism. The potential longevity-enhancing role of the *SIR2* gene family seems, therefore, to be preserved in mammals. But not surprisingly in larger and more complex organisms, the pathways by which Sirtuins achieve their effect have grown considerably more complicated as well.

Increased Sirt1 in mice and rats, for example, allows some of the animals' cells to survive in the face of stress that would normally trigger their programmed suicide. Sirt1 does this by regulating the activity of several other key cellular proteins, such as p53, FoxO and Ku70, that are involved either in setting a threshold for apoptosis or in prompting cell repair. Sirt1 thus enhances cellular repair mechanisms while buying time for them to work.

Over the course of a lifetime, cell loss from apoptosis may be an important factor in aging, particularly in nonrenewable tissues such as the heart and brain, and slowing cell death may be one way Sirtuins promote health and longevity. A striking example of Sirt1's ability to foster survival in mammalian cells can be seen in the Wallerian mutant strain of mouse. In these mice, a single gene is duplicated, and the mutation renders their neurons highly resistant to stress, which protects them against stroke, chemotherapy-induced toxicity and neurodegenerative diseases.

In 2004 Jeffrey D. Milbrandt of Washington University in St. Louis and his colleagues showed that the Wallerian gene mutation in these mice increases the activity of an enzyme that makes NAD, and the additional NAD appears to protect the neurons by activating Sirt1. Moreover, Milbrandt's group found that STACs such as resveratrol conferred a protective effect on the neurons of normal mice similar to the Wallerian mutation.

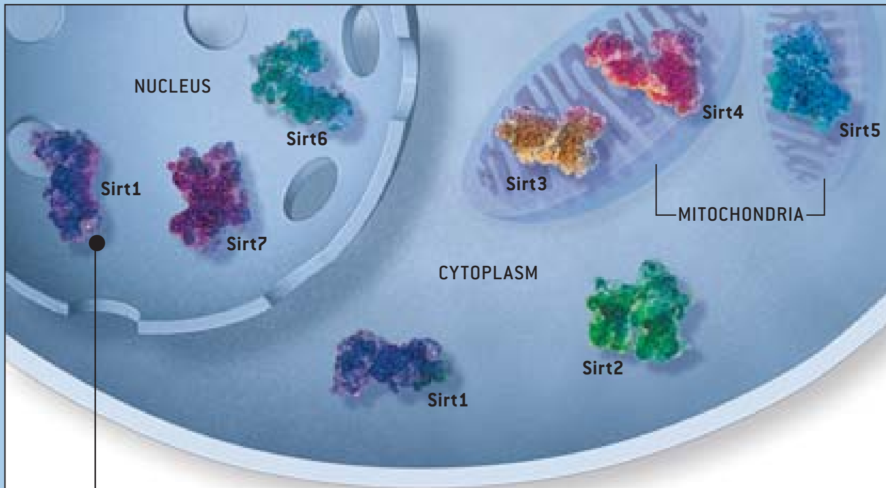
In a more recent study by Christian Néri of the French National Institute of Health and Medical Research, resveratrol and another STAC, fisetin, were shown to prevent nerve cells from dying in two different animal models (worm and mouse) of human Huntington's disease. In both cases, the protection by STACs required Sirtuin gene activity.

TAMI TOLPA (*illustration*); LUCY READING-IKKANDA (*diagram*)

SIRTUINS IN THE CELL

The Sirt1 enzyme is the best characterized Sirtuin, but it is not the only one found in mammals. Genes related to *SIRT1* give rise to similar enzymes that act in various locations within cells. Sirt1 operates both in the nucleus and in the cytoplasm, deacetylating other proteins and thereby altering their behavior. Many of its targets are transcription factors that directly activate genes or are regulators of those factors [*see examples below right*], giving Sirt1 regulatory control over a wide range of critical cellular functions.

Scientists are just beginning to identify the roles of the other Sirtuins and to determine whether they, too, influence longevity. Sirt2, for example, is known to modify tubulin, a component of a cell's internal scaffolding, and may affect cell division. Sirt3 is active in the cell's energy generators, the mitochondria, and appears to participate in regulating body temperature. The functions of Sirt4 and Sirt5 are unknown. Mutations in the gene encoding Sirt6 have been associated with premature aging.



SOME PROTEIN TARGETS OF SIRT1

FoxO1, FoxO3 and FoxO4: Transcription factors for genes involved in cell defenses and glucose metabolism

Histones H3, H4 and H1: Control DNA packing in chromosomes

Ku70: Transcription factor that promotes DNA repair and cell survival

MyoD: Transcription factor that promotes muscle development and tissue repair

NCoR: Regulator that affects multiple genes, including those involved in fat metabolism, inflammation and the functioning of other regulators, such as PGC-1α

NF-κB: Transcription factor that controls inflammation, cell survival and cell growth

P300: Regulator that causes acetyl tags to be added to histones

P53: Transcription factor that triggers programmed cell death in damaged cells

PGC-1α: Regulator that controls cellular respiration and appears to play a central role in muscle development

The protective effect of Sirtuins in individual cells is becoming increasingly clear. But if these genes are the mediators of calorie restriction's benefits, an unsolved puzzle remains how diet can regulate their activities and thus the rate of aging in an entire animal. Recent research by Pere Puigserver of the Johns Hopkins University School of Medicine and his colleagues has shown that NAD levels rise in liver cells under fasting conditions, prompting increased Sirt1 activity. Among the proteins Sirt1 acts on is an important regulator of gene transcription called PGC-1α, which then causes changes in the cell's glucose metabolism. Thus, Sirt1 was found to act both as a sensor of nutrient availability and a regulator of the liver's response.

Similar data have given rise to the idea that Sirt1 is a central metabolic regulator in liver, muscle and fat cells because it senses dietary variations via changes in the NAD/NADH ratio within cells and then exerts far-reaching effects on the pattern of gene transcription in those tissues. This model would explain how Sirt1 may integrate many of the genes and pathways that affect longevity described on page 54.

More than one mechanism may mediate Sirt1's bodywide activities, however. Another appealing hypothesis is that mammals register their food availability by the amount of energy they have stored in the form of body fat. Fat cells secrete hormones that convey signals to the other tissues in the body, but their message depends on the levels of fat stored. By reducing fat stores, calorie restriction may establish a pattern of hormone signals that communicates "scarcity," which activates cell defenses. Consistent with this idea is the fact that mice genetically engineered to be extra lean regardless of their food intake tend to live longer.

This possibility led us to wonder

THE AUTHORS

DAVID A. SINCLAIR and *LENNY GUARENTE* began working together to identify longevity-regulating genes and unravel their mechanisms in 1995, when Sinclair became a postdoctoral fellow in Guarente's lab at the Massachusetts Institute of Technology. Sinclair is now director of the Paul F. Glenn Laboratories for the Biological Mechanisms of Aging at Harvard Medical School and an associate of the Broad Institute in Cambridge, Mass. Guarente, Novartis Professor of Biology, has been on the faculty of M.I.T. for 25 years. His lab first identified the *SIR2* gene as the controller of life span in yeast and showed that the enzyme it encodes is responsible for the beneficial effects of calorie restriction in that organism. Today both authors are investigating the mammalian version of the gene, *SIRT1*. Sinclair's company Sirtis and Guarente's firm Elixir are each developing Sirtuin-activating molecules for pharmaceutical use.

TAMI TOLPA

whether Sirt1, in turn, also regulates fat storage in response to diet. Indeed, Sirt1 activity is increased in fat cells after food limitation, causing fat stores to move from the cells into the bloodstream for conversion to energy in other tissues. We surmise that Sirt1 senses the diet, then dictates the level of fat storage and thus the pattern of hormones produced by fat cells. This effect on fat and the signals it sends would, in turn, set the pace of aging in the entire organism and make Sirt1 a key regulator of the longevity conferred by calorie restriction in mammals. It would also closely link aging and metabolic diseases, including type 2 diabetes, associated with excess fat. Intervening pharmacologically in the Sirt1 pathway in fat cells might therefore forestall not only aging but also specific ailments.

Another critical process modified by Sirt1 is inflammation, which is involved in a number of disorders, including cancer, arthritis, asthma, heart disease and neurodegeneration. Recent work by Martin W. Mayo and his colleagues at the University of Virginia has shown that Sirt1 inhibits NF-κB, a protein complex that promotes the inflammatory response. The Sirt1-activating compound resveratrol has the same effect. This finding is particularly encouraging, both because the search for molecules that inhibit NF-κB is a highly active area of drug development and because another well-known effect of calorie restriction is its ability to suppress excessive inflammation.

If *SIR2* is thus the master controller of a regulatory system for aging that is activated by stress, it may function by acting as the conductor of an orchestra of players that includes hormonal networks, intracellular regulatory proteins and other genes associated with longevity. One of the more notable discoveries in recent years was that Sirt1 regulates production of insulin and insulinlike growth factor 1 (IGF-1) and that those two powerful signaling molecules, in turn, seem to regulate Sirt1 production as part of a complex feedback loop. The relation between Sirt1, IGF-1 and insulin is intriguing because it explains how Sirt1 activity in one tissue might be communicated to other cells in the body. Moreover, circulating levels of insulin and IGF-1 are known to dictate life span in various organisms—worms, flies, mice, possibly ourselves.

From Defense to Advance

BECAUSE PEOPLE have sought to slow aging for tens of thousands of years without success, some may find it hard to accept that human aging might be controlled by tweaking a handful of genes. Yet we know it is possible to forestall aging in mammals with a simple dietary change: calorie restriction works. And we have shown that Sirtuin genes control many of the same molecular pathways as calorie restriction. Without actually knowing the precise, and potentially myriad, causes of aging, we have already demonstrated in a variety of lifeforms that it can be delayed by manipulating a few regulators and letting them take care of the organisms' health.

We also know that the *SIR2* family of genes evolved far back in time because today they are found in organisms ranging from baker's yeast, *Leishmania* parasites and roundworms to flies and humans. In all these organisms but the last, which has not yet been tested, Sirtuins dictate length of life. This fact alone convinces us that human Sirtuin genes probably hold the key to our health and longevity as well.

Both our labs are running carefully controlled mouse experiments that should soon tell us whether the *SIRT1* gene controls health and life span in a mammal. We will not know definitively how Sirtuin genes affect human longevity for decades. Those who are hoping to pop a pill and live to 130 may have therefore been born a bit too early. Nevertheless, those of us already alive could live to see medications that modulate the activity of Sirtuin enzymes employed to treat specific conditions such as Alzheimer's, cancer, diabetes and heart disease. In fact, several such drugs have begun clinical trials for treatment of diabetes, herpes and neurodegenerative diseases.

And in the longer term, we expect that unlocking the secrets of longevity genes will allow society to go beyond treating illnesses associated with aging and prevent them from arising in the first place. It may seem hard to imagine what life will be like when people are able to feel youthful and live relatively free of today's diseases well into their 90s. Some may wonder whether tinkering with human life span is even a good idea. But at the beginning of the 20th century, life expectancy at birth was around 45 years. It has risen to about 75 thanks to the advent of antibiotics and public health measures that allow people to survive or avoid infectious diseases. Society adapted to that dramatic change in average longevity, and few people would want to return to life without those advances. No doubt, future generations accustomed to living past 100 will also look back at our current approaches to improving health as primitive relics of a bygone era. SA

Without actually knowing the precise causes of aging, we have demonstrated that it can be delayed.

MORE TO EXPLORE

Ageless Quest: One Scientist's Search for Genes That Prolong Youth. Lenny Guarente. Cold Spring Harbor Laboratory Press, 2002.

The Secrets of Aging. Sophie L. Rovner in *Chemical & Engineering News*, Vol. 82, No. 34, pages 30–35; August 23, 2004. **http://pubs.acs.org/cen/coverstory/8234/8234aging.html**

Calorie Restriction, Sirt1 and Metabolism: Understanding Longevity. Laura Bordone and Leonard Guarente in *Nature Reviews Molecular and Cell Biology*, Vol. 6, pages 298–305; April 2005.

Toward a Unified Theory of Caloric Restriction and Longevity Regulation. David A. Sinclair in *Mechanisms of Ageing and Development*, Vol. 126, No. 9, pages 987–1002; September 2005.

The Dangers of Ocean Ac

FRED BAVENDAM *Minden Pictures*

Much of the carbon dioxide given off from the burning of acid balance of seawater. The repercussions for marine

In 1956 Roger Revelle and Hans Suess, geochemists at the Scripps Institution of Oceanography in California, pointed out the need to measure carbon dioxide in the air and ocean so as to obtain "a clearer understanding of the probable climatic effects of the predicted great industrial production of carbon-dioxide over the next 50 years." In other words, they wanted to figure out how dire the situation would be today. That they had to argue the importance of such observations now seems astonishing, but at the time scientists did not know for certain whether the carbon dioxide spewing out of tailpipes and smokestacks would indeed accumulate in the atmosphere. Some believed that it would all be absorbed benignly by the sea or be happily taken up by growing plants on land.

Revelle and the young researcher he hired for this project, the late Charles David Keeling, realized that they had to set up equipment at remote locations, far from local sources and sinks of carbon dioxide, which would cause the measurements to vary erratically. One spot they chose was about as far from industrial activity and vegetation as anyone could get: the South Pole. Another was at a newly established weather station atop Mauna Loa in Hawaii.

The Mauna Loa monitoring has continued (with just one brief interruption) from 1958 to this day. Being not so remote as Antarctica, Hawaii sees carbon dioxide levels rise and fall sharply in step with the Northern Hemisphere's growing season, but at the end of each and every year, the concentration of this heat-trapping gas always ends up higher than it was 12

idification

fossil fuels goes into the ocean, where it changes the
life may be enormous **BY SCOTT C. DONEY**

months before. So it did not take long for the scientific community to realize that Revelle was right—much of the carbon dioxide released into the atmosphere was destined to remain there. But his calculations were also correct in showing that a substantial fraction would end up in the sea. And it was clear to Revelle long ago that the part that went into the ocean would fundamentally alter the chemistry of seawater. Unlike some aspects of climate change, the reality of this effect—essentially the acidification of the ocean—is not much debated, although its full implications are just now being revealed.

CORAL REEFS—and the extraordinary biodiversity they support—are under siege from many forces, including exposure to toxic chemicals and direct physical destruction. A less known but perhaps greater threat is the change in ocean chemistry caused by the burning of fossil fuels. Today one third of the carbon dioxide given off in that process enters the ocean, reducing its naturally alkaline pH. This shift toward more acidic conditions diminishes the ability of corals (and many other marine organisms) to grow.

How Unnatural?

THE HALF-CENTURY RECORD that Keeling produced is extremely valuable, but it is too short to place the current situation in context. Scientists have, however, been able to obtain a longer-term perspective by measuring air bubbles trapped in ice cores. From this natural archive they have learned that the atmospheric concentration of carbon dioxide was approximately constant for several thousand years and then began to grow rapidly with the onset of industrialization in the 1800s. This gas is now about 30 percent more abundant than it was a few hundred years ago, and it is expected to double or triple its former level by the end of this century.

This burgeoning supply of carbon comes largely from the burning of fossil fuels—coal, oil and natural gas. (Cement production and the burning of tropical forests add some, too, but to simplify things, let me gloss over such secondary contributions for the sake of clarity.) Unlike the constituents of living organisms, fossil fuels contain little or none of the radioactive form of carbon: the carbon 14 isotope, which has eight neutrons in the nucleus rather than the usual six. Fossil fuels also display a unique ratio of the two stable isotopes of carbon (carbon 12 and 13). The combustion of these fuels thus leaves a distinctive isotopic signature in the atmosphere. So no one can question where the growing surplus of carbon dioxide comes from.

The ocean has absorbed fully half of all the fossil carbon released to the atmosphere since the beginning of the Industrial Revolution.

Absorption rates can vary, but today about 40 percent of the carbon dioxide derived from fossil fuel remains in the atmosphere; the rest is taken up by vegetation on land or by the ocean, currently in about equal proportions. The injection of fossil-fuel carbon into the sea is, as of yet, a relatively small addition to the ocean's huge natural reservoir of this element. Detecting and quantifying the uptake, therefore, requires especially precise measurements, ones good to at least one part in 1,000. And because the amounts vary substantially from place to place, the task also demands the resources and perseverance to map carbon concentrations throughout the world.

Overview/CO_2 in the Ocean

- About a third of the carbon dioxide (CO_2) released by the burning of fossil fuels currently ends up in the ocean.
- Absorbed CO_2 forms carbonic acid in seawater, lowering the prevailing pH level (which is slightly alkaline) and changing the balance of carbonate and bicarbonate ions.
- The shift toward acidity, and the changes in ocean chemistry that ensue, makes it more difficult for marine creatures to build hard parts out of calcium carbonate. The decline in pH thus threatens a variety of organisms, including corals, which provide one of the richest habitats on earth.
- Within a century, the surface of the Southern Ocean will become corrosive to the shells of tiny snails that form a key link in the marine food chain within this highly productive zone.

Oceanographers did exactly that in the late 1980s and 1990s, as part of a global assessment that went by two acronyms: JGOFS (for Joint Global Ocean Flux Study) and WOCE (for World Ocean Circulation Experiment).

Those surveys, however, did not in themselves identify what part of the carbon measured is natural and what part derives from the carbon dioxide that people have dumped into the air. In 1996 Nicolas Gruber, now at the University of California, Los Angeles, and two of his colleagues developed an innovative technique for doing so. The application of Gruber's method to all the JGOFS and WOCE data, an exercise completed in 2004, suggests that the ocean has absorbed fully half of all the fossil carbon released to the atmosphere since the beginning of the Industrial Revolution.

Another way to document this process is to make repeated measurements of carbon on the same piece of ocean. One must be careful to distinguish the fossil carbon from the various biological sources of this element in the sea. And the observations need to span a decade or more to reveal the overall trend brought on by the burning of fossil fuels against the background of natural variability. Last year Rik Wanninkhof of the National Oceanic and Atmospheric Administration's Atlantic Oceanographic and Meteorological Laboratory and I led a research expedition to do just such an experiment.

With a party of 31 scientists, technicians and students onboard our research vessel, we spent almost two months sampling the physical and chemical properties of the western South Atlantic, from top to bottom, starting near Antarctica and ending near the equator. This is the very same slice of ocean that I and other scientists had first measured in 1989, when I was a graduate student.

When we compared our observations from 2005 with those made 16 years earlier, we found that the upper few hundred meters of the South Atlantic in general have higher carbon concentrations today than in the recent past, which is consistent with the notion that the sea is taking in atmospheric carbon dioxide. Other oceanographers have found similar trends in the Pacific and Indian oceans as well. But what exactly does this change portend for the marine environment?

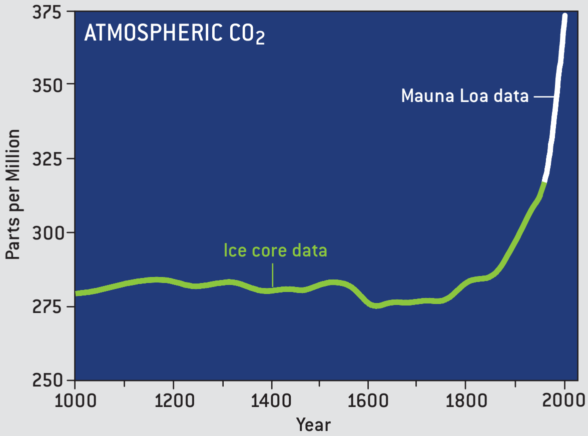
Ocean Chemistry 101

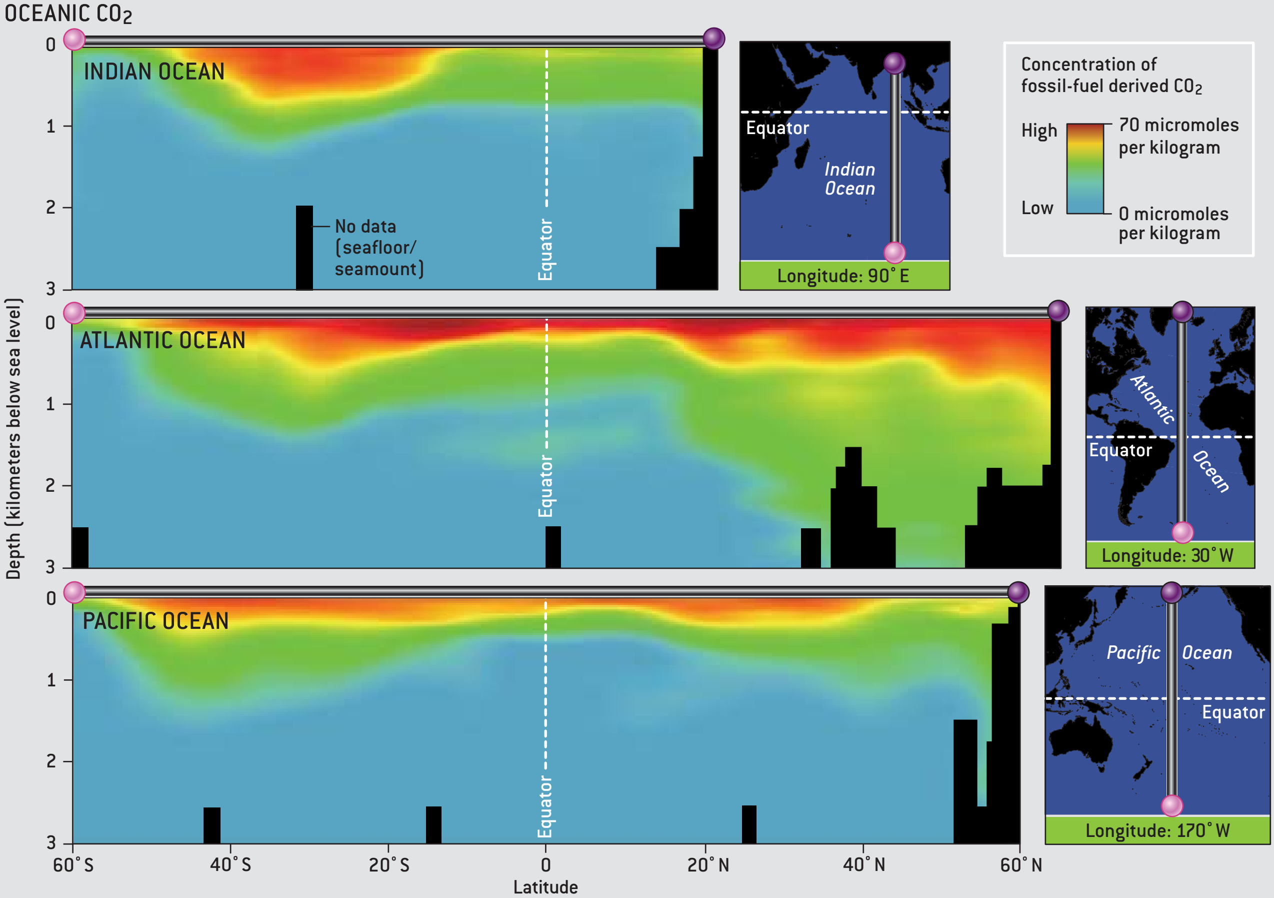
EXPLAINING THE IMPLICATIONS of these shifts in oceanic conditions requires, alas, a review of some freshman chemistry. But bear with me; it is really not that painful. Carbon dioxide (CO_2) combines with water to form carbonic acid (H_2CO_3), the same weak acid found in carbonated beverages. Like all acids, this one releases hydrogen ions (H^+) into solution, in this case leaving both bicarbonate ions (HCO_3^{-1}) and, to a lesser extent, carbonate ions (CO_3^{-2}) also swimming around. A small fraction of the carbonic acid remains in solution without dissociating, as does a little carbon dioxide. The resulting brew of carbon compounds and ions is thus rather complex.

