Winter 2004

The BRIDGE

LINKING ENGINEERING AND SOCIETY

NATIONAL ACADEMY OF ENGINEERING
OF THE NATIONAL ACADEMIES

Promoting the technological welfare of the nation by marshalling the knowledge and insights of eminent members of the engineering profession.

The
BRIDGE

NATIONAL ACADEMY OF ENGINEERING

The Bridge (USPS 551-240) is published quarterly by the National Academy of Engineering, 2101 Constitution Avenue, NW, Washington, DC 20418. Periodicals postage paid at Washington, DC.

Vol. 34, No. 4, Winter 2004

Postmaster: Send address changes to *The Bridge*, 2101 Constitution Avenue, N.W., Washington, DC 20418.

A complete copy of each issue of *The Bridge* is available in PDF format at *http://www.nae.edu/TheBridge*. Some of the articles in this issue are also available as HTML documents and may contain links to related sources of information, multimedia files, or other content.

The BRIDGE

Volume 34, Number 4 • Winter 2004

LINKING ENGINEERING AND SOCIETY

Editor's Note

Features

NAE News and Notes

(continued on next page)

Editor's Note

Pablo G. Debenedetti is the Class of 1950 Professor in Engineering and Applied Science at Princeton University and an NAE member.

Cutting-Edge Research for a Changing World

Every year, NAE sponsors a Frontiers of Engineering (FOE) Symposium, which brings together some 100 outstanding, competitively selected, young (ages 30–45) engineering leaders from academia, industry, and government laboratories for three days of sharing ideas and learning about cutting-edge research on a broad range of engineering topics. Modern engineering is characterized by globalization, rapid technological change, and the crossing of interdisciplinary boundaries, and the emerging engineering leaders who attend FOE symposia represent a wide spectrum of backgrounds, interests, and talents. The event offers a unique opportunity for them to learn about the frontiers in engineering areas other than their own. Six of the papers from this year's symposium are included in this issue of *The Bridge*. For the past two years, it has been my privilege to chair the FOE organizing committee, which selects the speakers and topics for the symposium.

The tenth FOE Symposium was held September 9–11, 2004, at the Beckman Center in Irvine, California. The program encompassed four themes: engineering for extreme environments, designer materials, multiscale modeling, and engineering and entertainment.

The opening session, organized by Mary Kae Lockwood and John Weatherly, focused on engineering for extreme environments. Laura Ray, whose paper appears in this issue, described the engineering challenges in designing scalable mobile robots for deploying instrument networks on the Antarctic plateau, an ideal location for studying the upper atmosphere at high magnetic latitudes. John Berkoe illustrated the role of simulation and modeling in extreme engineering projects through three examples, including the Tacoma Narrows Bridge mooring system and the conceptual design of the Chernobyl New Safe Confinement. Tommaso Rivellini (paper published in this issue) traced the evolution of landing technology over the last 40 years, culminating in the sky-crane system that will be used in the 2009 Mars Science Laboratory mission. The session closed with Kent Joosten's talk on accessing the lunar poles in future human exploration. Recent robotic and earth-based studies suggest that the lunar poles may offer advantages in terms of thermal conditions, availability of solar energy, and access to resources.

Kristi Anseth and Diann Brei organized the session on designer materials. Greg Carman explained the development and potential uses of thin films of materials that exhibit energy coupling. Thin-film active materials are still in their scientific infancy but may someday be used as sensors and actuators in applications ranging from drug delivery to microfluidics. Leslie Momoda, whose paper is published here, described ongoing research on performance-tailored structures that can adapt their performance or morphology on demand and the multifunctional materials that underlie such systems. Jennifer West discussed recent advances in vascular tissue engineering. Biomimetic strategies, including genetic modification of vascular cells and pulsatile stressing of smooth muscle cells, offer hope that novel substitutes for blood vessels may be fabricated for use in coronary artery bypass graft surgery, where treatment options are severely limited.

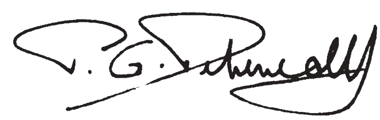
Multiscale modeling, that is to say, the computational analysis of systems with structures and dynamics that span many length and time scales, was the topic of the third session, organized by Dimitrios Maroudas and Grant Heffelfinger. Yannis Kevrekidis discussed the "equation-free" modeling of complex systems, whereby information from microscopic (e.g., atomic-level) dynamics is used, not to derive macroscopic equations, but to perform coarse-grained computational experiments that probe the system's behavior over macroscopic lengths and times. Rob Phillips' presentation focused on the computational modeling of complex problems, such as protein/DNA interactions. In his paper (published here), he gives examples of cases where an all-atom approach is inadequate because it generates an enormous amount of data without providing a concomitant increase in knowledge. His elegant solution involves maintaining the microscopic physics only to the extent it is needed. Adam Arkin's talk addressed biological models of events that occur on different time scales, from single-enzyme kinetics to the evolutionary time scale. He described how improvements in measurement technology and the abundance of genetic-sequence information are beginning to link disparate time scales. Bjorn Stevens discussed the computational challenges to simulating the climate system arising from interactions between small-scale processes, such as aerosol formation, that occur over the micrometer scale, and the planetary scale.

The relationship between engineering and entertainment was the subject of the fourth session, chaired by Chris Kyriakakis. Paul Debevec (paper published here) discussed the state of the art in the simulation of illumination in computer graphics. The accuracy of the simulation depends on re-creating the way light permeates a given scene, as it is reflected by some objects and travels through others. William Gardner addressed the challenges of accurately reproducing the spatial properties of sound, as well as progress toward the development of individualized binaural technology. His paper is also published in this issue.

The technical talks are always followed by extended, lively Q&A sessions with broad audience participation. The program surrounding this basic framework changes from year to year. The 2004 program featured presentations by representatives of three government agencies (Carey Schwartz of DARPA, Kenneth Harwell of the U.S. Department of Defense, and Doug Stetson of NASA's Advanced Planning and Integration Office) that fund cutting-edge research. Their descriptions of current projects elicited an excellent discussion and audience participation.

FOE symposia traditionally have an evening speaker, and this year was no exception. Alex Singer, a film director who has directed more than 280 television shows and five theatrical features during his 40-year career, spoke about the integration of technology and entertainment. In his lively talk, entitled "Unlikely Partners: DARPA and Me," he described an ongoing DARPA-funded short-film project envisioning a future of augmented cognition in the year 2030.

I know of no meeting as interdisciplinary, diverse, and stimulating as FOE, and I hope the six papers included in this issue convey some of the excitement we experienced in September at the 2004 symposium.



Low-cost mobile robots can advance scientific research on the Arctic plateau.

Cool Robots: Scalable Mobile Robots for Deployment in Polar Climates

Laura R. Ray, Alexander D. Price, Alexander Streeter, Daniel Denton, and James H. Lever

Laura R. Ray

Alexander D. Price

Alexander Streeter

Daniel Denton

James H. Lever

The Antarctic plateau is a unique location to study the upper atmosphere at high magnetic latitudes because it provides a stable environment for sensitive instruments that measure interactions between the solar wind and the earth's magnetosphere, ionosphere, and thermosphere. Existing stations on the edge of the continent and at the South Pole, and six low-power (50 W) automatic geophysical observatories, have demonstrated the value of distributed, ground-based observations of solar-terrestrial physics. Increasing the spatial density of these observations offers great scientific opportunities. A National Research Council report, *The Sun to the Earth and Beyond*, emphasizes the need for mobile instrument networks. The report recommends

---

Laura R. Ray is associate professor of engineering, Alexander D. Price is a master of engineering management candidate, Alexander Streeter is an M.S. candidate, and Daniel Denton is an A.B. candidate at the Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire. James H. Lever is a mechanical engineer with the U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire.

"comprehensive new approaches to the design and maintenance of ground-based, distributed instrument networks, with proper regard for the severe environments in which they must operate" (NRC, 2002).

This paper describes scalable mobile robots that can enable the deployment of instrument networks in Antarctica. The drivetrain, power system, chassis, and navigation algorithms can be scaled for payloads of roughly 5 to 25 kilograms. One can envision deploying robots from the South Pole to locations on the plateau for long-term or short-term observation and retrieving or repositioning the network through iridium-based communication. Potential missions include: deploying arrays of magnetometers, seismometers, radio receivers, and meteorological instruments; measuring disturbances in the ionosphere through synchronization of Global Positioning System (GPS) signals; using ground-penetrating radar (GPR) to survey crevasse-free routes for field parties; and conducting glaciological surveys with GPR. Robot arrays could also provide high-bandwidth communications links and power systems for field scientists.

Based on this concept, a single robot is under construction that would carry a triaxial fluxgate magnetometer, an iridium modem, and a modest set of weather instruments. The magnetometer will be a payload test case. Magnetometer arrays are already used in low latitudes and midlatitudes, but polar regions provide unique windows for observing the effects of solar wind on the Earth's magnetosphere. With mobile networks, sensor locations could potentially be tuned to events in the magnetosphere. Also, synchronized data from polar networks could potentially discover spatial characteristics of narrow-band spectral features in geomagnetic field data, identify magnetospheric boundaries, and refine models accordingly (Lanzerotti et al., 1999).

The deployment of a remote observatory on the Antarctic plateau via transport and small aircraft is expensive and entails dangerous takeoffs and landings at remote sites. For large-scale, widely distributed (>500 km radius) networks, relatively low-cost mobile robots could substantially reduce per-instrument deployment costs. Semiautonomous network deployment would also free limited aircraft and human resources for other missions.

The harsh weather of polar environments, long-range requirements, navigation issues, and variable terrain pose significant design challenges for inexpensive unmanned vehicles. Instruments would be deployed for long periods of time in drifting snow and will require a stable environment with low vibration and electromagnetic noise. Robots and deployed sensors should be retrievable with high reliability to minimize environmental impact and cost. In this paper, related robotics research is summarized, including mobile robot design concepts for polar environments, technical challenges associated with their development, and enabling technologies for cost-effective mobile robots.

State-of-the-Art Robots for Extreme Environments

NOMAD, a gasoline-powered robot for polar and desert environments, was developed at Carnegie Mellon University (Apostolopoulos et al., 2000; Carnegie Mellon University, 2004b). NOMAD, which is 2.4 × 2.4 × 2.4 m in size and weights 725 kg (Figure 1), can travel up to 50 centimeters per second (cm/s) and deploy instruments, such as a magnetometer. In 1997, NOMAD executed its first mission in the Atacama Desert of southern Chile, traversing 223 km through teleoperation. Subsequently, NOMAD successfully found and classified five indigenous meteorites on an Antarctic mission. For our purposes, however, NOMAD's large size, high cost, and nonrenewable fuel are disadvantages. In addition, its deployment experience suggests that navigation cameras may work poorly in polar climates due to reflection of sunlight off of the snowfield.

Spirit and Opportunity are Mars exploration rovers developed by the National Aeronautics and Space Administration Jet Propulsion Laboratory. Each 2.3 × 1.6 × 1.5-m rover weighs 174 kg and has a top speed of

FIGURE 1 NOMAD gasoline-powered robot. Source: Carnegie Mellon University, 2004b.

5 cm/s (NASA/JPL, 2004). The power source for Mars rovers, a multipanel solar array and two rechargeable lithium-ion batteries, enables the rover to generate 140 W of power for four hours per sol, when the panels are fully illuminated. The warm electronics box, which contains the batteries, electronics, and computer, can only operate in the range of –40°C to +40°C. Gold-painted, insulated walls, solid silica aerogel, thermostat and heaters, and a heat rejection system protect the body from the 113°C temperature swing during the Martian day. The payload includes a panoramic camera, a miniature thermal emission spectrometer, a Mössbauer spectrometer, an alpha-particle X-ray spectrometer, and a rock abrasion tool. Each of the six wheels is driven by its own in-wheel motor, and the two front and two rear wheels have steering motors for point turns. The rovers are well suited for their mission on Mars but are too expensive for the deployment of instrument networks.

Hyperion (Figure 2), designed for sun-synchronous exploration, is a 157-kg, 2 × 2.4 × 3-m vehicle that includes a 3.45-m^2 nearly vertical solar panel; its maximum speed is 30 cm/s (Carnegie Mellon University, 2004a; Wettergreen et al., 2001). The chassis is intentionally simple—a 1.5 N-m, 150 W brushless DC motor combined with a harmonic drive for an 80:1 reduction ratio drives a wheel through a bicycle chain. A passively articulated steering joint provides two free rotations, enabling moderate maneuverability and mechanical simplicity. This design has many appealing features for instrument-network deployment, including renewable fuel, simplicity, and potentially low cost.

Commercial all-terrain robots made by iRobot and ActivMedia weigh 39 to 100 kg and carry payloads of 7 to 100 kg. Powered by two DC servomotors and a 4-wheel differential-drive system, these battery-operated robots run for 2 to 6 hours at speeds between 1 and 2 m/s. With no navigation instruments or software, these robots cost from $7,000 to $22,000. They could potentially be retrofitted for solar-power operation, but they are not designed for low-temperature operation, and the solar panels alone could comprise the entire payload budget.

Cool Robot Concept

Our robot is designed to operate in interior Antarctica, which is characterized by low snowfall, moderate winds, and extreme cold. We envision networks of robots, guided by GPS and onboard sensors, that are launched and retrieved from the South Pole Station during the austral summer. Key design issues are outlined below.

FIGURE 2 Hyperion rover. Source: Carnegie Mellon University, 2004a.

Figure 3 shows a satellite photo of Antarctica. The vast central plateau covers more than five million square kilometers of relatively flat, crevasse-free terrain. A second large area of operation is the Ross Ice Shelf. Generally, Antarctic snowfields consist of dense, wind-blown snow. Aside from wind-sculpted *sastrugi*, dune-like features that are identifiable on satellite imagery, there are few obstacles. The central plateau receives less than 50 mm precipitation (<500 mm snowfall) in an average year. During summer months at the South Pole, wind speed averages 2 m/s (Valenziano and

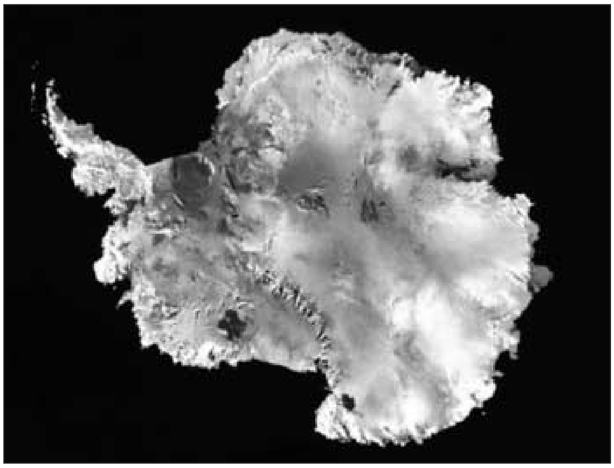


FIGURE 3 Satellite photo of Antarctica. Source: USGS, 2004.

Dall'Oglio, 1999), and the five-year maximum speed is 20.5 m/s (CMDL, 2004). The average daily temperature is –20°C to –40°C.

An Antarctic robot must traverse firm snow and occasional softer drifts, sustain mobility in windy conditions, have a minimal environmental impact, and operate in temperatures down to –40°C. We envision a lightweight, solar-powered, wheeled robot that can be transported in a Twin Otter aircraft and is capable of traversing 500 km within two weeks. After reaching a target location, the robot could collect data over a period of two to three months before returning to the South Pole for the winter. The design includes a low center-of-gravity vehicle with four direct-drive brushless electric motors, an enclosed thermally controlled volume for instrumentation and batteries, and a solar panel "box" for renewable energy. Table 1 provides design specifications for a wheeled robot and a price point for economic viability for the deployment of networks of such robots.

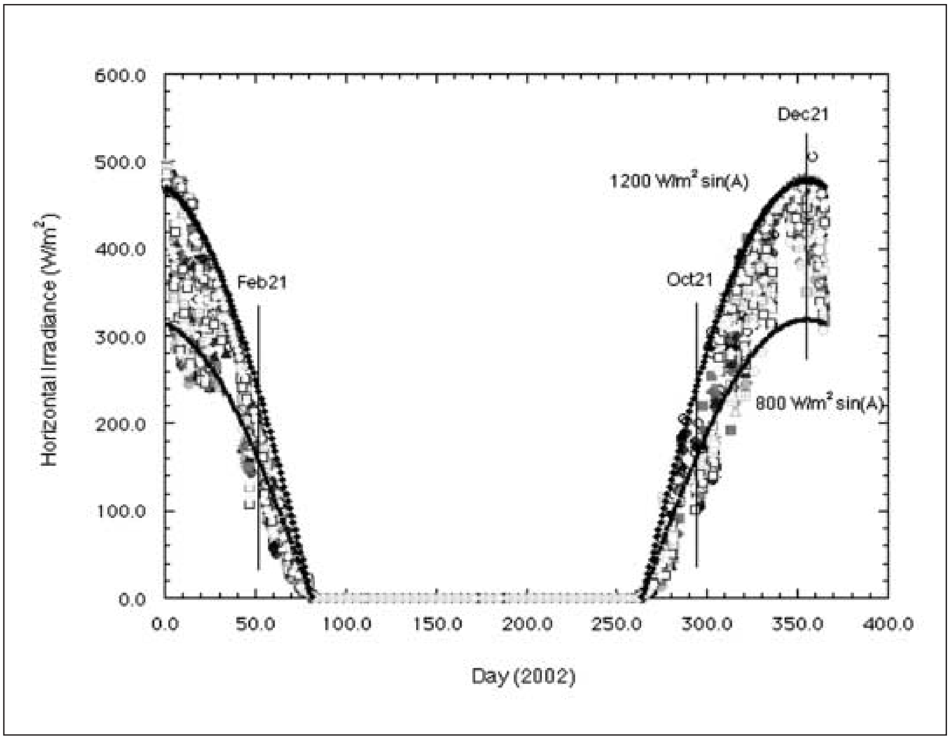


FIGURE 4 Daily average horizontal insolation at the South Pole, 2002. Source: CMDL, 2004.

Motion resistance in snow is attributable to sinkage and varies with the firmness of the snow pack immediately in front of the wheel, the length of the tire in contact with the snow, and the width of the tire (Richmond et al., 1995). Given the target ground pressure (< 20 kPa) in the dense snow of the Antarctic plateau, sinkage should be small. The estimated total resistance of 0.25 for a 90-kg vehicle will require a net traction force of 221 N. Travel of 500 km in two weeks will require an average speed of 0.41 m/s, with an average power requirement of 90 W and maximum power of 180 W for the top speed of 0.8 m/s. Allowing up to 40 W for housekeeping power and power-system efficiencies, the target power budget is approximately 220 W.

TABLE 1 Robot Specifications

Maximum speed	≥ 0.80 m/s
Mass (excluding payload)	≤ 75 kg
Payload mass	≥ 15 kg
Ground pressure	≤ 20 kPa
Operating temperature range	0°C to –40°C
Dimensions	≤ 1.4 × 1.15 × 1-m
Cost	≤ $20,000

Despite a harsh climate and low sun angles, Antarctica is an ideal place for a solar-powered robot. The summer sun provides a 24-hr energy source, and the central plateau receives scant precipitation and infrequent fog. The Antarctic plateau is nearly completely covered in snow, with albedo averaging 95 percent across visible and ultraviolet wavelengths (Grenfell et al., 1994) and fairly uniform scattering in all directions (Warren et al., 1998). The high altitude and dry air block less incoming radiation, and there is a small benefit due to the proximity of the Earth to the sun during the summer. The sun remains at approximately the same elevation throughout the day (especially near the South Pole), resulting in relatively constant energy input. With low elevation angles and significant reflected solar energy, nearly vertical solar panels will be optimal. Also, the efficiency of solar cells increases as temperature decreases.

Average horizontal insolation data for 2002 (Figure 4) show a range of horizontal irradiance of 300 to 500 W/m^2 at the South Pole (CMDL, 2004). Adjusting for sun elevation angle gives a net insolation between 800 W/m^2 and 1,200 W/m^2, which is consistent with earlier studies (Hanson, 1960). At other locations in Antarctica where cloud cover and fog are more frequent, the average insolation is about half this, but the sunny days are almost as bright. For comparison, consider a clear winter day in New England. At a sun elevation of 20 degrees, the total insolation is between 600 and 800 W/m^2. Table 2 summarizes insolation data at various locations on the Antarctic continent and elsewhere. The average solar energy input during the November to February operating window is approximately 1,000 W/m^2, with an average sun elevation of about 20 degrees.

TABLE 2 Insolation for Various Sun Conditions

Condition/Location	Nominal Insolation
Maximum for Antarctica, continent	1,200 W/m^2
Average for South Pole, Nov–Feb	1,000 W/m^2
Average for South Pole, at solstice	1,100 W/m^2
Average for south polar plateau	> 800 W/m^2
Average for Ross Ice Shelf	400 W/m^2
January 2004 measurements, Hanover, New Hampshire	660 W/m^2

To determine the optimum size of the solar panel, we developed a model to predict power as a function of solar insolation, sun elevation, and azimuth for solar panels in a snowfield. The model assumes diffuse reflection from the snow at a specified albedo. We validated the model using data collected with a commercial 20-W panel during January-February 2004 in Hanover, New Hampshire. Figure 5 shows the resulting robot design concept—a wheeled chassis enclosed by a five-panel box—along with predicted panel capacities extrapolated from the model for nominal Antarctic solar radiation, 20 degree sun elevation, and 90 percent albedo. The panel outputs are reported as a percentage of their standard capacities (rated at 1,000 W/m^2 insolation). The front panel (directly facing the sun) has a capacity of 128 percent (more than 100 percent due to reflected energy). Significantly, the top and side panels contribute nearly as much power as the panel facing the sun. Even the back panel receives substantial radiation because the robot's shadow is not as large as the area of snow that reflects light to the panel.

Enabling technology for the robot is the affordable, 20 percent efficient, A-300 solar cell by Sunpower, Inc., which became available in 2003. Figures 6 and 7 show predicted power available to the motors for a robot using 54 of these cells per panel (each cell is 12.5 × 12.5 cm) at 1,000 W/m^2 insolation and 90 percent albedo. The resulting robot will fit in the Twin Otter cargo bay. These results include efficient maximum-power-point-tracking (MPPT) circuits for each panel and subtract housekeeping power. The robot can drive at full speed even in below-average insolation. Under minimal insolation, the robot still has enough power to drive slowly or charge the batteries and drive in short bursts on battery power. Diffuse incoming radiation—light scattered by the atmosphere—is an unmodeled benefit because diffuse light is expected to be received perpendicular to the panels from all directions. The total diffuse radiation is expected to be 50 to 100 W/m^2 providing 40 to 80 W, enough to drive the robot in bursts and maintain instrument operation (Hanson, 1960).

A positive feature of the deployment of instruments

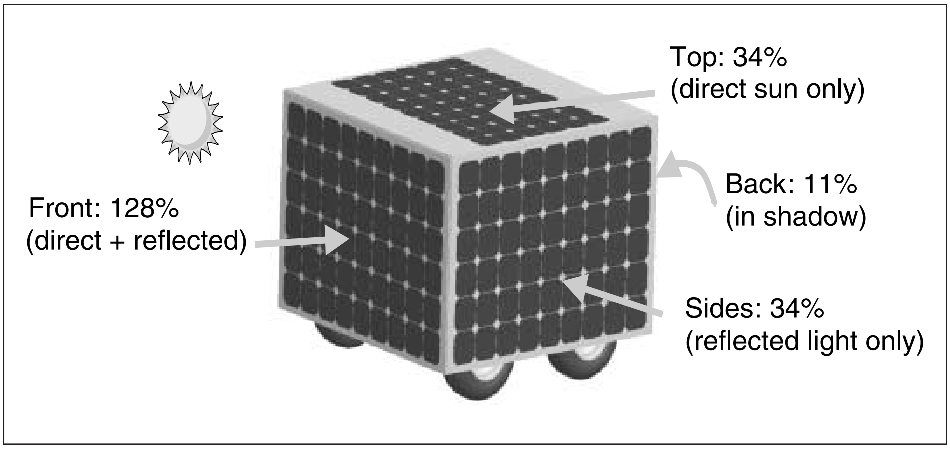


FIGURE 5 Panel power capacities in nominal Antarctic sun.

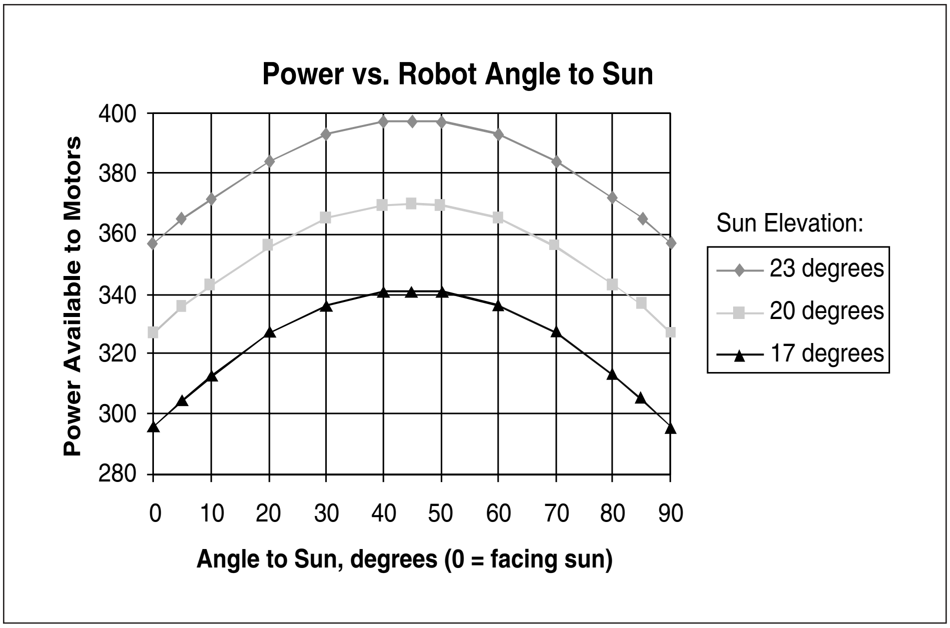


FIGURE 6 Power availability as a function of sun elevation angle and the robot azimuth angle.

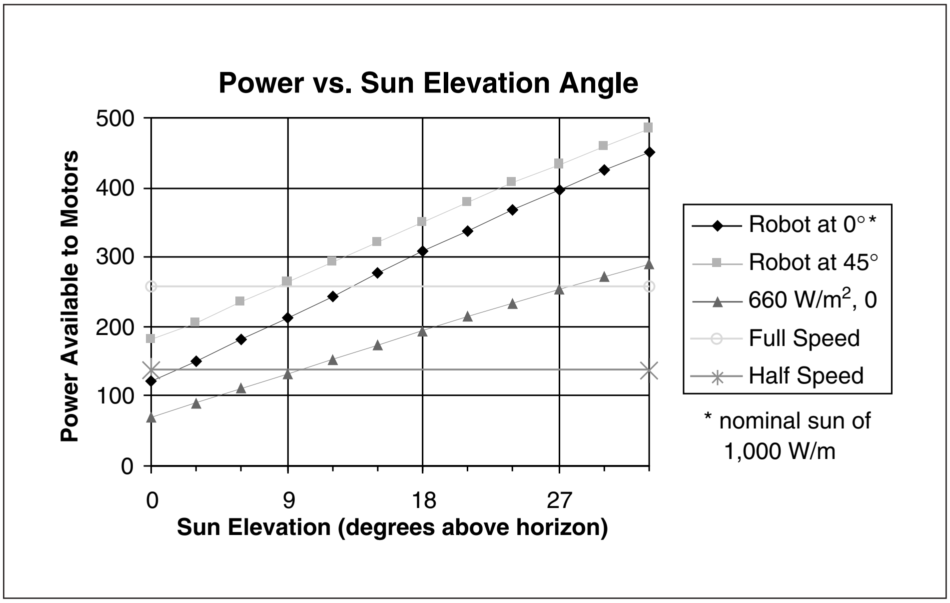


FIGURE 7 Power available and required at various sun angles.

on the Antarctic plateau is relatively flat, straight paths on dense snow, which minimizes planning the complexity of the path. One of the most promising navigation architectures for such paths is "mixed-mode" operation, which mimics human behavior (e.g., in hiking a known path over a long distance) (Simmons et al., 1995). The global objective is to stay on the path, but a local mode, in which the hiker goes around unanticipated objects (e.g., downed trees), is in force for short periods, after which the hiker returns to the path. In the initial stages of research, global navigation is being used, primarily through GPS and speed control, with sensors to detect unbalanced wheel speeds and hence potential problems with traction, and low-bandwidth path-correction algorithms to reduce "dither" around the path. In the traversal of long distances, GPS-induced path deviations are tolerable. Traction control can be layered onto the basic global path-following algorithm, along with sensors for tilt and wheel slip. Local-mode navigation would be invoked if sensors detect extreme tilting or slippage.

