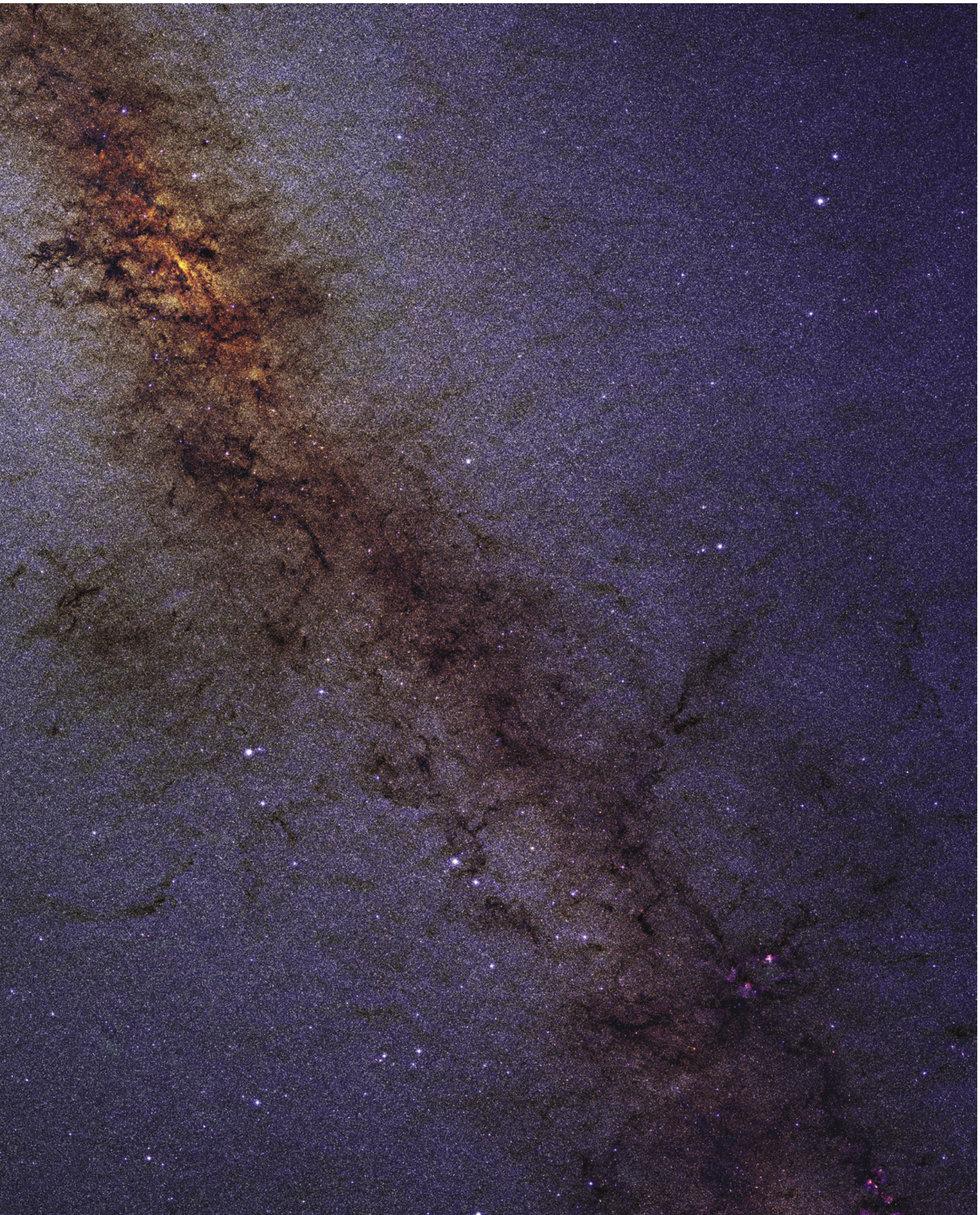
California
Institute of
Technology

E&S

ENGINEERING AND SCIENCE

Volume LXIV,
Number 1,
2001

IN THIS ISSUE

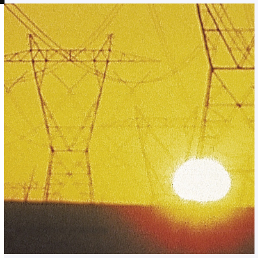


Peering through the dust that enshrouds our galaxy, the Two Micron All Sky Survey (2MASS) provides a clearer view of the center of the Milky Way than ever before seen. The actual center in this mosaic of more than a million stars is the bright reddish spot near the upper left. An article describing more of what this survey saw at near-infrared wavelengths begins on page 20.

California Institute
of Technology

ENGINEERING AND SCIENCE

Volume LXIV, Number 1, 2001



On the cover: Small snow crystals, each about one millimeter in diameter, grow on the ends of thin ice needles in Ken Libbrecht's lab. For more on how an infinite variety of shapes can come from a very simple molecule, see the story on page 10.

Engineering & Science (ISSN 0013-7812) is published quarterly at the California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125. Annual subscription $10.00 domestic, $20.00 foreign air mail; single copies $3.00. Send subscriptions to Caltech 1-71, Pasadena, CA 91125. Third class postage paid at Pasadena, CA. Published by Caltech and the Alumni Association. Telephone: 626-395-3630.

PICTURE CREDITS: Cover, 10, 14–19 — Ken Libbrecht; inside front cover, 20–27 — 2MASS; 2 — Robert Tindol; 4, 11 — Mark Garcia; 4, 7, 10–11, 12, 14–16 — Doug Cummings; 6 — Bob Paz; 7 — Mika Nyström; 8 — Nathan Litke; 9 — Erik Winfree; 28, 34–35 — Edison International; 30, 31 — Doug Smith; 37 — Craig Schwartz; 38 — Ken Jacques; 40–41 —Shannon Banks; 43 — Joe Umbro; inside back cover — Margaret DePrima

Ted Jenkins
President of the Alumni Association
J. Ernest Nunnally
Vice President for Institute Relations
Robert L. O'Rourke
Associate Vice President for Institute Relations

STAFF: *Editor* — Jane Dietrich
Managing Editor — Douglas Smith
Writer/Editor — Barbara Ellis
Contributing Writers — Robert Tindol, Mark Wheeler
Copy Editors — Emily Adelsohn, Michael Farquhar, Elena Rudnev
Business Manager — Debbie Bradbury
Circulation Manager — Susan Lee
Photographer — Robert Paz
Graphic Artist — Douglas Cummings

Visit Caltech on the Web at **http://www.caltech.edu**

CALIFORNIA: WINE, MOVIES, AVOCADOS, AND NOBEL LAUREATES

Not only does the State of California have the fifth largest economy in the world (having recently passed up France), it also has the largest concentration of Nobel laureates. And, because the year 2001 marks the 100th anniversary of the Nobel Prize, California will be celebrating and paying tribute to its laureates and to the significance of science and technology in the state. Which no doubt has something to do with its economy.

Of California's 94 Nobelists, Caltech counts 27 among its faculty and alumni (that's actually 28 prizes, since Linus Pauling won two). Three of the four faculty members still active on campus will be taking part in the Nobel Centennial Symposium in Beckman Auditorium on the afternoon of October 24. President David Baltimore (physiology/medicine '75), Rudy Marcus, the Noyes Professor of Chemistry (chemistry '92), and Ed Lewis, the Morgan Professor of Biology, Emeritus (physiology/medicine '95) will introduce speakers at the symposium. (Ahmed Zewail, the Pauling Professor of Chemical Physics and professor of physics, who won the chemistry Nobel in 1999, will unfortunately be out of town.) Other Caltech speakers at the event include Richard Andersen, the Boswell Professor of Neuroscience, and Professor of Physics Andrew Lange. A reception at the Athenaeum will follow.

Caltech's symposium is only a portion of the California Nobel Prize Centennial 2001, which will be attended by members of the Swedish royal family and the Swedish ambassador. It's preceded by a morning symposium at UCLA, which has a few Nobel laureates of its own. On October 25 a Centennial Celebration luncheon will be held at the California Science Center in downtown Los Angeles, chaired by Gayle Wilson, the state's former first lady and a member of Caltech's board of trustees. On October 26, the celebration moves northward to San Francisco, where there will be a symposium at the Exploratorium. A banquet at City Hall follows on the 27th.

The Nobel Centennial is also sponsoring an essay contest for junior- and senior-high-school students and establishing Centennial Scholarships. □

FLY LIKE AN EGYPTIAN

Although blustery winds added some excitement—the chief flyer, kite-surfing instructor Eric May (middle of picture, dark shorts), got lifted off the ground on the first attempt—the obelisk went upright in just 25 seconds on the second try. Graff (plaid shirt) handles a control line. The obelisk was provided by Daniel Correa (center, foreground) of Incablock, which makes concrete blocks.

When you think about building the Egyptian pyramids, you probably have a mental image of thousands of slaves laboriously rolling massive stone blocks with logs and levers. But as one Caltech professor is demonstrating, the task may have been accomplished by just four or five guys who flew the stones into place with a kite.

On June 23, Professor of Aeronautics Morteza (Mory) Gharib (PhD '83) and his team raised a 6,900-pound, 15-foot reinforced-concrete obelisk into a vertical position in the desert near Palmdale, using nothing more than a kite, a pulley system, and a support frame. The team eventually hopes to show that even 300-ton monuments—not to mention the far-less-massive building blocks of Egypt's 90-odd pyramids—could have been raised with a fraction of the effort that modern researchers have assumed.

Gharib, whose primary research interest is the nature of fluid flow, has been working on the project since local business consultant Maureen Clemmons contacted Caltech two years ago. Clemmons had seen a picture in *Smithsonian* magazine in 1997 of an obelisk being raised, and thought that the Egyptians could have used kites to accomplish the task more easily. All she needed was an aeronautics expert with the proper credentials to field-test her theory. It is a credit to her determination that the tests are occurring—with no scientific or archaeological training, she has managed to marshal the efforts of family, friends, and fellow enthusiasts.

Even today, moving heavy stones without power equipment is quite labor-intensive. In 1586, the Vatican moved a 330-ton Egyptian obelisk to St. Peter's Square. Lifting the stone upright took 74 horses and 900 men. For Gharib, the idea of accomplishing heavy tasks with limited manpower is appealing from an engineer's standpoint because it makes more logistical sense. "It's one thing to send thousands of soldiers to attack another army on a battlefield," he says. "But an engineering project requires everything to be put precisely into place. I prefer to think there were relatively few people involved."

The concept Gharib developed with SURF (Summer Undergraduate Research Fellowship) student Emilio Graff, a senior in aeronautics, is to build a simple tower around the obelisk, with a pulley system mounted somewhat forward of the stone's tip. That way, the base of the obelisk will drag the ground for a few feet as the kite lifts the stone, and it will be quite stable once it reaches the vertical. If the obelisk were raised with the base as a pivot, the stone would tend to swing past vertical and fall the other way. The kite rope is threaded through the pulleys and attached to the obelisk's tip. A couple of fliers steer the kite with guide ropes, moving it in figure-eights for maximum sustained lift.

Of course, no one has any idea if the ancient Egyptians actually moved stones or anything else with kites and pulleys, but Clemmons has found some tantalizing hints that the project is on the right track. On a building frieze now displayed in a Cairo museum, there is a wing pattern in bas-relief that does not resemble any living bird. Directly below are several men standing near vertical objects that could be ropes. And she has discovered that a brass ankh—long assumed to be merely a religious symbol—makes a very good carabiner for controlling a kite line.

Gharib next plans to raise a 10-ton stone, then perhaps a 20-ton one. Eventually, they hope to receive permission to raise one of the obelisks still lying in an Egyptian quarry. "The whole approach has been to downgrade the technology," Gharib says. "We first wanted to show that a kite could raise a huge weight at all. Now that we're raising larger and larger stones, we're also preparing to replace the steel scaffolding with telephone poles and the steel pulleys with windlasses like the ones that may have been used on Egyptian ships. In fact, we may not even need a kite. It could be we can get along with just a drag chute." Steady winds of up to 30 miles per hour are not unusual in the areas where the pyramids and obelisks are found. The wind in Palmdale gusted to over 20 m.p.h., although a 12-m.p.h. breeze would have sufficed. ☐—*RT*

PUT SOME CESIUM IN YOUR TANK

Gasoline averaging $3 per gallon? Oil drilling in an Alaskan wildlife reserve? A need to relax air quality standards? It seems the long-term future of fossil fuels is bleak. One promising solution scientists have been studying is fuel cells, which have their limitations too. But in the April 19 issue of *Nature*, Caltech Assistant Professor of Materials Science Sossina Haile reports on a new type of fuel cell that may resolve these problems.

Unlike an automobile engine, where a fuel is burned and expanding gases do the work, a fuel cell converts chemical energy directly into electrical energy. Fuel cells are pollution-free, and silent. The fuel cells in today's prototype cars are usually based on polymer electrolytes. (An electrolyte is a nonmetallic substance that conducts electricity.) Polymer electrolytes must be humidified in order for the fuel cell to function, and can only operate over a limited temperature range. Thus these fuel cell systems require many auxiliary components and are less efficient than other types of fuel cells.

Haile's laboratory has developed an alternative type of fuel cell based on a so-called "solid acid." Solid acids are chemical compounds, such as $KHSO_4$ (potassium hydrogen sulfate), whose properties are intermediate between those of a normal acid, such as H_2SO_4 (sulfuric acid), and a normal salt, such as K_2SO_4 (potassium sulfate). Solid acids can conduct electricity at rates similar to polymers, they don't need to be hydrated, and they can function at temperatures up to 250 °C. They are also typically inexpensive and easy to manufacture. But until now solid acids had not been examined as fuel-cell electrolytes because they dissolve in water; worse, they can lose their shape at even slightly elevated temperatures. To solve these problems, Haile and her colleagues operated the fuel cell at a temperature above the boiling point of water, and used a solid acid, $CsHSO_4$ (cesium hydrogen sulfate), that is not very prone to shape changes.

The next challenges, says Haile, are to reduce the electrolyte's thickness, improve the catalyst's performance, and, most importantly, prevent the reactions that can occur upon prolonged exposure to hydrogen. Still, she says, solid-acid fuel cells are a promising development. "The system simplifications that come about [in comparison to polymer electrolyte fuel cells] by operating under essentially dry and mildly heated conditions are tremendous. While there is a great deal of development work that needs to be done before solid-acid-based fuel cells can be commercially viable, the potential payoff is enormous." ☐—*MW*

In mid-May, the Division of Geological and Planetary Sciences commemorated the 75th anniversary of its founding (as the department of geology and paleontology under John Buwalda) with a reunion. About 100 alumni came for two days of talks by division faculty on their current research, as well as other sorts of celebrating, including this spectacular cake with a very literal "ring of fire." The following month, on June 24, Bob Sharp (BS '34, MS '35), the Sharp Professor of Geology, Emeritus, celebrated his 90th birthday. Sharp has been a member of the faculty since 1947 and was chair of the division from 1952 to 1968.

2001—A MARS ODYSSEY

As the Red Planet looms large in the midnight sky this summer, another JPL spacecraft has set out to meet it. The 2001 Mars Odyssey lifted off from Cape Canaveral on April 7, and is slated to arrive on October 24. Its primary mission, to run from January, 2002 through July, 2004, is to map the composition of the planet's surface, to detect water and shallow subsurface ice, and to study the radiation environment. The spacecraft carries three primary instruments.

THEMIS (Thermal Emission Imaging System) is a visible/infrared camera provided by Arizona State University that by day will map the distribution of minerals, particularly those that can form only in the presence of water. At night, it will look for active thermal regions that are warmer than their surroundings—potential "hot springs," at least by Martian standards.

The GRS (Gamma Ray Spectrometer), provided by the University of Arizona, will scout out 20 chemical elements, including silicon, oxygen, iron, magnesium, potassium, aluminum, calcium, sulfur, and carbon, on the Martian surface. A pair of neutron detectors in the GRS package (built by the Los Alamos National Laboratory and Russia's Space Research Institute) will allow the amount of hydrogen present in the top meter of soil to be calculated, which in turn acts as a proxy for determining the potential amount and distribution of Mars' water ice.

And MARIE (Mars Radiation Environment Experiment), led by NASA's John-

Not featured at the GPS reunion was the fine vintage at right, bottled for the neighboring Division of Physics, Mathematics and Astronomy in honor of SIRTF (the Space InfraRed Telescope Facility; see page 26). Infra Red wine comes in three "wavelengths"—3.6 microns, 30 microns, and 160 microns—indicating a range from light bodied to full bodied. It's reportedly "pretty good." Professor of Physics Tom Soifer (BS '68), director of the SIRTF Science Center, has enough Infra Red in his wine cellar to last until the satellite's launch in 2002.

son Space Center, will assess the radiation hazards to future human explorers. Cosmic rays emitted by the sun and other stars can trigger cancer or damage the central nervous system; similar radia-tion monitors have been flown on the Space Shuttles and on the International Space Station, but none has ever ventured beyond Earth's protective magnetosphere, which shields us from much of this radiation.

The orbiter is also designed to act as a communications relay for future Mars landers, including JPL's pair of Mars Exploration Rovers, to be launched in 2003. ☐

A REALLY NEAT DISCOVERY

JPL's NEAT (Near-Earth Asteroid Tracking) program's newest telescope, the just-refurbished 1.2-meter-diameter (48-inch) Oschin telescope at Caltech's Palomar Observatory, officially bagged its first asteroid on May 16. Better yet, the catch was a Potentially Hazardous Asteroid—one of about 300 now known whose orbit crosses Earth's. Provisionally named 2001JV1, it is about 0.7 kilometers (0.4 miles) in diameter, so it could leave a nasty welt. But don't get out the tinfoil hats or the beach umbrellas just yet—"potentially hazardous" means it would have to be significantly deflected from its current orbit to do us any harm.

Since its inception in December, 1995, NEAT has found roughly 100,000 asteroids, including about 100 near-Earth asteroids; NASA's goal is to find all of the estimated 700 to 1,500 asteroids larger than 1 kilometer (0.6 mile) that approach within 48 million kilometers (30 million miles) of Earth, and to do so by 2008. About 500 have been detected so far. The vast majority of these are harmless, but a tiny percentage have orbits that could eventually put them on a collision course with Earth.

The Oschin telescope, built in 1947, has been used for two landmark sky surveys, the second of which was completed in 2000. Its half-million-dollar upgrade, sponsored by NASA/JPL, has turned it into a fully automated facility with a computerized pointing system and a state-of-the-art CCD camera. ☐

BINOCULAR VISION: "FIRST FRINGE" AT THE KECKS

Astronomers' vision got a whole lot sharper on March 12, when at 10:40 p.m. Hawaiian Standard Time the W. M. Keck Observatory became an interferometer. The two 10-meter Keck telescopes atop the summit of Mauna Kea successfully pooled the light received from a star in the constellation Lynx known as HD 61294, attaining what astronomers refer to as "fringes." Interferometry, which has long been a staple of radio astronomy, means that the signals from two or more telescopes are combined to create a virtual telescope whose dish—or in this case mirror—is the size of the separation between the instruments, enabling you to discriminate between objects that are exceedingly close together. In this case, the telescopes are 85 meters (93 yards) apart, and the goal is to see warm, Jupiter-sized planets orbiting around nearby stars directly, rather than inferring their existence from the wobbles their gravitational tugs induce on their parent stars. This doesn't mean we'll be able to take the planet's picture, but with luck, the distinction between the star and the planet will be clean enough that the planet can be studied spectroscopically, giving information on its temperature, pressure, and atmospheric composition.

An interferometer has to align the incoming signals so that their peaks and troughs match up to within a very small fraction of a wavelength. This is easy enough with radio waves, which are a centimeter or more in length, but a very tough challenge with light, where the waves are measured in millionths of a meter. And it's further complicated by atmospheric turbulence, which causes the star (and its thousandfold dimmer companion) to shimmy disconcertingly. Each image wanders around on its detector independently from its twin at the other telescope, continuously altering the baseline separation.

The Keck Interferometer sends an image of the target star (or a bright, nearby guide star) from each telescope to a fast-readout infrared camera called the Keck Angle Tracker that sends commands back to adaptive-optics systems to compensate. The two star images are thus kept centered on a fiber-optic line that feeds another fast readout IR camera called the fringe tracker. The fringe tracker adjusts the "fast delay lines" that bounce the light through a system of adjustable prisms and mirrors to align the waves, while simultaneously compensating for the earth's rotation. A peak occurs in the fringe tracker when the paths are identically matched, and a minimum occurs when the paths are different by one-half the wavelength of the light. When the peak-to-minimum ratio exceeds a certain threshold value, fringes have been seen.

HD 61294 was the first of about 20 stars that were locked on to during the engineering run, which consisted of the first halves of the nights of March 12 to 14. The fringes would last for up to 10 seconds at a time, and for about 10 percent of the total duration that each star was tracked, which varied from 10 to 30 minutes. Now the challenge is to fine-tune the system to lock onto the fringes for long enough to make useful measurements, a teething period that is expected to take the rest of this year. Limited science operations, including looking at these planets and the dust rings from which planets condense, may begin early next year.

