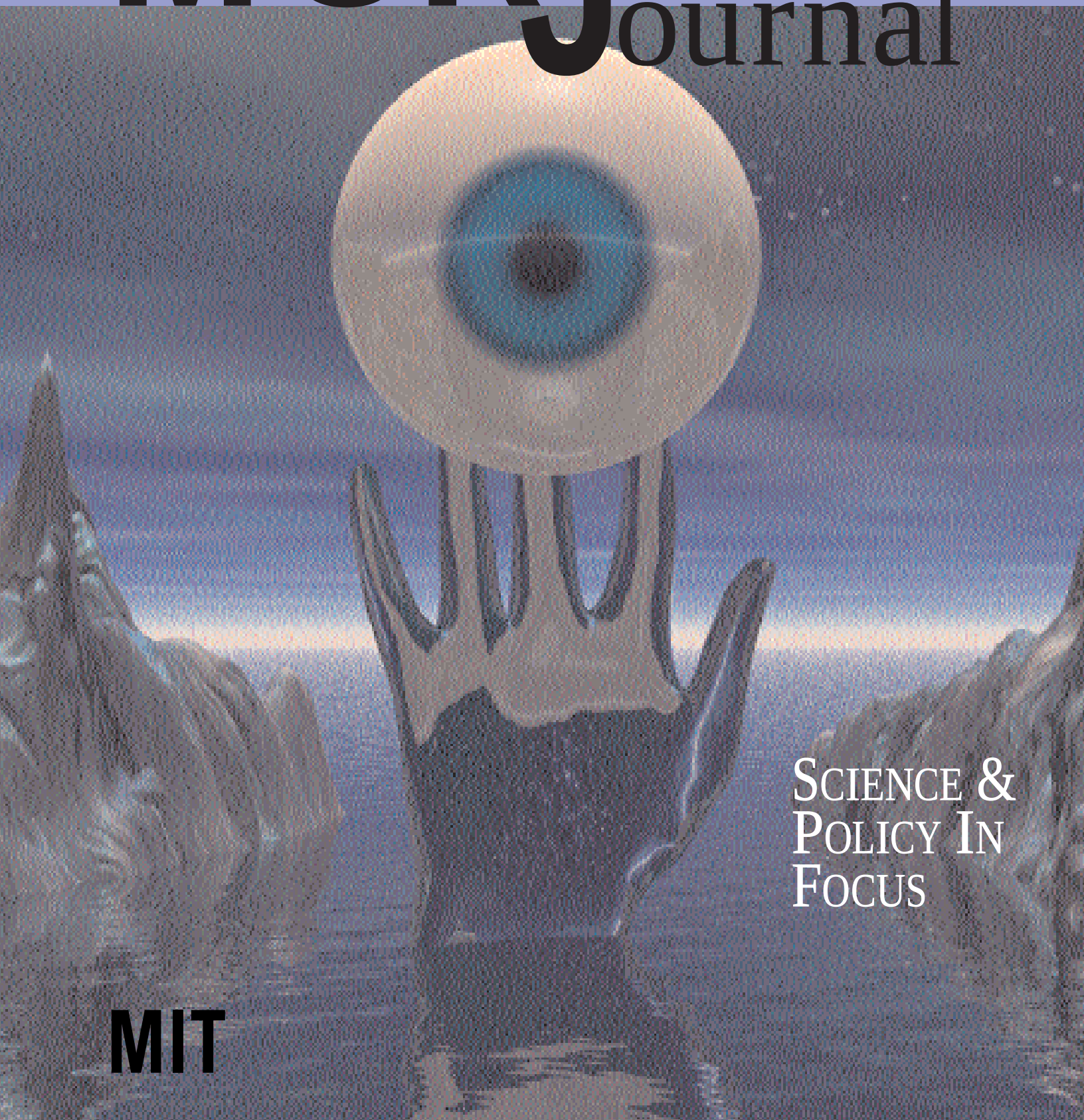
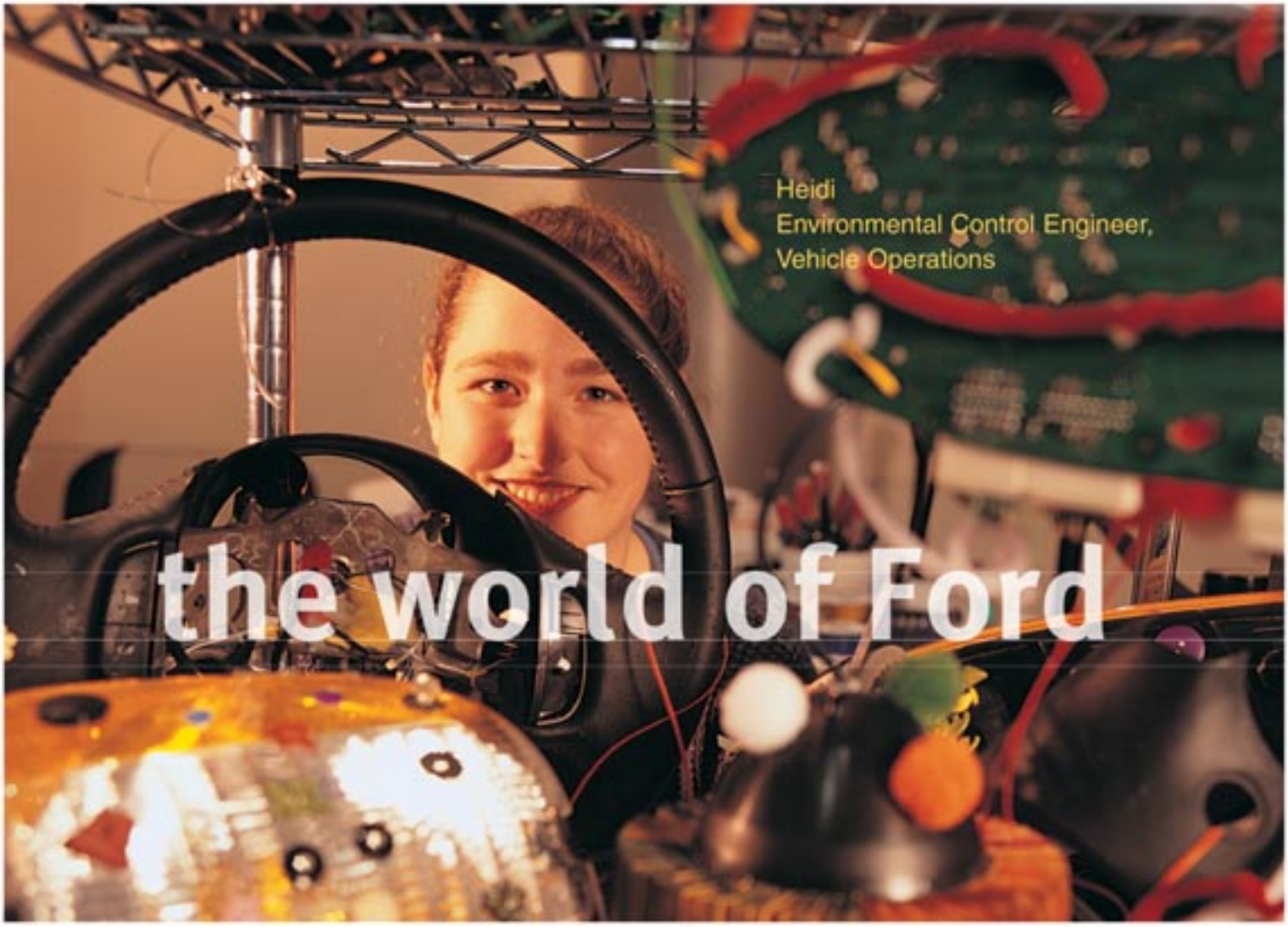
MURJ Journal

Massachusetts Institute of Technology Undergraduate Research Journal



Science & Policy In Focus

MIT



At Ford Motor Company, protecting the environment is one of our most important initiatives. That's why Heidi is so important to our future.

As the daughter of an engineer, Heidi grew up in the Ford family. From day one, she knew about our commitment to automobiles. She didn't know that our commitment to the environment was just as impressive.

The fact is, Heidi's environmental goals and ours are one in the same.

Through the Ford College Internship Program, Heidi had the opportunity to experience our environmental protection efforts first-hand. From closing underground storage tanks to taking our waste minimization efforts to new levels, her ideas made a real impact.

As it turns out, her internship was just the beginning.

Since joining the company full-time in 1998, Heidi has coordinated important environmental projects with a team of engineers - including the restoration of local waterways. Today she serves as the environmental contact for seven assembly plants and four other facilities located in the U.S., Canada, and Mexico.

Thanks to people like Heidi, we're working smarter and safer than ever before. Where will your ideas take us? Visit us online and see for yourself. Go to mycareer.ford.com. By choice, we are an Equal Opportunity Employer committed to a culturally diverse workforce.

Ford Motor Company

VOLVO mazda LINCOLN Mercury JAGUAR

MIT

UNDERGRADUATE RESEARCH JOURNAL
MIT Undergraduate Research Journal

Issue 3, Fall 2000

Undergraduate Research Journal,
77 Massachusetts Avenue,
Cambridge, MA 02139

00-07-0543

Welcome to the fall 2000 issue of the MIT Undergraduate Research Journal (MURJ). In this edition of our journal, we present to you a series of science and policy papers discussing often-controversial issues arising in academia. We have, as always, included along with our informational essays a series of research reports from students investigating questions from fields as distinct as anthropology and physics.

Those undergraduates who are familiar with MURJ will know that this issue's focus on science and policy reflects our wider goal to present national and international topics of concern to the MIT audience, allowing students to read about and discuss important ethical issues arising in science. In this issue, we explore topics as varied as Internet security and global warming, and urge students to read essays that affect more than just their primary field of study.

Given the specialization of many students, however, we have also chosen to continue presenting original research in each edition of our journal; what is included in our Reports section, however, has been written in a style comfortable to all undergraduates, unlike conventional research reports. As always, MURJ editors have attempted to orient the content of this journal toward the common undergraduate reader, allowing undergraduates studying nearly any discipline to read the research and analysis presented in our publication.

While introducing this third issue of MURJ, it is also essential that we give credit not only to broad national and international topics we discuss in this edition, but also to those that strike close to home. Science and public policy intersect right here in Cambridge, and in recognition of this fact, MURJ editors will be preparing a series of talks and conferences over the next few months to discuss science policy as it moves across the Cambridge-World axis. These talks and conferences will be announced to the entire MIT community, and all members of the public are welcome to participate.

Of course, this effort would not be complete without the assistance of many individuals who have guided us and have made MURJ a sustainable production. Thanks to our advisor, Dean Les Perelman, for his continued support, and to those professors who have spent their time reviewing our research submissions.

After reading this issue of MURJ, we encourage you to consider joining our group or perhaps submitting your own research to our next edition, which will appear on newsstands at the end of the spring term.

Sanjay Basu,
Editor-in-Chief

Solutions for Genomic Research

Incyte Genomics is a leading provider of an integrated platform of genomic technologies designed to aid in the understanding of the molecular basis of disease. Our platform includes database products, genomic data management software, microarray-based gene expression services, and related reagents and services. These are information management tools critical to the pharmaceutical and biotechnology industries for drug discovery and development. The employees of Incyte provide the tremendous energy, talent and expertise that will help us revolutionize the world of healthcare. We're looking for individuals who share our vision and want to play a role in defining a new industry as we build our global organization.

- **Software Development**
- **Information Technology**
- **Bioinformatics**
- **Production/Operations**
- **Research**
- **Science**
- **Engineering**
- **Technology Development**
- **Finance**
- **Administration**
- **Human Resources**
- **Legal**
- **Marketing**
- **Sales**

www.incyte.com

Incyte offers competitive salaries, an outstanding benefits package and significant opportunities for professional growth. Send your resume to: Incyte Genomics, Inc., Human Resources, fax 650-845-4176; email murj@incyte.com, or mail to 3160 Porter Dr., Palo Alto, CA 94304. Incyte recognizes the talent of its diverse workforce. EOE F/M/V .

IncyteGenomics
For Life.

Table of Contents

News

Genetics in Focus

Technology in Focus

Reports

Environment in Focus

Health Care in Focus

MIT Science News in Review

BIOCHEMICAL SCIENCES

Whitehead triumphs over human genome project

On June 26th, the Human Genome Project and Celera Genomics announced the primary assemblies of the human genome, a feat accomplished with the help of the Whitehead/MIT Center for Genome Research. Whitehead was the single largest contributor to the project, accounting for nearly a third of the sequences assembled by an international consortium of sixteen laboratories.

Researchers at the Whitehead Institute were responsible for laying the groundwork for the project and were also crucial in completing its final phase, producing over one billion DNA base pairs in the past year.

The consortium's goal for the spring of 2000 was to produce a "working draft" version of the human genome, an assembly containing overlapping fragments of roughly ninety percent of the genome, with some gaps and ambiguities. The ultimate goal is to produce a "finished" sequence with no gaps and 99.99 percent accuracy.

The international Human Genome Project consortium centers have produced far more sequence data than expected (approximately 22.1 billion bases of raw sequence data), resulting in a working draft substantially closer to the ultimate finished form than expected at this stage.

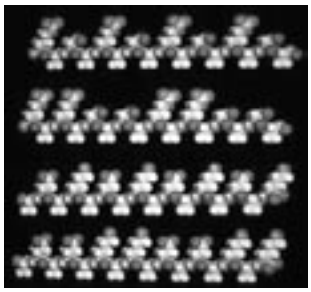
This "finished" sequence will be crucial in obtaining all possible information from human sequence data. As of now, this has already been achieved for about 24 percent of the genome, including chromosomes 21 and 22.

In a similar announcement, Celera Genomics announced the completion of its first assembly of the human genome. While both the private and public projects used similar sequencing technology, the way in which each approached sequencing the human genome is quite different. The Celera project uses a "whole genome shotgun" approach, in which the entire genome is shredded into small fragments that are subsequently put back together on the basis of sequence overlaps.

Whitehead scientists received accolades from President Clinton and Francis Collins, director of the Human Genome Research Institute, and tremendous media coverage. Eric Lander, professor of biology and director of the Whitehead Institute, and Lauren Linton, co-director of its sequencing center, were both present at the White House and recognized for their accomplishments.

"We deserve to be proud of our accomplishments and bask in this glory as the world's attention focuses on us. The credit goes to all the individuals at the Whitehead Genome Center who have worked hard to make us the flagship center of the Human Genome Project Consortium. Everyone associated with this project should feel proud," said Professor Lander.

Unique biomaterial may revolutionize neural regeneration



A long-sought goal of molecular neuroscience has been to perform the seemingly impossible: to regenerate nerve cells after damage. Since the accident of actor Christopher Reeve, the concept has gained widespread attention in the biology community, but few scientists have presented definitive results to suggest that the feat could be easily accomplished.

In the June 6 issue of the *Proceedings of the National Academy of Sciences*, MIT researchers collaborating with New York University scientists reported that they have designed a new material that may function as a scaffolding for regenerating nerve cells. The scaffold, made of proteins, acts as a fence to a vine, allowing neurons to grow fibers along it and ultimately reconstructing functional synapses.

The biomaterial brings researchers one step closer to the ultimate goal of growing replacement nerve cells for victims of spinal cord injuries and other nerve disorders

"Further development of biological materials and related cell-based therapies may bring us closer to the elusive goal of repairing the damaged nervous system," wrote Melitta Schachner, a researcher in molecular neurobiology at the University of Hamburg in Germany.

"The reason this material is so interesting and unique is that we can individually tailor it to grow virtually every type of cell in the body," said Shuguang Zhang, associate director of the Center for Biomedical Engineering at MIT. Zhang and NYU professor Todd Holmes are the first to use a peptide-based biomaterial at the molecular level.

Zhang and Holmes reported that the type of technology they used to make the biomaterial "will become increasingly important in developing approaches for a wide range of innovative medical technologies, including controlled drug release, new scaffolds for cell-based therapies, tissue engineering and biomineralization."

The duo also says that peptide-based materials might be used as an alternative to animal-derived substances that could potentially pass viruses onto cells. The peptide materials can be constructed without using animal cells, and can thus provide a safe means to treat patients without eliciting an immune response.

Genetic culprit infecting animals apprehended in plant cousins

In the March 31 issue of *Science,* researchers from MIT and the Louisiana State University Health Sciences Center report finding surprising similarities between two bacteria with widely differing lifestyles in how they establish chronic infections in mammals and plants.

Upon first glance, agriculturally important bacterium *Rhizobium meliloti* and pathogen *Brucella abortus* have no features in common. However, Graham Walker, professor of Biology, reports that both bacteria interact with their hosts similarly.

While all organisms require nitrogen, plants cannot utilize it until it has been converted into ammonia. This conversion is accomplished by leguminous plants with the help of *rhizobia*, bacteria residing in the soil. The bacteria invade the nodules on the roots of the plants and are then taken up by the plants. "You can look at it as a chronic infection of the plant, although unlike other infections, this is beneficial to the organism," said Walker. "The plant allows itself to be invaded by the bacteria. In this way the plant gets ammonia and the bacteria gets carbon."

On the other hand, *Brucella abortus* infects mammals by being ingested in the form of infected milk or animal tissue, resulting in brucellosis, a debilitating disease accompanied by symptoms that include fever, malaise, and weight loss. The disease is difficult to eradicate because the bacteria reside inside cells.

According to MIT postdoctoral fellow Kristin Levier, the bacterial gene bacA is crucial to both the *rhizobia* and the mammalian pathogen *Brucella abortus.* "The bacA function seems to be necessary for the bacteria to overcome the defense response of the host that would otherwise prevent a chronic infection," said Walker, as removal of the bacA gene made it impossible for the bacteria to infect its host.

Thus, this scenario, which involves removing the bacA gene from the bacteria to allow the human immune system to recognize and kill the intruder, is a perfect starting point for the development of a vaccine against *brucellosis.*

EyeGen wins $50K for the dye to die for

Although MIT's $50K Entrepreneurship Competition usually features a new marketing marvel, this year's competition has a biological base: a dye that makes DNA visible to the naked eye. The winners, members of a company named EyeGen, formulated a safe alternative to radioactive and fluorescent probes, which researchers commonly use to visualize DNA strands. The EyeGen group, which plans a $15 billion market for their product, may be revolutionizing bench-top biology with a simple dye called EyeGenRad.

"Since DNA manipulation is at the core of most genomic, medical and biotech research, EyeGenRad has broad application

as an invaluable tool," the team wrote in its official press release.

"Now we'll see what happens with the funding, find some lab space and start our operations," said chemistry graduate student Zoran Zdraveski, the team's scientific officer and one of the founders. "We planned on moving this company ahead, but the $50K shoved us along at a faster rate. We met so many people who've opened doors for us."

MIT applies 'Star Wars' technology to breast cancer

Technology developed 10 years ago to detect missiles for 'Star Wars' is now being applied to the treatment of breast cancer. Dr. Alan J. Fenn, senior staff member in the Advanced Electromagnetic Systems Group at Lincoln Laboratory, invented the technique.

The treatment uses focused microwave radiation to "heat — and kill—cells containing high amounts of water," Fenn said. Cancer cells have a high water content—around 80 percent—while healthy breast tissue contains much less.

The microwave treatment lasts for 20 minutes and heats cancer cells to 115 degrees Fahrenheit. Tumors typically shrink by 50 percent or more within a week. However, as part of this FDA-approved study, the patients had mastectomies 7 to 10 days after the thermotherapy treatment, making it impossible to watch the tumor shrink away completely.

In future clinical trials the doctors may apply a second treatment to eliminate the cancer faster.

"Our goal is to destroy all visible and microscopic cancer cells and pre-cancerous cells in the breast," Fenn said.

That would eliminate the need for surgery and conventional radiation treatments. If the cancer spreads, however, the patient would still need chemotherapy.

The procedure uses two tiny probes to sense and measure parameters during treatment. "Patients treated so far have gone home with only one or two band-aids," Fenn said.

He expects that in the future the procedure will be noninvasive. The technique could also be applied to other cancers. The next target will be prostate cancer.

Neural receptor affects breathing and memory

MIT researchers reported in May 1 of *Neuroscience* that the chemical receptor N-methyl-D-aspartate (NMDA), usually associated with memory and learning, might also play a role in the development of the fetal respiratory system.

Researchers found that the absence of this receptor led to long-term synaptic depression (LTD), an activity that aids in learning and memory building. Unexpectedly, the increased LTD occurred in the brainstem, an area normally associated with autonomic vital functions.

Further research on NMDA led to the discovery that a lack of NMDA receptors during prenatal development was lethal in lab rats.

"This is the first indication that prenatal development of specific regions in the brain controlling vital functions is very much dependent on NMDA receptor activity, and a lack of NMDA receptor activity in the fetus could affect newborns' breathing after birth," said Dr. Chi-Sang Poon, a Principal Research Scientist in the Harvard-MIT Division of Health Sciences and Technology (HST). In addition, the unexpected increase in LTD in the brainstems of the mutant mice "shows that learning and memory at a subconscious level could profoundly influence our vital functions," Poon said.

The research suggests that pregnant women should not have prolonged exposure to substances that affect NMDA's activities. Alcohol, PCP, and some anesthetics and analgesic drugs are among the list of substances that interfere with NMDA's normal functionality.

Miller models memory one neuron at a time

Among the greatest challenges in cognitive science is the attempt to map learning and memory. While neuroscientists have attempted the feat from nearly every angle, few claim to have accomplished the task. The goal, however, may be

one step closer to completion because of the work of Earl Miller, associate professor of brain and cognitive sciences in MIT's Center for Learning and Memory. Miller attempts to determine how our brains guide us through long-term projects, despite the everyday distractions and disappointments we face.

By studying the prefrontal cortex of the brain (crucial to high-level cognitive function), that is enormous in humans and appears to arbitrate voluntary activity, Miller thinks he might achieve the feat.

"The main thing we are interested in is the deepest issue in cognitive science, called executive or cognitive control," Miller said. "It governs how you decide what behaviors to engage in, how you make decisions, how you decide what to pay attention to, what to do with your life."

Miller, believe it or not, has begun to model some of these profound behaviors.

"When I started graduate school, if you had told somebody, 'I'm going to tell you a little bit about how voluntary, goal-directed behavior works,' he or she would have said, 'You can't study that. That's a crazy topic.' But just in the past few years, we've actually developed some real hypotheses about how this all works."

Miller began by looking at single neurons, "from the premise that the most important feature for neurons is their ability to extract information from experience…They need to be sculpted by experience, because virtually all voluntary behavior—everything we're saying and doing right now—is learned."

Miller has recorded the electrical activity of groups of neurons in rhesus monkeys to determine the capacity of the prefrontal cortex to assist the monkeys' ability to direct their attention to a visual image and putting it into long-term memory.

Miller has begun to identify which neural groups of neurons contribute to this activity; opening the door to potential pharmaceutical therapies that can prevent the loss of memory.

Physical Sciences

Ion-collision experiment uncovers Big Bang secrets

Most of us remember learning about Rutherford's nucleus scattering experiment. Recently, MIT scientists again collided gold nuclei—this time, for an altogether different purpose: to study the state of matter immediately after the Big Bang.