Navigation and motion control also depend on the power system. The robot will move along the path only when adequate solar power is available. In cloudy conditions, the robot will move under battery power if necessary to prevent drifting in, and in windy conditions, the robot may face a direction that minimizes drag and the potential for tipping. We do not anticipate the robot having a vision system because sastrugi are visible on satellite imagery and can be avoided through route selection. The lack of contrast in snow-covered terrain would make onboard navigation challenging and potentially expensive.

Design Embodiment and Enabling Technologies

We have attempted to minimize vehicle mass by using stiff honeycomb composites for the solar panels and chassis and custom-designed wheel rims and hubs. The enabling technologies and cost-design trade-offs and an estimate of the costs of parts and the mass for the prototype robot are highlighted below.

Because of their newness, complete panels for A-300 solar cells are not yet available. Moreover, traditional panel construction with its steel backing is not viable for the robot. Thus, we will construct the solar panels in house, using quarter-inch honeycomb sandwich panels (Nomex core with fiberglass facing). The cell will be encapsulated in silicone. Similar honeycomb composite will be used to make the chassis box (Figure 8). The construction and joinery of honeycomb panels are mature technologies in the aerospace field, and these materials supply area densities of 1.4 kg/m^2 for solar panels and 2.5 kg/m^2 for chassis walls.

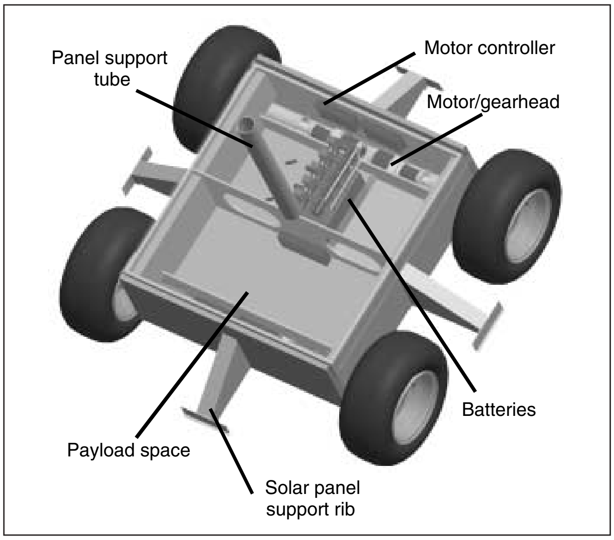


FIGURE 8 Internal chassis and components.

The wheels are sized to provide adequate ground clearance and low rolling resistance. We considered many different tires and trade-offs between traction, weight, pressure, and rolling resistance. Taking into account availability and cost, a 16 × 6-8 ATV tire, which has low mass and excellent traction, was selected. Low-mass rims and hubs are not available for ATV tires, however, so these were custom designed and machined in house to meet the strength and deflection requirements for a 90-kg robot. For a mass-produced robot, rims and hubs could be cast or stamped at low cost.

High-efficiency, brushless motors with 90 percent efficient gear trains and lubricant for operation at –50°C drive the wheels directly. Each motor has a controller that can be configured for speed or torque control. We tested a single motor-gearbox combination in a cold room in a box insulated as configured on the robot. The results showed that both long-term operation and start-stop operation would be efficient, and controllability would be maintained at cold temperatures.

The power-system architecture includes three lithium-ion batteries in series and five solar panels, each of which can operate under varying insolation and temperature conditions. To deliver power to a common power bus, each panel requires an MPPT, a device that allows each panel to operate at the bus voltage established by the batteries while meeting power demands for the motors. Custom-designed MPPTs that weigh less than 250 grams and have better than 97 percent efficiency have been designed and constructed.

We estimate that the five-panel robot, without payload, will weigh ~73 kg, with a total material cost of less than $15,000. The design is relatively insensitive to payload up to about 20 kg. The robot will be equipped with instruments to assess power-input and mobility models during field trials in Antarctica, anticipated for the 2005–2006 austral summer.

Conclusion

Solar-powered mobile robots for operation on the Antarctic plateau are feasible in terms of power availability, mechanical design, and power-system design. Waypoint navigation on the relatively obstacle-free plateau through GPS can provide long-distance travel appropriate to the scientific missions envisioned. Mobile robots capable of reliable, long-term operation on the Antarctic plateau can potentially advance scientific research through instrument deployment, mapping, and the provision of portable mobile power.

References

Apostolopoulos, D., M.D. Wagner, B. Shamah, L. Pedersen, K. Shillcutt, and W.L. Whittaker. 2000. Technology and field demonstration of robotic search for antarctic meteorites. International Journal of Robotics Research 19(11): 1015–1032.

Carnegie Mellon University. 2004a. Hyperion: Sun Synchronous Navigation. Available online at: *http://www.ri.cmu.edu/projects/project_383.html* (June, 2004).

Carnegie Mellon University. 2004b. Robotic Antarctic Meteorite Search: The NOMAD Robot. Available online at: *http://www.frc.ri.cmu.edu/projects/meteorobot/Nomad/Nomad.html#Mechanical* (June, 2004).

CMDL (Climate Monitoring and Diagnostic Laboratory). 2004. Daily Average Horizontal Insolation at the South Pole. Available online at: *http://www.cmdl.noaa.gov/info/ftpdata.html* (March, 2004).

Grenfell, T.C., S.G. Warren, and P.C. Mullen. 1994. Reflection of solar radiation by the Antarctic snow surface at ultraviolet, visible, and near-infrared wavelengths. Journal of Geophysical Research 99(D9): 18-669–18-684.

Hanson, K.J. 1960. Radiation studies on the south polar snowfield. IGY Bulletin 31: 1–7. Washington, D.C.: National Academy of Sciences.

Lanzerotti, L., A. Shona, H. Fukunishi, and C.G. Maclennan. 1999. Long-period hydromagnetic waves at very high geomagnetic latitudes. Journal of Geophysical Research 104(A12): 8, 423.

NASA/JPL (National Aeronautics and Space Administration Jet Propulsion Laboratory). 2004. Spacecraft: Surface Operations: Rover. Available online at: *http://www.marsrovers.jpl.nasa.gov/mission/spacecraft_rover_energy.html* (June, 2004).

NRC (National Research Council). 2002. The Sun to the Earth and Beyond: A Decadal Research Strategy in Solar and Space Physics. Washington, D.C.: National Academy Press.

Richmond, P.W., S.A. Shoop, and G.L. Blaisdell. 1995. Cold Regions Mobility Models. CRREL Report 95-1. Hanover, N.H.: Cold Regions Research and Engineering Laboratory. Available online at: *http://www.stormingmedia.us/82/8273/A827392.html.*

Simmons, R., E. Krotkov, L. Chrisman, F. Cozman, R. Goodwin, M. Hebert, L. Katraqadda, S. Koenig, G. Krishnaswamy, Y. Shinoda, W. Whittaker, and P. Klarer. 1995. Experience with Rover navigation for lunar-like terrains. Journal of Engineering and Applied Science 1: 441–446.

USGS (U.S. Geological Survey). 2004. Satellite Image Map of Antarctica. Available online at: *http://terraweb.wr.usgs.gov/TRS/projects/Antarctica/AVHRR.html* (June, 2004).

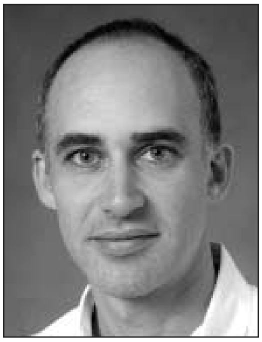
Valenziano, L. and G. Dall'Oglio. 1999. Millimetre astronomy from the high Antarctic plateau: site testing at Dome C. Publications of the Astronomical Society of Australia 16: 167–174.

Warren, S.G., R.E. Brandt, and P. O'Rawe Hinton. 1998. Effect of surface roughness on bidirectional reflectance of Antarctic snow. Journal of Geophysical Research 103(E11): 25-789–25-807.

Wettergreen, D., B. Shamah, P. Tompkins, and W.L. Whittaker. 2001. Robotic Planetary Exploration by Sun-Synchronous Navigation. In Proceedings of the 6th International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS 2001), Montreal, Canada, June 2001.

Each generation of landing technology addresses the challenges posed by the previous generation.

The Challenges of Landing on Mars



Tommaso Rivellini is the lead mechanical engineer for Mars Science Laboratory Entry, Descent, and Landing, at the Jet Propulsion Laboratory, California Institute of Technology.

Tommaso Rivellini

People have been fascinated with the idea of exploring Mars since the very beginning of the space age. Largely because of the belief that some form of life may have existed there at one time, surface exploration has been the ultimate ambition of this exploration. Unfortunately, engineers and scientists discovered early on that landing a spacecraft on the surface of Mars would be one of the most difficult and treacherous challenges of robotic space exploration.

Upon arrival at Mars, a spacecraft is traveling at velocities of 4 to 7 kilometers per second (km/s). For a lander to deliver its payload to the surface, 100 percent of this kinetic energy must be safely removed. Fortunately, Mars has an atmosphere substantial enough for the combination of a high-drag heat shield and a parachute to remove 99 percent and 0.98 percent respectively of the kinetic energy. Unfortunately, the Martian atmosphere is not substantial enough to bring a lander to a safe touchdown. This means that an additional landing system is necessary to remove the remaining kinetic energy.

On previous successful missions, the landing system consisted of two major elements, a propulsion subsystem to remove an additional 0.002 percent (~50 to 100 meters per second [m/s]) of the original kinetic energy and a dedicated touchdown system. The first-generation Mars landers used legs to accomplish touchdown. The second generation of touchdown systems used air bags to mitigate the last few meters per second of residual velocity. The National Aeronautics and Space Administration (NASA) is currently developing a third-generation landing system in an effort to reduce cost, mass, and risk while simultaneously improving performance as measured by payload fraction to the surface and the roughness of accessible terrain.

Legged Landing Systems

The legs of the 1976 Viking mission lander represent the first-generation landing system technology (Pohlen et al., 1977). Basic landing-leg technology was developed for the lunar Surveyor and Apollo programs in the early 1960s. In conjunction with a variable-thrust liquid propulsion system and a closed-loop guidance and control system, legs represented an elegant solution to the touchdown problem. They are simple, reliable mechanisms that can be added to an integrated structure that houses the scientific and engineering subsystems for a typical surface mission (Figure 1).

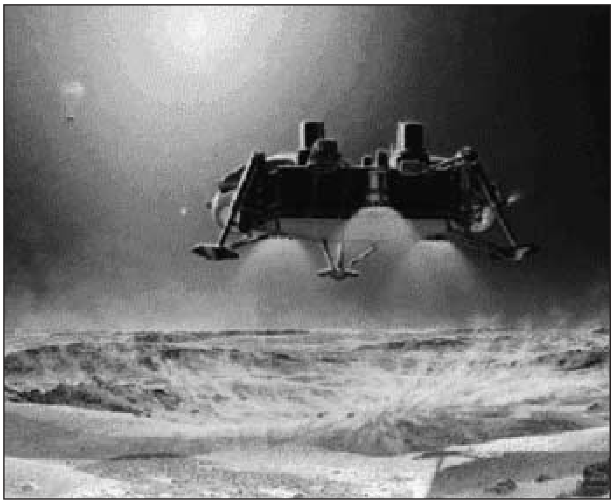


FIGURE 1 First-generation landing system used on the Viking lander, which landed on Mars in 1976.

The first challenge for a legged system is to enable the lander to touch down safely in regions with rocks. For this the legs must either be long enough to raise the belly of the lander above the rocks, or the belly of the lander must be made strong enough to withstand contact with the rocks. Neither solution is attractive. Either the lander becomes top heavy and incapable of landing on sloped terrain or a significant amount of structural reinforcement must be carried along for the remote chance that the lander will directly strike a rock. The decreased stability because of the high center of mass is exacerbated if a mission carries a large rover to the surface. Because of the rover's configurational requirements, it is typically placed on top of the lander. The Soviet Lunokhod lunar landers (Figure 2) are an excellent example of this type of configuration.

A second major challenge of the legged-landing architecture is ensuring safe engine cutoff. To prevent the guidance and control system from inadvertently destabilizing the lander during touchdown, contact sensors have been used to shut down the propulsion system at the moment of first contact. On sloped terrain, this causes the lander to free fall the remaining distance, which can significantly increase the total kinetic energy present at touchdown and, in turn, decrease landing stability and increase mission risk. Implementation and testing of fault protection for engine cutoff logic has been, and continues to be, a difficult problem.

The first in-flight problem associated with engine shut off occurred on the lunar Surveyor lander mission when the propulsion system failed to shut off at touchdown, resulting in a significant amount of postimpact hopping. Fortunately, the terrain was benign, and the problem was not catastrophic. The second in-flight problem occurred on the Mars 98 lander mission when the engines were inadvertently shut off prematurely because of a spurious contact signal generated by the landing gear during its initial deployment. This problem resulted in a catastrophic loss of the vehicle. As a result, the Apollo missions all reverted to a man in the loop to perform engine shut off.

A third major challenge with a legged landing system for missions with rovers is rover egress. Once the lander has come to rest on the surface, the rover must be brought to the surface. For legged landers, a ramped egress system is the most logical configuration. Because rovers are bidirectional, the most viable arrangement has been considered two ramps, one at the front and one at the rear of the lander. The Soviet Lunokhod missions landed in relatively benign terrain, and in all cases, both ramps were able to provide safe paths for the rover. In the Mars Pathfinder mission, one of the two ramps was not able to provide a safe egress path for the Sojourner rover, but the second ramp did provide safe egress. For vehicles designed to explore a larger fraction of the Martian surface and, therefore, land in more diverse terrain, combinations of slopes and rocks could conceivably obstruct or render useless the two primary egress paths.

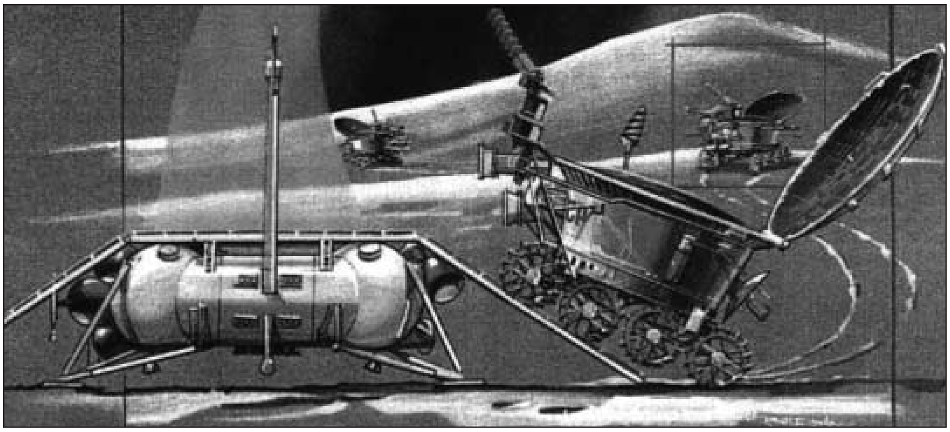


FIGURE 2 Lunokhod Soviet lunar rover leaving the legged lander that delivered it to the surface.

Air-Bag Landing Systems

The second-generation landing system was developed for the Mars Pathfinder mission and subsequently improved upon for the Mars Exploration Rover (MER) missions (Figure 3). These second-generation systems have a combination of fixed-thrust solid rocket motors and air bags to perform the touchdown task. The solid rocket motors, which are ignited two to three seconds prior to impacting the surface, slow the lander down to a stop 10 meters above the surface, from an initial velocity of approximately 120 meters/second. The lander is then cut away from the over-slung rockets and free falls for the remaining distance.

The air-bag system, which was developed to reduce cost and increase landing robustness, is designed to provide omnidirectional protection of the payload by bouncing over rocks and other surface hazards. Because the system can also right itself from any orientation, the challenge of stability during landing has been completely eliminated. Because the lander comes to rest prior to righting itself, the challenge of rock strikes has been reduced to strikes associated with the righting maneuver, which are significantly more benign. The challenge of thrust termination, in this case cutting the lander away from the rockets, remains but has been decoupled from the problem of landing stability. The problems of rover egress were addressed systematically on the MER missions; a triple ramp-like system provided egress paths in any direction, 360 degrees around the lander.

Although the air-bag landing system has addressed some of the challenges and limitations of legged landers, it has also introduced some challenges of its own. Horizontal velocity control using solid rockets and air-bag testing were significant challenges for both the Mars Pathfinder and MER missions.

The Sky-Crane Landing System

As Mars surface explorations mature, roving is becoming more important in the proposed mission architectures. The MER missions demonstrated the value of a fully functional rover not reliant on the lander to complete its surface mission. In the 2009 Mars Science Laboratory (MSL) and other future missions, the rover's capabilities and longevity will be extended. Future missions are also being designed to access larger areas of the planet and, therefore, will require more robust landing systems that are tolerant to slope and rock combinations that were previously considered too hazardous to land or drive on. The third-generation landing system, the sky-crane landing system (SLS), currently being developed for the MSL mission, will directly address all of the major challenges presented by

FIGURE 3 Second-generation landing system used on the Mars Pathfinder and MER landers.

the first- and second-generation landing systems. It will also eliminate the problem of rover egress.

SLS eliminates the dedicated touchdown system and lands the fully deployed rover directly on the surface of Mars, wheels first. This is possible because the rover is no longer placed on top of the lander. In the SLS, the propulsion module is above the rover, so the rover can be lowered on a bridle, similar to the way a cargo helicopter delivers underslung payloads (Figure 4).

The landing sequence for future missions will be similar to the Viking mission, except for the last several seconds when the sky-crane maneuver is performed (Figure 5). After separating from the parachute, the SLS follows a Viking-lander-like propulsive descent profile in a one-body mode from 1,000 meters above the surface down to approximately 35 meters above the surface. During this time, a throttleable liquid-propulsion system

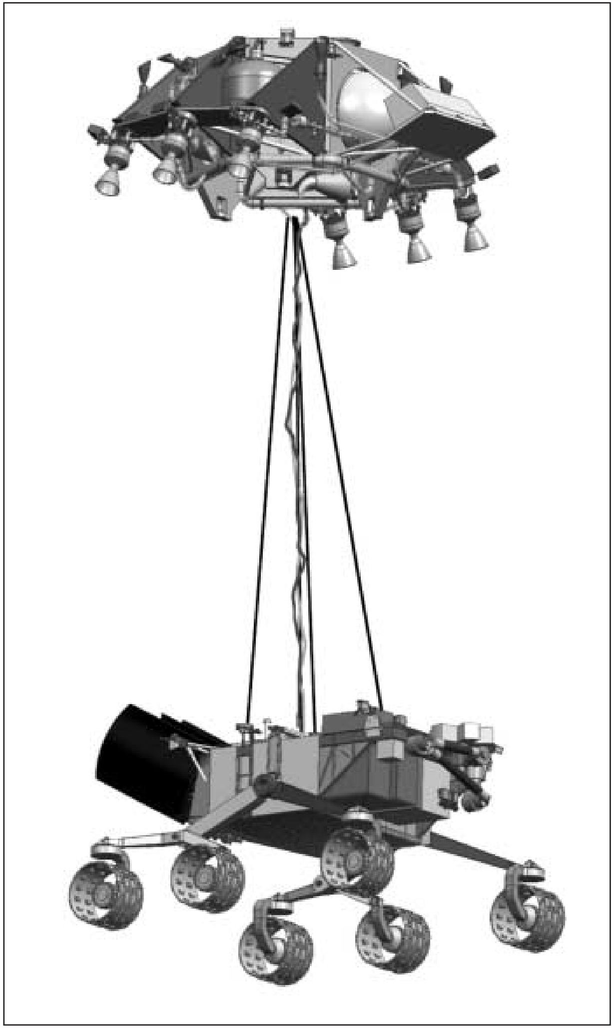


FIGURE 4 Sky-crane landing system shown with the rover already deployed.

coupled with an active guidance and control system controls the velocity and position of the vehicle. At 35 meters, the sky-crane landing maneuver is initiated, and the rover is separated from the propulsion module. The rover is lowered several meters as the entire system continues to descend. The two-body system then descends the final few meters to set the rover onto the surface and cut it away from the propulsion module. The propulsion module then performs an autonomous fly-away maneuver and lands 500 to 1,000 meters away.

The central feature of the SLS architecture is that the propulsion hardware and terrain sensors are placed high above the rover during touchdown. As a result, their operation is uninterrupted during the entire landing sequence. One important result of this feature is that the velocity control of the whole system is improved, and, therefore, the rover touches down at lower velocity. Thus, there is no last-meter free fall associated with engine cutoff, and, because dust kick-up is minimal, the radar antennas can continue to operate even while the rover is being set down on the surface.

The lower impact velocity has two effects. First, the touchdown velocities can now be reliably brought down to the levels the rover has already been designed for so it can traverse the Martian surface. Second, the low velocity, coupled with the presence of bridles until the rover's full weight has been transferred to the surface, results in much more stability during landing.

Because the rover does not have to be protected from the impact energy at landing and because there is no need to augment stability at landing, there is no longer a need for a dedicated touchdown system. This, in turn, eliminates the need for a dedicated egress system. The SLS takes advantage of the fact that the rover's mobility system is inherently designed to interact with rough, sloping natural terrain. Rovers are designed to have high ground clearance, high static stability, reinforced belly pans, and passive terrain adaptability/conformability. These are all features of an ideal touchdown system.

Touchdown Sensing

Touchdown sensing can be done in several ways. The simplest and most robust way is to use a logic routine that monitors the commanded up-force generated by the guidance and control computer. The landing sequence is specifically designed to provide a constant descent velocity of approximately 0.75 m/s until touchdown has been declared. Prior to surface contact, the commanded

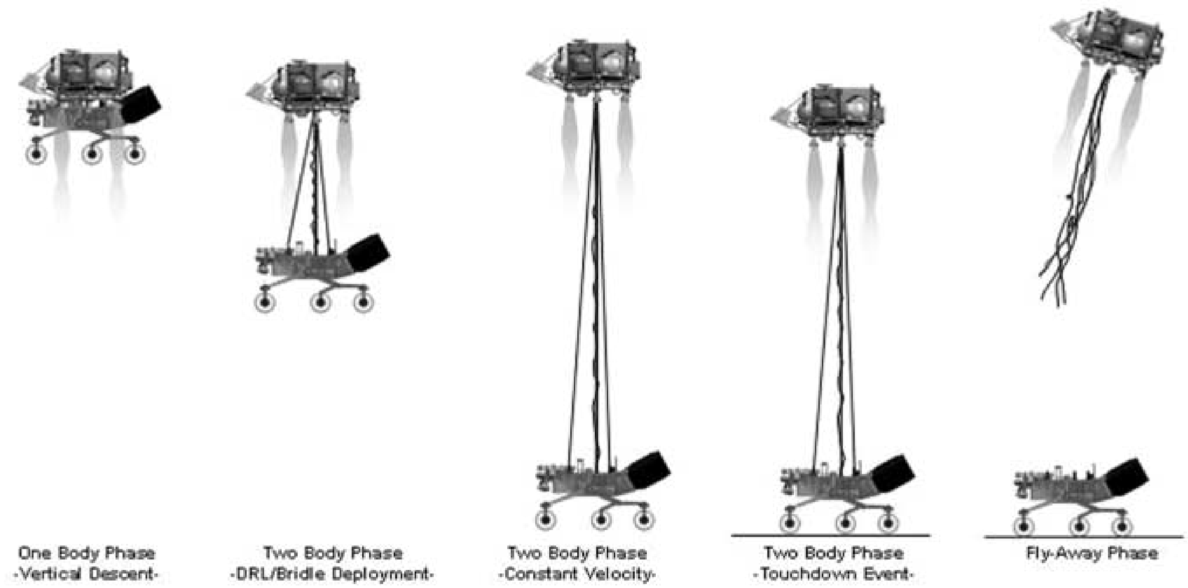


FIGURE 5 Sky-crane landing sequence showing the three main phases and events. Nominal touchdown velocities are 0.75 m/s vertically and 0.0 m/s horizontally.

up-force is equal to the mass of the rover plus the mass of the descent stage (which are roughly equal) times the gravity of Mars. During the touchdown event, the commanded up-force fluctuates depending on the specific geometry of the terrain.

Once the weight of the rover has been fully transferred to the surface of Mars, the commanded up-force takes on a new steady-state value equal to the mass of the descent stage times the gravity of Mars, approximately one-half of its pretouchdown magnitude. The system declares touchdown after the new lower commanded up-force has lasted for at least 1.5 seconds. This approach provides an unambiguous touchdown signature without the use of dedicated sensors.

The fly-away phase of the landing sequence is initiated when touchdown has been declared. During the fly-away phase, separation of the rover is accomplished by the pyrotechnic cutting of the bridle and umbilical lines connecting the rover and descent stage. The descent stage then uses its onboard computer to guide the propulsion module up and away from the rover and land it several hundred meters away.

Conclusion

As Mars explorers have learned the hard way, it's not typically the fall that kills you, it's the landing. Landing technology has matured significantly in the 40 years since NASA began exploring extraterrestrial surfaces. Each generation of landing technology has attempted to resolve the challenges posed by the previous generation. The SLS represents the latest stage in that evolution.

Acknowledgment

The research described in this paper was carried out at the Jet Propulsion Laboratory of the California Institute of Technology under a contract with the National Aeronautics and Space Administration.

Reference

Pohlen, J., B. Maytum, I. Ramsey, and U.J. Blanchard. 1977. The Evolution of the Viking Landing Gear. JPL Technical Memorandum 33-777. Pasadena, Calif.: Jet Propulsion Laboratory.

Multifunctional materials are emerging as a new interdisciplinary field.

The Future of Engineering Materials: Multifunction for Performance-Tailored Structures

Leslie A. Momoda is director of the Sensors and Materials Laboratory at HRL Laboratories, LLC, in Malibu, California.

Leslie A. Momoda

In the future, new functional and reduced-scale materials that are currently in the forefront of technology will be hybridized into designer materials that can perform dramatic "tailorable" functions in large engineered systems. These performance-tailored structures will have the ability to change or adapt the performance or style of a structure on demand. Today, engineers can imagine designing adaptive flight profiles from morphing aircraft-wing structures; comfort-tailored performance, such as active structural vibration and noise suppression or temperature compensation, from louvered or pore-based "smart skins"; energy-efficient structures, such as tropical-plant-inspired solar structures; adaptive structures that can compensate for distortion or heal themselves; and structures reconfigured to satisfy style preferences. Imagine, for example, being able to commute to work in a stately professional car that can be reconfigured into a sportier car for the weekend.

As system-operating scenerios become more constrained by space and logistics limitations, the ability to adapt a structure's performance at will is becoming increasingly attractive. Currently multi-mission objectives are met with multiple structures (i.e., one car to drive to work and another one for the weekend). These solutions work if there is excess capacity in the system (e.g., a two-car garage), but as the number of mission objectives increases, the procurement, storage, and maintenance of a large number of structures become prohibitive. As a consequence, engineered subsystems that provide structural adaptability are under development in several programs, such as the Defense Advanced Research Projects Agency (DARPA) Morphing Aircraft Structures Program (Wax et al., 2003), the General Motors Autonomy Concept (Burns et al., 2002), and many structural health-monitoring programs. All of these programs are designed to provide tailored performance in large multicomponent system structures.