The Keck Interferometer is funded by NASA's Origins program, and is a collaboration between Caltech, the University of California, and JPL/NASA. ☐—*DS*

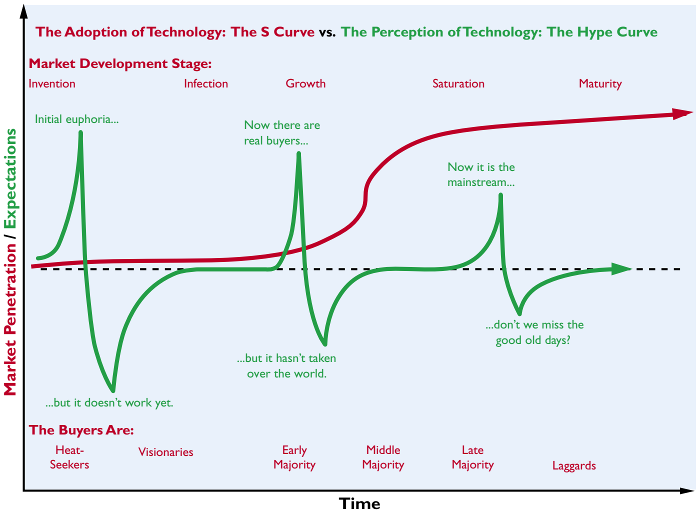
As new graduates sweltered under a broiling sun and parents and other wellwishers huddled under a sea of parasols, speaker Gordon Moore, PhD '54, former chair of Caltech's board of trustees and Intel Corporation, provided some advice from snowboarding: "Stay low and be confident as you move forward." At Caltech's 107th commencement on June 15, 204 new owners of BS degrees, 120 MS degrees, one Engineer, and 159 PhDs were ushered out into the real world.

FROM PUNCH CARDS TO PALM PILOTS

Caltech's computer-science option turned 25 this year. At Caltech, a birthday party means a symposium, so there were two days' worth of speakers on fields where Caltech has left its mark: chip design, parallel supercomputers, networking, and computer graphics, as well as a peek at what may lie beyond silicon. Here are some highlights.

Option cofounder Ivan Sutherland (MS '60), then Jones Professor of Computer Science, now vice president of Sun Microsystems, talked about the early days. In the Caltech tradition, the fledgling program had to "pick one thing and do it well." They opted for VLSI, or Very Large Scale Integration, because cofounder Carver Mead (BS '56, MS '57, PhD '60), Moore Professor of Engineering and Applied Science, Emeritus, was one of its fathers and would shortly write The Book on the subject. VLSI, which allows you to put an entire circuit—or, these days, millions of transistors—on a silicon chip, is now such a basic part of life that it doesn't merit a second thought. But back then, it was revolutionary, and the notion that grad students could actually design, build, and test several generations

One of the symposium's sessions was a panel discussion on "Entrepreneurship and Computer Science," at which Phil Neches (BS '73, MS '77, PhD '83), founder of Teradata Corporation, compared the standard business model's S-shaped curve of market penetration to the EKG-like "hype curve" of expectations typically found in the technology sector.

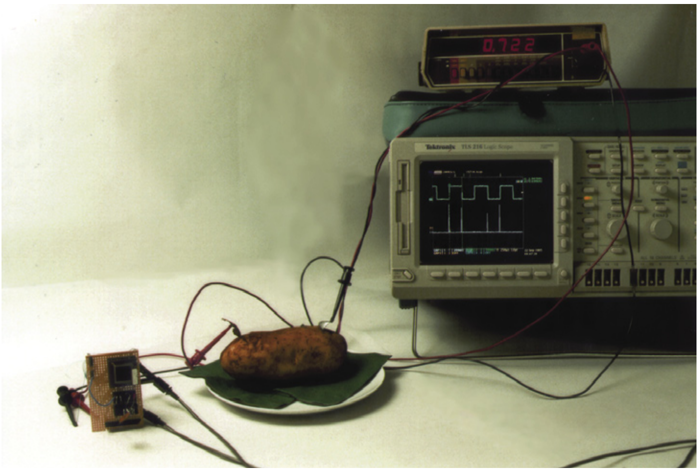


of chips in time to write a thesis was even more so.

At the dawn of the computer age, said Sutherland, logic elements (in the form of vacuum tubes) were expensive and unreliable. Wires, on the other hand, were cheap and reliable. Today, logic (transistors) is cheap and very reliable, but wires are expensive and very, very bulky. But "we're still tied to the mindset of 1950s programming, using detailed instruction sets. Instead, we need to put the programmer in charge of moving the data around. Addition is simple. Getting the operands to the adder is the hard part." He drew an analogy to another technological transition: "What we're doing now is copying the Roman stonemason's arched bridges in wrought iron. We need to begin building suspension bridges."

Most of the rest of the speakers talked about these suspension bridges. Stanford's William Dally (PhD '86) discussed designing chips based on a "parallel stream architecture." Such chips are broken up into many self-contained regions, each with its own memory and processor units. The idea is to handle as much of the computation as possible as locally as possible, by breaking the data up into streams and assigning each stream to its own region or regions; all the regions can then chew on pieces of the problem simultaneously without having to talk to each other very much. This reduces the torrent of data through the wires (both within and between chips) to a manageable flow, alleviating the traffic jams that would otherwise occur as more and more components are packed on a chip. Dally is building a one-teraflops (trillions of floating-point mathematical operations per second) machine that fits on a single shelf and draws less than a kilowatt of power. In contrast, the teraflops supercomputers up at Lawrence Livermore National Laboratory are the size of Beckman Institute Auditorium and use many megawatts of electricity. "These days, most of the power consumption isn't in the processing, it's in data movement." By 2011, he predicted, using such parallel stream architectures, we could have five teraflops on a chip and "a machine that won't require its own [power] substation." A handful of those chips could get you 100 teraflops, easy—a staggering amount of computing power. (When asked what home or business applications he anticipated for such a monster, he replied, "*Really* cool video games.")

Asynchronous chips, in which each circuit operates at its own pace rather than to the tick of a master clock, offer another way around the wiring bottleneck by leveling out the communications traffic, said Professor of Computer Science Alain Martin. Each asynchronous processor forwards its results the moment it finishes, rather than flooding the network when the clock says "SEND." Such chips also draw less power—as much as a quarter of the power consumed goes to running the clocks. Asynchronous chips are inherently faster, because the chip's speed is the average of all the processors' speeds instead of the speed of the slowest one. The downside is that since each processor is primed to accept a message all the time, any "glitch," or spurious signal, will be interpreted as data. (Synchronous systems are immune to this, as only a glitch at the exact moment that data is due would be taken seriously.) Clever communications protocols are needed to keep the data real and to manage the data flow in general. In fact, Martin titled his talk "Delays Have Dangerous Ends"—a quote from *Henry VI*—and was going to subtitle it, "A Shakespearean Approach to VLSI Design," but "decided it would not be appropriate for a Frenchman to do so."



In the lower left corner of the picture above is a silicon potato chip. It's also the world's first asynchronous microprocessor—a 16-bit chip with roughly 23,000 transistors, designed in Martin's lab in 1989. It ran at a then-respectable 17.5 MHz with a 5-volt power supply, drawing 230 milliwatts of power. But it will also run happily, albeit some 500 times more slowly, on the 0.9 volts and 40 microwatts obtainable from a nice, fresh, juicy potato, as Mika Nyström (MS '97, PhD '01) demonstrated.

To which an audience member replied, "So is it a tragedy or a comedy?"

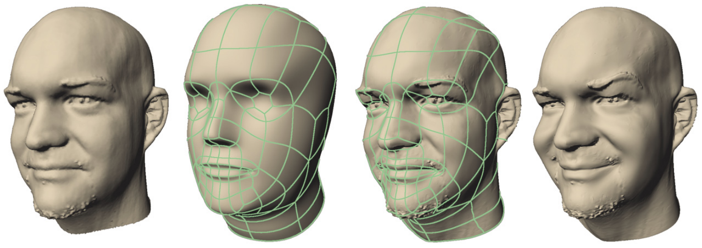
It was actually a history, beginning with the 1979 Caltech Conference on VLSI, where concurrent processing emerged as a discussion topic; Martin's lab built the world's first asynchronous microprocessor a decade later. A demonstration chip built in 1998 ran four times faster than the commercial two-million-transistor chip it was mimicking, and his lab still holds the record for the fastest working asynchronous circuits. (He quoted Carver Mead: "There is nothing more useless than a fast circuit that doesn't work.") He's trying to make them even more efficient. "If you track how the variables move—the atomic structure of the program, if you will—and then do a spectral analysis to see which operations 'cluster,' according to some measure, you can optimize your design based on the clustering." Martin, Dally, and a host of others are creating programming and chip-design tools that automatically deal with the complex details, leaving the humans free to draw the big pictures.

Another session covered how you link devices, asynchronous or otherwise, together. Caltech has been a pioneer in distributed computing since the early '80s, when Charles Seitz (a CS faculty member until 1994, when he left to found Myricom, a high-speed networking company) wired together a boatload of off-the-shelf PC processor boards to attain supercomputer performances at Radio Shack prices, and Geoffrey Fox (then professor of theoretical physics and dean of educational computing, now at Florida State) used it to tackle a gnarly quantum-field problem whose immensity had previously deterred all comers. Called the Cosmic Cube because its processors were connected like the vertices of an *n*-dimensional cube, it spawned an industry.

In the early '90s, JPL Senior Scientist and Caltech Faculty Associate (at CACR, Caltech's Center for Advanced Computing Research) Thomas Sterling, then at NASA Goddard, and his colleagues went one better by using commercial network technology and operating systems as well to create Beowulf—the model for today's teraflops computers. Sterling foresees a petaflops (quadrillion flops) machine by 2010. He proposes using Beowulfs—on a slightly smaller scale—in spacecraft. "NEAR [the Near-Earth Asteroid Rendezvous mission] sent back some 160,000 pictures. But the data product the scientists really wanted was a 3-D map of the asteroid. What if the spacecraft could do the advanced processing and you could just download the map? That would be equivalent to a 10^6 data compression."

The new frontier lies in linking computers that need to share some data, but don't want to get too intimate, in networks that evolve as the task dictates. Mani Chandy, Ramo Professor and professor of computer science, has come up with one such system, called Infospheres, which he plans to donate to humanitarian agencies doing disaster relief. Red Cross field workers could use their palmtop computers to coordinate food and clothing shipments, for instance, by tapping into various donors' inventory systems, railroad and airline schedules, and the truckers' dispatch centers. "Even poorer countries like India are heavily wireless," says Chandy. "Technology leapfrogs, so they've skipped right past telephone lines." The system employs what Chandy calls "screen scraping"—reading data from someone else's Web display that isn't formatted the way your computer likes it, but interpreting it anyhow and replying appropriately. Warming to that theme, Hewlett-Packard's Rajiv Gupta (MS '87, PhD '91), a protégé of Chandy's and Mead's, talked about having a chip in your car determine that your transmission is about to blow and automatically using your cell phone to make an emergency appointment at the closest dealership, displaying directions to it on your in-dash GPS unit, and booking a taxi to be there when you arrive so that you still get to your Very Important Meeting on time. "You don't know or care how the other components work, or even whose they are. You just need their output. The industrial revolution removed people as the bottleneck in the production of goods; this removes people as the bottleneck in the production of services."

Yaser Abu-Mustafa (PhD '83), professor of electrical engineering and computer science, designs systems called neural networks that

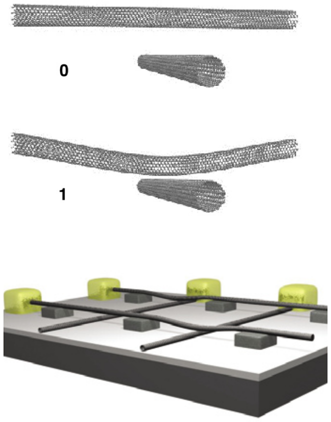


Schröder's lab develops methods for compressing and manipulating 3-D geometric information. Once properly encoded, you can treat the data like any other kind of signal, filtering out noise or changing some aspect of the object. Here, pushing up two critical grid points at the corners of the mouth puts a smile on someone's face. Schröder sees geometry as the fourth wave of multimedia, following on sound, still photos, and video, all of which are now routinely transmitted over the 'net. Fields as diverse as archaeology and biomedicine will benefit, and let's not forget the ubiquitous e-catalogs.

learn from hint and example rather than having to be programmed. For years, Abu-Mustafa has been training them to predict trends in the foreign-exchange market—a task chosen because "it is rich in data, is very noisy, and for which there is no mathematical model," thereby inadvertently launching the discipline of computational finance. "If I get a good system, you'll hear about it. If I get a *very* good system, you won't hear about it," he joked, going on to note that the results have, in fact, been published.

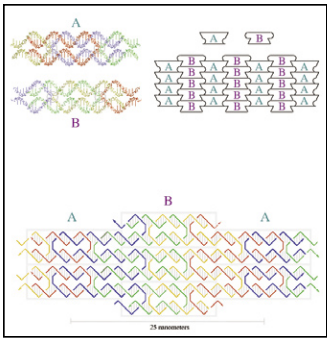
Jim Kajiya, a CS professor from 1978 to '94 and now assistant director of research at Microsoft, spoke of the emergence of computer graphics as a medium in its own right. "Applied computer sciences extend human capabilities: robotics, our muscles; memory and computation, our brains; and graphics, our imagination. So are graphics just for games for 14-year-old boys?" No. The confluence of computer vision and computer graphics becomes computer video—malleable, editable objects you can manipulate on your screen. Peter Schröder, professor of computer science and applied and computational mathematics, picked up the theme in describing his 3-D modeling research. How do you store and transmit geometric information so that a coarse but usable rendition of the object shows up almost immediately, with progressively finer detail filling in afterwards? How do you search a collection of shapes for common themes—to find all the animals, perhaps, or to recognize a face? And can you put digital objects that were scanned as separate entities together, reassembling a vase from its shards?

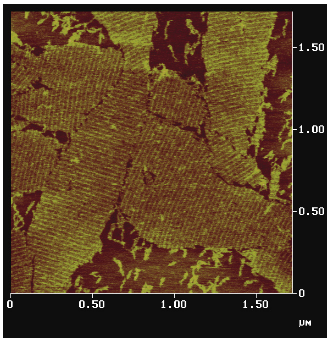
The final session looked at means of computing beyond silicon. According to Assistant Professor of Computer Science Andre DeHon, a silicon wire slated for production around 2005 is about 100 nanometers, or 200 silicon atoms, wide; about 44,000 square nanometers will be needed to encode one bit of information. But a variety of other technologies offer the possibility of building molecular wires and switches one nanometer wide; a bit might occupy 400 square nanometers. DeHon's own work centers on carbon nanotubes, which with the right electrostatic charge will flex to form binary (on/off) logic elements. "This will require a paradigm shift from the top-down methods of bulk carving and etching to bottom-up strategies for self-assembly, taking into account the specific characteristics of individual atoms." Even more exotically, quantum computing will attempt to exploit the bizarre properties of quantum-mechanical systems to solve problems that ordinary computers can't crack, said Associate Professor of Computer Science Leonard Schulman. The NSF recently established an Institute for Quantum Information at Caltech, to which Schulman and several colleagues, including one of Caltech's two most recent MacArthur "genius" Fellows, Assistant Professor of Physics Hideo Mabuchi (PhD '98), belong. And finally, Caltech's other MacArthur Fellow, Assistant Professor of Computer Science and Computation and Neural Systems Erik Winfree (PhD '98), described the first steps toward computing using DNA molecules whose structures encode a "program," and whose self-assembly into an array reads out the "answer." Look for more on Mabuchi and Winfree in a future issue. ☐—*DS*

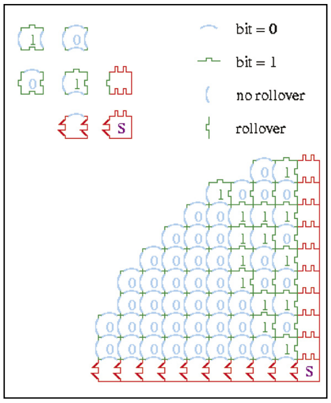


After Lieber, et al, *Science*, volume 289, page 95, July 7, 2000.

Single-walled carbon nanotubes are essentially soda straws whose walls are one carbon atom thick. Two perpendicular tubes, suspended two billionths of a meter apart, can flex under the influence of static cling to encode zeroes and ones (top). Chips built from arrays of such junctions (bottom) would be more compact than silicon ones. Computer architect DeHon is collaborating with Harvard chemist Charles Lieber to try to build practical circuits this way.



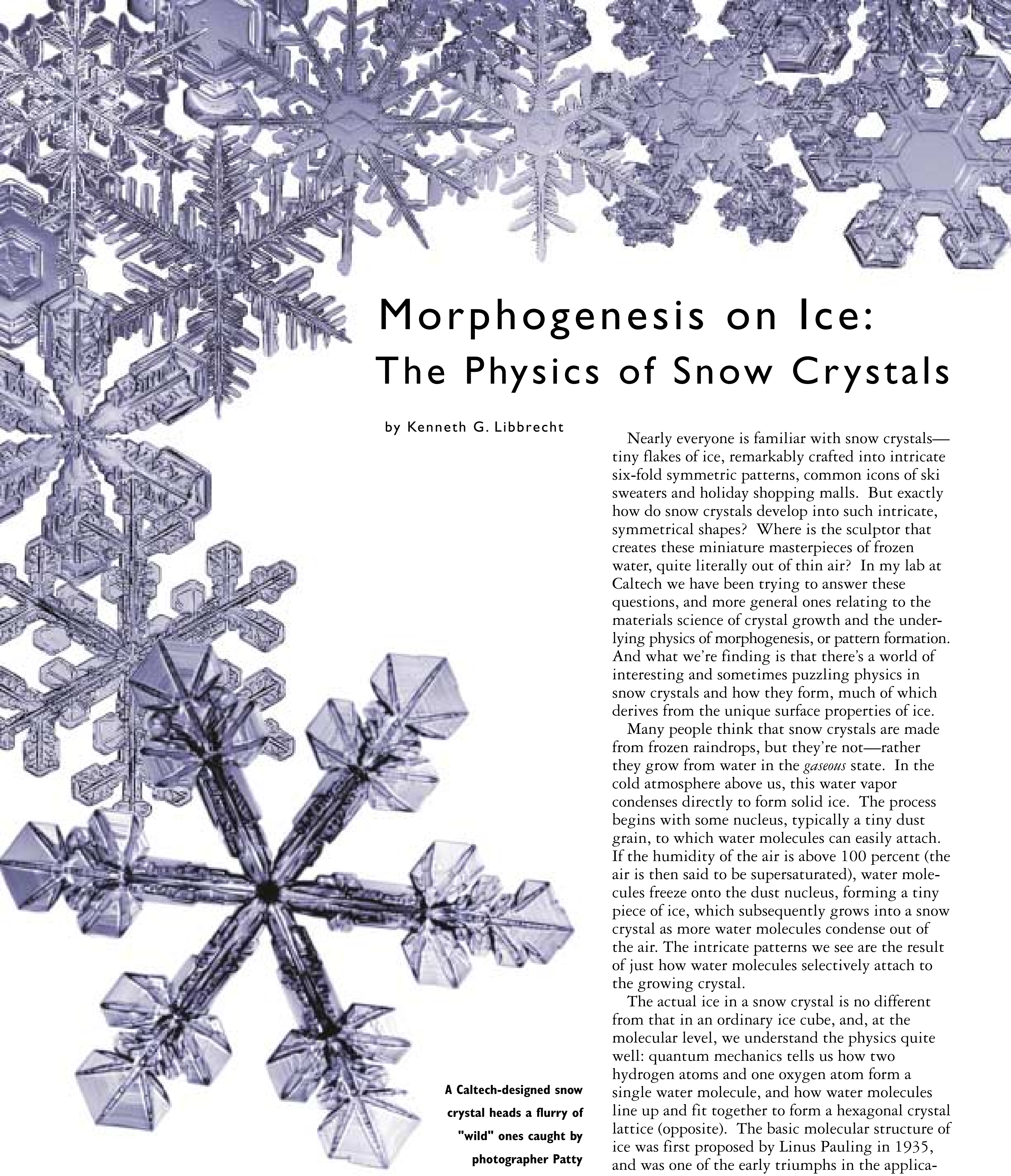




Left: A DNA computer would use "tiles" that interlock like puzzle pieces. In this example, the computer counts upward from 1 (in binary numbers) as the tiles fall into place, beginning from the one labeled "S" in the bottom right corner. Above, left: The first step is to create a lexicon of rigid, two-dimensional tiles that act as logic elements by binding to one another only in specific configurations. Here, four bits of single-stranded DNA self-assemble into double-crossover units that can take either of two forms, "A" or "B." A and B, in turn, fit together only one way to form a repeating pattern. Above, right: An atomic-force microscope scan of the actual A-B crystal. The B tiles stick up higher, giving it a corduroy look.