On June 12, physicists at the Brookhaven National Laboratory created a head-on collision between gold nuclei travelling near the speed of light in what they called the Relativistic Heavy Ion Collider (RHIC).

RHIC is a system used to model the early universe as it existed less than a second after the Big Bang. By accelerating gold and other nuclei to near the speed of light, Brookhaven scientists believe they can visualize energy densities associated with the Big Bang, perhaps uncovering clues to the plasma thought to exist during the event. That plasma consisted primarily of high-energy

quarks and gluons, which quickly cooled down to create the distinct particles that we now identify as composing matter.

By reversing the process through high-speed collisions, RHIC scientists attempted to fracture nuclei to their previous hot and dense states. When they finally observed these nuclei in June, they confirmed that plasma exists. The work may also determine the timing of such plasma formation during the Bang.

"We are excited, relieved and slightly tired," said Wit Busza, the Francis Friedman Professor of Physics and a MacVicar Faculty Fellow. Busza commented that "it is very important that we identify that moment and that we show that we have compressed normal matter so much that it produced the plasma."

"With RHIC, we are making an analogy to what happened just after the Big Bang. We are replaying history to study the process of transition similar to a phase transition that occurs when water vapor turns into fog," said Busza. "And it's a relief to know the equipment works to do that. But we still had to show that the events occurred and analyze what happened."

Rocket team fires up

Inspired by a $250,000 cash prize to launch a suborbital rocket 120 miles into the atmosphere, MIT undergraduates, graduates, alumni and faculty collaborated to create an elegantly simple rocket. The Cheap Access to Space (CATS) competition, funded by the Foundation for International Non-governmental Development of Space and the Space Frontier Foundation, is designed to challenge the current suborbital record height of approximately 25 miles.

Carl Dietrich (SB 1999) founded the Rocket Team in his senior year after hearing about the contest from a friend and has applied for a patent for a kerosene/liquid oxygen engine he designed in 1998. He worked full-time on the project with Westinghouse Science Award winner Andrew Heifitz (SB 1991), who was intrigued by the challenge of working with unmanned aerial robotics and impressed by the simplicity of the design.

The rocket was 14 feet tall and eight inches in diameter, a size dictated by the facilities available to construct and test vehicle components. It weighed approximately 200 pounds. The body consisted of a lightweight carbon fiber composite, while the engine was made up of aluminum, stainless steel and a silica-pheno composite material. As such, the rocket was significantly smaller, simpler, and cheaper than existing government and commercial liquid-fueled launch vehicles.

A prototype of the model was launched on Briggs Field this past spring. Team members have already met with NASA engineers and administrators at the launch/test facility at Wallops Island, VA, along with members of the Virginia Spaceport Authority. The design and safety review process with NASA will be initiated shortly to obtain a launch license from the Federal Aviation Administration and approval to launch at Wallops Island. If successful, this rocket could introduce cheaper, simpler commercial space launch vehicles to the skies.

Nuclear waste containers convert to clay given time

As the use of nuclear fuel skyrockets around the world, a team of MIT scientists led by Civil and Environmental Engineering Professor Franz-Josef Ulm has been asking a simple question: where is the waste going to go? Ulm and others have raised concerns about the ability of cement containers to hold nuclear waste over the long term. But rather than watch bad environmental policy take shape, Ulm and his colleagues took action; they designed a test to observe the degradation of nuclear waste containers, which commonly occurs over 300 years, in one day. Using a laboratory model based on miniaturized concrete slabs surrounded by a liquid bath, Ulm and his colleagues began to subject weathered materials to a variety of stresses to discover how the containers could be improved.

"Our goal is to go back to real-life structures, monitor the environment around them, and predict by model-based simulation what the concrete will do over extended periods of time," Professor Ulm said. And if a bad scenario plays-out, "we'll be able to predict its eventual effect, and intervene in time to slow down or reverse the aging," said Ulm.

Ulm is now attempting to export his model to those who need it most: the US government.

"When, and if, spent nuclear fuel from the US is buried in the Department of Energy proposed repository at Yucca Mountain, it will be placed in concrete casks that are supposed to maintain integrity at least 300 to 1,000 years," said Mujid Kazimi, an MIT professor of nuclear engineering. But what Ulm and others are finding is that the concrete containers can weaken over time, causing the concrete to dissolve. When he applied a shear stress to his laboratory containers, slivers of the material slipped apart.

"So a material that is originally very strong ultimately ages to one that behaves like a weak low-friction soil, such as clay," Ulm concluded. The findings may lead to a renewed emphasis to discover how the waste containers can be improved.

MIT scientists provide new spin on unique pulsars

Like a beacon in the night sky, a star spins, peacefully pulsing with x-rays until a sudden disturbance causes it to speed up its rotation.

On March 23 at the Rossi 2000 conference held at NASA Goddard Space Flight Center, assistant professor of physics Victoria M. Kaspi of MIT's Center for Space Research reported that an Anomalous X-ray Pulsar (AXP) experienced a seismic shift reminiscent of those found commonly in neutron stars. This confirmed that the AXP is truly a neutron star.

"We have thought that earthquake-like events might occur on these stars for some time now, but we have lacked the right

instrument to check this," said Kaspi. "No X-ray astronomy satellite in the past had the agility to observe these objects as often and as regularly as we needed. Thanks to [the Rossi X-ray Timing Explorer satellite], we now know for certain that glitches occur in AXPs, and we can study the interiors of these unusual objects using a form of 'seismology,' like the way geologists study the Earth from earthquakes."

Neutron stars are the remains of massive stars that explosively shed their outer shell upon exhausting their nuclear fuel. In the wake of these violent explosions, the solar core collapses into a spinning sphere nearly 10 miles in diameter. If the neutron star "pulses" with radiation as it spins, it is called a pulsar. Most pulsars emit both radio waves and X-rays. A select few, AXPs, emit X-rays exclusively.

Studying the relation between conventional radio pulsars and AXPs, Kaspi and collaborators searched for "glitches" in AXPs, when radio pulsars suddenly spin faster than normally observed, a likely result of earthquake-like phenomenon called vortex line upspinning deep within the stellar core.

After two years of observing, Kaspi and collaborators watched as one of their target AXPs, 1RXS J1708-4009, began to spin faster.

"We were delighted! The glitch we saw was practically identical to glitches that have been seen in the Vela radio pulsar," said Kaspi. "This is clear evidence that the AXP is a neutron star with an internal structure just like the radio pulsars."

Such discoveries lend support to the magnetar hypothesis, which predicts the existence of stars a thousand times more magnetic than the overpowering neutron star. Kaspi hopes that defining the nature of AXPs will aid the understanding of other rare and mysterious objects such as the magnetars and soft gamma-ray repeaters.

ENGINEERING AND TECHNOLOGY

Discover Magazine awards technological innovation

Three MIT innovations are among this year's 19 *Discover Magazine* Awards for Technological Innovation. The three inventions: sneakers that compose music, a cancer diagnosis technique, and lights that talk, were all featured in the July 19 edition of *Discover Magazine.*

Joseph Paradiso, technology director of the Things That Think group at the Media Laboratory, won a prize for sneakers he developed that let children create music as they move. Heralded as "a new twist in interactive entertainment," Paradiso's invention superseded all others in the entertainment category.

Dr. Todd Golub of the Whitehead Institute developed a DNA Chip Cancer Diagnosis technique. Doctors can now deposit RNA taken from a tumor biopsy onto probes affixed to small glass wafers. When scanned with a laser, a DNA profile of the tissue is automatically created, allowing physicians to customize cancer treatments for individual patients.

Customizing is also at the heart of Professor Steven Leeb's invention. The Associate Professor from the Department of Electrical Engineering and Computer Science produced "talking lights" that convert light sockets into information transmitters. Leeb says that the invention can allow a person to encode information in a sort of Morse code by modifying the rate at which a light flickers. Flickering at a rate faster than the eye can see, the lights trigger a speech mechanism that could, for example, direct a person to the bathroom in a building.

Leeb and his two fellow award winners were acknowledged at a June ceremony in Disney's Epcot Center. Their awards are given to "honor the unsung technological heroes whose creative genius improves the quality of our everyday lives and alerts us to what lies ahead in the frontiers of human achievement and ingenuity".

Manus robot rehabilitates stroke victims

An MIT robot has proven itself valuable in the physical therapy of recovering stroke victims. The May 23 issue of Neurology featured a report in which a second clinical study demonstrated that the robot significantly improved patients' recovery of arm movement.

The robot, dubbed Manus for MIT's motto, 'mens et manus' (mind and hand), has been in development for eleven years now and focuses on arm therapy. Two clinical trials have already been conducted. A third was initiated in January 2000 at the Spaulding Rehabilitation Hospital in Boston; the purpose being to "see if stroke recovery is limited to central nervous system reorganization, and if it can be influenced by more peripheral systems."

During the robot-aided therapy, the patient sits at a table and places the lower arm and wrist into a brace attached to the robot's arm. A video screen then prompts the patient to perform arm exercises. The robot facilitates arm movement by providing adjustable levels of guidance.

The study consisted of 56 patients split among two groups. Each patient received standard therapy. In addition, the experimental group spent an additional daily hour with MIT-Manus, while the control group used a "sham" robot once a week that did not guide patients through exercises. The experimental group demonstrated twice as much of an increase in arm movement as patients using conventional physical therapy.

Researchers are now focusing on creating additional robots for therapy of the legs and wrist/hand. The applications of this robot are far-reaching, as it can become an important teaching tool as well as an aide to doctors in assessing a patient's condition and prescribing optimum therapy.

Surgical disaster sparks innovation

Had the surgery turned out differently, it is very possible that Professor Ernesto Blanco of the Mechanical Engineering department would still have been interested in the surgical instruments. As it turned out, the surgery gone awry guaranteed that Blanco would modify the instrument at fault.

The trocar is used in endoscopic, minimally invasive surgery to make the initial puncture through skin. Once a hole has been made, other instruments can be inserted through the opening to complete the surgery.

In Blanco's surgery, the trocar cut too deeply due to the change in resistance once the skin had been pierced. Because the endoscopic surgery had to be truncated after the slip of the trocar, Blanco needed an open surgery. This resulted in a longer hospital stay, stitches (which incidentally inspired another invention) and an increased risk of infection.

The trocar Blanco has designed should, however, be able to circumvent these problems by addressing the initial difficulty of piercing the skin without injuring organs. The new design is shaped in an hollow X rather than the traditional solid V. The new design always has a protective guard on the blade within the body, a major improvement on trocars that can plunge 3/4 of an inch in the body before the safety guard is deployed.

In addition, the new trocar removes a step in endoscopic surgeries by filling the body cavity with carbon dioxide in order to separate organs from the abdominal walls. A patent has been applied for the device and will take approximately a year to arrive on the market.

Microvillages solve housing shortages in Turkey

600,000 people were left homeless in Turkey after August 17, 1999, when an earthquake registering 7.4 on the Richter Scale leveled many apartments and housing. Many of the victims of the earthquake are still living in tent cities and other temporary lodging.

Hearing of the devastation, Professor Jan Wampler of MIT's architecture department immediately collaborated with MIT alumni in Turkey to devise a solution to the housing shortage. The result was a plan to create microvillages—a plan that will be carried out by Habitat for Humanity, MIT and Mercy Corps International.

The proposed microvillages will be two-story housing units to be built in part by its future residents. The microvillages will center upon communal living and focus upon sustainability. Locally farmed wood as building material, harnessed wind and sun energy as power sources, rainwater for reservoirs, and treatment of sewage as fertilizer are just some of the ideas that will be utilized to create the microvillages. The microvillages will enable people to be active participants in their living unit through community maintenance.

The future of microvillages will be bright. In the past, people who left rural areas had to live in busy cities; microvillages will provide the conveniences of the city while maintaining the community of the more rural areas.

Coaxial cables succeed to perfection through perfect mirrors

In 1998, MIT scientists created a perfect mirror—now they are building on perfection by researching coaxial cables which can shoot light quickly and efficiently over long distances without distortion.

The commonplace metallic mirror reflects light from every angle while absorbing incident light. Dielectric mirrors (unlike metallic mirrors) do not conduct electricity, are made from layers of transparent dielectric materials, and are used in lasers, which require low loss of light and high reflectivity.

The perfect mirrors created in 1998 have the reflective properties of a metallic mirror and the low loss property of the dielectric mirror. An omnidirectional waveguide was constructed from a tube of the perfect mirror. The all-dielectric coaxial cable outperforms the optical fibers currently used in data transmission because it does not lose its light polarization as optical fiber does. The cable can create high-bandwidth and higher information capacity and will be researched for Internet applications.

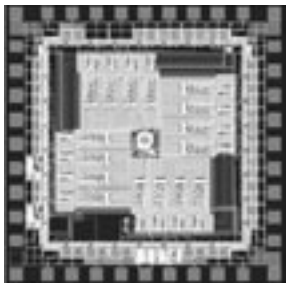
MIT researchers who have worked on the perfect mirror technology launched a new company OmniGuide to explore the coaxial cable's practicality. Currently a prototype and test of the coaxial cable is underway.

Feedback loops enhance electronic circuits to brain circuits

An ultimate challenge of artificial intelligence researchers is to create an electronic circuit that mimics the pathways of the brain. Researchers at MIT and Lucent Technologies' Bell Labs may have accomplished that very feat, as they report in the June 22 issue of *Nature*.

MIT researchers Richard Hahnloser, a postdoctoral fellow in the Department of Brain and Cognitive Sciences; Rahul Sarpeshkar, assistant professor of electrical engineering and computer science; and H. Sebastian Seung, assistant professor of computational neuroscience, say that they have created an electronic circuit that mimics that circuitry of the cerebral cortex, using feedback loops to copy the brain's design.

"Like electronic circuits, the neural circuits of the cortex contain many feedback loops," said Seung. "But neuroscientists have found that cortical feedback seems to operate in a way that is unfamiliar to today's electronic designers. We set out to mimic this novel mode of operation in an unconventional electronic circuit."



The prospects of such a circuit are great for artificial intelligence scientists, who expect to develop computer hardware capable of cognitive tasks. But the physical structure of the circuit itself is quite unique; the MIT group generated it by producing artificial neurons that communicate with each other via artificial synapses made from tran-

sistors grafted onto a silicon integrated circuit. The group also added inhibitory and excitatory feedback loops to their circuit to mimic the activities of the brain.

In the brain, feedback controls repetitive motion, such as the beating of the heart, or more advanced motion, such as movement of the limbs.

"Biologists like to focus on simple linear pathways through this network, ignoring the tangled web of feedback loops, which seem too complex to even contemplate," Seung said. "But it seems unlikely that we could ever understand intelligence or consciousness without understanding the role of feedback in the neural networks of the brain."

Although electrical engineers have been tempted to use common electronic feedback mechanisms to mimic the brain's circuitry, Seung and other have said that the brain uses altogether different rules for feedback. The group's new circuit has the ability to combine both digital and analog-type behaviors to truly imitate the brain's behavior.

Powerful pulsed superconducting magnet passes initial testing

MIT engineers and colleagues from Japan and the US reported that a 150-ton magnet, significant to fusion energy research, has passed its initial operating test.

The magnet, located in Japan, produced a magnetic field 260 times more powerful than that of the Earth, at 13 tesla, with a stored energy of 640 megajoules at a current of 46,000 amperes.

Two modules make up the magnet: one designed and built in the United States, the other in Japan. Dr. Raghavan Jayakumar, a visiting scientist at MIT, spearheaded the US magnet program. The magnet was built by engineers from MIT, Lawrence Livermore National Laboratory (LLNL) and the Lockheed Martin Corporation, the prime industrial contractor. The two modules were combined last year in Naka, Japan.

Joseph V. Minervini, principal researcher at MIT's Plasma Science and Fusion Center (PSFC) and Department of Nuclear Engineering, said, "It's the world's most powerful pulsed superconducting magnet." The researchers seek to demonstrate "superconducting performance parameters" and manufacturing methods for larger magnets planned for the International Thermonuclear Experimental Reactor (ITER), whose goals include demonstrating the feasibility of nuclear fusion as an energy source.

Perhaps the most noteworthy result of the initial testing was the stability demonstrated by the magnet while charging to full capacity. No training or quenching was required, common procedures in the early operation of other superconductors.

According to the researchers, future testing will involve trying to increase the speed at which they can charge up to 13 tesla and discharge.

World Science News in Review

BIOLOGICAL SCIENCES

Genetic basis to alcoholism revealed

In a recent report in the Journal of Neuroscience, Todd E. Thiele of the University of Washington, and colleagues, state that children of alcoholics are less sensitive to the "biochemical, motor, and perceptual changes induced by intoxicating levels of ethanol relative to children without a family history of alcoholism." It was also noted that these children are at an increased risk for developing alcoholism.

New studies by Thiele and colleagues on mice indicate a strong genetic basis to alcoholism. Mice, bred with behavior remarkably close to that of alcoholism and lacking a gene encoding a specific segment of an enzyme called Protein Kinase A, suffered from deficits in biochemical activity involving Protein Kinase A. Protein Kinase A helps convey chemical signals from the environment to the parts of cells responsible for rendering a gene active or inactive.

The missing gene resulted in a number of metabolic and behavioral effects, including the ability to consume more alcohol than normal mice, to drink significantly stronger alcohol solutions, and to recover far more rapidly from alcohol-induced stupor.

Similar studies on flies reflected a genetic basis for alcohol sensitivity as well. Thiele also showed that mice lacking a substance called Neuropeptide Y drink increased amounts of alcohol and can withstand alcohol-induced sedation. Neuropeptide Y, an important brain chemical, appears to be involved in the same biochemical pathway as Protein Kinase A. Thus, gross alterations in this pathway are responsible for alcoholic behavior.

Future research will uncover whether variations in the biochemical behavior of this pathway are responsible for alcoholism in humans.

Key brain receptor controls response to opiates

Neuroscientists have often coveted the infamous "substance P", a molecule in the brain involved in emotional responses as worrisome as depression and anxiety. The molecule, long believed to hold keys to the understanding of pleasure and pain, now appears to play a central role in the brain's reaction to opiates: drugs like morphine and heroin.

As Stephen Hunt and colleagues report in the May 11 edition of Nature, mice bred to lack a substance P receptor known as neurokinin-1 (NK-1) are less responsive that normal mice to pain. These mice also seem to lack the basic behaviors associated with stress. The mice even fail to act aggressively.

Hunt shows, in his most recent paper, that mice lacking NK-1 fail to respond to substance P and do not experience the pleasurable effects of morphine. The results may have implications for human therapy. Emotional problems like stress or anxiety could be alleviated. But even seemingly unrelated syndromes, like eating disorders, have been correlated to stress and may be resolved through inhibition of NK-1 and its numbing effect on stress and pain.

But wait: the elimination of NK-1 is no hedonist's dream-come-true. The processing of substance P, in an area of the brain known as the nucleus accumbens is associated with the motivation factors for eating and other essential behaviors. And pain is—no matter how unpleasant—an important component of a functioning individual's sensory shield against harm.

Nevertheless, Hunt suggests that NK-1 inhibitors could be used to help drug addicts. Mice lacking the receptor failed to experience the physical effects of morphine withdrawal, suggesting that humans addicted to drugs might be treated in a more efficacious manner. If the link between substance P and stress is indeed a strong one, an NK-1 dampening drug could also eliminate the craving for drugs altogether.

Deadly spider toxins may protect plants

Researchers in Australia and America recently suggested that cereal crops should be made as poisonous as one of the deadliest spiders on Earth. Toxins from the Australian funnel web spider, they report, could make "ideal, and environmentally friendly, plant-produced pesticides".

According to Glenn F. King at the University of Connecticut Medical Center, and colleagues, "crops given the genetic instructions to make the spider toxins would only be poisonous to pests".

Since insects are gradually evolving resistance to conventional pesticides, stronger pesticides are definitely needed.

The venom of the Blue Mountain Spider (Hadronyche versuta) contains over 100 toxins. While all of the toxins combined can kill insects and humans alike, individually the toxins are not all poisonous to us. That is, some are specific for killing insects.

Three years ago, King's group in Australia found one such toxin. In a recent issue of Nature Structural Biology, King reported another more potent insect-specific toxin, which sends the nervous system of cockroaches into overdrive, leading to

paralysis and death.

The toxin consists of 36 or 37 amino-acid building blocks, carefully folded in a specific three-dimensional shape. King named the new toxin Janus-faced atracotoxin, as these new proteins feature "two distinct molecular faces".

Since these venom toxins are too expensive to purify or mass-produce, King suggests "that cereal plants could be given the gene to produce the spider pesticides for themselves, making them toxic to the bite of a hungry caterpillar." This will admittedly be a very controversial practice in Europe, says King, as some researchers fear that the venom gene may leak into the environment.

King also suggested splicing the gene into viruses infecting moths and butterflies, delivering the toxin directly to its target. He is currently negotiating with companies to develop the technology further. The selectivity of the virus would make it incredibly safe, he states, although convincing the public of that is another story.

Natural antibiotics diversify protection

Decades of overuse of antibiotics have rendered many strains of bacteria resistant to the most common antibiotic agents used to combat infection. In 1992 alone, over 13,000 people died in US hospitals resulting from infections caused by resistant bacteria.

A team of US chemists reported in the Journal of the American Chemical Society that natural antibiotics could lead to a varied arsenal of anti-bacterial drugs.

Natural antibiotics have a wide range of structures, including macrolides and polyketides. One of the most common polyketides in use today is erythromycin, which contains carbon and oxygen atoms arranged in a ring-like structure. Erythromycin, like penicillin, is produced by a mold and is used to treat a variety of bacterial infections, including bronchitis.

Polyketide antibiotics are unusual in that a single synthase enzyme facilitates successive steps in the synthetic pathway, whereas in most natural antibiotics a different enzyme mediates each step. The polyketide sythases consist of a number of modules, each responsible for adding a few atoms to the polyketide chain.

Chaitan Khosla of Stanford University in California, and colleagues, are redesigning these synthases in the hopes of constructing a more potent group of polyketide antibiotics. By changing the order of steps in the synthetic pathway or altering the starting material, the researchers plan to generate libraries of polyketide molecules that can then be tested for antibiotic activity.

Khosla's group studied the modular enzyme DEBS, responsible for synthesizing the molecule from which erythromycin is ultimately made. They presented lone modules of DEBS each with different starting material. Surprisingly, the modules preferred non-natural substances.

These results indicated "the most important factor driving the evolution of the modules is not that they work as efficiently as possible but that they make the most biologically effective product—the best antibiotic agent."

Thus, redesigned polyketide synthases may be able to operate on non-natural starting materials, but their ability to do so remains difficult to predict on the basis of their normal function in the cell.

Synthetic molecule mimics hormones

The human body is able to regulate itself by a kind of molecular communication known as signal transduction, in which a molecule in a target cell will respond to a chemical messenger by altering its conformation to trigger new chemical processes.

German chemists Ulrich Koert and colleagues at Humbolt University have designed a synthetic messenger molecule that acts according to the same principles as hormones—it grasps a target, changes shape, and passes a chemical signal.