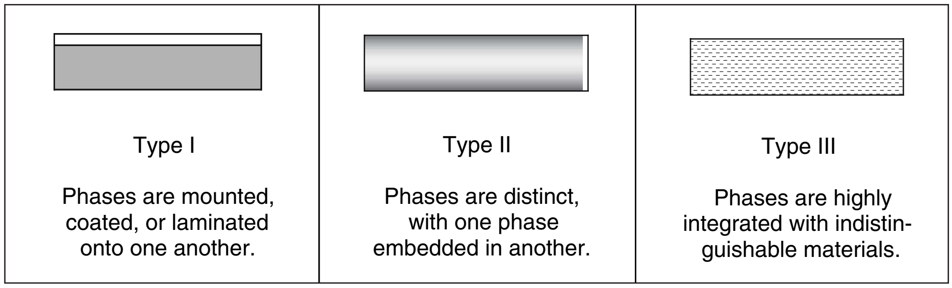


FIGURE 1 Schematic representations of the three types of multifunctional materials. Source: Matic, 2003.

Researchers are now thinking about how these same functionalities can be achieved in materials used to construct the system themselves, for example, a thin, smart-material skin that undergoes a radical but controlled change under mechanical strain; a coating that changes color on demand; a shell that reconfigures its shape to meet styling or performance criteria. These materials would enable the same system-level goals that are currently designed as subsystems, but could be more easily integrated into larger engineered structures because they will be lighter, smaller, less difficult to interface, and easier to maintain than current subsystems. Driven by recent advances in biomaterials and nanotechnology, multifunctional materials are emerging as a new interdisciplinary field that promises to provide a new level of functionality, adaptability, and tailorability for future engineered systems.

Multifunctional Material Systems

A multifunctional material is typically a composite or hybrid of several distinct material phases, in which each phase performs a different but necessary function, such as structure, transport, logic, or energy storage. Because each phase of the material performs an essential function, and because there is little or no parasitic weight or volume, multifunctional materials promise more weight-efficient, volume-efficient performance flexibility and potentially less maintenance than traditional multicomponent brass-board systems. The finer and more distributed the integration scale in the material, the faster and more autonomous the reaction times.

Multifunctionality in a material can be integrated on several dimensional scales with increasing interconnectivity between phases and engineering difficulty as the scale decreases. Matic (2003) has categorized these scales as different "types" (Figure 1). Type I material is comprised of phases in which one function is simply mounted, coated, or laminated to another, usually a structural component. Type II materials are comprised of distinct phases in which one function is embedded in another, usually a structural component. Type III materials are truly integrated; the phases are intermeshed, blurring the physical distinctions between them. The true promise of multifunctional materials for performance-tailored structures is most likely to be realized by Type III materials.

Multifunctional materials are designed for improved overall system performance. Thus, their performance metrics are inherently different from their single component phases, in which improvement of a single function, such as electrical conductivity, mechanical strain or force, or energy density, is maximized or minimized. A multifunctional material requires a new design methodology in which system-level performance is emphasized over the optimization of individual functions. Optimizing system-level performance involves optimization methodologies that are not commonly used in materials science.

Frontiers in Multifunctional Materials Technology

Materials Technology

The goal of multidisciplinary research on multifunctional materials, much of it under the auspices of the DARPA Synthetic Multifunctional Materials Program (Christodoulou and Venables, 2003), is to demonstrate weight and volumetric efficiencies and performance enhancements. Most research is focused on integrating two functions, usually a transport and structural function, and typically with low interconnectivity (e.g., Type I or Type II). Many of these studies rely heavily on inherently two-phase structural materials, such as fiber composites, laminates, foams, and other porous structures, as the matrix for multifunction. Even at this early stage, however, system-level benefits have been noted.

Structural batteries that reduce weight and complexity by integrating energy storage directly into the load-bearing structure have been developed by several teams using fibers, laminate, and nanotube construction. The energy density of the storage medium, such as a battery or supercapacitor, is reduced as the result of the incorporation of less conductive structural materials. However, the decrease in parasitic structure results in an overall weight savings and, therefore, improved energy density for the system as a whole.

The integration of actuation or sensing mechanisms into tailorable structural materials, which is essential for mechanical reconfigurability and structural morphing, is another area under investigation. Research using metallic foams or highly engineered mesostructured materials (dos Santos e Lucato et al., 2004) and elastomeric polymers (Pei et al., 2002) has shown tremendous promise for producing large structural motions along with the capacity to integrate actuation for feedback and control. Research conducted at HRL Laboratories, LLC, using hybrid, shape-memory materials (Figure 2), has demonstrated the feasibility of using integrated actuating tendons in a variably stiff matrix to produce low-energy flexure during reconfiguration but shape fixity for structural hold without energy when the motion is complete (W. Barvosa-Carter, personal communication). This concept has been extended to a hybrid laminate (Figure 3), which can produce a two-orders-of-magnitude change in stiffness, essentially from the stiffness of rubber to the hardness of steel, to accompany a shaping actuation force (G.P. McKnight, personal communication).

Self-healing composite structures are also under development (Chen et al., 2002). In these studies, a second phase, such as an adhesive or toughening agent, is added into the structure to ensure resealing on impact. Materials that can integrate other functions into structures, such as electromagnetism, thermal management, and optics capability, are also being investigated.

Optimization and Computational Design

Advances have been made in the development of optimization tools for designing integrated multifunctional materials. Sigmund and Torquato (1999) have done extensive work on topological optimization methods to determine the best morphological materials architectures to optimize performance from a highly integrated Type III material embedding of very dissimilar physical mechanisms. Many functional combinations, with as many as three phases, have been simulated. Although their work is purely theoretical, the results of their simulations have been validated by the similarity of the optimized topological solutions to microstructures and mesostructures found in biological systems. Macroscopic optimization tools for the design of less integrated Type I and II multifunctional materials have also been developed (Qidwai et al., 2002).

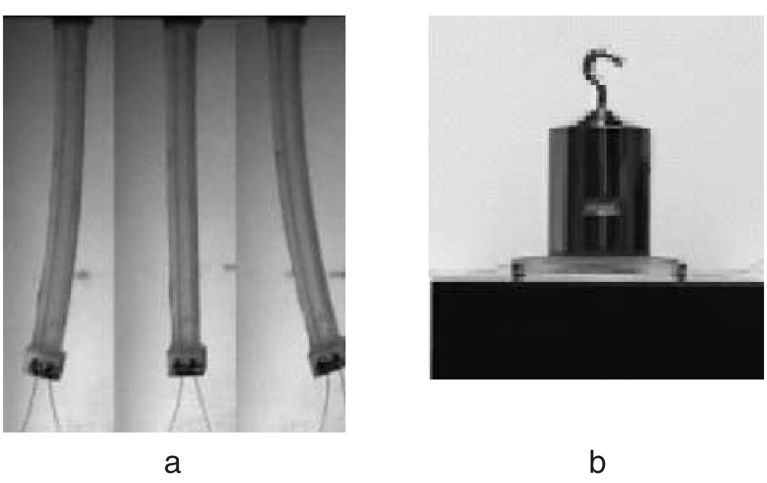


FIGURE 2 Hybrid shape-memory materials approach to reconfigurable structures. (a) Shape-memory tendons actuate a variably stiff shape-memory polymer beam that can be flexed with low-input energy. (b) Upon cooling, the beam becomes stiff, locking shape deformations in place to handle exterior loads (here 500 g).

Challenges and Prospects for the Future

Two-phase multifunctional systems show the promise of true materials integration. However, truly smart materials systems, analogous to biological systems, will require a combination of three or more functions, including logic, sensing, energy storage, structure, and actuation. Biological systems have perfected multifunction on a small scale. With the design of a priori multiple

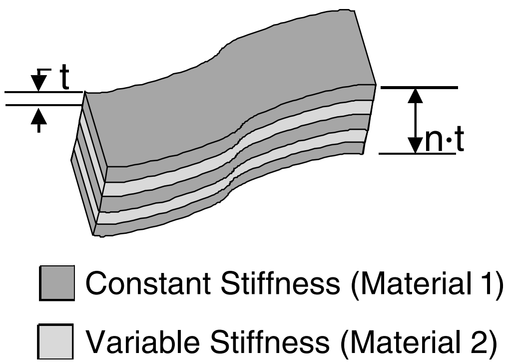


FIGURE 3 Variable stiffness laminate materials concept for producing large structural reconfigurations.

functions into a materials system, these concepts will be extended into large-scale structures. The complexities of these higher order systems will require a sophisticated understanding of how basic physical mechanisms can be manipulated to create new, potentially less singly optimal means of achieving function and multivariable optimization tools. For example, to increase the mechanical strength of energy storage systems, either the ionic conduction mechanism could be investigated for new logic capabilities or electrical conductivity mechanisms could be investigated to determine how they might influence mechanical strength.

As our understanding of materials on the nanoscale improves, we will be able to improve our control and increase the range of physical properties of materials even as we decrease the integration scale. Right now, while we are on the verge of understanding and harnessing physics on the nanoscale, we have a tremendous amount of work to do to learn how to fabricate large-scale materials from nanoscale elements. Although self-assembly and biological processing techniques look promising, they are not yet mature enough for the fabrication of multicomponent systems.

New system-level design methodology for materials will change not only the tools a materials scientist must know and understand, but also the role of the materials scientist in the system-design process. Typically, system designers choose from a toolbox of materials that have already been developed, and the materials scientist pre-designs these materials to improve a single function. Often the materials scientist, who acts independently of the design team, is present only to characterize data or troubleshoot a problem after design. In the new paradigm, however, the materials scientist must be actively involved from the inception of the system design, providing a finely engineered material on the meso-, micro-, or nanoscale to meet the overall system goals. This will require that the materials scientist be completely familiar with system design tools and computational tools, from the system scale to the micro- or nanoscale. In the future, the design of a new car, airplane, or satellite will truly begin on the atomic scale.

References

Burns, L.D., J.B. McCormick, and C.E. Barroni-Bird. 2002. Vehicles of change. Scientific American 287(4): 64–73.

Chen, X., M.A. Dam, K. Ono, A. Mal, H. Shen, S.R. Nutt, K. Sheran, and F. Wudl. 2002. A thermally re-mendable cross-linked polymeric material. Science 295(5560): 1698–1702.

Christodoulou, L., and J.D. Venables. 2003. Multifunctional material systems: the first generation. JOM: The Member Journal of The Minerals, Metals & Materials Society 55(11): 39–45.

dos Santos e Lucato, S.L., J. Wang, P. Maxwell, R.M. McMeeking, and A.G. Evans. 2004. Design and demonstration of a high authority shape morphing structure. International Journal of Solids and Structures 41(13): 3521–3543.

Matic, P. 2003. Overview of Multifunctional Materials. Pp. 61–69 in Smart Structures and Materials 2003, edited by D.C. Lagoudas. Active Materials: Behaviors and Mechanics. Proceedings of SPIE vol. 5053. Bellingham, Wash.: SPIE.

Qidwai, M.A., J.P. Thomas, and P. Matic. 2002. Design and Performance of Composite Multifunctional Structure-Battery Materials. Pp. 1–8 in Proceedings of the American Society for Composites Seventeenth Technical Conference, edited by C.T. Sun and H. Kim. Boca Raton, Fla: CRC Press.

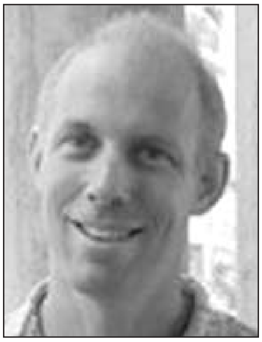
Pei, Q., R. Pelerine, S. Stanford, R. Kornbluh, M. Rosenthal, K. Meijer, and R. Full. 2002. Multifunctional Elastomer Rolls. Materials Research Society Symposium Proceedings 698: 165–170.

Sigmund, O., and S. Torquato. 1999. Design of smart composite materials using topology optimization. Smart Materials and Structures 8(9): 365–269.

Wax, S.G., G.M. Fisher, and R.R. Sands. 2003. The past, present and future of DARPA's investment strategy in smart materials. JOM: The Member Journal of The Minerals, Metals & Materials Society 55(11): 17–23.

From a model-building perspective, the goal is to "make things as simple as possible, but no simpler."

Modeling the Stuff of the Material World: Do We Need All of the Atoms?



Rob Phillips is a professor of mechanical engineering and applied physics at the California Institute of Technology.

Rob Phillips

The advent of computers ushered in a new way of doing science and engineering in which a host of complex problems ranging from weather prediction to the microstructural evolution of multiphase alloys to the DNA/protein interactions that mediate gene regulation could be explored explicitly using computer simulation. Indeed, some say that the physical sciences are now based on a triumvirate of experiment, theory, and simulation, with simulation complementing more traditional techniques for understanding problems involving many interacting degrees of freedom. One class of problems for which simulation is increasingly important is associated with the understanding and control of materials. When we speak of materials, we mean "stuff" as diverse as the materials of which man is made (soft, squishy stuff) and technologies (stuff with desirable properties, such as strength or conductivity) (Amato, 1997).

Clearly, the use of computation to understand and even design complex materials is one of the major challenges that will make it possible to replace the enlightened empiricism that gave rise to the great material ages (e.g., the Iron Age, Bronze Age, and silicon-based Information Age) with rational design. Similar roles are anticipated for simulation in many other fields as well. One of the flagship techniques for examining problems involving complex materials is molecular dynamics in which the microscopic trajectories of each and every atom are followed explicitly. Despite their promise, however, these simulations sometimes generate enormous quantities of information (i.e., terabyte data sets) without necessarily delivering the promised concomitant increase in understanding.

Terabyte data sets engendered by simulations represent a staggering quantity of information. A simple estimate reveals that the entire 10 floors worth of books in the Caltech Millikan Library corresponds roughly to a terabyte of information. More impressively, the genomes of many viruses have an information content that can be stored comfortably on a 256-megabyte memory stick alongside the genomes of even more complex organisms from bacteria to yeast. Indeed, even organisms as complex as humans have genomes that are much smaller than a terabyte. And yet our computers are overflowing with terabyte data sets, and worse yet, discussions of petabyte data sets are becoming routine. For example, a molecular dynamics calculation on a 100,000-atom system run for only 10 nanoseconds, woefully inadequate for accessing most materials processes, already generates a terabyte worth of data. Clearly, there is a mismatch between the quality of information generated in our simulations and the information present in genomes and libraries.

The question of how to build quantitative models of complex systems with many interacting degrees of freedom is not new. Indeed, one of the threads through the history of physics, the development of continuum theories, resulted in two compelling examples of this kind of theory—elasticity and hydrodynamics. These theories share the idea of smearing out the underlying discreteness of matter with continuum field variables. In addition, with both theories, material properties are captured in simple parameters, such as eleastic moduli or viscosity, which reflect the underlying atomic-level interactions without specifically mentioning atoms.

One lesson of these examples is that "multiscale modeling" is neither the exclusive domain of computational model building nor a fundamentally new idea. Indeed, in the deepest sense, the sentiment that animates all efforts at model building, whether analytical or computational, is of finding a minimal but predictive description of the problem of interest.

One feature that makes problems like those described here especially prickly is that they often involve multiple scales in space or time or both. An intriguing response to the unbridled proliferation of simulated data has been a search for streamlined models in which there is variable resolution. Many of the most interesting problems currently being tackled in arenas ranging from molecular biology to atmospheric science are those in which structures or processes at one scale influence the physics at another scale. As a response to these challenges, modelers have begun to figure out how to construct models in which the microscopic physics is maintained only where needed. One benefit of these approaches is that they not only reduce the computational burden associated with simulations of complex systems, but they also provide a framework for figuring out which features of a given problem dictate the way the "stuff" of interest behaves. Several examples of this type of thinking are described below.

"Multiscale modeling" is not a fundamentally new idea.

Before embarking on a discussion of case studies, it is worth discussing the metrics that might be used in deciding whether or not a particular coarse-grained model should be viewed as a success. From the most fundamental point of view, the job of theoretical models is to provide a predictive framework for tying together a range of different phenomena. For example, in the case of elasticity described above, there are vast numbers of seemingly unrelated problems (from flying buttresses to the mechanical response of ion channels) that may be brought under the same intellectual roof through reference to Hooke's law. With elasticity theory, we can *predict* how the cantilever of an atomic-force microscope will deflect when tugging on a protein tethered to a surface. In this sense, elasticity theory has to be viewed as an unqualified success in the coarse-grained modeling of materials and shows just how high the bar has been raised for multiscale models worthy of the name.

A Case Study in Multiscale Modeling: The Quasicontinuum Method

One of the computational responses to problems involving multiple scales is multiresolution models that attempt to capture several scales at the same time. There has been great progress along these lines in recent years from a number of different quarters, and presently we will consider one example, namely the quasicontinuum method that permits the treatment of defects in crystalline solids.[1] The main idea of the quasicontinuum method is to allow for atomic-level detail in regions where interesting physical processes, such as dislocation nucleation, dislocation intersections, and crack propagation are occurring, while exploiting a more coarse-grained description away from the key action. The motivation for the method is based on a recognition that when treating defects in solids there are both long-range elastic interactions between these defects and atomic-scale processes involving the arrangements of individual atoms. What makes these problems so difficult is that both the short-range and long-range effects can serve as equal partners in dictating material response.

The numerical engine that permits a response to problems of this type is finite elements that allow for nonuniform meshes and introduce geometric constraints on atomic positions through the presence of interpolation functions (so-called finite-element shape functions). Just as those of us who learned how to interpolate on logarithm or trigonometric tables remember, the key idea of the finite-element procedure is to characterize the geometric state of the system by keeping track of the positions of a few key atoms that serve as nodes on the finite-element mesh. The positions of all other atoms in the system can be found, if needed, by appealing to simple interpolation.

To simulate material response, geometry is not enough. We not only have to know where the nodes are, but also what forces act on them. To that end, the quasicontinuum method posits that the forces on the nodes can be obtained by appealing to interatomic potentials that describe interactions between individual atoms. Using the interpolated atomic positions, a neighborhood of atoms around each node is constructed, and the energies and forces are then computed using standard atomistic techniques. This is an elegant prescription because it ensures that the material response is strictly determined by the underlying microscopic physics without any *ad hoc* material assumptions. Once the geometric mesh has been constructed and the forces on the nodes computed, the simulation itself can take place by either minimizing the energy with respect to nodal coordinates or by using $F = ma$ physics to compute the trajectories of the system over time.

[1] There are many articles on the quasicontinuum method, but the interested reader is invited to consult *www.qcmethod.com,* which has an extensive list of papers dealing with this method.

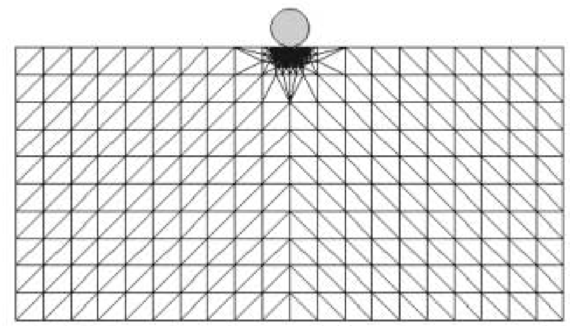


Figure 1 Schematic illustration of a multiresolution mesh used to describe nanoindentation of a crystalline solid. In the region just beneath the indenter, the mesh has full atomic resolution. In the "far fields," the mesh is much coarser.

For a concrete example, consider a crystalline solid subjected to external loading in the form of an indenter like the one shown in Figure 1. The quasicontinuum philosophy is to discretize the system in such a way that there is full atomic resolution where the action is (such as beneath the indenter) and a select, representative subset of atoms that serve as nodes of the finite-element regions where all-atom resolution is surrendered. For the particular case of two-dimensional dynamical nanoindentation considered here, the calculation involves a total of 5,000 nodes as opposed to the 10^7 atoms that would be needed in a full atomistic calculation. This point is driven home in an even more compelling fashion in the case of a fully three-dimensional calculation for which the full atomistic calculation would have implicated in excess of 10^{11} atoms (Knap and Ortiz, 2003).

The main point of this example is to reveal the kind of thinking now being used to address complex problems, such as material deformation. As exemplified by the quasicontinuum method, the underlying microscopic physics of bond stretching and bond breaking is treated explicitly where needed, and only approximately elsewhere.

The Problem of Living Materials

Understanding the workings of living materials presents even more compelling multiscale challenges than those encountered in the traditional materials setting. Indeed, we are now realizing that almost no individual macromolecule in the living world acts alone. Rather, the cell can be viewed as "a collection of protein machines," assemblies of individual macromolecules

(Alberts, 1998). One of the pressing challenges to have arisen from the stunning successes of structural biology is the study of assemblies, such as viruses and the many *"SOMES"* (i.e., macromolecular assemblies that work in concert to maintain the mass and energy budget of the cell), such as nucleosomes, ribosomes, proteosomes, and assemblies that mediate gene expression, such as replisomes, spliceosomes, and so on.

Models of the function of assemblies such as *"SOMES,"* will remain out of reach of traditional atomic-level techniques for the foreseeable future. Consider the process of translation mediated by the ribosome. Even if we very generously assume that a new amino acid is added only once every millisecond, a molecular dynamics simulation of translation would have to be run for 10^{12} time steps for the addition of even a single amino acid to the nascent polypeptide. The number of atoms (including the surrounding water) engaged in this process is well in excess of 100,000, implying a whopping 10^{17} positions corresponding to all of the atoms during the entire molecular dynamics trajectory.

Similar estimates can be made for the workings of many other macromolecular assemblies that mediate the processes of a cell. All of these estimates lead to the same general conclusion—that even as we continue to pursue atomic-level calculations, we must redouble our efforts to understand the workings of *"SOMES"* from a coarse-grained perspective.

So the hunt is on to find methods of modeling processes of biological relevance involving assemblies of diverse molecular actors, such as proteins, lipids, and DNA, without having to pay the price in excessive data of all-atom simulation. One way to guarantee a rich interplay between experiment and models is through the choice of case studies that are well developed from the standpoint of molecular biology and for which we have compelling quantitative data.

One example of great importance is the *lac* operon, which has served as the "hydrogen atom" of gene regulation. This gene regulatory network, which controls the digestion of the sugar lactose in bacteria, has been the cornerstone of the development of our modern picture of gene regulation. An intriguing history of this episode in the history of molecular biology can be found in the books of Judson (1996) and Echols (2001).

The basic idea is that only when a bacterium is deprived of glucose and has a supply of lactose does the bacterium synthesize the enzymes needed to digest lactose. The "decisions" made by the bacterium are mediated by molecules, such as *lac* repressor, a protein that sits on the DNA and prevents the genes responsible for lactose digestion from being expressed. *Lac* repressor binds to several sites in the vicinity of the promoter for the genes responsible for lactose digestion and prohibits expression of those genes while simultaneously creating a loop of DNA between the two repressor binding sites.

From a modeling perspective, a minimal description of this system involves the DNA molecule itself, RNA polymerase, *lac* repressor, and an activator molecule called CAP. The kinds of questions that are of interest from a quantitative modeling perspective include the extent of gene expression as a function of the number of copies of each molecular actor in this drama, as well as the distance between the DNA binding sites for *lac* repressor and other features.

Multiscale methods are taking root in the biological setting.

One recent multiscale attempt to simulate the interaction between DNA and *lac* repressor uses a mixed atomistic/continuum scheme, in which the *lac* repressor and the surrounding complement of water molecules are treated in full atomistic detail while the looped DNA region is treated using elasticity theory (Villa et al., 2004). The advantage of this approach is that it permits the DNA to present an appropriate boundary condition to the *lac* repressor simulation without having to do a full atomistic simulation of both DNA and the protein. Figure 2 shows an example of the simulation box and the elastic rod treatment of DNA. The key point of this example is not to illustrate what can be learned about the *lac* operon using mixed atomistic-continuum methods, but to illustrate how multiscale methods have begun to take root in the biological setting, just as they have in the conventional materials setting.

A second scheme, even more coarse-grained than the multiscale simulations of the *lac* repressor, is a statistical mechanics treatment of molecular decision makers, such as the repressor and its activator counterpart CAP. The relevant point is that all of the atomic-level specificity is captured by simple binding energies that reflect the affinity of these molecules for DNA and for each other.

This statistical mechanics perspective is a natural quantitative counterpart to the cartoons describing gene regulation used in classic texts of molecular biology. As shown in Figure 3, these cartoons depict various states of occupancy of the DNA in the neighborhood of the site where RNA polymerase binds. The statistical mechanics perspective adds the ability to reckon explicitly the statistical weights of each distinct state of occupancy of the DNA. From these statistical weights, a quantitative prediction can be made of the probability that a given gene will be expressed as a function of the number of molecules of each species.

Ultimately, one of the primary ways of judging a model must be by its ability to make predictions about as-yet undone experiments. The outcome of the so-called "thermodynamic models" (Ackers et al., 1982) described above is a predictive framework that characterizes the extent to which genes are expressed as a function of concentrations of the relevant decision-maker molecules (i.e., the transcription factors), the distance between the looping sites on the DNA, etc. Paradoxically, as a result of the great successes of structural biologists in determining the atomic-level structures of important complexes, such as DNA and its binding partners, we are now faced with the challenge of eliminating molecular details in models of their function.

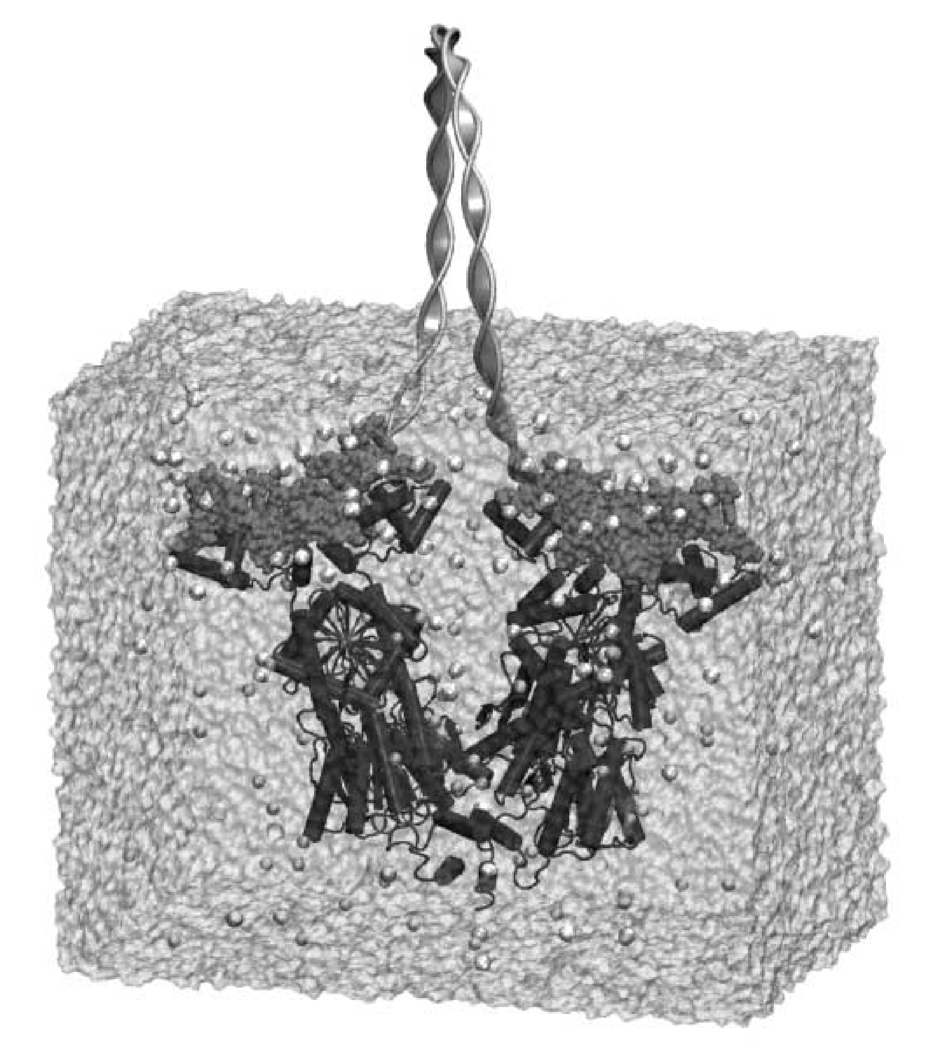


Figure 2 Illustration of a mixed atomistic/continuum description of the interaction of *lac* repressor protein with DNA. The *lac* repressor molecule is shown in dark gray, the part of the DNA treated explicitly is shown in medium gray, and the part of the DNA treated via continuum mechanics is shown as a ribbon. Water molecules surrounding the protein are shown in light gray. Source: Courtesy of Klaus Schulten and Elizabeth Villa.