Morphogenesis on Ice:
The Physics of Snow Crystals

by Kenneth G. Libbrecht



A Caltech-designed snow crystal heads a flurry of "wild" ones caught by photographer Patty Rasmussen.

Nearly everyone is familiar with snow crystals—tiny flakes of ice, remarkably crafted into intricate six-fold symmetric patterns, common icons of ski sweaters and holiday shopping malls. But exactly how do snow crystals develop into such intricate, symmetrical shapes? Where is the sculptor that creates these miniature masterpieces of frozen water, quite literally out of thin air? In my lab at Caltech we have been trying to answer these questions, and more general ones relating to the materials science of crystal growth and the underlying physics of morphogenesis, or pattern formation. And what we're finding is that there's a world of interesting and sometimes puzzling physics in snow crystals and how they form, much of which derives from the unique surface properties of ice.

Many people think that snow crystals are made from frozen raindrops, but they're not—rather they grow from water in the *gaseous* state. In the cold atmosphere above us, this water vapor condenses directly to form solid ice. The process begins with some nucleus, typically a tiny dust grain, to which water molecules can easily attach. If the humidity of the air is above 100 percent (the air is then said to be supersaturated), water molecules freeze onto the dust nucleus, forming a tiny piece of ice, which subsequently grows into a snow crystal as more water molecules condense out of the air. The intricate patterns we see are the result of just how water molecules selectively attach to the growing crystal.

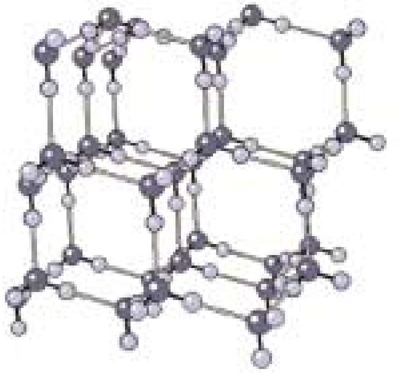
The actual ice in a snow crystal is no different from that in an ordinary ice cube, and, at the molecular level, we understand the physics quite well: quantum mechanics tells us how two hydrogen atoms and one oxygen atom form a single water molecule, and how water molecules line up and fit together to form a hexagonal crystal lattice (opposite). The basic molecular structure of ice was first proposed by Linus Pauling in 1935, and was one of the early triumphs in the application of quantum physics to materials science.

Build a better snowflake, and the world will shovel a path to your door.—KGL

It's this hexagonal symmetry of the ice crystal that is ultimately responsible for the six-fold symmetry of the snow crystals that fall from the sky. But just how do molecular forces, acting at the subnanometer scale, control the shape of a snow crystal ten million times larger? This same question applies to all crystals that form facets—flat surfaces that define the crystal shape—such as the mineral specimens shown below.

Facets have an interesting history in human society, which has led to some confusing expressions in our language. Many minerals in the earth grow into beautiful faceted shapes, and in early times these crystals were quite valuable. Since nice mineral specimens are rare, people naturally started carving facets into other materials, particularly glass (in which the molecules are randomly arranged, in contrast to the regular arrangement of atoms in a crystal). So now one can go to Macy's and buy a piece of fine "crystal"—which is in fact glass—with facets cut into it. Diamonds and other gemstones are crystalline materials, but here again the facets are usually of human origin. An amusing recent development has come from the superstition that quartz (and other) crystals possess mystical healing powers, which has greatly increased the demand for attractive mineral specimens. Demand begets supply, and lately I've seen "fake" quartz crystals in stores—real quartz, but with artificially made facets, cut to look like natural faceted quartz. I can't help but speculate that fake facets must diminish the healing powers, but I haven't explored the matter further.

Getting back to the natural world, the reason many crystals grow into faceted shapes is simply because some crystalline surfaces grow more slowly than others. And this in turn arises from the molecular structure of the crystal. For example, if we imagine beginning with a small, round ice crystal, we would find that the surface was quite rough on a molecular scale, with lots of dangling chemical bonds. Water molecules from the air can readily attach to these rough surfaces, and they grow relatively quickly. The facet planes are special, however, in that they tend to be smoother on a molecular scale, with fewer dangling bonds. Water molecules cannot so easily attach to these smoother surfaces, and so the facets grow more slowly. After all the rough surfaces have grown out, what's left are the slow-growing faceted surfaces.

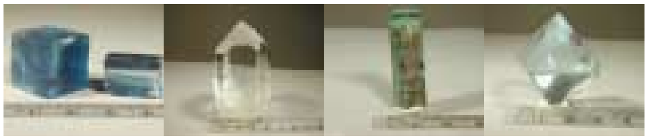


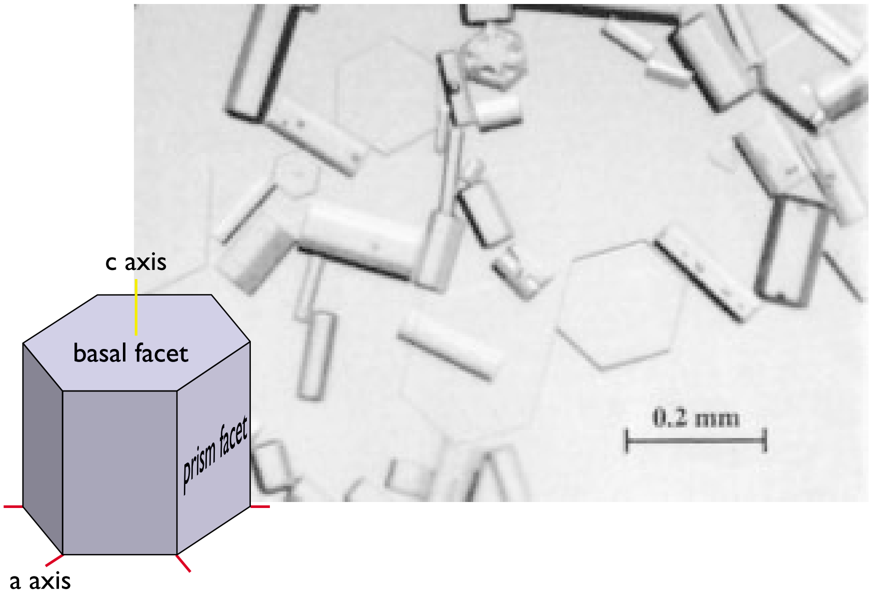
Snow crystals are six-sided because water molecules bond to form a hexagonal ice crystal lattice, above.

A big problem in crystal growth, and not just for ice, is that quantum mechanics cannot really tell us how *fast* a given surface will grow. It could in principle; but in practice, for essentially all real surfaces, the problem is exceedingly difficult. Supercomputer simulations can produce molecular models of crystals, but typically these elaborate models are not useful for modeling growth rates, because the thermal motions of molecules are very fast, while crystal growth is quite slow. A typical timescale for molecules in a crystal to jiggle back and forth is on the order of picoseconds (10^{-12} seconds), whereas the timescale for crystal growth is far longer—typically microseconds (10^{-6} seconds) or more.

Of course we don't need to keep track of every molecular wiggle, and over the past five decades many excellent theoretical techniques have been developed to describe the statistical mechanics behind crystal growth. Much of this has been driven by commercial concerns, as semiconductor crystal growth is the foundation of a very, very big business, and ever more sophisticated electronic and optical devices require the growth of ever more complex layered crystals. However, despite this huge, commercially funded research effort, crystal growth remains largely an

The crystals below show the natural faceting that is also a feature of snow crystals. From left, blue halite, (genuine) quartz, tourmaline, and fluorite.





W. Tape, *Atmospheric Halos*, p. 21, © 1994 by the American Geophysical Union.

The basic ice crystal shape, a hexagonal prism, can be seen in very simple crystals that fall at the South Pole, above.

empirical science, even for simple and well-studied materials like silicon. And we're now finding that ice has its own very different and quite fascinating story to tell, mainly because, unlike silicon, it has a very high vapor pressure. To tell this story, we must first see what ice crystals growing from water vapor look like.

Twelve-sided snow crystals form via a twinning mechanism and are actually fairly common. (Photo, P. Rasmussen)

We can learn a great deal simply by observing the great bounty of snow crystals that appear out of thin air, in snowfall. The earth's atmosphere is not a bad laboratory for the study of snow-crystal structure, and it's all for free. Natural snow crystals exhibit a remarkable variety of crystal shapes. The photos on the first two pages of this article show a collection of beautiful snow crystals that fell from the skies over Wisconsin, and these beauties demonstrate nicely the intricate structure that most of us associate with snowflakes. (A brief note on meteorological terms: a snow *crystal* refers to a single ice crystal, while snow*flakes* are clumps of snow crystals that stick together and fall to earth as little puffballs). While such fancy specimens are clearly the favorite of snow-crystal photographers, there are many other common shapes that tell us about the physics of snow crystals. The most basic ice-crystal shape is a hexagonal prism, which has two "basal" facets and six "prism" facets (see above). Very simple crystals like these are actually quite common, and the photo shows some examples that were collected at the South Pole. Because the conditions at the Pole are both very cold and very dry, snow crystals in this environment grow extremely slowly and typically don't become very large. Slow growth is the key factor for making simple snow crystals, and these tiny gems—in snowfall circles called "diamond dust"—probably floated through the Antarctic air for hours before growing as large as the thickness of a human hair. Not exactly ski sweater material, but these examples show the simple faceting that results from the underlying hexagonal symmetry of the ice crystal.

In more hospitable climates, snow crystals tend to grow much more quickly, and have a greater variety of forms. Simple prisms can grow into long, thin columns, which are usually hollow. Faster-growing columns branch into clusters of long, thin, needlelike crystals. Sometimes a snowfall can consist entirely of these columnar and needlelike crystals—quite painful! On another day, a snowfall could drop a preponderance of platelike crystals: small examples are often just simple hexagonal plates, but larger ones tend to have more structure, such as sectored plates or stellar dendrites ("dendrite" means "treelike," describing the branched structure of these crystals—these are the ones that come to mind when thinking of snow crystals). Platelike crystals are quite common in the wild, and they are the biggest specimens—often several millimeters across and easily seen with the naked eye.

Permutations of the simpler shapes give us a nearly infinite variety of complex snow crystal shapes—dendrites with sectored plate extensions, sectored plates with dendritic extensions, plates with platelike extensions, and even such exotica as "tsuzumi" crystals, which consist of columns capped with plates (named after a Japanese drum with a similar appearance). Some examples of the main types are shown on the opposite page. Twelve-sided snowflakes (left) can also be found floating through the air, and are not even that uncommon. These are actually two separate platelike crystals joined together at the center, one rotated 30 degrees relative to the other in a process called crystal "twinning." For the sake of anatomical correctness, however, I must point out that although *eight*-sided snowflakes are often seen on holiday wrapping paper, on ski sweaters, and even on soup cans (right), they are never found in nature.

Snow-crystal watching is incredibly easy, and can be quite fascinating. Simply pack a magnifying glass on your next ski trip, or whenever you might encounter some snowfall (a small pocket-size fold-up plastic model is perfect—you don't need the bulky Sherlock

T. Kobayashi & T. Kuroda, *Morphology of Crystals Part B*, ed. I. Sunagawa, Terra Scientific Publishing Co., 1987.

Snow crystals don't just come in the ski-sweater variety. Plate-like forms (top, from left to right) include a simple plate, sectored plate, stellar dendrite, and endless variations. Columnar forms range from simple hollow columns (far left) to thin needles (middle). Capped columns, or tsuzumi crystals, can also be seen (right).

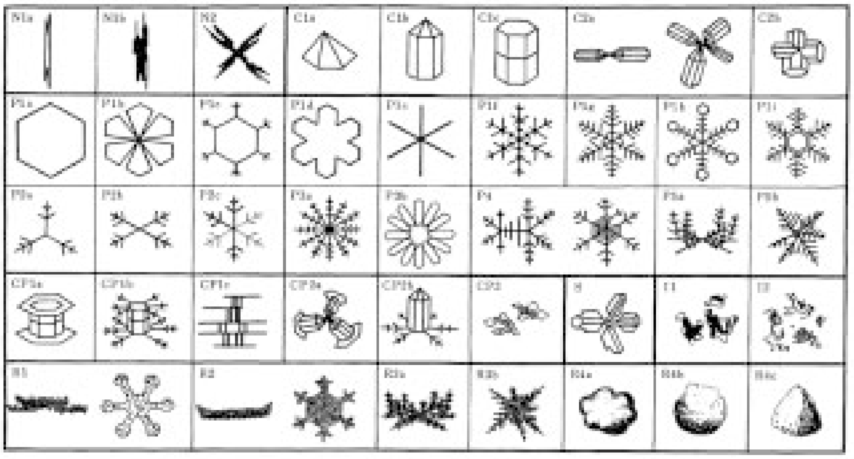
Holmes variety). On a good day you can see some beautiful crystals—and as with gemstones, photographs rarely do them real justice. You may find yourself echoing the sentiments of Thoreau, who remarked, "How full of the creative genius is the air in which these are generated! I should hardly admire them more if real stars fell and lodged on my coat."

So how do we make sense of all these different snow-crystal shapes? The first real scientific steps were taken by Japanese physicist Ukichiro Nakaya in the 1930s. Nakaya trained as a nuclear physicist, but on graduation found that no suitable jobs were available in his field. He eventually took a professorship at Hokkaido University in northern Japan, even though they had no nuclear physics department. Undaunted, Nakaya turned his attention to snow crystals, which were available locally in great abundance. After observing and categorizing natural snow crystals in great detail for the first time, Nakaya then developed techniques to grow artificial snow crystals in his laboratory, so that he could study their properties under controlled conditions.

One of Nakaya's early problems was that it was very difficult to grow individual snow crystals out of supersaturated air—the usual result was a mixed-up jumble of crystals that formed what was essentially a frost. Nakaya experimented with many thin fibers on which he tried to grow individual snow crystals—cotton fiber, silk fiber, metal wire, and even spider's web. Nothing worked, until he finally happened upon rabbit hair, and on this he was able to grow single crystals that were a great deal like natural specimens. (We tried this briefly in my lab, but without much success—I suspect we were using the wrong kind of rabbit!)

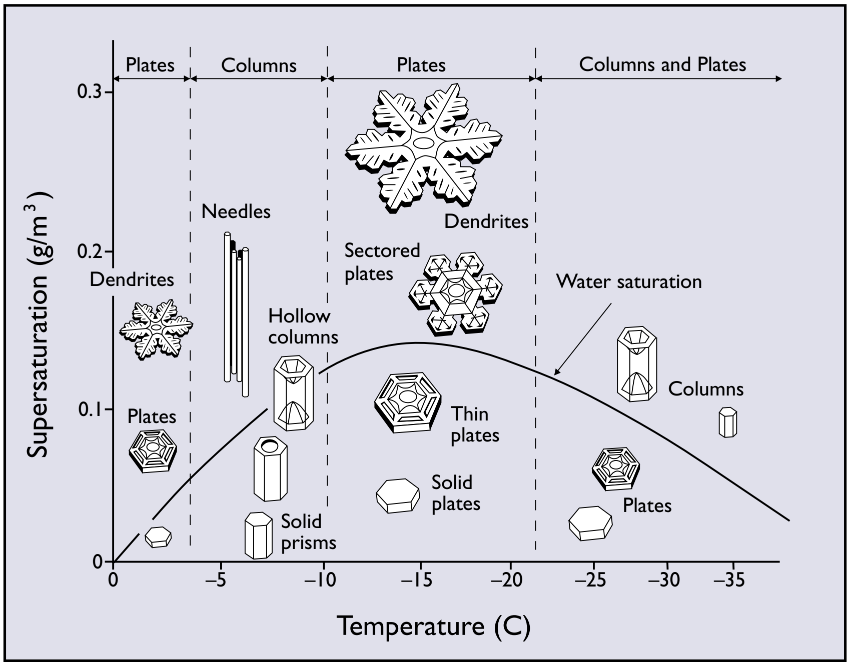
Nakaya discovered that snow crystals grow in different shapes, or morphologies, depending on the conditions in which they grow—in particular on the temperature and supersaturation of the air. Detailed measurements by Nakaya, and subsequently by others, were used to produce a snow crystal "morphology diagram" (next page) showing the crystal shapes that grow under different conditions.

This interesting diagram tells us a great deal about snow-crystal physics. It shows that the



Above, Nakaya in his laboratory. Right, part of his classification of natural snow crystals into 7 major groupings subdivided into 41 morphological types.

Reproduced by permission from *Snow Crystals: Natural and Artificial* by U. Nakaya, © Harvard University Press, 1954.



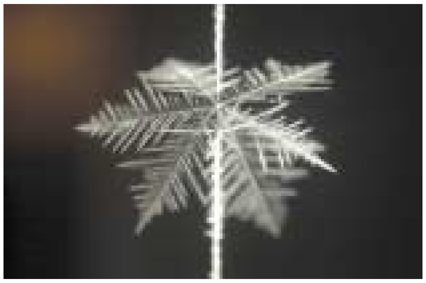
Snow crystals grow into different shapes depending on the temperature and supersaturation levels in the clouds in which they form. The morphology diagram, left, shows that some temperature ranges produce columns, others plates. Shapes get larger and more intricate the higher the supersaturation levels. (Adapted from a diagram by Y. Furakawa.)

basic *shape* of a growing crystal depends mainly on temperature: plates form at around –2 °C, columns at –5 °C, plates again at –15 °C, and either columns or plates below –25 °C. The appearance of *structure* depends more on the level of supersaturation, and therefore on growth rate: when the humidity is high, rapidly growing columns become feathery needle crystals, and hexagonal plates grow into stellar dendrites. An especially remarkable thing about snow crystal growth is the way the shape, or morphology, changes back and forth between plates and columns several times as a function of temperature. And the changes are large: within a few degrees the morphology changes from very long, narrow, needle-like crystals (at –5 °C) to very thin, flat, platelike crystals (at –15 °C). This is pretty bizarre—other materials don't change their growth morphology with temperature nearly as much as ice does. So what's going on?

Two main factors produce what we see in the morphology diagram: the intrinsic growth rates of the crystal facets, and diffusion. Let's look at crystal facet growth first. When the prism facets grow more slowly than the basal facets, we get columnar crystals; when the basal facets grow more slowly than the prism facets, we get platelike crystals. So the morphology diagram already tells us that the prism and basal facets grow at different rates, and both rates depend sensitively on temperature.

In my lab we've been working on a series of quantitative measurements to try to figure out in more detail what's really behind all this. The physics we're exploring is the statistical mechanics of crystal growth, which in turn depends on the detailed surface properties of ice. Ice is a better material than you might think for doing basic materials physics—it's a relatively simple substance, and, because of its environmental importance, ice has been studied using practically every experimental and theoretical tool known, so that the material itself is well characterized. Ice also turns out to be very convenient to work with in the lab—the stuff is cheap, it freezes at an easily accessible temperature, and as a chemical it poses no health hazards whatsoever (aside from drowning, which usually isn't much of a problem).

The basic idea for our experiments is quite simple: we grow snow crystals under controlled conditions, and measure the crystal dimensions as a function of growth time. In these experiments we usually grow only very tiny crystals, smaller even than the Antarctic ones. Smaller crystals grow mainly as simple prisms, which makes the growth much easier to model theoretically. After some roundabout mathematical machinations we've been able to infer the growth rates of the different ice surfaces as a function of temperature, supersaturation, and other conditions. And as you



The world's largest stellar dendrite? The crystal above grew to a diameter of more than 1 inch in 1.5 hours. Right: a hollow column, looking like a small shot glass, grows on the end of an ice needle.

Ice crystals, left, grown at controlled temperatures and humidities in a crystal growth chamber, show the same variations in size and shape predicted by the morphology diagram on the opposite page. From the top, with two examples for each group: small plates formed at –2 °C with low supersaturation; hollow and filled columns at –5 °C with moderately high supersaturation; large plates and stellar dendrites at –15 °C with high supersaturation; and small, fairly thick plates at –30 °C with low supersaturation.

might expect from looking at the morphology diagram, the results are a bit odd.

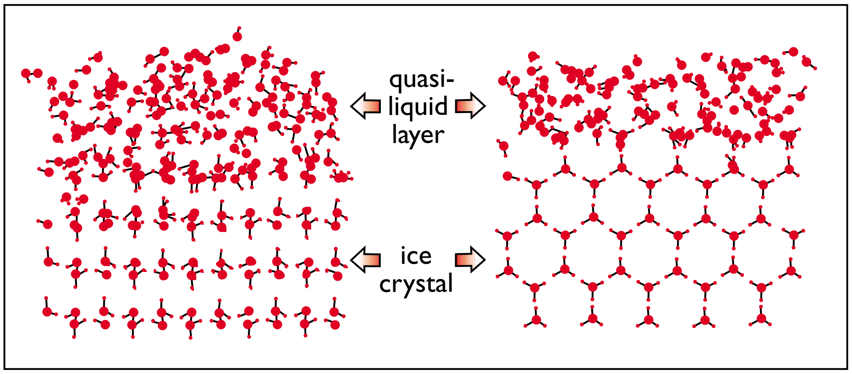
The growth rates of the basal and prism facets depend strongly on supersaturation as well as on temperature, and some recent measurements show that the rates even depend strongly on the background gas in which the crystals are grown. Snow crystals grown in pure water-vapor conditions (i.e., in a vacuum chamber containing no gases other than water vapor) do not grow into the same thin plates or long needles that we see in air, but rather into more nearly isometric simple prisms. Air is relatively inert, and helium gas is even more so, but both nevertheless slow the surface growth rates substantially, even after factoring out the effects of water-vapor diffusion.