CO_2: FROM ATMOSPHERE TO OCEAN

The concentration of carbon dioxide in the atmosphere has mounted considerably over the past century or so. This worrisome trend is well documented (*right*) by a combination of two techniques: the examination of air bubbles trapped in glacial ice (*green segment*, which shows 75-year averages) as well as direct measurements of the atmosphere (*white segment*, which reflects the annual average determined at a weather station situated atop Mauna Loa on the big island of Hawaii).

Large as it is, the increasing concentration of carbon dioxide in the atmosphere would have been even greater had not much of it been absorbed by the sea, a phenomenon that detailed oceanographic surveys have now documented. The cross sections below show where about half of this fossil-fuel effluent now resides—in the upper portions of the world's oceans.

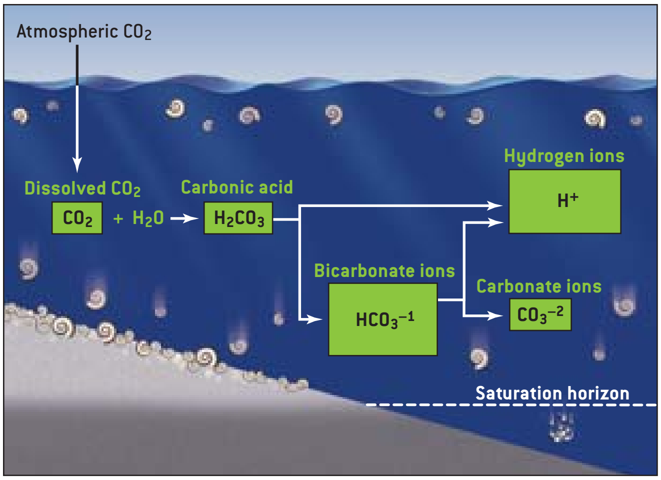




JEN CHRISTIANSEN; SOURCES: CARBON DIOXIDE INFORMATION ANALYSIS CENTER AT OAK RIDGE NATIONAL LABORATORY (http://cdiac.ornl.gov/ftp/trends/co2/lawdome.combined.dat and http://cdiac.ornl.gov/ftp/ndp001/maunaloa.co2) (*atmospheric CO_2 data*); ROBERT KEY *Princeton University* (*oceanic CO_2 cross section data*)

One simple result of all this dissolving and dissociating is the increase in hydrogen ion concentration, which chemists normally quantify in terms of the familiar pH scale. A drop by one pH unit corresponds to a 10-fold increase in the concentration of hydrogen ions, making the water more acidic, whereas a change of one unit upward corresponds to a 10-fold decrease, making the water more alkaline. Neutral pH (that of pure water) is 7. The pH of pristine seawater measures from 8 to 8.3, meaning that the ocean is naturally somewhat alkaline.

The absorption of carbon dioxide has already caused the pH of modern surface waters to be about 0.1 lower (less alkaline) than it was in preindustrial times. Unless civilization modifies its appetite for fossil fuels soon and in a significant way, ocean pH will fall an additional 0.3 by 2100. In a troubling prediction of the more distant future, Ken Caldeira, an oceanographer at the Carnegie Institution of Washington, suggests that ocean pH several centuries from now will be lower than at any time in the past 300 million years.



CARBON DIOXIDE absorbed from the air combines with water to form carbonic acid. A portion of this compound persists in the ocean, but most of it dissociates into acidifying hydrogen ions along with bicarbonate ions. Some of the latter also dissociate, forming carbonate ions and yet more hydrogen ions. These chemical changes cause an upward shift in the "saturation horizons" for calcite and aragonite—the water levels deep in the sea below which shells of marine organisms made of these minerals dissolve.

These shifts in pH may seem small, but they provide ample cause for alarm. Notably, recent experiments indicate that the change will prove harmful to some forms of marine life—in particular, to organisms that depend on the presence of carbonate ions to build their shells (or other hard parts) out of calcium carbonate ($CaCO_3$).

At first this concern appears paradoxical. After all, if some of the carbon dioxide absorbed by the sea dissociates into carbonate ions, one might expect that there should be plenty to go around, even more than would have been available otherwise. That logic, though, is flawed because it neglects the effect of all the hydrogen ions that are also created, which tend to combine with carbonate ions, forming bicarbonate ions. The net result is thus a reduction in the concentration of carbonate ions.

The worry is that a lowering of pH (and thus of carbonate ion concentration, which is expected to drop by half over this century) will hamper the ability of certain organisms to make calcium carbonate, so much so that these organisms will then have difficulty growing. Some of the most abundant lifeforms that could be affected in this way are a type of phytoplankton called coccolithophorids, which are covered with small plates of calcium carbonate and are commonly found floating near the surface of the ocean (where they use the abundant sunlight for photosynthesis). Other important examples are planktonic organisms called foraminifera (which are related to amoeba) and pteropods (small marine snails). These tiny creatures constitute a major food source for fish and marine mammals, including some species of whales.

Biologists also fear what might happen to corals, which despite their plantlike appearance are actually colonies of small animals related to sea anemones. They feed by filtering plankton out of the water, and they secrete calcium carbonate skeletons, which accumulate over time to form coral reefs—some of the most productive and biologically diverse ecosystems in the ocean. Coralline algae (algae that also secrete calcium carbonate and often resemble corals) contribute to the calcification of many reefs, too. The Great Barrier Reef off the coast of Australia, for instance—the largest biological structure in the world—is simply the accumulation of generation after generation of coral and coralline algae. Less obvious examples occur deeper down in the sea, where cold-water coral communities dot continental margins and seamounts, forming important fish habitats.

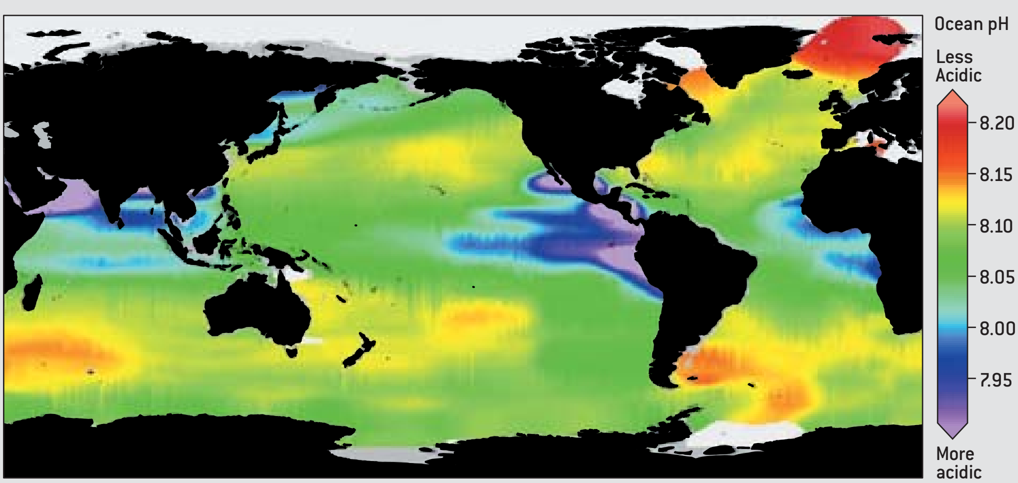
Shallow-water corals owe their beautiful colors in part to symbiotic algae, which live inside the coral cells. In response to various forms of environmental stress, these algae sometimes leave their hosts, exposing the white calcium carbonate skeleton underneath. Such "bleaching" events can be brought on by extreme warmth, for example. And some scientists sus-

JEN CHRISTIANSEN (top); CHRIS SABINE *NOAA Pacific Marine Environment Laboratory* (bottom)

THE OCEAN'S CHANGING ACIDITY

Measurements taken in the top 50 meters of the ocean reveal that pH varies considerably from place to place. Scientists expect oceanic pH to decrease in the years ahead.

Areas of relatively low pH (signifying more acidic conditions) arise mostly through the natural upwelling of deeper waters. Those zones, such as in the east equatorial Pacific, might be good places for scientists to study the effects expected to prevail over wider areas in the future.



pect that the acidification of the ocean (or more properly, the reduction in the ocean's slightly alkaline state) also tends to prompt such episodes.

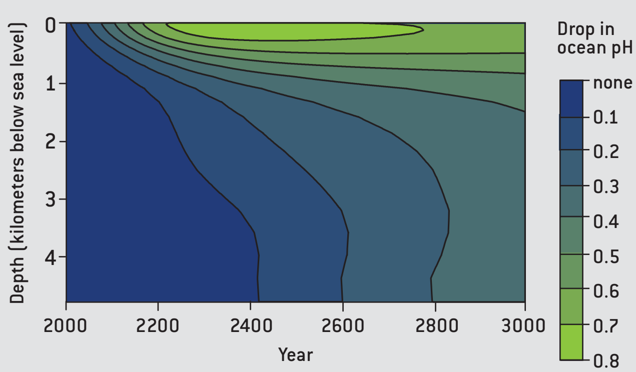
Survival of the Thickest?

BUT CORALS and other calcifying marine organisms could be affected by acidification in even more significant ways—their shells might actually disintegrate. Drop a piece of chalk (calcium carbonate) into a glass of vinegar (a mild acid) if you need a demonstration of the general worry: the chalk will begin dissolving immediately. Gaining a fuller understanding of which life-forms are most at risk of such a fate requires another short chemistry lesson.

The calcium carbonate in corals or in the shells of other marine creatures comes in two distinct mineral forms: calcite and aragonite. And some calcite-secreting organisms also add magnesium to the mix. Aragonite and magnesium calcite are more soluble than normal calcite. Thus, corals and pteropods, which both produce aragonitic shells, and coralline algae, which manufacture magnesium calcite, may be especially susceptible to harm from ocean acidification.

As the ocean becomes more and more acidic, the upper, shell-friendly portion will become thinner.

The solubility of calcium carbonate depends fundamentally on the carbonate ion concentration (and therefore indirectly on pH), but it also hinges on several other variables, including temperature and pressure. Under modern conditions, many deep, cold waters are acidic enough to dissolve calcium carbonate shells. They are said to be "undersaturated." Shallow, warm surface waters are described as "supersaturated" with respect to both calcite and aragonite, meaning that these minerals have no tendency to dissolve. The transition between supersaturated and undersaturated conditions is referred to as the saturation horizon: the level below which things begin to dissolve.



Although acidification of the sea (as measured by the drop in pH) has so far been small, scientists expect larger shifts in the future. The greatest changes will take place near the surface, but over time the full depth of the ocean will feel these effects.

JEN CHRISTIANSEN; SOURCE: "ANTHROPOGENIC CARBON AND OCEAN pH," BY K. CALDEIRA AND M. E. WICKETT IN *NATURE*, VOL. 425; SEPTEMBER 25, 2003

The influx of carbon dioxide from the atmosphere has caused the saturation horizons for aragonite and calcite to shift closer to the surface by 50 to 200 meters compared with where they were positioned in the 1800s. And recent studies indicate that further widespread shoaling will take place over the next several decades. Thus, as the ocean becomes more and more acidic, the upper, shell-friendly portion will become thinner. That is to say, less and less of the sea will remain hospitable for calcifying organisms.

Early on, many scientists reasoned that ocean acidification would pose only a minor problem, because surface waters would remain supersaturated—at least with respect to calcite, the robust form of calcium carbonate. In the late 1990s Christopher Langdon, a marine biologist at the University of Miami, conducted an elegant experiment to test this surmise: he manipulated the water chemistry over an artificial coral reef that had been set up in a huge tank at Columbia University's Biosphere II laboratory (which, rather incongruously, is located in the middle of the Arizona desert). Strikingly, he found that the rate of calcium carbonate production in the corals declined with lower pH, although the water remained considerably supersaturated with respect to aragonite. Shortly afterward, Ulf Riebesell of the Alfred Wegener Institute for Polar and Marine Research in Germany and his colleagues demonstrated a similar stunting of planktonic coccolithophorids. Laboratory experiments are now available to show the deleterious effects of increased carbon dioxide (and the lower pH that results) for all the major groups of marine organisms that have hard parts made of calcium carbonate.

Because cold waters are naturally less supersaturated than warm ones for the various forms of calcium carbonate, high-latitude and deep water ecosystems may be the first to suffer from ocean acidification. Polar surface waters most likely will

THE AUTHOR

SCOTT C. DONEY is a senior scientist in the department of marine chemistry and geochemistry at the Woods Hole Oceanographic Institution. He began pursuing oceanographic studies while an undergraduate at the University of California, San Diego, and ultimately obtained a doctorate in chemical oceanography in 1991, after completing a program jointly administered by the Massachusetts Institute of Technology and the Woods Hole Oceanographic Institution. Among his other professional activities, Doney serves on the NASA Orbital Carbon Observatory science team and is chair of the Ocean Carbon and Climate Change Scientific Steering Group, which is a part of the U.S. Global Change Research Program.

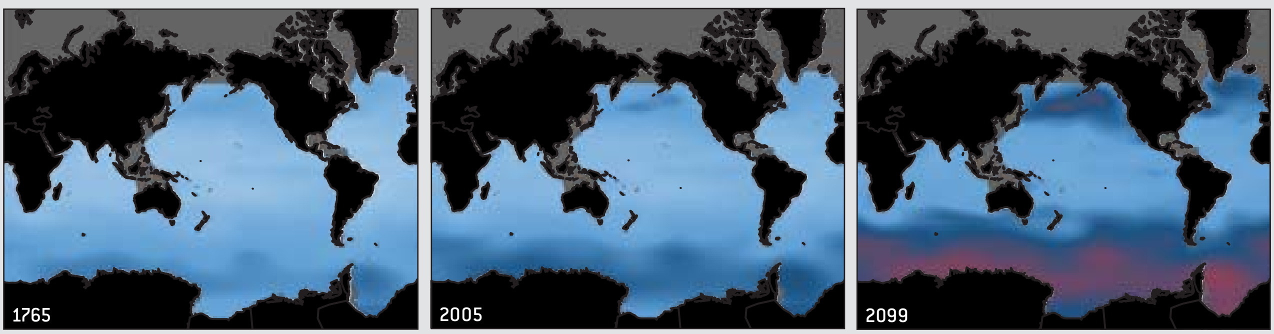
MARINE ORGANISMS at risk from the increasing acidification of the ocean include the corals and coralline algae commonly found in reef communities, as well as foraminifera and coccolithophorids, which are abundant in most surface waters. Also under threat, particularly in cold, polar waters, are a variety of small marine snails called pteropods.

become undersaturated for aragonite before the end of this century. One worrisome possibility, based on the work of Victoria J. Fabry of California State University, San Marcos, is that polar pteropods will simply disappear altogether. Or perhaps they will be forced to migrate to lower, warmer latitudes, assuming that they can adapt to those environments. No one knows how a sharp decrease in pteropod numbers will affect other parts of the marine ecosystem. But the fact that these small snails are a key link in the food chain in the Southern Ocean (which supports large populations of fish, whales and seabirds) is ample cause for concern.

High-latitude calcareous phytoplankton and zooplankton might share a similar fate, although their declines would come decades later because their shells are formed from calcite, the less soluble form of calcium carbonate. Deep coral communities will probably suffer, too, particularly those in the western North Atlantic along the path of water that contains high concentrations of carbon from fossil-fuel emissions.

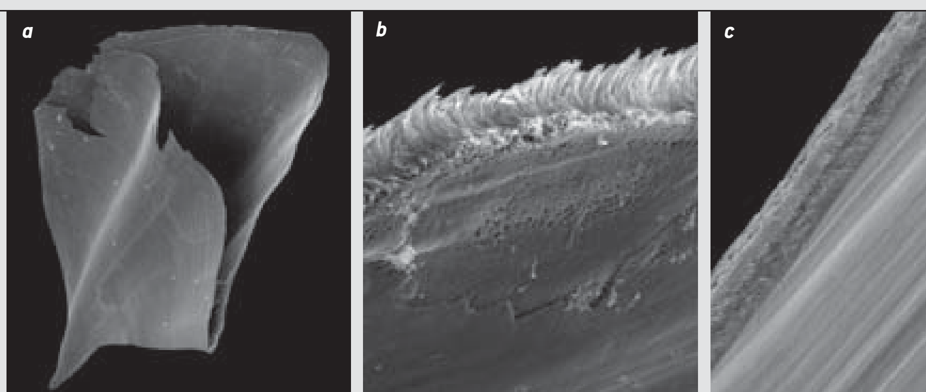
THE (RAGGED) FUTURE OF ARAGONITE

Diminishing pH levels will weaken the ability of certain marine organisms to build their hard parts and will be felt soonest and most severely by those creatures that make those parts of aragonite, the form of calcium carbonate that is most prone to dissolution. The degree of threat will vary regionally.



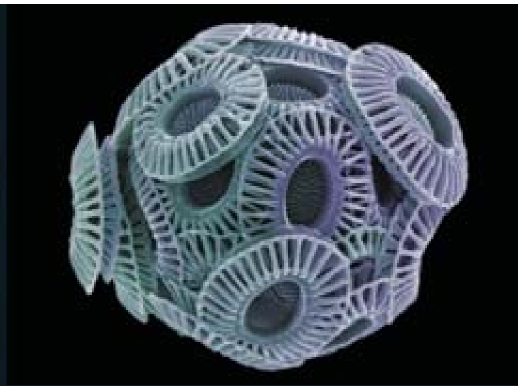
Before the Industrial Revolution (*left*), most surface waters were substantially "oversaturated" with respect to aragonite (*light blue*), allowing marine organisms to form this mineral readily. But now (*center*), polar surface waters are only marginally oversaturated (*dark blue*). At the end of this century (*right*), such chilly waters, particularly those surrounding Antarctica, are expected to become undersaturated (*purple*), making it difficult for organisms to make aragonite and causing aragonite already formed to dissolve.

Pteropods form a key link in the food chain throughout the Southern Ocean. For these animals (and creatures that depend on them), the coming changes may be disastrous, as the images at the right suggest. The shell of a pteropod kept for 48 hours in water undersaturated with respect to aragonite shows corrosion on the surface (*a*), seen most clearly at high magnification (*b*). The shell of a normal pteropod shows no dissolution (*c*).



ALEXIS ROSENFELD *Photo Researchers, Inc.* (*fire coral*); KEOKI STENDER *Fishpics Hawaii* (*coralline algae*); WOODS HOLE OCEANOGRAPHIC INSTITUTION (*foraminifer*); STEVE GSCHMEISSNER *Photo Researchers, Inc.* (*coccolithophorid*); RUSS HOPCROFT *University of Alaska–Fairbanks/NOAA* (*pteropod*); JAMES C. ORR *Laboratory of the Science of Climate and the Environment, UMR CEA-CNRS, France* (*map data and pteropod shell images*)

Foraminifer (*Globigerina bulloides*)



Coccolithophorid (*Emiliania huxleyi*)

Pteropod (*Limacina helicina*)

The outlook for coral reefs is even bleaker. For those precious ecosystems, ocean acidification is but one of many environmental stresses, an onslaught that includes greenhouse warming, local pollution, overfishing and habitat destruction. Many coral reefs are already in decline, and ocean acidification may push some over the edge into nonexistence.

Coming Sea Change

AS BAD AS CONDITIONS are expected to be for many marine organisms, there will be some winners, too. Right now very little of the carbon in seawater takes the form of dissolved carbon dioxide, and this scarcity limits the growth of some types of phytoplankton. Many of these species devote precious energy to concentrate carbon dioxide inside their cells, so one might guess that increases in dissolved carbon dioxide will be beneficial to them. Perhaps that will be the case. Not enough is known, however, about this "fertilization" effect to make firm predictions for the future of phytoplankton or to say whether higher carbon dioxide levels will benefit the photosynthetic algae that live inside corals. Many species of marine phytoplankton use HCO_3^{-1} for photosynthesis. And because the concentration of this ion will remain largely unchanged, biologists do not expect that these organisms will experience a significant boost. Some higher plants (sea grasses, for example) use dissolved carbon dioxide directly and probably will benefit from its rising levels, just as plants on land are expected to gain as the atmospheric concentration of this gas increases.

Many coral reefs are already in decline, and ocean acidification may push some over the edge into nonexistence.

How can scientists better gauge the response of ocean ecosystems to acidification? Most current efforts in this area involve short-term laboratory experiments on single species. Scientists have also mounted small-scale field studies to examine the acute effects that would accompany the deliberate disposal of atmospheric carbon dioxide in the deep sea, one of the various strategies being considered to sequester carbon dioxide and keep it out of the air [see "Can We Bury Global Warming?" by Robert H. Socolow; SCIENTIFIC AMERICAN, July 2005]. Although this work is informative, the results do not translate easily into an understanding of the consequences of long-term, chronic exposure to modestly lower pH. Nor is it straightforward to extrapolate from laboratory studies to whole ecosystems, where many different organisms interact.

One possibility for gaining a more realistic assessment of the problem would be to elevate carbon dioxide levels artificially for months to years in a patch of the ocean or on a coral reef. Experiments involving the large-scale manipulation of carbon dioxide levels are commonly carried out on land, but marine scientists and engineers are just now beginning to explore the logistics for extending this approach to the ocean. Another tactic is to study how marine organisms fare in regions that have long had lower pH, such as the Galápagos Islands, which are bathed in waters that are naturally rich in carbon dioxide.

Yet a third strategy might be to examine the geologic record of times when carbon dioxide concentrations reached much higher levels than that of the present and when ocean pH was presumably much lower—such as during an interval of anomalously warm climate that took place some 55 million years ago (the Paleocene-Eocene Thermal Maximum), when many marine organisms died off. The concern of many scientists today is that the current episode of acidification is taking place more rapidly than anything that has transpired in the past, leaving oceanic species no time to adapt. Although the effects may be hidden from people's view, dramatic alterations in the marine environment appear to be inevitable. SA

MORE TO EXPLORE

Anthropogenic Carbon and Ocean pH. Ken Caldeira and Michael E. Wickett in *Nature*, Vol. 425, page 365; September 25, 2003.

Anthropogenic Ocean Acidification over the Twenty-First Century and Its Impact on Calcifying Organisms. James C. Orr et al. in *Nature*, Vol. 437, pages 681–686; September 29, 2005.

Ocean Acidification Due to Increasing Atmospheric Carbon Dioxide. Royal Society, 2005. Available at **www.royalsoc.ac.uk/displaypagedoc.asp?id=13314**

Cognitive Radio

Smart radios and other new wireless devices will avoid transmission bottlenecks by switching instantly to nearby frequencies that they sense are clear

By Steven Ashley

Your favorite radio station transmits on a specific frequency. When you set your receiver to so many cycles per second, you tune the antenna circuit to pluck that station's frequency out of the ether. If other transmitters interfere with your reception, your only real option is to wait out the problem. In the best of all worlds, though, your receiver would respond by switching immediately to an open backup frequency that carries your station's broadcast. Such a solution is beyond today's radio technology, and perhaps that example makes the problem seem trivial. But now imagine that interference is interrupting an urgent, emergency cell-phone call. In that case, rapid transfer of the call to a clear cell channel would be more than merely convenient—it might save a life.