Hormones such as adrenaline and insulin act as chemical messengers. Synthesized in specialized glands, they are released into the bloodstream where they bind to a receptor on a target cell. These receptor proteins alter their shape, triggering molecules called G proteins to initiate a cascade of chemical processes that transmit the message to the cell.

The molecule synthesized by Koert consists of three units: a receptor, a transducer, and an effector, as reported in the journal Angewandte Chemiel.

The receptor, a "pair of molecular claws", grasps the target, a zinc atom. The transducer unit consists of a backbone of three carbon rings linked side-to-side, attached to the receptor at one of the rings. Once the zinc atom is captured, the rings are turned inside out and the entire backbone changes shape.

The effector unit is comprised of two arms with fluorescent groups at their tips, absorbing light of one wavelength and emitting it at a slightly changed wavelength. Initially close together, the arms swing open upon the change in backbone configuration, thereby causing the groups to emit light of a shorter wavelength. UV light is transmitted as opposed to blue-green light, this change representing the "output signal" for the molecule.

While the principles behind this synthetic transducer are not new, it has paved the way for novel ideas. Perhaps researchers can synthesize a molecule with arms at each end, capable of responding to one chemical signal by releasing another. Or perhaps because of its strong resemblance to the structure of cholesterol, embedded in the cell membrane, this new molecule may be inserted into cell membranes as a means of signal transduction. Regardless of what researchers choose to do with this newfound molecule, the future seems bright.

Stem cells clear away the clouds for cornea patients

The hoopla concerning stem cell research over the past few years has produced the idea that tissues—perhaps even entire organ systems—might be produced for those in need of transplants. Although the prospect seemed far away, two teams

now report that they have artificially constructed one of the most complex portions of the eye: the human cornea.

Ray Jui-Fang Tsai of the Chang Gung Memorial Hospital of Taoyun, Taiwan and Ivan R. Schwab and R. Rivkah Isseroff of the University of California, Davis both describe the bioengineering of the cornea, producing a clear, thin film of epithelial tissue for use in transplants. The fragile cornea, which sometimes becomes clouded or cracks, often results in vision loss in the elderly. And normal transplants mostly fail.

The two teams now report that, using different approaches, they have been able to successfully transplant bioengineering corneas into the victims of vision loss.

Those patients had damaged corneas with depleted stem cells, a situation that would lead to scarring and opacity if a normal transplant were performed.

Tsai solved the problem by growing some corneal stem cells from the patient's healthy eye and feeding them into an amniotic membrane matrix. The cells grew into a layer of tissue that could then be stitched into an impaired eye, replacing damaged corneal tissue. Of the six patients Tsai treated, all regained vision.

But Schwab and Isseroff used a different approach, which reversed vision loss in 10 of 14 patients. They harvested stem cells from the donor, too; but the team first divided them into groups, some of which were placed on amniotic membrane, others frozen for later use. That way, if the transplant failed, the group could provide further treatment without reharvesting the stem cells.

Unfortunately, the results of these two experiments have not been assessed on the long term. But the biomaterials used will already be available in clinics for patients in need of immediate care. "The really exciting thing is where this can take us," Schwab said to reporters at Nature. "Replacing diseased tissues and organs with bioengineered tissue is rapidly moving from the realm of science fiction to reality."

Cell fusions create genetically engineered sheep

The same researchers who created the cloned sheep Dolly and Polly now introduce the genetically modified sheep Cupid and Diana—a major step in the history of genetic engineering. Previously, genetic engineering was carried out by mutating the stem cells of mouse embryos. Modification could result in 'knock-outs' of genes to mimic human disorders. However, mutation of stem cells from other mammals proved unsuccessful.

Alexander Kind and his group at PPL Therapeutics in Edinburg used a simple variation of their cloning procedure to obtain the genetically engineered sheep. Instead of mutating stem cells of an embryo, the group fused a genetically modified sheep cell with sheep eggs whose nuclei had been removed. The egg cell then matured normally. The engineered sheep contained the human gene for alpha-1-antitrypsin along with a gene triggering production of the enzyme only in sheep with mammary glands.

Fourteen clone sheep were created, of whom three (Cupid, Diana, and an unaffected sheep) survive. Diana's milk contains high levels of alpha-1-antitrypsin. Dave Solter, a researcher at the Max Planck Institute of Immunobiology in Freiburg, German, points out that "It was inevitable that this step forward would be taken as soon as the nuclear transfer technique was developed."

While this technique will not, obviously, be used on humans, there is a wide range of alternative uses for this technique." All inbreed dogs seem to have some kind of inherent genetic disorder and these could be corrected," said Solter. Other options include genetically modified farm animals and research animals.

Waste becomes clean energy source

In the age of recycling, 'waste not' is a defining rule. But now, some scientists suggest that waste might be recycled itself, used as a biogas for fuel. Chemists at the Queens University of Belfast, Northern Ireland, revealed in early July that they have devised a way to burn biogas without producing polluting by-products.

Bacteria in swamps and other areas in fact, regularly produce biogas. But most of this fuel is not used because it produces polluting products. The biofuel can also be made by using anaerobic bacteria grown on waste material in airtight tanks.

Robert Burch and Barry Southward now propose that biogas from waste treatment plants in municipal areas or from agricultural sources be used as a renewable energy source. If the fuel is simply burned, it produces ammonia, which generates noxious gases upon further reaction. The two scientists have, however, reported that they have found a catalyst to efficiently convert ammonia into nitrogen gas instead of nitrous oxides. The catalyst is the first to allow for conversion at low temperatures, and consists of a simple mixture of copper oxide and platinum on aluminum oxide.

The report, in July's "Chemical Communications", suggests that use of this catalyst might be the first step to using waste material as a clean, efficient energy source.

Antibodies fight chemical warfare

Despite seemingly endless debate and political controversy surrounding the use of chemical weapons, some biologists have taken no risks. Plotting ahead with plans to avert major catastrophe, a team of French researchers reported in July that they plan to destroy the toxic substances used in chemical weapons from the inside out: using the body's own antibodies.

Currently, chemical weapons are interdicted using artificial enzymes that hydrolyze, or split, common nerve agents called "organophosphorus" compounds. But for some nerve gases, enzymes don't work and may not even be readily available to treat nerve gas victims.

Pierre-Yves Renard of the CEA Service des Molicules Marquies at Saclay, France, and colleagues have reported that to destroy these nerve gases, one could use a modified antibody. Renard reported in The Proceedings of the National Academy of Sciences USA that antibodies can be tagged to attack specific nerve gas molecules.

Antibodies are normally used by the body's immune system to protect us from infections by attaching themselves to invading particles and marking these invaders for destruction by immune cells. But when nerve gas is attached to these antibodies, the antibodies can be used as enzymes to catalyze the destruction of the nerve gas molecules.

The antibodies constructed by Renard were raised in mice to correspond to the transition states of VX nerve gas. Renard then established that these catalytic antibodies neutralized VX itself.

The possibility of using such antibodies in actual nerve gas cases is strong. Although several nations signed a chemical weapons treaty the Chemical Weapons Convention, many did not, and so are not prohibited from using chemical weapons.

Killer T cells lose their bite when HIV hits

Victor Appay of the Institute of Molecular Medicine at the John Radcliffe Hospital in Oxford, UK, and his colleagues explain in the Journal of Experimental Medicine that certain anti-viral "killer T cells" in people infected with HIV lose their sting over time.

CD8 T lymphocytes, commonly called "killer T cells," are part of the body's front-line of defense. They kill virus-infected cells and produce anti-viral proteins that interfere with virus multiplication. In CD8 T cells that recognize HIV antigens and attack HIV-infected cells, HIV hinders the former function.

In patients who have the virus but have yet to develop AIDS, CD8 T lymphocytes produce small amounts of perforin, a protein that leads to the death of an infected cell. The CD8 T lymphocytes continue to release normal amounts of cytokine proteins, which help kill the virus directly.

To fully mature, CD8 T cells need the help of CD4 T cells. However, HIV directly infects and switches off CD4 cells. Their gradual loss brings on the severely immunodeficient state of AIDS. It is theorized that the disappearance of CD4 T cells also strands the HIV-responsive CD8 T cells at an immature stage of development. The stunted CD8 T cells continue to produce anti-viral proteins, however.

For years, scientists have wondered why the immune systems of HIV-infected patients hold the virus in check for years without eradicating it completely. These findings may help explain what happens during this silent period of chronic HIV infection.

'Super-aspirin' inhibits cancer

One in a hundred colorectal cancer patients lives in constant fear of a relapse. These patients are genetically predisposed to growing clumps of precancerous cells, termed polyps, in their colon. If left untreated, the polyps become cancerous and lethal. The disease, termed familial adenomatous polyposis (FAP), could only be cured until recently by surgery.

Work by Gideon Steinback and his colleagues of the University of Texas in Houston, however, suggests that ingestion of celecoxib may reduce the number and size of polyps—and, correspondingly, the incidence of colon cancer.

Colecoxib, also known as 'super-aspirin', is a non-steroidal anti-inflammatory drug (NSAID), like aspirin. While most NSAIDs inhibit the cancer-produced enzymes *Cox1* and *Cox2*, celecoxib inhibits only *Cox2*. By doing this, celecoxib avoids the side effects that result from NSAID inhibition of *Cox1*: gastric ulcers and hemorrhage. *Cox2* is a formative agent in polyp formation.

This work represents the first time a drug could be used to affect polyps. In this, Mayeto Mark Taketo of the Graduate School of Pharmaceutical Sciences in the University of Tokyo, Celecoxib is extremely useful. "It would be very beneficial to young FAP patients if we could postpone surgery even for 5, 10 or 15 years while they are taking celecoxib, because radical operations of the colon can affect not only the quality of their lives but also their psychological development."

Gene driving metastasis discovered

Cancer-treating physicians have long been frustrated by the process of metastasis, through which cancer cells are able to break off tumors and travel to other parts of the body. The behavior, which can have fatal consequences and is responsible for 90 percent of cancer deaths, is reportedly controlled by a single gene identified by MIT researchers and discussed in a report in the August 3 issue of Nature.

Edwin A. Clark, who, while conducting his work, was a post-doctoral fellow in the Hynes laboratory at MIT's Center for Cancer Research, reported the research. Clark used mouse and human models of melanoma (skin cancer) along with data from the Human Genome Project and gene expression tools called DNA arrays to identify genes that might be responsible for regulating metastasis.

Clark and colleagues examined variants of melanoma cells, classifying them according to their ability to metastasize. They then used DNA arrays to determine how expression patterns of certain genes varied between the metastasizing cells and non-metastasizing cells. They found that 32 genes were expressed more strongly in metastatic cells than in non-metastatic controls.

Clark noticed that the genes were all either cell adhesion, cytoskeletal, or blood cell forming genes. One of these, *rhoC*, is a regulator of both cytoskeletal structure and cell movement, making it a highly likely candidate for metastasis. Clark investigated the gene and found that it played a central role in the metastasis process.

"If we overexpress *rhoC* in poorly metastatic cells, they become highly metastatic. If we express an inhibitor of it in highly metastatic cells, then they show reduced metastasis," said Richard Hynes, Director of the Center for Cancer Research. Hynes also suggested that the discovery of rhoC's involvement in metastasis could provide the basis for new cancer therapies.

"I believe that invasion and metastasis are now ripe for a concerted attack because of two scientific advances: our deeper understanding of the molecular basis of cell adhesion based on

the past couple of decades of cell biology and the availability of mammalian genomic sequences and methods for exploiting that information."

Chocolate keeps arteries clean

As if one needed yet another reason for eating chocolate, yet another has been created: Chocolate is good for your heart. Chocolate, according to a recent report in the American Journal of Clinical Nutrition, contains flavanoids.

Flavanoids, which may also be found in red wine, tea, onions, and parsley, protect the heart in two ways.

They reduce the amount of low-density lipoproteins (the 'bad' cholesterol) in the blood. Flavanoids also reduce the clumping of blood platelets, a symptom that sometimes leads to atherosclerosis.

Carl Keen from the Department of Nutrition of the University of California at Davis found that subjects who drank cocoa, as opposed to caffeinated drinks and water, had lower platelet activation. Blood took longer to clot. Despite this, the benefits of chocolate in combating atherosclerosis is debatable.

Tissa Kappagoda of the Preventive Cardiology Unit of the University of California at Davis warns that chocolate "contains a significant proportion of fat calories and it may not be the best option for patients with coronary atherosclerosis".

Depression influences developing fetuses

Prenatal depression affects the developing fetus in many negative ways, including premature birth, lowered IQ in later life, and slower response time.

Fetuses and newborns have levels of cortisol that reflect maternal levels. High levels of this hormone are often found in depressed people.

Tiffany Field of Nova Southeastern University in Fort Lauderdale, Florida, notes that "If you measure the cortisol level in the mother at 28 weeks gestation, you can predict with 98% certainty if the baby will be born premature." Massage and exercise that lowers cortisol levels seem to counteract this effect.

The brain activity of two-day-old babies of depressed mothers is concentrated in the right side, often associated with negative emotions. These babies respond poorly to the "Brazelton assessment," a test for newborns that analyzes facial expression, movement and response to sound.

Depressed babies also have erratic sleep patterns in which they frequently alternate between deep and active sleep, which is called "indeterminate sleep." It turns out that the only neonatal variable that has any bearing on the predicted IQ of 12-year-olds was whether or not they had had indeterminate sleep.

Another group looked into whether or not full-term newborns of depressed mothers differ in their behavior and responses to sensory stimulation. All the babies could distinguish differences in the weight and roughness of objects, but they took twice as long as normal to do so. These babies also showed no preference for their mother's faces and spent less time exploring new objects.

PHYSICAL SCIENCES

Laser shines from chemicals in a beaker

Lasers are now so commonplace that the average student can hardly escape being annoyed by the sight of a red laser dot flash across a lecture hall. Now scientists are planning to dispense them all over: so ubiquitously, in fact, that bacteria might have them.

In the May 22 issue of Applied Physics Letters, Hui Cao and colleagues from Northwestern University describe a new laser no bigger than a microbe. And this is no standard lecture pointer. Mixing chemicals in a beaker produces the laser, reports Cao.

Cao devised a chemical synthesis to generate particles of the semiconductor zinc oxide. The particles, which are a few nanometers in size, can be made to clump together in aggregates, creating a random jumble of nanoparticles inside a beaker. The tiny zinc oxide particles scatter short-wavelength ultraviolet light, much like water droplets in clouds scatter sunlight. The scattering can become so extreme, in fact, that light cannot travel far without being sent in a new direction—causing the light to localize to what Cao calls 'optical cavities'.

When Cao then stimulated his zinc oxide mix with a laser, he found that his concoction emitted ultraviolet light with a very narrow range of wavelengths—the characteristic signature of a laser beam. Normally, laser light's wave-like oscillations are all in sequence with one another, unlike ordinary light. First, light is emitted from a substance that has been 'excited'; then the light stimulates the emission of more light, and a chain reaction results. The light produced is confined to a chamber by mirrors, which cause the light to bounce back and forth, stimulating more emission with each bounce.

The tiny beaker-laser made by Cao might be easier to make than the common mirror-guided laser. So does Cao call for mass production? Not so fast, say his colleagues. One disadvantage of the system is that its properties are hard to control. The emission wavelengths are also randomly generated, and so the laser beam produced might be quite unfocused. Nevertheless, Cao's discovery opens the door to the prospect of true nano-laser light.

Lightning misses the point

New research indicates that the common belief that sharp rods make the best lightning conductors is false, as blunt tips attract lightning more than sharp ones.

According to legend, Benjamin Franklin established the electrical nature of electricity by flying a kite in a thunderstorm. In another experiment, Franklin discharged electrified objects with a metal needle, concluding that a sharp iron rod on rooftops may similarly discharge thunderclouds when lightning strikes. Franklin had successfully invented the lightning conductor.

Recently, however, Charles Moore from the New Mexico Institute of Mining and Technology, and colleagues, discovered that a blunt-tipped rod is in fact a better conductor. Moore set up an array of aluminum rods ranging from half an inch to two inch-

es in diameter at the summit of South Baldy Peak, a 3,288-meter mountain in New Mexico. Blunt and sharp tips were placed approximately six meters apart, thereby giving the lightning a clear choice.

This study was continued for seven years, during which no sharp rods were hit, but twelve blunt rods were struck, as reported in the journal Geophysical Research Letters.

Moore's group also tested new lightning-protection devices called early streamer emitters (ESEs), which apparently reduce the severity of a lightning strike by emitting a "rapid return stroke" (streamer) back up to the thundercloud. These devices are said to provide a greater area of protection and typically use sharp-tipped rods. The researchers concluded that ESEs do not emit "early streamers" during thunderstorms and that they, in fact, do not offer any more protection that any ordinary rod.

Thus, the researchers concluded, "Franklin's method for providing [lightning] protection has been made less effective than it could be, by his urging that the tip of lightning rods be sharpened."

Potent sensor amplifies signal

The best sensors generate large signals when it detects a small amount of the substance of interest, but amplification of the signal is usually done electronically. However, a molecule that acts as both a sensor and an amplifier has been developed by the groups of Vincenzo Balzani at the University of Bologna in Italy, and Fritz Vvgtle, a chemist at the University of Bonn in Germany. Their findings were recently published in Chemical Communications.

Balzani's team has created molecules that signal when they attach themselves to target molecules. Most sensors work in 1:1 relationships; that is, for every target, you need one sensor. However, researchers have made a molecule with 32 light-emitting units that all turn off when the molecule binds a single target, an ion of cobalt.

The sensor is a dendrimer, or cascade of branching molecular chains that form a multi-armed core. The chains clump together to form a roughly spherical shape. In the dendrimer created by Balzani's and Vvgtle's groups, all 32 branch tips glow green under ultraviolet light. The inside chains contain nitrogen atoms that interact with the cobalt ions. When the dendrimer traps one cobalt ion, the 32 branch tips stop fluorescing.

Twenty-five micrograms of cobalt in a liter of water, three to five times the safe daily intake for humans, can produce a significant change in intensity of the sensor.

The value of this research is in its principles, not in immediate practical applications. Designer dendrimer branches may be able to recognize and trap other targets selectively or may become better at amplification. Scientists are currently exploring the possibility of using dendrimers as drug molecule transports or as catalysts.

Diamond yields to beta-carbon nitride in hardness

In 1989, beta-carbon nitride was predicted to be possibly harder than diamond, the hardest known substance. Physicist Marvin Cohen of the University of California first proposed beta-carbon nitride in 1985.

Knowing that hard materials contain atoms linked by short, strong bonds, Cohen realized that similar bonds form between carbon and nitrogen atoms. He believed that these atoms could form a stable crystalline solid where the expected ratio of nitrogen to carbon would be 1:3:1.

Yip-Wah Chung at Northwestern University in Illinois and colleagues report in Applied Physics Letters that they have something matching the description of beta-carbon nitride.

A superlattice of alternating layers of zirconium nitride (ZrN) and a mixture of carbon and nitrogen, each a few atoms thick, has features expected of Cohen's hypothetical material. It is tentatively called CNx, and the precise composition and structure is not fully known.

The strongest evidence comes from measuring X-rays kicked out of the CNx layers by the energetic electron beam, and from the way that the material absorbs infrared radiation. These measurements suggest that the bonds between the carbon and nitrogen atoms have the proper predicted characteristics. And finally, the ratio of nitrogen to carbon in the films matches the prediction of 1:3:1.

Bouncing electrons off the layers shows that CNx films contain orderly, crystalline regions. The density is also roughly the same as that predicted for beta-carbon nitride.

Though promising, this evidence stops short of providing definitive evidence that beta-carbon nitride has been found.

Synthetic surface binds proteins at will

University of Chicago researchers have developed a surface capable of binding and releasing proteins at will. The 'self-assembled monolayer' is created from a molecule-thick coating of 'alkylthiol' organic molecules on a surface of gold. Alkylthiol molecules are composed of an oil-like carbon tail with a sulfur head that binds to the gold. Different ligands may modify the properties of the surface.

Drs. Hodneland and Mrksich reasoned that a surface capable of releasing bound ligands at a trigger would be useful in studying cell adhesion. The ligand chosen was biotin, to which the protein streptavidin binds.

To create the surface, the researchers tipped the carbon tails of alkylthiols with a quinonepropionic ester (QPE), to which the biotin ligand was attached. The QPE responds to voltage, freeing the biotin ligand from the alkylthiol surface. The surface functioned effectively in practice, binding streptavidin to the surface, then releasing it with biotin when voltage was run through the gold surface.

Theoretically, any ligands may be connected to a similar surface, making a number of applications feasible. In addition to studying cellular adhesion, such a surface may be eventually used to control the release of drugs.

Additives make carbon dioxide better solvent

Non-toxic carbon dioxide in its compressed or "supercritical" state is a promising, cheap alternative to organic solvents now used in chemical industry. The gas is colorless, odorless, non-flammable and can be easily depressurized and released after use. It is also a picky solvent. For example, it can strip away caffeine without disturbing the chemicals that give coffee its flavor.

However, as Walter Leitner of Max Planck Institute in Mulheim, Germany, notes, the "poor solubility of many interesting target substances places severe restrictions on the widespread use of carbon dioxide [as a solvent]." Dangerously high pressure or expensive additives must be used to dissolve insoluble catalysts into carbon dioxide.

Fortunately, researchers have developed a suitable, chemical go-between that could lead to widespread carbon dioxide use. Eric J. Beckman and colleagues at the University of Pittsburgh, Pennsylvania, report in Nature that they have a cheap additive that easily dissolves in low-pressure carbon dioxide.

Beckman's group has linked together two distinct chemical groups (polyether and carbonyl). Formerly insoluble catalysts and reagents can be attached to this low-cost chain that falls apart in carbon dioxide. This high solubility helps to dissolve the attached groups. Other functional groups may stabilize useful emulsions between water and carbon dioxide.

This development could dramatically lower the cost of using carbon dioxide in commercial applications such as dry-cleaning and pharmaceutical production.

Scientist plumbs gravity at atomic level

Decades ago, Bose and Einsten proposed the existence of a new state of matter: the Bose-Einstein condensate (BEC), in which atoms are sustained at such low temperatures that they exist in the same quantum ground state. BEC could provide an ideal system from which to study basic quantum phenomena.

In the late June edition of Physics Review Letters, Gershon Kurizki and Duncan O'Dell of the Weizmann Institute report that BEC could also be used to examine atomic responses to forces like gravity.

Kurizki and O'Dell propose that if they were to induce long-range attraction between atoms, the force would produce a gravity-like state; unlike gravity, the force could be made strong enough to allow for an examination of the behavior of atoms at a microscopic level.

The proposal is the first to concern the examination of gravitational forces using BEC. Previously, atoms in BEC could interact only at very short range. But by streaming intense laser light into the BEC, O'Dell and Kruizki say they can induce a long-range force, producing an artificial gravity-like attraction. Their proposed apparatus uses 18 different lasers focused at unique angles to induce the appropriate force.

But MIT Prof. Wolfgang Ketterle told Nature reporters, "The laser power is at the limit of what is available," he said. "To use 18 such beams is extremely difficult."

Alternatively, one could conduct the experiment using lasers of shorter wavelength, creating small condensates with just a few atoms. "The experimental challenges are considerable," admits Duncan O'Dell of Kurizki's group.