We still don't understand why ice does all of what it does, which is why we continue studying it. But one feature of ice that almost certainly plays a big role in its crystal growth is a phenomenon called surface melting. For any crystal, the surface molecules are not as tightly bound as the molecules deep inside, since surface molecules don't have so many neighbors to hang on to. Thus, the surface molecules can sometimes jostle loose while still remaining attached to the solid, forming what's called a quasiliquid layer (see diagram). This layer disappears at very low temperatures, when it's so cold that the molecules don't have much jostle in them, and gets thicker at higher temperatures, eventually increasing to effectively infinite thickness at the melting point. Surface melting is present on a large variety of materials, but the physics of this phenomenon is particularly puzzling for high-vapor-pressure materials like ice.

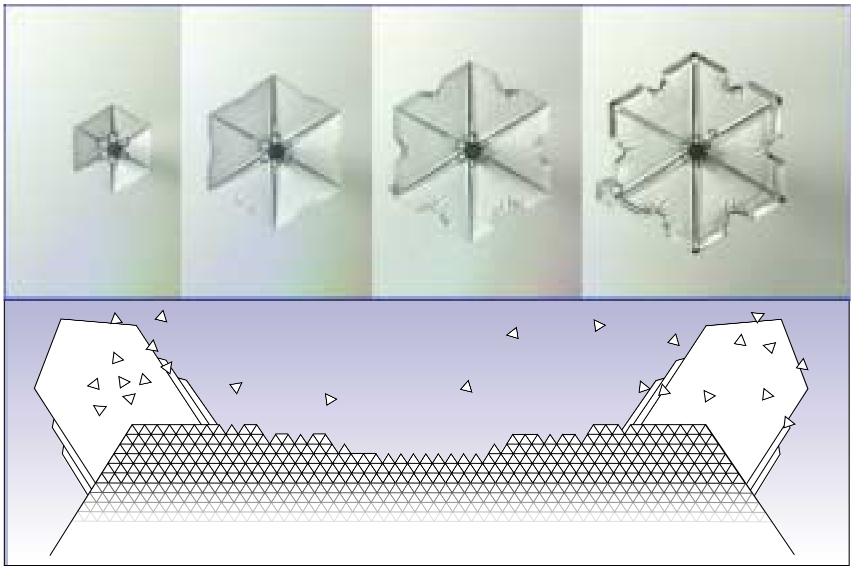
The quasiliquid layer on ice has been studied for a long time (it was first proposed by Faraday in the 1850s to explain some unusual properties of ice), and it seems to play a role in many disparate environmental phenomena. Lightning, for example, is known to arise from collisions between ice particles in clouds—tropical clouds with no ice seldom produce lightning. The charge transfer during these collisions is thought to depend on the details of the ice surface structure, and thus on the quasiliquid layer. On a different front, some of the chemistry that takes place in high-altitude clouds, and which is important for ozone depletion, also occurs on the surface of ice crystals, and is affected by the presence of a quasiliquid layer. And back down on earth, it's the quasiliquid layer that makes ice especially slippery and helps snow stick together in snowballs.

The quasiliquid layer can affect crystal growth in a number of ways. At low temperatures, when the layer is gone, the facet surface can be very smooth, and so grows slowly. It's said that the surface has a high nucleation barrier, since molecules on the smooth surface don't have much to hold on to. At higher temperatures, however, when the quasiliquid layer first starts to form on a facet surface, molecules begin to jostle loose, so that the surface cannot be so smooth any more, and the growth rate shoots up. But then, at still higher temperatures, the quasiliquid layer is so thick it starts to look like a real liquid, and the solid/quasiliquid boundary can itself become smooth, like that between a solid and its melt. Thus there is a nucleation barrier at the solid/quasiliquid interface, and the growth rate goes down again. There's considerably more to the story, but you get the idea.

When you consider that the properties of the quasiliquid layer, and how they depend on temperature, can be different for the different ice facets, it becomes plausible that ice growth shows the unusual characteristics that it does. But the devil is in the details, and our current research is



A quasiliquid layer forms on the surface of ice when water molecules at the edges of the crystal jostle loose from the tightly bound inner lattice. The thickness of this layer depends on temperature and is generally different for the basal (left) and prism (right) crystal faces.



How snow crystals get their arms. This photo series of a growing crystal shows how a hexagonal sectored plate develops six distinct arms. The hexagonal growth is eventually unstable because the corners of the plate stick farther out into the supersaturated air and thus collect more water molecules. On the nanoscale, below, what appears to be a straight facet actually contains many molecular steps. Diffusing molecules are more likely to hit the corners, but molecules are more likely to stick where the surface is rough, in between the corners.

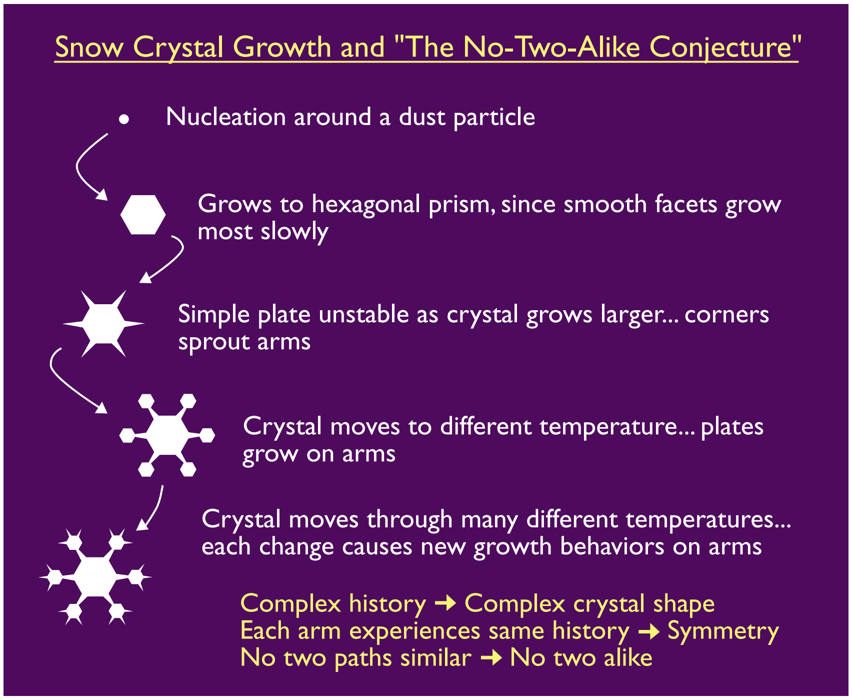
aimed at reaching a better understanding of just what's going on. This is all pretty basic stuff—the structure of ice surfaces, and how it affects crystal growth—but we are surprisingly ignorant of even these fundamental issues, not only for ice but also for most real surfaces.

The second big factor affecting snow crystal shapes is the diffusion of water molecules through the air, and here at least we know the underlying physics fairly well. A growing crystal captures water molecules from the air right around the crystal, and additional water molecules have to make their way through the surrounding air molecules to reach the surface. Diffusion tends to *hinder* crystal growth, and this produces what turns out to be a fairly common effect, known as the Mullins-Sekerka instability, that goes far in explaining why snow crystals grow into such intricate structures.

If you imagine that our newborn snow crystal is a hexagonal plate, you can see that the corners of the crystal stick out a bit farther into the air than the center faces in between, making it a bit easier for water molecules to diffuse to them. The corners should therefore grow a bit faster than the rest of the crystal surface. However, as the corners grow, they leave molecular steps behind them, which make the center faces a bit rougher than the smooth corners (see diagram above).

Since rough surfaces grow more quickly than smooth ones, the overall result is that the center faces grow just as fast as the corners, so that the

The growing snow star on the opposite page was created by cycling it between –15 °C, to develop dendritic spikes, and –12 °C, to develop sectored plates, mirroring the way natural snow crystals form complex shapes as they move through different temperature zones in a snow cloud, right.



Snow Crystal Growth and "The No-Two-Alike Conjecture"

Nucleation around a dust particle

Grows to hexagonal prism, since smooth facets grow most slowly

Simple plate unstable as crystal grows larger... corners sprout arms

Crystal moves to different temperature... plates grow on arms

Crystal moves through many different temperatures... each change causes new growth behaviors on arms

Complex history → Complex crystal shape
Each arm experiences same history → Symmetry
No two paths similar → No two alike

A normal ice dendrite (right) grew at a steady 3 microns per second until the application of 1400 volts. Then the growth developed into a thin needle morphology (inset). These electric needles have been clocked at up to a staggering 200 microns per second.

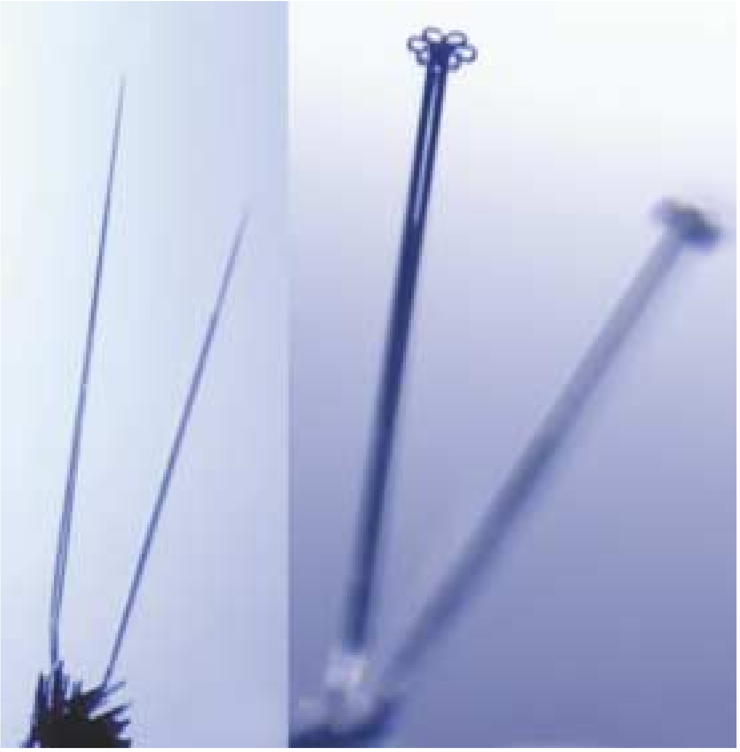
overall shape of the crystal doesn't change as it grows. And since the molecular steps are very tiny, the crystal continues to look like a simple hexagonal prism as it grows.

This goes well for a while, but to continue this shape-preserving growth, the center faces become rougher and rougher as the crystal grows larger, until eventually they become so rough that they can't be any rougher. At this point the center faces can no longer keep up, and the corners really do grow faster, with the end result that our hexagonal snow crystal sprouts six tiny arms. As the arms themselves grow longer, they too come under the influence of the Mullins-Sekerka instability, and can sprout their own side branches. In the end we find that a very complex crystal shape results from a rather simple physical phenomenon. Growth instabilities like these are quite common in nature, and are responsible for a great deal of pattern formation, in the biological world as well as the physical world.

With these tools, we can now can pretty much explain why there's such a rich variety of symmetrical snow-crystal shapes in nature, or what I like to call the "no-two-alike conjecture." After a snow crystal is born, it quickly grows into a small hexagonal prism, with the facet surfaces growing more slowly than the other surfaces. Then, as the crystal grows larger, the Mullins-Sekerka instability often kicks in, causing the corners to sprout arms. Exactly how fast the arms grow depends on the local temperature and supersaturation experienced by the crystal, since, as we've seen from the morphology diagram, snow-crystal growth is extremely sensitive to environmental conditions.

It's important to note that the local conditions are essentially the same for each arm on a tiny snow crystal, so that the arms all grow in the same way. Then, as this growing crystal travels through regions of the atmosphere with different temperatures and supersaturations, its growth will change with the conditions, and the arms will all change their growth in unison. The final crystal shape can be very complex, reflecting the complex path the crystal followed through the atmosphere. Yet since the arms all grow at the same time, they tend to look alike, so that the crystal has a sixfold symmetric appearance. And since no two crystals follow exactly the same path through the sky as they fall, each grows into a slightly different shape. So we end up with a myriad of complex, symmetric patterns, with no two alike. In principle, one could look at a snow crystal and decipher the conditions under which it grew. Or, to quote Nakaya, "Snow crystals are hieroglyphs sent from the sky."

Another area we've been exploring in my lab is that of using high electric fields to enhance and control the dendritic instabilities in crystal growth. We started by growing normal dendritic crystals like the one in the photo above, which are very easy to grow at –15 °C when the supersaturation is high. The tip of such a crystal grows at a steady 3–4 microns per second, and side branches appear in a semiregular pattern as dictated by the Mullins-Sekerka instability. Dendritic growth like this has been studied for quite a while (it has applications in metallurgy and the structure of alloys), and we know how to calculate the simpler properties of the growth, such as the steady-state tip velocity, which is the rate at which the tip elongates.

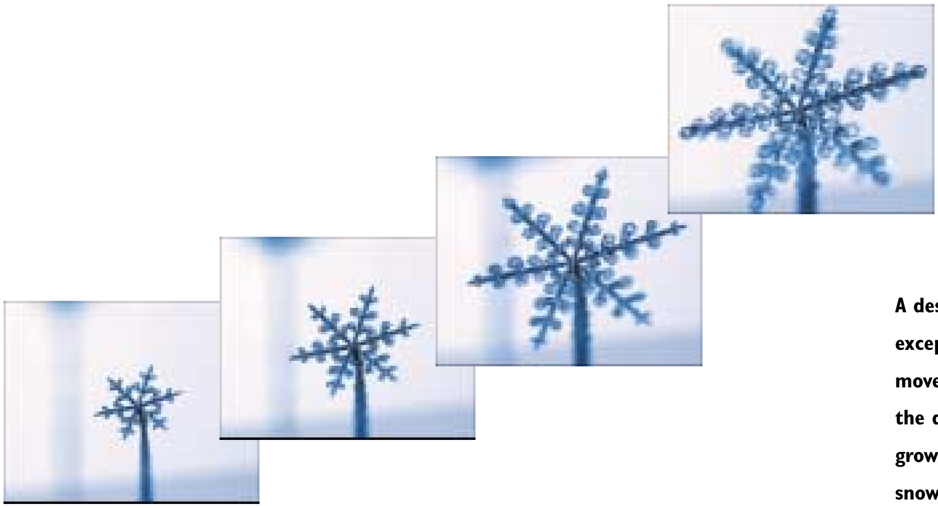


Electric needles, left, were grown at −5 °C with the application of 2000 volts and chemical vapor additives to induce growth along the c-axis of the crystal. When the voltage was removed and the needles moved to −15 °C, plates grew. Right: needle growth at −5 °C.

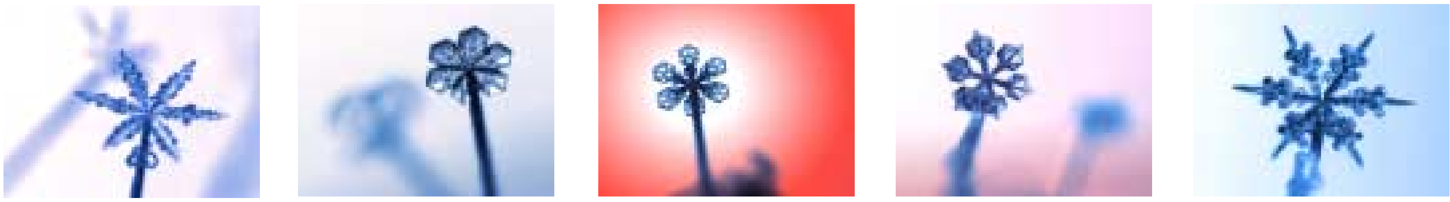
To make life a bit more interesting, we began growing these crystals on the end of a wire connected to a high-voltage power supply. As we turned on the voltage, the electric field near the tip of the growing crystal polarized the water molecules in the surrounding air, and these polarized molecules were attracted to the tip. This enhanced the normal diffusion rate, and the tip velocity increased accordingly—all nicely according to theory. But when the tip velocity was about twice as fast as normal, whoosh! The tip just took off, growing as much as 200 microns per second—greased lightning in the crystal-growth game! The crystals grew into long, thin needles, with tip radii as small as 100 nanometers.

The theory we developed to explain the physics of all this suggested that electrically enhanced growth would work even with nonpolar molecules (water is highly polar), and we recently demonstrated this in the lab by growing iodine needle crystals. Metallic needle crystals have also been grown from certain metal-carbonyl vapors using related techniques, and it's tempting to speculate that diamond needles could be produced one day using electrically enhanced growth (although I suppose it's too late to corner the market in phonograph needles).

In our lab, we have a special fondness for growing "c-axis" electric needles of ice (think tiny ice versions of a standard hexagonal pencil). For a long time we found that growing this particular type of needle crystal was a hit-or-miss proposition; some days it worked well, some days it didn't work at all. At first we thought we were having contamination problems from solvent vapors floating around in our chamber (the laboratory version of air pollution), so we decided to give the chamber a good, long bake to clean it out. To our surprise, when we started our experiments again, we couldn't get any good needles at all—not one. It turned out we weren't being *hurt* by contamination in the chamber; we were actually being *helped*



A designer snowflake. This crystal was grown at −14 °C, except that at approximately periodic intervals it was moved to −7 °C for just a few seconds. Each move induced the development of side branches at the tips of the growing arms. A movie of this can be seen at snowcrystals.net.



A gallery of designer snow stars—tiny ice flowers on thin ice stems.

PICTURE CREDITS:
10, 14–19 — Ken Libbrecht; 11 — Mark Garcia; 10–11, 12, 14–16 — Doug Cummings

by it. So we tried adding various solvent vapors, at concentrations of only a few parts per million in the air, and *Voilà*!—we were soon producing beautiful c-axis needles again, very reliably. Acetic acid, the main ingredient in vinegar, seems to work especially well, and we're still puzzling over exactly what physical and chemical mechanisms are controlling the needle growth.

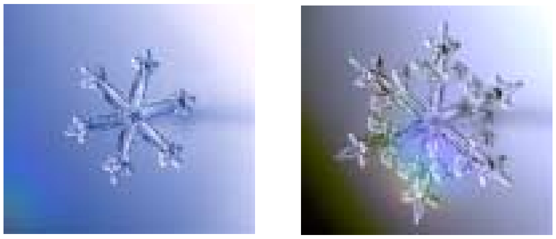
A wonderful feature of c-axis needles is that after growing a long ice needle, we can turn off the voltage and grow a normal platelike snow crystal on its end. Most of this work is done in our vertical diffusion chamber with a convenient temperature gradient—warm at the top, cold at the bottom. By moving the needle from –5 °C, where the c-axis needles grow best, to –15 °C, where plates grow best, we can grow a beautiful snow crystal on the end of the needle, like a tiny ice flower on a thin ice stem.

We've developed these techniques to the point where we can now create "designer" snow crystals, growing shapes of our own choosing by controlling the humidity and temperature of the growth, and some examples are shown in the photos on these pages. It's something of a new art form—miniature ice sculpture. Instead of cutting away material, we design and fabricate using the natural rules of pattern formation. Stellar dendrites, sectored plates, hollow columns can all be made relatively easily, and now we're even learning to control side branching and other features to create more complex and unusual snow-crystal shapes.

So we see that snow-crystal growth is governed by some sophisticated physics, mathematics, and chemistry, all working in concert to create these tiny, filigreed ice sculptures that fall down from the sky. Snow crystals are not only beautiful to look at, but they also teach us about surface physics, the statistical mechanics of crystal growth, and the intricacies of pattern formation processes in nature. ☐

Growing up in Fargo, North Dakota—where it's said there's 10 months of snow and two months of poor sledding—may have inspired Ken Libbrecht's current interest in snow crystals, but for most of his early career he worked at the other end of the temperature spectrum, investigating the internal structure of the sun. A Caltech graduate (BS '80), he returned to join the faculty in 1984, and has been Executive Officer for Physics since 1997. Nowadays, as well as investigating the physics of snow crystals, Libbrecht is involved in advanced detector development for LIGO, the Laser Interferometer Gravitational-wave Observatory. Check out his award-winning Web site, snowcrystals.net, for much, much more on snow crystals, including how to make your own in the kitchen, and how to help locate the snow crystal capital of the world. This article is adapted from a Watson lecture given in January 2001.