Summary

The critical question for building models of the material world is the extent to which we can suppress an atom-by-atom description of the function of materials. As emergence of "multi-scale modeling" reveals, even with increasing computational power, a host of important problems remain out of reach of strictly brute-force approaches. In the analysis of the material world, whether of the complex, rigid metallic structures used to construct cities or the soft, squishy materials that make up the organisms that populate them, the key action often takes place at the level of individual atoms—whether of a bond breaking at a crack tip or the active site of an enzyme. Many of the atoms that are part of these processes are interlopers, however, that seem to be little more than passive observers that provide boundary conditions for the atoms actively involved in the process of interest.

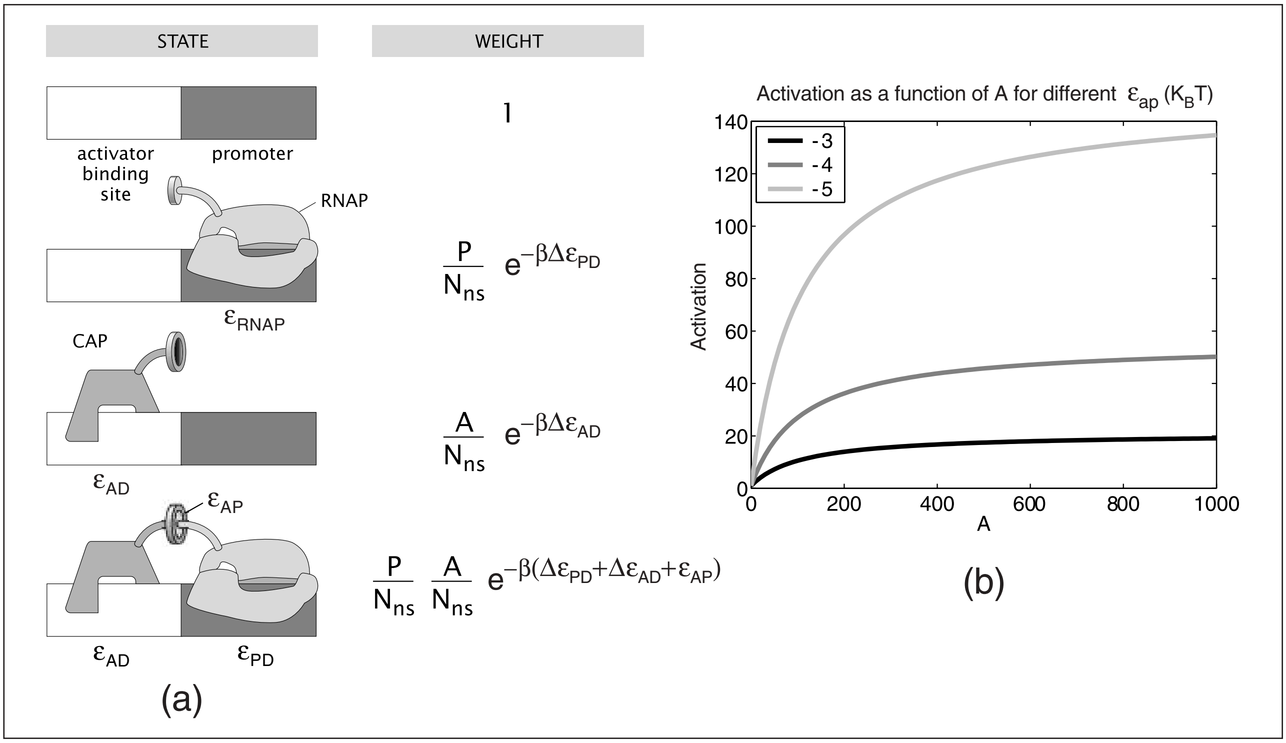


Figure 3 a. Schematic illustration showing the relation between cartoon models of various states of the genetic network and their corresponding weights in the statistical mechanics framework. b. Graph showing activation as a function of the number of activator molecules. The three curves correspond to different strengths for the interaction between the RNA polymerase and the activator.

From a model-building perspective, the goal is to “make things as simple as possible, but no simpler” (i.e., to eliminate as many molecular details as possible). In my opinion, this is one of the key design criteria for multiscale models. Although multiscale computational models are receiving most of the effort and attention right now, I believe the hunt should continue for *analytic* models that can capture the key features of complex materials and lead to the kind of insight that can be discussed at a blackboard.

Acknowledgments

I am grateful to Jané Kondev, Michael Ortiz, Ellad Tadmor, Ron Miller, Vijay Shenoy, Vivek Shenoy, David Rodney, Klaus Schulten, Darren Segall, and Laurent Dupuy for collaboration and conversation. All of them played a key role in the development of my understanding of multiscale modeling. Farid Abraham first illustrated the meaning of a terabyte to me by comparing it to the informational content of a library.

References

Ackers, G.K., A.D. Johnson, and M.A. Shea. 1982. Quantitative model for gene regulation by l phage repressor. Proceedings of the National Academy of Sciences 79(4): 1129–1133.

Alberts, B. 1998. The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell 92(3): 291–294.

Amato, I. 1997. Stuff: The Materials the World Is Made Of. New York: Basic Books.

Echols, H. 2001. Operators and Promoters. Berkeley: University of California Press.

Judson, H.F. 1996. The Eighth Day of Creation. Plainview, N.Y.: Cold Spring Harbor Laboratory Press.

Knap, J., and M. Ortiz. 2003. Effect of indenter-radius size on Au(001) nanoindentation. Physical Review Letter 90(22). Article No. 226102.

Villa, E., A. Balaeff, L. Mahadevan, and K. Schulten. 2004. Multi-scale method for simulating protein-DNA complexes. Multiscale Modeling and Simulation 2(4): 527–553.

Someday we may be able to make a photoreal computer model of anything—no matter what it is made of or how it reflects light.

Capturing and Simulating Physically Accurate Illumination in Computer Graphics

Paul Debevec is executive producer of graphics research at the Institute for Creative Technologies, University of Southern California, Marina del Rey, California.

Paul Debevec

Anyone who has seen a recent summer blockbuster has seen the results of dramatic improvements in the realism of computer-generated graphics. Visual-effects supervisors now report that bringing even the most challenging visions of film directors to the screen is no longer in question; with today's techniques, the only questions are time and cost. Based both on recently developed techniques and techniques that originated in the 1980s, computer-graphics artists can now simulate how light travels within a scene and how light reflects off of and through surfaces.

Radiosity and Global Illumination

One of the most important aspects of computer graphics is simulating the illumination in a scene. Computer-generated images are two-dimensional arrays of computed pixel values, with each pixel coordinate having three numbers indicating the amount of red, green, and blue light coming from the corresponding direction in the scene. Figuring out what these numbers should be for a scene is not trivial, because each pixel's color is based both on how the object at that pixel reflects light and the light that illuminates it. Furthermore, the illumination comes not only directly from the light sources in the scene, but also indirectly from all of the surrounding surfaces in the form of "bounce" light. The complexities of the behavior of light—one reason the world around us appears rich and interesting—make generating "photoreal" images both conceptually and computationally complex.

As a simple example, suppose we stand at the back of a square white room in which the left wall is painted red, the right wall is painted blue, and the light comes from the ceiling (Figure 1). If we take a digital picture of this room and examine its pixel values, we will indeed find that the red wall is red and the blue wall is blue. But when we look closely at the white wall in front of us, we will notice that it isn't perfectly white. Toward the right it becomes bluish, and toward the left it becomes pink. The reason for this is indirect illumination: toward the right, blue light from the blue wall adds to the illumination on the back wall, and toward the left, red light does the same.

Indirect illumination is responsible for more than the somewhat subtle effect of white surfaces picking up the colors of nearby objects—it is often responsible for most, sometimes all, of the illumination on an object or in a scene. If I sit in a white room illuminated by a small skylight in the morning, the indirect light from the patch of sunlight on the wall lights the rest of the room, not the direct light from the sun itself. If light did not bounce between surfaces, the room would be nearly dark!

In early computer graphics, interreflections of light between surfaces in a scene were poorly modeled. Light falling on each surface was computed solely as a function of the light coming directly from light sources, with perhaps a roughly determined amount of "ambient" light added irrespective of the actual colors of light in the scene. The groundbreaking publication showing that indirect illumination could be modeled and computed accurately was presented at SIGGRAPH 84, when Goral et al. of Cornell University described how they had simulated the appearance of the red, white, and blue room example using a technique known as *radiosity*.

Inspired by physics techniques for simulating heat transfer, the Cornell researchers first divided each wall of the box into a 7×7 grid of patches; for each patch, they determined the degree of its visibility to every other patch, noting that patches reflect less light if they are farther apart or facing away from each other. The final light color of each patch could then be written as its inherent surface color times the sum of the light coming from every other patch in the scene. Despite the fact that the illumination arriving at each patch depends on the illumination arriving (and thus leaving) every other patch, the radiosity equation could be solved in a straightforward way as a linear system of equations.

The result that Goral et al. obtained (Figure 1) correctly modeled that the white wall would subtly pick up the red and blue of the neighboring surfaces. Soon after this experiment, when Cornell researchers constructed such a box with real wood and paint, they found that photographs of the box matched their simulations so closely that people could not tell the difference under controlled conditions. The first "photoreal" image had been rendered!

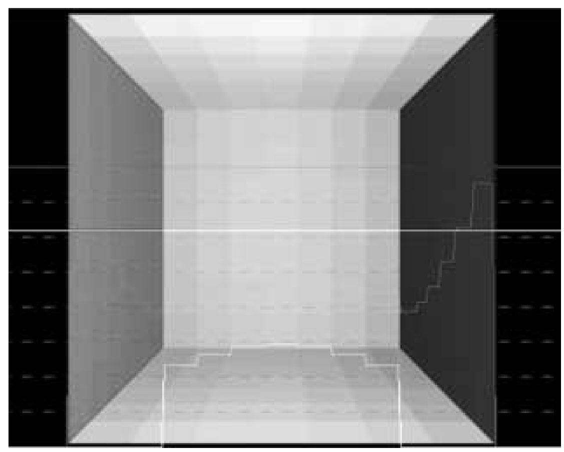


FIGURE 1 A simulation of indirect illumination in a scene, known as the "Cornell box." The white wall at the back picks up color from the red and blue walls at the sides. Source: Goral et al., 1984. (The Cornell box may be seen in color at *http://www.graphics.cornell.edu/online/box/history.html*.)

One limitation of this work was that the time required to solve the linear system increased with the cube of the number of patches in the scene, making the technique difficult to use for complex models (especially in 1984). Another limitation was that all of the surfaces in the radiosity model were assumed to be painted in matte colors, with no shine or gloss.

A subsequent watershed work in the field was presented at SIGGRAPH 86, by Jim Kajiya from Caltech, who published the "The Rendering Equation," which generalized the ideas of light transport to any kind of geometry and any sort of surface reflectance. The titular equation of Kajiya's paper stated in general terms that the light leaving a surface in each direction is a function of the light arriving from all directions upon the surface, convolved by a function that describes how the surface reflects light. The latter function, called the bidirectional reflectance distribution function (BRDF), is constant for diffuse surfaces but varies according to the incoming and outgoing directions of light for surfaces with shine and gloss.

Kajiya described a process for rendering images according to this equation using a randomized numerical technique known as *path tracing*. Like the earlier fundamental technique of *ray tracing* (Whitted, 1980), path tracing generates images by tracing rays from a camera to surfaces in the scene, then tracing rays out from these surfaces to determine the incidental illumination on the surfaces. In path tracing, rays are traced not only in the direction of light sources, but also randomly in all directions to account for indirect light from the rest of the scene. The demonstration Kajiya produced for his paper is shown in Figure 2. This simple scene shows realistic light interactions among both diffuse and glossy surfaces, as well as other complex effects, such as light refracting through translucent objects. Although still computationally intensive, Kajiya's randomized process for estimating solutions to the rendering equation made the problem tractable both conceptually and computationally.

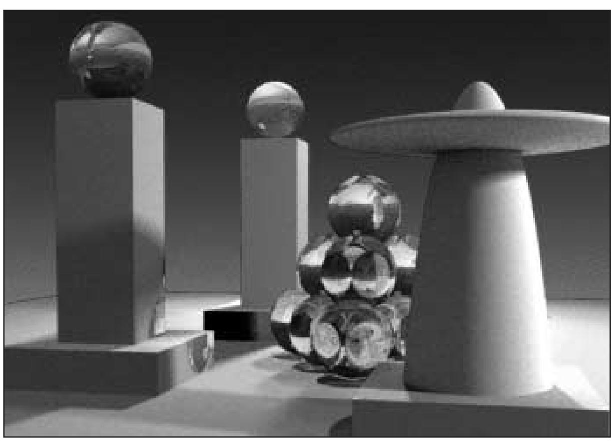


FIGURE 2 A synthetic scene rendered using path tracing. Source: Kajiya, 1986.

Bringing Reality into the Computer

Using the breakthroughs in rendering techniques developed in the mid-1980s, it was no simple endeavor to produce synthetic images with the full richness and realism of images in the real world. Photographs appear "real" because shapes in the real world are typically distinctive and detailed, and surfaces in the real world reflect light in interesting ways, with different characteristics that vary across surfaces. And also, very importantly, light in the real world is interesting because typically there are different colors and intensities of light coming from every direction, which dramatically and subtly shape the appearance of the forms in a scene. Computer-generated scenes, when constructed from simple shapes, textured with ideal plastic and metallic reflectance properties, and illuminated by simple point and area light sources, lack "realism" no matter how accurate or computationally intensive the lighting simulation. As a result, creating photoreal images was still a matter of skilled artistry rather than advanced technology. Digital artists had to adjust the appearance of scenes manually.

Realistic geometry in computer-generated scenes was considerably advanced in the mid-1980s when 3-D digitizing techniques became available for scanning the shapes of real-world objects into computer-graphics programs. The Cyberware 3-D scanner, an important part of this evolution, transforms objects and human faces into 3-D polygon models in a matter of seconds by moving a stripe of laser light across them. An early use of this scanner in a motion picture was in *Star Trek IV* for an abstract time-travel sequence showing a collage of 3-D models of the main characters' heads. 3-D digitization techniques were also used to capture artists' models of extinct creatures to build the impressive digital dinosaurs for *Jurassic Park*.

Digitizing and Rendering with Real-World Illumination

Realism in computer graphics advanced again with techniques that can capture illumination from the real world and use it to create lighting in computer-generated scenes. If we consider a particular place in a scene, the light at that place can be described as the set of all colors and intensities of light coming toward it from every direction. As it turns out, there is a relatively straightforward way to capture this function for a real-world location by taking an image of a mirrored sphere, which reflects light coming from the whole environment toward the camera. Other techniques for capturing omnidirectional images include fisheye lenses, tiled panoramas, and scanning panoramic cameras.

The first and simplest form of lighting from images taken from a mirrored sphere is known as *environment mapping*. In this technique, the image is directly warped and applied to the surface of the synthetic object. The technique using images of a real scene was used independently by Gene Miller and Mike Chou (Miller and Hoffman, 1984) and Williams (1983). Soon after, the technique was used to simulate reflections on the silvery,

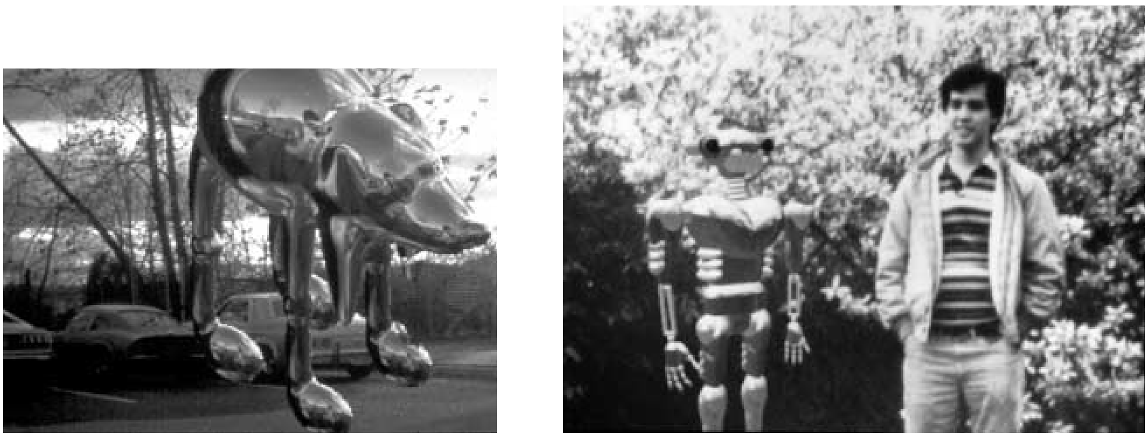


FIGURE 3 Environment-mapped renderings from the early 1980s. (a) An environment-mapped shiny dog. Source: Miller and Hoffman, 1984. (b) An environment-mapped shiny robot. Source: Williams, 1993.

computer-generated spaceship in the 1986 film *Flight of the Navigator* and, most famously, on the metallic T1000 "terminator" character in the 1991 film *Terminator 2*. In all of these examples, the technique not only produced realistic reflections on the computer-graphics object, but also made the object appear to have truly been in the background environment. This was an important advance for realism in visual effects. Computer-graphics objects now appeared to be illuminated by the light of the environment they were in (Figure 3).

Environment mapping produced convincing results for shiny objects, but innovations were necessary to extend the technique to more common computer-graphics models, such as creatures, digital humans, and cityscapes. One limitation of environment mapping is that it cannot reproduce the effects of object surfaces shadowing themselves or of light reflecting between surfaces. The reason for this limitation is that the lighting environment is applied directly to the object surface according to its surface orientation, regardless of the degree of visibility of each surface in the environment. For surface points on the convex hull of an object, correct answers can be obtained. However, for more typical points on an object, appearance depends both on which directions of the environment they are visible to and light received from other points on the object.

A second limitation of the traditional environment mapping process is that a single digital or digitized photograph of an environment rarely captures the full range of light visible in a scene. In a typical scene, directly visible light sources are usually hundreds or thousands of times brighter than indirect illumination from the rest of the scene, and both types of illumination must be captured to represent the lighting accurately. This wide dynamic range typically exceeds the dynamic range of both digital and film cameras, which are designed to capture a range of brightness values of just a few hundred to one. As a result, light sources typically become "clipped" at the saturation point of the image sensor, leaving no record of their true color or intensity. This is not a major problem for shiny metal surfaces, because shiny reflections would become clipped anyway in the final rendered images. However, when lighting more typical surfaces—surfaces that blur the incidental light before reflecting it back toward the camera—the effect of incorrectly capturing the intensity of direct light sources in a scene can be significant.

We developed a technique to capture the full dynamic range of light in a scene, up to and including direct light sources (Debevec and Malik, 1997). Photographs are taken using a series of varying exposure settings on the camera; brightly exposed images record indirect light from the surfaces in the scene, and dimly exposed images record the direct illumination from the light sources without clipping. Using techniques to derive the response curve of the imaging system (i.e., how recorded pixel values correspond to levels of scene brightness), we assemble this series of limited-dynamic-range images into a single high-dynamic image representing the full range of illumination for every point in the scene. Using IEEE floating-point numbers for the pixel values of these high-dynamic-range images (called HDR images or HDRIs), ranges exceeding even one to a million can be captured and stored.

The following year we presented an approach to illuminating synthetic objects with measurements of real-world illumination known as image-based lighting (IBL), which addresses the remaining limitations of environment mapping (Debevec, 1998). The first step in IBL is to map the image onto the inside of a surface, such as an infinite sphere, surrounding the object, rather than mapping the image directly onto the surface of the object. We then use a global illumination system (such as the path-tracing approach described by Kajiya [1986]) to simulate this image of incidental illumination actually lighting the surface of the object. In this way, the global illumination algorithm traces rays from each object point out into the scene to determine what is lighting it. Some of the rays have a free path away from the object and thus strike the environmental lighting surface. In this way, the illumination from each visible part of the environment can be accounted for. Other rays strike other parts of the object, blocking the light it would have received from the environment in that direction. If the system computes additional ray bounces, the color of the object at the occluding surface point is computed in a similar way; otherwise, the algorithm approximates the light arriving from this direction as zero. The algorithm sums up all of the light arriving directly and indirectly from the environment at each surface point and uses this sum as the point's illumination. The elegance of this approach is that it produces all of the effects of the real object's appearance illuminated by the light of the environment, including self-shadowing, and it can be applied to any material, from metal to plastic to glass.

We first demonstrated HDRI and IBL in a short computer animation called *Rendering with Natural Light*, shown at the SIGGRAPH 98 computer-animation festival (Figure 4, left). The film featured a still life of diffuse, shiny, and translucent spheres on a pedestal illuminated by an omnidirectional HDRI of the light in the eucalyptus grove at UC Berkeley. We later used our lighting capture techniques to record the light in St. Peter's Basilica, which allowed us to add virtual tumbling monoliths and gleaming spheres to the basilica's interior for our SIGGRAPH 99 film *Fiat Lux* (Figure 4, right). In *Fiat Lux*, we used the same lighting techniques to compute how much light the new objects would obstruct from hitting the ground; thus, the synthetic objects cast shadows in the same way they would have if they had actually been there.

The techniques of HDRI and IBL, and the techniques and systems derived from them, are now widely used in the visual-effects industry and have provided visual-effects artists with new lighting and compositing tools that give digital actors, airplanes, cars, and creatures the appearance of actually being present during filming, rather than added later via computer graphics. Examples of elements illuminated in this way include the transforming mutants in *X-Men* and *X-Men 2*, virtual cars and stunt actors in *The Matrix Reloaded*, and whole

FIGURE 4 Still frames from the animations *Rendering with Natural Light* (left) and *Fiat Lux* (right) showing computer-generated objects illuminated by and integrated into real-world lighting environments.

cityscapes in *The Time Machine*. In our latest computer animation, we extended the techniques to capture the full range of light of an outdoor illumination environment—from the pre-dawn sky to a direct view of the sun—to illuminate a virtual 3-D model of the Parthenon on the Athenian Acropolis (Figure 5).

FIGURE 5 A virtual image of the Parthenon synthetically illuminated with a lighting environment captured at the USC Institute for Creative Technologies.

Applying Image-Based Lighting to Actors

In my laboratory's most recent work, we have examined the problem of illuminating real objects and people, rather than computer-graphics models, with light captured from real-world environments. To accomplish this we use a series of light stages to measure directly how an object transforms incidental environmental illumination into reflected radiance, a data set we call the *reflectance field* of an object.

The first version of the light stage consisted of a spotlight attached to a two-bar rotation mechanism that rotated the light in a spherical spiral about a person's face in approximately one minute (Debevec et al., 2000). At the same time, one or more digital video cameras recorded the object's appearance under every form of directional illumination. From this set of data, we could render the object under any form of complex illumination by computing linear combinations of the color channels of the acquired images. The illumination could be chosen to be measurements of illumination in the real world (Debevec, 1998) or the illumination present in a virtual environment, allowing the image of a real person to be photorealistically composited into a scene with the correct illumination.

An advantage of this photometric approach for capturing and rendering objects is that the object need not have well defined surfaces or easy-to-model reflectance properties. The object can have arbitrary self-shadowing, interreflection, translucency, and fine geometric detail. This is helpful for modeling and rendering human faces, which exhibit all of these properties, as do many objects we encounter in our everyday lives.

Recently, our group constructed two additional light stages. Light Stage 2 (Figure 6) uses a rotating semicircular arm of strobe lights to illuminate the face from a large number of directions in about eight seconds, much more quickly than Light Stage 1. For this short a period of time, an actor can hold a steady facial expression for the entire capture session. By blending the geometry and reflectance of faces with different facial expressions, we have been able to create novel animated performances that can be realistically rendered from new points of view and under arbitrary illumination (Hawkins et al., in press). Mark Sagar and his colleagues at Sony Pictures Imageworks used related techniques to create the digital stunt actors of Tobey Maguire and Alfred Molina from light stage data sets for the film *Spider-Man 2*.

For Light Stage 3, we built a complete sphere of 156 light sources that can illuminate an actor from all directions simultaneously (Debevec et al., 2002). Each light consists of a collection of red, green, and blue LEDs interfaced to a computer so that any light can be set to any color and intensity. The light stage can be used to reproduce the illumination from a captured lighting environment by using the light stage as a 156-pixel display device for the spherical image of incidental illumination. A person standing inside the sphere then becomes illuminated by a close approximation of the light that was originally captured. When composited over a background image of the environment, it appears nearly as if the person were there. This technique may improve on how green screens and virtual sets are used today. Actors in a studio can be filmed lit as if they were somewhere else, giving visual-effects artists much more control over the realism of the lighting process.

In our latest tests, we use a high-frame camera to capture how an actor appears under several rapidly cycling basis lighting conditions throughout the course of a performance (Debevec et al., 2004). In this way, we can simulate the actor's appearance under a wide variety of different illumination conditions after filming, providing directors and cinematographers with never-before-available control of the actor's lighting during postproduction.

FIGURE 6 Light Stage 2 is designed to illuminate an object or a person from all possible directions in a short period of time, allowing a digital video camera to capture directly the subject's reflectance field. Source: Hawkins et al., in press.

varying color and shininess components (Gardner et al., 2003). We found that by moving a neon tube light source across a relatively flat object and recording the light's reflections using a video camera we could independently estimate the diffuse color and the specular properties of every point on the object. For example, we digitized a 15th-century illuminated manuscript with colored inks and embossed gold lettering (Figure 7a). Using the derived maps for diffuse and specular reflection, we were able to render computer-graphics versions of the manuscript under any sort of lighting environment with realistic glints and reflections from different object surfaces.

A central complexity in digitizing reflectance properties for more general objects is that the way each point on an object's surface responds to light is a complex function of the direction of incidental light and the viewing direction—the surface's four-dimensional BRDF. In fact, the behavior of many materials and objects is even more complicated than this, in that incidental light on other parts of the object may scatter

A Remaining Frontier: Digitizing Reflectance Properties

Significant challenges remain in the capture and simulation of physically accurate illumination in computer graphics. Although techniques for capturing object geometry and lighting are maturing, techniques for capturing object reflectance properties—the way the surfaces of a real-world object respond to light—are still weak. In a recent project, our laboratory presented a relatively simple technique for digitizing surfaces with

a b c

Figure 7 (a) A digital model of a digitized illuminated manuscript lit by a captured lighting environment. Source: Gardner et al., 2003. (b) A synthetic model of a face using a simulation of subsurface scattering. Source: Jensen et al., 2001. (c) A digital model of a translucent alabaster sculpture. Source: Goesele et al., in press.

within the object material, an effect known as subsurface scattering (Hanrahan and Krueger, 1993). Because this effect is a significant component of the appearance of human skin, it has been the subject of interest in the visual-effects industry. New techniques for simulating subsurface scattering effects on computer-generated models (Jensen et al., 2001) have led to more realistic renderings of computer-generated actors (Figure 7b) and creatures, such as the Gollum character in *Lord of the Rings*.