But "the ability to emulate gravitational interactions in the lab is fascinating," says Kurizki.

Telescope data confirms Euclidean nature of universe at the Big Bang

The history of cosmology is filled with a spirit of revision. Models are postulated, new observations recorded, old models discarded, and new theories constructed. But in violation of this seemingly impenetrable historical rule, cosmologists recently received data that confirmed an old theory.

"The Boomerang results fit the new cosmology like a glove," Michael S. Turner of the University of Chicago. Turner had been referring to data from the Boomerang telescope, which, along with the telescope Maxima, has been used to confirm postulates about the very beginning of the universe.

To discover fundamental events making up the big bang, both telescopes were used to track levels of cosmic microwave background radiation—radiation sources believed to have been left as remnants of the big bang.

Boomerang flew over Antarctica and tracked just 3 percent of the sky; Maxima covered even less area, flying over Texas, but had greater resolution than its counterpart. By constructing data from the two telescopes, cosmologists determined approximately when the radiation might have been released. It is believed that the photons of background radiation were bound in a plasma state; some parts of the plasma were denser than others, attracting particles toward them. Pressure, on the other hand, pushed the particles apart, creating a battle between pressure and inertia. The plasma oscillated, creating vibrations like sound waves.

Picking up those vibrations in the cosmic atmosphere, Boomerang and Maxima were able to confirm what physicists expected: that oscillations would be found at a particular frequency, then half of that frequency, and so on. Both telescopes confirm that the vibrations existed, and lead to a framework to discuss the very early geometry of the universe. Because of the specific patterns detected by the two telescopes, cosmologists were able to demonstrate that the early universe was in fact Euclidean in form: the rules of Euclidean trigonometry applied. But odd data from the two telescopes also show that the universe is slightly spherical, meaning that it was likely to be very close to completely flat originally and has now been distended by forces of gravity.

Scientists indirectly measure 'dark' matter

Cosmologists have calculated how much matter should be in the Universe, but they can't find it. The mass of all the stars, galaxies, black holes, and light particles, gives only a few per cent of the total. However, David Wittman of Bell Labs, Murray Hill, New Jersey, and his colleagues have a solution. Dark matter can now be measured indirectly, based on the shape of distant galaxies.

Large objects, such as galaxies, bend passing light slightly towards themselves. The heavier they are, the more the light deviates. This "weak gravitational lensing" may be used to estimate how much mass lies between us and distant galaxies. As light from far-off nebulae comes towards us, intervening mass and dark matter causes its path to deviate and distorts the image. The team found that adjacent galaxies appear to be squashed by approximately 1%. These distortions are used to make the calculations.

Their result may knock out the popular "cold dark matter" model, which claims that sub-atomic particles make up most of the missing mass. However, these new findings support the "vacuum-energy" theory—the idea that space itself has a dark energy that may oppose gravity.

The team plans to look at approximately 100 million sources across the sky to develop a picture of the distribution of matter. "We now have a completely independent, direct way of measuring mass distribution—it has a fantastic future," says Anthony Tyson, a member of Wittman's team.

Brown dwarves flare like other stars

Celestial brown dwarfs, though bigger than the biggest planets, are much smaller than the smallest stars, which themselves maintain just enough mass to be able to sustain thermonuclear reactions in their cores. Also known as failed stars, brown dwarfs can only shine about one tenth of a percent as brightly as the sun.

Recently, however, as reported in the July 20 issue of Astrophysical Journal Letters, scientists working with NASA's Chandra X-ray Observatory saw the brown dwarf LP 944-20 unleash a bright x-ray flare. "We were shocked," said Robert Rutledge, professor at the California Institute of Technology. "This is really the mouse that roared. We didn't expect to see flaring from such a lightweight object."

Researchers expected detecting a few photons per hour in observing LP 944-20. However, LP 944-20, like all brown dwarfs, lacks the ability to support strong fusion reactions, obtaining most of its energy from release of gravitational energy as it contract a few inches per year.

"It was as if we were searching for a dim bulb and instead found a bright flash of light," added Lars Bildsten, a researcher from the University of California at Santa Barbara. Moreover, the energy contained in the flare was comparable to a small solar flare and approximately one billion times greater than x-rays emitted from Jupiter.

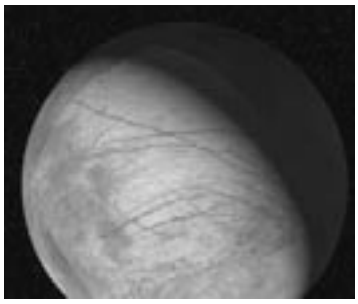
This dynamic flare also offers some evidence that brown dwarfs and perhaps even giant planets have magnetic fields. Principal investigator Gibor Basri suggests that the flare may have resulted from turbulent magnetized material below the surface of LP 944-20, producing a twisted magnetic field. "A subsurface flare could heat the atmosphere, allowing currents to flow and give rise to the x-ray flare-like a stroke of lightning," he remarked.

This x-ray flare, although the first ever observed from a brown dwarf, will most likely not be the last. Basri confirms, "New sky surveys show that the objects may be as common as stars."

The search for aliens turns to Jupiter

Although extraterrestrial life seekers have often looked to Mars for signs of water, spectacular images from NASA's Galileo spacecraft have redirected focus elsewhere: Jupiter's moon Europa. Galileo has taken pictures of what scientists conclude is a liquid water ocean on the Jovian moon hidden beneath surface ice several hundreds of meters thick. But believers in extraterrestrial creatures haven't won this battle: the buried ocean is, unfortunately for them, probably too dark to support complex photosynthetic organisms.

But the latest theory in a hotly contested debate on Europa's frozen ocean suggests that the water body may have indeed supported photosynthetic life at some time in the past. The theory comes from Eric Gaidos of the Jet Propulsion Laboratory in Pasadena and Francis Nimmo of the University of Cambridge, UK. The duo suggest that Jupiter's gravity could cause ice flow on Europa, creating frictional heat and producing warm, soft ice. The warmer ice, they say, explains "ridges" in the moon observed by the Galileo spacecraft after its Jupiter fly-by.



If Gaidos and Nimmo are correct, Europa's ocean could have small pockets of salty water, melted by friction and sufficiently warm to allow "liquid water and sunlight coexist—a requisite combination for photosynthetic life," according to Gaidos. "Terrestrial organisms may be able to survive many thousands of years without a source of energy, frozen in ice or permafrost," he said. But these lifeforms are probably far from the quintessential alien characters invading 1950's movie screen. If they do exist, the organisms are probably frozen bacteria, given a chance to grow because of the combination of sun and water on Europa's warm ice pockets.

The theory, however, has not received a warm welcome from planetary scientists. Gaidos and Nimmo, themselves, admit that a melted water pocket will probably quickly lose heat and may only last for a few years. The problem was discussed by Astrobiologist

Bruce Jakosky of the University of Colorado, who said that "People are trying to understand the liquid water budget, what energy sources there are, and whether there is there adequate energy to support life…we don't know what life would look like in such a different environment."

Mother Earth may wreck anti-pollution drive

Global warming is steadily increasing and shows no sign of stopping. An increase of 0.1 to 0.3 degrees C in the next ten to fifteen years may conclusively prove that the global warming is manmade, but an untimely volcanic eruption may cool the world and stunt current efforts to contain pollution. The most significant attempt to control the greenhouse-gas emissions was the Kyoto Protocol, developed in 1997. In order to be effective, however, the treaty must be ratified with 55 signatures. As of yet, only 22 countries have signed the protocol.

Island and low-lying countries whole-heartedly approve of the protocol, but countries with less to lose are less enthusiastic. Developing countries blame developed nations for the bulk of greenhouse gases, and developed countries are stalling until underdeveloped nations take the first step. This political bickering is only helped by the argument that the greenhouse effect may not be manmade.

William Hyde and Thomas Crowley of Texas A&M University state that a volcano's cooling effect on the earth within the next decade "would likely confuse the debate about temperature trends and impede detection of the climate change signal."

Volcanoes cool the earth by shooting sulfate aerosols into the atmosphere that reflect the Sun's heat and cool the earth. Hyde and Crowley investigated ice cones recording climate-changing eruptions over the past six centuries.

Using this history, they concluded that the possibility of another major eruption within the next decade is 37%. The possibility of two eruptions within a decade is 15%, and the possibility of a single eruption with three times as much cooling effect as a major volcano is 20%.

While these percentiles are crude approximations, they do provide a starting point from which to start working with future climatic changes.

Volcanoes boil sulfur away

New research shows that sulfur, also known as brimstone, may burn up in volcanoes. Sulfur has been long associated with volcanoes, both in ancient mining sites and in the delectable sulfuric stench of rotten eggs around some volcanic regions. Geologists have also sighted lakes of liquid sulfur around a volcano in Costa Rica. Despite this, frozen greenish-yellow flows of sulfur are rarely found around volcanoes. Geologists assumed that this was due to the lack of sulfur around active volcanoes. The only previous sighting of a sulfur flow, documented by Japanese researchers in 1936, supported the assumption that any sulfur present would have solidified.

Andrew Harris and Robert Wright from the Open University, UK, along with Sarah Sherman from the University of Hawaii witnessed a significantly different scene. The molten, dark red sulfur lakes they found burned with a bluish flame, releasing sulfur dioxide. The lake burned away within four hours, leaving a characteristic trench that the researchers found elsewhere around the volcano. The lack of sulfur flows around volcanoes thus seems to be due to the heat of the volcanoes and not to the lack of sulfur itself. This new evidence may force a reconstruction of volcanic deposits on Earth as well as Io, Jupiter's closest moon.

ENGINEERING AND TECHNOLOGY

Molecular transistors show some 'gain'

Miniaturized computers are certainly in demand. The only problem is that silicon chips can only get so small before they start malfunctioning. In response, some scientists have turned to the construction of chips through small molecules with the ultimate goal of producing a molecular transistor.

Transistors are essentially on-off switches, determining whether or not current will flow through a wire by applying a voltage to a 'gate' electrode that sits between the input and output electrodes. Transistors commonly have three terminals: an input ('source') electrode, an output ('drain'), and a control electrode ('gate'). But a transistor is not only a gate: it can also act as an amplifier, boosting weak currents passing through it. This amplifying property, or "gain", is an essential component of any modern information processing circuit.

Although attempts to achieve gain through molecular devices had often resulted in failure, Massimiliano di Ventra of Vanderbilt University and his colleagues report in June 2 edition of Applied Physics Letters that gain may very well be achieved in a molecular device. Di Ventra used computer calculations to predict the behavior of a three-terminal molecular device: a benzene molecule connected with sulfur atoms that acted as drains and sources. Two tiny metal disks placed above and below the ring served as a gate. Using his computer simulation, di Ventra was able to show that the conductance of his model molecule increased sharply if the voltage on the gate was adjusted to a particular value.

A channel through the molecule acted like a bridge for current, an effect known as 'resonant tunneling'. The result demonstrates that a molecule-sized transistor may indeed be a feasible creation—perhaps one arriving to us in the near future.

Metal-binding peptides attract nanocrystals

There are two ways to build an object: carve it from a singular block of material or build it from its many parts. Traditionally, electronic circuits are created using the former method. Flat sheets of silicon with dopant atoms are cut into sandwich-like pieces and carved into useful structures. Researchers now are working to further miniaturize the assem-

bly process by using 'nanocrystals' of semiconducting materials to build circuits.

Nanocrystals are small items, only a few nanometers across. This makes them several times smaller than the transistors in regular integrated circuits. Nanocrystals can then plausibly be used as switching mechanisms or memory element.

The small crystals are organized by metal-binding peptides. Our cells move materials on the same small scale as the nanocrystals. Nature is used to work with organic-based compounds, and is not adapted to semiconductor materials.

Previous research done at the University of Copenhagen revealed that proteins have chemical groups that bond to certain metals. Scientists found peptides that have the right molecular structure to fit on the surface atoms of the metals.

At the University of Texas at Austin, Angela Belcher and other researchers have found the analog of the peptides and metals for semiconducting materials. After 5 rounds of trials of running different peptides chemically linked to protein coats of colphage they discovered which phage particles were capable of sticking to the metals, and from that were able to isolate the useful peptides.

Lasers stretch cells

Researchers at the University of Texas used Newton's third law to create an optical stretcher. Optical stretchers are used to manipulate cells and will allow cancer researchers increased flexibility to handle cells.

The predecessor to the optical stretcher is the optical tweezers. The tweezers consist of two equal and opposite laser beams that hold a cell in between them as the forces from the moving photons cancel each other out. This device is regularly used to manipulate cells, particles, and DNA.

Josef Kas (UT) discovered that by utilizing stronger but less focused laser beams, that cells could be stretched without being damaged.

Before this was revealed, scientists had problems trying to use the tweezers to stretch cells because they were not able to grab the cell strongly enough to pull it. Handles had to be attached to the cells by means of particle attachments to the cells. This technique was limited by how strongly the handles could be attached.

The advent of the optical stretcher will lead to more explorations in cell mechanics and complex deformation of complex cells.

Optical torch to distinguish single molecules soon

Leeuwenhoek's optical microscope is far from out of date in current research. With the development of optical torches (essentially tiny light sources) researchers can now view objects 500 nanometers across lit up by a single-molecule probe.

Vahid Sandoghar and colleagues at the University of Kanstanz (Germany) built their probe based upon an old technique called scanning near-field optical microscopy (SNOM). The restriction of the old technique was that the resolution of the old microscopes could not be focused on a point finer than its own wavelength (diffraction limit).

The optical torch developed by Sandoghar consists of a tapering optical fiber that has a single molecule on its tip. On the molecule, there is a single organic crystal, which has a few molecules of fluorescent terrylene in it. The laser light absorbed by the terrylene molecule can be tuned to excite just one of the terrylene molecules. Thereby, creating a singular molecule of light: the optical torch.

Researchers expect soon to be able to achieve resolutions capable of picking out individual molecules.

Atomic manipulation reveals possibility for quantum chips

The prospect of producing an electronic circuit out of just a few atoms has received hot debate—but with few results. The idea proposed is to use atoms instead of electrons; and while few have produced a chip itself, researchers in Austria recently demonstrated that atoms could in fact be guided along tiny wires, perhaps forming the basis for entirely new types of computers.

Jvrg Schmiedmayer of the University of Innsbruck, Austria and colleagues report in the journal Physical Review Letters that atoms can be moved down chip-sized wires just a few thousandths of a millimeter wide. Schmiedmayer believes that he can use this movement of atoms just like a conventional electron current, performing the computational operations that occur in our everyday computers at an even smaller level. The atomic wires, says Schmiedmayer, could act as magnetic guides to direct atoms in a chip.

Schmiedmayer used gold-plated semiconductor chips etched with 'ditches' to allow electrical current to pass and induce a magnetic field. Combining this magnetic field with fields produced by nearby, thicker wires, the researchers create a magnetic "canyon" to move magnetic atoms along a defined path.

Why use atoms instead of electrons? Unlike electrons, atoms can be moved into quantum states known as Bose-Einstein condensates, allowing for scientists to use the laws of quantum theory to achieve a computing power greater than that in conventional computers.

Quantum logic gate structure proposed

As silicon-base hardware rapidly nears its limits, quantum computers—capable of processing information billions of times greater than conventional machines—seem to be the holy grail of the computer industry. The idea of storing information in the quantum properties of particles was thought up by Nobel Laureate Richard Feynman. A quantum computer would represent data, termed qubits, though the quantum state of particle, such as the spin of the nucleus.

Carlo Jacoboni's group in the University of Modeno, Italy, has proposed a structure for a quantum logic gate—the basic component of a quantum computer. The logic gate, consisting of a pair of superconducting wires separated by a control wire, regulates the passage of electrons through the gate by using

interference properties of quantum particles in the control wire. The setup may be viewed as a preferential capacitor, which allows passage only if the control and data electrons meet and interact. Then, the data electron is pulled to the end of the control wire, where it travels on.

Unfortunately, this system is not robust. Any uncontrolled interaction of the data electron with the environment (termed quantum decoherence) would trigger an unpredictable effect on the system, corrupting the information. In addition, the system must operate at extremely low temperatures to ensure that the electrons pass along the wires without interacting with the outside.

Nonetheless, Jacoboni insists that the logic gate has potential. "Our proposed system allows an arbitrary number of qubits in a very limited space," says Jacoboni. It is also "integrable with the conventional electronics needed to interact with the end user."

The quantum computer is definitely in the future, but Carboni's quantum logic gate may be a large step towards creating it.

Device detects the sound of breaking bonds

Chemists at Cambridge University, UK, have developed a super-sensitive listening device capable of picking up the faintest of noises. The device, invented by David Klenerman and colleagues, is able to detect even the attractive forces between atoms.

Chemical bonds make molecules behave differently from the individual atoms contained within them. Bonds range from weak forces to strong covalent bonding that holds together brittle minerals. Klenerman's group investigated bond strengths using microscopic latex balls stuck to a shaking gold surface of a quartz crystal microbalance (QCM). The surface and the spheres were coated with different chemicals, thus creating stronger bonds. As the surface was shaken, the bonds holding together the balls and the surface ruptured, allowing the balls to roll away.

The rupturing of these chemical bonds generates noise. The gold platform, also vibrating, can be used as a microphone that can detect the range of weak sounds created upon breaking chemical bonds. Using this device, Klenerman's group was able to conclude that a covalent bond is about 600 times stronger than weak, non-specific attraction and ten times stronger than a hydrogen bond.

The group, reporting their results in the journal Langmuiri, states that despite being unable to pinpoint exact bond strengths, this new method could offer revolutionary ways of separating chemicals and produce biosensors capable of detecting biological molecules from the way they link together.

Miniature balance counts bacteria

While it is relatively easy for tests to reveal the presence of specific bacteria in large quantities, it is far more difficult to detect and, furthermore, count individual bacteria of a given species.

Harold G. Craighead and his colleagues from Cornell University at NY and the Institute of Microelectronics at Singapore have created an instrument capable of doing this.

The instrument resembles a miniature set of scales. It consists of a springboard-like cantilever made from silicon nitride. When the lever moves, it deflects a laser aimed at the unattached end. The oscillation frequency of the cantilever, derived from the light deflections, is dependent on the cantilever's mass. Since the cantilever is so small, a relatively small number of bacteria is needed to start oscillation. For example, the researchers demonstrated that a mere 16 E. coli bacteria were necessary to induce a detectable oscillation.

Specificity of the machine for a distinct species is achieved by attaching antibodies to the cantilever. The antibody guarantees specificity not only of species but also of particular strains of bacteria.

The organism used to test the cantilever was *O157:H7*, a toxic strain of E. coli. The scale was able to differentiate between the toxic E. coli and Salmonella bacteria.

Gene Patenting
A Threat to Science and More?

Martin Kurtev

On June 26, 2000, scientists made the announcement that they are in the final phase of mapping the human genome. The event received great media attention, and the area of genetic medicine was once again in the spotlight. The possible benefits to society from the practical applications of the complete sequencing of all human genes appeared immeasurable—new knowledge was to be gained, new drugs were to be discovered, old diseases were to be cured. Beneath the apparent excitement over this great accomplishment of science, however, lay the controversial issue of gene patenting. Long before the genome was anywhere close to completion, a number of public and private sectors had already filed for gene and sequence patents, which gave them exclusive rights over that particular piece of genetic material. As the number of patents, which appeared to be granted too easily, increased, many scientists and clinicians found their research thwarted by various restrictions imposed by commercial, and in some instances, academic, patent holders. Just like any controversial issue, gene patenting has its critics and supporters, but the issue is important not just for the way in which it affects science but also for the away it affects each one of us. It is a debate over whether we can patent knowledge and also whether someone can have exclusive rights over something that belongs to all of us.

"Gene patent" is a broad term that refers to the patenting of either a process that involves isolation of DNA (where DNA refers to either DNA or associated materials such as RNA) as well as to a chemical substance related to DNA. As more information becomes available through genome mapping, the total number of "working" human genes will become known. This number is expected to be somewhere between 28,000 and 100,000. This represents only about 4 percent of the total human genome. The remainder of the genetic information is made up of sequences that do not presently have any known function. These areas of the genome will likely have future uses, which suggests that having ownership of those future uses will confer eventual benefit.

The United State Patent and Trademark Office (USPTO) allows patenting of genetic discoveries from nature as long as the inventors (1) identify novel genetic sequences, (2) specify the sequence's product, (3) specify how the product functions in nature, and (4) enable one skilled in the field to

use the sequence for its stated purpose. The USPTO has issued patents on more than 6,000 genes, about 1,000 of them human genes, dating back more than 10 years. The Office has also issued patents for gene fragments. Full sequence and function often are not known for gene fragments. On pending applications, their utility has been identified by such vague definitions as providing scientific probes to help find a gene or to help map a chromosome. Questions have arisen over the issue of when, from discovery to development into useful products, exclusive right to genes could be claimed. The 300- to 500-base gene fragments, called expressed sequence tags (ESTs), represent only 10 to 30% of the average cDNA, and the genomic genes are often 10 to 20 times larger than the cDNA. A cDNA molecule is a laboratory-made version of a gene that contains only its information-rich (exon) regions; these molecules provide a way for genome researchers to fast-forward through the genome to biologically important areas. The original chromosomal locations and biological functions of the full genes identified by ESTs are unknown in most cases. Patent applications for such gene fragments have sparked controversy among scientists, many of whom have urged the USPTO not to grant broad patents in this early stage of human genome research to applicants who have neither characterized the genes nor determined their functions and uses.

The patenting of gene fragments is controversial. Some say that patenting such discoveries is inappropriate because the effort to find any given EST is small compared with the work of isolating and characterizing a gene and gene products, finding out what it does, and developing a commercial product. They feel that allowing holders of such "gatekeeper" patents to exercise undue control over the commercial fruits of genome research would be unfair. Similarly, allowing multiple patents on different parts of the same genome sequence—say on a gene fragment, the gene, and the protein—adds undue costs to the researcher who wants to examine the sequence. Not only does the researcher have to pay each patent holder via licensing for the opportunity to study the sequence, he also has to pay his own staff to research the different patents and determine which are applicable to the area of the genome he wants to study. Research scientists who work in public institutions often are troubled by the concept of intellectual property, because their norms tell them that science will advance more rapidly if researchers enjoy free access to knowledge. By contrast, the law of intellectual property rests on an assumption that, without exclusive rights no one will be willing to invest in research and development (R&D).

In order to better understand the implications of gene patents, we must examine closely the arguments that are presented in favor and against their existence. The supporters of the idea think that patenting provides a strategy for protecting inventions without secrecy. A patent grants the right to exclude others from making, using, and selling the invention for a limited term, 20 years from application filing date in most of the world. To get a patent, an inventor must disclose the invention fully so as to enable others to make and use it. Within the realm of industrial research, the patent system promotes more disclosure than would occur if secrecy were the only means of excluding competitors. This is less clear in the case of public-sector research, which typically is published with or without patent protection. The argument for patenting public-sector inventions is a variation on the standard justification for patents in commercial settings. The argument is that post-invention development costs typically far exceed pre-invention research outlays, and firms are unwilling to make this substantial

investment without protection from competition. Patents thus facilitate transfer of technology to the private sector by providing exclusive rights to preserve the profit incentives of innovating firms. Patents are generally considered to be very positive things. Furthermore, patenting prevents wasteful and duplication of effort and focuses research into new, unexplored areas. In the case of genetic patenting, it is the scope and number of claims that has generated controversy.