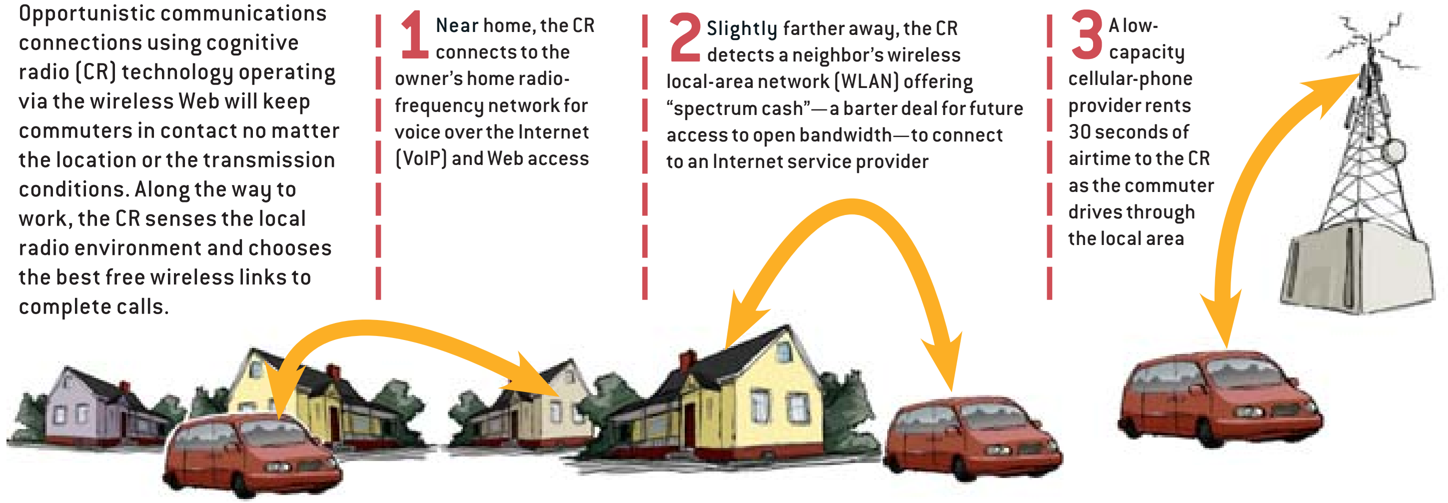
Engineers are now working to bring that kind of flexible operating intelligence to future radios, cell phones and other wireless communications devices. During the coming decade, cognitive radio technology should enable nearly any wireless system to locate and link to any locally available unused radio

WIRELESS SIGNALS jump automatically to an available, open frequency in cognitive radio. The result would be much more reliable transmissions—and maybe lower communications costs in the future.

MATT VINCENT

CAN YOU HEAR ME NOW?

Opportunistic communications connections using cognitive radio (CR) technology operating via the wireless Web will keep commuters in contact no matter the location or the transmission conditions. Along the way to work, the CR senses the local radio environment and chooses the best free wireless links to complete calls.



MATT VINCENT

spectrum to best serve the consumer. Employing adaptive software, these smart devices could reconfigure their communications functions to meet the demands of the transmission network or the user.

Cognitive radio technology will know what to do based on prior experience. On the morning drive to work, for instance, it would measure the propagation characteristics, signal strength and transmission quality of the different bands as it rides along with you [*see box above*]. The cognitive radio unit would thus build an internal database that defines how it should best operate in different places and at specific times of day. In contrast, the frequency bands and transmission protocol parameters of current wireless systems have been mostly fixed.

As cognitive radios send and receive signals, they will nimbly bound in and out of free bands as required, avoiding those that are already in use. This lightning-fast channel jumping should permit cognitive radio systems to transmit voice and data streams at reasonable speeds. By making much more efficient use of existing radio-frequency (RF) resources to work around spectrum-availability traffic jams, wireless communications should become far more dependable and convenient and perhaps considerably cheaper than it is today. Indeed, if cognitive radio technology progresses as its developers hope, a glut of RF-spectrum options may actually arise in time. The airwaves will never be the same again.

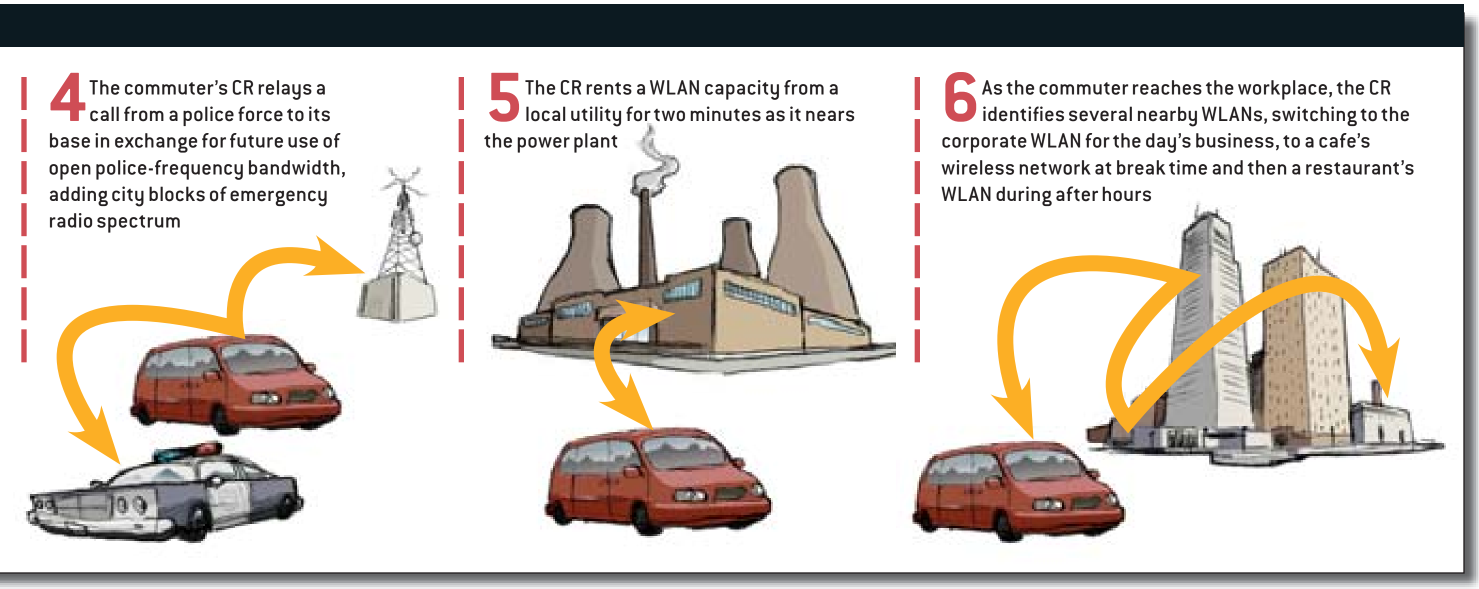
Overview/*Intelligent Radios*

- Cognitive radio is an emerging smart wireless communications technology that will be able to find and connect with any nearby open radio frequency to best serve the user. Thus, a cognitive radio should be able to switch from a band of the radio spectrum that is blocked by interference to a free one to complete a transmission link, a capability that is particularly important in an emergency.
- Adaptive software will enable these intelligent devices to reconfigure their functions to meet the demands of communications networks or consumers as needed. These alterations will be based on the ability to sense and remember various factors such as the radio-frequency spectrum, user behavior, or network state in different transmission environments at any one place and time. As a result, wireless communications should become far more dependable and convenient.
- The new flexibility afforded by cognitive radio may also eventually enable consumers to take advantage of cheaper wireless network paths available locally to make calls, a feature that would do much to revolutionize the communications business.

No Room on the Air

UNFORTUNATELY, those airwaves are all too crowded nowadays. Some bands are so overloaded that long waits and interference are the norm. The availability of these transmission links depends on the wireless systems in use. The radio spectrum—the segment of the electromagnetic continuum containing waves in the radio-frequency range—accommodates countless communications devices today. In the U.S., the Federal Communications Commission assigns users to specific frequencies. These include the well-known AM, FM, shortwave and citizens bands and VHF and UHF television channels, as well as hundreds of less familiar bands that serve cellular and cordless telephones, GPS trackers, air traffic control radars, security alarms, radio-controlled toys and the like [*see box on page 71*].

The present shortage of radio spectrum results in large part from the cost and performance limits of legacy hardware established during the past century. By the late 1950s, for instance, the holdover designs of vacuum-tube TV sets forced new transistor-based models to receive only VHF signals until engineers could revamp the sets some years after-



ward. Such hardware-related inflexibility is now being addressed by adaptive, software-based wireless designs.

This next-generation wireless technology, called software-defined radio (SDR), uses both embedded signal-processing algorithms to sift out weak radio signals and reconfigurable code structures to receive and transmit new radio protocols. Experts anticipate that in the relatively near term this software-driven advance will produce a seismic shift in radio design.

The change means, for example, that SDR code and other programmable radio-frequency front-end interface technologies running on a standard laptop computer (fitted with a small RF peripheral component interconnect card) could receive TV signals and display them. If the laptop were then equipped with an analog RF SDR card, it could upload software programming that would allow it to behave as a cellular handset or base station, a wireless personal organizer or even a military-frequency radio—whatever is required (and permitted) for the task at hand. Although few know it yet, the world has just entered the era of SDR wireless communications.

Cognitive radio is arriving on the heels of SDR technology and building on it. This new wireless paradigm involves SDR systems that can reconfigure their analog RF output and that incorporate "self-awareness" and knowledge of transmission protocols, etiquette and procedures. These developments will yield a cognitive radio able to sense its RF environment and location and then alter its power, frequency, modulation and other operating parameters so as to dynamically reuse whatever spectrum is available.

Self-awareness refers to the unit's ability to learn about itself and its relation to the radio networks it inhabits. Engineers can implement these functions through a computational model of the device and its environment that defines it as an individual entity ("Self") that operates as a "Radio"; the model also defines a "User" about whom the system can learn.

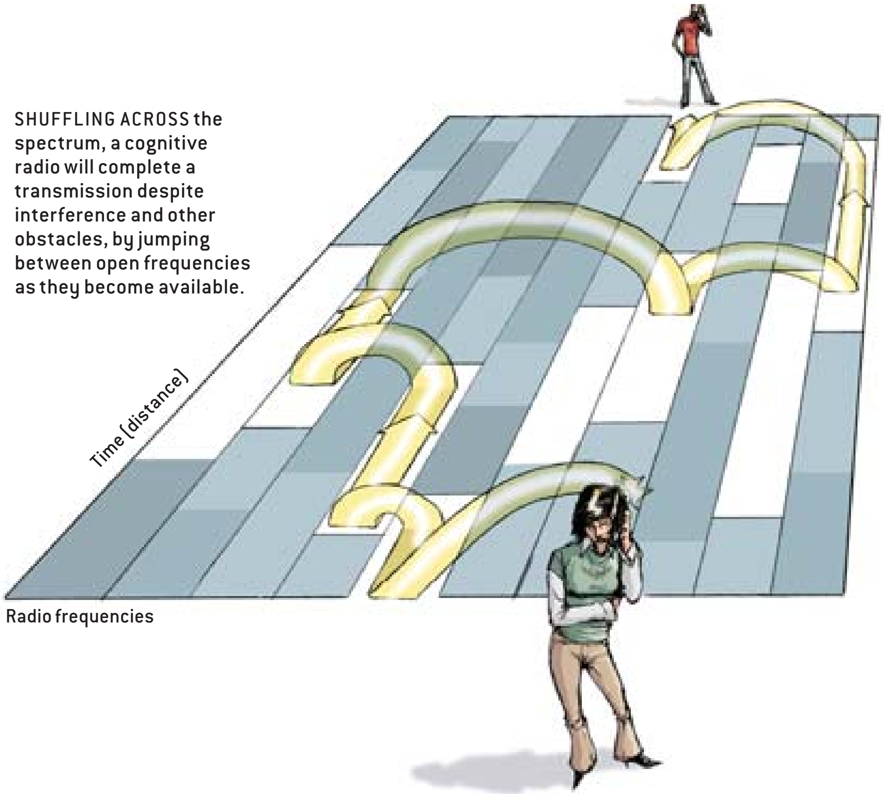
A cognitive radio will be able to autonomously sense how its RF environment varies with position and time in terms of the power that it and other transmitters in the vicinity radiate. These data structures and related software will enable a cognitive radio device to discover and use surrounding networks to the best advantage while avoiding interference from other radios. In the not too distant future, cognitive radio technology will share the available spectrum optimally without instructions from a controlling network, which could eventually liberate the user from user contracts and fees.

The potential for cognitive radio technology to redefine existing wireless services becomes clear when one considers their economics. A monthly cellphone service bill, for instance, contains charges for leasing radio spectrum, renting cell towers and purchasing the handset, as well as the amortization of the hardware at the cell base site, the cost of interconnections among cell sites, billing expenses and network operator profit. These fees pay for the investments that cellular service providers make to create and operate dedicated RF networks.

Such costs could drop dramatically, and service quality could improve greatly, when cognitive radio is finally unleashed in the marketplace. Think about the best advanced-technology cell phone now being sold. More than one gigahertz (GHz) of useful but underutilized radio spectrum is available to that handset. At any instant, however, the device employs at most 10 megahertz (MHz)—a hundredth of what there is—and even that is selected from only about 100 MHz of fixed spectrum allocations that the phone's circuits can access.

Furthermore, a typical cell phone incorporates several hundred million instructions per second of processing capacity that is largely dedicated to unique cellular standards. The service provider uploads these standards for its own purposes, such as fixing bugs in the software; they are not necessarily for the customer's immediate benefit. Case in point: that capacity could be used to securely upload third-party software

SHUFFLING ACROSS the spectrum, a cognitive radio will complete a transmission despite interference and other obstacles, by jumping between open frequencies as they become available.



that would permit the phone to link to a free wireless local-area network (WLAN). At a technical conference on mobile communications in 2004, a senior executive at Motorola stated that the WLAN-based telephone has been technically feasible for years but that cellular service providers do not want such a device. And it is little wonder: such a phone could automatically switch to corporate WLANs during the workday, depriving service providers of fees for hours every day.

But the cognitive radio genie is out of the bottle. SDR's entrée to little-used radio spectrum, together with the cognitive radio's autonomous control software (acting to the consumer's advantage), frames a business path toward the adoption of the technology.

Free Spectrum Abounds

EXCLUDING high-frequency and microwave bands greater than 6 GHz, about 2.8 GHz of currently allocated radio spectrum between 28 and 5,600 MHz is underutilized but accessible to cognitive radio. (That estimate derives from the nominal sensitivity of receivers and the gain levels of existing antennas.) Meanwhile the bands for cellular phones and wireless Internet services are often oversubscribed. Myriad electronic widgets, from keyless automobile entry fobs to garage-door openers to radio controls for toys, use those bands for short-range data communications; a gathering of users, such as a meeting of radio-controlled model-airplane enthusiasts, can swamp the allocated spectrum. Likewise, cellular bands that are typically almost empty at 3:30 in the morning are completely jammed at the peak calling time of 10 A.M. or during the evening homeward commute, particularly if road traffic is heavy.

Above 6 GHz, humidity and precipitation greedily absorb radio-frequency signals; even in dry air, absorption peaks near 20 and 60 GHz. Nevertheless, certain short-range data links (often categorized as "campus" or military "up the hill" links) now achieve megabit-per-second transmission rates at frequencies near 34 and 70 MHz. Rising computer power has recently enabled wireless devices operating in these upper bands to offer gigabits per second of instantaneous bandwidth within very small areas, or "picocells," of coverage. This technology could be helpful to mobile users communicating between vehicles on a highway, or among pedestrians, or in fixed wireless systems inside buildings.

Jens Zander, an authority on radio systems at the Royal Institute of Technology in Stockholm, argues that there is no shortage of radio spectrum, only a dearth of affordable communications infrastructure. Cellular-telephone towers, interconnections to the public switched-telephone network and billing systems, and so on make up the substantial and expensive underpinnings essential for leased spectrum. Since the 1990s, as cell phones shrank from the brick-size "bag phone" to the Motorola StarTac and on to today's multifunction clamshell device, building and maintaining a dedicated infrastructure was the only way to proceed. In early 2005, however, Vanu, Inc., demonstrated the first Global System for Mobile Communications (GSM) SDR base station with an RF converter that makes the radio signal processable by essentially a high-performance laptop without a keyboard or display. Only five years before, the GSM transcoder and rate adaptation unit alone needed its own server rack and kilowatts of operating power. During that period, semiconductor advances reduced the cost of a base station for an affordable small cell site, so that nowadays it could be a laptop or home computer.

Changes in the Air

THE ONGOING microelectronics and computer technology revolution has thus altered the fundamental limits of radio hardware during the past decade, cutting the costs of cellular infrastructure systems to less than 1 percent of what they were. The impact of these transformations on advanced wireless technology and their markets is just now being felt.

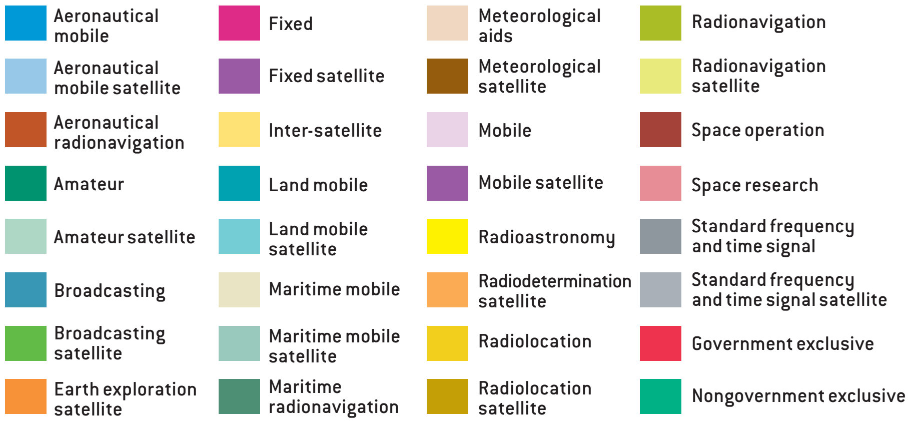
In earlier years analog TV (using specialized hardware and with 6 MHz bandwidth) was the largest practical consumer of the radio spectrum. At present, high-definition digital TV delivers the equivalent of nearly 100-megabit-per-second transmission rates in the same 6 MHz band. An Intel Pentium-powered laptop can now generate pictures and sounds using software and a digitized version of the analog TV signal from an RF converter unit. This converter changes the carrier frequency of a radio signal from RF at the antenna to some intermediate

MATT VINCENT

RADIO-FREQUENCY SPECTRUM

Radio spectrum in the U.S. is allocated by the Federal Communications Commission to a large variety of users and applications, with each assigned to a specific bandwidth located between frequencies of nine kilohertz—9,000 cycles per second—and 300 gigahertz. Here is a simplified representation of the radio-frequency spectrum and its allocations in the U.S.

RADIO SERVICES



SOURCE: U.S. DEPARTMENT OF COMMERCE

frequency that an analog-to-digital conversion chip can transform into a software-processable format. High-speed analog-to-digital conversion chips can thus exploit hundreds of megahertz of RF spectrum simultaneously. Some such chips are powered by microelectromechanical system (MEMS) circuits—semiconductors that incorporate micron-scale mechanical devices—such as a digitally reconfigurable analog RF capacitor. In production quantities, MEMS-based RF peripheral cards can access tens of megahertz of radio spectrum anywhere between 30 and 5,600 MHz for about the same price as a present-day cell phone.

MEMS-based RF devices have been slow to enter the market because they cost more than the less capable fixed-RF chip sets. A landmark ruling by the FCC in 2004 that favored the development of cognitive radio, however, offers new incentives for manufacturers to adopt RF MEMS products. The government agency recommended cognitive radio technology for use in low-power ad hoc networks in unused TV bands. This decision released more than 100 MHz for cognitive radio in typical urban markets. The emergence of RF MEMS and the FCC endorsement combine to push for greater spectrum sharing in days to come. Operating in the low and middle bands of the radio spectrum, one or two RF MEMS analog channels can create short-range ad hoc networks in any band where licensed users agree to rent, share or otherwise barter for radio spectrum.

An RF MEMS cognitive radio card can therefore turn a cell phone into a WLAN, a laptop into a cell phone or a cordless telephone into a picocell "tower." From such a picocell, a home computer fitted with a cognitive radio control system could rent airtime to passersby, billing for secure wireless voice or data through the associated Internet service provider.

Remaking the Wireless Web

IN TRADITIONAL cell-phone systems, most of the intelligence for efficient operation resides in the network. Although newer cell technologies feature greater processing capacity, they are really not that much smarter than their predecessors. Customers still need a contract with a service provider to obtain access to the network and then to the public switched-telephone network. Cognitive radio technology, in contrast, embeds the intelligence required to connect to wireless networks in the radio handset, laptop or wireless organizer. Because a cognitive control subsystem governs the SDR capabilities, a unit can detect RF-networking opportunities wherever it finds itself.

At present, 90 percent of new laptop computers contain WLAN capabilities. Home and business WLANs and related hot spots are multiplying exponentially. Cognitive radio will have the operational intelligence to rent or borrow WLAN and other RF spectrum quickly for seconds or minutes at a time in exchange for "spectrum cash," a verifiable promise to loan the cognitive radio's own picocell capabilities to another cognitive radio in the future. From these wireless Web access points, the Internet service provider would then transfer the user's data or call to anyone, anywhere in the world. One can see then that cognitive radio does not need a dedicated cellular network to connect a user via wireless and the Internet to other devices. In addition, as the cognitive radio interactions with the wireless Web expand, the need for a long-term contract with a cellular-service provider diminishes.

Spatial Radio Knowledge

WHEN A TYPICAL CONSUMER makes use of a wireless network employing current commercial electronics, the system does its best to consume as much

ALL COMMUNICATIONS LINKS, ALL THE TIME

Fast, reliable battlefield communications—a key to victory in modern warfare—could be guaranteed by cognitive radio technology. Whereas different forces and weapons systems today operate radio systems that can be incompatible, next-generation smart radio technology could help military commanders stay apprised of the latest situation in (and above) the combat zone with real-time voice, data and video links that reliably connect all friendly forces, despite interference caused by the fog of war. Future military radios could use cognitive radio technology to maintain these crucial communications lifelines.

MATT VINCENT

scarce spectrum as it can while simultaneously jamming other nearby radios. Cognitive radio will be smart enough to introduce etiquette—sensible transactional practices—into RF-spectrum operations. It will also intelligently detect and interact with nearby picocells to keep the cognitive radio user connected by the means that best serve his or her needs, which may differ among various times and situations.

To accomplish these tasks, a cognitive radio unit requires several things. First, it must "know" how radiated RF power at its location varies with distance along the ground, among obstructions and up in the air. Cell phones do not need this information because the fixed network employs dedicated radio spectrum that has been previously calibrated for existing radiated power patterns. Cognitive radios instead sense the entire local RF environment of low, medium and high bands, mapping its features as a function of space, time and frequency propagation. The development of spectrum-sensing cognitive radio will require the design of high-quality sensor devices and practical algorithms for exchanging spectrum-monitoring data between cooperating communications nodes. Systems that feature multiple-input/multiple-output capabilities will direct transmissions along complex multipath components—thereby accounting for reflections of signals from objects such as buildings and vehicles—and away from other potentially interfering radios.