Right, two photos of a growing stellar dendrite, taken 5 minutes apart, show the development of side branching. The ice needle on which this c-axis star was grown is hidden from view below it.



A Billion Points of Heat

by Jane Dietrich

Infrared radiation—that segment of the electromagnetic spectrum with wavelengths longer than visible light, between visible light and microwaves—can pierce through the dusty universe to disclose quantities of unseen stars and galaxies; it reveals cool stars that radiate heat but no visible light, failed stars, and stars just being born in gaseous clouds. Although water vapor and carbon dioxide in the atmosphere absorb much of the infrared, satellites above the atmosphere, such as the Hubble Space Telescope and the Infrared Astronomical Satellite (IRAS), have extended our sight spectacularly. But there's also one window, at about 2 microns in what is called the near infrared (the infrared extends from wavelengths of 1 to 300 microns), where waves can pass relatively easily through the atmosphere and be detected from the ground. And at 2 microns, you're seeing light given off directly by stars, whereas at longer infrared wavelengths you see starlight that has been absorbed and reradiated by dust glowing in the space between stars.

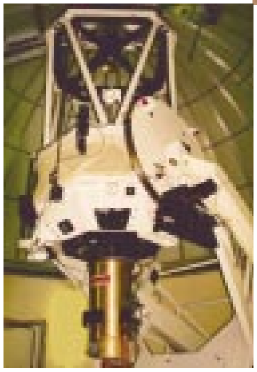
Astronomers first surveyed the sky through the 2-micron window in the 1960s. But dramatic advances in detector technology in the last 20 years have made it possible to detect objects more than 80,000 times fainter than those discovered by the first 2-micron sky survey, and in the 1990s the National Aeronautics and Space Administration (NASA) and the National Science Foundation (NSF) funded the Two Micron All Sky Survey (2MASS), a collaboration between the University of Massachusetts and Caltech, that just finished gathering its data in February 2001.

The first survey was also a Caltech undertaking. Back in the early 1960s, Bob Leighton and Gerry Neugebauer aimed a telescope equipped with infrared detectors to scan as much of the sky as could be seen from Mount Wilson—about 70 percent of it. Leighton designed the 62-inch telescope's epoxy reflecting dish based on a principle he had first observed in his mother's mop bucket as a child, and built it in the back of his office (see *E&S,* No. 4, 1998). The infrared detectors were samples donated to Neugebauer by a friend in the defense industry; they had been developed after World War II for the heat-seeking guidance system of the Sidewinder missile. Their Two-Micron Sky Survey (TMSS), published in 1969, noted 20,000 infrared sources and cataloged about 5,700 previously unseen celestial objects.

"We thought it was a fun thing to do," says Neugebauer, now the Robert Andrews Millikan Professor of Physics, Emeritus. (Leighton, BS '40, PhD '47, the Valentine Professor of Physics, Emeritus, died in 1997.) In an era when infrared

Left: The Orion Nebula, or Messier 42, one of 2MASS's greatest hits (which required only 10 minutes of observing time). The Trapezium Cluster in the center contains more than 3,000 bright, hot stars, the densest concentration of young stars in our solar neighborhood. Even younger, protostellar, objects appear in the small red region at the top. The bright, wispy clouds come from molecular hydrogen in dust-scattered starlight, detectable at two microns.

In April 1964, Bob Leighton's 62-inch two-micron infrared telescope (right) prepares to move into its "dome" on Mount Wilson, which is shown under construction in September 1963 at far right. The carpenter on the left is Jerry Nelson, '65, who later went on to design the 10-meter Keck Telescope's segmented mirror.



Top: The 2MASS 1.3-meter telescope in the southern hemisphere sits high in the Andes at the Cerro Tololo Interamerican Observatory. Inside the dome (above) the camera, with its massive arrays of detectors and filters for three wavelengths, is attached to the back of the telescope.

astronomy was starting to grow rapidly, they also "happened to be there first," according to Leighton's oral history. Leighton's telescope is now a piece of history enshrined in the Smithsonian Institution.

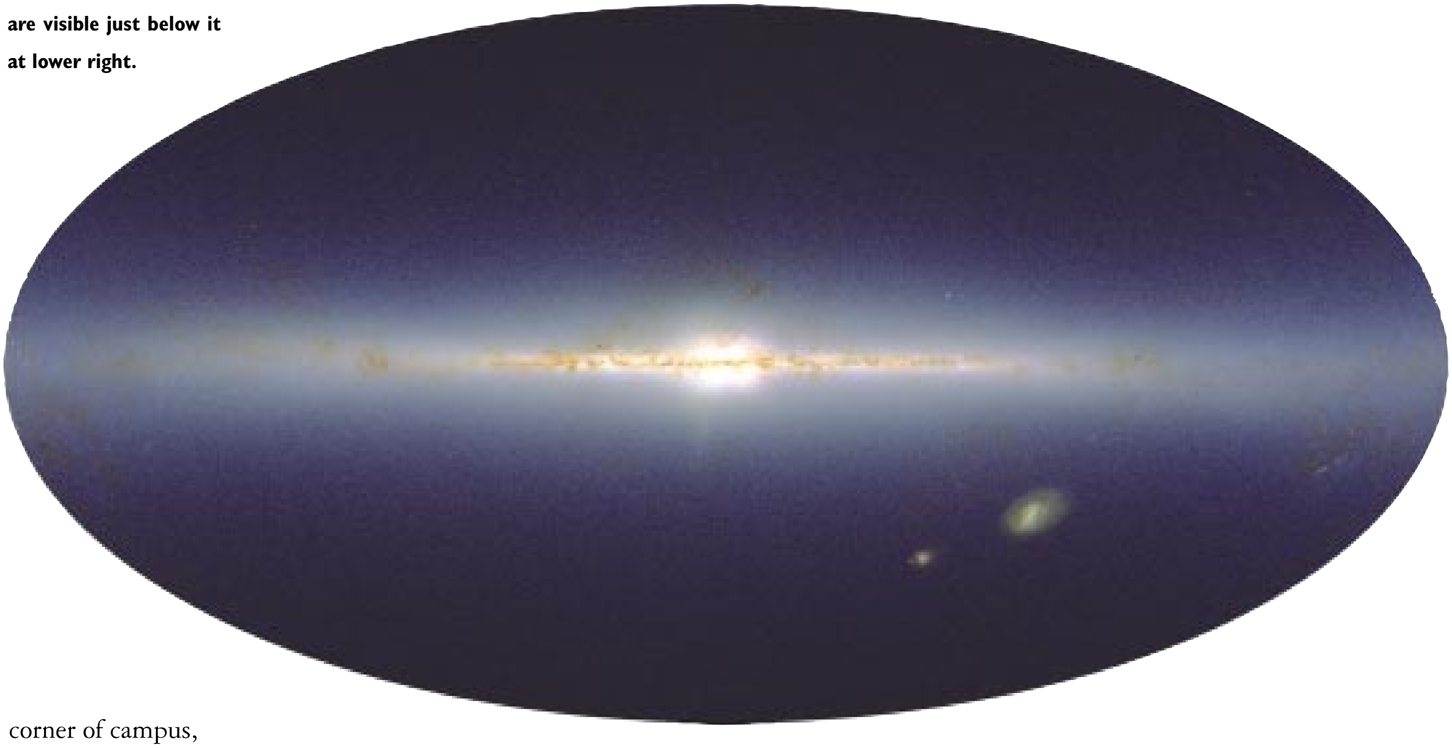
Infrared astronomy has long since left behind seat-of-the-pants technology and missile leftovers. 2MASS, says Neugebauer, "is everyone's dream of how you should do something, really do it right." The completely automated twin telescopes—one for each hemisphere—were designed and constructed by the University of Massachusetts team under astronomer Michael Skrutskie, principal investigator of the sky survey. At 51 inches (1.3 meters) they're a bit smaller than Leighton's telescope; smaller telescopes with large fields of view are better suited for covering the larger swaths of the heavens needed for a survey of the whole sky; huge mirrors like the 10-meter Keck Telescope, which gather faint light from the edges of the universe, peer through a pinhole in comparison. The northern-hemisphere instrument, on 8,550-foot Mount Hopkins (even at the 2-micron window, less atmosphere is better) in Arizona, started sweeping the sky in June 1997 and finished its observations in November 2000; its twin at the Cerro Tololo Interamerican Observatory, at about 7,000 feet in the Chilean Andes, started up in March 1998 and sent its last data to Caltech this past February. Each telescope's camera scanned strips of sky running 8.5 arcminutes wide and 6 degrees long before moving on to the next strip, or "tile." The whole sky is tiled by 59,650 such strips.

The heart and soul of the telescopes are their sensitive detectors, which represent the greatest leap in technology. Where the Mount Wilson model contained eight detector elements, a 2MASS telescope camera sports three arrays of 64,000 such elements, each with a filter for a particular wavelength. Like the earlier telescope, the new arrays are also a legacy, if not an actual hand-me-down, of defense research (for sensing heat from Earth rather than the sky). In their transfer from the military to astronomy, such infrared detectors made possible the Hubble Space Telescope's spectacular high-resolution images of infant stars forming in glowing, billowing clouds of space dust.

Each telescope has been producing a 20-gigabyte tape per night, and once a week for the past four years a stack of these digital linear tapes has been arriving (via Fed Ex) at the Morrisroe Astroscience Laboratory, home of Caltech's Infrared Processing and Analysis Center, or IPAC. Altogether there are hundreds of tapes, amounting to 25 terabytes (10^{12} bytes) of data. For comparison, the photographic plates of the 1998 Palomar optical sky survey (this was also the second one, the first having been done in the 1950s), when digitized, comprised 3 terabytes of data, already several hundred times more than what was collected by IRAS, which was at that time one of the largest collections of data ever. How do you take such an incomprehensibly huge data set and turn it into something comprehensible: high-resolution images and an accurate, reliable catalog of half a billion objects? You turn it over to IPAC, which was created to analyze the IRAS data.

Caltech had stayed involved in infrared astronomy after Leighton and Neugebauer's pioneering work. Neugebauer, in fact, was U.S. cochairman of the joint science working group (which also included scientists from the Netherlands and the United Kingdom) for IRAS, which was launched in 1983 and collected data for 10 months. IRAS was managed for NASA by the Jet Propulsion Laboratory, but when it came to analyzing the mountains of data—a task that was to take several years—it became more convenient to move the project to campus as IPAC because of JPL's access restrictions, especially for noncitizens, and because of the great interest on campus in the data and their scientific interpretation. In the southwest

Below: The whole sky, as seen in a 2MASS three-color, composite image. The Milky Way, edge-on, stretches almost all the way across, its center the bright bulge. Filaments of dust cut through the galactic plane, while the Magellanic Clouds, our nearest galaxy neighbors, are visible just below it at lower right.



corner of campus, behind Braun gym, and lying low to appease its residential neighbors along Arden Road, the astroscience laboratory has been in operation since 1986, largely a mystery to the rest of campus.

"IPAC is unique as a piece of campus with special connections to JPL," says George Helou, IPAC's director. Although the funding comes from NASA, and the lab carries a bit of JPL "flavor," all of the lab's approximately 80 staff members are employees of the Division of Physics, Mathematics and Astronomy. "IPAC straddles the two cultures," Helou says, "the science and research culture of campus and the project and engineering culture of JPL. And 2MASS combines the science and engineering in unique ways."

2MASS is the first digital, electronic survey of the sky at relatively high resolution. While the recent Palomar Sky Survey turned to JPL's Artificial Intelligence group to develop a program to turn its photographic images into computer-digestible data for cataloging, 2MASS was born digital. What this project needed was robust software to turn its terabytes into images, as well as into useful and reliable catalog information. The task fell to Roc Cutri, who, as task leader as well as project scientist, actually does live in both cultures. Cutri oversaw the design and implementation of the automated software pipeline, called 2MAPPS, which stands for 2MASS Production Processing System. To convert the raw digital data into calibrated images in three colors and to extract the quantitative information of the position and brightness of each object, Cutri and his team wrote 300,000 lines of computer code—not the usual pastime for an astronomer, but intellectually challenging all the same, says Cutri.

The three colors represent three different wavelengths in the near infrared: the J band at a wavelength of 1.25 microns, the H at 1.65 microns, and the K_s at 2.17. (The infrared's photometric "bands"—the natural windows astronomers make use of—follow a rather loose alphabetical scheme; the longer wavelengths after K continue in a more regular fashion: LMNOP.) Classes of astronomical objects can vary in brightness at the different wavelengths, so simultaneously observing them at all three wavelengths makes identifying and interpreting them easier, says Cutri.

Below: The Mount Hopkins telescope begins its night's work surveying the northern sky.

Right: In the Flame Nebula, NGC 2024, part of the Orion Molecular Cloud Complex, the near infrared reveals a dense stellar cluster, comprising stars likely surrounded by the accretion disks that could be sites of where planets form.

Below: The Cat's Paw, or Bear Claw, Nebula (NGC 6334).

But before the identifying and interpreting begin, the data must be fed through the software pipeline. While the telescopes finished their observations this past February, it will take six months to a year to finish crunching through the raw data to produce the visual and quantitative results. The full set of images—the whole sky—and catalog will be published next spring. Analyzing the data will take decades; Cutri reckons there's enough fodder for a 50-year legacy. "2MASS offers the statistical context to study a very large number of objects, and also the detail to follow up the interesting ones," he says.

In the meantime, half the sky has been accessible on line since March 2000, the single largest collection of astronomical data from a NASA mission. The final catalog will contain half a billion objects. The team actually tallied one and a half billion sources, but have winnowed down the number to those whose brightness and position can be guaranteed with great precision. "Astronomers need to be able to depend on the accuracy of the data," says Cutri, who as project scientist is responsible for ensuring the scientific quality of the data. He compares it to a phone book, which would be useless if one out of ten digits were wrong. With 2MASS, extraterrestrials could phone home—confidently—to any of nearly half a billion stars.

The best way to insure the scientific quality of the data is "to do the science yourself," according to Cutri. He was a collaborator in one of the early discoveries to come out of 2MASS: brown dwarfs, cooler than any ever before seen. Brown dwarfs are stars that never ignited, and therefore give off no visible light. They do give off some heat, however, which can be detected in the infrared. Already in 1998, in only one percent of what would be the total data, a team of 2MASS researchers led by Davy Kirkpatrick at IPAC found a number of these objects, which they named L dwarfs, revising a century-old alphabetic system

Emissions from the planetary nebula NGC 3132 (above) are mostly in the K_s band, which gives the ring around the wind-blown center its reddish color. In the center is a low-mass star like our sun, losing its outer envelope as it dies.

Top left: The remnant of a supernova (IC443) exhibits a bright blue arc of excited iron in the J band and a red ribbon of hydrogen in the K_s band.

Top right: 2MASS reveals the nearby spiral galaxy Maffei 2, which lies in the "zone of avoidance" and was previously almost lost in the dust.

(even more mysterious than the designation of wavelengths) in which stars are ranked from hottest to coolest as OBAFGKM—with L now at the end). Then came Caltech grad student Adam Burgasser's discovery of still cooler, fainter, methane-rich brown objects called T dwarfs. (Make that OBAFGKMLT.) Together, L and T dwarfs, scarcely bigger than Jupiter and collapsing under their own weight, are probably the most populous stellar objects in our galaxy, outnumbering real stars two to one, and some of them may also be our solar system's nearest neighbors.

Because the Milky Way itself contains so much dust, it has been heretofore impossible to see neighboring galaxies in the wonderfully named "zone of avoidance," which extends to 30 degrees on either side of the galactic plane. But 2MASS has cut through that murky zone and revealed many more galaxies in our local universe. It has also produced a high-resolution map of our galaxy, including the galactic center, which at other wavelengths is obscured by dust—and an exquisite census of the Milky Way.

One of the challenges of such an immense data set, says Helou, "is how to search through a billion sources to find the really interesting objects—the unusual stars and quasars." So perhaps the greatest importance of 2MASS lies in overlapping it with surveys at other wavelengths (for example the very short X rays on one end of the spectrum and the much longer radio waves on the other), the "synergy," as Cutri describes it, "of putting data sets together. The value of combined surveys goes far beyond that of a single one."

Postdoc Robert Brunner agrees. "The sum is greater than the parts," says Brunner. "You gain more from data by joining them together than using them independently." As project scientist for the Digital Sky Project, he's trying to mesh a bunch of surveys at different wavelengths—a computationally challenging task matching billions of sources at different resolutions all over the sky. He and his group are creating software to improve the accuracy of the matches. Brunner is also currently working with Microsoft Research on a Sky Server to merge data from 2MASS, the Palomar Sky Survey, and other surveys, approaching the problems less from a computer-science angle and more from the viewpoint of the scientific user. "We're starting to move on to something of service to the broader community." Ultimately, the information-rich dataset of these combined digital surveys will be accessible on line in a National Virtual Observatory.

Sometimes the overlap hasn't been foreseen. Richard Ellis, professor of astronomy and director of Palomar Observatory, is part of an international team conducting the 2-degree-Field Galaxy Redshift Survey (2dF) of a swath of the southern sky, using a novel instrument built by Keith Taylor (now a member of the professional staff at Caltech) at the Anglo-Australian Telescope near Coonabarabran, Australia. The team is seeking to measure the stellar population density and mass distribution of galaxies on large scales. They mapped the positions of 170,000 galaxies, but optical wavelengths cannot accurately discern the stellar content of galaxies. Infrared radiation provides a much better handle on how many stars are in a galaxy, because the infrared output is directly proportional to the number of stars. But until 2MASS, no infrared telescope had seen deep enough into space to get a fair sample. Combining the digital catalogs of both surveys (a process that was accomplished literally overnight) matched redshifts/distances from 2dF with 2MASS luminosity/population data for 17,000 galaxies. With that large a sample Ellis and his group could determine how many stars there are in the universe per unit volume of space, and concluded that the amount of mass in stars is only about 0.2 percent of the total mass needed to stop cosmic expansion. "The overlap led to the most accurate census of stars in the local universe," says Ellis,

The infrared shows up a bright core in the radio source Cepheus A. Its massive young stars and molecular gas are completely invisible to optical wavelengths.

Opposite page: The Carina, or Keyhole, Nebula (NGC 3372) contains an unusually high concentration of young massive stars and stars that are still forming.

"an achievement that neither survey ever imagined when it was originally conceived."

"It's gratifying to develop a product that other scientists can use," says Cutri. And not just scientists. 2MASS can turn anyone's computer into a "desktop observatory" through IPAC's Infrared Science Archive (IRSA), the Web-based system for accessing the databases. Besides the catalog of data, about 5 million images in the digital Image Atlas will also be available on line. Just enter the name of your favorite galaxy or pick a point on the sky. You can start off at http://www.ipac.caltech.edu/2mass/ and be sure to catch "2MASS Galactic Center: The Movie" on the home page. This is not its only role in show business: the *Star Trek Voyager* TV series has also snapped up some 2MASS images for space backdrops.

The 2MASS team in the galactic center. Front row, from left: Schuyler Van Dyk, Diane Engler, Ron Beck, Eugene Kopan, Roc Cutri, Tracey Evans, William Wheaton, Robert Hurt, Sherry Wheelock, and Jeonghee Rho. Back row: Ken Marsh, Cong Xu, Howard McCallon, Tom Jarrett, Laurent Cambresy, J. Davy Kirkpatrick, John Gizis, and Raymond Tam. Not pictured: Brant Nelson, Helene Huynh, Tom Chester, and John Fowler.

A topical session ("The Big Picture: Latest Science Results from 2MASS") at the American Astronomical Society's meeting in early June, held conveniently in Pasadena, highlighted the already substantial science that is emerging from the data on half the sky. The session spanned a range of topics including—besides brown dwarfs—asteroids, stellar populations in the Milky Way and the Large Magellanic Cloud (its nearest neighboring galaxy), other nearby galaxies, distant active galactic nuclei, and the cosmic near-infrared background radiation. As of the beginning of June, 174 papers had been published using 2MASS data.

Funding for 2MASS will continue for another year and a half, when the project will deliver its final products and then fold, says IPAC's Helou. "Then we'll find something else interesting and worth doing, something with a specifically Caltech point of interest." Already well under way is SIRTF, the Space InfraRed Telescope Facility, a satellite to be launched in July 2002 to observe the sky at wavelengths from 3 to 180 microns. 2MASS data are critical in laying the groundwork for the mission, which is administered by JPL for NASA. There's considerable Caltech scientific interest in this new venture, and the SIRTF Science Center will be located on campus, sharing the Keith Spaulding building with the remaining segments of Business Services that have not moved elsewhere. And IPAC will be analyzing the data. □

The information revolution has had a profound impact on the economy but very little impact on economic policy, which is largely still generated by 19th-century ideas.... Having a 21st-century economy based on laws designed for the manufacturing sector is really quite ridiculous.