Gene patenting has its critics, however, and the arguments against it are just as convincing as those for it. In general, patents of partial and uncharacterized cDNA sequences will reward those who make routine discoveries but penalize those who determine biological function or application. Thus, patents give inappropriate reward to the easiest step in the scientific process. Patents could impede the development of diagnostics and therapeutics by third parties because of the financial costs associated with using patented research data or procedures. Patent stacking, which means allowing a single genomic sequence to be patented in several ways, such as an EST and a gene, may discourage product development because of high royalty costs owed to all patent owners of that sequence; these are costs that will likely be passed on to the consumer. Because patent applications remain secret until granted, companies may work on developing a product only to find that new patents have been granted along the way, with unexpected licensing costs and possible infringement penalties. Costs increase not only for paying for patent licensing but also for determining what patents apply and who has rights to downstream products. Private biotech companies who own certain patents can monopolize certain gene test markets. Finally, patent filings are replacing journal articles as places for public disclosure, thus reducing the body of knowledge in the literature.

As hard as it may seem to believe it, gene patents may soon come to influence very important aspects of our lives. On gene patents will depend how fast new drugs are developed and how accessible they are to the people who need them. New genetic screening tests may be available only at certain places, with the patient being subjected to the pricing policies, however high they may be, of the institution owning the exclusive rights to the test. The issue becomes increasingly important when faster advances in a given research area or the quicker development of a new drug may save the lives of thousands of people affected by a certain ailment. In addition, gene patents have halted ground-breaking research in many laboratories and thus prevented clinicians from practicing the best medicine possible, even the things that are standard care. One possible resolution to the conflict between academic and commercial interests is the latter's absolute commitment not to charge exorbitant fees for testing, allowing clinical services to conducts experiments in-house and continue to do research.

The scientific and financial aspects of the issue have received great publicity but the humanistic and ethical implications of the controversy have remained in the shadow. The one ethical question that must be answered is whether someone can have exclusive rights over a part of nature—a basic constituent of life—and therefore own all or a part of another organism. The controversy over gene patenting does not relate just to scientific research and the development of new drugs and gene tests. On a more global scale, gene patenting has been used by developed countries to rob Third World countries of their natural resources. A famous case of such "genetic colonialism" is the Indian Neem tree. It has been known for centuries that extracts from the tree can cure various wounds and skin diseases. Knowledgeable of this fact, Western pharmaceutical companies were quick to improve the extraction and purification procedures through which these balms were prepared, and thus obtained a monopoly over most of the products that come from the tree. Such patents remove old methods of extraction and use from the public domain for the benefit of large first world corporations. They have also forced Indian farmers to pay exorbitant royalties for products that were formerly free. This case is one example of how the scientific and commercial interests of large biotech companies have combined to threaten the livelihood of citizens of Third World countries.

Another global threat that gene patents pose is to gene diversity. Whenever a valuable cash crop is discovered, the most useful strains gradually become dominant, while the "wild" strains - having wide genetic diversity—cease to be cultivated and may be lost. Over-reliance on a single strain puts the consumer in danger should that crop fail. The Irish potato famine remains a chilling example of loss of both intraspecies and interspecies genetic diversity. When a virus destroyed the potato crop, millions of Irish died of starvation, because they had planted nothing else, neither other crops nor a variety of potato strains, some of which might have resisted infection. On the other hand, the availability of wild-type strains for genetic research and manipulation can yield improved strains with high nutritional values and pest and disease resistance.

Just like cloning, gene patenting remains one of the most controversial areas in molecular biology and genetic medicine. The most fundamental question that needs to be answered is whether institutions can claim exclusive rights over something that they did not create but merely discovered that exists. Patenting of knowledge should not be confused with the patenting of the application of knowledge—the invention. To patent such a discovery may well profit that individual or the company for whom that individual works but it cannot possibly profit the human race; quite the opposite. The number of people that seem to be harmed by gene patenting far outweighs the number that seem to profit from it. People's livelihoods are threatened and so is the existence of some organisms. It seems that the soul of Science has been possessed by the demon of Competition and has all but lost sight of the spirit of Co-operation. In order to make the best, quickest and most profitable use of knowledge for humanity, it must always be shared as freely as the oxygen we breathe. Just think what would have happened if Newton patented gravity or Watson and Crick the double helix. 

A Database Nation?

Security and New Challenges to Privacy

Minhaj Siddiqui

When you enter a department store, a camera pointed at the door most likely records your face. Imagine a world in which this surveillance doesn't end there. A computer matches your face to your name using records that the Department of Motor Vehicles sold to commercial interests. It then searches for your name in databases obtained from major credit card companies to see what your consumer preferences are and to check your credit history. Your income bracket can be estimated by searching for your neighborhood of residence in a phone directory. With all this information at hand, salespeople can now decide not only whether it is worth their time to approach you, but what products to push at you. The same system can consult police records to check if you have broken the law in any way, ranging from murder to an unpaid parking ticket. The police can be notified at that point to arrest you.

All this because a camera is pointed at the doorway.

Granted, talk of privacy issues tends to take on a tinge of paranoia, but the scenario may become reality in the near future. The potential for abuse of such invasive technology is great. With so much to gain from increased knowledge of its consumers, commercial interests are sure to take advantage of consumer profile databases as they are compiled. Similarly, other industries such as health care providers and even the government have much to gain from the knowledge found in these databases.

Similar issues are raised by Simson Garfinkel in *Database Nation: The Death of Privacy in the 21st Century.* In this book, Garfinkel, an MIT and Columbia graduate with a strong background in computer privacy issues, illustrates the dangers associated with a society that does not act to control abuses in invasion of privacy. He then calls for action from society to avoid such a future. The concept of privacy is important to define in this discussion. Many will hear talk of protecting privacy and feel as if they need not partake because they have nothing to hide. The privacy being discussed here is of a very broad nature encompassing the characteristics that allow individuals to

distinguish themselves from others. The threat to privacy is not a loss of privacy in the sense that criminals can no longer run illegal drug cartels from inside VCR repair shops; it is the loss of the right to keep any type of knowledge, whether it be a disease or a political thought, away from those who would use that information against the individual. This loss of privacy under the most extreme circumstances would allow HMOs to deny service because the person's complete medical history is an open book. Further, HMOs could access that person's DNA profile, obtained when they donated blood, with its interpretation as to what ailments may later affect the individual.

The threat, as described by Garfinkel, only grows from there. A loss of individual privacy may not only mean others knowing too much, indeed maybe even more than the person knows about him or herself, but as the use of this knowledge becomes more universal, a basic database entry error could destroy a person's lifestyle. Even today there exist a few very central information-gathering services that have the sole purpose of compiling large databases about people. Key examples are the three major credit report bureaus, Equifax, Experian, and Trans Union, all of which maintain extensive databases of credit information (credit cards, bank accounts, loans, etc.) for everyone in the United States. If someone develops a "bad credit rating," these are the companies from which that knowledge is found. In the field of medicine, the Medical Information Bureau (MIB) is in charge of maintaining a database of all major diagnoses that a person has had in his or her lifetime. Each person has a series of codes associated with his or her name, with diagnoses ranging from heart problems or AIDS to codes indicating dangerous lifestyles such as "adverse driving records, hazardous sports, or aviation activity."

Now suppose that some day a person's record, out of the millions that pass through these corporations every day, has an error: instead of "stomachache," the code for HIV was recorded. In a future of no privacy, this person may encounter discrimination, having sudden difficulties obtaining any type of health or life insurance or being harassed in the workplace. Management may refuse a promotion for fear that he or she will not live much longer and thus would be a poor investment to train. Other companies may refuse to hire the person for similar reasons. If this person's career requires continuous and close interaction with others, those who could now very easily know the medical condition of this individual may stop coming to this person because of the stigma sometimes associated with AIDS. All this because of an inputting error.

To exacerbate the problem, it is very difficult to change negative reports once established, even if caused from a clerical error. In some cases, people do not even have access to their own files, cannot find out what is being said about them, and thus cannot even check whether erroneous data is being recorded.

After presenting the problems that a future without privacy may pose, Garfinkel presents possible ways to combat these problems. Privacy-conscious individuals should avoid volunteering information to data-collecting companies by not participating in random sweepstakes that require them to supply a large amount of personal information for the remote chance of winning a prize. Other proposed solutions to protect consumers involve very large-scale national actions with calls to the government urging careful consideration of the issues and the passing of legislation to protect individuals. For example, laws could be passed forbidding any type of database to withhold information from the individual if other corporations use it to judge the person. Similarly, legislation could establish formalized procedures for fighting data that may be recorded incorrectly in some database.

Loss of privacy does not have to be an inevitable result of technological advances in society. This is not to say that privacy must be maintained at all costs; loss of some privacy can be helpful, such as increased ability of police forces to capture and prosecute real criminals. However, especially in the case of commercial interests amassing information about people, a balance needs to be established. It is important that we are cognizant that our privacy can be infringed upon and that the consequences of uncontrolled invasion are indeed great. Act accordingly.

On the Structure & Discovery of Facial Processes: Hidden Markov Models

Development through Entropy Minimization Algorithms

Jane H. Maduram

Ultimately, the aim of science is to understand systems in a micro and macro sense. This broadly involves deciphering the structure of and the relationships between hidden processes and observable events in a given system, which may be done most succinctly by analyzing data sets. As data sets are products of the original process, they indirectly reveal the structural components and processes within the process that contribute to specific outcomes. A mathematical model of a data set must manipulate information in a manner similar to that found in the original process in order to accurately replicate the data set, resulting in an involuntary mirroring of the internal structure of the original process.

An efficient model must, in addition, trim the mechanisms such that no unnecessary commands or steps are made to produce a compact model. The most compact model achievable, according to a theorem in Kolmogorov complexity, will be the best predictor of a complex process, with probability one.[1] While the task of finding a compact model may seem a relatively casual problem, there is no method for deriving the most compact model possible. The closest approach available is to search restricted spaces for a compact space via algorithms. This method becomes increasingly limited, however, as the complexity of the model increases.

$$P(A_i \mid A) = \frac{P(A_i)\,P(A \mid A_i)}{\sum_{j=1}^{N} P(A_j)\,P(A \mid A_j)}$$

The face is a case of a complex space that is extremely difficult to model. Facial processes encode speech and emotion, two processes not fully understood by psychologists or physiologists. This adds to the difficulty of determining which mechanisms operate in the face. In order to address

References

1. Vitányi P & Li M. Minimum description length induction, Bayesianism, and Kolmogorov complexity. Submitted to IEEE Transactions on Information Theory, 1997.
2. Ekman P, Huang T, Sejnowski T, and Hager J. (Eds). Final Report to NSF of the Planning Workshop on Facial Expression Understanding. Technical report. National Science Foundation, Human Interaction Lab., UCSF, 1993
3. Blair C. Interdigitating muscle fibers throughout orbicularis oris: preliminary observations. Journal of Speech and Hearing Research, 1986; 2: 256-'6.
4. Blair C & Smith A. EMG recording in human lip muscles: can single muscles be isolated? Journal of Speech and Hearing Research, 1986; 2: 256-'6.
5. Gentil M. Variability of motor strategies. Brain Lang (B5H) 1992; 42(1): 30-7.
6. Fisher CG. Confusions among visually perceived consonants. Journal of Speech and Hearing Research, 1968; 11: 796-804.

this difficulty, this preliminary study investigates a novel method of approximating facial processes to examine both the feasibility and success of current models in structure discovery of the face in contrast to an HMM, which gives biological impetus to the mathematical claim supported by Kolmogorov complexity that the most compact model imitates the internal structure of the original process.

$$P_e(\theta) \propto e^{-H(\theta)}$$

Background

The face contains 44 muscles that play a direct role in displaying emotion.[2] Muscle types present in the face include linear, sheet, and sphincter muscles, each of which respond differently to stimuli. Muscles in the face are highly integrated,[3] making it extremely difficult to accurately stimulate or record from individual muscles through nonintrusive or intrusive means.[4] An additional problem is that the sets of muscles working in conjunction for specific actions may vary from person to person,[5] further complicating the task of determining what simulated muscles should be used for a specific task within a general model.

The elemental mechanics of speech are summarized by phonemes, visemes, and coarticulation. A phoneme is a unit of speech that is audibly unique. A viseme[6] is a unit of speech that is visibly unique. Complications arise when attempting to establish a correspondence between visemes and phonemes. The mapping from phonemes and visemes is many-to-many.[7] This poses problems in systems that simplify the relationship to a one-to-one correspondence such as keyframing,[8] as it ignores one of the most fundamental parts of speech: coarticulation. Coarticulation is the act of physically preparing for a phoneme before or after the word is actually spoken. Visemes produced by phonemes are context-dependent and actively change, depending on the phoneme(s) preceding (carry-over coarticulation) and following (anticipatory coarticulation) it.[9,10,11] The rules defining coarticulation are further complicated by the varying roles of dominance that phonemes possess.[12]

The workings of emotion are difficult to define. Psychological evidence shows that the dynamics and interactions of facial movement are more important than the static expressions themselves.[13,14,15,16] This concept is dramatically shown in the contrast of anger and happiness, where motion is key in distinguishing the two emotions.[17] As with speech, facial movement must be considered within context.

FACS

The Facial Action Coding System (FACS) has long been the only method available for fragmenting facial expressions.[18] The quantitative aspects of the FACS have evolved in two different directions, each direction coping differently with the same qualitative criterion. One approach effectively bypasses the quantitative nature of most facial animation by focusing on appearances rather than on the mechanics of facial change. The face is simplified to a texture map that may be stretched at action centers, points that the FACS specifies as being primary to movement. The other evolved approach works with the muscles directly connected to these action centers to create a quantitative model. The selected muscles are then added to a model that integrates functions for the bone, muscle, fat, and skin, thus creating a highly realistic and individualistic model that sacrifices a reasonable number of calculations in return for high anatomical realism. In both cases, FACS forces an emphasis on local facial movement.[19] The incompatibility of these two models may reflect in some part the present inability of the computer sciences to satisfactorily resolve the qualitative nature of the FACS.

Algorithm

In place of the qualitative techniques used by the FACS to approximate facial movements, an algorithm was used during this study to train the Hidden Markov Model, which was then compared against the FACS-dependent model. The model was based on Bayesian probability, a statistical procedure that estimates the parameters of an underlying distribution based on the observed distribution. Bayes's rule is stated in Equation 1. The parameter estimate of Bayesian probability requires the input of a prior distribution, customarily a Gaussian distribution, from which a posterior distribution is derived. The posterior distribution is the normalized product of the likelihood and the prior distribution, and the parameter estimate is the mode of the posterior. While Bayesian analysis is efficient, it is controversial because the validity of the parameter estimate relies entirely on the validity of the prior distribution, which cannot be assessed statistically.

The algorithm used[19] modifies standard Bayesian elements in three integral parts: the prior, a set of maximum a posteriori (MAP) parameters, and trimming functions. While most standard Bayesian priors are typically set to an uninformative value on the premise that the data alone should affect the posterior, the algorithm

7. Owens E & Blazek B. Visemes observed by hearing-impaired and normal-hearing adult viewers. Journal of Speech and Hearing Research, 1985; 28: 381-393.

8. Parke F. Parameterized modeling for facial animation. IEEE Computer Graphics and Applications, 1982; 2(9): 61-68.

9. Henke WL. Dynamic articulatory model of speech production using computer simulation. Ph.D. dissertation, Massachusetts Institute of Technology, 1966.

10. Huffman MK. Patterns of coarticulation in English. UCLA Working Papers in Phonetics, 1986; 63: 26-47

11. Whalen DH. Coarticulation is largely planned. Journal of Phonetics, 1990; 18: 3-35.

12. Fowler CA. Production and perception of coarticulation among stressed and unstressed vowels. Journal of Speech and Hearing Research, 1981; 46: 127-139.

13. Bassili JN. Facial motion in the perception of faces and of emotional expression. Journal of Experimental Psychology 1978; 4: 373-379.

14. Bassili JN. Emotion recognition: The role of facial motion and the relative importance of upper and lower areas of the face. Journal of Personality and Social Psychology, 1989; 37: 2049-2059.

15. Bruce V. Recognizing faces. London: Lawrence Erlbaum Associates, 1988: 1-50.

argues that an uninformative prior underestimates the intrinsic knowledge present of what learning is defined by and of what defines a good model. In order to account for this, an entropic prior is introduced (Equation 2) such that Pe represents a partial differential function with an entropy term and represents the parameters of the model. This allows the prior to push parameters to strong, information-containing parameters. In contrast to Maximum Entropy methods, which search for the most noncommittal model that fits the data, the entropic prior forces a sparse, structured, and deterministic model for the data.

The maximum a posteriori estimator produces parameters that maximize the probability of the model given the data. The posterior, given the entropic prior, is Equation 3. To determine the MAP of the posterior, the derivative of the

$$P_e(\theta) \propto P(w|\theta)P_e(\theta)$$

$$\theta_i = \frac{-w_i}{W\left(-w_i e^{1+\lambda}\right)}$$

log-posterior is set to zero. This ultimately equals Equation 4, where W is the Lambert W function.[20] In terms of entropy, the MAP estimator is Equation 5. These estimates drive irrelevant parameters to zero. While it may be viewed as a coarse-grade process-the estimator may overlook patterns in small data sets, when patterns may be viewed as incidental-the estimator is able to reliably determine structure in large data sets by selecting the strongest hypothesis consistent with the data rather than the most unbiased model. Whereas the MAP estimator gradually reduces selected parameters to zero, trimming allows a rapid elimination of those selected parameters. Trimming is typically completed at or near a convergence when a loss in the likelihood is balanced by a gain in the prior. Its purpose is to produce a matrix that consists of independent parameters. In entropic training, parameters are trimmed when a change will increase the entropy faster than the log-likelihood. Trimming operates as a driving function towards a skeletal model whose surviving parameters become increasingly well-supported and accurate. In honing the matrix, the parameter subspace is transferred to a simpler dimensionality of geometry, where it retrains.

Training is especially useful when used with the Markov Model (MM), where the probability of a future occurrence is not changed by additional knowledge, given the exact state that the model is in.[21] The MM is a mathematical function that defines states and transitional probabilities between those states to model processes. Transitional probabilities determine the path of the model from one state to another, regardless of previous occurrences, allowing the MM to retain context. In irregular and HMM, the latter of which is used in this study, paths between states can be eliminated during training. This allows the removal of states. Following entropic training results in an HMM that is compact and sparse, with an increased capacity to carry contextual information and an increased specificity between states. The model also encourages a dynamically simple model of data in contrast to a statically simple model. Due to its construction, the algorithm functions best as a method of structure discovery; its strengths are prediction and classification.

Procedure

Landmark features were painted on the face to outline the hairline, eyebrows, nasolabial line, and lips. The data was taken from a single person reciting "The Jabberwocky" by Lewis Carroll in 55 seconds over 1,640 frames. The reading was taped onto a broadcast quality tape at 29.97 fps. The features were then tracked by a program based on Hager's SSD texture-based tracker.[22] Spring tensions were assigned to edges connecting a pair of trackers, thus enabling the program to correct itself when trackers blundered due to rapid movement or temporary invisibility of the landmark feature. Landmark features outlining the three points on the hairline and the bridge of the nose were used to filter out movements of the face and/or the video camera. Visual data consisted of coordinates, each with a corresponding velocity. The velocity was added to preserve the dynamics of facial expression. Sounds were processed by a mix of LPC and RASTA-PLP

$$-\max_{\theta} \log P_e\left(\theta|w\right) = \min_{\theta} H(\theta) + D\left(w\|\theta\right) + H(w)$$

16. Bruner JS. & Tauguiri R. The perception of people. In Handbook of Social Psychology. Houston: Addison-Wesley, 1954.
17. Minsky M. The Society of Mind. New York: Simon and Schuster Inc., 1985: 1-336.
18. Ekman P & Friesen W. Manual for the Facial Action Coding System. Palo Alto: Consulting Psychologists Press, 1978.
19. Brand M. Structure learning in conditional probability models via an entropic prior and parameter extinction. Neural Computation (in press) 1998.
20. Essa IA & Pentland AP. Coding, Analysis, Interpretation, and Recognition of Facial Expressions. MIT Medial Laboratory Perceptual Computing Section Technical Report No. 325, 1995.
21. Corless RM, Gonnet GH, Hare DEG et al. On the Lambert W function. Advances in Computational Mathematics, 1996; 5: 329-359.
22. Taylor HM & Karlin S. An Introduction to Stochastic Modeling. Boston: Academic Press, 1984.
23. Hager G & Toyama K. The XVision system: A general-purpose substrate for portable real-time vision applications. Computer Vision and Image Understanding, 1997.
24. Hermansky H & Morgan N. Rasta processing of speech. IEEE Transactions on Speech and Audio Processing, 1994; 2(4): 578-589.

features, which is known to be robust to environmental variations.[23] In addition to this, amplitudes, pitch, frequencies, and energy parameters of the first three formants were used.

$$Sx'_n(t) = Sx_n(t) - \frac{\sum_{\tau=1}^{T} Sx_n(\tau)}{T}$$

$$Sy'_n(t) = Sy_n(t) - \frac{\sum_{\tau=1}^{T} Sy_n(\tau)}{T}$$

An additional set of recordings were performed on a 12-second recording of the phonemes "ooh," "aah," "shh," and "ssss." Five points were outlined around the lips and on the jaw. Three hundred frames were obtained at 25 frames per second, and energy channels with 9PLP bands were analyzed to provide audible information. HMM models developed with and without the algorithm were trained on the data, then compared. Once the data set was documented, it was subjected to Principal Component Analysis (PCA) to remove redundancies. The remaining dimensions were selected to preserve approximately 90 percent of the variance, simplifying the data to increase the probability of producing a statistically significant HMM.

The data was split into a visual data set and an acoustical data set. The visual HMM was split into two primary components: a finite state machine modeling the dynamics of the face and a set of Gaussian distributions that isolated and associated facial configurations specific to those states. The facial dynamics were integrated with processed modeling from the acoustic data set to create an HMM that responded to vocal and facial input. The animation utilizing the vocal/facial model uses Candide, a model that contains 100 triangles with most of the action units stipulated by the FACS[24]. It is considered to be a standard in the model-based image coding community.[25] The HMM worked directly with the points, avoiding the action units. The FACS-dependent model relied on linear algebra to manipulate individual action units, approximating the dissection of phrases into phonemes and visemes.

Error estimation was performed in one of two ways. Ambiguity of each state, resulting in a confidence estimate, was calculated by the average possible states per frame. Actual error was calculated using the normalized coordinates Sx'n(t), Sy'n(t), Rx'n(t), and Ry'n(t), where these functions are defined by Equations 6 through 9. Here, coordinates Sx'n(t) and Sy'n(t) represent the coordinates of the training data at the nth synthesized point of frame t. Rx'n(t) and Ry'n(t) represent the coordinates of the trained model, HMM or FACS. T is the total number of frames. The normalized error for p points is shown in Equation 10.