Obtaining models of how real people and objects scatter light in their full generality is a subject of ongoing research. In a recent study, Goesele et al. (in press) used a computer-controlled laser to shine a narrow beam onto every point of a translucent alabaster sculpture (Figure 7c) and recorded images of the resulting light scattering using a specially chosen high-dynamic-range camera. By making the simplifying assumption that any point on the object would respond equally to any incidental and radiant light direction, the dimensionality of the problem was reduced from eight to four dimensions yielding a full characterization of the object's interaction with light under these assumptions. As research in this area continues, we hope to develop the capability of digitizing anything—no matter what it is made of or how it reflects light—so it can become an easily manipulated, photoreal computer model. For this, we will need new acquisition and analysis techniques and continued increases in computing power and memory capacity.

References

Debevec, P. 1998. Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. Pp. 189–198 of the Proceedings of the Association for Computing Machinery Special Interest Group on Computer Graphics and Interactive Techniques 98 International Conference. New York: ACM Press/Addison-Wesley.

Debevec, P., A. Gardner, C. Tchou, and T. Hawkins. 2004. Postproduction re-illumination of live action using time-multiplexed lighting. Technical Report No. ICT TR 05.2004. Los Angeles: Institute for Creative Technologies, University of Southern California.

Debevec, P., T. Hawkins, C. Tchou, H.P. Duiker, W. Sarokin, and M. Sagar. 2000. Acquiring the reflectance field of a human face. Pp. 145–156 of the Proceedings of the Association for Computing Machinery Special Interest Group on Computer Graphics and Interactive Techniques 97 International Conference. New York: ACM Press/Addison-Wesley.

Debevec, P.E., and J. Malik. 1997. Recovering high dynamic range radiance maps from photographs. Pp. 369–378 of the Proceedings of the Association for Computing Machinery Special Interest Group on Computer Graphics and Interactive Techniques 97 International Conference. New York: ACM Press/Addison-Wesley.

Debevec, P., A. Wenger, C. Tchou, A. Gardner, J. Waese, and T. Hawkins. 2002. A lighting reproduction approach to live-action compositing. Association for Computing Machinery Transactions on Graphics 21(3): 547–556.

Gardner, A., C. Tchou, T. Hawkins, and P. Debevec. Linear light source reflectometry. 2003. Association for Computing Machinery Transactions on Graphics 22(3): 749–758.

Goesele, M., H.P.A. Lensch, J. Lang, C. Fuchs, and H-P. Seidel. In press. DISCO—Acquisition of Translucent Objects. Paper from the Association for Computing Machinery Special Interest Group on Computer Graphics and Interactive Techniques 31st International Conference, August 8–14, 2004.

Goral, C.M., K.E. Torrance, D.P. Greenberg, and B. Battaile. 1984. Modelling the interaction of light between diffuse surfaces. Association for Computer Machinery Computer Graphics (3): 213–222.

Hanrahan, P., and W. Krueger. 1993. Reflection from layered surfaces due to subsurface scattering. Pp. 165–174 of the Proceedings of the Association for Computing Machinery Special Interest Group on Computer Graphics and Interactive Techniques 1993 International Conference. New York: ACM Press/Addison-Wesley.

Hawkins, T., A. Wenger, C. Tchou, A. Gardner, F. Goransson, and P. Debevec. In press. Animatable facial reflectance fields. Paper from the 15th Eurographics Symposium on Rendering in Norrkoping, Sweden, June 21–23, 2004.

Jensen, H.W., S.R. Marschner, M. Levoy, and P. Hanrahan. 2001. A practical model for subsurface light transport. Pp. 511–518 of the Proceedings of the Association for Computing Machinery Special Interest Group on Computer Graphics and Interactive Techniques 2001 International Conference. New York: ACM Press/Addison-Wesley.

Kajiya, J.T. 1986. The rendering equation. Association for Computing Machinery Special Interest Group on Computer Graphics and Interactive Techniques Computer Graphics 20(4): 143–150.

Miller, G.S., and C.R. Hoffman. 1984. Illumination and Reflection Maps: Simulated Objects in Simulated and Real Environments. Course Notes for Advanced Computer

Graphics Animation. New York: Association for Computing Machinery Special Interest Group on Computer Graphics and Interactive Techniques Computer Graphics.

Whitted, T. 1980. An improved illumination model for shaded display. Communications of the Association for Computing Machinery 23(6): 343–349.

Williams, L. 1983. Pyramidal parametrics. 1983. Association for Computing Machinery Special Interest Group on Computer Graphics and Interactive Techniques Computer Graphics 17(3): 1–11.

Sound is inherently a spatial perception.

Spatial Audio Reproduction: Toward Individualized Binaural Sound

William G. Gardner is founder and president of Wave Arts Inc. located in Arlington, Massachusetts.

William G. Gardner

The compact disc format, which records audio with 16-bit resolution at a sampling rate of 44.1 kHz, was engineered to reproduce audio with fidelity exceeding the limits of human perception. And it works. However, sound is inherently a spatial perception. We perceive the direction, distance, and size of sound sources, and reproducing the spatial properties of sound accurately remains a challenge. In this paper, I review the technologies for spatial sound reproduction and discuss future directions, focusing on the promise of individualized, binaural technology.

Hearing

People hear with two ears, and the two audio signals received at the eardrums completely define the auditory experience. An amazing feature of the auditory system is that with only two ears sounds can be perceived from all directions, and the listener can even sense the distance and size of sound sources. The perceptual cues for sound localization include the amplitude of the sound at each ear, the arrival time at each ear, and the spectrum of the sound, that is, the relative amplitude of the sound at different frequencies.

The spectrum of a sound is modified by interactions between sound waves and the torso, head, and external ear (pinna). Furthermore, spectral modification depends on the location of the source in a complex way. The auditory system uses spectral modifications as cues to the location of sound, but because the complex shape of the pinna varies significantly among individuals, the cues for sound localization are idiosyncratic. Each individual's auditory system is adapted to the idiosyncratic spectral cues produced by his or her head features.

Binaural Audio

Binaural audio refers specifically to the recording and reproduction of sound at the ears. Binaural recordings can be made by placing miniature microphones in the ear canals of a human subject. Exact reproduction of the recording is possible through properly equalized headphones. If the recording and playback are for the same subject and there are no head movements, the results are stunningly realistic.

Many virtual-reality audio applications attempt to position a sound arbitrarily around a listener wearing headphones. They rely on a stored database of head-related transfer functions (HRTFs), that is, mathematical descriptions of the transformation of sound by the torso, head, and external ear. HRTFs for the left and right ears of a subject specify how sound from a particular direction is transformed en route to the ear drums. A complete description of a subject's head response requires hundreds of HRTF measurements from all directions surrounding the subject. Any sound source can be virtually located by filtering the sound with the HRTFs corresponding to the desired location and presenting the resulting binaural signal to the subject using properly equalized headphones. When this procedure is individualized by using the subject's own HRTFs, the localization performance is equivalent to free-field listening (Wightman and Kistler, 1989a,b).

Figure 1 shows the magnitude spectra for right ear HRTFs measured for three different human subjects with a sound source located on the horizontal plane at 60 degrees right azimuth. Note that the spectra are similar up to 6 kHz; the significant differences in HRTFs at higher frequencies are attributable to variations in pinna shape. Figure 2 shows the magnitude spectra of HRTFs measured from a dummy head microphone for all locations on the horizontal plane. Note how the spectral features change as a function of source direction.

Most research has been focused on localization; subjects presented with an acoustic stimulus are asked to report the apparent direction. Their localization is then compared to free-field listening to assess the quality of reproduction. But this method does not account for many attributes of sound perception, including distance, timbre, and size. In an experimental paradigm developed by Hartmann and Wittenburg (1996), the virtual stimulus is reproduced using open-air headphones that allow free-field listening. Thus, real and virtual stimuli can be compared directly. In these experiments, subjects are presented with a stimulus and asked to decide if it is real or virtual. If a virtual stimulus cannot be distinguished from a real stimulus, then the reproduction error is within the limits of perception. When this experimental paradigm was used to study

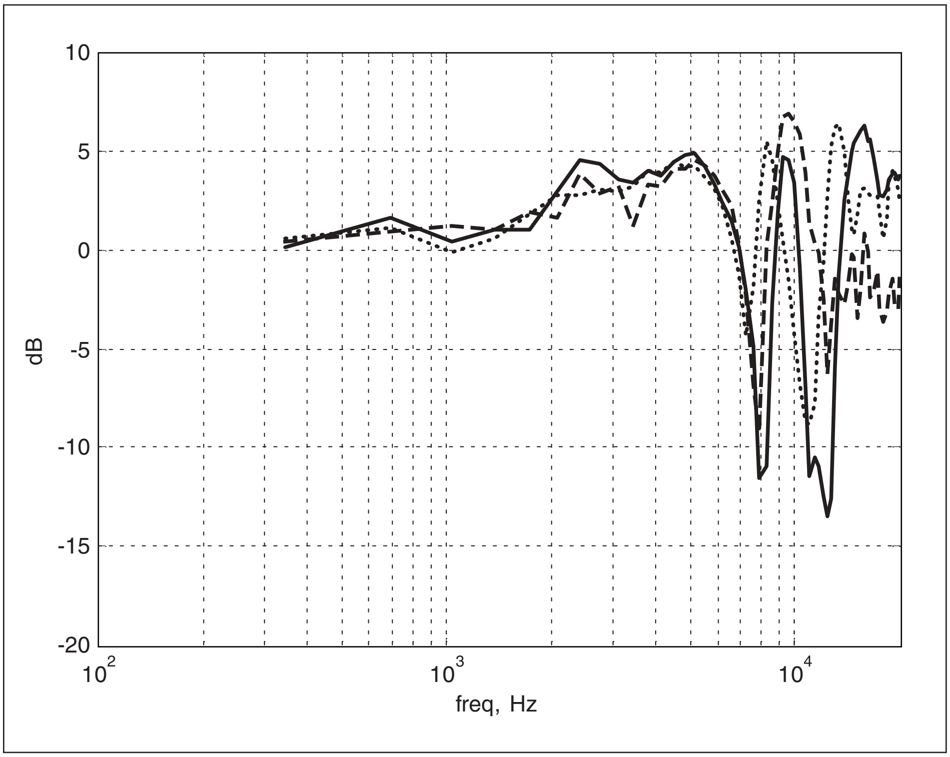


FIGURE 1 Spectrum magnitude for right-ear HRTFs for three different human subjects with a sound source at 60 degrees right azimuth on the horizontal plane. Due to variations in ear shape, the HRTFs differ significantly above 6 kHz.

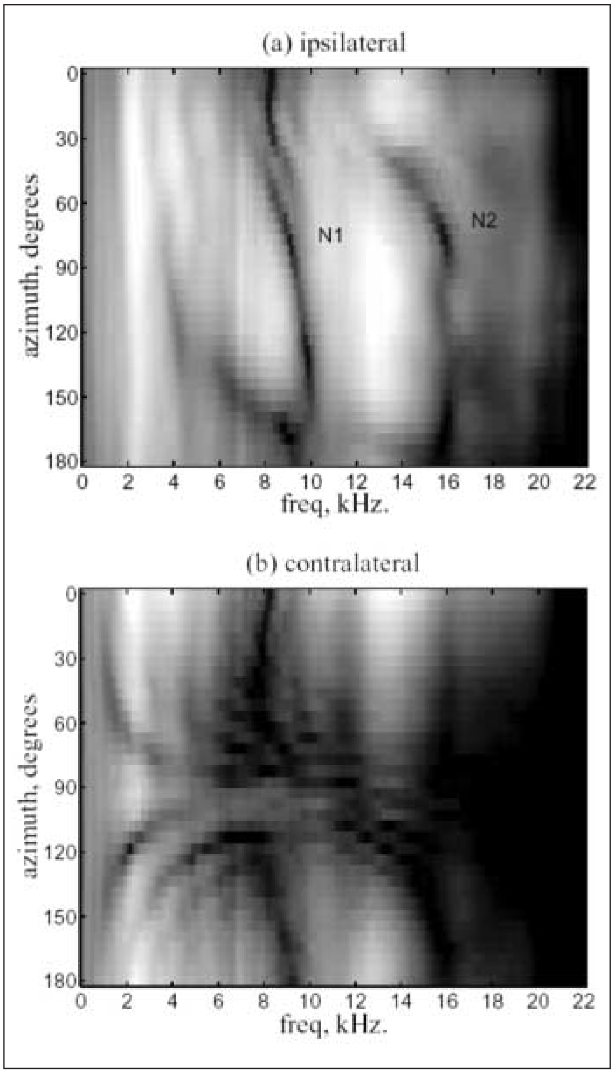


FIGURE 2 Magnitude spectra of KEMAR dummy-head HRTFs as a function of azimuth for a horizontal source (Gardner, 1998). 2a. Ipsilateral (same side) ear. 2b. Contralateral (opposite side) ear. White indicates +10 dB; black indicates –30 dB. Notch features are labeled in (a) according to Lopez-Poveda and Meddis (1996). The figure shows the complex, yet systematic, variation in spectrum as a function of source direction.

the externalization of virtual sound, the results demonstrated that individualized spectral cues are necessary for proper externalization.

The major limitation of binaural techniques is that all listeners are different. Binaural signals recorded for subject A may not sound correct to subject B. Nevertheless, by necessity, binaural systems are seldom individualized. Instead, a reference head, often a model that represents a typical listener or HRTFs known to perform adequately for a range of different listeners, is used to encode binaural signals for all listeners. This is called a "non-individualized" system (Wenzel et al., 1993).

The use of non-individualized HRTFs is limited by a lack of externalization (the sounds are localized in the head or very close to the head), incorrect perception of elevation angle, and front/back reversals. Externalization can be improved somewhat by adding dynamic head tracking and reverberation. Nevertheless, the lack of realistic externalization is often cited as a problem with these systems.

The great challenge in binaural technology is to devise a practical method by which binaural signals can be individualized to a specific listener. There are several possible approaches to meeting this challenge: acoustic measurement, statistical models, calibration procedure, simplified geometrical models, and accurate head models solved using computational acoustics.

With the proper equipment, measuring the HRTFs of a listener is a straightforward procedure, although not practical for commercial applications. Microphones are placed in the ears of the listener, either probe microphones placed somewhere in the ear canal or microphones that block the entrances to the ear canals. Measurement signals are produced from speakers surrounding the listener to measure the impulse response of each source direction to each ear. Because tens or hundreds of directions may be measured, the listener is positioned either in a rotating chair or in a fixed position surrounded by hundreds of speakers. The measurements are often made in an anechoic (echo-free) chamber.

Various statistical methods have been used to analyze databases of HRTF measurements in an effort to tease out some underlying structure in the data. One important study applied principal component analysis (PCA) to a database of HRTFs from 10 listeners at 256 directions (Kistler and Wightman, 1992). Using the log magnitude spectra of the HRTFs as input, the analysis indicated that 90 percent of the variance in the data could be accounted for using only five principal components. The study tested the localization performance using individualized HRTFs approximated by weighted sums of the five principal components. When the listener's own HRTFs were used, the results were nearly identical. The study gathered only directional judgments, and externalization was not considered. The study showed that a five-parameter model is sufficient for synthesizing individualized HRTF spectra, at least in terms of directional localization and for a single direction. Unfortunately, the five parameters must be calculated for each source direction, which means individualized measurements are still necessary.

One can imagine a simple calibration procedure that would involve the listener adjusting knobs to match a parameterized HRTF model with the listener's characteristics. The listener could be given a test stimulus and asked to adjust a knob until some attribute of his perception was maximized. After adjusting several knobs in this manner, the parameter values of the internal model would be optimized for the listener, and the model would be able to generate individualized HRTFs for that individual. Some progress has been made in this area. For example, it has been demonstrated that calibrating HRTFs according to overall head size improves localization performance (Middlebrooks et al., 2000). However, to date, detailed methods of modeling and calibrating the data have not been found.

Many researchers have developed geometrical models for the torso, head, and ears. The head and torso can be modeled using ellipsoids (Algazi and Duda, 2002), and the pinna can be modeled as a set of simple geometrical objects (Lopez-Poveda and Meddis, 1996). For simple geometries, the acoustic-wave equation can be solved to determine head response. For more complicated geometries, head response can be approximated using a multipath model, wherein each reflecting or diffracting object contributes an echo to the response (Brown and Duda, 1997). In theory, head models should be easy to fit to any particular listener by making anthropometric measurements of the listener and plugging these into the model. However, studies have shown that although simplified geometrical models are accurate at low frequencies, they become increasingly inaccurate at higher frequencies. Because of the importance of high-frequency localization cues for proper externalization, elevation perception, and front/back resolution of sound, simplified geometrical models are not suitable for creating individualized HRTFs.

A more promising approach has been to use a three-dimensional laser scan to produce an accurate geometrical representation of a head as a basis for computational acoustic simulation using finite-element modeling (FEM) or boundary-element modeling (BEM) (Kahana et al., 1998, 1999). With this method, HRTFs can be determined computationally with the same accuracy as acoustical measurements, even at high frequencies. Using a 15,000-element model of the head and ear (Figure 3), Kahana demonstrated computation of HRTFs that match acoustical measurements very precisely up to 15 kHz.

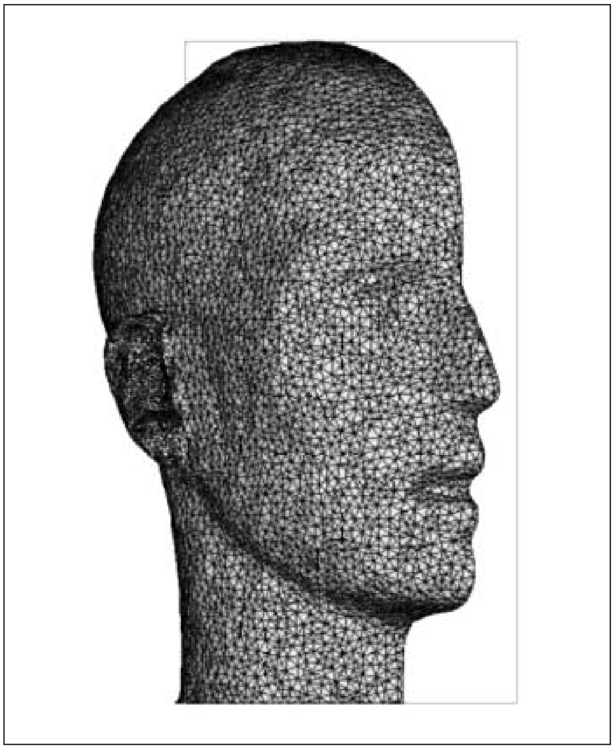


FIGURE 3 Mesh model of one-half of a KEMAR dummy head using 15,000 elements. The model is sufficiently detailed to produce HRTFs that match acoustical measurements, even at high frequencies. Source: Kahana, 1999.

There are, however, a number of practical difficulties with this method. First, scanning the head is complicated by the presence of hair, obscured areas behind the ears, and the obscured internal features of the ear. Second, replicating the interior features of the ear requires making molds and then scanning them separately. Third, after the various scans are spliced together, the number of elements in the model must be pruned to computationally tractable quantities without compromising spatial resolution. Finally, solution of the acoustical equations requires significant computation. For all of these reasons, this approach currently requires more effort and expense than acoustical measurement of HRTFs.

The technique does suggest an alternative approach to determining individualized HRTFs. A deformable head model could be fashioned from finite elements and parameterized with a set of anthropometric measurements. After making head measurements of a particular subject and plugging these into the model, the model head would "morph" into a close approximation of the subject's head. At that point, the computational acoustics procedure could be used to determine individualized HRTFs for the subject. Ideally, the subject's measurements could be determined from images using computer vision techniques. The goal would be a system that could automatically determine individualized HRTFs based on a few digital images of the subject's head and ears. The challenges will be to develop a head model that can morph to fit any head, to obtain a sufficiently accurate ear shape, and to develop ways to estimate the parameters from images of the subject.

Cross-talk–Cancelled Audio

Binaural audio can be delivered to a listener over conventional stereo loudspeakers, but each loudspeaker (unlike headphones) creates significant "cross-talk" to the opposite ear. The cross-talk can be cancelled by preprocessing the speaker signals (called cross-talk cancellers) with the inverse of the 2 × 2 matrix of transfer functions from the speakers to the ears. Cross-talk cancellers use a model of the head to anticipate what cross-talk will occur, then add an out-of-phase cancellation signal to the opposite channel. Thus, the cross-talk is acoustically cancelled at the listener's ears. If the head responses of the listener are known, and if the listener's head remains fixed, an individualized cross-talk cancellation system can be designed that works extremely well.

Non-individualized systems are effective only up to 6 kHz and then only when the listener's position is known (Gardner, 1998). However, despite poor high-frequency performance, cross-talk–cancelled audio is capable of producing stunning, well externalized, virtual sounds to the sides of the listener using frontally placed loudspeakers. As a result of the listener's pinna cues, the sounds are well externalized. The sounds are shifted to the side as a result of the dominance of low-frequency time-delay cues in lateral localization; the cross-talk cancellation works effectively at low frequencies to provide this cue.

Multichannel Audio

The first audio reproduction systems were monophonic, reproducing a single audio signal through one transducer. Stereophonic audio systems, recording and reproducing two independent channels of audio, sound much more realistic. With two loudspeakers, it is possible to position a sound source at either speaker or to position sounds between the speakers by sending a proportion of the sound to each speaker. Stereo has a great advantage over mono because it reproduces a set of locations between the speakers. Also with stereo, uncorrelated signals can be sent to the two ears, which is necessary to achieve a sense of space.

Multichannel audio systems, such as the current 5.1 surround systems, have continued the trend of adding channels around the listener to improve spatial reproduction. 5.1 systems have left, center, and right frontal speakers, left and right surround speakers positioned to the sides of the listener, and a subwoofer to reproduce low frequencies. Because 5.1 systems were designed for cinema sound, the focus is on accurate frontal reproduction so that movie dialogue is spatially aligned with images of the actors speaking. The surround speakers are used for off-screen sounds or uncorrelated ambient effects. The trend in multichannel audio is to add more speaker channels to improve the accuracy of on-screen sounds and provide additional locations for off-screen sounds. As increasing numbers of speakers are added at the perimeter of the listening space, it becomes possible to reconstruct arbitrary sound fields within the space, a technology called wave-field synthesis.

Stereophonic audio systems reproduce a set of locations between the speakers.

Ultrasonic Audio

Ultrasonics can be used to produce highly directional audible sound beams. This technology is based on physical properties of air, particularly that air becomes a nonlinear medium at high sound pressures. Hence, it is possible to transmit two high-intensity ultrasonic tones, say at 100 kHz and 101 kHz, and produce an audible 1 kHz tone as a result of the intermodulation between the two ultrasonic tones. However, the demodulated signal will be significantly distorted, so the audio must be preprocessed to reduce the distortion after demodulation (Pompei, 1999). Although this technology is impressive, it cannot reproduce low-frequency sounds effectively, and it has lower fidelity than standard loudspeakers.

Summary

Binaural audio has the potential to reproduce sound that is indistinguishable from sounds in the real world.

However, the playback must be individualized to each listener's head response. This is currently possible by making acoustical measurements or by making geometrical scans and applying computational acoustic modeling. A practical means of individualizing head responses has yet to be developed.

References

Algazi, V.R., and R.O. Duda. 2002. Approximating the head-related transfer function using simple geometric models of the head and torso. Journal of the Acoustical Society of America 112(5): 2053–2064.

Brown, C.P., and R.O. Duda. 1997. An efficient HRTF model for 3-D sound. Pp. 298–301 in Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. New York: IEEE.

Gardner, W.G. 1998. 3-D Audio Using Loudspeakers. Boston, Mass.: Kluwer Academic Publishers.

Hartmann, W.M., and A. Wittenberg. 1996. On the externalization of sound images. Journal of the Acoustical Society of America 99(6): 3678–3688.

Kahana, Y., P.A. Nelson, and M. Petyt. 1998. Boundary element simulation of HRTFs and sound fields produced by virtual acoustic imaging. Proceedings of the Audio Engineering Society's 105th Convention. Preprint 4817, unpaginated.

Kahana, Y., P.A. Nelson, M. Petyt, and S. Choi. 1999. Numerical modeling of the transfer functions of a dummy-head and of the external ear. Pp. 330–334 in Proceedings of the Audio Engineering Society's 16th International Conference. New York: Audio Engineering Society.

Kistler, D.J., and F.L. Wightman. 1992. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. Journal of the Acoustical Society of America 91(3): 1637–1647.

Lopez-Poveda, E.A., and R. Meddis. 1996. A physical model of sound diffraction and reflections in the human concha. Journal of the Acoustical Society of America 100(5): 3248–3259.

Middlebrooks, J.C., E.A. Macpherson, and Z.A. Onsan. 2000. Psychophysical customization of directional transfer functions for virtual sound localization. Journal of the Acoustical Society of America 108(6): 3088–3091.

Pompei, F.J. 1999. The use of airborne ultrasonics for generating audible sound beams. Journal of the Audio Engineering Society 47(9): 726–731.

Wenzel, E.M., M. Arruda, D.J. Kistler, and F.L. Wightman. 1993. Localization using nonindividualized head-related transfer functions. Journal of the Acoustical Society of America 94(1): 111–123.

Wightman, F.L., and D.J. Kistler. 1989a. Headphone simulation of free-field listening I: stimulus synthesis. Journal of the Acoustical Society of America 85(2): 858–867.

Wightman, F.L., and D.J. Kistler. 1989b. Headphone simulation of free-field listening II: psychophysical validation. Journal of the Acoustical Society of America 85(2): 868–878.

NAE News and Notes

NAE Newsmakers

Frances E. Allen, IBM Fellow Emerita, IBM Thomas J. Watson Research Center, was presented the **Anita Borg Award for Technical Leadership** by the Anita Borg Institute for Women and Technology at a banquet on October 7, 2004.

W. Gene Corley, senior vice president, Construction Technology Laboratories Inc., received the **ASCE President's Award** for distinguished leadership and service to the United States. The award, which was established to honor George Washington, was presented to Dr. Corley in recognition of his service as the ASCE representative on the building-performance assessment teams organized to investigate how well the Alfred P. Murrah Federal Building stood up to the 1995 bombing and the performance of the World Trade Center in response to the terrorist attacks in 2001.

Edith M. Flanigen, retired fellow, UOP LLC, won the **Lemelson-MIT Lifetime Achievement Award** in April 2004 for four decades of pioneering work in chemistry and materials science that has resulted in more than 100 patents and revolutionized molecular sieve materials (porous crystals that can separate molecules on the basis of size). Her inventions have helped make the petroleum-refinement process cleaner, safer, and more efficient. As part of the award, Dr. Flanigen will participate in a congressional briefing about the future of scientific exploration and invention in the United States. The Lemelson-MIT Lifetime Achievement Award includes a cash prize of $100,000.

Nick Holonyak Jr., John Bardeen Chair Professor of Electrical and Computer Engineering and Physics, University of Illinois, won the prestigious **Lemelson-MIT Prize** for 2004 for his invention of the first practical LED (light-emitting diode), the first visible-spectrum semiconductor laser, and the household dimmer switch. His discoveries have had a major impact on the lighting industry, global communications, and consumer products. The Lemelson Prize, which includes a cash award of $500,000, is the world's largest single cash prize for invention and is awarded annually to an individual who demonstrates remarkable inventiveness and creativity and who inspires others. Dr. Holonyak received the award on April 23 in a ceremony at the National Academy of Sciences.

Louis J. Lanzerotti, consultant, Bell Laboratories, Lucent Technologies, and distinguished research professor, New Jersey Institute of Technology, received the Committee on Space Research (COSPAR) **William Nordberg Medal** in recognition of his "exceptional contributions to the practical application of fundamental research in space physics to space and ground-based communications." The medal was presented to Dr. Lanzerotti at the 35th COSPAR Scientific Assembly in Paris, France, on July 19, 2004.