Economic Policy in the Information Age

by Simon J. Wilkie

The standard economic model, in which many buyers and many sellers compete with one another in the marketplace, works well for tangible goods such as bananas. But you get into deep trouble when you try to apply it to network commodities such as electricity.

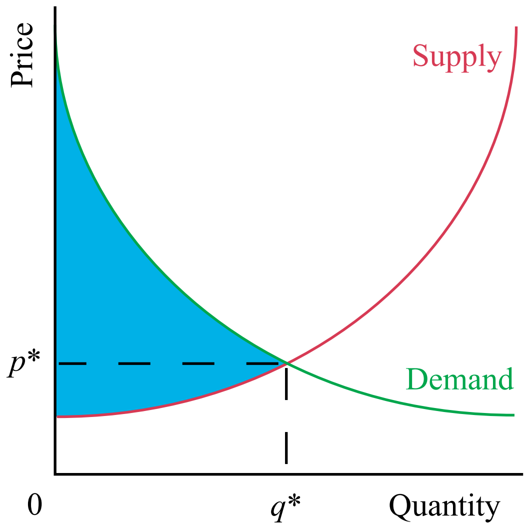
The United States is in the throes of the third industrial revolution—the first one, of course, being the harnessing of mechanical power, and the second one being the harnessing of electrical power. This one is the harnessing of information, and has been sparked by the biggest capital investment in the history of humankind. The U.S. spent roughly $4 trillion on information technology, broadly defined, from 1960 to 1994, and it's expected that we'll be spending a trillion a year by 2005, even with the current slowdown. Like the other industrial revolutions, it has taken 30 to 50 years for the results to show up. The productivity gains we've seen in the economy in the last few years have their origins in the early investments that are just now starting to kick in.

The information revolution has had a profound impact on the economy but very little impact on economic policy, which is still largely generated by 19th-century ideas. The legal framework was set by the Sherman Act and the Clayton Act, which were developed in the 1890s to bust trusts such as Standard Oil. The notion of regulating utilities appeared in the early part of the 20th century, culminating in the Telephone Regulation Act of 1933, which established AT&T as a monopoly. Manufacturing, however, which was the dominant paradigm at the turn of the century, is now less than 17 percent of our economy. It's going the way of agriculture. Health care is now almost 15 percent of the economy, and in a couple of years, it's going to be bigger than manufacturing. Having a 21st-century economy based on laws designed for the manufacturing sector is really quite ridiculous.

There are two main points I want to make. The first is that minutiae are important in the design of economic institutions—that is, the details matter. And they matter a lot. Which is kind of ironic because most economic-policy debate is big-think debate: should we have a market or not? Should an industry be regulated or unregulated? These are the wrong questions. The important questions are really in the small details of how a market is structured. Second, I want to make a case for the fundamental importance of abstract economic theory—the type of arcane research, divorced from the real world, that we do here at Caltech. If you don't pay attention to these extremely mathematical, abstract models, you're bound to make disastrous policy mistakes. So minutiae are important, and the boring questions are really the interesting ones. I'm going to apply this perspective to the Microsoft antitrust case and to California's electricity-deregulation debacle, which is an endless source of fun until your bill comes at the end of the month.

To see where we need to go, we first need to know where we are. The standard economic model says we have a market in which many buyers and many sellers compete with one another. Each buyer has a personal valuation for each good. When you buy bananas, you know how much you're willing to pay for any given quantity. And each seller knows his or her production costs, and sets an individual minimum price accordingly. The market adjusts supply to be equal to demand, as that old axiom says. This is summed up on the next page, in the economist's favorite diagram. The vertical axis is price, and the horizontal axis is quantity. The downward-sloping demand curve says that the lower the price is, the more of a product people are willing to buy. The supply curve slopes upward, saying that the higher the market price, the larger the quantity that sellers will want to sell. At the intersection lies the equilibrium, which is the quantity (q^*) that is actually traded at the market price (p^*).

This competitive market has several desirable properties, the first of which is efficiency, or **E**. That is, the market maximizes the sum of the valuations minus the sum of the costs. This maximizes social welfare—the greatest good for the greatest number. (However, it doesn't mean



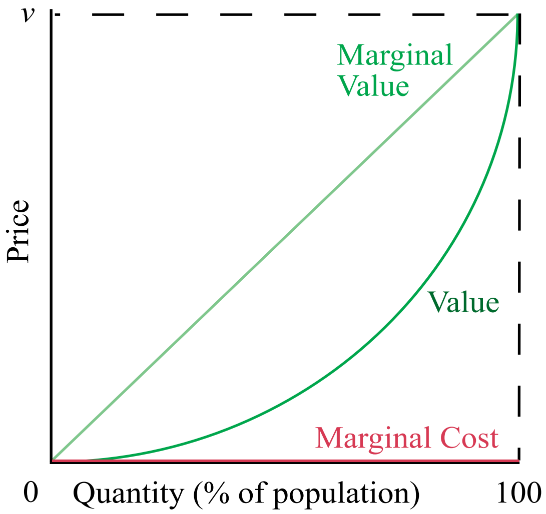
The competitive-market supply and demand curves as drummed into the heads of generations of business majors. The supply curve (more properly, the marginal cost curve) rises while the demand (or marginal value) curve falls. They intersect at point (p^*, q^*), and the shaded region is the social surplus, or value in excess of cost, that people get from trading in the market.

it maximizes *your* welfare. For that, you'd need to do a weighted sum, where your weight was higher than anybody else's.) The distance between the marginal cost (the cost of producing one more unit of the product) and the marginal valuation (the price the consumer is willing to pay for one more unit) is the amount of what we call surplus, or benefit, that people get from trading in the market. Mathematically, the market starts at quantity $q = 0$, and runs until q equals q^*, at which point the market price has dropped to p^* and the sellers are selling at cost, so they quit. But they made a profit on all the previous sales. And at p^* the buyers, who have been buying the product for less than their personal valuations, are paying as much for the last unit of the product as they think it's worth, so they quit. But they got a bargain on all the previous units. So the market is efficient because it maximizes the integral—the shaded region between the two curves—without even knowing what that integral is. Adam Smith discovered the magic of this "invisible hand" a couple of centuries ago.

The market has three other remarkable properties that we don't talk about as much. First, notice that there's no Tony Soprano—we don't need coercion to get people to use the market. Participation is voluntary, because until the market hits equilibrium and shuts down, every buyer leaves with a bargain and every seller leaves with a profit. We call this property **V**. The next property is **B**, for balance. Supply equals demand without needing an infusion of cash or goods external to the system—unlike, say, Russia, which is kept afloat by large amounts of funds flowing in from the World Bank and the United States. Or your kids might trade toys, but that only works so long as there's a perpetual infusion of new toys from the parents. But a normal adult market has the miracle of always being balanced. The last property, **S**, is the most important one—this market is strategy-proof. People have no incentive to game the market. If I go to Von's supermarket to buy three pounds of bananas, and they're a dollar a pound, I have no incentive to buy four pounds, and I don't think, "Ha! I'll fool them and only buy two pounds!" If I want three pounds, I buy three pounds. There is absolutely no benefit to me from strategic behavior. To sum up, the competitive-market model has four really nice properties: it's efficient, it's voluntary, it's balanced, and it's strategy-proof. That's why policymakers tend to be opposed to monopolies and market regulation, which short-circuit the market's functioning.

But the general paradigm for the 21st century is a network-market model. Network markets are ubiquitous—the telephone system, the Internet, and the power grid are obviously networks. But HMOs are actually networks, too. When you join a primary-care physician's group, you're also signing up for the set of specialists affiliated with that group. Banks and ATMs are networks. A country club is a network. When you join the club, you get the use of their facilities; you also get to enjoy, or unenjoy, the company of the other members of the club. In fact, another word for "schmoozing" *is* "networking." The eBay Web site is a network. Network markets work precisely

The telephone system, the Internet, and the power grid are obviously networks. But HMOs are actually networks, too.... A country club is a network. When you join the club, you get the use of their facilities; you also get to enjoy, or unenjoy, the company of the other members of the club. In fact, another word for "schmoozing" *is* "networking."



Things look quite different in a network market. Now the marginal-cost curve is flat, at next to nothing, while the product's value increases as more people buy it. The marginal value, which is the derivative of the value, increases even faster until 100 percent of the population owns the product and its value is maximized. This means that a monopolistic seller can name any price up to the maximum value, *v*, and people will still buy the product. In this case, a monopoly can be efficient, in the economic sense of the word.

because of the mass of users that they attract.

In a network market, neither the supply curve nor the demand curve behaves as expected. A supply curve only slopes upward when the marginal cost, the derivative of the cost function, is increasing. And the marginal cost *does* go up, for traditional commodities. Most people think of the marginal cost as going down as more units are produced, but that's only true up to a point. If I'm growing bananas, I'll cultivate my most fertile land first. As demand grows, I'll use increasingly poorer land, and I'll have to buy more fertilizer and more water to produce a crop. The same is true of steel—if a blast furnace runs around the clock, labor costs will skyrocket. The foundry either has to hire more people or pay massive overtime. So there's actually a dis-economy of scale. But let's think about software for a moment: The marginal cost to Microsoft of me buying an extra copy of their operating system is pretty much zero. They have the fixed cost of developing the product, and then the cost of burning one more CD is virtually nil. That's also true of the Internet—the cost of setting it up was huge, and the cost of adding an extra unit is minuscule by comparison. Most network-structured economies have this fundamental problem that supply tends not to be upward-sloping. They really *do* make it up on volume.

Even more troublesome, the demand curve slopes downward only when the *quality* of the product is inherent and is independent of its *quantity*. But network commodity's quality is systemic; it's not inherent in the commodity itself, as it is in a banana. In the electricity market, if I decide to flip on my air conditioner, it affects the voltage—minusculely, but it affects the quality of the service that everybody else in my neighborhood gets. And a network's value to a user also depends on its quantity, but not on the quantity you buy as an individual—rather, the value depends on the number, or the identity, of users. If I'm the only person with a telephone, it's worthless. Its value to me increases as other people buy it, because then I have more people I can talk to. And if I were the only person in the country club, I wouldn't be willing to pay very much to join it. The value typically increases with the number

> Von Neumann thought he was the smartest person in the world, so he couldn't understand why he kept losing at poker. Being von Neumann, he decided to figure it out, and he developed the theory of strategic interaction between individuals.

of users, but not necessarily. If you have a cable modem, your downloading speed is divided by the number of users who share your connection. Now you have a positive benefit proportional to the total number of users, plus a negative benefit proportional to the number of users who live on your block. The net benefit, if you'll pardon the pun, of having your neighbor in the system may be negative. We call such attributes externalities, because now the product's value is external to the product itself.

So we have to throw our beloved picture out and go to a more abstract framework to analyze network markets. We use game theory, a mathematical technique developed in the 1940s by John von Neumann and Oskar Morgenstern at Princeton. Von Neumann thought he was the smartest person in the world, so he couldn't understand why he kept losing at poker. Being von Neumann, he decided to figure it out, and he developed the theory of strategic interaction between individuals. Game theory lay dormant until the 1970s, when it was seized upon by economists starting to get interested in the economics of information. In fact, a lot of the pioneering work on game theory in economics was done here at Caltech by some of my colleagues—Matt Jackson, John Ledyard, Dick McKelvey, Tom Palfrey (PhD '81), and Charlie Plott. When we model the market as a game, we ask: can we design an economic mechanism, like we would design an engineering device, that has the attributes we want and solves the problem we want to solve? Is it mathematically possible to construct such a thing?

The network math works like this. We have a set of alternatives, which we call **A**'s. For the electricity grid, each **A** would be a possible topology of the network: the capacities of the transmission lines, how they are connected, and so on. The benefit I get from being in this network depends on the choice of **A**—if we're talking about the Internet, I'd like a high-speed connection better than a low-speed connection, for instance—and it depends on **U**, the particular group of users. It might depend positively on **U**, as in a telephone network; or it might depend negatively on **U**, as in the cable modem example, or in a swimming pool—the more people in the water, the larger the negative impact when I jump in.

We can describe our four goals mathematically. Efficiency requires us to choose the **A** and the **U** that maximizes the sum of our individual welfares, minus the cost of providing network configuration **A**. Voluntariness means that if we pay for the network by charging each user an amount, T_i, which could be a flat fee or a function of some sort, the benefit you get minus the money you pay has to be nonnegative. In other words, everybody comes out ahead from being part of the network, or at least breaks even. And the network itself breaks even—that's balance. The sum of the T_i's has to equal the cost of the network. If the T_i's are insufficient, we need an external infusion of cash or assets. If they exceed the cost, we have to decide what to do with the surplus. And finally, strategy-proofness says that your benefit from (**A**,**U**) minus T_i has to be at least as great as the benefit you could get by lying, by manipulating the system to induce some other (**A**′,**U**′) minus your T_i' for that different choice of network conditions and users. For instance, if I'm in a rural area, it might be very expensive to connect me to the network. But if I lie and say I have an enormous value, then it's still efficient to connect me, so it's in the interests of rural users to manipulate their values upwards to ensure that they're connected. The only way to stop that from happening is to make their T_i's extremely high. In fact, U.S. policy is exactly the opposite—we subsidize rural users to help them connect to the network.

Game theory has led to several relevant theorems. The fundamental one, published independently in the early 1970s by Allan Gibbard (then

It's a holdup in both senses—it's an impediment to your use of the network, and it's highway robbery.

at Chicago, now at Michigan) and Mark Satterthwaite (BS '67) at Northwestern, stated that it's impossible to find a mechanism that satisfies our four requirements that it be efficient, voluntary, strategy-proof, and balanced. However, the approach was so abstract that the possibility remained that for some class of network models one could, in fact, have all four. This hope was dashed in 1979, when Harvard's Jerry Green and J. J. Laffont (Laffont is now at the University of Toulouse) revisited the issue. Their work was done in the context of providing a public good, such as building a bridge, but applies to networks as well. It says that no general network-market model can satisfy all four of our desiderata.

Caltech's Matt Jackson, then at Northwestern, in collaboration with Salvador Barberà at the Universitat Autonoma de Barcelona, found that if we're willing to chuck out our beloved, slavish devotion to efficiency, we can come up with a mechanism that will satisfy the other three requirements. We won't need the Sopranos, people won't game the system, and it requires no external infusion of funds. However, the mechanism looks a lot like price caps, which makes industry very nervous. This result was also obtained independently by Hervé Moulin (then at Duke, now at Rice) and Scott Shenker (a computer scientist then at Xerox PARC, now at Berkeley, who was interested in network protocols). And a theorem by Ted Groves at UC San Diego says, when applied to this context, that we can keep efficiency while getting strategy-proofness and voluntary participation, if we're willing give up balance.

And finally, several other people and I have shown that we can get efficiency, voluntary participation, and balance if we're willing to give up strategy-proofness as a global concept and replace it with a local concept. That is, instead of it not being in anybody's interest to game the market *ever*, it's not in my interest to game the market as long as nobody else is gaming the market. If everyone else is playing fair, the system enforces fair play on my part. But if a group of people collude and try to game the market, they can do it. This local strategy-proofness is called the Nash equilibrium, because mathematician John Nash developed the idea at Princeton in about 1950.

So the state of the art in network models is that we can get three out of four. It's mathematically impossible to achieve all four. This means that we have to tailor each market to the particular characteristics of the network it serves—there is no one-size-fits-all optimal policy. And yet, our economic policy is still largely driven by the standard, competitive-market goal of four out of four. But in fact, when we look at the best way to handle different network markets, we may arrive at conclusions that are polar opposites of each other, as I'll demonstrate with a couple of examples—Microsoft and electricity.

Microsoft's operating system is a network externality, because the more people that use it, the more products are developed for it, and the more benefit you get from it. (I use a Mac, myself, so I'm denied a bunch of software that other people have; but for some reason I get more benefit from having a Mac, and fewer friends, than other people do from enjoying more friends and cheaper products. Go figure.)

The problem with this network is a really subtle one, but it's very interesting. It's what economists call the holdup problem, and it occurs when you have a network made of different components that are priced separately. Imagine that you're in New York City and you want to travel down to Washington, D.C., to protest some issue. You jump in your car and you get on the New Jersey Turnpike, which is a toll road. You drive through Jersey to the Delaware Pike, you pay a second fee to Delaware, and you get to Washington. Here's the holdup problem: suppose it's worth a dollar to you to take the trip, and the Jersey Turnpike charges 50 cents. A dollar minus 50 cents leaves 50 cents. You drive down the road and you get to the Delaware Turnpike. What if the Delaware Turnpike hits you for a buck? They have you over a barrel—you've already spent 50 cents, but you're closer now than ever and it's still worth a dollar to you to finish the trip. So you go on to Washington, you eat the other 50 cents, and you mutter to yourself, "What a rip-off. I'll never do that again." So it's a holdup in both senses—it's an impediment to your use of the network, and it's highway robbery.

Now, imagine I'm Microsoft and I've got a nice little monopoly going, with all these peripheral products adding value to my network, and then somebody new comes along with something as essential as my operating system. Suddenly, in order to get the full benefit of your computer, you have to buy my product plus this other guy's product. But you have to buy *my* product first.

A browser isn't much good without an operating system to run it on—at the moment. This means that the industry overall faces a holdup problem. Microsoft knows that if it charges a high price for its product, then the browser company, which now has a captive market, can also charge a high price. But then nobody will buy either product. So the first firm in—I hate to use this expression—the value chain really *is* threatened by the firms farther down the line. In the big picture, it might actually be mathematically efficient for the first firm in line to kill off the second firm and integrate the two products. Predatory pricing—selling the second good for free, or below cost, in order to kill a competitor—is illegal, but it solves the holdup problem. Maybe Microsoft *should* be allowed to decide whom to subsidize and whom to kill, given that it already has the operating-system market tied up. If Microsoft had a viable competitor in that market, that might not be true, because then people would have an alternative route to make the journey.

Let's move on to the mother of all mess-ups: electricity "deregulation." I've got that in quotes because people usually think of deregulation as removing regulations, but this "deregulation" produced a new set of rules the size of a phone book.

Current federal policy is driven by the idea that we want to break Microsoft's monopoly because monopolies are bad. We outlaw predatory pricing, because predatory pricing enhances monopoly. We want to have open systems, open network platforms, to guarantee the largest amount of access and the largest amount of product development. But in the network model, none of those things can be shown to always be efficient. Microsoft's strategy has encouraged innovation in some areas and thwarted it in others, so the overall effect on efficiency isn't clear. We don't know what the efficient policy actually is, and it might not be that the efficient policy is the best policy. For example, if the efficient mechanism's not balanced, the social cost—the flow of money in or out of the system—might outweigh the benefits. It's a very complex issue, and we need to spend a lot of time modeling the minutiae of the industry in order to get the right solution.

Let's move on to the mother of all mess-ups: electricity "deregulation." I've got that in quotes because people usually think of deregulation as removing regulations, but this "deregulation" produced a new set of rules the size of a phone book. Electricity is another pervasive network externality. Electric power follows Kirchhoff's law, as you may remember from Phys 1, so we don't know where the individual electrons are going but we know systemically what's going to happen. Electricity users are very sensitive to fluctuations in voltage—I have a set of expensive tube amps in my stereo at home; I'm really unhappy when they blow. And there are lots of computers containing lots of business records in lots of offices. So it's essential that the quality of the service is held constant; that is, the voltage fluctuations must be kept within tolerable levels.

The way this was traditionally dealt with was by having a monopoly; the monopoly solved the systemic problems; we regulated the monopoly. But a monopoly could charge a high price and be inefficient. Under regulation, it turns out it was *still* inefficient. Under the old regulatory system, we had balance—the system broke even; the price regulations ensured that. We had voluntary participation. The Supreme Court ruled in the Hope Natural Gas Company case of 1904 that a regulated firm was entitled to a fair return on its investment. So the utilities weren't coerced—Edison voluntarily sold electricity under regulation, and made money as a result. The argument

on the consumer side is a bit more subtle: nobody *forced* me to be part of the network. In theory, I could have always gone "off the grid" and put photovoltaic panels on my roof, or a windmill in my backyard. It wouldn't be cheap, but I could do it, and as electric bills spiral upward, a number of people are. Or I could have renounced my TV, microwave oven, air conditioner, computer, etc. and lived like the castaways on *Gilligan's Island*. It wasn't likely to happen, but nobody was stopping me. And, finally, the price caps made the system relatively strategy-proof. We had three out of four, and the downside was that we lost efficiency.

Deregulation was enacted under political constraints, so they went for four out of four. The theorem says you can't do it—unfortunately, nobody read the theorem in policyland. They intended to lower prices, so consumers would benefit and would join voluntarily; they were going to induce efficiency by relaxing producer price controls, so producers would join voluntarily, too; and the mechanism was set up to break even, so it had balance. And they relied on competition to make the mechanism strategy-proof.