Results

$$Ry'_n(t) = Ry_n(t) - \frac{\sum_{\tau=1}^{T} Ry_n(\tau)}{T}$$

Of the 42 action units proscribed by the FACS, 11 are defined in CANDIDE. Of these 11 action units, only seven were needed to describe the range of motion present in video. The existing points for CANDIDE did not correlate precisely with the landmark features used, so points were added to precisely match landmark features. These action units were modified so as to be controlled by these added points. The average error derived for the FACS was .6881.

The algorithm-developed HMM trained with five facial points was representative of the larger

$$Rx'_n(t) = Rx_n(t) - \frac{\sum_{\tau=1}^{T} Rx_n(\tau)}{T}$$

25. Rydfalk M. CANDIDE: A parameterized face. Ph. D. Thesis, Linköping University, Department of Electrical Engineering., October 1987.

26. Li H, Roivainen P & Forchheimer R. 3-D motion estimation in model-based facial image coding. IEEE Transactions in Pattern Analysis and Machine Intelligence, 1993; 15(6): 545-555.

27. Schwartz R, Chow Y, Kimball O et al. Context-dependent modeling for acoustic-phonetic recognition of continuous speech. IEEE International Conference on Acoustics, Speech and Signal Processing, 1985; (): 1569-1572.

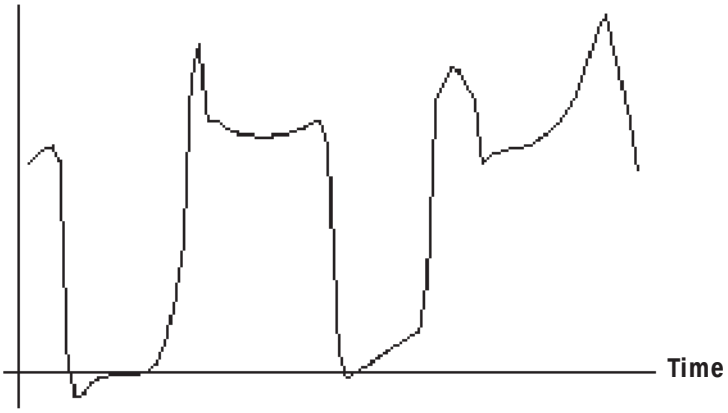
28. Lee KF & Hon HW. Speaker-independent phone recognition using Hidden Markov Models. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1989; 37:1641-1646.

$$\frac{\sum_{\tau=1}^{T}\sqrt{\sum_{i=1}^{p}\left(\left(Sx'_i(\tau)-Rx'_i(\tau)\right)^2+\left(Sy'_i(\tau)-Ry'_i(\tau)\right)^2\right)}}{\sum_{\tau=1}^{T}\sqrt{\sum_{i=1}^{p}\left(\left(Sx'_i(\tau)+Rx'_i(\tau)\right)^2+\left(Sy'_i(\tau)+Ry'_i(\tau)\right)^2\right)}}$$

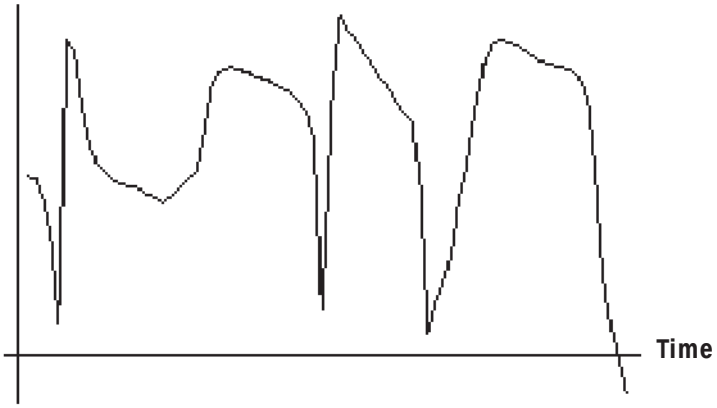
HMMs trained on the entire face. Training data consisted of alternative x and y coordinates of individual points, as demonstrated by the point found on the jaw. This is displayed in Table 1. The movement of the jaw shown in Figure 1 demonstrates the unique properties of the visemes "ooh," "aah," "shhh," and the sibilant "ssss." Because the 12-second training sequence contained four phonemes, a relatively small 8-state HMM combining auditory and visual information was used to model the 20-dimensional observation vector. The 34 transitions in the sequence produced an average entropy of 0.0204. Each state had an average of one "plausible alternative." Conventionally trained HMMs produced an average entropy of 1.296, allowing each state in the sequence 1.8 "plausible alternatives." The visemes clearly visible in the composite graphs of the face (see Figure 2) were effectively synthesized by the algorithm-developed HMM utilizing five landmark features. As Figure 2 demonstrates, horizontal movement of the mouth is more subtle than vertical movement of the mouth and jaw, leading to a necessity of greater detail around the face. This need was addressed in a more extensive series of experiments. (Figure 1, showing movement of the jaw separated into horizontal and vertical components; Figure 2, showing the horizontal and vertical movement of the points surrounding the lips and the jaw).

HMMs integrating 21 landmark features of the face adopted a total of 315 HMMs to the upper and lower parts as well as the entirety of the face to investigate the efficiency of face-to-voice training. The number of states used over these trials varied between 18 and 24 states. Training was completed for 105 different HMMs on the upper face. The most efficient HMM developed contained 23 states, resulting in the low error estimate of .1865. Each state in the optimal sequence had an average of .033 possible alternatives. This resulted in a confidence estimate of 96.7 percent. The lower face, as well, was trained 105 different times. The average HMM among these contained 24 independent states.

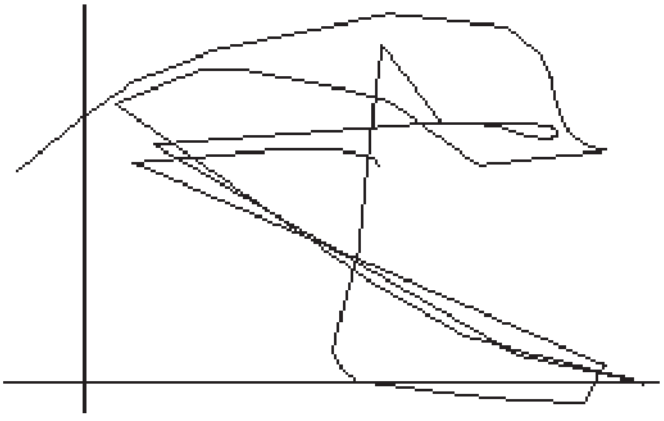
Point 1: Vertical Movement

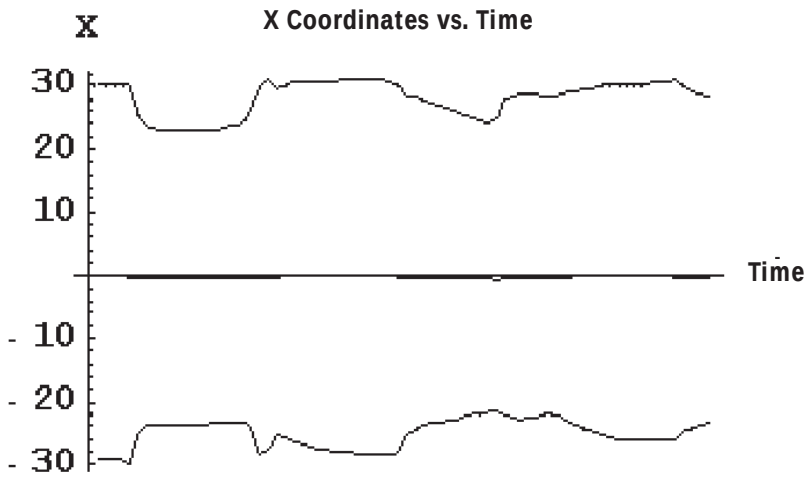


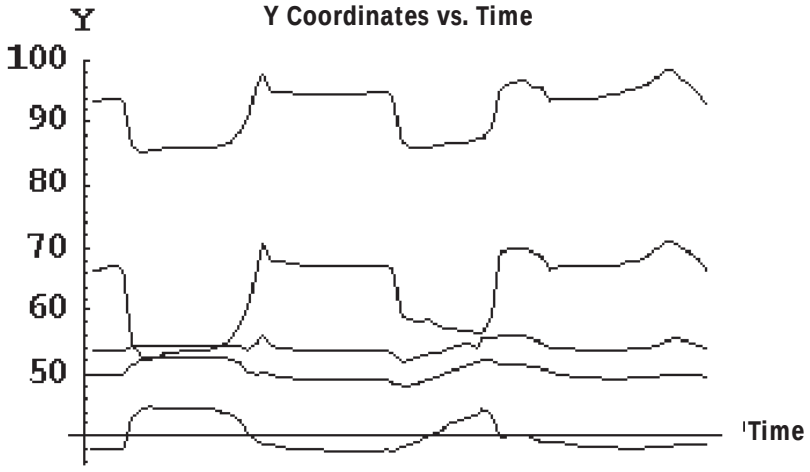
Point 1: Horizontal Movement



Point 1: Composite Movement







The error produced by the average HMM was .3289. Each state had, in the optimal sequence, an average of .052 possible alternatives, producing a confidence estimate of 94.2 percent.

The entire face was, once again, trained 105 different times. The smallest average error of .3618 was obtained with an HMM that was composed of 23 states and five dimensions. Each state possessed an average of .042 alternatives from the optimum path of probabilities, allowing a confidence estimate of 95.8 percent. Statistical analysis of variance in the holistic model was p = 10-5.

Discussion

The compactness of any model is based on the amount of data needed to run it. In the case of the FACS, the action units needed, the extent to which each will be used, the direction of the move, and the duration of the movement must be specified for each viseme. This large amount of information required reveals that the FACS-derived model is not compact. In contrast, the HMM model developed from the algorithm requires only two sets of time-independent information to simulate a given sequence: initial probabilities and transitional probabilities. Given these, it is possible for the HMM to perform a fast and efficient sequencing of facial processes. While the HMM uses an average of 23 states in contrast to the FACS-derived model's use of seven action units, this does not impact the complexity of the FACS vs. the HMM as the units are not equivalent. Each state of the algorithm-derived HMM compresses two or more fully defined action units into simpler terms of coordinate movement and velocity, allowing for a compact representation of the same material in flexible terms dependent only on the vocal and facial peculiarities of the subject.

Compact models of a process should imitate the internal structure of the facial processes. This is revealed primarily through three major components of muscles: time, subtlety, and interdependencies. The FACS-derived model stresses static expressions by ignoring the time component of facial expression. This implies that facial expression is solely dependent on the position of facial features when, in fact, motion plays a significant role in distinguishing between emotions, as stated before. Sequencing and timing of expressions are especially critical to the recognition of emotions. These facts are acknowledged by the HMM, which includes velocity in its assessment of states. By doing so, it bypasses the inability of the FACS-derived model to deal with the timing of the separate types of muscles found in the face: linear, sheet, and sphincter. This ability of the HMM to specify the velocity as well as the position of landmark features promotes natural movement as velocities are assigned to the appropriate regions of the face within the proper context.

In addition to neglecting the dynamics of muscle motion, the FACS neglects to address the subtlety of muscle motion. The FACS stresses large, local, and isolated spatial patterns. It is unable to describe fine eye or lip motions, concentrating primarily in gross motion. In addition, it is unable to predict the physiological effect that muscles have on each other. The need for fine detail is emphasized by the lip movement shown by the 8-state HMM, which demonstrates the fine detail needed in differentiating phonemes. The HMM predicts the face holistically, defining each state for the entire face-not for a portion of it-so logical side-effects are encoded along with the intended movement. In addition to simplifying simulation, the trained HMM unobtrusively

absorbs mechanisms, hidden or known, that regularly correspond to given movements.

The inability of FACS-dependent models to address holistic movements is displayed prominently in the discussion of coarticulation. Since coarticulation, as previously mentioned, involves the modification of visemes based on its context, it poses a uniquely difficult problem for FACS-derived models. The need for context has been well documented for its ability to assure natural movement, but the FACS is unable to provide for context due to its emphasis on static poses. [26,27] The HMM, however, is able to meet this need. By definition, transitional probabilities preserve context within HMMs by determining the circumstances under which a designated state should appear. As the decision of a succeeding state is entirely independent of the preceding sequence, context is the only deciding factor that affects the choice of the successor. This property of HMM allows it to effectively mimic the mechanics of facial processes. The entropy rates of the HMMs developed indicate that context is used both before and after designated sequences in a time frame similar to that found in facial and visual coarticulation. The FACS-derived model copes by assigning elaborate algorithms and dominance rankings to deal with coarticulation (MIKETALK) or by simplifying the viseme-to-phoneme relationship to a one-to-one arrangement. Those methods, however, cannot overcome this internal defect inherent in the FACS.

Qualitatively, the velocity trajectories developed for each state were smooth and similar to natural facial movement. Resulting transitions from state-to-state were geometrically and temporally well-behaved.

A compact model of a process that contains mechanisms intrinsic to an original process must be able to extrapolate sets consistent with output from that process. The FACS-derived model does not contain many of the mechanisms found in the face, and it has a correspondingly high error estimate of .6881. The HMM, however, contains some of the mechanisms found in the face, and has an overall low error. The error estimate for the entire face, .3618, was the sum of the error estimate of the lower face (.3289) and the upper face (.1865). The confidence estimate for states was 95.9 percent for the entire face, which was the average of the lower face (94.2%) and upper face (96.7%) confidence estimates. On the whole, these low estimates confirm that the common mechanism found in the face and the model produce similar data sets, as would be expected.

Conclusions and Open Questions

Previously, Kolmogorov complexity supported the claim that a compact model is capable of reproducing the internal structure of a process. This was validated by the study, which verified the quantitatively superior capability of the HMM model in accuracy, compactness, and pattern to a model using standard encoding sequences (the FACS). The developed HMM allows for a greater understanding of facial processes and a viable process for quantitatively simulating the face. Further experiments extending the range of subjects and expressions may be useful.

The HMM model has numerous applications. By determining the structure of the face, it may allow an understanding of the neuromuscular structure and additional hidden mechanisms of the face. This may prove useful in cases of facial paralysis, such as that found in strokes and in Moebius Syndrome. In addition, it may provide further information as to the correlation between vocal information and facial expressions. Its unique probability distributions specific to training data and facial peculiarities will also allow further applications in data compression.

SPHERES:

A Testbed for Spacecraft Formation Flight Research in Microgravity Conditions

Allen Chen and Stuart Jackson

Space-based telescopes require large, heavy mirrors to deliver the angular resolution and light-collecting power necessary for applications such as interplanetary observation or ground surveillance. The higher the resolution desired, the larger the primary mirror, or aperture, must be. Large mirrors are difficult to deploy in space. The size and weight of the spacecraft scale with the size of the mirror, making it prohibitively expensive to develop, manufacture, and launch large single-aperture telescopes. Instead of developing these expensive single-aperture telescopes, both NASA and the Air Force are studying the use of separated spacecraft interferometers, which could deliver the performance provided by very large single-aperture telescopes. These systems use a number of smaller, cheaper telescopes that work collectively to synthesize a much larger aperture at a great reduction in overall system cost and in individual spacecraft size and complexity. Two missions designed to use space based interferometers or distributed arrays are shown in Figure 1. The resolution provided by an interferometer is a function of the distance between the apertures, or baseline: The longer the baseline, the higher the resolution. As a result, it is desirable to have as long a baseline as possible. One possible implementation would be to structurally connect the apertures; however, this would result in large supporting structures and a complicated deployment system. Additionally, large structures tend to vibrate, especially in space, where damping sources are not noticeably present. Vibrations degrade the performance of interferometers, since high precision optical alignment is needed.

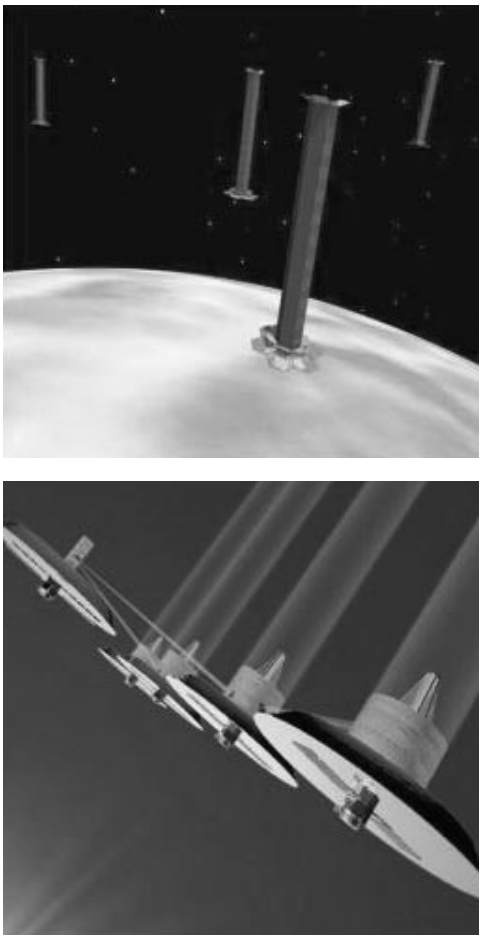


Figure 1 The Air Force's TechSat 21 Mission (top) and NASA's Terrestrial Planet Finder Mission (bottom) hope to utilize space based interferometery and distributed arrays

By contrast, interferometers made up of separated spacecraft have simplified dynamics, since structural vibrations have been eliminated. They also have increased system redundancy. Because of the high optical precision needed, a separated spacecraft interferometry system would require precise control of both the position and the attitude of all the constituent spacecraft. To achieve the required control precision, new formation flight control and metrology algorithms must be developed and tested. SPHERES is a formation flight testbed

developed over the past two years by the Space Systems Laboratory and the Department of Aeronautics and Astronautics at MIT. The SPHERES program is driven by the need for testbeds to validate algorithms pertinent to formation flying spacecraft. The testbed provides a low-risk, cost-effective, long-duration, replenishable, and easily reconfigurable platform with representative dynamics for the development and validation of algorithms relevant to formation flight. More specifically, SPHERES will allow testing of (1) relative altitude control and station-keeping between satellites, (2) retargeting and image plane filling maneuvers, (3) collision avoidance and fuel balancing algorithms, and (4) array geometry estimators. The use of SPHERES in a microgravity environment closely resembles the dynamics of future missions, except for the orbital dynamics. Still, it does allow the designers of the missions to validate the algorithm design, initialization, debugging, and refinement process. The validation of control algorithms and their development processes is an essential step in reducing the risk of future formation flight missions.

Figure 2 Stanford formation flight testbed using blimps

Other Research in Formation Flight

Currently, a number of formation flight tools and testbeds are being developed and/or operated by a variety of groups in government, industry, and university settings. These testbeds have different strengths and weaknesses in their ability to aid the development of formation flight technologies. The SPHERES project looks to complement these other testbed development programs. NASA's Goddard Space Flight Center (GSFC) is developing a computer-based formation flight simulation tool.[1] The tool is based on the commercial-of-the-shelf (COTS) product VirtualSat Pro, produced by the Hammers Company. VirtualSat is a real-time simulator that can be used to model a spacecraft formation or array. This software-based testbed allows easy expansion and refinement to the simulation, while also providing a cost-effective platform for developing formation flight algorithms; however, it relies solely upon software simulation to accurately depict the dynamics of the system and space environment.

Stanford University's Formation Flying Testbed (FFTB) is a ground-based testbed that consists of three free-floating vehicles that move on a granite table by firing compressed air thrusters.[2] The vehicles are autonomous and self-contained. The testbed was created to study guidance navigation and control issues by simulating microgravity dynamics in a 2-D plane; the 2-D nature of the testbed limits its ability to capture dynamics that are closely representative of space-borne vehicles.

A second Stanford University ground-based testbed allows demonstration of formation flight in three dimensions by using blimps as autonomous, self contained vehicles.[3] This testbed, shown in Figure 2, has been used to demonstrate robust Carrier-Phase Differential GPS (CDGPS) formation sensing and investigate basic formation control strategies; these investigations are limited by the blimp dynamics, which can vary widely from the dynamics of space-borne vehicles.

The New Millennium Program EO-1 mission, scheduled for launch in 2001, will demonstrate autonomous navigation and formation flying between EO-1 and Landsat.[4,5,7] The mission will validate autonomy algorithms and software tools, provided by GSFC, the Jet Propulsion Laboratory, and the Air Force Research Lab, that are key to formation flight. EO-1's testing of various algorithms in space provides realistic dynamics and results that can be directly applied to future space missions; however, EO-1's imaging mission limits the types of maneuvers that can be performed.

The Orion mission, currently being developed by Stanford University, seeks to validate key formation flight control and sensing issues with the use of low-cost microsatellites.[6] Orion will provide an on-orbit testbed for the development of autonomy and other distributed spacecraft technologies; however, the space-borne nature of the program limits the lifetime of the mission, increases the risk associated with testing control algorithms, and also increases the cost of developing and operating the testbed.

SPHERES will complement these testbeds by providing a cost-effective bridge between ground-based testbeds, whose dynamics are not truly representative of space-borne vehicles, and space-based testbeds, which are limited by lifetime, configuration, and cost. SPHERES provides full six degree of freedom (DOF) dynamics in a microgravity environment, while remaining easily replenishable and reconfigurable.

SPHERES Systems Overview

The SPHERES satellites are self-contained with on-board propulsion, processing, RF communication, and metrology.[7] In addition, the testbed has four metrology transmitters that provide global metrology reference and a laptop computer that acts as a "ground station" and provides experiment control.

A partially assembled SPHERES unit is shown next to a fully assembled unit in Figure 3. Some panels and hardware have been removed from the partially assembled unit to provide a clearer view of the components. The partially assembled SPHERES unit photo is of a prototype model; many of the cables and hoses have been shrunk and shortened for the flight models.

The SPHERES satellites and related testbed equipment were designed to fit in a Space Shuttle middeck locker, with room for expendables like propellant tanks and batteries. Each individual satellite is roughly the size of a volleyball, but more dense. The physical properties and other specifications for each satellite can be found in Table 1. Many of the purchased components have been modified for SPHERES use. Figure 4 shows and identifies many of the various subsystem components. In order to demonstrate formation flying, a series of maneuvers must be performed. These complex maneuvers can be viewed as a series or combination of less complicated maneuvers. This flow-down continues until each operation is a sum of basic level operations. In order to achieve the ultimate goal of formation flying, each tier of operations must be accomplished by steadily increasing the difficulty of the subsequent maneuver. Below are examples of each category of maneuver.

Basic Vectored Maneuvers
- Maintaining position/orientation
- Point-to-point translation/rotation
- Combination rotation-translation

Time-Dependent Maneuvers
- Tracing out a specified path
- Disturbance rejection

Coordinated Maneuvers
- Maintaining position/orientation relative to another satellite
- Follow the leader

Formation Flying
- Multiple satellites flying in rigid body formation
- Position swapping
- Simulated satellite failure

During the preliminary design phase, a set of "strawman" operations was defined. This strawman, which flows down from the SPHERES Requirements Document 8, defines the minimum requirements for mission operations. These minimum requirements ensure that meaningful maneuvers can be performed within the 20-second limitation of the KC-135:

- The satellites must be able to translate 1 m in 5 s, from start to stop.
- The satellites must be able to rotate 360° in 5 s, from start to stop.
- The tolerances on the satellites are .5 cm on position and 3° on attitude.