A fully functional cognitive radio system will be smart enough to sense the local RF "scene," to choose the radio band, mode and service it needs as well as the SDR upload connections to the selected band and mode. It will then direct its transmission energy toward the intended receiver while minimizing interference with other radios, including cognitive ones. Thus, it will display a high level of spectrum etiquette and connect the user securely and privately.

The accuracy of such operations could be improved by the development of three-dimensional computer representations of the full local cityscape stored on gigabyte hard drives, which would be accessed wirelessly as needed. Predictions of received signal strength based on these models would allow cognitive radios to avoid most interference. Having standardized broadcast channels over which cognitive radios experiencing interference could "complain" without jamming others would complete the radio etiquette cycle.

The ideal of cognitive radio etiquette is complicated by the variation over time of the aggregate interference produced by the environment, which includes that created by natural electrical noise (from lightning), electrical power generators, electric motors, automobile ignition systems and radio transmitters. The effects of these RF sources change over time. At night, for example, few elevators are active, so their electric drive motors produce little noise, but during rush hour that noise grows. The total power radiated by all sources tends to be greatest in urban centers at about 10 A.M., with less activity occurring in rural settings and at night. Although the statistical complexity of such aggregate sources makes them difficult to predict, cognitive radio will learn the patterns at important locations (such as at the workplace and home) for known users.

The Future of Smart Radio

AFTER ASCERTAINING the varying energy patterns in each band, cognitive radio devices will be able to use Semantic Web technology [see "The Semantic Web," by Tim Berners-Lee, James Hendler and Ora Lassila; SCIENTIFIC AMERICAN, May 2001] to exchange this information freely with others. It will help to optimize each unit's search for underused and rentable spectrum. Cognitive radios can thereby avoid jamming other users yet transmit with sufficient power to overcome ambient interference and cooperate creatively.

The decisions concerning the future of cognitive radio technology are shaping up as a battle between two giant business sectors: the cell-phone and telecommunications industry versus "the Internet industry," which includes Microsoft, Intel, Google, Internet service providers and consumer computer firms. Although entrenched interests may resist it, progress toward cognitive radio seems likely because the relative chaos and inflexibility of unregulated radio bands (such as those now handling instrumentation, scientific and medical devices) could be avoided. Ultimately, smart operator etiquette based on cognitive radio technology will turn gigahertz of underused spectrum into wideband connectivity for many users.

If FCC regulators continue on the current path, they will make huge single-use swaths of spectrum hundreds of megahertz wide available for shared use. The long-touted scarcity of radio spectrum in the future may be replaced by a surfeit of available frequencies. Rather than a cell phone needing a minute to upload a compressed megapixel-size image, it might be able to handle 10 such images a second.

Just as the emergence of cell-phone technology has led to wide social and business consequences, cognitive radio's adoption will induce similar changes as advanced devices exploit the wireless Web to displace now traditional cell phones. The growth of cognitive radio will take some time to occur, but the effect on all our lives will be significant. SA

Steven Ashley is a staff editor.

MORE TO EXPLORE

Cognitive Radio Shows Great Promise. Bruce Fette in *COTS Journal;* October 2004. **www.cotsjournalonline.com/home/article.php?id=100206**

Berkeley Wireless Research Center Cognitive Radio Workshop, November 1, 2004. **bwrc.eecs.berkeley.edu/Research/MCMA/**

Cognitive Radio Architecture. Joseph Mitola III. John Wiley & Sons, April 2006.

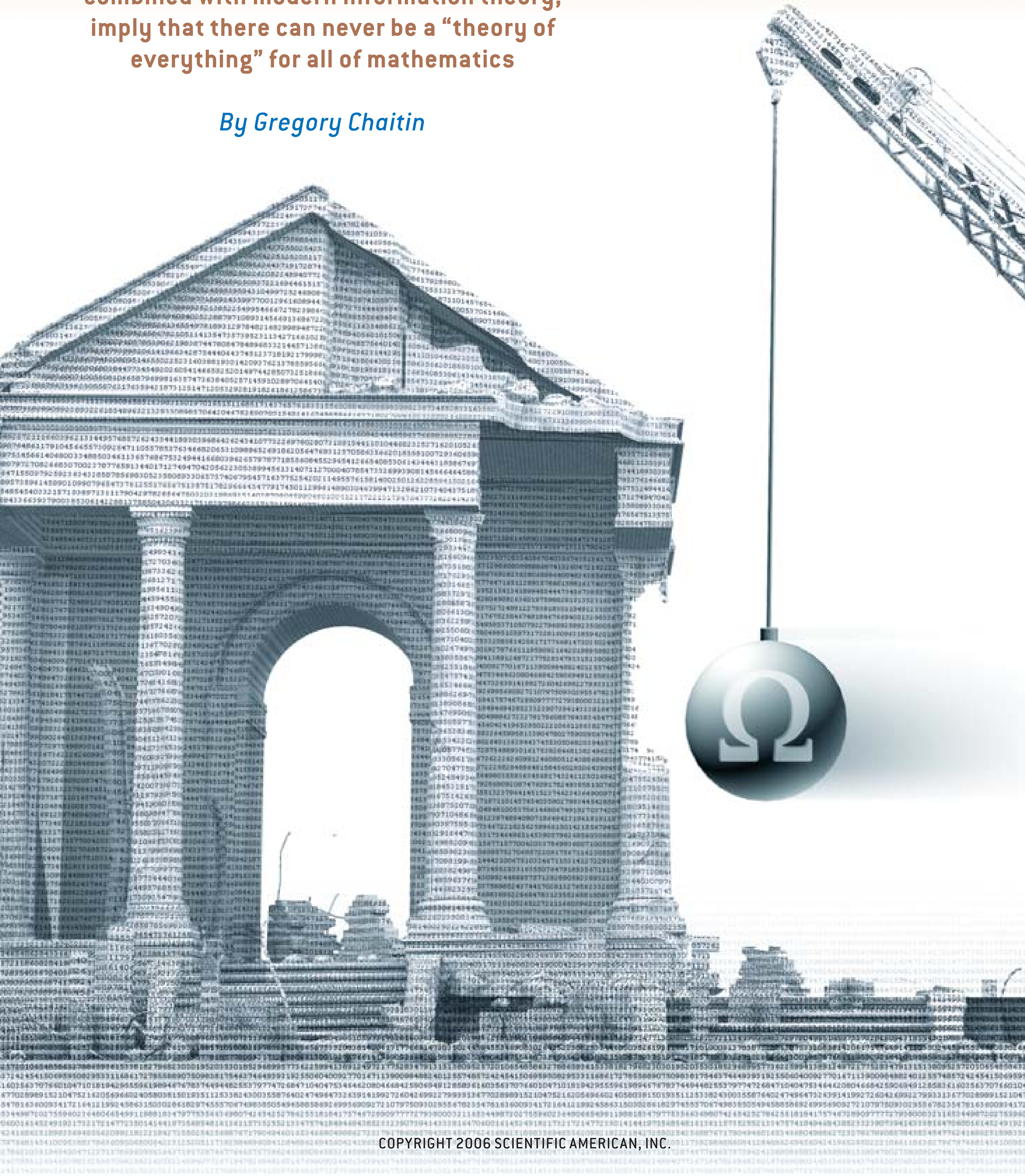
European Community's End-to-End Reconfigurability (E2R) radio project: **phase2.e2r.motlabs.com**

Joseph Mitola's Web site is at the KTH (Royal Institute of Technology, Sweden): **www.it.kth.se/~jmitola**

Software Defined Radio Forum: **www.sdrforum.org**

Ideas on complexity and randomness originally suggested by Gottfried W. Leibniz in 1686, combined with modern information theory, imply that there can never be a "theory of everything" for all of mathematics

By Gregory Chaitin



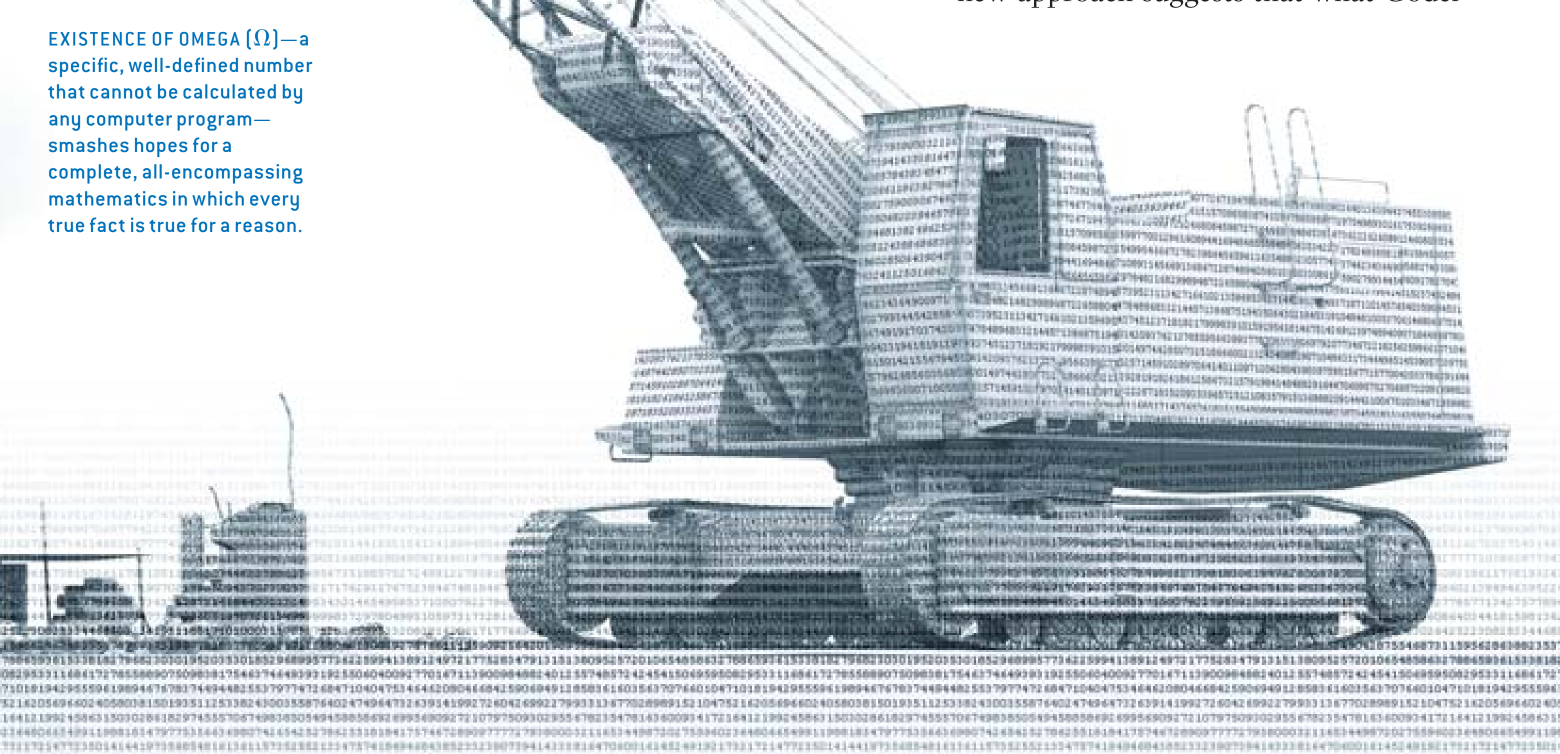
The Limits of Reason

In 1956 *Scientific American* published an article by Ernest Nagel and James R. Newman entitled "Gödel's Proof." Two years later the writers published a book with the same title—a wonderful work that is still in print. I was a child, not even a teenager, and I was obsessed by this little book. I remember the thrill of discovering it in the New York Public Library. I used to carry it around with me and try to explain it to other children.

It fascinated me because Kurt Gödel used mathematics to show that mathematics itself has limitations. Gödel refuted the position of David Hilbert, who about a century ago declared that there was a theory of everything for math, a finite set of principles from which one could mindlessly deduce all mathematical truths by tediously following the rules of symbolic logic. But Gödel demonstrated that mathematics contains true statements that cannot be proved that way. His result is based on two self-referential paradoxes: "This statement is false" and "This statement is unprovable." (For more on Gödel's incompleteness theorem, see www.sciam.com/ontheweb)

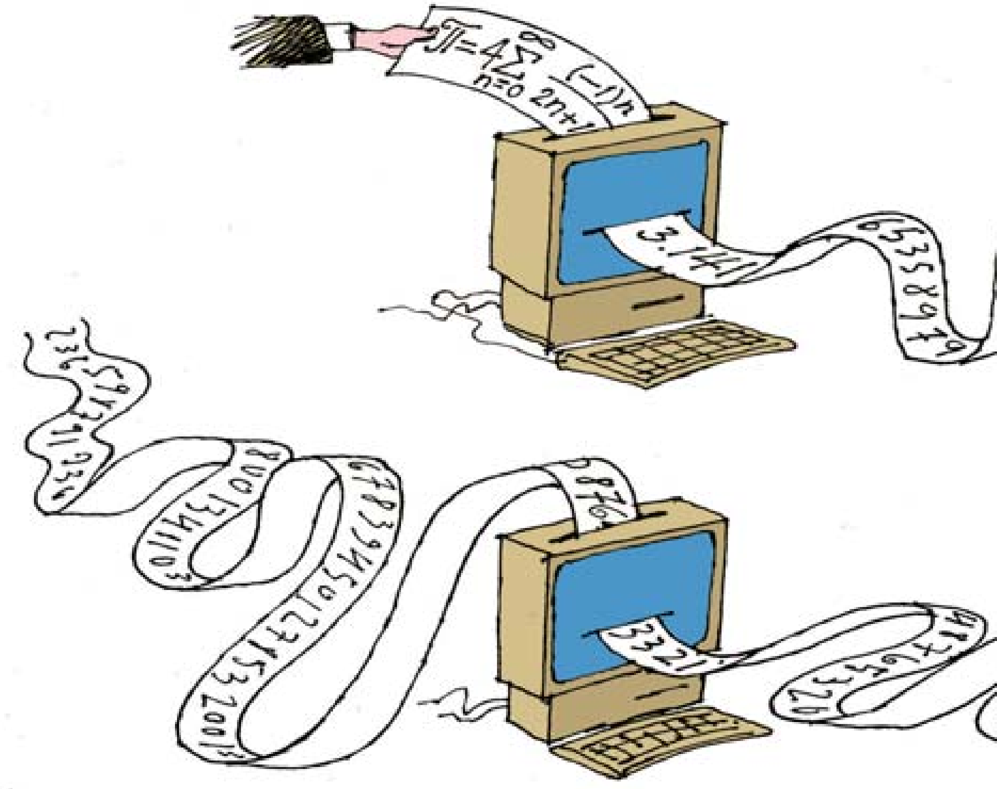
My attempt to understand Gödel's proof took over my life, and now half a century later I have published a little book of my own. In some respects, it is my own version of Nagel and Newman's book, but it does not focus on Gödel's proof. The only things the two books have in common are their small size and their goal of critiquing mathematical methods.

Unlike Gödel's approach, mine is based on measuring information and showing that some mathematical facts cannot be compressed into a theory because they are too complicated. This new approach suggests that what Gödel



EXISTENCE OF OMEGA (Ω)—a specific, well-defined number that cannot be calculated by any computer program—smashes hopes for a complete, all-encompassing mathematics in which every true fact is true for a reason.

discovered was just the tip of the iceberg: an infinite number of true mathematical theorems exist that cannot be proved from any finite system of axioms.



Complexity and Scientific Laws

MY STORY BEGINS in 1686 with Gottfried W. Leibniz's philosophical essay *Discours de métaphysique* (*Discourse on Metaphysics*), in which he discusses how one can distinguish between facts that can be described by some law and those that are lawless, irregular facts. Leibniz's very simple and profound idea appears in section VI of the *Discours,* in which he essentially states that a theory has to be simpler than the data it explains, otherwise it does not explain anything. The concept of a law becomes vacuous if arbitrarily high mathematical complexity is permitted, because then one can always construct a law no matter how random and patternless the data really are. Conversely, if the only law that describes some data is an extremely complicated one, then the data are actually lawless.

Today the notions of complexity and simplicity are put in precise quantitative terms by a modern branch of mathematics called algorithmic information theory. Ordinary information theory quantifies information by asking how many bits are needed to encode the information. For example, it takes one bit to encode a single yes/no answer. Algorithmic information, in contrast, is defined by asking what size computer program is necessary to generate the data. The minimum number of bits—what size string of zeros and ones—needed to store the program is called the algorithmic information content of the data. Thus, the infinite sequence of numbers 1, 2, 3, … has very little algorithmic information; a very short computer program can generate all those numbers. It does not matter how long the program must take to do the computation or how much memory it must use—just the length of the program in bits counts. (I gloss over the question of what programming language is used to write the program—for a rigorous definition, the language would have to be specified precisely. Different programming languages would result in somewhat different values of algorithmic information content.)

To take another example, the number pi, 3.14159…, also has only a little algorithmic information content, because a relatively short algorithm can be programmed into a computer to compute digit after digit. In contrast, a random number with a mere million digits, say 1.341285…64, has a much larger amount of algorithmic information. Because the number lacks a defining pattern, the shortest program for outputting it will be about as long as the number itself:

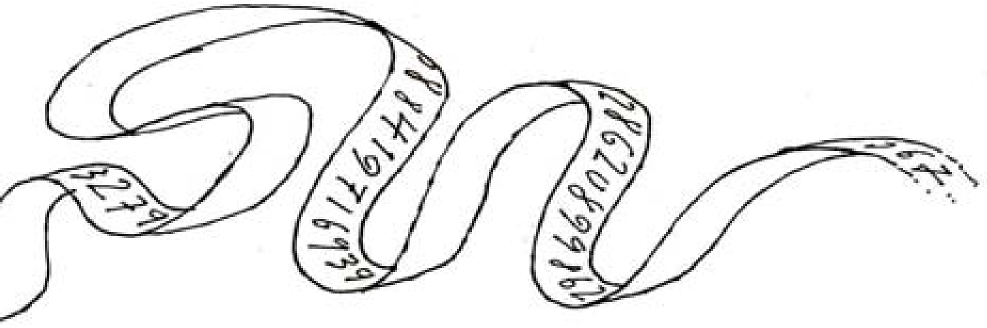
```
Begin
Print "1.341285…64"
End
```

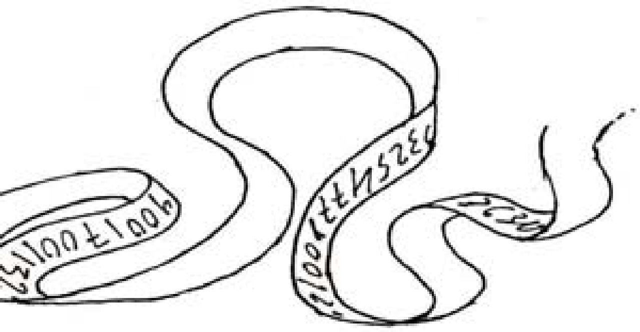
(All the digits represented by the ellipsis are included in the program.) No smaller program can calculate that se-

Overview/*Irreducible Complexity*

- Kurt Gödel demonstrated that mathematics is necessarily incomplete, containing true statements that cannot be formally proved. A remarkable number known as omega reveals even greater incompleteness by providing an infinite number of theorems that cannot be proved by any finite system of axioms. A "theory of everything" for mathematics is therefore impossible.
- Omega is perfectly well defined [*see box on opposite page*] and has a definite value, yet it cannot be computed by any finite computer program.
- Omega's properties suggest that mathematicians should be more willing to postulate new axioms, similar to the way that physicists must evaluate experimental results and assert basic laws that cannot be proved logically.
- The results related to omega are grounded in the concept of algorithmic information. Gottfried W. Leibniz anticipated many of the features of algorithmic information theory more than 300 years ago.

KENN BROWN; CONCEPT BY DUSAN PETRICIC (*preceding pages*); DUSAN PETRICIC (*above*)





ALGORITHMIC INFORMATION quantifies the size of a computer program needed to produce a specific output. The number pi has little algorithmic information content because a short program can produce pi. A random number has a lot of algorithmic information; the best that can be done is to input the number itself. The same is true of the number omega.

quence of digits. In other words, such digit streams are incompressible, they have no redundancy; the best that one can do is transmit them directly. They are called irreducible or algorithmically random.

How do such ideas relate to scientific laws and facts? The basic insight is a software view of science: a scientific theory is like a computer program that predicts our observations, the experimental data. Two fundamental principles inform this viewpoint. First, as William of Occam noted, given two theories that explain the data, the simpler theory is to be preferred (Occam's razor). That is, the smallest program that calculates the observations is the best theory. Second is Leibniz's insight, cast in modern terms—if a theory is the same size in bits as the data it explains, then it is worthless, because even the most random of data has a theory of that size. A useful theory is a compression of the data; comprehension is compression. You compress things into computer programs, into concise algorithmic descriptions. The simpler the theory, the better you understand something.

Sufficient Reason

DESPITE LIVING 250 years before the invention of the computer program, Leibniz came very close to the modern idea of algorithmic information. He had all the key elements. He just never connected them. He knew that everything can be represented with binary information, he built one of the first calculating machines, he appreciated the power of computation, and he discussed complexity and randomness.

If Leibniz had put all this together, he might have questioned one of the key pillars of his philosophy, namely, the principle of sufficient reason—that everything happens for a reason. Furthermore, if something is true, it must be true for a reason. That may be hard to believe sometimes, in the confusion and chaos of daily life, in the contingent ebb and flow of human history. But even if we cannot always see a reason (perhaps because the chain of reasoning is long and subtle), Leibniz asserted, God can see the reason. It is there! In that, he agreed with the ancient Greeks, who originated the idea.

Mathematicians certainly believe in reason and in Leibniz's principle of sufficient reason, because they always try to prove everything. No matter how much evidence there is for a theorem, such as millions of demonstrated examples, mathematicians demand a proof of the general case. Nothing less will satisfy them.

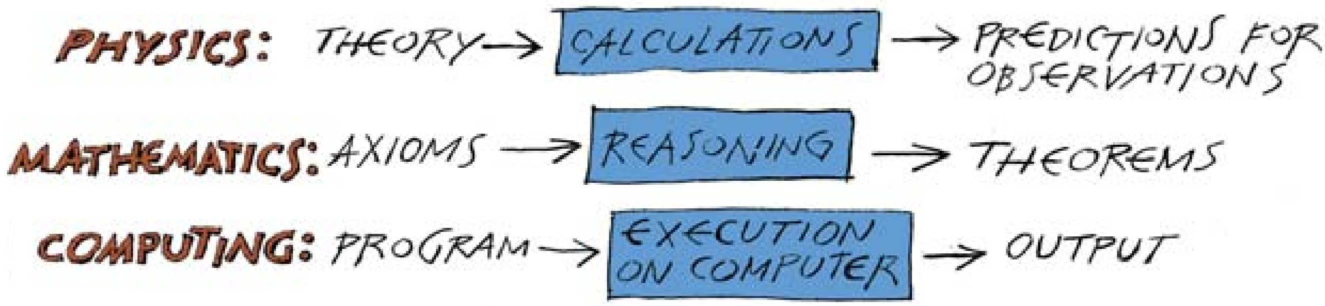
And here is where the concept of algorithmic information can make its surprising contribution to the philosophical discussion of the origins and limits of knowledge. It reveals that certain mathematical facts are true for no rea-

How Omega Is Defined

To see how the value of the number omega is defined, look at a simplified example. Suppose that the computer we are dealing with has only three programs that halt, and they are the bit strings 110, 11100 and 11110. These programs are, respectively, 3, 5 and 5 bits in size. If we are choosing programs at random by flipping a coin for each bit, the probability of getting each of them by chance is precisely $1/2^3$, $1/2^5$ and $1/2^5$, because each particular bit has probability $1/2$. So the value of omega (the halting probability) for this particular computer is given by the equation:

$$\text{omega} = 1/2^3 + 1/2^5 + 1/2^5 = .001 + .00001 + .00001 = .00110$$

This binary number is the probability of getting one of the three halting programs by chance. Thus, it is the probability that our computer will halt. Note that because program 110 halts we do not consider any programs that start with 110 and are larger than three bits—for example, we do not consider 1100 or 1101. That is, we do not add terms of .0001 to the sum for each of those programs. We regard all the longer programs, 1100 and so on, as being included in the halting of 110. Another way of saying this is that the programs are self-delimiting; when they halt, they stop asking for more bits. *—G.C.*



PHYSICS AND MATHEMATICS are in many ways similar to the execution of a program on a computer.

son, a discovery that flies in the face of the principle of sufficient reason.