Gordon E. Moore, chairman emeritus, Intel Corporation, received the Society of Chemical Industry (SCI) America Section **2004 Perkin Medal**, on September 14, 2004. The Perkin Medal is the highest honor given for outstanding applied chemistry in the United States. Dr. Moore, one of the pioneers of the semiconductor industry, was honored for helping to create the world's first integrated circuit, the world's first microprocessor, and several other cornerstone products. On the same day, SCI honored Dr. Moore, widely known for "Moore's law," by announcing a new medal named for him. The **SCI Gordon E. Moore Medal** will be awarded for innovation by an industrial scientist under 45 years of age.

Two NAE members have been elected to the National Academies Institute of Medicine (IOM). **Frances H. Arnold**, IBM Fellow Emerita, IBM Thomas J. Watson Research Center, and **Edward H. Kaplan**, William N. and Marie A. Beach Professor of Management Sciences and professor of public health, Yale School of Management, were elected to the IOM on October 18, 2004.

NAE President Honored by ASME

President **Wm. A. Wulf** was presented with the American Society of Mechanical Engineering's **Ralph Coats Roe Medal** on November 17. The medal honors individuals who have raised public understanding and appreciation of engineers and engineering. Dr. Wulf was recognized for his many achievements as NAE president, particularly the initiative to survey activities designed to raise public awareness of engineering. The medal was presented during the 2004 International Mechanical Engineering Congress and RD&D Expo in Anaheim, California.

2004 Annual Meeting

NAE members, foreign associates, and guests gathered in Washington, D.C., this October for the 2004 NAE Annual Meeting. The meeting began on Saturday, October 2, with the orientation of new members. This was followed by the NAE Council Dinner honoring the 75 new members and 11 foreign associates.

The public session on Sunday, October 3, was opened by NAE Chair **Craig R. Barrett**, who called for more government funding for R&D to keep the United States competitive in the global economy (see p. 45). President **Wm. A. Wulf** then addressed the group on the question of outsourcing and other problems confronting engineers in the twenty-first century (see p. 47). The induction of the NAE Class of 2004 followed President Wulf's address.

The program continued with the presentation of the 2004 **Founders Award** to **Eli Ruckenstein**, Distinguished Professor of Chemical Engineering at the State University of New York, Buffalo, who was recognized "for leadership in modernizing research and development in key areas of chemical engineering." Dr. Ruckenstein was not able to attend the meeting, so his award was accepted by NAE member **Howard Brenner**, who also read his acceptance speech (see p. 51).

The 2004 **Arthur M. Bueche Award** was presented to **John B. Slaughter**, president and CEO of the National Action Council for Minorities in Engineering, "for support of engineering research and education within the National Science Foundation, many contributions to the development of science and technology policy, and lifelong dedication to increasing diversity in the disciplines of science and engineering." Dr. Slaughter's remarks are reprinted on p. 53.

Following the presentation of the awards, the recipient of the 2004 **Gordon Prize**, **Frank S. Barnes**, lectured on "Opportunities and Challenges for Engineering Education as Seen from an Interdisciplinary Telecommunications Master of Science" (p. 55). The Gordon Prize, which is awarded for new modalities and experiments in education that develop effective engineering leaders, was presented to Dr. Barnes, Distinguished Professor in the Department of Electrical and Computer Engineering, University of Colorado at Boulder, at an awards ceremony in Washington, D.C., last February.

Dr. Melody Moore was presented with the Gilbreth Lecture honor for "Frontiers of Human-Computer Interaction: Direct Brain Interfaces." (Left to right) Craig Barrett, Melody M. Moore, Wm. A. Wulf.

Dr. Alan Russell delivered the Gilbreth Lecture on "Using Biotechnology to Detect and Counteract Chemical Weapons."

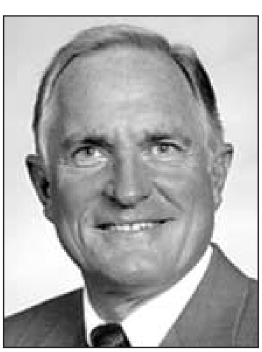
After a break, the two recipients of the Lillian M. Gilbreth Lectureships, which recognize outstanding young engineers, presented talks. The first speaker, Dr. Melody Moore, assistant professor in the Computer Information Systems Department at Georgia State University, spoke on frontiers of human-computer interaction. The second speaker, Dr. Alan J. Russell, director of the McGowan Institute for Regenerative Medicine at the University of Pittsburgh, spoke about using biotechnology to detect and counteract chemical weapons.

At the Annual Business Session on Monday morning, members were asked to vote on proposed amendments to the NAE Articles of Organization and the NAE bylaws. Members who were not present had been given an opportunity to vote by mail. The amendments to both documents were approved. (The revised documents can be found on the NAE website, *http://www.nae.edu.*) The business session was followed by a symposium entitled "A Century of Innovation in Manufacturing." Speakers included Pamela A. Drew, vice president and deputy general manager, Airborne Intelligence, Surveillance and Reconnaissance, Integrated Defense Systems, The Boeing Company; **James Padilla**, chief operating officer, Ford Motor Company; Kenneth Flamm, Dean Rusk Chair in International Affairs, Lyndon B. Johnson School of Public Affairs, University of Texas at Austin; Alfonso Velosa, associate director, Gartner, Inc.; Delcie R. Durham, program director, Division of Design, Manufacture and Industrial Innovation, National Science Foundation; Lawrence J. Rhoades, president and chief executive officer, Extrude Hone Corporation; and Sidney Perkowitz, Charles Howard Candler Professor, Department of Physics, Emory University. On Monday afternoon, members and foreign associates participated in NAE section meetings at the Keck Center.

The final event of the meeting was the annual reception and dinner dance, held for the first time at the Andrew W. Mellon Auditorium. Entertainment was provided by the Gross National Product and music was provided by the Radio King Orchestra.

The next annual meeting is scheduled for October 9 and 10, 2005.

Chairman's Remarks



Craig R. Barrett

These remarks were delivered October 3, 2004, at the NAE Annual Meeting.

It's a pleasure to be here today to celebrate the success of engineering—a celebration on two fronts. First, we're celebrating the success of the new inductees to the National Academy of Engineering. Second, in a larger sense, we're celebrating engineering as a whole and the contributions of engineering to the successes of the United States—the success of the U.S. economy and the continued increases in capability around the world.

This celebration comes at an interesting time. As we look around the world today, a significant change is taking place. In fact, it's a once-in-human-history change, something that has never happened before. In the last 10 years, roughly half of the world's population has joined the free economic system. India, China, Russia, and the Eastern European countries have a combined population of approximately three billion people, roughly half the world's population. Almost overnight in economic terms, half of the world's population has suddenly become available to purchase goods and services from the rest of the world and contribute their labors and expertise to the global economic system.

Bringing three billion people into the free economic system is bound to cause dislocations, competition, and fundamental changes in the way the system operates. We've seen a little bit of that in the national debate here in the United States, which has centered on phrases such as offshoring; outsourcing, the movement of jobs out of the United States; and competition for jobs. If any of you watch Lou Dobbs on *CNN Tonight*, you've seen him call me, and others, the Benedict Arnold CEOs of the world for moving jobs offshore. Lou often forgets to mention that my company does 70 percent of its business outside the United States. I have yet to figure out how we're supposed to do 70 percent of our business overseas but not have any foreign employees.

The interesting thing about this change, though, is that the countries that have just come into the free economic system come with rich educational heritages. Russia, China, India, and the Eastern European countries have excellent universities, excellent university graduates, and knowledge-based workers who can do just about any job anyone else in the world can do. Even if you discount 90 percent of those three billion people as agrarian and uneducated, that still leaves 300 million who are educated. That's equivalent to the U.S. workforce and the Western European workforce. There's bound to be competition.

You can look at another aspect of the situation. Consider the impact of Taiwan with a population of fewer than 25 million or the impact of South Korea with a population of fewer than 50 million. Look at their impact on the world's economic system, and then make a rough extrapolation, as any good engineer would do, of the impact of three billion. We should expect some strong competition going forward for engineering-based jobs and knowledge-based jobs.

The potential response of the United States to these new circumstances could be shaped by engineering fundamentals. First, any competition for economic growth, any competition for standard of living, typically goes with the educational level of the workforce. The average education of a population is a good predictor of the standard of living of the country. If you want the highest standard of living in the world, as the United States, Western Europe, and Japan currently enjoy, you have to assume that you need an educated workforce that is second to none.

There's been a lot of discussion on education in the United States, specifically K through 12 education, and, frankly, if you captured all of the hot air that's been expended on this topic and funneled it down turbines, we'd probably be energy independent. We also have discussions on No Child Left Behind. Senator Glenn ran a commission a few years ago that issued a report, *Before It's Too Late*, which basically focused on how to interest children in K through 12 in mathematics and science and on improving the teaching of mathematics and science, and hence engineering, in the United States. But I think there is probably enough dialogue on education already.

The second engineering metric you can apply to economic competitiveness is research and development (R&D). If you do a statistical study on the standard of living relative to the percentage of gross domestic product invested in R&D, you find there's a very strong correlation. The higher the investment in R&D, the greater the generation of new ideas and new concepts for goods, services, and businesses. Investment in R&D has driven the established economies around the world today.

But if you look at U.S. investment, specifically in the physical sciences and engineering, particularly in basic R&D, which is the domain of the U.S. government (primarily the National Science Foundation, U.S. Department of Defense, and U.S. Department of Energy) and track that investment over the past 20 or 25 years, you find a surprising result. Investment has been basically flat in terms of absolute dollars of today. It's been at about $5 billion for the last 25 years with very little variation.

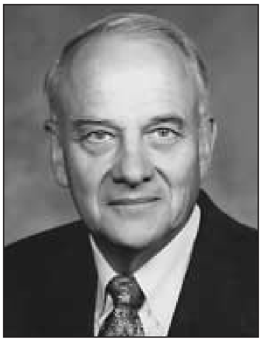
I am consistently amazed by that statistic. I can put it in terms of two comparisons for you. First, I work for Intel Corporation. Our R&D budget, for just one company—mostly for product-oriented applied R&D rather than basic research—is approximately $5 billion a year, which is exactly the total investment by the U.S. government in basic R&D.

The other comparative statistic is the amount of money the U.S. government spends on agricultural subsidies. Last time I looked, agriculture was the industry of the nineteenth century, whereas engineering and technology represent the industries of the twenty-first century. The amount the United States government spends on agricultural subsidies each year, direct agricultural subsidies, not food stamps, is roughly $25 billion. I am continually perplexed that we spend five times as much supporting industries of the nineteenth century as we do investing in the industries of the twenty-first century.

So I leave you with one thought this morning as we celebrate engineering accomplishments and the ideals of the National Academy of Engineering and the new inductees. If we want to be successful in the future, we have to make an investment. There is no way around that. Competition has increased in the world—to a level that has never been seen before. Many other countries around the world are investing heavily in knowledge-based industries and universities. The basic research laboratory in the United States is not an industrial research laboratory; it is the collection of tier-one and tier-two university research laboratories.

I encourage all of us to raise our voices on one topic. I desperately hope that when the two presidential candidates meet to talk about domestic issues at least one of them mentions investment in R&D as a way to ensure our future economic health and economic well-being. I'd love us all to raise our voices and ask the U.S. government, Congress, the President, and the presidential candidates about investing in the future. In fact, there is no future unless we invest today. I hope you will join me and Bill Wulf and the rest of the Academy in raising our voices in that direction.

President's Address



Wm. A. Wulf

These remarks were delivered October 3, 2004, at the NAE Annual Meeting.

It is an immense honor to welcome once again our new members and foreign associates. This talk each year gives me an opportunity to reflect on problems and opportunities for engineering. In reviewing the possibilities this year, I realized that there is an elephant in the living room, and I had better not ignore it. I am referring to the outsourcing (or offshoring) of engineering jobs to India and China and other places. The outsourcing issue is connected to other issues that, taken together, form a "mosaic," an image of troubling proportions for our society, not just for engineering. So before I turn to outsourcing per se, let me put it in context.

First, everyone who has looked into the problem agrees that the data on outsourcing are poor! Last year, at an NAE "summit" on the engineering workforce, an incredible range of opinions were expressed—from shortage to oversupply—and each opinion was supported by the selective use of data. A Rand report released this June makes the same point, namely, that the data are inadequate. As a result, the situation is now rife with hype, the sort of situation the media love to amplify. Let's try to avoid that.

Second, the physical arrangement today, that is, me at the podium and you in the audience, is not conducive to a dialogue. But dialogue is what we need, specifically a dialogue about what we are going to do (or *not* do). It's too easy to fob this problem off onto politicians. As the most respected assembly of engineers in the United States, NAE should be front and center on this issue, which is critical to our profession. Today, I am setting a context for that dialogue, a "discussion database" on our website (*www.nae.edu*), with a link at the top level. I encourage you to engage in the discussion.

Third, I refer you to the 20 Greatest Achievements Project. For the benefit of new members and guests, this is a project we undertook in 1999 to come up with a list of the 20 greatest engineering achievements of the twentieth century; "greatest" was defined in terms of impact on our quality of life. The full list is available on our website, and a "coffee table" book is for sale by the National Academies Press. The important point for this discussion is how profoundly engineering has affected society in the last 100 years—arguably more than any other single thing. Remember, in 1900, few people had electricity or telephones or automobiles. There were no airplanes; the average life span in the United States was 46, whereas it is now 77 (it has been estimated that clean water added 20 of those 31 years); 50 percent of the U.S. population lived on farms, whereas now only 2 percent do because of the increase in productivity brought about by mechanization; there were no antibiotics, radios, TVs, Internet, and so on.

Fourth, I am not an economist—whether that's a good thing or not we'll have to see—and this is clearly an economic issue. So I won't pretend that I can see all of the policy levers that may be applicable here.

In that context, I will tell you some of the things I *know* and then move on to some of the things I *believe* but can't prove. I will then suggest some things we ought to be doing and some linkages to other issues that form the "mosaic" I referred to earlier.

Some Things I *Know*

First, I know that the plural of anecdote is not data! The popular press likes to tell stories about individuals: "John Jones says he spent four years being educated for a three-year engineering career, but now he is leaving engineering . . . ." That's not data. The story is meant to suggest that the problem is rampant and to make you feel personally threatened. But, in fact, it provides no basis for knowing anything.

Second, I know that conflicting definitions of "engineer" by the National Science Foundation (NSF) and the Bureau of Labor Statistics confuse the situation. In my view, the term "computer science" further confuses things (even though it is *called* a science, I believe it is [mostly] an engineering discipline). We need to be very careful to use consistent, comparable data.

Third, I know much of the hard data we have is reported two years late. That really matters in the current context because of the "dot com" boom/bust.

Fourth, I know that the number of new engineers per year in the United States declined from about 80,000 in 1985 to about 65,000 in the mid-1990s; it is back up to about 75,000 in the latest data. To put this in context, the number is more than 200,000 per year in China and India, and about one million per year worldwide. Thus, new U.S. engineers account for only about 7.5 percent of the world total.

Fifth, I know we were importing engineers on H1B visas, about 20,000 to 25,000 per year at the peak. I have seen this figure reported as 195,000 in some places, but that is the total for all professions, not just engineering. This is the sort of hype that annoys me.

Sixth, I know we are creating engineering jobs in India and China; one need only drive down a road in Bangalore to see that. The question is to what extent these jobs are replacing jobs in the United States. Surely they are to some degree, but I have seen no reliable data on numbers. There are reports of good and bad experiences with outsourcing, and there are also cases of "insourcing," in the pharmaceutical industry, for example.

Seventh, starting salaries for new bachelor-level engineers are not moving up significantly (which would imply a shortage) nor down significantly (which would imply an oversupply). They *are* 1.5 to 2 times the starting salaries for new B.A.s, and the average engineering salary is comparable to a new lawyer's starting salary. This suggests to me that employers see value in U.S. engineers.

Eighth, engineering unemployment is up from 0.5 percent to around 6 percent—at least in some fields. That is comparable to unemployment in the rest of the workforce—which is unusual and worrisome because historically engineering unemployment has been lower than in the workforce as a whole! Does that tell us something has changed structurally? We must be careful in interpreting this because of the crazy economic boom/bust we've experienced. It's also worth noting that 6 percent unemployment used to be considered full employment and that less than 6 percent was considered to cause inflation. Is the situation today an anomaly, or was the situation in the 1990s an anomaly? It's hard to tell!

Finally, I know we need to keep things in perspective. Six percent unemployment among engineers is about 120,000 individuals. That should be compared to the 500,000—half a million—Americans who lose their jobs each *week*. In a healthy economy, jobs "churn," and that's a *good* thing because that's how things improve, how we lose low-end jobs and acquire high-end jobs.

Here's an example. In 1970, there were 421,000 switchboard operators supporting 9.8 billion phone calls. Today there are 78,000 operators (less than one-fifth as many) supporting 98 billion calls (10 times as many). Should we have protected those jobs? I don't think so. Although we can have empathy for those who lost their jobs, and we perhaps have a societal responsibility to retrain them, consumers are the big winners—a 40-fold decrease in the cost of telephony and a host of new services.

Some Things I *Believe*

First, I believe that engineers are *not* telephone operators! U.S. prosperity, security, and health depend on technology created by engineers, and that dependence is evident in the 20 great achievements I mentioned earlier. I also believe that this dependence will increase. And this dependence on engineering fundamentally changes the nature of the discussion, from empathy for individuals to concern for our economy, our security, our health, the environment, etc.

Second, I strongly believe two things that are superficially incompatible. One is that human capital is precious—when we lose an engineer we diminish that stock. The other is that leaving engineering is not, in itself, a failure of either the individual or the profession. An engineering education is great preparation for many careers. In fact, I wish we had more engineers in other professions; from where I sit these days, I especially wish we had more engineers in politics!

Third, I believe that the aggregate data and discipline-specific data are quite different. That matters because engineering is changing, and we need to be able to disaggregate the trends.

Fourth, I don't believe in a protectionist approach to solving the problem of access to engineering talent. Protectionism might be OK in the short run, but it's not in the long-term interest of the country. It would not have been good for the country to protect those telephone operators' jobs I mentioned earlier, for example. We need to think of positive, constructive alternatives to ensure the nation's access to the engineering talent it needs.

Finally, I believe that, for the most part, managers make rational decisions. If they can get comparable talent at one-fifth the cost in India, and *if* the start-up cost is small, and *if* the cost stays small, and *if* the productivity per unit cost is high enough, and *if* they can manage from 10,000 miles and 12 times zones away, then they *will* outsource—and they should! The problem is the *nation's* access to engineering talent, and the individual manager in an individual company is not responsible for solving that problem. But those are all big "ifs," and there is some evidence that each has been wrong in specific cases.

What bothers me most is the long-term protectionist approach advocated by some, even some professional societies. Some form of protection might or might not be appropriate in the short term, but it is not a long-term solution in my view. The world is changing—or, perhaps more accurately, has changed. The familiar and comfortable world of mass production manufacturing by predominantly nationally owned and operated companies is gone. In the long term, we have to prepare our engineers to operate in this new environment in a way that the "ifs" I listed are not true. We must also have empathy for those who are displaced from their jobs and help them prepare for new ones.

The Mosaic

The proposed protectionist approach as a long-term solution to the problem of outsourcing is part of the mosaic I mentioned earlier, a mosaic that presents an image of increasingly short-term thinking and a lack of balance. Let me suggest some other pieces of the mosaic that relate to engineering.

First, the absolute number of engineering graduates has declined. As I said earlier, we graduated about 85,000 in the mid-1980s; the number dropped to about 65,000 in the mid-1990s, but has recently risen to about 75,000. One must remember, however, that the total number of undergraduates has increased in the intervening 20 years. So, although the absolute number is up somewhat, it is still a decreasing fraction of the total number of undergraduates.

Second, the percentage of undergraduates in engineering (between 4 and 5 percent) is the second lowest in the developed world. By way of contrast, the percentage is 12 percent in the United Kingdom and more than 40 percent in China.

Third, as NAE members already know from my newsletter, I am deeply concerned about the increasing difficulty of students and senior scientists and engineers getting visas to enter the United States. Very briefly, in the wake of 9/11, the U.S. government instituted new policies for issuing visas, which was both inevitable and appropriate in some ways. However, I believe the way this has been done has been counterproductive.

In the past, the best and brightest in the world came to the United States to study, and many remained here to contribute their talents to improving our way of life. In the past, senior and respected scholars came to visit U.S. universities, to lecture, and to share their knowledge. In the past, major international conferences were hosted in the United States, giving our scientists and engineers privileged access to the most recent developments around the world. In the past, the image of the United States was of a free and open society—a land of opportunity.

We have not completely eliminated those "in the past" items, but we are coming perilously close! In an attempt to make ourselves safer, we have both prevented the world's best and brightest from attending our universities and alienated them and senior scientists and engineers in the process.

Fourth, after three decades of steady progress in engineering for women and underrepresented minorities, in the early 1990s, the percentage of women and underrepresented minorities declined. It then rebounded a bit . . . but has essentially been flat for more than a decade now.

Fifth, for most of our history, higher education was considered a public good, that is, it was considered good for the country to increase the number of educated citizens. Before becoming president of NAE, I was a professor at the University of Virginia, which was founded by Thomas Jefferson, who believed that we could not have a democracy without an educated citizenry. In the middle of the nineteenth century, our national leaders created the land grant colleges, because they felt we could not have prosperity without an educated workforce. Until recently, states have supported state universities, provided scholarships, and kept tuitions low to make higher education accessible to all.

But something has changed! We no longer seem to consider higher education a public good, but a private one. State support for universities is dropping like a stone, tuitions are rising, and student loans have replaced scholarships in many states. All of these trends reflect a largely unstated view that shifting costs to students is OK because *only* students benefit from higher education. It's the "only" in that last sentence that worries me.

Sixth, funding at NIH has doubled to almost $30 billion; at the same time, funding at NSF is $4.25 billion. The last yearly *increase* at NIH was bigger than the entire NSF budget. As I grow older, I do not begrudge NIH a nickel of that. But I think the imbalance reflects a lack of appreciation of how much our society—even medical technology—depends on the contributions of the physical sciences and engineering. I also think the growing imbalance reflects a short-term focus on the "disease du jour" and a failure to appreciate the impact of long-term increases in knowledge and skills in all of the sciences on our quality of life, including the eradication of disease.

Finally, we live in the most technologically dependent society that has ever existed. Yet the vast majority of our population is technologically illiterate—not dumb, just uninformed. I mentioned Thomas Jefferson's opinion about the need for an informed citizenry—well, I think he would be worried. Sitting, as I do, at the nexus of science, engineering, and public policy, *I* am worried!

Despite our efforts here at the National Academies, I see policy makers every day expounding on national and homeland security, energy, and dozens of other technical issues about which they haven't a clue! I interact with journalists who, in my view, miss the real issues in a story because they have no knowledge of its technical dimensions. I put up with advocates of all stripes proposing technologically nonsensical solutions to very real problems. I worry about people having to make personal decisions—whether or not to disarm an air bag, for example—about which they have fragmentary information and no disciplined thought process on which to rely. Mostly, these are good, intelligent people who want to do the right thing, but they simply do not have the knowledge they need to function in a technological society.

I could describe more pieces in my mosaic, but I will get too depressed.

The Whole Picture

Each piece in my mosaic tends to be treated as a separate problem—and taken alone is, perhaps, not all that important a problem. But they all have one property in common, short-term thinking. None of them addresses the long-term ramifications of the overall problem. For me, viewing them in this overall context makes them all seem urgent.

Some of these problems we can do something about and some we can't—at least not directly. For example, we can help make engineering more attractive to U.S. students. NAE has established the Public Understanding of Engineering Program to create a better (and more accurate) image of engineering. We can also do something about engineering education to improve the curriculum and pedagogy. It is a disgrace, for example, that only about half of the students who enter engineering programs finish them. The students who leave are not poor students; in fact, in terms of GPA, grades in math and science, SAT scores, rank in high school, they are indistinguishable from those who stay. They are being driven out by our curriculum and pedagogy. And that is entirely fixable!

We have little direct leverage over the visa situation, but we have direct and personal knowledge of the benefits foreign students and scholars bring to the United States. As an organization, NAE is trying to convince policy makers of that. In addition, each and every one of us can write to our representatives and demand action.

We can do much the same with respect to other issues, like the imbalance in funding between the life sciences and the physical sciences and engineering. And each of us, in our companies or schools, can take ownership of the diversity issue.

I want to make a special plea on the issue of technological literacy. Engineering schools tend to see themselves as professional schools, and they only offer courses for engineering majors. So just how do we expect the other 95 percent of students to become technically literate? I believe engineering schools must accept the responsibility for the technological literacy of all students on campus. We cannot force them to take engineering courses, of course, but we can at least make the courses that are available accessible to the average liberal arts major. In fact, I think we have an obligation to do so.

Conclusion

Although I have touched on a broad range of topics and not gone into depth on them individually, I see them as pieces of a mosaic. And the composite image of the mosaic is one of disinvestment in the future. Each piece of the mosaic raises a topic of concern to engineers, and directly or indirectly, we can do something about them. Engineers have had a profound impact on society through the products, processes, and infrastructure we developed. Now it is time for us to contribute in another way.

Remarks by Eli Ruckenstein, 2004 Founders Award Recipient

(Left to right) Wm. A. Wulf, Eli Ruckenstein, Craig Barrett, Christopher Magee.

The 2004 Founders Award was presented to Eli Ruckenstein, Distinguished Professor of Chemical Engineering, State University of New York, Buffalo, for "leadership in modernizing research and development in key areas of chemical engineering." These remarks were delivered October 3, 2004, during the NAE Annual Meeting.

I feel honored, humbled, and at the same time overwhelmed to receive this year's Founders Award, which puts me in very distinguished company.

I was born in a small town in northern Romania in an atmosphere that might remind you of the novels of Gogol or Turgenev, and I was the first one in my family to go to college. The early days of my career coincided with the early days of communism in Romania. I was not a Party member, and I did not adjust easily to the very political tone of the times—which will not surprise many of you. I worked in a scientific community that had few if any connections to the outside world, and I would never have imagined, even in my most ambitious dreams, that I would find myself experiencing a moment like this, addressing you on this occasion.

Over the course of my career I worked on a broad range of problems, many of which are conventionally associated with subfields of physics, chemistry, and engineering. The fact that I am here addressing you on this special occasion underscores the interdisciplinary nature of our field. Engineering research is a synthesis, which implies learning whatever experimental or theoretical techniques are required to solve the problem at hand, many times forays into distant disciplines.

My own way of doing research was to a large extent defined by the fact that I am basically self-taught and that I don't belong to any school. I was educated during a special time in the history of Romania, my native country, immediately after the Second World War. A tradition in pure mathematics had existed for some time, and physics and chemistry benefited from the return of a few distinguished researchers who had been trained in the West. However, Romania had no tradition in modern engineering, and much of my education consisted of studying the available original literature on my own.

My early work on heat and mass transfer—which didn't get to the West until 1958—owes much to the encouragement and understanding of my teacher and mentor, Professor Bratu. In spite of uncertainties under the Communist regime and the fact that I had no formal authority, I managed to attract a group of somewhat younger colleagues with whom I cooperated on a variety of problems, from mass and heat transfer to kinetics of gene expression to interfacial phenomena and thermodynamics of small systems.

When I arrived in the United States at the age of 45, everything I had learned about the West was from a limited number of journals and an even more limited number of contacts with researchers from outside Romania. I was immediately struck, overwhelmed, by the amount and breadth of information that became available to me overnight. At that point, I really started a new career.

Since then, my work has extended in roughly five different directions: catalysis, colloids, separation of proteins, polymers, and material science. Don't worry, I'm not going to give a lecture on each of these directions. I just want to highlight the contributions closest to my heart.

In the area of catalysis, my students and I studied theoretically and experimentally the stability of small metallic clusters on catalytic supports. The other area of catalysis I was fascinated with was the kinetics of selectivity of supported catalysts. I also formulated a theory for the mechanism of oxidation by mixed oxides.

In the area of colloids, we developed a hydrodynamics of colloidal particles that accounted for double-layer and van der Waals interactions. We introduced the concept of "interaction force boundary layer." We also proposed thermodynamic theories of surfactant aggregation, microemulsions, and liquid crystals. I feel happiest about two contributions: (1) the reformulation of the classic theory of double-layer forces to include in a unified way the interplay between double-layer hydration forces and the hydration of ions; and (2) a unified kinetic approach to nucleation and growth.