Based on these requirements, the subsystem teams completed their initial designs.

Subsystems Overview

The structure is an aluminum frame covered by Lexan panels. The frame provides numerous mounting locations for the various subsystems. The structure was designed using ProEngineer and procured professionally. Figure 5 shows a ProEngineer schematic of the internal aluminum frame accompanied by a photograph of the resulting assembled structure.

The propulsion subsystem consists of twelve solenoid-actuated valves that expel carbon dioxide through machined nozzles. The thrusters are grouped in six pairs of opposing thrusters to provide attitude and station-keeping control. The propellant is stored in Department of Transportation (DoT)-approved steel tanks that hold 74 grams of liquid CO_2 at 860 psig. A regulator drops this pressure to 70 psig prior to distribution via a manifold. The tanks are replaceable and provide between 20 seconds and 30 minutes of operations, depending on the maneuvers. Figure 6 shows an overview of the propulsion subsystem components and outlines how the CO_2 propellant is distributed through the pressure system and out to the various thrusters.

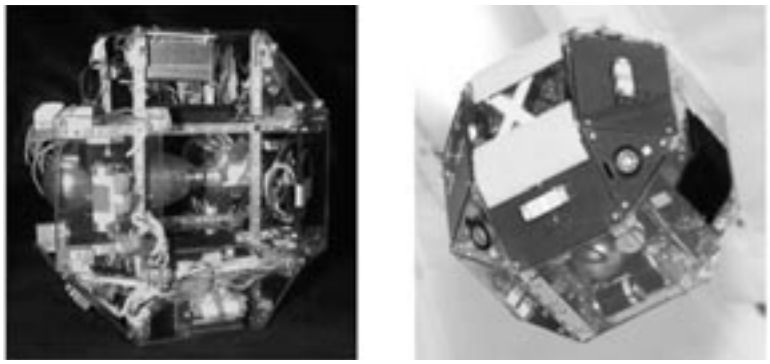


Figure 3 Photographs of partially assembled SPHERE (left) and flight model SPHERE (right)

The metrology subsystem has global and

Table 1

SPHERES SPECIFICATIONS

Diameter	0.2 m
Mass	3.4 kg
Max. Linear Acceleration	0.17 m/s^2
Max. Angular Acceleration	3.5 rad/s^2
Battery Life	60-90 min
Power	6.2 W
Baud Rate	19200 baud
Metrology Resolution	2.0 cm
Tank Life	20 s – 30 min depending on usage

inertial elements. The global metrology system measures the time of flight of 40 kHz ultrasonic pulses emitted by four externally located metrology transmitters. Infrared transmitters provide precise synchronization pulses. Eight ultrasonic microphones distributed on the surface of the satellite are used to derive a total of thirty-two propagation delays, which are then used to derive position and attitude. Additionally, each satellite has an internal inertial measurement unit (IMU) consisting of a three-axis accelerometer and three rate gyros. The independent global and local systems can be configured to work in concert to provide real-time metrology data. The power subsystem consists of two battery packs and four voltage regulators. The battery packs each contain six Duracell Ultra AA alkaline batteries and are located on opposite sides of the satellite. Voltage is regulated to 3.3V, 5V, 12V, and 22V. The 5V and 12V regulators are COTS, while the 3.3V and 22V regulators are custom circuitry. Each SPHERES unit consumes six to nine watts under nominal operation. Given this power consumption, each SPHERES unit has a lifetime of approximately 60-90 minutes.

The avionics subsystem consists of a TI C40 DSP, a Motorola Tattletale computer, a propulsion circuit board, a metrology board, a power distribution board, a UART internal digital communications board, two external RF communications circuits, and eight boards to connect the metrology infrared and ultrasound receivers. The custom boards were designed using OrCAD and procured from professional board manufacturers. The boards are placed around the internal structural supports and are interconnected to one another.

The software subsystem is multirate. The real-time software executes a 1.0 kHz interrupt to actuate the thrusters via pulse width modulation. The control algorithms are updated using a 50 Hz interrupt. State estimation occurs at a rate of 50 Hz from the IMU with 1 Hz updates from global metrology.

The communications subsystem is multi-channel. Satellite-to-satellite (STS) communications as well as satellite-to-ground (STG) communications are updated at one to ten hertz rates. The communications system uses a token ring architecture for both the STG as well as the STS

References

1. Bauer F, Hartman K, How J, et al. Enabling Spacecraft Formation Flying through Spaceborne GPS and Enhanced Automation Technologies. Proc ION GPS-00 Conf; Sept 1999.
2. Corazzini T, Robertson A, Adams J, et al. GPS Sensing for Spacecraft Formation Flying. Proc ION GPS-97 Conf; Sept 1997.
3. Olsen E, Park C, How J. 3-D Formation Flight Using Differential Carrier-Phase GPS Sensors. Proc ION GPS-98 Conf; Sept 1998.
4. http://eo1.gsfc.nasa.gov/Technology/FormFly.html

communications. STG communication is used to archive measured telemetry and to send commands to the SPHERES units.

Operational Environments

The SPHERES testbed is operable in three different environments: a 2-D laboratory platform, the KC-135, and the ISS. Each of these environments provides unique advantages that can be utilized in different phases of the testing of control algorithms with the SPHERES system. The advantages and disadvantages of each environment are summarized in Table 2. Duration refers to the length of the experiment that can be performed in each environment. Contact refers to the level of interaction possible between the algorithm designer and the actual experimental run being performed. Dynamics and DOF assess how well the environmental conditions match those found in a potential formation flying mission. Frame rotation considers the motion of the global metrology frame, the experiment environment, with respect to the inertial frame.

2-D Laboratory Platform

The SPHERES prototypes can be tested on a 2-D frictionless table at the MIT Space Systems Laboratory. This allows for some simulated effects of microgravity while in a 1-g environment. It provides 3 degrees of freedom-2 translational and 1 rotational. The SPHERES system, as well as each individual unit, can be tested before going into the expensive environments of the KC-135 and ISS. First-run tests of control algorithms can also be implemented using the 2-D testbed. It allows for long-duration testing and gives the researchers direct contact with the SPHERES system while running tests. It is a cost-effective device to work out any unforeseen problems in both the SPHERES system itself and in any control algorithms being tested.

The design utilizes pucks that "float" by expelling compressed gas downward onto the testbed's glass surface. The finalized design is an air-bearing levitation vehicle, in which three CO_2 tanks feed three pucks via a single regulator. The pucks lift the vehicle off the glass surface on which it sits. A SPHERE satellite sits atop the vehicle for testing. Figure 7 shows the air-bearing vehicle and a SPHERE structure mounted to the vehicle. The testbed allows for three SPHERES units to operate simultaneously on the ground in a 1-g, 3-DOF environment.

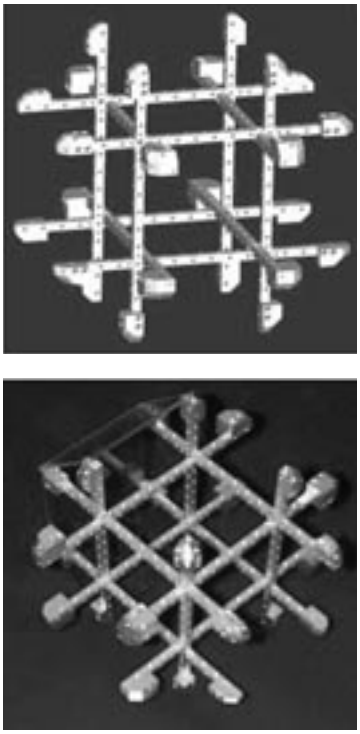


Figure 5 ProEngineer schematic and photo of partially assembled structure

KC-135

Unlike the 2-D laboratory environment, which is limited in its ability to provide realistic dynamics, the KC-135 environment provides six DOF dynamics that more closely resemble real missions, at a fraction of the cost of placing SPHERES on the ISS. Unfortunately, each experiment is limited to the duration of the microgravity segment of one parabola, approximately 20 seconds. Additionally, the parabolic motion of the aircraft causes the fuselage and global metrology frame to rotate ~900 in 30 seconds.

This causes a discrepancy between the global metrology reference frame and the inertial reference frame. Additionally, turbulence and imperfect parabolas due to pilot error provide a very active disturbance environment. The KC-135 provides a cost-effective option for first-run tests and operational checkouts in 6-DOF. The KC-135, shown in Figure 8, has been used to test and verify that the SPHERES system is functional and operational. Tests run verified autonomy, control in 6-DOF, and metrology systems. In addition, basic control algorithms were tested; however, more complicated control algorithms will require more than 20 seconds to verify. The results of the initial testing are discussed below.

International Space Station

Free of the time constraints imposed by the KC-135's microgravity parabolas, the SPHERES testbed will provide the most research capability in the shirt-sleeve environment onboard the International Space Station (ISS). This environment will provide long-term, hands-on, microgravity research. The experiment can run on the ISS for several hours at a time by simply replacing exhausted tanks and/or batteries when needed. The procedures for replacing both propellant

5. Folta D, Newman L, Gardner T. Foundations of Formation Flying for Mission to Planet Earth and New Millennium. AIAA/AAS Astrodyn Spec Conf; July 1996.

6. How J, Twiggs R, Weidow D, et al. Orion: A Low-Cost Demonstation of Formation Flying in Space Using GPS. AIAA/AAS Astrodyn Conf; 1998.

7. SPHERES Design Document, ver. 2.0. SSL Internal Doc; 1999.

8. SPHERES Requirement Document, ver. 2.1. SSL Internal Document; 1999.

9. Miller D, Saenz-Otero A, Wertz J, et al. SPHERES: A Testbed for Long-Duration Satellite Formation Flying in Microgravity Conditions. AIAA/AAS Space Flt Mech Mtg AAS 00-110; Jan 2000.

tanks and batteries are simple and quick. Since astronauts will be working directly with the SPHERES testbed, troubleshooting and maintenance become very simple through direct interaction with the experiment. Algorithm errors or mechanical problems that could cause catastrophic failures in space-based platforms are easily managed in the SPHERES testbed: The astronaut need only reach out to stop and reset the experiment. Additionally, the hands-on

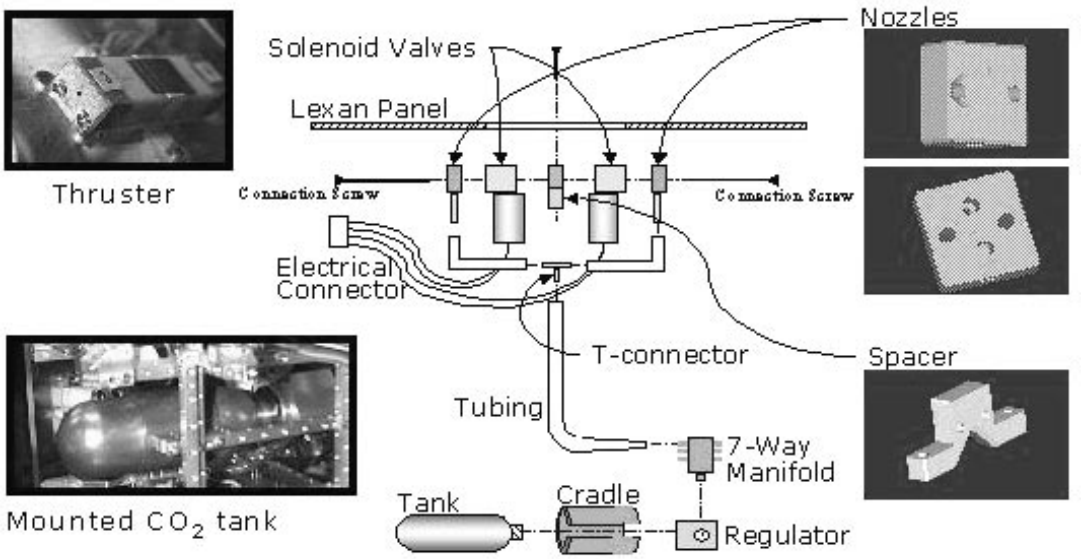


Figure 6 Overview of the propulsion system

nature of the testbed makes downloading data and uploading new control algorithms trivial. The SPHERES testbed is also cost-effective on the ISS because it can be run indefinitely with the resupply of expendables, such as batteries and compressed gas tanks, and is run inside the shirt-sleeve environment of the station rather than outside. While the inside of the station still provides the needed 6-DOF microgravity environment, it eliminates the need for expensive thermal and radiation-hardened spacecraft that can survive the harsh environment of space. The ISS is the best microgravity environment for a formation flying testbed, because it allows all 6 DOF and closely simulates the dynamics of a satellite system, while still remaining an environment where long-term, hands-on research can be conducted.

KC-135 Operational Testing

In February and March 2000, the SPHERES testbed was flown on board the KC-135 to demonstrate the operability of the testbed in a microgravity environment. Figure 9 shows the SPHERES team operating the testbed on the KC-135. The tests performed included assessing the satellites' ability to propel themselves in full 6

Table 2

COMPARISON OF ENVIRONMENTS

	Duration	*Contact*	*Dynamics*	*DOF*	*Frame Rotation*	*Cost*
2-D Platform	Long	High	Average	3	Slow	Low
KC-135	Short	High	Good	6	Fast	Medium
ISS	Medium	Low	Good	6	Moderate	High

DOF, verifying the testbed's ability to obtain metrology readings correctly, verifying the satellites' ability to communicate while rotating and translating in all directions, and demonstrating the ability to operate fully untethered for reasonable periods of time.

The results of the tests proved full operation of the majority of SPHERES systems. The propulsion system demonstrated the ability to maneuver in a microgravity environment, though it sometimes failed to overcome turbulence effects experienced by the airplane. Propellant tanks lasted for approximately ten parabolas, though the exact thrusting lifetime is difficult to calculate given the variety of maneuver profiles.

The metrology system's inertial and global systems were tested separately. The rate gyroscopes were shown to be reliable and sufficiently sensitive to allow rotational control of the satellites in all three rotational DOF. The global metrology tests indicated that the system could locate the 3-D position of the SPHERES, but only within a limited section of the test area. The limitation has been attributed to the field of view of the ultrasound receivers, indicating a need to increase their field of view.

Satellite-to-ground and intersatellite communications were successfully demonstrated, with only a few breaks during tests. Most problems were encountered during the initialization of the communications units, indicating a need to correct the initial synchronization of the communications system. The power, software, avionics, and structures subsystems demonstrated full functionality with virtually no problems. Battery packs lasted in excess of one hour. A single C40 processor was able to command a variety of maneuvers. The avionics system performed well, with the exception of one wiring defect. The SPHERES structures proved able to survive both the microgravity environment at the top of the parabolas and the high-gravity (~1.8 g) pull-out at the bottom of the parabolas.

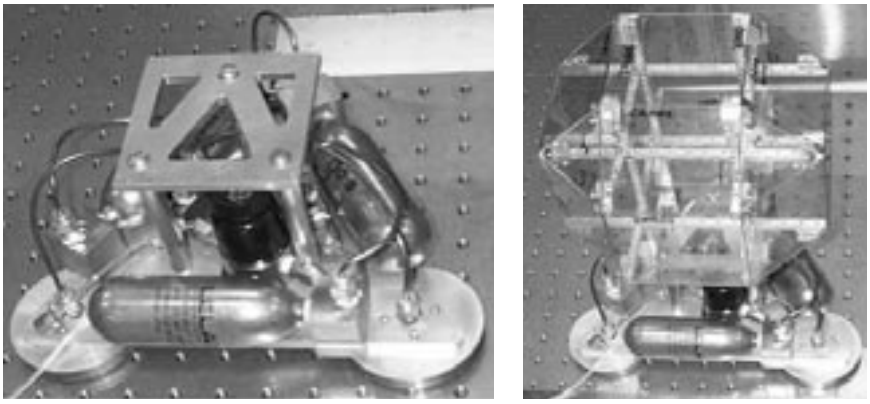


Figure 7 2-D testbed air bearing with and without a SPHERE structure mountd to the top

Limited algorithm testing was also performed using the testbed. A Master/Slave architecture, where one satellite follows the motions of the "master" satellite, was successfully tested to demonstrate the use of SPHERES as a formation flight testbed. Two different tests were performed. In the first, a team member manually maneuvered the master SPHERE and the motion of the slave was observed. The test clearly revealed the "slave" emulating the motion of the master. In the second type of test, the master satellite was attached to the KC-135 frame. Thus, the slave followed the rotation of the KC-135 as it pitched over the top of the parabolas at a moderately high rate. This made the slave SPHERE appear to hold orientation with the rest of the airplane, regardless of turbulence and imperfect parabolas.

Figure 8 NASA's KC-135 Reduced Gravity Aircraft (the "Vomit Comet")

Future Work

Following last spring's KC-135 flights, the project entered a period of transition and refinement where it was handed over by the undergraduates of the CDIO class to the graduate students and researchers of the MIT Space Systems Lab. The metrology, propulsion, and communications subsystems are the primary focus of the SPHERES hardware improvements now underway. In addition, control algorithms from outside agencies such as NASA Goddard Research Center will be tested along with new MIT-developed algorithms. Optical components will also be added to the SPHERES 2-D testbed to aid in evaluating the performance of these algorithms. Additional SPHERES hardware upgrades will also be necessary before the testbed can be flown on the space shuttle or International Space Station in the next two to five years.

MIT SPHERES was developed through a unique educational experiment at the Massachusetts Institute of Technology (MIT). Undergraduate Aerospace Engineering students are exposed to the full life cycle of an aerospace product through Conception, Design, Implementation, and Operation (CDIO) phases. Students not only learn about design, teamwork, and communication, but also interact with potential customers from government and industry, appreciate the constraints of integrating to a carrier, exercise professional computer-aided design tools, and struggle with the iterative process of design improvement and system-wide

Acknowledgments

The SPHERES project is the result of a team effort by graduate and undergraduate students and staff of MIT, with the support of staff from Payload Systems, Inc.

The members of the team are:
Chief Systems Engineer
Prof. David Miller
Undergraduate Students
(Class of 2000):
George Berkowski
Chad Brodel
Sarah Carlson
David Carpenter
Allen Chen
Stephanie Chen
Shannon Cheng
Daniel Feller
Stuart Jackson
Fernando Perez
Brad Pitts
Jason Szuminski
Julie Wertz
Teaching Assistant
Alvar Saenz Otero
MIT Staff
Paul Bauer
Dr. Ray Sedwick
Prof. Dava Newman
Payload System Staff
Dr. Javier deLuis
Steve Sell
Graduate Students
Mark Hilstad
Mitch Ingham
Edmund Kong
Alice Liu

Specificity in Context:
Selected Studies in Architecture and Technology for Education

David Foxe

The architecture of higher education institutions for health care programs is an example of a diverse and highly specialized niche within the architecture profession. Architectural design in this realm encompasses a wide variety of aesthetic, educational, technological, and logistical issues relating to the person in roles as learner, teacher, facilitator, and administrator, as well as patient, doctor, etc…

For eleven weeks in the summer of 2000, I was an intern with the higher education team of Kahler Slater Architects in Milwaukee, WI, which is a 130-person firm with offices in Milwaukee and Madison, WI. Guided by a list of topics I was assigned to research, I organized the existing information from locations within the office, researched information to be incorporated from external sources, and developed a cohesive way to combine the existing and new information in an eminently expandable organizational system. Although the system can constantly evolve and gain new information, this analysis takes a written "snapshot" of the information as it existed in August, 2000. The analysis was originally directed primarily towards the Kahler Slater higher education team while being as understandable and usable as possible for others within and beyond the firm; the original report has been reformatted and shortened to the present form.

A primary idea that permeates these issues of spaces, technologies, and trends are that the designs are characterized by specificity in context; the research files consist of highly specific information that is shown within the context of individual projects and/or locations, and the organization of the files puts this specificity into a more usable context within the specifics of higher education architecture. Consequently, the three major issues that the research files and this analysis address are: (1) design recommendations for health- and education-related spaces, (2) the architectural implications of various technologies, and (3) the larger architectural, business, and educational trends which have an effect on designs for higher education.

One must first understand the business, educational, and planning trends that have an effect within the context of higher education in order to transfer them into specific, meaningful, technology-enhanced health- and education-related spaces.

Business and Education at the turn of the 21st Century

Educational facilities are increasingly following financially lucrative fields such as biomedical technology, internet startups and networking, and materials science. Areas of the US that do not seem receptive to new business innovations will have a hard time keeping graduates and attracting businesses towards a good economy, and are in danger of being bypassed as locations for new and expanding companies. In the case of Wisconsin, the state needs to build on the intellectual capital of a major research university (UW-Madison), to extend regional and local technical training and distance education to local and rural businesses, and to change the state's perception as an old, blue collar economy.

Business trends indicate that there will be fewer workers in standard full-time jobs and that traditional white-collar professions may likely decline in the future. Business is becoming less interested in specific skills and more interested in talent; one can be trained in job-specific tasks, but not in being a better thinker or a more creative problem-solver. Highly specialized blue-collar skills will become more important, however, as white-collar jobs lessen in a virtual world.[1]

Additionally, as federal funds and taxpayer support decrease, institutions develop partnerships with corporations for applied research, sponsorships, endorsements, and naming rights. As research grants go to teams rather than individuals, spaces must facilitate team research both within the university and for working with businesses. These extensive partnerships are hailed by some as the educational establishment finally coming to terms with the real world, while others insist that institutions should not "sell out" to business. With tuition increasing more than twice as fast as inflation and nearly one third of students transferring between institutions, universities are turning to market economy and tuition-driven models where tuition is the student's investment, and the campus is key to being marketable.

The proper role of a business in an educational institution is not yet consistently well defined, so a variety of public/private/academic partnerships have begun to flourish independently. The Center for Environmental Sciences and Technology Management in Albany, NY is a public/private/academic partnership as a research center and marketing facility of products, including offices of the National Weather Service. The light, circular rotunda is the gathering spot for the partners, with discussion spaces spreading to adjacent conference rooms and circulation spaces. The Michigan Virtual Automotive College (MVAC) is collaboration between higher education, the state, training providers, and the automotive industry that has internet-based courses.

In the business world, the line between services and products is becoming increasingly vague in an information-based economy. Architecture is expanding far beyond the design process because firms make money by staying involved during the entire life of a building. According to Richard Hobbs, FAIA, the vice president of professional practice for the American Institute of Architects (AIA), "What we're seeing is a shift in the role played by some architects from one that's project-oriented to one that's relationship-oriented…The new paradigm is for architects to serve clients as facilitators and integrators." Architects are now in an information business, providing continued expert advice before and after the design process. Businesses and educational institutions are also faced with becoming "information enterprises," with long-term clients, especially for people in technology areas whose skills become obsolete in only a few years. Thanks to the Internet, the college or university that an teenager selects can be viewed as the lifetime provider of educational services.[2] "Education will go from being benefit driven, like a manufactured product, to value driven, where the importance of an institution is its core values as a lifelong learning partner.