Indeed, as I will show later, it turns out that an infinite number of mathematical facts are irreducible, which means no theory explains why they are true. These facts are not just computationally irreducible, they are logically irreducible. The only way to "prove" such facts is to assume them directly as new axioms, without using reasoning at all.

The concept of an "axiom" is closely related to the idea of logical irreducibility. Axioms are mathematical facts that we take as self-evident and do not try to prove from simpler principles. All formal mathematical theories start with axioms and then deduce the consequences of these axioms, which are called theorems. That is how Euclid did things in Alexandria two millennia ago, and his treatise on geometry is the classical model for mathematical exposition.

In ancient Greece, if you wanted to convince your fellow citizens to vote with you on some issue, you had to reason with them—which I guess is how the Greeks came up with the idea that in mathematics you have to prove things rather than just discover them experimentally. In contrast, previous cultures in Mesopotamia and Egypt apparently relied on experiment. Using reason has certainly been an extremely fruitful approach, leading to modern mathematics and mathematical physics and all that goes with them, including the technology for building that highly logical and mathematical machine, the computer.

So am I saying that this approach that science and mathematics has been following for more than two millennia crashes and burns? Yes, in a sense I am. My counterexample illustrating the limited power of logic and reason, my source of an infinite stream of unprovable mathematical facts, is the number that I call omega.

The Number Omega

THE FIRST STEP on the road to omega came in a famous paper published precisely 250 years after Leibniz's essay. In a 1936 issue of the *Proceedings of the London Mathematical Society,* Alan M. Turing began the computer age by presenting a mathematical model of a simple, general-purpose, programmable digital computer. He then asked, Can we determine whether or not a computer program will ever halt? This is Turing's famous halting problem.

Of course, by running a program you can eventually discover that it halts, if it halts. The problem, and it is an extremely fundamental one, is to decide when to give up on a program that does not halt. A great many special cases can be solved, but Turing showed that a general solution is impossible. No algorithm, no mathematical theory, can ever tell us which programs will halt and which will not. (For a modern proof of Turing's thesis, see www.sciam.com/ontheweb) By the way, when I say "program," in modern terms I mean the concatenation of the computer program and the data to be read in by the program.

The next step on the path to the number omega is to consider the ensemble of all possible programs. Does a program chosen at random ever halt? The probability of having that happen is my omega number. First, I must specify how to pick a program at random. A program is simply a series of bits, so flip a coin to determine the value of each bit. How many bits long should the program be? Keep flipping the coin so long as the computer is asking for another bit of input. Omega is just the probability that the machine will eventually come to a halt when supplied with a stream of random bits in this fashion. (The precise numerical value of omega depends on the choice of computer programming language, but omega's surprising properties are not affected by this choice. And once you have chosen a language, omega has a definite value, just like pi or the number 3.)

Being a probability, omega has to be greater than 0 and less than 1, because some programs halt and some do not. Imagine writing omega out in binary. You would get something like 0.1110100.... These bits after the decimal point form an irreducible stream of bits. They are our irreducible mathematical facts (each fact being whether the bit is a 0 or a 1).

Omega can be defined as an infinite sum, and each N-bit program that halts contributes precisely $1/2^N$ to the sum [*see box on preceding page*]. In other words,

THE AUTHOR

GREGORY CHAITIN is a researcher at the IBM Thomas J. Watson Research Center. He is also honorary professor at the University of Buenos Aires and visiting professor at the University of Auckland. He is co-founder, with Andrei N. Kolmogorov, of the field of algorithmic information theory. His nine books include the nontechnical works *Conversations with a Mathematician* (2002) and *Meta Math!* (2005). When he is not thinking about the foundations of mathematics, he enjoys hiking and snowshoeing in the mountains.

DUSAN PETRICIC

each N-bit program that halts adds a 1 to the Nth bit in the binary expansion of omega. Add up all the bits for all programs that halt, and you would get the precise value of omega. This description may make it sound like you can calculate omega accurately, just as if it were the square root of 2 or the number pi. Not so—omega is perfectly well defined and it is a specific number, but it is impossible to compute in its entirety.

We can be sure that omega cannot be computed because knowing omega would let us solve Turing's halting problem, but we know that this problem is unsolvable. More specifically, knowing the first N bits of omega would enable you to decide whether or not each program up to N bits in size ever halts [*see box on page 80*]. From this it follows that you need at least an N-bit program to calculate N bits of omega.

Note that I am not saying that it is impossible to compute some digits of omega. For example, if we knew that computer programs 0, 10 and 110 all halt, then we would know that the first digits of omega were 0.111. The point is that the first N digits of omega cannot be computed using a program significantly shorter than N bits long.

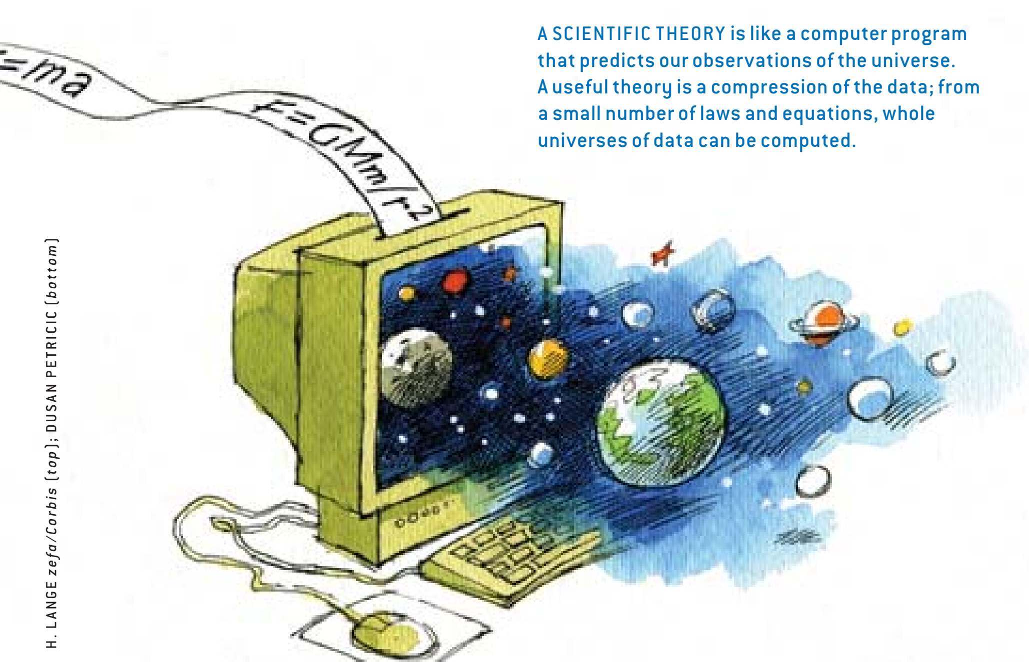
Most important, omega supplies us with an infinite number of these irreducible bits. Given any finite program, no matter how many billions of bits long, we have an infinite number of bits that the program cannot compute. Given any finite set of axioms, we have an infinite number of truths that are unprovable in that system.

Because omega is irreducible, we can immediately conclude that a theory of everything for all of mathematics cannot exist. An infinite number of bits of omega constitute mathematical facts (whether each bit is a 0 or a 1) that cannot be derived from any principles simpler than the string of bits itself. Mathematics therefore has infinite complexity, whereas any individual theory of everything would have only finite complexity and could not capture all the richness of the full world of mathematical truth.

This conclusion does not mean that proofs are no good, and I am certainly not against reason. Just because some things are irreducible does not mean we should give up using reasoning. Irreducible principles—axioms—have always been a part of mathematics. Omega just shows that a lot more of them are out there than people suspected.

So perhaps mathematicians should not try to prove everything. Sometimes they should just add new axioms. That is what you have got to do if you are faced with irreducible facts. The problem is realizing that they are irreducible! In a way, saying something is irreducible is giving up, saying that it cannot ever be proved. Mathematicians would rather die than do that, in sharp contrast with their physicist colleagues, who are happy to be pragmatic and to use plausible reasoning instead of rigorous proof. Physicists are willing to add new principles, new scientific laws, to understand new domains of experience. This raises what I think is an extremely interesting question: Is mathematics like physics?

GOTTFRIED W. LEIBNIZ, commemorated by a statue in Leipzig, Germany, anticipated many of the features of modern algorithmic information theory more than 300 years ago.



A SCIENTIFIC THEORY is like a computer program that predicts our observations of the universe. A useful theory is a compression of the data; from a small number of laws and equations, whole universes of data can be computed.

H. LANGE *zefa/Corbis* (*top*); DUSAN PETRICIC (*bottom*)

Mathematics and Physics

THE TRADITIONAL VIEW is that mathematics and physics are quite different. Physics describes the universe and depends on experiment and observation. The particular laws that govern our universe—whether Newton's laws of motion or the Standard Model of particle physics—must be determined empirically and then asserted like axioms that cannot be logically proved, merely verified.

Mathematics, in contrast, is somehow independent of the universe. Results and theorems, such as the properties of the integers and real numbers, do not depend in any way on the particular nature of reality in which we find ourselves. Mathematical truths would be true in any universe.

Yet both fields are similar. In physics, and indeed in science generally, scientists compress their experimental observations into scientific laws. They then show how their observations can be deduced from these laws. In mathematics, too, something like this happens—mathematicians compress their computational experiments into mathematical axioms, and they then show how to deduce theorems from these axioms.

If Hilbert had been right, mathematics would be a closed system, without room for new ideas. There would be a static, closed theory of everything for all of mathematics, and this would be like a dictatorship. In fact, for mathematics to progress you actually need new ideas and plenty of room for creativity. It does not suffice to grind away, mechanically deducing all the possible consequences of a fixed number of basic principles. I much prefer an open system. I do not like rigid, authoritarian ways of thinking.

Another person who thought mathematics is like physics was Imre Lakatos, who left Hungary in 1956 and later worked on philosophy of science in England. There Lakatos came up with a great word, "quasi-empirical," which means that even though there are no true experiments that can be carried out in mathematics, something similar does take place. For example, the Goldbach conjecture states that any even number greater than 2 can be expressed as the sum of two prime numbers. This conjecture was arrived at experimentally, by noting empirically that it was true for every even number that anyone cared to examine. The conjecture has not yet been proved, but it has been verified up to 10^{14}.

I think that mathematics is quasi-empirical. In other words, I feel that mathematics is different from physics (which is truly empirical) but perhaps not as different as most people think.

I have lived in the worlds of both mathematics and physics, and I never thought there was such a big difference between these two fields. It is a matter of degree, of emphasis, not an absolute difference. After all, mathematics and physics coevolved. Mathematicians should not isolate themselves. They should not cut themselves off from rich sources of new ideas.

New Mathematical Axioms

THE IDEA OF CHOOSING to add more axioms is not an alien one to mathematics. A well-known example is the parallel postulate in Euclidean geometry: given a line and a point not on the line, there is exactly one line that can be drawn through the point that never intersects the original line. For centuries geometers wondered whether that result could be proved using the rest of Euclid's axioms. It could not. Finally, mathematicians realized that they could substitute different axioms in place of the Euclidean version, thereby producing the non-Euclidean geometries of curved spaces, such as the surface of a sphere or of a saddle.

Why Is Omega Incompressible?

I wish to demonstrate that omega is incompressible—that one cannot use a program substantially shorter than *N* bits long to compute the first *N* bits of omega. The demonstration will involve a careful combination of facts about omega and the Turing halting problem that it is so intimately related to. Specifically, I will use the fact that the halting problem for programs up to length *N* bits cannot be solved by a program that is itself shorter than *N* bits (see www.sciam.com/ontheweb).

My strategy for demonstrating that omega is incompressible is to show that having the first *N* bits of omega would tell me how to solve the Turing halting problem for programs up to length *N* bits. It follows from that conclusion that no program shorter than *N* bits can compute the first *N* bits of omega. (If such a program existed, I could use it to compute the first *N* bits of omega and then use those bits to solve Turing's problem up to *N* bits—a task that is impossible for such a short program.)

Now let us see how knowing *N* bits of omega would enable me to solve the halting problem—to determine which programs halt—for all programs up to *N* bits in size. Do this by performing a computation in stages. Use the integer *K* to label which stage we are at: $K = 1, 2, 3, \ldots$

At stage *K*, run every program up to *K* bits in size for *K* seconds. Then compute a halting probability, which we will call omega_K, based on all the programs that halt by stage *K*. Omega_K will be less than omega because it is based on only a subset of all the programs that halt eventually, whereas omega is based on *all* such programs.

As *K* increases, the value of omega_K will get closer and closer to the actual value of omega. As it gets closer to omega's actual value, more and more of omega_K's first bits will be correct—that is, the same as the corresponding bits of omega.

And as soon as the first *N* bits are correct, you know that you have encountered every program up to *N* bits in size that will ever halt. (If there were another such *N*-bit program, at some later-stage *K* that program would halt, which would increase the value of omega_K to be greater than omega, which is impossible.)

So we can use the first *N* bits of omega to solve the halting problem for all programs up to *N* bits in size. Now suppose we could compute the first *N* bits of omega with a program substantially shorter than *N* bits long. We could then combine that program with the one for carrying out the omega_K algorithm, to produce a program shorter than *N* bits that solves the Turing halting problem up to programs of length *N* bits.

But, as stated up front, we know that no such program exists. Consequently, the first *N* bits of omega must require a program that is almost *N* bits long to compute them. That is good enough to call omega incompressible or irreducible. (A compression from *N* bits to almost *N* bits is not significant for large *N*.) *—G.C.*

OMEGA represents a part of mathematics that is in a sense unknowable. A finite computer program can reveal only a finite number of omega's digits; the rest remain shrouded in obscurity.

KENN BROWN; CONCEPT BY DUSAN PETRICIC

Other examples are the law of the excluded middle in logic and the axiom of choice in set theory. Most mathematicians are happy to make use of those axioms in their proofs, although others do not, exploring instead so-called intuitionist logic or constructivist mathematics. Mathematics is not a single monolithic structure of absolute truth!

Another very interesting axiom may be the "P not equal to NP" conjecture. P and NP are names for classes of problems. An NP problem is one for which a proposed solution can be verified quickly. For example, for the problem "find the factors of 8,633," one can quickly verify the proposed solution "97 and 89" by multiplying those two numbers. (There is a technical definition of "quickly," but those details are not important here.) A P problem is one that can be solved quickly even without being given the solution. The question is—and no one knows the answer—can every NP problem be solved quickly? (Is there a quick way to find the factors of 8,633?) That is, is the class P the same as the class NP? This problem is one of the Clay Millennium Prize Problems for which a reward of $1 million is on offer.

Computer scientists widely believe that P is not equal to NP, but no proof is known. One could say that a lot of quasi-empirical evidence points to P not being equal to NP. Should P not equal to NP be adopted as an axiom, then? In effect, this is what the computer science community has done. Closely related to this issue is the security of certain cryptographic systems used throughout the world. The systems are believed to be invulnerable to being cracked, but no one can prove it.

Experimental Mathematics

ANOTHER AREA of similarity between mathematics and physics is experimental mathematics: the discovery of new mathematical results by looking at many examples using a computer. Whereas this approach is not as persuasive as a short proof, it can be more convincing than a long and extremely complicated proof, and for some purposes it is quite sufficient.

In the past, this approach was defended with great vigor by both George Pólya and Lakatos, believers in heuristic reasoning and in the quasi-empirical nature of mathematics. This methodology is also practiced and justified in Stephen Wolfram's *A New Kind of Science* (2002).

Extensive computer calculations can be extremely persuasive, but do they render proof unnecessary? Yes and no. In fact, they provide a different kind of evidence. In important situations, I would argue that both kinds of evidence are required, as proofs may be flawed, and conversely computer searches may have the bad luck to stop just before encountering a counterexample that disproves the conjectured result.

All these issues are intriguing but far from resolved. It is now 2006, 50 years after this magazine published its article on Gödel's proof, and we still do not know how serious incompleteness is. We do not know if incompleteness is telling us that mathematics should be done somewhat differently. Maybe 50 years from now we will know the answer. SA

MORE TO EXPLORE

For a chapter on Leibniz, see **Men of Mathematics.** E. T. Bell. Reissue. Touchstone, 1986.

For more on a quasi-empirical view of math, see **New Directions in the Philosophy of Mathematics.** Edited by Thomas Tymoczko. Princeton University Press, 1998.

Gödel's Proof. Revised edition. E. Nagel, J. R. Newman and D. R. Hofstadter. New York University Press, 2002.

Mathematics by Experiment: Plausible Reasoning in the 21st Century. J. Borwein and D. Bailey. A. K. Peters, 2004.

For Gödel as a philosopher and the Gödel-Leibniz connection, see **Incompleteness: The Proof and Paradox of Kurt Gödel.** Rebecca Goldstein. W. W. Norton, 2005.

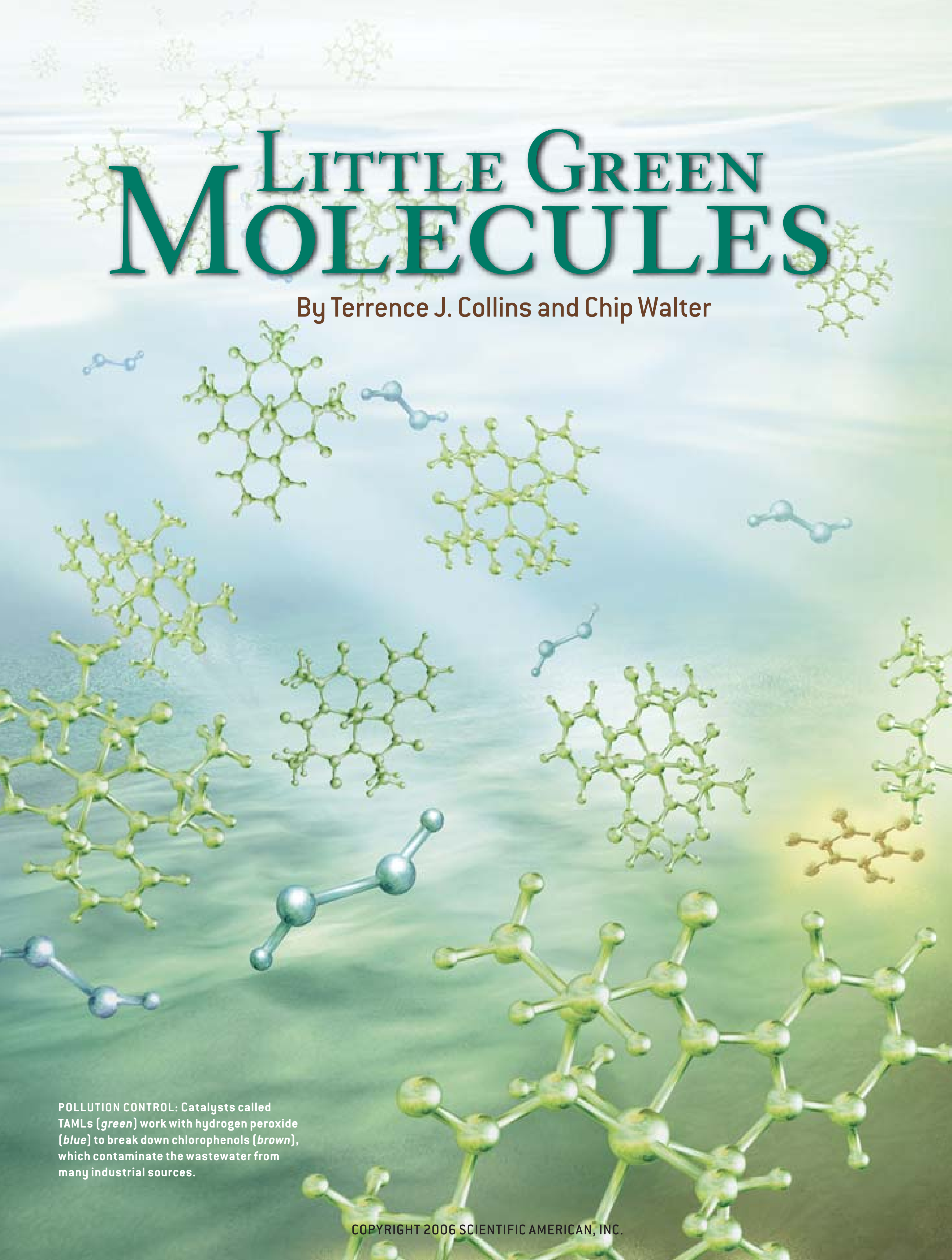
Meta Math!: The Quest for Omega. Gregory Chaitin. Pantheon Books, 2005.