Rather than boring you with descriptions of other theoretical efforts, I will mention some of our more recent technological innovations: the development of a solid-solution catalyst for CO_2 reforming of methane; the preparation of some interesting compounds for the storage of H_2; and the preparation of a paste with high thermal conductivity, which is now used in all IBM computers. In addition, from concentrated emulsions we developed various technologies for preparing polymers, conductive polymers, and membranes for separation processes.

My career has been and remains a great source of satisfaction to me. For this, I must acknowledge my students, postdocs, and collaborators who, through their hard work, patience, and dedication, have taught me, inspired me, and stimulated me. I would like to share this award with them.

I consider myself a very lucky man who has been surrounded by many guardian angels. I would not have survived my early days as a young assistant professor at the Polytechnic University in Bucharest without Professor Bratu, who protected me from my own inability to fit into the politically driven academic environment of communist Romania, where I was considered a dangerous reactionary.

I also want to thank Jim Davis, Bill Gill, and the Chemical Engineering Department at Clarkson University. They are responsible for my initial move to the United States, and I owe them all a debt of gratitude.

I also owe a debt of gratitude to Art Metzner and my colleagues at the University of Delaware who welcomed me and my family warmly, made us feel at home, and helped me through the years of adjustment to the American way of life and the American academic system. I will never forget their moral support for me and my wife during the two and one-half years we struggled to get our then-teenage children out of Romania.

I also want to acknowledge George Lee, Bill Gill, Ralph Yang, Carl Lund, and my colleagues at Buffalo for providing the resources and helping create the supportive atmosphere that made my last 25 years the most productive, enjoyable, and rewarding of my career.

Someone who deserves a special mention is Howard Brenner, who was instrumental in my getting my first position in the United States. I am deeply thankful to Howard for his help, encouragement, and friendship over the years.

Finally, I would like to thank my family, especially the person who has been the greatest influence in my life, my wife Velina, for her constant support, love, and selfless dedication during our 56 years together. Without her, I would not have the wonderful children I have, I would not have as many friends as I have, I would not have had the career I have had, and I certainly would not be here today.

Science has been my life, and I have been lucky to have the opportunity to wake up every morning excited about my work and curious about the next adventure. On my next birthday, I will turn 80, and I still hope to find that intangible Holy Grail. Your recognition tonight energizes me and encourages me to continue searching. I started my career in those dark days in Romania with little but ambition and youthful exuberance on my side. I now live in a free country, and I still have the ambition—and on a good day a little exuberance.

Thank you again for this honor, and thank you for your attention.

Remarks by John Brooks Slaughter, 2004 Bueche Award Recipient

2004 Bueche Award presented to John B. Slaughter. (Left to right) Christopher Magee, Wm. A. Wulf, John B. Slaughter, Craig Barrett.

The 2004 Bueche Award was presented to John Brooks Slaughter, president and CEO, National Action Council for Minorities in Engineering, for "support of engineering research and education within the National Science Foundation, many contributions to the development of science and technology policy, and lifelong dedication to increasing diversity in the disciplines of science and engineering." These remarks were delivered on October 3, 2004, during the NAE Annual Meeting.

I wish to express my appreciation to Dean Terry King and his associates from the College of Engineering at my alma mater, Kansas State University. Their persistence and tenacity made this award to me possible. I also want to thank everyone who wrote letters and supported my candidacy so earnestly. Most of all I want to thank my wife, Bernice, and my children, John Jr. and Jacqueline, whose sacrifices, support, and love have sustained me throughout my professional career. Bernice has been my cornerstone, my fountain of inspiration, and my unending source of strength; I dedicate this award to her.

I had the privilege of getting to know Art Bueche during the last few years of his life. He was extremely helpful and supportive when I took on the responsibilities of the National Science Foundation, and he was a wise and generous mentor. I cannot begin to express my humility as I receive this award recognizing his contributions as a statesman and an advocate for science and technology.

It is also with a sense of humility that I join the ranks of distinguished engineers and public servants who received the Arthur M. Bueche Award in the past. I am truly honored to have my name added to the list of luminaries from our profession who preceded me. Although I hardly deserve to be in their company, I am grateful to the National Academy of Engineering for this honor.

I am not joking when I say I was the first black engineer I ever met. African Americans and Latinos in Topeka, Kansas, in the 1940s and 1950s were hardly encouraged to achieve in science or engineering. Nor was Topeka a city with much that could be characterized as scientific or technological, with the exceptions of Menninger's for the former and the Santa Fe Railroad for the latter. Children in Topeka were unlikely to have neighbors or uncles who were scientists or engineers to inspire them to take courses in math and science or introduce them to the mysteries of nature and machines. Consequently, it was uncommon for youngsters, especially minority youngsters, to want to become engineers. My hardworking, loving parents kindled and kept alive my spark of interest in engineering. They kept it from being snuffed out by the indifference and low expectations of a public school system whose leaders were consumed with defending at all costs the tradition of segregated schools. This year we have been reminded of the demise of legally sanctioned segregated schools as we celebrate the 50th anniversary of Brown vs. the Board of Education of Topeka.

My eyes were opened to the opportunities and excitement of engineering at Kansas State University. In 1956, I was the only African American graduate in engineering at the university. On a recent visit to the campus, I was pleasantly surprised to meet a large group of African American and Latino engineering students who are enjoying their experiences, are supported and encouraged by a dedicated faculty and staff, and are succeeding. Although much remains to be done to eliminate disparities in access and opportunities, the same thing is happening on many campuses across the country.

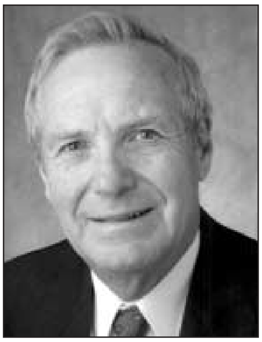
My organization, the National Action Council for Minorities in Engineering (NACME), is working to increase the number of educational institutions committed to improving their capacity to identify, enroll, educate, retain, and graduate minority engineering students. In 2003, NACME joined with the National Academy of Engineering, the National Academy of Sciences, MIT, Stanford, IBM, and DuPont in filing an *amicus* brief in the University of Michigan affirmative-action case before the U.S. Supreme Court. The statement expressed our conviction that "access to the educational opportunities available in our nation's selective universities is essential if underrepresented minorities are to have an opportunity to contribute to strengthening America's scientific and technological capabilities." The statement went on to say, "We can no longer afford the potential loss of creativity, productivity, and talent that results from policies and practices that inhibit the participation of any of our country's most valuable resource—our youth."

Although I am encouraged by the number of students in engineering from historically underrepresented minorities, I remain discouraged by the small number of minority faculty members in science and engineering. The argument often advanced by colleges and universities that there aren't enough minority Ph.D.s in science and engineering in the pipeline is not as tenable as it was a few years ago. Unfortunately, only about 1 percent of engineering faculty nationwide are African American or Latino, a situation that has not improved in the past 20 years, even though the number of black and brown Ph.D.s in science and engineering increases every year and more and more of them are available and fully prepared for faculty appointments.

Nevertheless, I am hopeful—no, I am confident—that the American higher education community will accelerate its efforts to create a more inclusive and pluralistic environment. Then, and only then, can it assist our nation in living up to its promise and ascend to a higher and nobler plateau, a place where both excellence and equity reside.

Thank you!

Opportunities and Challenges for Engineering Education as Seen from an Interdisciplinary Telecommunications Master's Degree Program



Frank S. Barnes

The 2004 Gordon Prize was presented in February 2004 to Frank S. Barnes, Distinguished Professor, University of Colorado at Boulder, for innovations in engineering education. This lecture was delivered on October 3, 2004, during the NAE Annual Meeting.

Today I wish to describe an interdisciplinary master's degree program in telecommunications that Professor George Codding and I started back in 1970. The program was designed to meet a significant worldwide need for students who not only knew the technology of telecommunications, but also had acquired a good understanding of the legal, political, and economic aspects of the industry. The program does not fit the normal mold for an engineering program, but it has enabled more than 2,000 students to play significant roles in the operation, management, and growth of the telecommunications industry over the last 34 years. It has also served as a model for many other programs both in the United States and around the world.

With this program as a model, I will describe some new opportunities and some serious problems in the way the current social structure of engineering schools limits the diversity and quality of the programs we can offer our students.

The Interdisciplinary Telecommunications Master's Degree Program at the University of Colorado at Boulder

In the late 1960s and early 1970s, many students considered engineering "bad" and responsible for many of the world's ills. Enrollments were down, and I was looking for a new way to attract students to electrical engineering. In a conversation with Dr. John Richardson, I learned of a need for people who not only understood the limits of communications technology, but also understood the economic and political realities of how the Federal Communications Commission (FCC) and others were determining the shape of the industry. John also pointed out to me that Professor George Codding, in the Political Science Department, was a world expert on the International Telegraph Union (ITU).

The result of George and my discussions was a proposal for a new M.S. degree funded in part by the National Science Foundation that included business, political science, sociology, and electrical engineering. The program had several new features. First, it was the first program that I know of that had a serious technical content for electrical engineers that also included law and policy as a fundamental part of the program. Second, none of us wanted to teach a service course for more than a few years, so the program had to be open to essentially any student with a bachelor's degree. This forced us to develop the program that had very minimum prerequisites and took advantage of the students' maturity. With the limit of one to two years for a full-time student to finish the program and considering the breadth of subject matter to be covered, the courses could not require four years of prerequisites in any subject. The required math background, for example, was limited to algebra, trigonometry, and calculus. To add depth, we required a thesis so that the student had some depth of knowledge in at least one subject and had shown an ability to write a coherent document.

In 1971 the program started with a single student with an undergraduate degree in international relations. That student is now an attorney with a practice in telecommunications. Over the next 30 years, the program grew to a peak enrollment of about 500 students—half of whom were participating via distance education. About half of the students have engineering degrees, and half have undergraduate degrees in a wide variety of other disciplines. The program now offers more than 20 courses, approximately half of which are technical and half of which cover economics, business, law, and policy. The current requirements are given in Table 1 below; brief course descriptions are given in Appendix I.

The electives include a wide range of topics, from wireless and optical fibers through protocols and security to standards and

TABLE 1

- The M.S. degree in telecommunications in the School of Engineering and Applied Sciences at the University of Colorado at Boulder requires 36–37 master's credits.
- Two courses: TLEN 5300 and 5836 are designed to improve the proficiency of students in math, writing, and presentations skills. These courses are prerequisites but may be waived by the faculty advisor. These credits are not included in the required total for the degree.

Degree requirements include:

1. Four technical courses for a total of 12 credits, including TLEN 5310 (Telecommunications Systems) and TLEN 5330 (Data Communications 1) and two additional 3-credit courses.
2. Two economics or business courses for a total of 6 credits, including TLEN 5010 (Telecommunications Finance 1) and one additional business or economics course.
3. One policy and law course for a total of 3 credits.
4. One programming course for a total of 3 credits.
5. Two telecommunications seminars for a total of 2 credits.

Two options for final requirements:

1. Two research seminars for a total of 2 credits.
2. Three additional 3-credit courses electives for a total of 9 credits.

   or

3. One elective course for a total of 3 credits.
4. TLEN 5700 research course and 6 thesis credits (may substitute 3-credit research methods course and 4 thesis credits).

telecommunications law. These options make it possible for students to develop strengths in their respective areas of interest.

Courses such as the required course for some programming can be waived if the student already has a background in that area. In general, I advise students with nontechnical backgrounds to take technical courses beyond the requirements—especially the lab. I advise students with engineering backgrounds to take additional business and policy courses.

In my view, this program has been remarkably successful in making it possible for a wide variety of students to pursue successful careers in fields that would not otherwise have been open to them. Examples include: a senior vice president at Cox with a nontechnical background who, more than 20 years later, is using material he learned in class on characteristics of communications systems; a young lady in Mexico who took my course in optical fibers who now holds a position corresponding to an FCC commissioner and who is one of a very few engineers in Mexico with some background in policy and economics; a recent graduate who has been implementing a telecommunications system that allows doctors at the Harvard Medical School to prescribe drugs directly to nurses in the slums of Lima, Peru, and has cut the error rate in drug delivery from 17 percent to less than 5 percent.

It is useful to reflect on some of the things that we have learned in the 34 years of teaching this M.S. program in telecommunications. First, many students do not know at age 18, when they leave high school, that they will need a technical background to access many of the jobs that could lead to satisfying careers for them. Even after they graduate from college, it may take several *more* years for them to learn that they need a technical background. But by then, they may not be able to take four years to start over as a freshman in engineering, math, or physics. However, we have found that they can learn, in just one or two years, much of what they need to operate and manage major communications systems. These students often bring verbal and people skills with them that many engineering students do not have, and these skills often lead them into top management positions.

Many of the toughest problems engineers face today involve dealing with the public to implement technical solutions to societal problems. Government bodies, including the FCC, the ITU, and state public utilities commissions, set the rules by which competitors can compete. These rules, in turn, often determine which technologies are used and developed and who gets rich or goes broke. For example, look at the current growth in mobile telephones. This technology was held up for more than 20 years in the United States because the bandwidth was allocated to UHF TV. Another example is the current rules on the universal service fund relating to subsidies for rural telephones and voice-over IP. Another class of problems shows up in locating cell phone base stations in places where there are public concerns about possible health effects, aesthetics, and possible reductions in property values.

Similar problems occur in the power industry related to who builds transmission lines in a competitive environment and where power plants are located. In civil engineering, there are related problems in managing river basins, sewer systems, and traffic. In all of these cases, engineers need to know more than the technology if they are going to take on leadership positions.

Some New Directions for Engineering Education

Based on our experience with the telecommunications program, I suggest that we have the ability to open new opportunities for our engineering undergraduates and M.S. students by providing them with a background in public policy and an overview of major worldwide problems in infrastructure. Engineers need more than technical competence to satisfy the needs for energy, communications systems, and control of environmental pollution. Courses on risk analysis, environmental law, resources and population growth, and security would all be good electives or aspects of the social science requirements. Many major challenges for engineers in the future will require solving political and economic problems in addition to technical problems. Technology will be only a part of the problem, and often the easy part. For example, building a nuclear power plant is much easier than persuading people to let you build it in their "backyard."

It is also much easier to get faculty agreement that a new course would be valuable for some of their students than it is to remove a course from the curriculum. The trick in adding any new material to the curriculum is fitting it in without decreasing the quality of the technical education already offered. As it is, many technical concepts students should know cannot be covered for lack of time. And students often do not master the material as completely as we think they should.

The key to adding new material is twofold: (1) improving our choice of materials so students leave our program with a solid technical and liberal arts foundation and the ability to acquire new skills rapidly on the job; and (2) improving the efficiency with which the students master the material presented in class. Each faculty will have to make its own choices about how to do this. However, I would like to suggest some options that I hope we will be able to implement, at least in part, at the University of Colorado.

A first step is motivating students to become deeply involved in a subject to the extent that they will spend more time on it than it takes to earn a passing grade. This can be done in many ways—by demanding more with tough tests and frequent quizzes or by getting students involved so that they are self-motivated by interest in the subject or by peer pressure.

One possible approach is to get students involved in teaching other students on a systematic basis, because you learn a subject more thoroughly when you try to teach it. Because students do not want to be embarrassed in front of their peers, they prepare more thoroughly for teaching; they also learn organizational and communications skills. Thus, we might have sophomores or juniors running help sections for the freshman calculus classes. This would give the freshmen more personal attention and force the upper-classmen to review and understand the fundamentals of classes they have already taken. Teaching could be carried all the way up through the senior year, with fifth-year students and graduate students helping teach upper-class courses. There will be fluctuations in quality, as always. However, seminars on teaching, how people learn, ethics, and sexual harassment based on problems encountered in actual teaching experiences might well be worked into the curriculum.

A second approach is for some classes to be taught jointly across continental boundaries. Our students are going to have to compete in a global economy, and more and more design teams are going to be geographically dispersed. A number of schools already team students from schools in Europe and the United States who must work together to solve a common design problem. At the University of Colorado, we have two start-up courses, one on advanced aircraft design and another on embedded systems. These courses will be jointly taught in Europe and at the University of Colorado with students in both places receiving some lectures from each school. Challenges in working across cultures and time zones should help prepare students to work on worldwide design teams after they graduate.

We are also developing labs that can be operated over the Internet so that students can do at least part of their lab work from anywhere and at times of their choosing. Many of the experiments in the telecommunications lab are being made available this way (the use of routers, etc.), and a lab in optical communications is being jointly developed at CU and the University of Houston.

These labs will have the advantage of increasing the hours available for using expensive computer-controlled equipment. In addition, no school can either afford, or find, all of the faculty they would like to have. Using telecommunications to share faculty and import course material may be one way to strengthen programs.

The next step in improving our teaching efficiency will require a better understanding of how students learn. This understanding is likely to be based on an improved understanding of how information is stored and processed in the brain, as well as by acquiring information from psychology and educational theory. For example, a Russian program that builds a profile of a student's learning style by looking not only at the kinds of wrong answers the student gives but also the time delay in their decision making caught my attention when the Russians claimed they could reduce the time required to learn a course like geography by a factor of three over standard results.

Problems in Developing New Educational Programs

1. Educational institutions and faculties are extremely conservative and it is very hard to initiate a program that does not fit into the current departmental structure. It is even more difficult to keep it going, even if the students are very successful, if the program covers a broad range of material and does not include a hot research topic that is well funded. In addition, faculty may have a difficult time getting tenure if they are not doing research that fits into a standard area that senior members in their home department consider important. Often the time taken to develop a course for another school or program is considered a waste of time. Finally, students in these courses are seldom in a position to contribute to Ph.D.-level research in the home discipline. Thus, working on an interdisciplinary program at either the undergraduate or master's level may be hazardous to an assistant professor's promotion and tenure.

2. A second major hazard is the amount of time it takes to raise research funding. With current funding rates for grants from NSF and NIH running at 3 to 20 percent, 80 to 97 percent of the faculty are wasting their time. Imagine going to class on the first day and telling students that, on the average, 80 to 97 percent of them are going to fail the course! To me this suggests that many faculty would be better off doing the research they can with limited or no funding or spending their time on improving the courses they teach.

   The audience here today is made up largely of those who have succeeded in raising the money they needed to do important things. In most cases, this has been done at a cost of lots of time. It is not unusual for successful senior faculty members to spend 50 percent of their time writing grant proposals and reports or attending contractor meetings just to keep their programs funded. This suggests that we have done a poor job of systems engineering in distributing research money and that some of the leadership in this room should make a serious effort to improve the system.

3. Using student evaluations to determine raises and promotions is often destructive. Student ratings are frequently a measure only of how well a course matches the student's expectations. If a course is much more demanding than expected or if one gives tests with low exam scores (independent of the grades at the end of the term), one has a high probability of receiving low student ratings. Students are seldom in a position to evaluate the importance of the material until well after they graduate, and the choice of material is one of an instructor's most important functions. In my view, student ratings can be used in about 10 percent of cases to improve instruction, but they should have little effect on promotion, tenure, and raises. In too many cases, instructors are less demanding than they should be, thus sending the wrong message about how courses should be taught.

Summary

We were fortunate to develop an interdisciplinary program in telecommunications at the right time that has enabled many technical and nontechnical students to pursue successful careers that would not have been available to them otherwise. I believe that this program can serve as a model for the development of programs in other fields that will also allow students to take leadership roles in solving major infrastructure problems throughout the world.

In closing, I would like to thank the National Academy of Engineering, Bernard Gordon, and Bill Wulf for giving me the opportunity to speak to you this afternoon. I also want to express my deep appreciation

to the many students and faculty members who have made our program a success. Without their accomplishments, there would have been nothing for me to report.

Appendix I

Telecommunications Economics

Introduces students to a range of microeconomic principles and models used to undertake economic and financial analyses of telecommunications and e-commerce networks.

International Telecommunications Strategy

Covers issues involved in developing a business plan for a telecommunications investment in markets outside the United States, including estimating revenues and costs, sources of data, network costing issues, and interconnection.

Telecommunications Business Strategy

Covers concepts, strategies, and practical implementation of market-oriented business strategy in the telecom industry with plenty of real-world examples.

Standardization and Standards Wars

Provides students with a tool set of economics, business, and legal principles involved in the technology standardization process and covers topics including the role of various standards models (e.g. open vs. closed, proprietary vs. non-proprietary, and de facto vs. de jure), as well as a survey of international standard-setting organizations and their respective roles.

Telecommunications Law

Examines laws governing telecommunications industries, including federal and state regulation and international aspects. Includes telephone, cable, satellite, cellular, and other wireless systems and the Internet.

Telecommunications Theory and Applications

Provides necessary mathematical and engineering background to study telecommunications, including physical units, numbering systems, trigonometric functions, sine waves, attenuation, logarithms, indices, decibels, complex numbers, calculus, elementary probability, and power and circuit analysis.

Telecommunications Systems

Provides an in-depth look at basic telecom terminology and concepts, introductions to voice and data networks, signaling, and modulation/multiplexing.

Data Communications 1

An up-to-date technical survey of data and computer communications starting with the basics and including details about Wireless, MAN, and WAN systems and standards that include T-carrier, SONET, HDLC, Frame Relay, and ATM.

Telephony

Covers fundamental requirements for voice communication and processing, multiplexing, synchronization, switch architecture, and subscriber access. Transport and access technologies discussed include ISDN, frame relay, asynchronous transfer mode, ADSL, SONET, and signaling system no. 7. Packet speech services, protocols, and operation over frame relay, ATM, and Internet protocol are also considered.

Satellite Communications

Discusses fundamental concepts of communication satellites including orbital mechanics, spacecraft and earth station system configuration, link budgets, propagation issues, modulation and multiplexing techniques, multiple access schemes (FDMA, TDMA, CDMA), orbit selection, error control coding, and satellite network architecture.

IP Routing

This course breaks IP routing technologies into two fundamental pieces: an in-depth study of Interior and then Exterior Gateway Protocols (IGPs and EGPs).

Data Communications 2

Expands on the topics associated with Internet protocols, processes, and selected applications. Introduces Internet-related protocols and applications, including the application layer (HTTP, FTP, SMTP, DNS, and sockets), the transport layer (UDP, TCP), routing (RIP, OSPF, BGP), IP version 6, multimedia (RTP), network security (IPSec), and network management (SNMP).

Multimedia Networking

Focuses on how multimedia services and networking technology are co-evolving to support convergence, with Voice-over-IP developed in detail.

Advanced Telecommunications Lab

Students work in teams to learn the fundamental techniques of signal transmission and impairment measurement, voice and data switching, systems administration, and the fundamental functions of data networking and services.

Signaling Systems

Develops an understanding of modern signaling protocols and differences among these protocols, including an appreciation of actually implementing signaling protocols in the Internet Protocol environment.

Wireless and Cellular Communications

Covers miniaturization, clever use of radio spectrum, the expanded roles of the cellular phone and pager, PDAs, the popularity of the IEEE802.11 networks, the new breed of low earth orbiting satellites, and even the Global Positioning Systems (GPS).

Telecommunications Seminar

National and local speakers give lectures on technical, business, policy, and regulatory issues in telecommunications.

Capstone 1

Capstone 1 develops basic quantitative and qualitative research techniques.

Network Management and Operations

Covers both technological and managerial aspects in network management through an examination of the capabilities and needs of organizational network and systems management, including network management standards such as SNMP, RMON, and CMIP.

Wireless LANs

Discusses packet wireless data communication with an emphasis on wireless LANs.

Network Security Lab

Lab work covering system hardening, firewalls, intrusion detection, vulnerability assessment, and investigation.

Foundations of Computer and Network Security

Introduces students to computer and network security, the branch of computer science that studies methods and techniques to protect critical information resources.

Spectrum Policy and Management

The radio spectrum is a natural resource that is increasingly crucial to the economic and social well-being of all people. This course focuses on how this resource is being managed in the face of escalating demand, rapidly changing technology and increasing globalization.

Privacy, Security, and Digital Rights Management

Introduces students to the laws that regulate the basic technologies of the Internet and the management of information in the digital age.

Capstone 2

Requires teams of 3 or 4 students to select a topic, research and document this topic through the guidance of a faculty member, and present their findings to their fellow students in the form of a 20-minute defense (including time for questions and discussion) and a supporting paper.

Telecommunications Application Programming

Students learn server-side, object-oriented programming as a cross-platform application development environment.

Unix, C, C++

This course is recommended for students who anticipate a need for knowledge of both UNIX/Linux and C programming in their future careers.

Applied Network Security

Technical discussion of threats, vulnerabilities, detection, and prevention is presented, including issues such as cryptography, firewalls, network protocols, intrusion detection, security architecture, security policy, forensic investigation, privacy, and the law.

Law for the Information Age

This seminar addresses the numerous legal issues that arise in relation to the Internet, in particular issues raised by telecommunications regulation, intellectual property, and antitrust and constitutional law, highlighting the continuing challenges posed by technological change for economic regulation.

Research Techniques

Provides both quantitative and qualitative techniques for use in capstone projects, master's theses, and Ph.D. dissertations.

Video and Technology

Students will obtain an in-depth knowledge of the specifics of various technical aspects of television and visual human perception and how image transmission and delivery fit into the telecommunications industry market mix.

Leadership and Management

Provides a background in leadership philosophy, principles, and practices to enable new managers to develop skills in leading and managing, including organizational leadership, leadership styles, leadership tools, roles of effective leaders, change management, and business ethics. Same as EMEN 5050.

Optical Communications

Lectures, readings, and projects in the field of optical communications systems theory and practice, including optical fiber properties, optical transmitters and receivers, Raman and Erbium optical amplifiers, modulation, optical link design, and wavelength division multiplexing optical networks (undersea, long haul, metro, local loop, and LAN).

Microeconomics (Telecom)

Introduces economic theories of consumer choice, production, firm and market organization, and general equilibrium and applies economic models to telecommunications demand, costs, pricing and strategic interaction, and network externalities.

IP Routing Design

Applies IP routing theory directly to design-based networking problems and real-world examples incorporating physical design, logical design, financial analysis, and laboratory implementation.

Telecommunications Finance 2

This course introduces a relatively new methodology, real options, to evaluate investments and projects to account for uncertainty, with application to telecommunications investments.

Independent Study

An opportunity to study an area of the student's interests under the guidance of a faculty member.

Thesis

A master's thesis is a substantial body of original scholarly work by an individual student.

Call for Nominations for 2005–2006 Awards

For more than 40 years, NAE has recognized outstanding engineers who have made a lifelong commitment to advancing the human condition through engineering achievement and/or innovation in engineering and technology education. NAE awards five prizes worth more than $2.5 million (biennially) for achievements in engineering and engineering education—the Founders Award, Arthur M. Bueche Award, Charles Stark Draper Prize, Fritz J. and Dolores H. Russ Prize, and Bernard M. Gordon Prize.

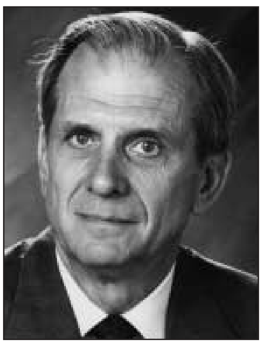
The Founders Award and Arthur M. Bueche Award are presented annually in October during the NAE Annual Meeting. The Founders Award is presented to an NAE member or foreign associate who has exemplified the ideals and principles of NAE through professional, educational, and personal achievement and accomplishment. The Bueche Award is awarded to an engineer who has been actively involved in determining U.S. science and technology policy, promoting U.S. technological development, and improving relations between industries, government, and universities.