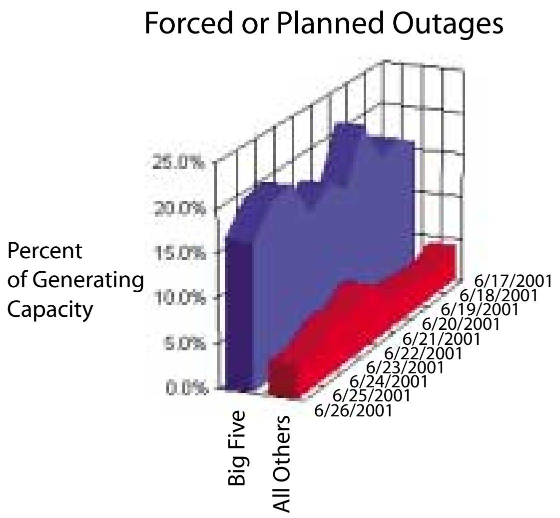
So what happened? There's actually a sequence of markets. The so-called day-ahead market, for delivery tomorrow, is by the hour: 10 o'clock, 11 o'clock, 12 o'clock, and so on. The day-ahead market matches expected supply and demand. But say the next morning it turns out that the day is going to be hotter than forecast, and people are going to crank up their air conditioners. So there's the morning market, which is for same-day delivery in 15-minute intervals, to fine-tune supply and demand. And as the delivery deadline approaches, there are many more markets: a market for spinning reserves—people being paid to keep their generators running in case they're called upon; for nonspinning reserves—people who have their plants fired up, but the generators aren't turning; and so on. There's this whole hierarchy of markets based on how quickly a particular plant can be called on to produce. Then, at the very last moment, there's a market that forces supply to be equal to demand.

The later markets are run by the Independent System Operator, or ISO, a sort of quasigovernmental operation that pays whatever price has to be paid to maintain our constant voltage. The ISO's mandate is to keep the lights on at all costs. The ISO loses money, but it has to break even because the system has to be balanced. So it allocates its cost to the users, which in this case include the Big Three utilities: Southern California Edison, Pacific Gas & Electric, and San Diego Gas & Electric. But it doesn't instantaneously know who caused the excess demand, because it's a systemic problem, so the cost is shared via some rule.

In May 2000, for reasons that I'll talk about shortly, prices on the day-ahead market jumped through the roof. Consumer demand was up, so if the mechanism was strategy-proof, Edison, PG&E, and SDG&E should just have increased their demand and paid the higher price. But the cost they could pass on to the consumers was fixed by the retail price caps, so they didn't want to do that—they'd take a bath if they did. (As you know, they took a bath anyhow, but I'll get to that in a moment.) On the other hand, if they reduced their demand a little bit, it would push the unfilled consumer demand into those last-ditch markets where the ISO would have to cover it. The ISO divides its cost between the users, so the logic was this: I could buy an extra dollar's worth of electricity today, but if I don't, the ISO will buy it tomorrow. It will cost the ISO two bucks, but if the ISO divides that equally among us, two over three is less than a buck. Unfortunately, this only works if I'm the only person who does it. If two people do it, the cost becomes two people times two dollars divided by three people, which is more than a buck; if all three do it, everyone winds up paying two bucks. There's a strong incentive to be the first to act, even though the advantage you get is fleeting, because if you are honest you are guaranteed to lose unless everyone else is equally honest. (This particular scenario is a staple of game theory, and is called the Prisoner's Dilemma, because it was originally couched in terms of two cellmates given the opportunity to rat each other out in exchange for a lighter sentence.) So when the price went up, the declared demand—the amount the utilities said they wanted to buy—went down, and the excess was pushed into the residual markets. Suddenly the ISO, which was intended to do the transactions needed to suppress the last little fluctuations in the system, was buying 15 percent of the power. It was never meant to do that. And it was allocating costs in a way that was completely non-strategy-proof.

The supply shock—that price jump—was set up, again, because there was no balanced, strategy-proof mechanism. The miracle is that it took a year for the flaw to become apparent. Anyway, when supply was withdrawn, the same thing happened—the demand was forced into the last-minute markets, where you can charge almost anything and the ISO has to pay it. All at once, the scheduled maintenance time, the downtime when generators were removed from the system, roughly doubled. Maintenance outages are a matter of public record—whenever a generator is out for part of a particular day (the data don't track duration) the operator has to file a report. Generators were going on the fritz left and right—on some days we had 30 percent outages. And one study estimates that the producers' profits went up by $6 billion. Once again, you can't get four out of four. If we'd settled for three out of four, we never would have had this problem.

So the combination of giving the producers an incentive to withhold supply, magnified by an incentive for the buyers to withdraw demand



In the current market, when supply is reduced the sellers earn a higher price on every megawatt sold. Thus, when a generator goes off line, the sellers that own many generators make higher profits. The above data from the ISO shows a week's worth of generator outages reported by the so-called Big Five producers (AES, Duke, Dynergy, Mirant, and Reliant) compared with the smaller, independent producers.

has proved to be disastrous. The system can't be balanced—it's running a giant deficit of about $15 billion, and the taxpayer has to pick up the bill. We have managed to achieve none out of four.

Since we can attain three out of four, there at least four possible solutions. We could renounce **E** (efficiency). That is, we can go back to a regulated monopoly, or we could introduce price caps. A lot of people are crying for that, because, well, things were bad in the old days but they weren't *this* bad! Or we could give up on **V** (voluntariness)—the state could seize the power plants through eminent domain, and force them to sell power to us at a fixed price. There have been a lot of calls for that, too. Alternatively, we could abandon **B** (balance). If we got rid of the balance requirement, we could assign long-term contracts for the delivery of a specified amount of power based on our best guesses for demand. We know we can engineer the awarding of these contracts in a strategy-proof manner, per Ted Groves's theorem I mentioned earlier. However, then the ISO will not always break even, and the taxpayer will have to foot the bill. And the imbalance could be large, in which case we're no better off than we are now.

My preferred solution is to design a better market; that is, we relax **S** (global strategy-proofness) and go for local strategy-proofness by giving people the correct incentives. In the previous market design, all the units of electricity that were bid for less than the market-clearing price—the price of the lowest unsuccessful seller (if you arranged all the bids from lowest to highest, the lowest unsuccessful seller would be the first seller whose bid was not taken)—were sold at the market-clearing price. We could stand that on its head, by breaking up the market into a set of smaller markets for each unit of capacity—per 100 megawatts, say. So there's a market for the first 100 megawatts, and the market sets a price. But then what we do is we award the sale—at the market-clearing price—to the generator who submitted the *lowest* bid. So it's to your advantage to bid a low price, because you'll get paid the highest price. Then there's another market for the next 100, and the process repeats. Now, if you try to withhold supply, you take yourself out of all the markets except for the last one, so your action benefits you only on the last 100 megawatts—unlike the exisiting situation, where you would affect the price of *all* the megawatts sold. And by taking yourself out of the previous markets, you lose all the business transacted therein; if demand is less than you predict and the market never gets to the 100 megawatts you're holding out, you're only hurting yourself. The more markets, the better this mechanism works—within computational reason, of course.

Actually, I think the best mechanism is to have a spot market with the no-**S** solution I've just described, coupled with the ISO using long-term contracts for power reserves that are awarded by what I call the Teacher's Pet method. Brownie points are given to producers who consistently have the lowest prices on the spot market, or who have the best reliability record—i.e., the fewest maintenance outages. Then if two bidders come in at the same price, the one with the most brownie points wins. Many government agencies already do this—the Defense Department, for example, puts performance incentives into its contracts, and bases future awards on the contractor's history of cost overruns and so forth. This system could also be used to reward whistle-blowing companies by giving them major brownie points in the next round of contracts. However, as I said before, when the ISO enters into long-term contracts, the balance requirement goes out the window. So I call this the no-**BS** solution.

These kinds of issues are going to get even bigger and more complex as our economy increasingly becomes a network of networks. So we really need to sit down now, and figure out the arcane details of how these markets work, in order to head off future missteps.

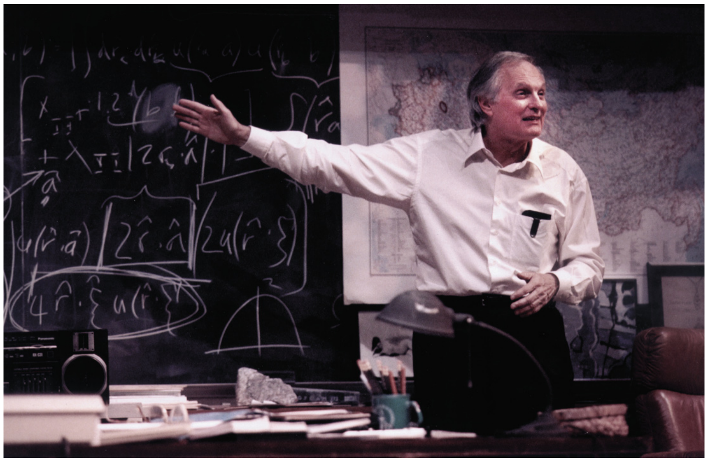
Assistant Professor of Economics Simon J. Wilkie earned his B. Comm. from the University of New South Wales in 1984, and his MS and PhD from the University of Rochester in 1988 and 1990, all in economics. He was a member of the technical staff at Bellcore, where he specialized in game theory as applied to telecommunications markets, before coming to Caltech as an instructor in economics in 1994. He was made an assistant professor the following year. This article is adapted from a Seminar Day talk.

PICTURE CREDITS:
28, 34–35 — Edison International; 30, 31 — Doug Smith

SCIENCE ON STAGE

***QED*, a play by Peter Parnell, produced by the Center Theatre Group at the Mark Taper Forum, Los Angeles, March 10–May 13, 2001**

***Oxygen*, a play by Carl Djerassi and Roald Hoffmann, produced by the San Diego Repertory Theatre at the Lyceum Theatre, San Diego, April 2–7, 2001 (published by John Wiley & Sons, 2001)**



Alan Alda plays Richard Feynman in *QED* at the Mark Taper Forum (Caltech physicists were consulted on the blackboard.)

by Jay A. Labinger
Administrator,
Beckman Institute

The 20th century has given us a number of encounters between the spheres of science and the theater—Brecht's *Galileo* (1939) and Dürrenmatt's *The Physicists* (1962), to cite just two examples. As we transit to a new century and millennium, the frequency and impact of such plays seem to be decidedly on the increase. The most visible has certainly been Michael Frayn's *Copenhagen*, an account of a World War II meeting between Niels Bohr and Werner Heisenberg, which has enjoyed lengthy, sold-out runs in London and New York, despite (or, possibly because of, though that doesn't seem very likely) a heavy dose of dense quantum mechanics. David Auburn's *Proof*, about a mathematician's daughter, recently won the Pulitzer Prize for best drama of 2000. Tom Stoppard's *Arcadia* (my personal favorite in this genre), featuring themes of chaos theory and thermodynamics, has been a continual success, with frequent revivals, since its first appearance in 1993.

Should this surprise us? Certainly science pervades our contemporary world, and equally certainly the theater must reflect that world to stay relevant. But it is not clear whether the aesthetic and intellectual demands of the two spheres are compatible—the encounter might be more of a collision!

At the very least, the playwright tackling a science-related theme will have problems to solve. How much of the scientific content must the audience understand, for the play to be fully effective? For example, scientists' motivations might well appear incomprehensible to an audience that doesn't appreciate the significance of their scientific work. On the other hand, one of the more basic rules of theater is "show, don't tell." How can that significance be adequately communicated, without violating that rule, and risking a complete breakdown of rapport?

Two plays with science connections have recently premiered in Southern California. The first, *QED*, features Alan Alda portraying the late Caltech physicist Richard Feynman. Apparently Alda himself was the prime initiator of the project, having been impressed by the dramatic potential of Feynman's life as depicted in Ralph Leighton's *Tuva or Bust!*, and recruited Parnell (previously best known for his adaptation of *The Cider House Rules*) as playwright. The play consists of Feynman talking—sometimes on the telephone, with his wife, friends, colleagues, and doctors, as well as with a student (the only other character in the play), but mainly directly to the audience—during a day and evening near the end of his life.

Feynman/Alda talks mostly about himself: his interests, his past life, his future—his science? We do get some, especially in the first act, but it is hardly integral to the play. We are treated to a number of platitudes about science; we are *told*, but hardly ever *shown*, how excited scientists are about their work. Alda tries to illustrate what doing Feynman's kind of physics might be like by means of an example from chess, not from science. On the occasions when real science is presented, it is at a level way over a nonphysicist's head, as when Alda starts sketching Feynman diagrams on a blackboard, explaining them in terms of virtual photons and the like.

This combination of vague generalities and arcane complexities, with little in between, has the effect (whether intended or not) of marginalizing the scientific theme. The audience is encouraged to take in what's easy and tune out what's hard, never challenged to work at making sense of unfamiliar ideas. Perhaps the clearest indication of how little is expected is that *every* time (it seemed like dozens, though I suppose it was only three or four) Alda says "quantum electrodynamics" he turns to the audience and repeats "QED." Couldn't they trust the audience to figure out the title's significance after the first time?

The net result is that Feynman the character is not a scientist with a personality; he's just a personality who happens to be a scientist. *QED* may well appeal to many—it does afford the opportunity to spend some time with an entertaining persona (though how much of that is Feynman, and how much Alda, is not easy to ascertain). But the problems of dealing with a scientific theme in a play have not been solved in any way, merely evaded.

Oxygen is a different matter. The playwrights are two well-known chemists, Nobel laureate Roald Hoffmann and National Medal of Science awardee Carl Djerassi. (Both are also well known outside of chemistry as prolific authors of fiction, nonfiction, and poetry.) The premise of *Oxygen* is that the Nobel Foundation has decided to institute a new program of "retro-Nobels," recognizing work done before the establishment of Nobel Prizes at the beginning of the 20th century. A committee for the retro-chemistry award quickly zeroes in on the discovery of oxygen as a worthy subject for the award. But who should receive it? Carl Wilhelm Scheele, a Swedish pharmacist, who was apparently the first to obtain a sample in the laboratory? Joseph Priestley, the first to publish his findings? Antoine Lavoisier, the first to understand what oxygen really is? All three?

Interwoven with the contemporary action is an account of a (fictional) 1777 meeting of the three chemists, invited to Sweden by King Gustav III to decide who should get credit. Each of the three is assigned his advocate on the committee, whose arguments in favor of their candidates echo not only those made by the candidates on their own behalf but also sad stories about priority claims and professional jealousy among the advocates themselves. This resonance is nicely reinforced by having a single actor play each candidate-defender pair; temporal scene shifts are signaled by minor costume changes. Another resonant device is the inclusion of a young historian of science writing her dissertation on "Women in the lives of 18th century scientists" as secretary to the Nobel committee; the wives attend and play important roles at the 1777 meeting, especially Mme. Lavoisier.

Evading the playwright's dilemma is not an option here as it was in *QED*: the scientific content is *central* to the dramatic argument. Lavoisier was the first to understand the role of oxygen in phenomena such as combustion and rusting, thereby overthrowing the phlogiston theory in which both Scheele and Priestley devoutly believed. Unless one appreciates the significance of that, the priority dispute makes little sense. So somehow it must be explained, without squelching the drama by a descent into didacticism. Hoffmann and Djerassi try hard to steer between the two looming cliffs (at one point they interpolate a stylized masque, performed by Lavoisier and his wife, to communicate some of the material) but their solution to the problem is not entirely satisfying.

In an interview with a San Diego paper before the premiere, Djerassi claimed that their writing about "a part of our culture which we did not have to absorb" was an advantage; but it may have also been somewhat of a disadvantage, making them a bit less sensitive to the needs of an audience that is unfamiliar with that culture. Similarly, the contemporary chemists are not so compelling characters as one might wish. They are obviously meant to be seen as passionate about their science, which carries over to the positions they take during the committee's deliberation, but we aren't really shown where such passion might come from. Perhaps the authors, as passionately committed scientists themselves, thought it would be obvious?

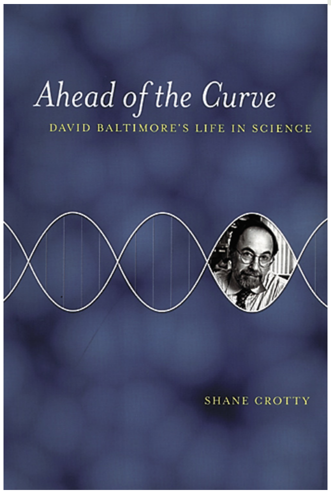
It seems likely that *Oxygen* was influenced by *Arcadia*: the two plays exhibit certain similarities (beyond the scientific themes), most prominently the use in both of alternating time frames. If the latter is more successful as a dramatic event (which it is), there is no shame in that for Hoffmann and Djerassi—Stoppard is, after all, one of the leading playwrights of our time. But possibly there is an instructive message, that one must be wary of being *too* close to one's subject. *Oxygen*, much more than *QED*, illustrates both the potential problems and rewards of dramatizing science. Let's hope that Hoffmann and Djerassi, and others as well, will keep on trying. ☐

Jay Labinger, a chemist and member of the professional staff, is coeditor of The One Culture? A Conversation About Science*, which has just been published by the University of Chicago Press.*

In the San Diego Repertory Theatre production of *Oxygen*, Lou Seitchik (left) stars as Priestley, Randall Dodge (center) as Lavoisier, and Jeff Anthony Miller as Scheele.

A Life in Science

by John Sutherland, Visiting Professor of Literature



Ahead of the Curve:

David Baltimore's Life in Science

270 pages

University of California Press, 2001

$29.95

What, one wonders, goes through David Baltimore's mind when, having picked up his espresso from the Red Door (as I've seen him do), he browses the Caltech Bookstore and passes a rack stacked high with two books about David Baltimore: Dan Kevles's *The Baltimore Case* (Norton, 1998) and now, three years later, Shane Crotty's biography.

Everything has happened fast in Baltimore's life. As Crotty records, he believed, in his early 20s, that if one did not make one's mark by 30, there would be no mark for posterity to admire. He got his Nobel aged (if that's the word) a prodigiously youthful 37. His career hit a wall, it seemed, in his 50s with "the case." Now, phoenixlike, he is arisen to lead Caltech into the New Millennium. "I live in the future, not in the past," he is quoted as saying. Having achieved so much, he has still, it seems, much to achieve. Nor will they be ordinary achievements. He is, he believes, "the only functioning scientist who is running a major university in the United States." Fast *and* two-fisted.

Having your biography written while you are still alive, the English poet Philip Larkin said, is like being measured up (still breathing) by the undertaker. I am not a scientist (even Crotty's accessible explanations about recombinant DNA and retroviruses are sometimes a bit beyond me). But I am a biographer. And it is the problems of the biographer's craft that primarily interest me in Crotty's enterprise.

It is difficult to write "authorized" biography about the living. Baltimore evidently sanctioned this book, although the interviews he gave his biographer seem to have been singularly unrevealing. Punches have to be pulled when dealing with a living subject. If they're not, authorization and "permissions" are yanked. And the libel lawyers are in the wings (you can't, as every biographer knows, libel the dead). But for the reader the pleasure in biography is, essentially, voyeuristic. We want to see what makes the person "tick." To do that, you have to take the back off the watch and do some prying.

In Baltimore's case, biographical prying is further discouraged by the fact that he is, manifestly, someone who values and protects his privacy. Crotty has been careful not to trespass. So much so, that at times he seems to be complicit with his subject in veiling what biography normally conceives its responsibility to uncover.

This is not to say that one wants *National Enquirer* or "blackwash" revelations (not that there would be any here). One can respect Crotty's decision not to press on personal but (in this context) irrelevant aspects of his subject's life. Baltimore's first marriage and divorce, for example, are dealt with in a sentence. His second marriage is recorded, but without any close-up detail.

Nonetheless, there are areas of private life that are relevant to the personal evolution of someone so extraordinarily distinguished as Baltimore. Most careers, even "a life in science," follow the rule As the twig is bent, so grows the tree. Childhood—the formative years—is important. Baltimore's first 20 years are summarized here in three pages. Crotty gives us a luxuriant and protracted word picture of the Swarthmore campus ("the dogwood trees and a thousand rhododendrons bloomed, carpeting the campus with red, cream, and pink petals") but nothing about the Baltimore home or even whether there were siblings other than a brother briefly mentioned. What did his parents do for a living?

We learn that Baltimore and Francis Ford Coppola "were the two-man tuba section in the Great Neck High School Marching Band" (a curious fact for which one is profoundly grateful). But I can find nothing in this biography about Baltimore's father. His mother is credited with being his lifelong inspiration. But she has no index entry, nor does she make any real presence in the narrative.

There is nothing, apart from one throwaway reference, to Baltimore's Jewishness. As someone born in 1938, he may conceivably as a young man have encountered prejudice. He is now, we learn, "unreligious." Was his upbringing secular, or did he lose his faith?

The lack of personal background is tantalizing. More so, since there are fleeting allusions to important aspects of his adult personality passed down from his family. There is, for example, a parenthetic reference to the Baltimores sympathizing with, for two generations before David, "leftists and socialists." In his thirties, we learn, Baltimore "hated Nixon," and thought his "War on Cancer" a sham. At this period of his life (when he was doing his most exciting scientific work) Baltimore "disdained capitalist society" and declared himself "an anticapitalist." When did his views change? Or have they?