Short biographies of mathematicians can be found at **www-history.mcs.st-andrews.ac.uk/BiogIndex.html**

Gregory Chaitin's home page is **www.umcs.maine.edu/~chaitin/**

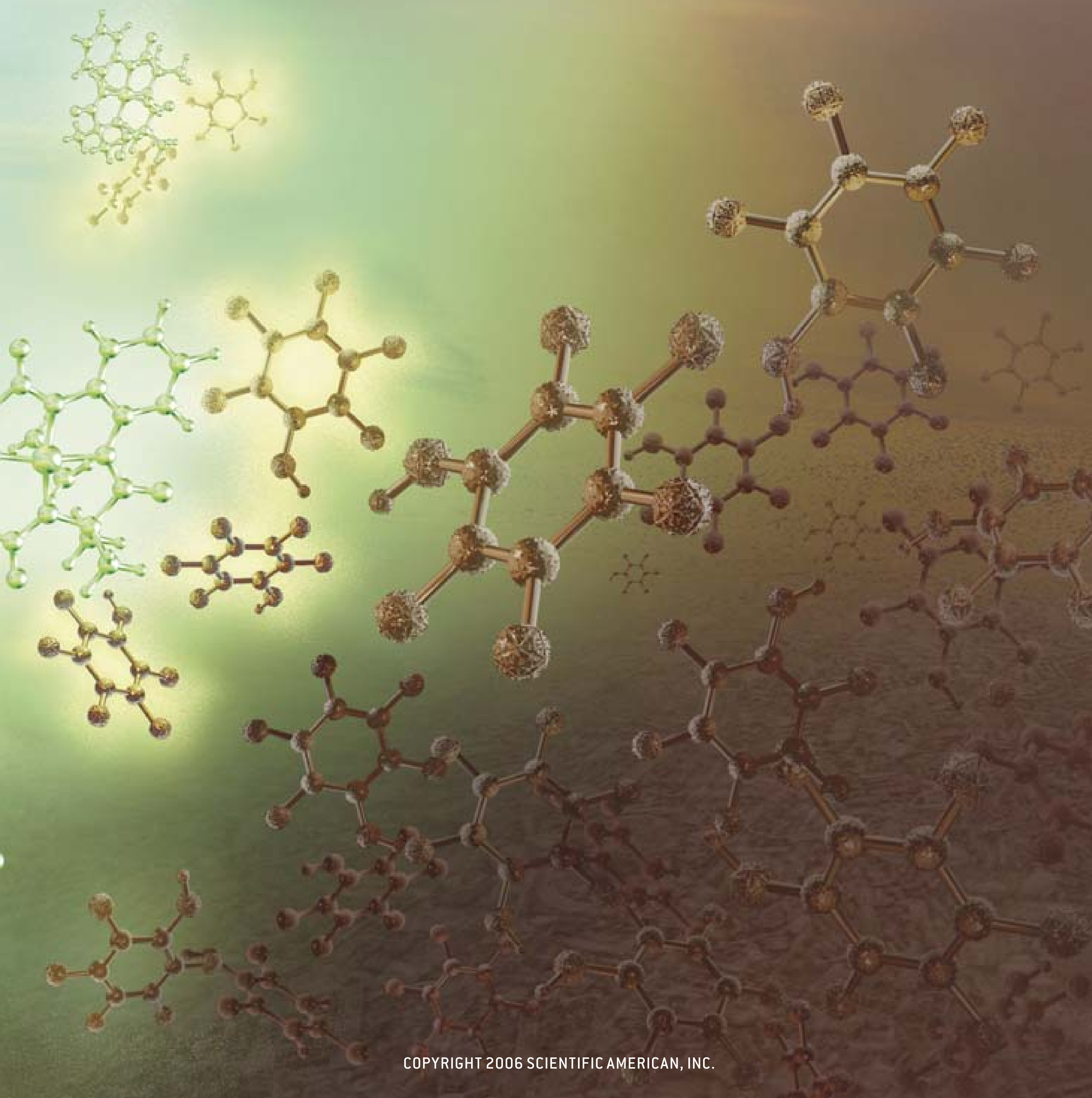
Little Green Molecules

By Terrence J. Collins and Chip Walter



POLLUTION CONTROL: Catalysts called TAMLs (*green*) work with hydrogen peroxide (*blue*) to break down chlorophenols (*brown*), which contaminate the wastewater from many industrial sources.

Chemists have invented a new class of catalysts that can destroy some of the worst pollutants before they get into the environment



The fish that live in the Anacostia River, which flows through the heart of Washington, D.C., are not enjoying its waters very much. The Anacostia is contaminated with the molecular remnants of dyes, plastics, asphalt and pesticides. Recent tests have shown that up to 68 percent of the river's brown bullhead catfish suffer from liver cancer. Wildlife officials recommend that anyone who catches the river's fish toss them back uneaten, and swimming has been banned.

The Anacostia is just one of dozens of severely polluted rivers in the U.S. The textile industry alone discharges 53 billion gallons of wastewater—loaded with reactive dyes and other hazardous chemicals—into America's rivers and streams every year. New classes of pollutants are turning up in the nation's drinking water: traces of drugs, pesticides, cosmetics and even birth-control hormones [*see illustration on opposite page*]. The amounts are often infinitesimal, measured in parts per billion or trillion (a part per billion is roughly equivalent to one grain of salt dissolved in a swimming pool), but scientists suspect that even tiny quantities of some pollutants can disrupt the developmental biochemistry that determines human behavior, intelligence, immunity and reproduction.

Fortunately, help is on the way. Over the past decade researchers in the emerging field of green chemistry have begun to design the hazards out of chemical products and processes. These scientists have formulated safer substitutes for harmful paints and plastics and devised new manufacturing techniques that reduce the introduction of pollutants into the environment. As outlined by the Green Chemistry Institute of the American Chemical Society, the first principle of this community is: "It is better to prevent waste than to treat or clean up waste after it has been created." As part of this effort, however, researchers have also made discoveries that promise cost-effective methods for purging many persistent pollutants from wastewater.

In one example of this work, investigators at Carnegie Mellon University's Institute for Green Oxidation Chemistry (one of us, Collins, is the institute's director) have developed a group of designer catalyst molecules called TAML—tetra-amido macrocyclic ligand—activators that work with hydrogen peroxide and other oxidants to break down a wide variety of stubborn pollutants. TAMLs accomplish this task by mimicking the enzymes in our bodies that have evolved over time to combat toxic compounds. In laboratory and real-world trials, TAMLs have proved they can destroy dangerous pesticides, dyes and other contaminants, greatly decrease the smells and color from the wastewater discharged by paper mills, and kill bacterial spores similar to those of the deadly anthrax strain. If broadly adopted, TAMLs could save millions of dollars in cleanup costs. Moreover, this research demonstrates that green chemistry can lessen some of the environmental damage caused by traditional chemistry.

Green chemistry can lessen some of the environmental damage caused by traditional chemistry.

Overview/*Catalysts for Cleaning*

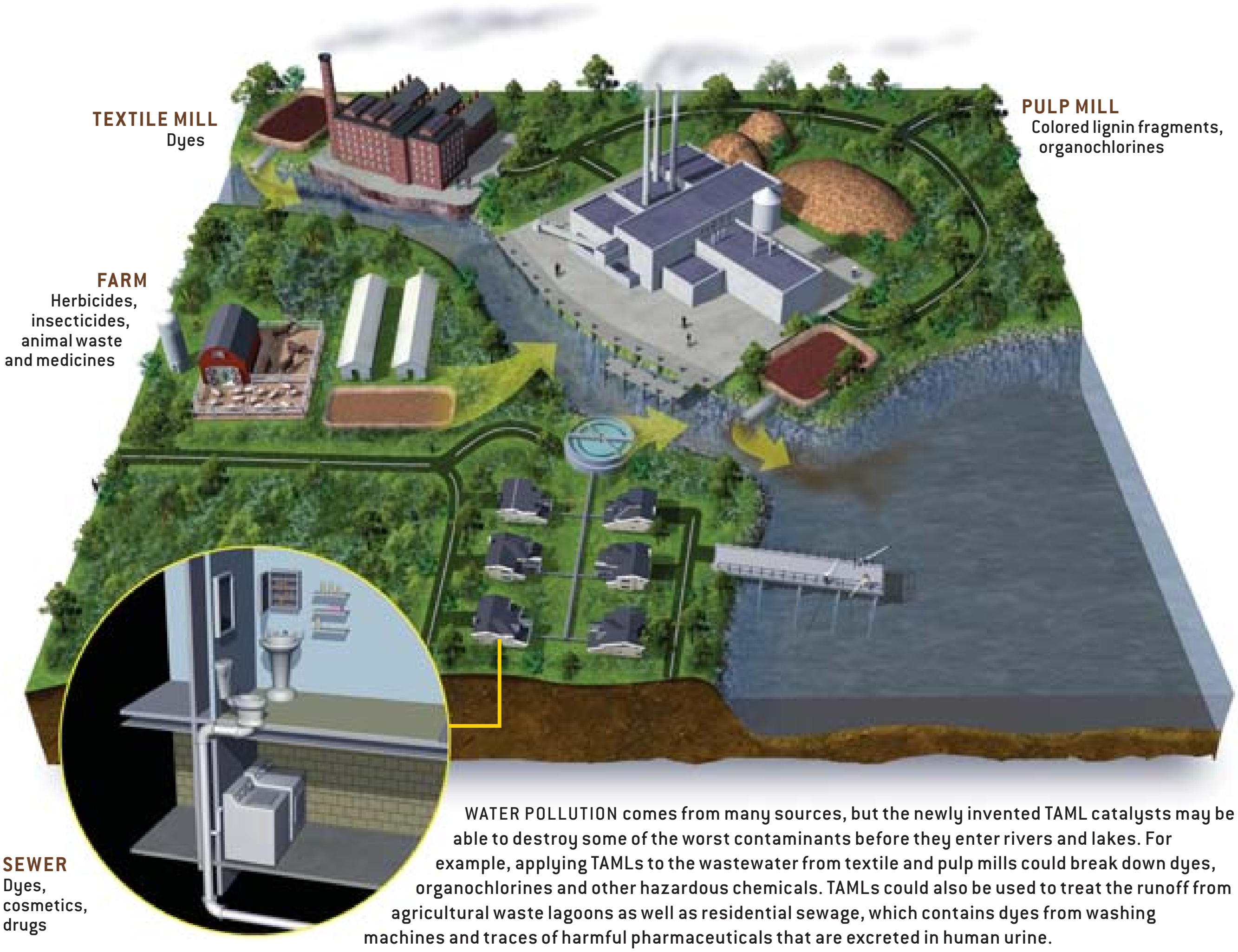
- Many pollutants released into waterways, such as dyes and pesticides, have become so omnipresent that they pose a serious threat to human health.
- Chemists have recently created enzymelike catalysts called tetra-amido macrocyclic ligand activators (TAMLs, for short) that can destroy stubborn pollutants by accelerating cleansing reactions with hydrogen peroxide.
- When applied to the wastewater from pulp mills, TAMLs have reduced staining and hazardous chemicals. The catalysts may also someday be used to disinfect drinking water and clean up contamination from bioterror attacks.

The Need to Be Green

A FUNDAMENTAL CAUSE of our escalating environmental plight is that people perform chemistry in different ways than Mother Nature does. For eons, biochemical processes have evolved by drawing primarily on elements that are abundant and close at hand—such as carbon, hydrogen, oxygen, nitrogen, sulfur, calcium and iron—to create everything from paramecia to redwoods, clown fish to humans. Our industries, in contrast, gather elements from nearly every corner of the planet and distribute them in ways natural processes never could. Lead, for example, used to be found mostly in deposits so isolated and remote that nature never folded it into living organisms. But now lead is everywhere, primarily because our paints, cars and computers have spread it around. When it finds its way into children, even at minuscule doses, it is severely toxic. The same can be said for arsenic, cadmium, mercury, uranium and plutonium. These elements are persistent pollutants—they do not degrade in animal bodies or in the surrounding environment—so there is a pressing need to find safer alternatives.

Some of the new synthetic molecules in medicines, plastics and pesticides are so different from the products of natural chemistry that it is as though they dropped in from an alien world. Many of these molecules do not degrade easily, and even some biodegradable compounds have become omnipresent because we use them so copiously. Recent research indicates that some of these substances can interfere with the normal expression of genes involved in the development of the male reproductive system. Scientists have known for several years that prenatal exposure to phthalates, compounds used in plastics and beauty products, can alter the reproductive tract of newborn male rodents; in

JEAN-FRANCOIS PODEVIN (*preceding pages*)



WATER POLLUTION comes from many sources, but the newly invented TAML catalysts may be able to destroy some of the worst contaminants before they enter rivers and lakes. For example, applying TAMLs to the wastewater from textile and pulp mills could break down dyes, organochlorines and other hazardous chemicals. TAMLs could also be used to treat the runoff from agricultural waste lagoons as well as residential sewage, which contains dyes from washing machines and traces of harmful pharmaceuticals that are excreted in human urine.

DON FOLEY

2005 Shanna H. Swan of the University of Rochester School of Medicine and Dentistry reported similar effects in male infants. Another study headed by Swan found that men with low sperm counts living in a rural farming area of Missouri had elevated levels of herbicides (such as alachlor and atrazine) in their urine. Starting from our factories, farms and sewers, persistent pollutants can journey intact by air, water and up the food chain, often right back to us.

To confront this challenge, green chemists at universities and companies are investigating the feasibility of replacing some of the most toxic products and manufacturing processes with more environment-friendly alternatives [*see box on page 88*]. The work of Collins's team at Carnegie Mellon traces its origins back to the 1980s, when public health concerns about chlorine were intensifying. Chlorine was then, and still is, often used for large-scale cleaning and disinfection in manufacturing, as well as for the treatment of drinking water. Although chlorine treatment is inexpensive and effective, it can create some ugly pollutants. The bleaching of wood pulp with elemental chlorine in paper mills had been a major source of cancer-causing dioxins until the Environmental Protection Agency banned the process in 2001. (Most mills now bleach wood pulp with chlorine dioxide, which reduces the production of dioxins but does not eliminate it.) By-products created by the chlorination of drinking water have also been linked to certain cancers. Chlorine in its common natural form—chloride ions or salts dissolved in water—is not toxic, but when elemental chlorine reacts with other molecules it can generate compounds that can warp the biochemistry of living animals. Dioxins, for instance, disrupt cellular development by interfering with a receptor system that regulates the production of critical proteins.

Rather than relying on chlorine, we wondered if we could put nature's own cleansing agents—hydrogen peroxide and oxygen—to the work of purifying water and reducing industrial waste. These cleansers can safely and powerfully obliterate many pollutants, but in nature the process usually requires an enzyme—a biochemical catalyst that vastly increases the rate of the reaction. Whether natural or man-made, catalysts act as old-fashioned matchmakers, except that rather than bringing two people together they unite specific molecules, enabling and accelerating the chemistry among them. Some natural catalysts can boost chemical reaction rates a billionfold. If not for an enzyme called ptyalin, found in our saliva, it would take several weeks for our bodies to break down pasta into its constituent sugars. Without enzymes, biochemistry would move at a numbingly slow pace, and life as we know it would not exist.

In nature, enzymes called peroxidases catalyze reactions involving hydrogen peroxide, the familiar household chemical used to bleach hair and remove carpet stains. In forests, fungi on rotting trees use peroxidases to marshal hydrogen peroxide to break down the lignin polymers in the wood, splitting the large molecules into smaller ones that the fungi can eat. Another family of enzymes, the cytochrome p450s, catalyzes reactions involving oxygen (also called oxidation reactions). Cytochrome p450s in our livers, for example, use oxygen to efficiently destroy many toxic molecules we inhale or ingest.

For decades, chemists have been struggling to build small synthetic molecules that could emulate these enormous enzymes. If scientists could create designer molecules with such strong catalytic abilities, they could replace the chlorine- and metal-based oxidation technologies that produce so many pollutants. In the early 1980s, however, no one was having much luck developing test-tube versions of the enzymes. Over billions of years of evolution, nature had choreographed some wonderfully elegant and extremely complex catalytic dances, making our efforts in the laboratory look clunky. Yet we knew that we could not achieve our goal of reducing pollution unless we found a way to mimic this molecular dance.

Catalytic Converters

CREATING SYNTHETIC enzymes also meant assembling molecules that would be robust enough to resist the destructive reactions they were catalyzing. Any chemistry involving oxygen can be destructive because the bonds it makes with other elements (especially hydrogen) are so strong. And because each molecule of hydrogen peroxide (H_2O_2) is halfway between water (H_2O) and molecular oxygen (O_2), this compound is also strongly oxidizing. In water, hydrogen peroxide often produces a kind of liquid fire that demolishes the organic (carbon-containing) molecules around it. A lesson from the enzymes was that a working catalyst would probably need to have an iron atom placed inside a molecular matrix of organic groups. So we had to toughen the molecular architecture of such groups to ensure they could survive the liquid fire that would result from the activation of hydrogen peroxide.

Borrowing further from nature's design, we eventually solved this problem by creating a catalyst in which four nitrogen atoms are placed in a square with a single iron atom anchored in the middle [*see box on opposite page*]. The nitrogen atoms are connected to the much larger iron atom by covalent bonds, meaning that they share pairs of electrons; in this kind of structure, the smaller atoms and attached groups surrounding the central metal atom are called ligands. Next we linked the ligands to form a big outer ring called a macrocycle. Over time we learned how to make the ligands and linking systems tough enough to endure the violent reactions that the TAMLs trigger. In effect, the ligands we invented became a kind of firewall that resisted the liquid fire. The longer it resisted, the more useful the catalyst. Of course, we did not want to create an indestructible catalyst, which could end up in effluent streams and perhaps produce a pollution problem of its own. All our existing Fe-TAML catalysts (TAMLs with iron as the central metal atom) decompose on timescales ranging from minutes to hours.

> Whether natural or man-made, catalysts act as **old-fashioned matchmakers.**

Building the ligand firewalls was not easy. It required developing a painstaking four-step design process in which we first imagined and then synthesized ligand constructions that we hoped would keep the firewall in place. Second, we subjected the catalyst to oxidative stress until the firewall disintegrated. Third, we looked for the precise location where the breakdown began. (We found that ligand degradations always start at the most vulnerable site.) And in the final step, once we had pinpointed the weakest link, we replaced it with groups of atoms we believed would hold up longer. Then we started the whole design cycle again.

After 15 years, we finally created our first working TAML. We knew we had succeeded one morning when Colin Horwitz, a research professor at our institute, showed off the results of a bleaching experiment that featured our most advanced design at the time. We looked at the results, and there it was: every time Horwitz squirted dark dye into a solution containing the TAML catalyst and hydrogen peroxide, the solution quickly turned colorless. We now knew that our firewalls were finally holding up long enough to allow the TAMLs to do their job. The molecules were acting like enzymes, and yet they were much, much smaller: the molecular weight of a TAML is about 500 daltons (a dalton is equal to one twelfth the mass of carbon 12, the most abundant isotope of carbon), whereas the weight of horseradish peroxidase, a relatively small enzyme, is about 40,000 daltons. The diminutive TAML activators are easier and cheaper to make, and much more versatile in their reactivity, than their natural counterparts.

Since then, we have evolved more than 20 different TAML activators by reapplying the same four-step design process that enabled us to create the first

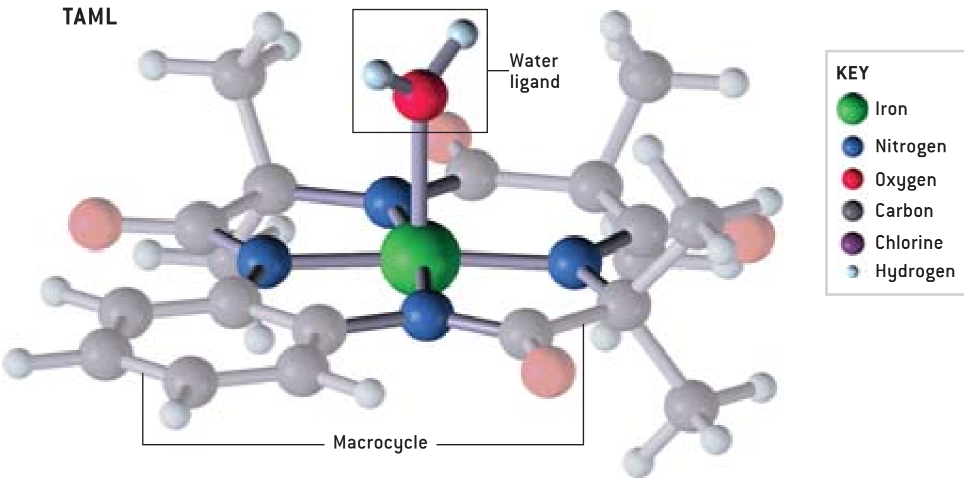
THE AUTHORS

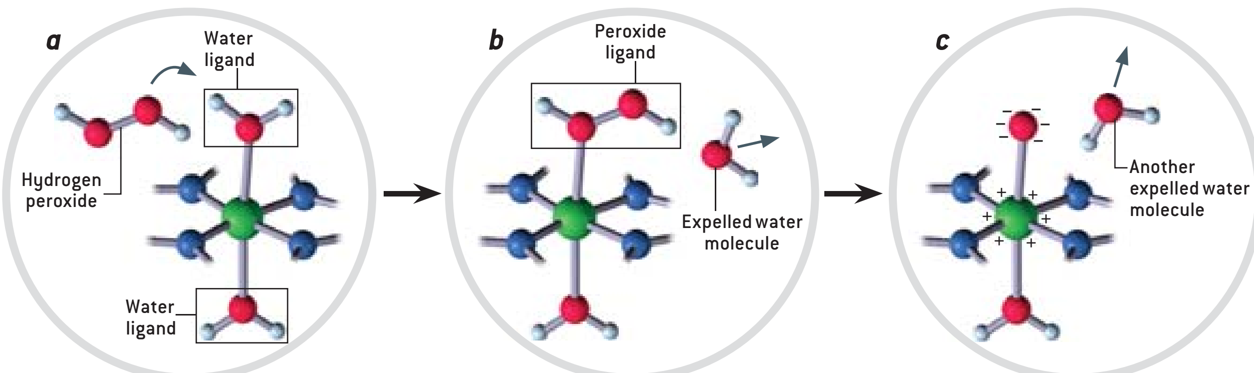
TERRENCE J. COLLINS and *CHIP WALTER* have worked together to educate the public about the challenges and possibilities of green chemistry. Collins is Thomas Lord Professor of Chemistry at Carnegie Mellon University, where he directs the Institute for Green Oxidation Chemistry. He is also an honorary professor at the University of Auckland in New Zealand. Walter is a science journalist and author of *Space Age* and *I'm Working on That* (with William Shatner). He teaches science writing at Carnegie Mellon and is a vice president of communications at the University of Pittsburgh Medical Center.

A MOLECULAR CLEANING MACHINE

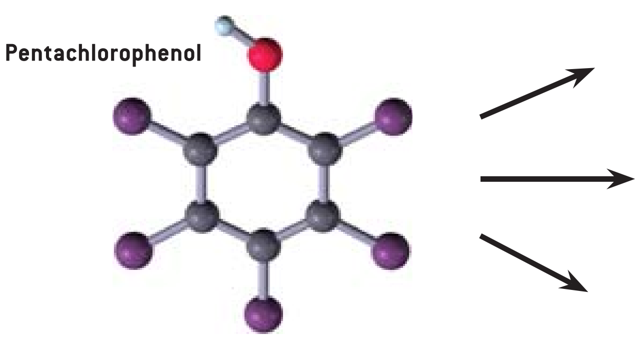
Chemists designed TAMLs to emulate the natural enzymes that catalyze reactions involving hydrogen peroxide. TAMLs, though, are hundreds of times smaller than enzymes, so they are easier and cheaper to manufacture.

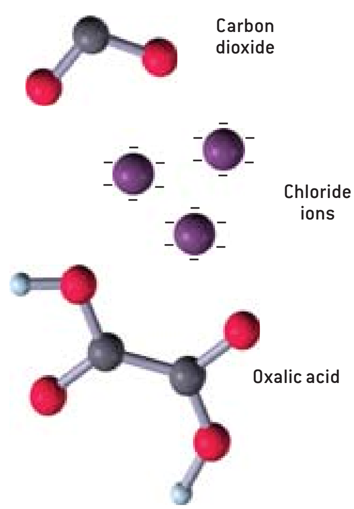
At the center of each TAML is an iron atom bonded to four nitrogen atoms; at the edge are carbon rings linked to form a big outer ring called a macrocycle. This linking system acts as a firewall, enabling the molecule to endure the violent reactions it triggers. In its solid state the TAML also has one water molecule (H_2O) attached to the iron atom. (The attached groups are called ligands.)





When a TAML dissolves in water, another molecule of H_2O connects to the catalyst (*a*). If hydrogen peroxide (H_2O_2) is also in the solution, it can replace one of the water ligands, which are loosely attached and easily expelled (*b*). The peroxide ligand then discards both its hydrogen atoms and one oxygen atom in the form of a water molecule, leaving one oxygen atom attached to the iron (*c*). The oxygen pulls electrons farther away from the iron atom, turning the TAML into a reactive intermediate.