The Charles Stark Draper Prize and Bernard M. Gordon Prize are presented annually, and the Fritz J. and Dolores H. Russ Prize is presented biennially, in February during the National Engineers Week celebration at the NAE Annual Awards Dinner and Presentation Ceremony. The Charles Stark Draper Prize is a $500,000 cash prize honoring an engineer whose contributions have significantly improved the quality of life, made it easier for people to live freely and comfortably, or provided access to information. The Fritz J. and Dolores H. Russ Prize is a $500,000 cash prize honoring achievements in bioengineering worldwide that have improved the human condition or had a significant impact on society. The Bernard M. Gordon Prize for Innovation in Engineering and Technology Education is a cash prize of $500,000, split equally between the recipient and his or her institution for the continuation of the award-winning program. The Gordon Prize honors technology educators whose innovative programs have strengthened the engineering workforce by cultivating students' communication skills, creativity, and teamwork.

NAE takes pride in identifying and saluting the best in engineering (see our website, *www.nae.edu/awards*, for a list of previous award recipients). We invite you to help us continue the tradition of honoring individuals or teams by nominating outstanding colleagues, professors, peers, and industry leaders for the upcoming prize cycle. Nominations for the 2005 Founders and Bueche awards and the 2006 Draper and Gordon prizes will be accepted **January 4, 2005**, through **April 8, 2005**.

Members and foreign associates will receive nomination materials by mail. Nonmembers who wish to submit nominations can obtain materials from Leila Rao at (202) 334-1237 or *lrao@nae.edu* or download them from our website, *www.nae.edu/awards*. Mail or fax your nomination to: NAE Awards, National Academy of Engineering, 500 Fifth Street, N.W., NAS 308, Washington, DC 20001. Fax: (202) 334-1595.

Report of the Home Secretary



W. Dale Compton

The proposed changes in the NAE Articles of Organization were approved by a vote of 798 to 40, with 1 abstention. The proposed changes in the NAE Bylaws were approved by a vote of 740 to 11, with 1 abstention.

The Membership Policy Committee (MPC) has concluded a two-year assessment of membership procedures and presented its conclusions to the NAE Council on May 13, 2004. The recommendations that were accepted are summarized below:

- Election to NAE membership should continue to be based on accomplishment, rather than other criteria, such as future promise. NAE should continue to develop its young engineer program and similar programs to provide visibility for new, promising technologies and technologists.
- NAE should continue to encourage NAE members to participate in NRC and NAE study task forces. Efforts will continue to identify members with expertise and interests in particular studies.
- The approximate distribution of members among business, academia, and other sectors is 5/4/1, which is considered appropriate.
- Available data indicate that efforts to encourage the nomination of underrepresented minorities and women have been effective. Rapid changes in the demographics of the profession are certain to continue, however, and will require continued efforts to encourage nominations of minorities and women, as well as members of other ethnic categories.
- Replacement nominations should be in the hands of a "nominating group" made up of the section chair, the section vice chair, and the peer committee chair from the previous election cycle. Replacement nominations will be vetted by the home secretary in consultation with the president prior to submission to the NAE Council.
- Data on the sections show that small sections are indeed growing, although perhaps not as fast as some would prefer. Experience has shown that management's encouragement of nominations of strong candidates is critical to section growth.
- After careful consideration of the effect of the number of references for nominees (currently a minimum of three and a maximum of four) on the probability of election, the impact on the peer committee of fewer than four references, and the opportunity, already in place, for one reference per nomination to be based on study of a nominee's accomplishments rather than on personal acquaintance with the nominee's work, the MPC unanimously concluded that the minimum number of references should not be changed.

The NAE Council has asked the current MPC to examine the following issues:

- The ultimate size of the Academy and actions that should be taken to achieve it.
- Actions that can be taken to increase the number of nominations of individuals from the business sector.
- Review of the experiences of the ad hoc Peer Committee 13 during the 2002 and 2003 elections to determine if the process, or a modification of it, would be useful for identifying emerging technologies.

W. Dale Compton
Home Secretary

New Program Officer for the Committee on Engineering Education

Richard Taber

Richard Taber joins NAE as the new program officer for the Committee on Engineering Education (CEE). Mr. Taber joins us after three years with Oak Ridge Associated Universities as a corporate and foundation relations consultant with the National Science Foundation. In that position, he was responsible for the operational management of an alliance of public and private funders of precollege and undergraduate education with the goal of identifying and disseminating effective funding strategies.

Mr. Taber's prior positions include senior research engineer with the Institute of Textile Technology and process engineer and production manager with Milliken & Company. He received his bachelor's degree from the Philadelphia College of Textiles and Science (now Philadelphia University) and his master's degree from the Institute of Textile Technology.

CEE is engaged in numerous activities, including the multiphase Engineer of 2020 Project. The Phase I report, which focused on defining the characteristics of desired student outcomes, has been immensely popular and is now in its second printing. This past summer a Phase II workshop was held on defining implementation strategies for achieving the desired student outcomes identified in the Phase I report. The report of the Phase II workshop is scheduled for publication in early 2005.

Education Research Scholar Joins CASEE

Dr. Barbara Lovitts has joined the Center for the Advancement of Scholarship on Engineering Education (CASEE) as a senior program officer. She is leading a project to develop a prototype database of educational practices in engineering and science education. Dr. Lovitts is a distinguished researcher, author, and invited speaker with expertise in quantitative and qualitative education research, project and program evaluation, and grant program management. She joins NAE from the College of Behavioral and Social Sciences at the University of Maryland, College Park, where she was a research associate in charge of activities of two Alfred P. Sloan Foundation-funded studies, including *Making the Implicit Explicit: A Blueprint for Creating Performance Expectations* and *Assessing the Outcomes of Doctoral Education* (Stylus, 2006).

Dr. Lovitts' experience includes senior research analyst with the American Institutes of Research and program director in the Research on Learning and Teaching Program of the National Science Foundation. Upon joining NAE, Dr. Lovitts stepped down as a member of the National Academy of Sciences Committee on Assessment for NIH Minority Research Training Programs. Dr. Lovitts is the author of the widely cited *Leaving the Ivory Tower: The Causes and Consequences of Departure from Doctoral Study* (Rowman and Littlefield, 2001). She received her doctoral degree in sociology from the University of Maryland, College Park.

CASEE Fall Student Intern

Sarah Hunt

Sarah Hunt, a Ph.D. candidate in ecological and environmental anthropology at the University of Georgia, is interning this fall with the Center for the Advancement of Scholarship on Engineering Education (CASEE). Ms. Hunt's primary project for the fall term is identifying literature on changes in student learning outcomes as a result of academic or co-curricular interventions. This literature will then be incorporated into a prototype database for use by engineering faculty who may have little experience with engineering education research but might benefit from the findings of such research.

Ms. Hunt earned her bachelor's degree in biology from Earlham College, where she was admitted to Phi Beta Kappa. She hopes her internship with the National Academies will provide experience to advance her long-term career goals, which include becoming a faculty member and researcher on educational activities that span anthropology, engineering, and environmental studies, with a focus on the nexus of technology, innovation, and policy.

Janice Tsai, NAE Intern

Janice Tsai

Janice Tsai will earn a master's degree in library and information science at Rutgers University next spring. As an Eagleton fellow at the Eagleton Institute of Politics at Rutgers, Ms. Tsai studied the practice of politics, agenda-setting, lobbying, and communications in government. She completed her B.A. in mathematical methods in the social sciences and political science at Northwestern University. Prior to her return to graduate school, she spent two years in Chicago working at IBM designing network architectures for web hosting systems.

During her three-month internship at NAE, Ms. Tsai worked with the Committee on Diversity in the Engineering Workforce on the Diversity Innovation Initiative to find ways to increase diversity in engineering. She also assisted the *EngineerGirl!* project by developing funding proposals to support the addition of interactive web content.

As a Christine Mirzayan Graduate fellow, Ms. Tsai says her experience at NAE "was invaluable in exposing me to the intersections of engineering and policy." Her career goals are to help develop and guide technology policy through research, quantitative analysis, and experience, especially in the areas of intellectual freedom and copyrights, information security, and collaborative research. She will pursue a master's or Ph.D. in public policy beginning next fall.

Tenth U.S. Frontiers of Engineering Symposium

Laura Ray (Dartmouth) demonstrates the design rationale for scalable mobile robots for use on the Antarctic plateau.

NAE hosted the tenth U.S. Frontiers of Engineering (FOE) Symposium at the Beckman Center in Irvine, California, September 9–11. This milestone meeting was attended by approximately 110 engineers, and talks covered topics in engineering for extreme environments, designer materials, multiscale modeling, and engineering and entertainment. In February 2005, NAE will publish a symposium volume containing extended summaries of the presentations.

On the second afternoon of the symposium, representatives of the Defense Advanced Research Projects Agency (DARPA), the U.S. Department of Defense (DOD) Office of Basic Science, and National Aeronautics and Space Administration (NASA) gave presentations to the group about their missions and work.

This year's dinner speaker was Alexander Singer, a film director who has served on two National Research Council committees, one on the convergence of telecommunications, computers, and entertainment and one on opportunities for collaboration between the defense and entertainment research communities. His talk, "Unlikely Partners: DARPA and Me," touched on numerous issues at the intersection of entertainment and engineering. He reminded Frontiers participants that their work is as creative as the work of artists.

Pablo P. Debenedetti, Class of 1950 Professor, Department of Chemical Engineering, Princeton University, chaired the organizing committee and the symposium. The 2005 organizing committee, also chaired by Dr. Debenedetti, has begun planning for the next U.S. FOE meeting, to be held September 22–24, 2005, at GE Global Research Labs in Nisakayuna, New York.

Greg Carman (UCLA) describes thin-film active materials at the U.S. Frontiers of Engineering Symposium.

Funding for the 2004 symposium was provided by the Air Force Office of Scientific Research, DARPA, DOD (DDR&E-Research), NASA, Eastman Kodak, Microsoft Corporation, Cummins Inc., ATOFINA Chemicals Inc., Air Products and Chemicals Inc., and Dr. Ruth M. Davis and other individual donors.

NAE has been hosting an annual U.S. FOE symposium since 1995. FOE also hosts bilateral programs with Germany and Japan. FOE meetings bring together outstanding engineers from industry, academia, and government at a relatively early point in their careers (all participants are 30 to 45 years old) and provide an opportunity for them to learn about developments, techniques, and approaches at the forefront of fields other than their own. As engineering becomes increasingly interdisciplinary, interactions with people in other disciplines are critical to researchers on the cutting edge of new concepts and technologies. The meeting also facilitates the establishment of contacts and collaborations among the next generation of engineering leaders.

For more information about the symposium series or to nominate an outstanding engineer to participate in future FOE symposia, contact Janet Hunziker at the NAE Program Office at (202) 334-1571 or by e-mail at *jhunziker@nae.edu*.

Calendar of Meetings and Events

2004

December 1–2	Invention Session, Committee on Diversity in the Engineering Workforce
December 4	NAE Committee on Membership Meeting
December 9	News and Terrorism: Communicating in a Crisis Workshop Miami, Florida
December 16	Committee on Engineering Education Meeting

2005

January 12	NRC Executive Committee Meeting
	NRC Governing Board Meeting
February 4	NAE Membership Policy Committee Meeting Irvine, California
February 7–8	NRC Governing Board Meeting Irvine, California
February 9–10	NAE Council Meeting Irvine, California
February 10	NAE National Meeting Irvine, California
February 16	NRC Executive Committee Meeting
	NRC Governing Board Meeting
February 20–26	National Engineers Week
February 21	NAE Awards Dinner Ceremony and Presentation
February 24	NAE Finance and Budget Committee Conference Call
March 16	NRC Executive Committee Meeting
	NRC Governing Board Meeting
March 22	NAE Regional Meeting University of Southern California, Los Angeles
April 21	NAE Regional Meeting University of Minnesota, Minneapolis

All meetings are held in the Academies Building, Washington, D.C., unless otherwise noted. For information about regional meetings, please contact Sonja Atkinson at *satkinso@nae.edu* or (202) 334-3677.

In Memoriam

JOHN H. ARGYRIS, 87, professor emeritus, University of Stuttgart, died on April 2, 2004. Dr. Argyris was elected a foreign associate of NAE in 1986 for pioneering and continuing contributions in computer mechanics over a period of more than 30 years.

RICHARD M. CARLSON, 79, retired chief, Advanced Systems Research and Analysis Office, U.S. Army Aviation and Troop Command, and consultant, died on July 12, 2004. Dr. Carlson was elected to NAE in 1990 for his contributions to the application of composite materials to operational helicopters.

HAROLD CHESTNUT, 83, retired consultant, GE Corporate Research and Development Center, died on August 29, 2001. Dr. Chestnut was elected to NAE in 1974 for contributions to the theory and practice of control systems and for systems engineering.

FLOYD L. CULLER, 81, president emeritus, Electric Power Research Institute Inc., died on September 29, 2004. Mr. Culler was elected to NAE in 1974 for contributions to the development of nuclear power.

BOB O. EVANS, 77, partner, Rocket Ventures, died on September 2, 2004. Mr. Evans was elected to NAE in 1970 for contributions to the development of nuclear power.

MAXIME A. FAGET, 83, retired director, Space Industries Inc., died on October 9, 2004. Dr. Faget was elected to NAE in 1970 for contributions to the design and engineering of the Mercury and Apollo spacecraft.

WILLIS M. HAWKINS, 90, consultant, Lockheed Martin Corporation, died on September 28, 2004. Dr. Hawkins was elected to NAE in 1966 for the design and development of aircraft, missile, and space systems.

BERTRAM WOLFE, 77, retired vice president and general manager, GE Nuclear Energy, and consultant, died on September 6, 2004. Dr. Wolfe was elected to NAE in 1980 for contributions to the development of advanced nuclear concepts and projects, particularly fast flux reactors.

Publications of Interest

The following reports have been published recently by the National Academy of Engineering or the National Research Council. Unless otherwise noted, all publications are for sale (prepaid) from the National Academies Press (NAP), 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055. For more information or to place an order, contact NAP online at <*http://www.nap.edu*> or by phone at (888) 624-8373. *(Note: Prices quoted by NAP are subject to change without notice. Online orders receive a 20 percent discount. Please add $4.50 for shipping and handling for the first book and $0.95 for each additional book. Add applicable sales tax or GST if you live in CA, DC, FL, MD, MO, TX, or Canada.)*

Accident Precursor Analysis and Management: Reducing Technological Risk through Diligence. In the aftermath of catastrophes, it is common to find prior indicators, missed signals, and dismissed alerts, that is, precursors that might have signaled impending danger and even averted the catastrophes. This report of an NAE workshop in July 2003 documents industrial and academic approaches to detecting, analyzing, and benefiting from accident precursors and examines public-sector and private-sector roles in the collection and use of precursor information. The report includes an analysis, findings, and recommendations by the authoring committee and 11 individually authored background papers on four topics: the opportunity of precursor analysis and management; risk assessment; risk management; and linking risk assessment and risk management. Paper, $35.00.

Air Quality Management in the United States. According to this report, the implementation of air quality regulations should be less bureaucratic—with more emphasis on results and less on process—and should protect ecosystems as well as people. The report recommends that the Environmental Protection Agency target groups of pollutants rather than individual pollutants and that revised or new regulations address air pollution that travels across state and international borders. This will require better tracking of emissions to assess which populations are at the highest risk of health problems from pollution and to measure the efficacy of pollution-control strategies. Hardcover, $49.95.

Analytical Methods and Approaches for Water Resources Project Planning. This review of analytical and planning methods used by the U.S. Army Corps of Engineers (USACE) focuses on the steps outlined in the federal "Principles and Guidelines" document and in the USACE "Planning Guidance Notebook." The study concludes that clear, comprehensive planning studies are often hindered by inconsistent legislation and limited interagency collaboration. The report recommends that the administration and Congress, in cooperation with the states, work toward reconciling inconsistencies in national water-related policies and create a body to coordinate water policies among federal agencies with water-management responsibilities. The committee also recommends that the federal "Principles and Guidelines" document, which was last revised in 1983, be revised and updated. Paper, $28.50.

Army Science and Technology for Homeland Security: Report 2—C4ISR. Shortly after the events of September 11, 2001, the U.S. Army asked the National Research Council to conduct a series of studies on how the Army could use science and technology to meet its homeland defense obligations. The first report, *Science and Technology for Army Homeland Security—Report 1*, included a survey of a broad range of technologies and recommended applying Future Force technologies to homeland security wherever possible. The committee concluded that the Army should play a major role in providing emergency command, control, communications, computers, intelligence, surveillance, and reconnaissance (C4ISR) capabilities and that the technology and architecture for homeland security C4ISR was compatible with the technology of the Army's Future Force. This second report focuses on how C4ISR can facilitate the Army's efforts to assist the U.S. Department of Homeland Security and emergency responders during or after catastrophic events. Paper, $36.50.

California Agricultural Research Priorities: Pierce's Disease. The glassy-winged sharpshooter is one of the recent invasive pests afflicting California agriculture. The sharpshooter transmits a bacterial pathogen that causes Pierce's disease, which impairs production of wine, table, and raisin grapes in California. The report recommends improving the process and setting priorities for research on Pierce's disease funded by state agencies and wine industry groups. Research should be focused on identifying feasible options for controlling the spread of the disease and designing sustainable approaches that are adaptable and affordable over the long term. Areas of research that should be pursued more intensely include: the genetic makeup of the pathogen that triggers Pierce's disease; mechanisms for making grapes resistant to the disease; the introduction of predators to the sharpshooter; and the management of crop planting to avoid spread of the disease. Paper, $35.00.

Charting the Future of Methane Hydrate Research in the United States. Methane hydrate is a natural form of clathrate—a chemical substance in which one molecule forms a lattice around a "guest" molecule with chemical bonding. In methane hydrate, the guest molecule is methane and the lattice consists of water that forms an ice-like solid. Methane hydrate has become the focus of international attention because of its vast potential for human use worldwide. If methane can be produced from hydrate, a reasonable assumption given that there are no obvious technical or engineering roadblocks to commercial production, the nation's natural gas supply could be extended for many years to come. This report reviews the U.S. Department of Energy (DOE) Methane Hydrate Research and Development Program, the project selection process, and projects funded to date. Key recommendations include focusing research on seven high-priority areas; ensuring scientific oversight in the selection, initiation, monitoring, and assessment of major projects funded by DOE; funding of fellowships by DOE to increase its contribution to education and training; and providing project applicants with instructions and guidelines that outline the requirements for timely, full disclosure of project results and spell out the consequences of noncompliance. Paper, $43.00.

Climate Data Records from Environmental Satellites: Interim Report. This report identifies the necessary elements in a program to create high-quality climate data from satellites. The study includes a review of past attempts to create climate data records; suggests steps for generating, re-analyzing, and storing satellite climate data; and points out the importance of partnering among agencies, academia, and industry. The National Oceanic and Atmospheric Administration will use this report—the first in a two-part study—to draft an implementation plan for creating climate data records. Paper, $27.25.

Computer Science: Reflections on the Field, Reflections from the Field. This report describes the intellectual character of computer science (CS) in accessible form and through examples. The object of the report is to prepare readers for what the future might hold and to inspire CS researchers to help create it. Rather than provide a comprehensive list of research topics or a taxonomy of research areas, the committee describes key ideas at the core of CS, but does not define boundaries. This volume also includes some two dozen essays on particular aspects of CS research and their results. Paper, $35.00.

Emerging Technologies and Ethical Issues in Engineering: Papers from a Workshop, October 14–15, 2003. A group of distinguished engineers and ethicists gathered for an NAE workshop to discuss the responsible development and uses of emerging technologies. Presentations were focused on four areas of engineering—sustainability, nanotechnology, neurotechnology, and energy—and the ethical issues engineers and society as whole are likely to face. Several approaches to tackling ethical issues were discussed, including: analyzing the factual, conceptual, application, and moral aspects of an issue; assessing the risks and responsibilities of a particular course of action; and considering theories of ethics or codes of ethics developed by engineering societies.

Ethics can be integrated into engineering education for students and professionals, either as an aspect of courses in the engineering curriculum or as components of engineering projects to be examined along with research findings. Engineering practice workshops can also be effective for confronting ethical issues, particularly when the participants include experienced engineers and young engineers, or even engineering students. Because of the enormous impact of engineering on the environment, health care, and quality of life, the explicit consideration of ethics should be an integral part of engineering education, as well as engineering practice. Paper, $36.00.

The Hydrogen Economy: Opportunities, Costs, Barriers, and R&D Needs. The announcement of a hydrogen fuel initiative in the president's 2003 State of the Union speech stimulated interest in the potential of hydrogen in the nation's long-term energy future. Prior to that time, the U.S. Department of Energy (DOE) had asked the National Research Council to conduct a study of the technical issues involved in the hydrogen economy to assist in the development of its hydrogen R&D program. This assessment includes the current status of technology; cost estimates; the status of CO_2 emissions; distribution, storage, and end use; and the DOE RD&D program. The report assesses hydrogen as a fuel for the future and identifies challenges that must be overcome to realize that future. Topics covered include hydrogen end-use technologies, transportation, hydrogen-production technologies, and transitional issues for hydrogen-fueled vehicles. Paper, $32.00.

Implementing Climate and Global Change Research: A Review of the Final U.S. Climate Change Science Program Strategic Plan. The report reviews a draft strategic plan from the U.S. Climate Change Science Program, which was established in 2002 to coordinate U.S. research on climate change and global change. The U.S. Climate Change Science Program combines the decade-old Global Change Research Program and the new Climate Change Research Initiative. The primary goal of the program is to "measurably improve the integration of scientific knowledge, including measures of uncertainty, into effective decision support systems and resources." Paper, $30.00.

Indicators for Waterborne Pathogens. With recent and forecasted advances in microbiology, molecular biology, and analytical chemistry, a reassessment is in order of the current practice of relying predominantly or exclusively on traditional bacterial indicators to detect and identify all types of waterborne pathogens. This report concludes that indicator approaches will be required for the foreseeable future because monitoring for the complete spectrum of microorganisms in water is not practical or even feasible and because many known pathogens are difficult to detect directly and reliably in water samples. This comprehensive report recommends the development and use of a "tool box" approach by the Environmental Protection Agency and others for assessing microbial water quality; this would mean available indicator organisms (and/or pathogens in some cases) and detection method(s) could be matched to the requirements of a particular situation. The report also recommends that a phased, three-level monitoring framework be developed to support the selection of indicators and indicator approaches. Paper, $54.00.

Managing the Columbia River: Instream Flows, Water Withdrawals, and Salmon Survival. Flows of the Columbia River, although modified substantially during the twentieth century, still vary considerably from season to season and year to year. The smallest flows tend to occur during the summer when the demand for irrigation water is highest and when water temperatures are highest. These periods of low flows, high demand, and high temperatures are critical for juvenile salmon migrating downstream through the Columbia River hydropower system. Although the impact on salmon of individual water withdrawals may be small, the cumulative effects of numerous withdrawals affect Columbia River flows and thus increase risks to salmon. The report concludes that, although our knowledge of the migratory behavior and physiology of salmon is imperfect, decision makers should acknowledge this and be willing to take action in the face of uncertainties. The report recommends that the state of Washington, Canada, other basin states, and tribal groups establish a basin-wide forum to consider future water withdrawal application permits. If the state of Washington issues additional permits for water withdrawals, those permits should include provisions for curtailing withdrawals in critical high-demand periods. Hardcover, $35.00.

Meeting the Energy Needs of Future Warriors. The central feature of combat soldiers' equipment is increasingly sophisticated sensing, communications, and related electronics for use on the battlefield. The most critical factor for their use is the development of power supply systems capable of operating those electronics effectively for up to 72 hours. The Army requested that the National Research Council review the state of the art in power systems and recommend technologies that could support the rapid development of power systems for future soldier systems. This report includes assessments of technological options for meeting different power-level requirements, power-system designs, and soldier-energy sinks. Future design concepts, particularly low-power systems, are also described. The recommendations are focused on technology development and system design. Paper, $23.00.

Naval Forces' Defense Capabilities against Chemical and Biological Warfare Threats. U.S. naval forces must be prepared to respond to a broad array of threats, including chemical and biological warfare (CW and BW) threats. To help review its current state of preparedness, the Chief of Naval Operations asked the National Research Council to assess the U.S. Navy's defense capabilities against CW and BW threats and the development of methods of sensing and analyzing chemical and biological agents, withstanding or avoiding exposure to such agents, and dealing with contamination under a broad range of operational conditions. This report provides an overview of potential threats and an evaluation of the Navy's operations, nonmedical programs, and medical countermeasures designed to confront CW and BW threats. The report also includes general and specific findings and recommendations based on these assessments. Paper, $28.00.

A Patent System for the 21st Century. This report focuses on how well the patent system encourages research, innovation, and the dissemination of knowledge and how well it is being adapted to rapid technological and economic changes. The panel concludes that the system has shown admirable flexibility in accommodating new technologies and reflecting the importance of intangible capital of all sorts. Nevertheless, there is reason to be concerned about the quality of the patents issued (whether they meet statutory standards of novelty, utility, nonobviousness, and adequate written description), the resources available to the U.S. Patent and Trademark Office to keep pace with changes and the volume of applications, features of U.S. law that limit the dissemination of information in patents and that raise the cost and uncertainty of litigation over patent validity and infringement, access to patented research technologies for basic noncommercial research, and redundancies and inconsistencies among national patent systems that raise the cost of global intellectual property protection. Paper, $35.00.

Preparing Chemists and Chemical Engineers for a Globally Oriented Workforce: A Workshop Report to the Chemical Sciences Roundtable. Globalization—the flow of people, goods, services, capital, and technology across international borders—significantly affects the chemistry and chemical engineering professions. Chemical companies are looking for new ideas, a trained workforce, and new market opportunities, regardless of geographic location. During an October 2003 workshop, leaders in chemistry and chemical engineering from industry, academia, government, and private funding organizations came together to explore the implications of the global research environment for the chemistry and chemical engineering workforce. The workshop presentations identified deficiencies in the current U.S. educational system and pointed out the need for the United States to create and sustain a globally aware workforce in the near future. The goal of the workshop was to inform the Chemical Sciences Roundtable, a science-oriented, apolitical forum for leaders in the chemical sciences to discuss chemically related issues affecting government, industry, and universities. Paper, $18.00.

Retooling Manufacturing: Bridging Design, Materials, and Production. As the U.S. Department of Defense (DOD) continues to pursue the development of the future warrior system, the difficulty of moving complex technologies rapidly from design to manufacturing is becoming a major concern. Communication gaps between designers and manufacturers have hindered the rapid development of new products for future military operations. DOD asked the National Research Council to develop a framework for "bridging" these gaps through data management, modeling, and simulation. This report provides a framework for virtual design and manufacturing and an assessment of necessary tools; an analysis of the economic dimensions of the problem; a discussion of barriers to virtual design and manufacturing in the DOD acquisition process; recommendations for addressing these issues; and a description of research needs. Paper, $28.50.

River Basins and Coastal Systems Planning within the U.S. Army Corps of Engineers. The emphasis of U.S. Army Corps of Engineers (USACE) water resources projects is changing. Traditional civil works construction is giving way to maintenance, operational, habitat restoration and rehabilitation, and recreational projects. Planning water projects requires an integrated systems approach capable of balancing all relevant issues and identifying unintended consequences or cumulative effects.

Congress requested that the National Research Council review USACE's planning, design, operation, and evaluation processes in the context of the nation's river basins and coastal systems. This report provides detailed recommendations for improvements in planning and management practices for river basins and coastal systems, including, but not limited to: modification and updating of planning guidance documents and the incorporation of new requirements into USACE's project process and guidelines (e.g., mandatory federal cost sharing or required assessments for stating how new projects will fulfill USACE's commitment to environmental stewardship). Paper, $39.00.

A Vision for the International Polar Year 2007–2008. In recognition of the International Polar Year (IPY), the United States plans to present an overview of potential science themes, enabling new technologies, and public outreach programs for individuals and scientific communities. The committee recommends that the U.S. scientific community and other agencies use the IPY to initiate environmental change and variability in polar regions; explore scientific frontiers, ranging from the molecular to the planetary scale; and engage the public (and ultimately advance general scientific literacy) by providing information about polar regions in the global system. Paper, $28.50.

The BRIDGE

(USPS 551-240)

NATIONAL ACADEMY OF ENGINEERING
2101 Constitution Avenue, NW
Washington, DC 20418

Periodicals
Postage
Paid