One of the more interesting human subplots to the narrative is Baltimore's impassioned resistance to the Vietnam War (had he been born five years later, Canada might have been able to claim him as its most distinguished scientist). Baltimore's truly eloquent and idealistic outburst against the ineffable John Dingell during "the case" reminds one of nothing so much as those gallant dissidents who stood up publicly to denounce HUAC and McCarthyist purges, 40 years earlier. (Crotty, incidentally, handles this episode very effectively.)

The aspect of Baltimore's intellectual character that emerges most clearly is that he is a loner. As a young scientist he was a self-made man. His alma mater will take no pleasure in Crotty's book. Baltimore, perhaps its most famous living alumnus, is quoted as saying: "At Swarthmore the teaching of biology was poor—at best. The courses were really generally *bad*." But perhaps genius needs to be left alone, to grow at its own rate in its own peculiar way. For students like David Baltimore, bad courses are the best courses. Would undergraduate education at Cambridge, MIT, or Caltech have crushed the original genius out of him?

Late-20th-century, laboratory-based science cannot be done at the highest level by "loners." It costs too much. Few biologists are born billionaires. Accommodations must be made: with institutions, with the state, and with "capital." As a young scientist, Baltimore apparently believed that if funds were needed for his kind of science it should ideally be from the taxpayer ("the only way to do research was on government money").

But when he made his pact with a large institution (with the ultracapitalist name, Rockefeller) did he have any twinges of "radical, leftist" conscience?

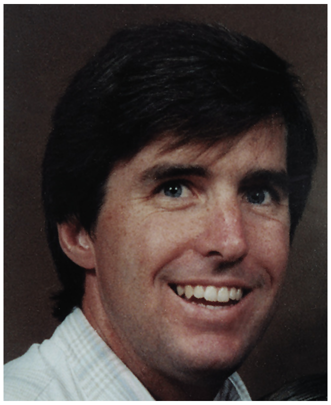
What went through Baltimore's mind, in August 1980, when Jack Whitehead offered Baltimore a research institute? He who sups with the devil should use a long spoon? Or, this is the only way forward for research, such as that into molecular biology, which needs unimaginably large sums of money? These are questions that the reader (legitimately, I think) asks. This biography gives hints, but no answers.

There is much to applaud in Crotty's book. I found his expositions of Baltimore's research for the layman (as a layman) admirably comprehensible. Crotty is good on the ethical problems raised by gene research, and Baltimore's (sensible, one apprehends) thinking on the Pandora's box his genius has opened.

This is an interesting study of a fascinating and important man. But, as biography, Crotty's book stimulates an appetite it signally fails to satisfy. There remain enigmas. For instance: the best prose in the book is Baltimore's (I would point to the witty summary of his "education in irrationality," in his inauguration address at Caltech, quoted here as epilogue). Baltimore is a brilliant scientist, yes. But he is also a highly cultivated man, with a love of theater, jazz, art, and literature. We do not learn from this book how he became that unusual man. The posthumous biography will doubtless tell us. And, by the time it comes along, there will, for a certainty, be much, much more for the biographer to record. ☐

John Sutherland taught at Caltech, in the Division of the Humanities and Social Sciences, from 1983 to 1992, and has visited quarterly since. He has written biographies of Sir Walter Scott and Mrs. Humphry Ward and is currently writing the authorized biography of the poet Stephen Spender.

JEFFREY SCOT BANKS
1958–2000



In 1996 Jeff Banks (right) received the National Academy of Sciences Award for Scientific Reviewing, presented by John Ferejohn, Munro Professor of Political Science at Stanford and senior fellow of the Hoover Insitute (Ferejohn taught at Caltech from 1971 to 1983). Banks's name appears on the wall with those of awardees from previous years.

Jeffrey Scot Banks, professor of political science, died December 21 at the age of 42 of complications of a bone-marrow transplant.

After earning his PhD from Caltech in 1986, Banks left to join the faculty of the University of Rochester and returned to Caltech in 1997. He taught and did research in the general field of political theory, including political economy, game theory, and social choice. He made significant contributions to a field of political science characterized by the use of formal mathematical and deductive methods to model political behavior—behavior such as strategic voting, bargaining, coalition formation, and jury decisions.

A conference in his honor was held on campus April 7; students and colleagues presented papers that drew on Banks's work on the role of incomplete information in models of political processes. Those colleagues, who had gathered from around the country, also joined friends for a memorial service in Dabney Lounge to remember and celebrate the life of Jeff Banks in a less scholarly fashion.

John Ledyard, professor of eonomics and social sciences and chair of the Division of the Humanities and Social Sciences, likened the empty seat at the conference to the hole in a pilots' formation or the empty barstool. Ledyard welcomed everyone and introduced the other speakers—Banks's teachers, colleagues, and students, who offered remarks in roughly the chronological order in which each speaker had encountered his career. Ledyard was slightly out of chronological order himself, arriving as a professor at Caltech just as Banks was finishing his PhD.

Born in San Diego, Banks graduated from UCLA in 1982. Richard McKelvey, the Wasserman Professor of Political Science, recalled hearing of this "really smart UCLA student" who had applied to Carnegie Mellon for graduate school. McKelvey set about explaining to Banks why he should come to Caltech instead. He did, earning his PhD in 1986 with a thesis on "Signaling Games: Theory and Applications," with McKelvey as his thesis adviser.

David Porter, currently on the staff of the Economic Science Laboratory at the University of Arizona, first met Banks as a fellow graduate student. "When I think about the wonderful qualities of Caltech, namely, cutting-edge research, innovative thinking, honesty, and cleverness, I think of Jeff," said Porter. "And if you worked with him, you knew you were in for a lot of laughs and fun."

In addition to numerous academic papers ("It's remarkable how much he accomplished in such a short span of time," said McKelvey), Banks coauthored a book, *Positive Political Theory I: Collective Preference*, with David Austen-Smith. Austen-Smith, now professor of political science and economics at Northwestern University, spoke of Banks's "evangelical zeal" for political science. "To Jeff, doing research was sheer pleasure." He also noted Banks's easy disposition and enthusiasm and his ability to introduce lines from the movie *This Is Spinal Tap* into seminar presentations.

In 1986, Eric Hanushek (now at Stanford's Hoover Institution) was chairman of the economics department at the University of Rochester and found himself for the first time in competition with the political science department for a faculty appointment. This led to Banks's unique appointment and ultimately tenure in two departments.

"He was a natural success at Rochester," said Hanushek. If it hadn't been for his illness and return to Caltech, he claimed, Banks would have substantially changed political science and political economy at the university. "There are some people with whom everyone identifies as a friend," he said. "Jeff was one of those people."

John Duggan, PhD '95, who followed Banks's path to the University of Rochester, where he is now associate professor of political science and economics, described him as a "really deep thinker, a careful and rigorous thinker, and he challenged you to be also."

"He just loved research so much," said Duggan. "His energy and enthusiasm were infectious, and that made working with him so much fun." Banks was productive even when he was ill, Duggan added, and left several papers that will be published posthumously. "In the profession, our debt to him is great."

Banks received numerous awards and recognition for his work. From 1989 to 1994, he was a National Science Foundation Presidential Young Investigator, and received the National Academy of Sciences Award for Scientific Reviewing in 1996. In 1996 he was also elected a fellow of the Econometric Society.

"Jeff was my teacher and my thesis adviser," said Daniel Diermeier, who studied with him at Rochester. Diermeier explained that the German term for thesis adviser is *Doktorvater*. "There's truth in this concept," he said, "which, as a father and a teacher now myself, I appreciate more. Teaching is about creating someone who is then creative in turn. We grow into our research."

Diermeier, who is now the IBM Professor of Regulation and Competitive Practice at Northwestern University, also appreciated the American informality he met at Rochester. "All my previous teachers had the same first name: *Professor Doktor*. And now here was 'Jeff' in his sneakers." Two qualities made him unique, said Diermeier: "the deep joy" that radiated from him and his deep commitment to research.

Banks was diagnosed with leukemia in 1995 and underwent a bone-marrow transplant in the summer of that year. In 1997, he returned to Caltech as professor of political science. He became executive officer for the social sciences in 1999, a post in which his dry wit and calm, easygoing nature, as well as his knowledge of voting theory, helped smooth many meetings.

Professor of Economics Matt Jackson came to Caltech at the same time that Banks returned. "He wanted everyone to enjoy life," Jackson said. "He could always see the humor in a situation." In his professional life, he taught others to "sweat the details; details matter." And even when his health deteriorated, "he came in, taught his courses, met with graduate students, and kept doing the day-to-day things, no matter how difficult. He made a big difference in the small things as well as the big ones."

On behalf of Banks's family (which includes sons Bryan, 15, and Daniel, 13), his wife, Shannon, thanked all those who had come that day "not just to mourn his passing but to celebrate his life." She thanked the anonymous bone-marrow donor "who allowed the extra time" and also all those in the audience who had signed onto the bone-marrow registry because of her husband's illness. "He fought long and hard to stay with us." She also presented to Ledyard and to Larry Rothenberg, director of the Wallis Institute at the University of Rochester, framed copies of Banks's Presidential Young Investigator Award and a photo of him receiving the National Academy of Sciences award.

In closing, Ledyard stated that the new seminar room in Baxter Hall would be named in Banks's memory and also announced the creation of the Jeff Banks Memorial Seminar Fund. Contributions to the fund may be sent to Susan Davis, Caltech 228-77, Pasadena, CA 91125. Checks should be made out to the California Institute of Technology. ☐

HONORS AND AWARDS

Professor of Biology Pamela Bjorkman is one of 72 American scientists elected this year to membership in the National Academy of Sciences (NAS). She's the first woman out of a total of 67 living Caltech faculty members elected to that honor. Bjorkman, who is also executive officer for biology and an investigator with the Howard Hughes Medical Institute, has been a member of the Caltech faculty since 1988. Her research focuses on molecules involved in cell-surface recognition, particularly molecules of the immune system.

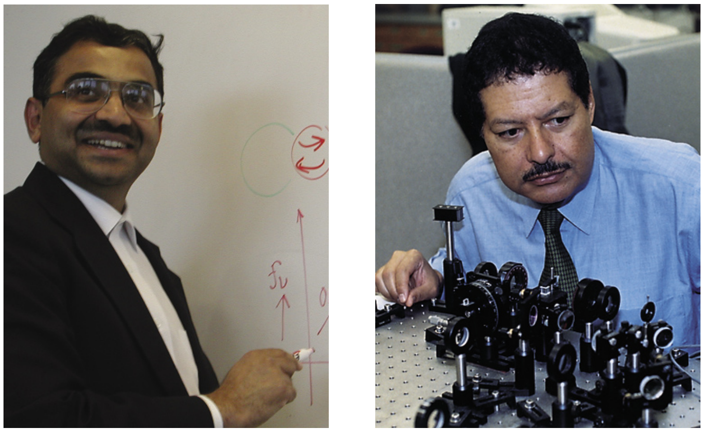
John Abelson, the Beadle Professor of Biology, has been elected to the American Philosophical Society.

Paul Bellan, professor of applied physics, has received one of two 2001 SPD Popular Writing Awards, given each year to a professional scientist and to a science writer or journalist by the Solar Physics Division of the American Astronomical Society.

Roger Blandford, the Tolman Professor of Theoretical Astrophysics, was named the Tetelman Fellow at Yale for 2001; he delivered the Tetelman lecture in mid-February. In June, he traveled to Munich to give the Siemens Lecture.

Mory Gharib, professor of aeronautics and faculty member in bioengineering, was invited by the American Association for Thoracic Surgery to give the Honored Speaker address to the 81st AATS conference on May 8 in San Diego. He discussed the challenges and rewards of applying bioengineering principles to space exploration.

William Goddard, the Charles and Mary Ferkel Professor of Chemistry and Applied Physics, has been selected by the Southern California Section of the American Chemical Society to receive the Richard C.



Shri Kulkarni, the MacArthur Professor of Astronomy and Planetary Science (left), and Ahmed Zewail, the Pauling Professor of Chemical Physics, have been elected to the Royal Society, established in England in 1660, the world's oldest scientific academy in continuous existence.

Tolman Medal. The medal was formally awarded at the Athenaeum on April 19.

Alan Hajek, associate professor of philosophy, has received a $10,000 grant from the Center for Theology and the Natural Sciences. He will develop a new course entitled Probability, the Philosophy of Religion, and the Philosophy of Science.

Linda Hsieh-Wilson, assistant professor of chemistry, has been selected to receive a 2001 Beckman Young Investigators award. The award program "helps provide research support to the most promising young faculty members in the early stages of their academic careers in the chemical and life sciences." This year marks the 10-year anniversary of the program, which is funded by the Arnold and Mabel Beckman Foundation, an independent, nonprofit foundation established in 1977.

Of 50 awards presented by the American Chemical Society at its April meeting, more than 10 percent went to Caltech faculty members. From left: the Nakanishi Prize to Jack Roberts, Institute Professor of Chemistry; the ACS Award for Creative Advances in Environmental Science and Technology to Michael Hoffmann, the Irvine Professor of Environmental Science; the Herbert D. Brown Award for Creative Research in Synthetic Methods to Bob Grubbs, the Atkins Professor of Chemistry; the ACS Award in Polymer Chemistry to David Tirrell, the McCollum-Corcoran Professor and professor of chemistry and chemical engineering; the George C. Pimentel Award in Chemical Education to Harry Gray, the Beckman Professor of Chemistry; and the ACS Award for Creative Innovation to John Baldeschwieler, the Johnson Professor of Chemistry and professor of chemistry, emeritus.

Philip Hoffman, professor of history and social science, has been named a Fellow of the John Simon Guggenheim Memorial Foundation. His project for the fellowship period will be "The Role of Crises in Economic and Financial Development," on which he will collaborate with UCLA professor of economics Jean-Laurent Rosenthal. Together with Gilles Postel-Vinay, Hoffman and Rosenthal coauthored *Priceless Markets: The Political Economy of Credit in Paris, 1660–1870* (University of Chicago Press, 2000).

Wolfgang Knauss has been named the Theodore von Kármán Professor of Aeronautics and Applied Mechanics, effective April 1. This title replaces that of professor of aeronautics and applied mechanics.

Shri Kulkarni, the MacArthur Professor of Astronomy and Planetary Science, has been invited to give this year's Sackler Lecture at Princeton University's department of physics.

Andrew Lange has been named the Marvin L. Goldberger Professor of Physics, effective July 1. This title replaces that of professor of physics.

Eliot Meyerowitz, professor of biology and chair of the biology division, has been named a Wilbur Lucious Cross Medal winner for 2001 from Yale University.

Paul Messina, director of the Center for Advanced Computer Research, received the Distinguished Associate award from the U. S. Department of Energy for his achievements in computational science and for his contributions to the DOE's Stockpile Stewardship program, designed to ensure the safety and reliability of the nuclear weapons arsenal.

John Preskill, professor of theoretical physics, has been invited to be the 2002 Lorentz Chair at the University of Leiden. Described as "the most prestigious visiting professorship in the Netherlands," the chair since its founding in 1955 has been held by 10 Nobel Prize winners.

Richard Roberts, assistant

professor of chemistry, has received a Presidential Early Career Award for Scientists and Engineers "for his innovative combinatorial method of selecting and designing protein motifs that specifically recognize biologically important RNA structures." The award recognizes outstanding young professionals at the outset of their independent research careers, providing up to five years of grant support. He has also been selected as an Alfred P. Sloan Research Fellow.

George Rossman, professor of mineralogy, has been selected to receive the Mineralogical Society of America's Dana Medal, which recognizes "continued outstanding scientific contributions through original work in the mineralogical sciences."

David Rutledge has been named the Kiyo and Eiko Tomiyasu Professor of Electrical Engineering. Effective April 1, this title replaces that of professor of electrical engineering.

Anneila Sargent, professor of astronomy and director of the Owens Valley Radio Observatory and of the Interferometry Science Center, has been elected a foreign associate of the Royal Astronomical Society "in recognition of her inspiring leadership and outstanding service to the promotion of astronomy."

Wallace Sargent, Bowen Professor of Astronomy, has been selected as the fourth Icko Iben, Jr., Distinguished Astronomy Lecturer at the University of Illinois at Urbana-Champaign, where he will deliver a public lecture, give a joint colloquium to the astronomy and physics departments, and interact with faculty, staff, and students. The lectureship brings world-renowned astronomers and astrophysicists to the university.

Edward Stone, the Morrisroe Professor of Physics, and director of the Jet Propulsion Laboratory from 1990 until May 2001, has received NASA's Distiguished Service Medal.

Keith Taylor, a member of the professional staff in astronomy, has received the Astronomical Society of the Pacific's 2001 Maria and Eric Muhlmann Award, for his "unique contributions to astronomical instrumentation at various observatories." □

Alexander Varshavsky, the Smits Professor of Cell Biology, has been named the co-recipient of the 2001 Wolf Foundation Prize in Medicine. He shares the $100,000 prize with Avram Hershko of the Technion, awarded for their discovery of the "ubiquitin system of intracellular protein degradation and the crucial functions of this system in cellular regulation." The Wolf Prize was established in 1978 to promote science and art for the benefit of mankind. Varshavsky was also recently elected to the American Philosophical Society.

Recipients of this year's ASCIT Teaching Awards are Oscar Bruno, professor of applied mathematics; George Cheron, lecturer in Russian; Kjerstin Easton, grad student in electrical engineering; Glenn George, lecturer in computer science and electrical engineering; Loren Hoffman, undergraduate; Dirk Hundertmark, Todd Instructor in Mathematics; Edward McCaffery, visiting professor of law; Thomas Neenan, lecturer in music; and Charles Peck, professor of physics.

The Graduate Student Council's Teaching Awards went to Hans Hornung, the Johnson Professor of Aeronautics; Julia Kornfield, associate professor of chemical engineering; and Brian Stolz, assistant professor of chemistry. Recipients of the GSC Mentoring Awards are Agustin Colussi, senior research fellow in environmental engineering science, and Brian Stolz.

FACULTY BOARD CHAIR ELECTED

For the first time, the faculty board chair will be occupied by a woman, Marianne Bronner-Fraser, the Billings Ruddock Professor of Biology. Melany Hunt, professor of mechanical engineering, was voted vice chair, and Ned Munger, professor of geography, emeritus, secretary, relieving Ward Whaling, professor of physics, emeritus, after a 16-year tour of duty.

Bronner-Fraser studies the development of neural crest cells in vertebrate embryos. These cells emerge from the neural tube shortly after neurulation and migrate to various parts of the body to establish diverse cell types such as neurons, glia, and pigment cells. Investigating what dictates the pattern of cell migration and what determines the type of cell they become could elucidate the causes of birth defects and cancers associated with neural crest cells, and suggest ways to prevent them.

There was another first in this year's election: the voting was conducted on line. As befits one of the partners in the Caltech/MIT Voting Technology Project, an electronic ballot was used, set up by project member Michael Alvarez, associate professor of political science, and Marionne Epalle, communications specialist in Engineering and Applied Science. □

MATHEMATICS RETREAT

Above: Annemarie DePrima at the Sea Ranch house.
Right: Charles and Margaret DePrima.

When the late Charles DePrima and his first wife Annemarie built their dream house in The Sea Ranch (a bucolic California coastal community about two hours north of the Golden Gate), they considered the possibility that it might someday become a retreat for Caltech's mathematics faculty. In fact, DePrima envisioned it as a smaller version of Oberwolfach, a mathematics retreat in Germany's Black Forest that he visited several times during his 40 years as a professor at Caltech. Annemarie passed away in 1984 after only four years in their new home; she and Charles had no children.

Charles and Margaret met in 1985 and married in 1987. Together they finalized the trust originally conceived by Charles and Annemarie, naming Caltech to receive the Sea Ranch property upon the death of the survivor. Charles DePrima died in 1991. Although she never knew Annemarie, Margaret has kept their dream alive by making their Sea Ranch home available one month each summer for use by the mathematics faculty.

DePrima was a distinguished mathematician who had become acquainted with Albert Einstein during his earlier years at the Courant Institute and the NYU Mathematics Institute. He was a dedicated teacher during his years at Caltech, and observed that there were very few special talks or seminars in mathematics designed for undergraduates. So, before his death, he and Margaret made a gift to Caltech to endow an undergraduate mathematics lecture series, which bears his name. His good friend and colleague Jack Todd, professor of mathematics, emeritus, delivered the first DePrima Lecture in 1991.

For information contact:

Chris Yates, JD
Susan A. Walker, CFP
Carolyn K. Swanson
Office of Gift and Estate Planning
California Institute of Technology
Mail Code 105-40
Pasadena, California 91125
phone: (626) 395-2927
fax: (626) 683-9891
planned_gifts@caltech.edu
www.gep.caltech.edu

ENGINEERING & SCIENCE

California Institute of Technology
Pasadena, California 91125