When in solution with molecules of pentachlorophenol—a toxic chemical used in wood treatment—TAMLs and hydrogen peroxide break down the pollutant to nontoxic compounds and ions. The strong positive charge of the iron atom in the reactive intermediate enables the molecule to destroy pollutants, although scientists have not yet determined the details of the process.

MELISSA THOMAS; SOURCE: ARANI CHANDA *Institute for Green Oxidation Chemistry*

CHEMISTRY GOES GREEN

The invention of TAML catalysts is just one of the many achievements of green chemistry, which strives to develop products and processes that reduce or eliminate the use and generation of hazardous substances. Some other accomplishments are listed below.

PROJECT	PARTICIPANTS	STATUS
Using plant sugars to create polylactic acids (PLAs), a family of biodegradable polymers that could replace many traditional petroleum-derived plastics	Patrick Gruber, Randy L. Howard, Jeffrey J. Kolstad, Chris M. Ryan and Richard C. Bopp, NatureWorks LLC (a subsidiary of Cargill)	NatureWorks has built a factory in Nebraska to manufacture PLA pellets, which are used to make water bottles, packaging materials and other products 
Discovering synthesis reactions that allow manufacturers to substitute water for many common organic solvents, some of which can cause cancer	Chao-Jun Li, McGill University	Pharmaceutical and commodity chemical companies are investigating the process
Developing metathesis chemistry, a method of organic synthesis that can produce drugs, plastics and other chemicals more efficiently and with less waste	Robert H. Grubbs, California Institute of Technology; Richard R. Schrock, Massachusetts Institute of Technology; Yves Chauvin, French Petroleum Institute	Widely applied in the chemical, biotechnology and food industries, this research was awarded the 2005 Nobel Prize in Chemistry 
Replacing toxic petroleum-based solvents with supercritical carbon dioxide, a high-temperature, high-pressure fluid that has the properties of both a liquid and a gas	Martyn Poliakoff, Michael George and Steve Howdle, University of Nottingham, England	Thomas Swan & Co., a British manufacturer of specialty chemicals, has built a plant that uses supercritical fluids
Inventing a new method for producing sertraline, the key ingredient in the antidepressant Zoloft	James Spavins, Geraldine Taber, Juan Colberg and David Pfisterer, Pfizer	The process has reduced pollution, energy and water use while improving worker safety and product yield

working model. Each TAML has its own reaction rate and lifetime, allowing us to tailor the catalysts to match the tasks we want them to perform. Most of the catalysts incorporate elements such as carbon, hydrogen, oxygen, nitrogen and iron, all chosen for their low toxicity. We call some of the molecules "hunter TAMLs" because they are designed to seek out and lock onto specific pollutants or pathogens, in much the same way that a magnetized mine seeks out the metal hull of a ship. Other TAMLs act as blowtorches that aggressively burn most of the oxidizable chemicals with which they come into contact. Still others are less aggressive and more selective, so that they will, for example, attack only certain parts of molecules or attack only the more easily oxidized molecules in a group. We expect to adapt TAMLs to advance green chemistry for decades to come. Although more toxicity testing must be done, the results so far indicate that TAMLs break down pollutants to their nontoxic constituents, leaving no detectable contamination behind. We now have more than 90 international patents on TAML activators, with more in the pipeline, and we also have several commercial licenses.

Interestingly, we still do not know all the details of how the TAMLs work, but recent studies have provided deep insights into the key reactions. In their solid state, Fe-TAMLs generally have one water molecule attached as a ligand to the iron atom, oriented perpendicularly from the four nitrogen ligands; when put in solution, another water molecule connects to the opposite side of the iron atom. These water ligands are very loosely attached—if hydrogen peroxide is also in the solution, a molecule of it easily replaces one of the water molecules. The peroxide ligand swiftly reconstitutes itself, expelling both its hydrogen atoms and one oxygen atom (which escape as H_2O, a water molecule) and leaving one oxygen atom attached to the iron at the center of the Fe-TAML, which is now called the reactive intermediate (RI).

Oxygen is much more electronegative than iron, which means that its nucleus pulls most of the electrons in the complex bond toward itself and away from the iron nucleus. This effect increases the positive charge of the iron at the center of the TAML, making the RI reactive enough to extract electrons from oxidizable molecules in the solution. We have not yet determined how the RI breaks the chemical bonds of its targets, but current investigations may soon reveal the answer. We do know, however, that we can adjust the strength of the TAMLs by changing the atoms at the head and tail of the molecule; putting highly electronegative elements at those locations draws even more negative charge away from the iron and makes the RI more aggressive.

SOURCE: TERRENCE J. COLLINS *Institute for Green Oxidation Chemistry* AND PAUL ANASTAS *Green Chemistry Institute*; NATUREWORKS LLC (*bottles*); AP PHOTO (*Nobel medal*)

Industrial Strength

BUILDING TAMLS in the laboratory is one thing; scaling them up for commercial use is another. So far the lab tests and field trials have been promising. Tests funded by the National Science Foundation, for example, demonstrated that Fe-TAMLs plus peroxides could clean up the contamination from a bioterror attack. We found that when we combined one TAML with tertiary butyl hydroperoxide—a variation of hydrogen peroxide that replaces one of the hydrogen atoms with a carbon atom and three methyl (CH_3) groups—the resulting solution could deactivate 99.99999 percent of the spores of *Bacillus atrophaeus,* a bacterial species very similar to anthrax, in 15 minutes. In another important potential application, we hope to use Fe-TAMLs and hydrogen peroxide to someday create an inexpensive disinfectant to tackle the infectious waterborne microbes that account for so much death and disease worldwide.

Building TAMLs in the laboratory is one thing; **scaling them up** for commercial use is another.

In three field trials, we explored how well TAMLs can alleviate the pollution created when paper is manufactured. Every year the paper and wood pulp industry produces more than 100 million metric tons of bleached pulp, which is turned into white paper. Besides generating dioxins, chlorophenols and other hazardous organochlorines, many pulp mills discharge a coffee-colored effluent that stains streams and rivers and blocks light from penetrating the water. The reduction of light interferes with photosynthesis, which in turn affects organisms that depend on plants for food. The sources of the staining are large colored fragments of lignin, the polymer that binds the cellulose fibers in wood. Bleaching with chlorine dioxide removes the lignin from the cellulose; the smaller lignin fragments are digested by bacteria and other organisms in treatment pools, but the larger pieces are too big to be eaten, so they end up in rivers and lakes.

We have tested the effectiveness of Fe-TAMLs at decolorizing these fragments at two pulp mills in the U.S. and one in New Zealand. In New Zealand we combined Fe-TAMLs and peroxide with 50,000 liters of effluent water. In the U.S. we directly injected Fe-TAMLs into a pulp-treatment tower or an exit pipe over the course of several days to bleach the wastewater. Overall, the Fe-TAMLs reduced the staining of the water by up to 78 percent and eliminated 29 percent of the organochlorines.

The development of other TAML applications also looks exciting. Eric Geiger of Urethane Soy Systems, a company based in Volga, S.D., has found that Fe-TAMLs do an excellent job processing soybean oil into useful polymers that display physical properties equal to, if not better than, those of current polyurethane products. TAMLs may even find their way into washing machines: in another series of tests, we found that a tiny quantity of catalyst in certain household laundry products eliminated the need to separate white and colored clothing. TAMLs can prevent staining by attacking dyes after they detach from one fabric but before they attach to another. We are also working on a new family of TAMLs that can break the very stable molecular bonds that allow drugs and agricultural chemicals to pass intact into drinking water.

Despite the success of these trials, we have not resolved all the questions about TAML activators. More testing on industrial scales remains to be done, and it is important to ensure that TAMLs do not create some form of pollution we have not yet observed. Too often chemical technologies have seemed completely benign when first commercialized, and the devastating negative consequences did not become clear until decades later. We want to do everything in our power to avoid such surprises with TAMLs.

Cost is also an issue. Although TAMLs promise to be competitive in most applications, large corporations are deeply invested in the chemical processes they currently use. Shifting to new systems and techniques, even if they work, usually requires significant investments. One great advantage of TAML technology, though, is that it does not require major retooling. What is more, TAMLs may ultimately save companies money by offering a cost-effective way to meet increasingly stringent environmental laws in the U.S., Europe and elsewhere.

The advances of green chemistry to date represent only a few interim steps on the road to dealing with the many environmental challenges of the 21st century. The deeper question is, Are we going to practice acute care or preventive medicine? Right now most chemists are still trained to create elegantly structured compounds that solve the specific problem for which they have been engineered, without regard to their broader impact. We are in effect performing global-scale experiments on our ecosystems and ourselves, and when these experiments fail the cost can be catastrophic. New green chemical techniques offer an alternative. The Industrial Revolution has unfolded, for the most part, without design or forethought. Perhaps now we can take some creative steps to reverse that trend and help make a world, and a future, that we can live with. SA

MORE TO EXPLORE

Toward Sustainable Chemistry. Terrence J. Collins in *Science,* Vol. 291, No. 5501, pages 48–49; January 5, 2001.

Rapid Total Destruction of Chlorophenols by Activated Hydrogen Peroxide. Sayam Sen Gupta, Matthew Stadler, Christopher A. Noser, Anindya Ghosh, Bradley Steinhoff, Dieter Lenoir, Colin P. Horwitz, Karl-Werner Schramm and Terrence J. Collins in *Science,* Vol. 296, pages 326–328; April 12, 2002.

More information can be found online at **www.cmu.edu/greenchemistry** and **www.chemistry.org/portal/a/c/s/1/acsdisplay.html?DOC=greenchemistryinstitute\index.html**

The Elusive Goal of

MACHINE TRANSLATION

Statistical methods hold the promise of moving computerized translation out of the doldrums

By Gary Stix

> Natrium Nepal Asia legend: The lion, the sorceress, the evil spirit wardrobe "already lack" the evil spirit abstains the trilogy "rich in poetic and artistic flavor, also has not let" the Harley baud "the series novel have the infinite pleasure the undercurrent to be turbulent.

The preceding gibberish was brought to you by a Chinese-to-English translation carried out by Altavista's Babelfish, the popular Internet-based translator. In coherent English, from a bilingual page on the Web site of Taiwan's *China Post,* it reads:

> "The Chronicles of Narnia" doesn't come near the poetic vision of "The Lord of the Rings" trilogy, and it doesn't have the dark undercurrents that makes the "Harry Potter" series endlessly fascinating.

This passage illustrates that machine translation, or MT, as it is known, remains one of the more challenged subdisciplines of the blighted field of artificial intelligence. A proper name or a few well-crafted phrases suffice to throw the software off track. In the past few years, though, a new research approach has fueled a revival for machine translation: brute-force computing methods—which gauge the probability that a word or phrase in one language matches that in another—are at last bringing MT closer to human performance, in the estimation of developers of this software.

Tougher Than Chess

THE EVER INCREASING POWER of hardware and software algorithms today has propelled the computer past the chess grandmaster. (Recall that IBM's Deep Blue supercomputer triumphed over Garry Kasparov in 1997.) But on the whole, machine translation has experienced only halting progress in achieving humanlike capabilities in its more than 50-year history—and some critics would classify even that characterization as overly generous.

In 1954 IBM and Georgetown University demonstrated the translation of more than 60 sentences from Russian into English. The IBM press release, dated January 8, 1954, glowed: "Russian was translated into English by an electronic 'brain' today for the first time." The military defense community and computer scientists expected routine machine translation within five years, but it never materialized.

In 1966 the U.S. government–sponsored Automatic Language Processing Advisory Committee reported that humans could perform faster, more accurate translation at half the cost. "There is no immediate or predictable prospect of useful machine translation," its study concluded.

Funding dried up, and only modest advances came in subsequent decades. In the late 1960s the U.S. Air Force supplied support to a small company that created the machine translator called Systran—the Internet version of which provided the first paragraph of this article—to cope initially with voluminous demands to translate Russian documents into English.

Systran is based on rules about the source and target languages, as was IBM's original "brain" system, which relied on six rudimentary rules that govern syntax, semantics and the like. For example, the word "*o*" in Russian could be translated by an IBM 701 computer as either "about" or "of." If "*o*" followed the word "*nauka*" (science), it looked for the appropriate rule that told it to translate "*o*" as "of"—in other words, the "science of," not the "science about."

The Paris-based Systran company ranks as the biggest machine translation company in the world. Even with customers that include Google, Yahoo and Time Warner's AOL, its annual revenues were just $13 million for 2004—in an overall market for translations of all varieties that is estimated worldwide to total nearly $10 billion. "We're so small, and we're the largest," says Dimitris Sabatakakis, Systran's chairman and chief executive officer.

SLIM FILMS

No More Rules

FOR RULE-BASED SYSTEMS, language experts and linguists in specific languages have to painstakingly craft large lexicons and rules related to grammar, syntax and semantics to generate text in a target language. Commercial systems contain tens of thousands of grammar rules for a corpus that is made up of hundreds of thousands of words.

Beginning in the late 1980s, IBM created a system for translating French into English called Candide that required knowledge of neither grammar nor syntax. It eschewed rules in favor of taking substantial bodies of already translated text, matching words between the two languages (more recent systems use whole phrases) and finally deriving probabilities—based on Bayes's theorem—to estimate whether an English word was a correct translation from the French.

Another analysis that relied solely on large English texts assessed whether the word translated into English fit in grammatically with surrounding words. The word or phrase in the target language accorded the highest probability could then be used to "decode" future texts—and multiple words could be linked to build entire documents. If the statistics showed that the word "*pouderie*" usually equated to "blowing snow," that, in principle, was all that was needed.

IBM eventually dropped its effort. At the end of the 1990s it could take an entire day for a machine translation of a single page. But then things began to stir. The Internet produced a rapid growth in the number of large, bilingual bodies of text. The Web also created demand for translation that could never be met by humans.

In 1999 the National Science Foundation held a workshop at Johns Hopkins University to construct a software tool kit that could be readily disseminated to the scientific community, an action that drew attention and spurred new activity. In 2002 one of the workshop organizers, Kevin Knight of the University of Southern California, and Daniel Marcu, also at U.S.C., founded Language Weaver, the only statistical machine-translation company. It now claims to be capable of translating at least 5,000 words a minute back and forth between English and Arabic, Farsi, French, Chinese and Spanish.

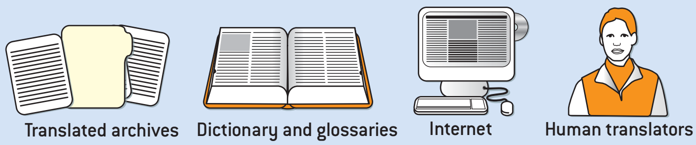
Google Is a Winner

ANOTHER ALUMNUS of both the workshop and U.S.C., Franz Och, was hired by Google. Last summer the still experimental Google system engineered by Och bested competitors such as IBM to win every category in a competition organized by the National Institute of Standards and Technology to translate 100 newswire documents from Arabic

STATISTICAL MACHINE TRANSLATION

INPUTTING ALREADY-TRANSLATED TEXTS

Existing translated texts from various sources form the foundation of the automated translations.



Statistical methods have proved to be more effective than other types of automated machine translations based on rules crafted by human translators.

The new methods take advantage of the brute-force calculating power of machines to crunch through existing translated texts to determine the probability that a word or a phrase in one language matches that in another.

PREPROCESSING

The texts are scanned, aligned and formatted.

`Que hambre tengo yo.`

↓ Preprocessor ↓

`que hambre tengo yo`

PHRASE MATCHING IN TRANSLATED TEXTS

A translation model picks out two- or three-word phrases from the source language (in this case, Spanish) that match the target language (English).

Source language: *Spanish*

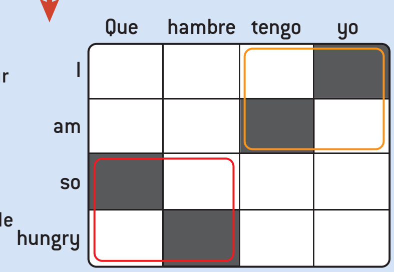
```
Este guiso tradicional se
ennoblece con el bogavante,
la viera y el rodaballo.
Que hambre tengo yo.
```

Target language: *English*

```
This traditional stew is
refined with scallops,
lobster and turbot.
I am so hungry
```

TRANSLATION MODEL

By using statistics to measure how often and where words occur in a given phrase in both languages, the model derives a template for word reordering. It also takes advantage of other techniques, such as reducing multiple Spanish words to a single translated word (*not shown*).



LANGUAGE MODEL

Working from its own statistical analyses of English-only texts, a language model attempts to predict the most likely word and phrase ordering for the already-translated text. Greater frequency of a phrase's occurrence increases the probability that it is correct.

```
I am so baffled by Modern and Postmodern art.
```

```
The boy is so thirsty and the mother so sad.
```

```
I am so pleased to meet you.
```

```
"I am so hungry to see everything, and to know everything," she said to herself.
```

```
He cried, "I am so hungry! Will you give me food?"
```

```
What strength have I that I should endure?
```

```
I have so many people to thank.
```

I am so > Have I that

I am so > I have so

So thirsty > Thirsty

Am so hungry > What hunger have

DECODER

When a new sentence gets inputted—one that can differ slightly or substantively from the text already processed (only *sed* substitutes for *hambre* here)—the decoder develops several hypothetical translations and picks the one with the highest probability.

INPUT TEXT

Que sed tengo yo.

DECODER

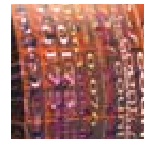
I am so thirsty	*P* = 0.13
What thirst have I	*P* = 0.09
Have I what thirst	*P* = 0.07
Thirsty I am so	*P* = 0.00

TRANSLATION

I am so thirsty.

LUCY READING-IKKANDA; SOURCES: KEVIN KNIGHT *Language Weaver* AND PHILIPP KOEHN *University of Edinburgh*

Can a machine translator provide more than "gisting," a rough idea of the contents of a foreign-language text?



or Chinese into English. Och has mentioned that feeding the machine-translation software with text that equated to one million books was key to performance improvements. He contrasted Google's current Chinese-to-English MT system (Systran) with the experimental statistical one crafted by him and his co-workers:

> Google/Systran: "Doctor indicates, the bright kernel prearranges recuperates the about one month."
>
> Google Research: "Doctors said Akihito is scheduled to rest for about a month."

The buzz about statistical machine translation has put Systran on the defensive. "You need rules when learning a foreign language," Sabatakakis comments. "You don't learn a language with statistical methods." Systran uses statistical techniques when creating systems in very narrow domains, such as translating patent documents. But the current embrace of statistical methods is somewhat of a marketing technique, he says. The company still employs 50 people in research and development, among them linguists. "The major difference between Systran and Google is that Google claims that it doesn't need native Chinese people to develop Chinese [applications] because of the magic and beauty of this stuff," Sabatakakis says, adding, "If we don't have some Chinese guys, our system may contain enormous mistakes."

The distinction between the two camps has begun to blur a little as statistical MT researchers have started to incorporate techniques that account for the syntactical structure of a sentence. These methods forgo the intervention of a human linguist: a syntactic model might estimate the chance that an English adjective-noun phrase gets reordered after translation into French. Knight of Language Weaver says that relying on phrases instead of single words allows the statistics to deal with semantics as well, avoiding, for instance, having his surname translated as "*Caballero*."

Microsoft Research has a substantial natural-language group, which for the past six years has also worked on MT. The group first focused on rule-based systems. But it is increasingly incorporating statistical techniques. Recently Microsoft used primarily statistical approaches when translating its online customer-support Web sites into 12 new languages, including Russian, Arabic and Chinese. The text does not get edited afterward. "Some of it is admittedly pretty rough; other parts of it are quite good," notes Steve Richardson, a senior researcher in the natural-language processing unit. "The quality of the more statistical approaches is comparable to or beginning to exceed that of the rule-based systems that we used before."

Getting the Gist

ALL THESE TECHNIQUES, however, raise the question of whether the machine-translation equivalent of a Deep Blue, the IBM chess computer, will ever beat humans at their own game. Can a machine provide more than mere "gisting," a rough idea of the contents of a foreign-language text? Kevin Hendzel, a spokesman for the American Translators Association, says that the current optimism only promulgates decades' worth of overhyped claims—FAHQT, the idea of "fully automatic high-quality translation," for instance. Gisting can help sort through massive amounts of foreign-language texts as long as it is understood to be inherently unreliable, he notes. Even a rough translation has its perils. He cites one Arabic-to-English translation that mentioned two sides "going at" each other, a fragment that caught the attention of security officials. The reference turned out to be for a soccer game, not a terrorist attack or imminent battle.

Keith Devlin, executive director of Stanford University's Center for the Study of Language and Information, remarks that machine-based systems will never equal the human linguist. "The use of statistical techniques, coupled with fast processors and large, fast memory, will certainly mean we will see better and better translation systems that work tolerably well in many situations," Devlin says, "but fluent translation, as a human expert can do, is, in my view, not achievable."

Knight, the pioneer in statistical translation, disagrees and points to the progress achieved during this decade. He foresees no limit to the technology, which will ultimately achieve human-level translations for everything except possibly poetry. He has shown blind examples of human translations alongside those from a machine, and audiences have confused the two. "Let's not kid ourselves—there are lots of mistakes in human-level translations. The bar is not as high as you would imagine," he says. To prove that this round of translation tools is more than the perennial sales pitch, the statistics jocks who now lead the field must demonstrate that this time FAHQT is real. Only then will the technology go beyond, as Microsoft's Richardson puts it, mere "MT promises." SA

MORE TO EXPLORE

The History of Machine Translation in a Nutshell. Online at John Hutchins's Web site: **http://ourworld.compuserve.com/homepages/WJHutchins/nutshell.htm**

A Statistical MT Tutorial Workbook. Kevin Knight. Online at **www.isi.edu/natural-language/mt/wkbk.rtf**

The Candide System for Machine Translation. Adam L. Berger et al. Online at **http://acl.ldc.upenn.edu/H/H94/H94-1028.pdf**

WORKINGKNOWLEDGE

TINY MOTORS

Spin and Swing

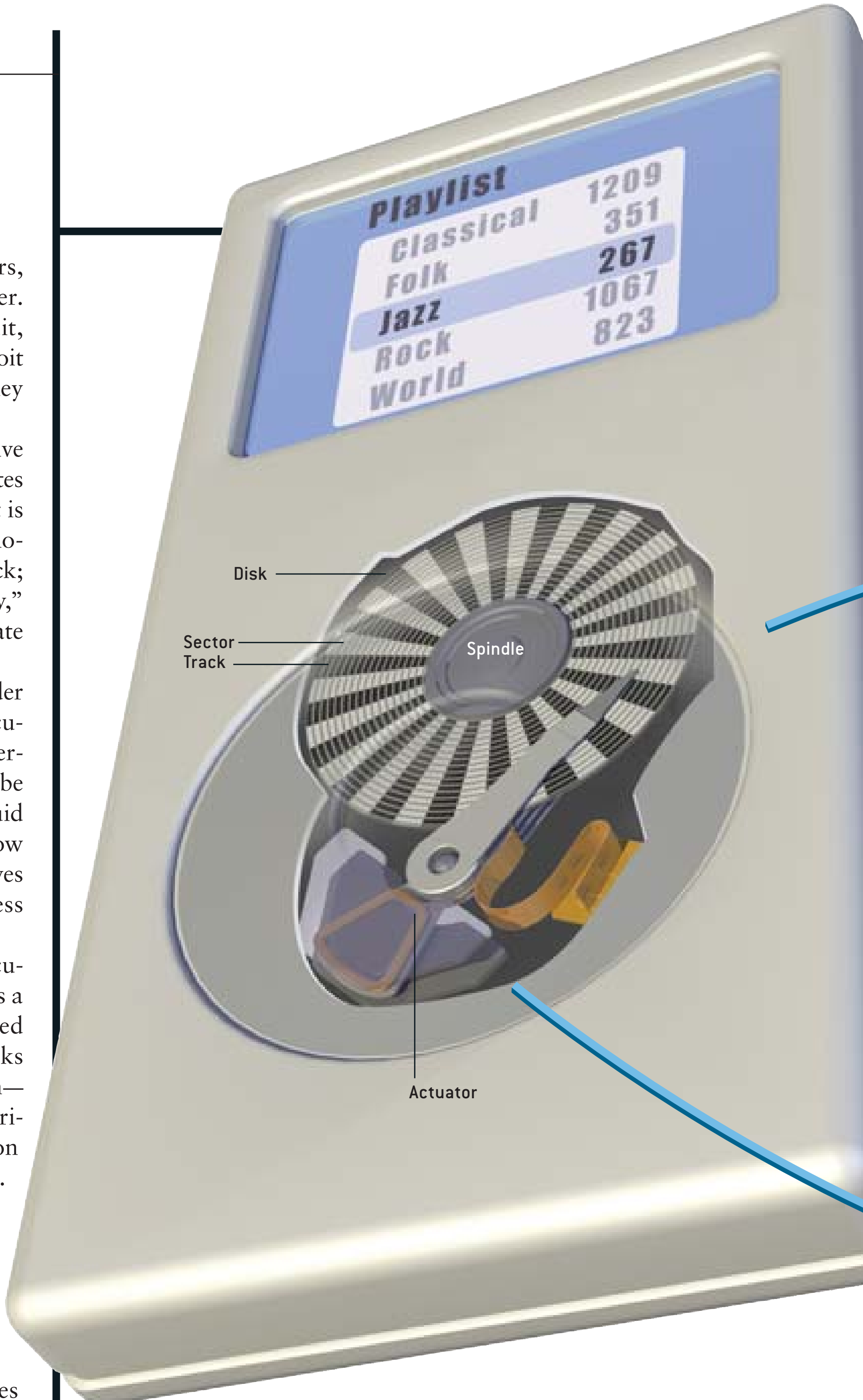
Portable consumer products such as music players, cameras and cell phones are becoming ever smaller. Miniaturized electronics deserve some of the credit, but so do ever shrinking motors. Most of them exploit the same physics principles as their bigger kin; the key is design and manufacturing ingenuity.

MP3 players such as the iPod provide an instructive example. They cram about 1,500 songs, or six gigabytes of data, per square inch of magnetic hard disk that is thinner than a dime. The disk is spun by a spindle motor. The rotor at its center is only two millimeters thick; it and other parts "would fit on the tip of your pinky," notes Hans Leuthold, head of motor R&D at Seagate Technology in Scotts Valley, Calif.

Most larger motors turn on small ball or cylinder bearings, but the disk drive must rotate with an accuracy of 0.05 micron. Crafting tiny ball bearings perfectly round enough to deliver such specificity would be prohibitively expensive. Instead the motor uses a fluid for bearings, typically an oil. Manufacturers are now experimenting with air bearings, because denser drives will require faster rotating speeds—demanding less friction than even slippery oil offers.

To read the densely packed bits quickly and accurately, a second motor, a voice-coil actuator, swings a read head back and forth over the disk's magnetized tracks. It is a variation of larger actuators. Yet tracks are becoming so tight—down to a tenth of a micron—that greater density may require a micromotor, fabricated using microelectromechanical technology, on the actuator's tip to fine-tune the head's placement. Other such motors are being designed that one day could scour plaque from clogged arteries or fix damaged cells.

Motors made with piezoelectric materials that bend to create motion are entering markets, too. Nevertheless, "spindle motors will probably always be there" to turn mini-disk drives, even denser ones that store video files, says Bi Chao, research scientist at the Data Storage Institute in Singapore. Suppliers are devising stronger yet smaller motor magnets to reduce power consumption, thus extending battery life. They continue to drive down expenses to lower the already cheap price of $1 to $3. —*Mark Fischetti*



GEORGE RETSECK

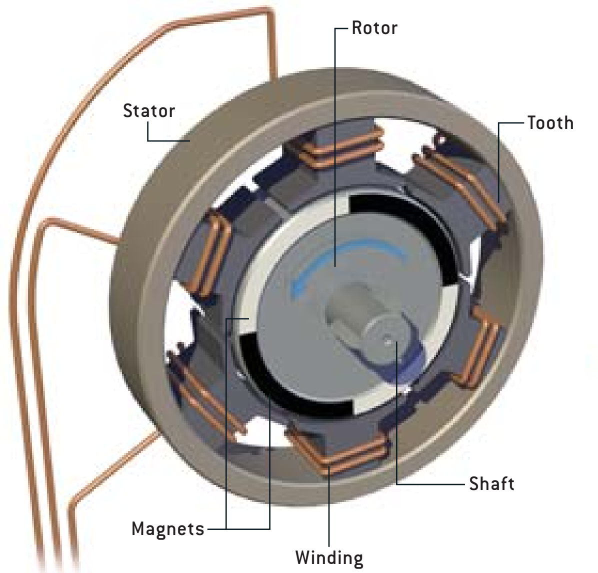
HARD-DISK DRIVES such as those in many MP3 players require one motor to spin the magnetic disk, so the correct sector can be read, and an actuator motor to move the read head back and forth to access the desired data track.

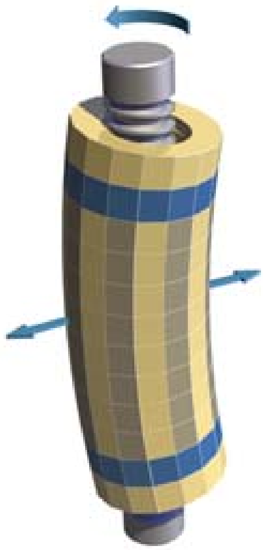
DID YOU KNOW...

BIOLOGY: Certain bacteria propel themselves using a flagellum, a whiplike tail. Molecules that surround its base release energy in a way that rotates the strand to and fro. Inspired, scientists are devising man-made molecular motors that can rotate at predetermined speeds or turn to preset positions when exposed to light. These could one day be inserted in living cells to manipulate internal processes.

CHEMISTRY: Liquids that remain stable across great temperature ranges are crucial to tiny motors. The machines are quiet and do not generate much heat, overcoming the noise and cooling troubles of larger designs. But they must operate in the deep freeze of a Minnesota winter and the oven of a car dashboard in an Arizona summer. The fluid film that forms the bearing cannot stiffen too much in extreme cold, because the motor cannot generate enough torque to overcome its resistance. When hot, the liquid cannot thin so much that moving parts touch and become damaged.

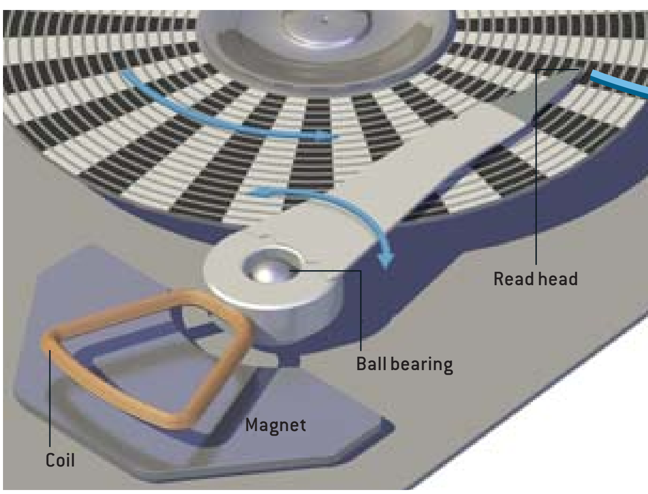
PHYSICS: Cell phones and pagers use a vibrator motor to signal an incoming call when the ringer is turned off. The little motor looks like a button-style watch battery. Incoming current interacts with the motor's permanent-magnet fields, causing the motor housing to vibrate.



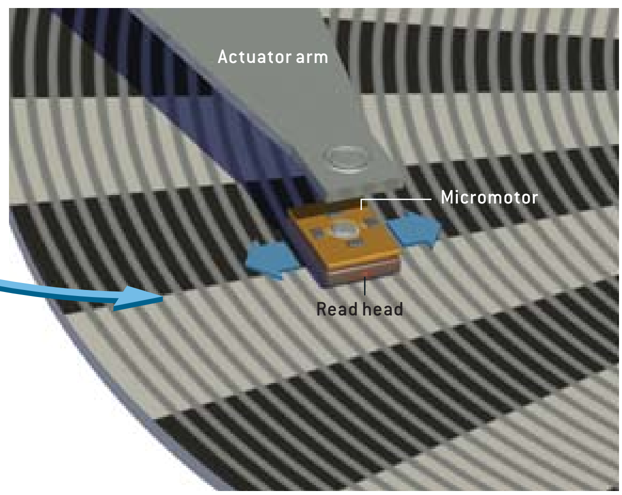


PIEZOELECTRIC motor materials change shape. Electric fields cause the outer ceramic bands to deform in precession, like a hula hoop winding up and down around a torso. Interior threads turn the ceramic screw out and in, an action that can zoom a lens or pump a fluid.

SPINDLE MOTOR is typically a permanent-magnet synchronous design. Windings generate a magnetic field around the top tooth of the stator and subsequently in each tooth; the advancing fields repel magnets on the rotor, causing it to spin, turning the shaft. A computer chip controls the current in the windings.



ACTUATOR motor is usually a voice-coil design. Magnets create a field; sending a current through the coil generates force that pivots the arm back and forth, moving the read head.



MICROMOTOR is made with microelectromechanical technology. Electric fields in the motor cause it to shift slightly back and forth, helping the read head discern tracks packed too tightly for the actuator arm to pinpoint.

Topic suggested by reader Philip Draper. Send ideas to workingknowledge@sciam.com

From Surmise to Sunrise

THE SCIENCE OF LIFE AND THE ART OF WRITING WELL **BY JONATHAN WEINER**

FROM SO SIMPLE A BEGINNING: THE FOUR GREAT BOOKS OF CHARLES DARWIN
Edited by Edward O. Wilson
W. W. Norton & Company, 2006 ($39.95)

DARWIN: THE INDELIBLE STAMP: THE EVOLUTION OF AN IDEA
Edited, with commentary, by James D. Watson
Running Press, 2005 ($29.95)

REEF MADNESS: CHARLES DARWIN, ALEXANDER AGASSIZ, AND THE MEANING OF CORAL
By David Dobbs
Pantheon Books, 2005 ($25)

"Great scientific discoveries are like sunrises," says E. O. Wilson. First they touch just the tips of a few peaks and steeples; then they illuminate the whole world.

The greatest discoveries change everything for us, says James D. Watson: not only our feelings about science, "but about existence."

No discovery has changed us more than Charles Darwin's. Now, thanks to a publishing coincidence, we have not one but two new Darwin readers, one introduced by Wilson, the other by Watson—two of the world's best-known biologists since Darwin. Each volume reprints the same classics: Darwin's *Voyage of the Beagle, On the Origin of Species, Descent of Man,* and *Expression of the Emotions in Man and Animals.* Each volume even carries the same epigraph. It is the final sentence from *Descent,* the sentence in which, after decades of dread and delay, Darwin finally allowed himself to drive his idea all the way home: "We must acknowledge, as it seems to me, that man with all his noble qualities, with sympathy which feels for the most debased, with benevolence which extends not only to other men but to the humblest living creature, with his god-like intellect which has penetrated into the movements and constitution of the solar system—with all these exalted powers—Man still bears in his bodily frame the indelible stamp of his lowly origin."

I've been dipping into these two volumes, rereading Darwin with Wilson's and Watson's commentaries. Like everyone who follows biology, I've read all three men for years. And to me the most interesting thing about this little publishing coincidence is the larger coincidence that these three biologists, who succeeded in opening up new views of life, have also been brilliantly successful popular writers.

Darwin's first book, the story of his five-year voyage around the world as a young man onboard the HMS *Beagle,* was a best-seller in 1845, even though it said virtually nothing about evolution by natural selection, the great discovery he had begun to incubate in secret. *Origin* was such a success in 1859 that the first printing sold out before publication, as Watson says in his introduction, "no fewer than one-third being bought by Mudie's Circulating Library, an endorsement of the likely popularity of the book equivalent to a recommendation today from Oprah Winfrey." Even Darwin's last book, a study of earthworms, which he published in 1881, the year before he

DARWINIAN METHOD: The evolutionary theorist wrote a series of great books to convince others of his powerful vision.

EMILY HARRISON (*photoillustration*); STAPLETON COLLECTION/CORBIS (*photograph*); PROFILES IN HISTORY/CORBIS (*writing*)

died, was a surprise best-seller—by Victorian standards, at least. "My book has been received with almost laughable enthusiasm," he wrote, "and 3500 copies have been sold."

In 1953, almost a century after *Origin,* the young James Watson stood over his sister's shoulder as she typed up a 900-word research paper he had written with his friend Francis Crick about the structure of DNA. "There was no problem persuading her to spend a Saturday afternoon this way," Watson wrote afterward, "for we told her that she was participating in perhaps the most famous event in biology since Darwin's book." *The Double Helix,* Watson's memoir of his discovery, was a best-seller, too.

As for Wilson, his book *Sociobiology*

caused a scandal in 1975 by speculating about the evolution of human instincts in the same style in which it analyzed the evolution of instincts in bees, wasps and ants. Since then, his skills as a writer have won him two Pulitzer Prizes and an audience as wide as Watson's. Not that the two men have always gotten along. In Wilson's memoir, *Naturalist,* he devotes a chapter to the period in the 1950s and 1960s when he and Watson were young biology professors at Harvard. Wilson writes, "I found him the most unpleasant human being I had ever met."

How can we explain this coincidence of scientific and literary success in Darwin, Watson and Wilson? (And in many other biologists—too numerous to list—who have been reviewed in a decades-long parade in the pages of *Scientific American.*)

One trivial factor has to do with mathematics. There's a publishing rule that every formula you use in a book cuts your potential readership in half. In *Origin,* Darwin was able to write profoundly about evolution without using a single formula. In *Principia,* on the other hand, Isaac Newton used so many formulas that if that publishing rule is correct, according to my rough calculations, the book shouldn't even have been read by Newton. Even today biologists such as Watson and Wilson can often do without formulas and still give us deep, wide views of nature. Other kinds of scientists manage that, too, but the trick is harder to bring off if you're a theorist of prime numbers, quantum states or string.

Then there's the power of biology as a subject. Like artists, philosophers, theologians and the rest of us, biologists think about life and death. Because these are subjects we all brood about, we can understand why biologists might care deeply enough to make the science of life their life's work. It is Darwin's love of life (a passion for which Wilson coined a word, "biophilia") that makes his *Voyage* a perennial pleasure to read. As Watson writes, *Voyage* is a story of discovery "not only of unknown lands and the organisms that inhabit them, but also of a young man's discovery of his own potential." Darwin's memoir of his years on the *Beagle* is, as Watson says, "an intimate book in which Darwin reveals much of his character." You could say the same of Watson's *Double Helix* or Wilson's *Naturalist.* Each memoir is about the wonders of life and about finding a wonderful life. We can feel Darwin's enthusiasm on his opening page, when he describes the *Beagle*'s first landing, on January 16, 1832, in the Cape Verde Islands. "The scene," Darwin writes, "as beheld through the hazy atmosphere of this climate, is one of great interest; if, indeed, a person, fresh from sea, and who has just walked, for the first time, in a grove of cocoa-nut trees, can be a judge of anything but his own happiness."

And there's the style in which Darwin worked toward his theory of the origin of species. He didn't gather a thousand and one facts and then invent a theory to explain them—which was the scientific style that had been urged on the world by the prophet of science Francis Bacon. Instead the young Darwin made a leap of imagination and then worked for decades to find out if his idea really held up. David Dobbs talks about Darwin's revolutionary style of doing science in *Reef Madness,* a book about Darwin's coral reef theory. (Dobbs's book is a gem; the title is its only flaw.) Darwin's leap of imagination is a feat that an artist can appreciate. He had a powerful vision of the way things are, of the way things go, and then he wrote a shelf of great books that convinced his readers of his vision. As Dobbs writes, "It was a move toward the power of story." Watson and Crick worked the same way in their discovery of the double helix: first the leap of intuition, then the tests.

Finally, of course, there is the power of the story itself. Darwin was born in the static world of scripture, and he left us a turbulent world of perpetual change. Ever since Darwin, we live in a world of stories. The story of that change will be told forever. We'll never get tired of reading and rereading it. First Darwin journeys alone from surmise to sunrise. Then the truth dawns on us all. SA

Jonathan Weiner won a Pulitzer Prize in 1995 for The Beak of the Finch. *He teaches science writing in Columbia University's Graduate School of Journalism.*

Annual Rapport

LIKE A GOOD WINE, 2005 HAD A FINE FINISH **BY STEVE MIRSKY**

One of the better gifts in Santa's bag is the year-end edition of the *British Medical Journal,* which brightens the cold days of late December with research reports that are actually fun to read. Take the report entitled "Harry Potter Casts a Spell on Accident Prone Children." The authors note that kid crazes, such as skating and scootering, often lead to emergency room visits. But traumatic injury as a result of widespread *Harry Potter* reading would presumably be rare, they say, "given the lack of horizontal velocity, height, wheels, or sharp edges associated with this particular craze."

So they tracked ER visits on weekends after the release of the past two *Harry Potter* books and found a "significant fall," but only of the kind in which no one gets hurt. Hospital trips were halved when kids were hunkered down with Harry. The authors muse: "It may therefore be hypothesized that there is a place for a committee of safety conscious, talented writers who could produce high quality books for the purpose of injury prevention." The Muggles who wrote the report point out, however, that the flip side could be "an unpredictable increase in childhood obesity, rickets, and loss of cardiovascular fitness."

Another study of particular interest during the celebratory season was on the "Shape of Glass and Amount of Alcohol Poured." Seems that our brains confuse height with volume—even experienced bartenders attempting to pour a standard shot misoverestimate and actually deliver about 20 percent more alcohol into a short, wide glass than into a tall, thin one. The scientists recommend going with highball glasses to avoid overpouring or to use "glasses on which the alcohol level is marked." Now, I can't see getting a special pair of glasses to wear just for pouring drinks, but I don't entertain that much. [*Editors' note: We agree—he barely entertains us at all.*]

Perhaps the most important research turned up in "The Case of the Disappearing Teaspoons: Longitudinal Cohort Study of the Displacement of Teaspoons in an Australian Research Institute." Noting that lounges at their institute always seemed to be short on teaspoons, which led to an inability to use sugar or instant coffee efficiently, which in turn must have had a negative effect on the quality of work, the researchers attempted to answer the question "Where have all the bloody teaspoons gone?"

The investigators planted 70 "discreetly numbered" teaspoons in the kitchen and did a weekly census. After five months, 80 percent of the teaspoons were gone, with teaspoons in common areas having a half-life of a mere 42 days. The authors propose three explanations. Quoting the classic 1968 *Science* article "The Tragedy of the Commons," which explains how individual overuse can destroy a community asset, they note that people may actually have taken the spoons. They also suggest the possibility of a distant planet populated by spoon life-forms, to which the spoons somehow migrated. Finally, they consider the theory of counterphenomenological resistentialism—"the belief that inanimate objects have a natural antipathy towards humans, and therefore it is not people who control things but things that increasingly control people." In other words, if you still think you're running the show, well, where *have* all the bloody teaspoons gone?

The *BMJ* was joined in spreading holiday cheer by Judge John Jones. On December 20 he issued his decision in the Dover, Pa., intelligent design trial that has been frequently mentioned on this page. The decision, available on the Web, is an instant classic of science writing, not to mention jurisprudence. In addition to delineating why untestable intelligent design is equivalent to the spoon-space-travel hypothesis above, Judge Jones harangued the school board for its "breathtaking inanity" in requiring an antievolution, pro-ID statement to be read in public school science classes. Local authorities are now looking into perjury charges against pro-ID board members who apparently lied in depositions and on the witness stand. Not very intelligent. SA

ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

How do electric eels generate a voltage, and why don't they get shocked?

Angel Caputi, senior scientist and head of the department of integrative and computational neurosciences at the Clemente Estable Institute for Biological Research in Montevideo, Uruguay, provides this answer:

The electric eel generates large electric currents by way of a highly specialized nervous system that synchronizes the activity of disk-shaped, electricity-producing cells packed into a specialized electric organ. The nervous system uses a command nucleus that tells the electric organ to fire. A complex array of nerves activates the thousands of cells at once.

Each electrogenic cell carries a negative charge of a little less than 100 microvolts on its outside compared with its inside. When the command signal arrives, the nerve terminal emits a minute puff of acetylcholine, a neurotransmitter. This release creates a transient path with low electrical resistance connecting the inside and the outside of one side of the cell. Thus, each cell behaves like a battery, with the activated side carrying a negative charge and the opposite side a positive one.

Because the cells are equally oriented inside the electric organ like a series of batteries piled into a flashlight, the current generated by an activated cell "shocks" any inactive neighbor into action, setting off a cascade that runs its course in just two milliseconds or so. If the eel lived in air, the resulting current could be as high as one ampere, turning the creature's body into the equivalent of a 500-volt battery. But water provides additional outlets, diminishing the current.

One explanation for why eels can shock other animals without zapping themselves could be that the severity of an electric shock depends on the amount and duration of the current flowing through any given area of the body. For purposes of comparison, an eel's body has roughly the same dimensions as an adult man's arm. To cause muscles in an arm to spasm, 200 milliamps of current must flow into them for 50 milliseconds. An eel generates much less energy than that with its two-millisecond current flow. Also, a large part of the current dissipates into the water through the eel's skin. The current discharged into the smaller bodies of prey is much stronger proportionally. For example, prey one tenth the eel's length is about one one-thousandth the animal's volume. Therefore, small animals close by get shocked, rather than the discharging eel.

What causes stuttering, and is there a cure? —Z. Schwartz, Vineland, N.J.

J. Scott Yaruss, associate professor in communication science and disorders at the University of Pittsburgh School of Health and Rehabilitation Sciences and co-director of the Stuttering Center of Western Pennsylvania, explains:

The classification of stuttering encompasses a number of communication disorders. Neurogenic stuttering and psychogenic stuttering, as their names imply, have a specific known cause—either a flaw in the makeup of the brain or a profound psychological challenge. Developmental stuttering, the most common version of this disorder, typically starts between the ages of two and a half and four. In addition to disruptions in their speech, people who stutter often experience negative emotional, cognitive or behavioral reactions that can further affect their ability to communicate.

Current theories suggest that stuttering arises from a combination of several genetic and environmental influences. Some of the elements now being examined are motor skills, language skills and temperament.

There is no known cure for stuttering, although many treatment approaches have proved successful for reducing its effects. Options include training to change speech patterns, counseling by speech-language pathologists to minimize negative reactions, pharmaceutical interventions and electronic devices that may enhance fluency. Self-help and support groups also play a prominent role in speech improvements for many people. SA

For a complete text of these and other answers from scientists in diverse fields, visit www.sciam.com/askexpert

MATT COLLINS