Educational "economics" are changing from labor intensive to capital intensive, from local to national, and from small to large scale production and evaluation, due to the expansion of distance learning and other programs. Although internet opportunities have high startup costs, the internet allows for low marginal costs of distributing educational content. Universities could lose a "local monopoly," and become part of a larger group of "degree-granting bodies," like the DeVry Institutes that teach technology and business without including dormitories or tenured faculty, and the distance-education-based University of Phoenix that gives adults advanced training but has no campus.[3] Ted Marchese, vice president of the American Association for Higher Education (AAHE), describes challenges of educational economics for universities to include threats to continuing education programs, questions about whether to replace existing, home-grown courses with nationally produced courseware, and seeing enrolled students appear in the registrar's office with brand-name Web course credits.[4]

Meanwhile, socioeconomic changes are reflected in the widespread dissolution of the nuclear family, the growing number of elderly people (who often need and want access to educational and health facilities), and the predicted growth of the wealthy and poor sectors while the middle class shrinks. Educational institutions are faced with developing facilities and curriculum that coherently address these demographic groups. In order to deal with demographic implications, funding issues, and competition from online resources, educational institutions may adopt a private sector model, where instead of having "selective" application and turning more students away, they try to let all the students come and become learning customers for life. Colleges don't tell students "goodbye" at graduation, and alumni become the market for distance learning and training options.

The aforementioned business trends that emphasize knowledge and talent over isolated skills have dramatic implications on education. Rather than more technically oriented students, we need, "broad-educated problem solvers who can acquire and apply knowledge in a wide range of ever-changing disciplines. In blunt terms, the Information Age may demand the primacy of the broad-based liberal arts education."[5] In this knowledge-based society, colleges and universities now have the central role to fill as the homes of intellect, knowledge creation, research, and lifelong learning.

Learning Trends: Constructivism, Technology, and Health Care

Current learning theories focus on analysis and application of information rather than the traditional teaching/learning model, with extensive implications for educational facilities. Constructivism, a widely acclaimed learning theory, is based on the innate human desire to make sense of the world through experiences and other methods. Constructivism uses contextualized learning in knowledge acquisition and construction, in contrast with the traditional "behavioralist" approach that uses rote learning and external reinforcement. Learners build procedural knowledge (techniques, skills, abilities) and propositional knowledge (facts, concepts, propositions), as well as dispositions (attitudes, values, interests) to decide if something is worth doing. Students construct knowledge with its usefulness, its functional context, and the social context of the student. Part of this social context is that learning is not just about the individual because the community of learners share a common physical—architectural—context. Learning is most effective using a realistic social and physical context as part of the process. By constructivist standards, "experts" organize information based on patterns, and through experience they amass a rich index of cognitive structures to easily recall and use. This structuring of knowledge parallels educational facility design, where the building is based on a fundamental organization of space patterns, so that the learner can experience spaces as part of the educational process.

Another venue for education, the workplace, has strengths as a learning environment because it has real goal-directed activities and problem solving, access to guidance from experts and coworkers, and intrinsic reinforcement, i.e., getting paid. Constructivist educators and businesspeople advocate that vocational education should change to incorporate approaches other than lectures and other situations that don't build experiential knowledge. Vocational education is also paralleling human resource development as both are tending towards a closer integration with applied learning and workplace preparation as the key to competitiveness and economic progress. The importance of "learning while doing" that is essential for occupational education programs is thus applicable to a vast variety of educational programs and paradigms. Vocational education and technology are getting rid of the distinction and inequality "between working with one's hands and working with one's head."[6] Two-year institutions of higher education are reemerging, serving government, community groups and businesses, working families, and other secondary and post-secondary education institutions. Furthermore, as occupations now require, education is becoming seamless and never-ending without strict phases such as elementary school, secondary school, higher education, and continuing education.

Some radical theorists predict that technology will cause content learning to become obsolete, but in order for educated people to focus on the process of applying knowledge, they must first learn content in an integrated manner, so that curriculum can then deal with constructing meaningful patterns, relationships, and other interdisciplinary links. In this manner, learning will move beyond only mastering content and skills, to incorporate lifelong-learning concepts that require continuous engagement in critical thinking, questioning, and contextual problem solving.

Technology is frequently a political panacea for these larger pedagogical and socioeconomic issues. Voters are encouraged to support educational technology over reducing class size or

repairing school buildings, because technology is not just an education issue, but part of preparation for the business world. President Clinton proclaimed that national education policy should include technology so that, "A child would be able to stretch a hand across a keyboard and reach every book ever written, every painting ever painted, and every symphony ever composed."[7] The gadgetry and infrastructure of technology is not a panacea for educational needs, nor will it function without human support. The same researchers that proclaimed that "Interactive multimedia will become the new tool of education…the traditional blackboard will be replaced by text, graphics, sound, animation, full-motion video, and other formats," later found that "old-fashioned methods of learning may be more effective than high-tech simulations…the realism of high-tech simulations does not necessarily or consistently translate into more effective learning and retention."[8] Technology can enhance the learning of those on campus and also reach students at a distance, and can better use faculty and student time, but "technology alone does not teach;"[9] the fundamental issue is still the students' motivation and the teachers' effectiveness. Information technology should be about community and conversation, but it is often used as a glorified workbook and a vehicle for submitting tests. Rather than merely repackaging "talking head" or "chalk and talk" styles of pedagogy, these tools should build real distributed learning communities to discuss and share ideas of real substance. According to Canadian educator Dr. Judith M. Newman, "The way we use technology in education is not the real issue; the important questions are really about curriculum and learning strategies."[10] Since software products and curriculum usages of computers have an inherent bias towards certain assumptions about learning, it is better for technology to be only one facet of the learning process, rather than an end in itself.

Health care education is inextricably linked to communication over distance, particularly in rural areas. Places ranging from the Australian Outback to Appalachian areas of West Virginia and elsewhere are pioneering systems to allow health care professionals to provide help and information without geographic impairment. Health care is interdisciplinary within and beyond medical professions to deal with issues of ergonomics, aesthetics, scientific research, etc.. Health care is moving away from acute illness care to health promotion and disease prevention, away from hospital settings to home and community settings with wider responsibilities. Health care reform is changing the way in which health care is provided and altering the role of allied health professionals, especially nurses. According to Purdue University[11], the continued streamlining of the American health care system is creating more demand for advanced practice nurses, and schools of nursing are answering the call by expanding both on campus and via distance education technology.

Health care education is following larger business trends with growing programs with partnerships, internships and "earn-while-you-learn" positions. As a result, educational opportunities and partnerships that are an attraction to technology companies also appeal to health care institutions. Another current interdisciplinary example is biotechnology, which has developments that are an example of science research fused with clinical medicine and business endeavors, and these partnerships are a new and different situation with which medical professionals will need to become educated.

Planning Makes Perfect (or at least better)

Historically, campus planning focused on laying out designs for which a more perfect institution could be created or expanded in a coherent fashion. While this role remains, the realm of planning is now faced with how to deal with an existing physical context, and continually work to adapt it to evolving programs, functions, and technologies.

Planning is not an event, an office, a resource to sit on a shelf, or a singular task, but rather a teaching and learning process that should result in controlled change, rather than lengthy documents, reports, and committee meetings. This process must be simple and unencumbered by excessive forms, committees, computer printouts, and rhetoric, so that time and energy can be concentrated towards the importance of people and data. The planning process should answer questions of what to achieve, how to accomplish it, and what criteria will measure achievement. The planning process should link facilities and capital plans, and should produce planning tools both for day-to-day activities as well as long-range capital projects, and tools for administrators to govern, monitor, and market the progress. This process is not inherently simple or linear; "…planning for higher education does not work with a "cookbook"-type approach; guiding institutions requires flexible leadership that can perpetually learn and grow."[12]

Traditional master plans tend to define the "destination" by giving a "twenty-year snapshot"

expressed through a rendered site plan or a pretty picture of proposed construction without giving the path to achieve the plan. Instead, master planning should emphasize connections with current facilities by being both a viable road map towards future plans as well as a practical, valuable guide for daily decision-making. Dealing with the existing overall context is also important, because approximately 80% of existing buildings will still exist in 20 years, and the plans shouldn't only focus on the other 20%.

Campus architecture should always operate within (and be subservient to) the larger context and environment. (In response to out-of-context designs, it has been said that campus architecture is "too important to be left to architects."[13]) Scale, not style, is the essential element in good campus design, so university leaders must insist that architects design campus facilities on a warm, human scale. Campuses should use a land use policy that addresses ecology, open space, buildings, roads, parking, pedestrian circulation, recreation, athletics, and utility infrastructure. The land use and open spaces can both enhance the environment and provide a coherent link between diverse functions. Outdoor spaces can have a sense of destination as an "outdoor room" as well as being circulation space to pass through. An institution's commitment to specific building projects or renovations does not, however, constitute a commitment to a land use policy. Within the land use policy, there should be a hierarchy of widths of circulation that accommodates necessary equipment, pedestrians, and vehicles without conflicts, and these paths should have a legible sense of direction and signage, particularly for impaired individuals. The pedestrian scale, and way of dealing with precedent, influences whether the campus has an image and feeling of an "academic village" with a unified sense of place, or a small "city" with many isolated places.

The campus image and overall massing need to be addressed carefully because visitors and prospective students often judge a campus based on first impressions. "Most high-school students decide if they like a university or not in 15 seconds,"[14] so the campus architecture and student facilities need to demonstrate attention to detail and commitment to student needs. The community environment, the sense of place at the heart of the campus, and whether the campus looks "traditional" with a park-like setting and fine architecture are all part of the overall ambiance. Besides the campus itself, skillfully edited photographs of the campus and its facilities will serve as the dominant marketing materials. The economics of the institution are improved through the appeal of the campus to prospective students. The physical planning can also deal with social aspects of the campus to improve interpersonal relationships and relationship of the campus to its neighbors with issues of housing, business partnerships, and community involvement.

Facility management becomes increasingly important, as the complex operations of university campus can be modeled by coordinating multiple departments, campuses, and standards for furniture, equipment, technology, and spaces. All of these facility management issues are able to become part of the actual strategic planning process so that quality campus maintenance is seen as part of an institution's image as well as for predicting facility expenses. Operational "life cycle" costs in a master plan include not just expenses for the initial building and daily operation, but also routine and preventative maintenance. Issues during the building's life cycle in the short term will deal with software and hardware, but since these structures are intended to last several decades, issues of infrastructure (power circuits, outlets, wiring, cabling), general cleanliness, and the overall spatial layout will become more important. Thus, the technology itself is not the entire issue; the architecture needs to facilitate long-term needs for flexibility. In terms of facility management for a research-oriented or other science building, there needs to be a body that coordinates use of facility, responds to changes in the research environment, and communicates building maintenance so as not to disrupt sensitive experiments.

Technology will impact campus planning in overall master planning, facility planning, and spatial layout, as well as specific spaces. Technology requires increasingly specialized space, with increased space demand for closets, storage, network infrastructure, utilities, and the information appliances themselves. Current interfaces such as the keyboard and mouse are transitory and will likely evolve or be replaced, but it is dangerous to design with fixed assumptions of the future, such as that ubiquitous wireless technology will eliminate all needs for cabling. Future use and demand for spaces and technology cannot be predicted by extrapolating current uses and trends. Building from effective use of the existing campus technology, new hardware/software should be selected using a defined process for supporting operations and maintenance. Ensure campus-wide technology standardization, balancing standards with individual preferences. Technology should address both instructional and non-instructional needs, such as providing for professional development and

training and collegial communication. Furthermore, it is incorrect to consider technology as a capital expense; the acquisition and renewal of technology must be integrated into the operating expenses in order for a facility to remain current.

A planner's most difficult challenge with respect to technology is "future proofing" buildings, but how should you plan "for something that you don't even know about?" What if the future arrives ahead of schedule? One answer is that instead of trying to have the latest gadgets in entirely customized areas, make sure that any construction and renovation is such that the building can be changed again in the future. One facet to remodeling and new construction that needs consideration, however, is the sensitivity of alumni with connections to "the old building that you just can't tear down"; the technology changes need to be subservient to the aesthetic physical environment.

When planning for specific structures, there must be great attention to the programming phase, including an analysis of needs with respect to present and future users. Participants in facilities programming meetings have other responsibilities and primary functions, so their participation should be as easy as possible, with easy access to pertinent data without searching through unneeded materials. Organize elements of the facility program in the written program document concisely; every word should have style and purpose that reflects quality architectural design. Make the information accurate and qualitative, and ensure the information is presented in an organized, concise, prioritized fashion. Programming and other planning processes need to be kept as brief as possible to hold faculty interest and improve enthusiasm.

Common errors that occur in programming (with possible solutions) include a lack of support for the program; stakeholders at the institution need to feel it is "their" document rather than "the architect's." Through workshops and presentations, the process should facilitate communication between administrative, instructional, and support services. Get a top-down approval of key administrators who will be stewards to champion and implement the plan, and make sure that all internal objectives and curriculum goals are clear before planning.

While it is good to involve as many users as possible, including community groups, make sure that there is a small core group of decision makers to allow for coherent leadership. Participants in the larger process should include a building user committee, project team, students, key administrators, trustees, and any other ways of involving stakeholders, but keep deans off of committees, making them instead part of an "executive oversight" group. The client team that works in programming should not try to design the building and be too prescriptive; program statements can deal with what, when, why, and how much, but the "how" needs to be left to the architect.

A common problem is insufficient gross area and inconsistent terminology regarding calculations of circulation space within departments; these need to be defined at the beginning so square footage arguments do not compromise the design. People don't understand square footage except when it's not enough. Administrators tend to focus on cost per square foot as an objective measure, but instructors and students perceive and use the educational space in very different ways independent of raw area.

The architect and clients may have different assumptions that cause incompatible quality goals; have key user groups tour comparable facilities in a range of prices, and clarify wants and needs. Part of this quality issue may be determining the value difference between new construction and reuse/renovation of existing facilities. Also, do not make unfounded assumptions of existing utility, subsurface, and building conditions, as well as site access and scheduling.

A primary issue of programming is that each department may be used to being the exclusive owner and user of large numbers of spaces. New campus facilities will be interdisciplinary, owned by the institution not the department, with diverse, collaborative funding and 24-hour access. Before adding space, use current buildings to the best and highest amount of use. When designing, have clear proposals signed off as to what really belongs to a department and what is considered "shared." For optimum efficiency in new designs, non-hazardous lab functions may be consolidated, and classroom space can become a shared, fluid, flexible resource. This can often be seen in existing centrally scheduled rooms, which have a much higher use/utilization rate than those under departmental control. By designing by function rather than academic discipline, disciplines can be combined even though programmatic spaces (classroom, office, lab, etc.) are more separate. In order to make these shared spaces usable, the physical adjacencies and zoning levels of departments are essential, to ensure that the rooms are situated in a practical fashion for all user groups.

Another departmental issue is unrealistic projected departmental expansion, a so-called

"Garrison Keillor Factor"[15] where all programs are "above average" and have 15-20% growth. Look at the history of program developments, departmental accomplishments, and the resulting faculty sizes at peer institutions to gauge what is realistic. Similarly, programs can reflect projected increases in student enrollment, but the program should aim to affect an increase quality of institution rather than mere number of students.

When attempting to convert curriculum and user needs into programming, educators and architects try to program every part of the built environment. By not designing spaces for spontaneous interaction, conversation, and reflection, planners actually compromise the effectiveness of "programmed" spaces. Gathering spaces need to be a priority in how a facility is organized.[16]

In the future, these changes with flexible, mixed-use programs may allow for vastly different functions to be combined into a new facility. For such a project, funding can be combined and mixed-use spaces can allow for technology-rich spaces, new meeting places for the entire institution, and a location for innovative academic programs. Educational facility planning may expand to incorporate public places with fused-use space to accommodate education, entertainment, training, and other functions into what we now see as fixed-use facilities such as malls, office buildings, and government facilities.

The process of master planning is expanding into the virtual realm, but tangible products are still a necessity. Once the master planning process has provided a framework for implementing change, make hard copies and electronic versions, packaging the information in manner suitable for promotional and fundraising purposes. These planning products will help to make the planning process more fluid and have real impact rather than remaining abstract plans.

Pedagogical Architecture: Spaces and Technologies

The realm of architecture for education is no longer one where the built environment is passive. The environment, described above as an integral part of new learning strategies, must now be part of the process as a teaching and learning tool. "No architect should be permitted to build for academe unless he or she fully appreciates that his or her building is an educational tool."[17] The learning environment should be rich in opportunities to discover, display, and demonstrate education, so that learning is a natural occurrence. The building must match and adapt to the user groups' educational preferences, and facilitate the incorporation of new spaces and technologies. A primary example of literally pedagogical architecture is the Lewis Center for Environmental Studies at Oberlin College by William McDonough and Partners, which demonstrates the sustainable technologies and design principles that are in the curriculum. In less extreme examples, buildings such as the Duke University Levine Science Research Center are designed "to encourage interaction among scientists from across the campus." The concept of "laboratory" or "library" can be extended so that the whole building has a larger role, and is filled with specific, concentrated, functional areas. The ideas of "classroom" and "computer lab" are eschewing their traditional form and needing new, practical, spaces.

Infused throughout these spaces are myriad technologies, whose design and implementation can no longer be considered peripheral to the architecture. Concentrated, technology-enhanced spaces are complemented by distributed areas for wired learning opportunities in the rest of the campus facilities. These technology systems, along with the structure and mechanical systems of the architecture, should contribute towards the experience of learning about technology. While the technological specifications change quickly, their design concepts have a more lasting meaning within the larger educational context.

Audiovisual (AV), visual projection, and other presentation technologies are a field that includes rapid changes in "state of the art" specifications for brightness, resolution, and size, and rapid changes with converging media. These technologies, previously limited to auditoriums and corporate boardrooms, are now becoming commonplace in medical imaging, international meeting facilities, and churches. Slides are becoming obsolete (except for art and architecture lecture situations) as slide images can be converted to computer formats. A document camera is preferred over the overhead projector or slides, because it can read a transparency or an opaque image and transmit it to the computerized projector. Presentation technology is loosely defined; to different faculty members it may mean sophisticated AV equipment, seamless computer interfaces, using video instead of slides, using an overhead projector, or using different colored chalk.

In education situations it is important to differentiate between equipment to deal with incorporation of physical demonstration material requiring video camera/display, versus digital

presentations from a computer or other source depending on different teaching strategies. Also, digital technologies often need to function along with quality analog means of display such as chalkboards, and whiteboards. The logistics of presentation technologies can become complex depending on the quantity of media being used within a single lecture. (An example is that within a 50-minute lecture at MIT, media may include nine motorized mechanical chalkboards, video close-ups of experiments, document cameras, overhead projectors, computer animations and internet documents.) Besides knowing what media need projecting, which need to be shown concurrently? Switching systems between several computers or computer and video may be cost-prohibitive. Does the presenter need to be videotaped while the projector is showing images, causing conflicting lighting issues? Will video production facilities for students and/or faculty need to occur in similar places? The placement and aesthetics of projection media will determine how interactive the learning environment feels, and these issues differ for lecture facilities and cooperative learning spaces.

For a video/graphics projector, consider image size, the type of input sources, light output, control, weight, power, cables, and ease of access for maintenance. For projectors, the required brightness (luminous intensity in lumens) is affected by the surrounding ambient light requirements; all facilities need to be able to eliminate natural light when necessary. Besides the lumens of a projector and the ambient light, the lumens per square foot, which is proportional to the inverse square of the distance to the screen, will determine the apparent image brightness.

Rear projection, where the image is projected onto the screen from behind, allows a presenter to walk in front of the image without casting a shadow, and the presenter does not have to look into a bright projector light (a health hazard) when addressing the audience. The image brightness is more intense, allowing more ambient light in the room, but the equipment and maintenance are quite expensive. If there isn't enough depth in a rear projection booth to shoot straight, mirrors are used. A typical rear screen projector is smaller, quieter, and lighter than a long-range front projector, and front-projector lamps are expensive and difficult to change, especially with a sloped floor. If front projection is used, especially for large lecture halls, the screen can be above the height of the presenter. For televisions and monitors, which have screen measurements given diagonally, the size should be more than one diagonal inch per foot from the farthest viewer. The farthest a person should be away from the screen is based on the size of the projected image and the amount of detail needed to be seen: for general viewing, image height equals one-eighth to one-sixth the distance from the screen; for highly detailed inspection, image height equals one-fourth the distance from the screen.

With so many technologies interacting, overall acoustic quality for the spoken voice alongside technologies becomes increasingly important. Depending on the space, the technologies may need to be more subservient to comfortable acoustics that allow for collaboration and discussion without difficulty hearing the presenter. Guidelines for audio and acoustics include consideration of analog and digital formats of audio and video, and where the equipment is accessible and controllable. Sound controls need to be located within the open portion of a room to set volume; lighting and projection can be controlled from a booth. Consider possibilities for multiple control options and locations depending on the presenters, lecturers, etc. needs and preferences, but controls such as touch screens or engraved panels need to be readable in low lighting and very user friendly. Microphones in the ceiling do not work well for teleconferencing, so microphone wiring needs to go to fixtures in the room. Use conduits to separate microphone and audio from power, because power cables may induce buzzing in the microphone signal.

Sound barriers, which control sound by absorbing (insulation, etc.) and reflecting (hard surfaces), can be ruined by even a small hole, so a sound barrier wall needs to extend from floor to roof deck. It does not work to place insulation above the ceiling or extend a wall of insulation up to the deck. To control sound between rooms, equalize the noise level heard from either side of a wall or floor, either by masking the sound or adding "white noise," muzak, etc. to create a consistent sound that becomes unnoticeable and forms a distraction from the noise you're trying to cover up.

In general, when monitors, computers, or audiovisual equipment racks are mounted in casework, they need separate mechanical ventilation or at least air circulation with air holes out to the adjacent room. Equipment racks need to be accessed from the front and back, and should be on wheeled or slide-out racks. Wall-mounted equipment is supported by studs and/or brackets between studs, and the same applies for joists supporting ceiling-mounted equipment. For wall- or floor-mounted interface boxes, there needs to be clarity as to the required depth, connections, size and location of plate, and the finish

for interior design. Make sure systems are developed in ways that are practically adaptable to changing communication lines such as copper cable, fiber optic, and other video control systems. Fiber optic capacities which currently approach 50 million bits per second will continue to increase, approaching 1 billion bits per second, facilitating enough bandwidth for increasing video and data demand, but the wiring systems will still need functional ways to be replaced. Consider below-floor-level conduit with wires going through furniture to avoid unsightly or dangerous cables dangling. Projection systems are often facility-dedicated, but some applications may require an integrated mobile solution for multiple videoconferencing capabilities; the AV cart may make a comeback after all.

Auditoriums, including lecture halls and performance spaces, are incorporating these presentation technologies as well as access to information technology networks. The physical environments of past lecture facilities are not conducive to more interactive teaching methods and multiple technologies, but the incorporation of technology should not eliminate traditional, spontaneous teaching formats. For example, if lecture halls have an AV booth that is separate and locked, this can cause difficulties with technologies not being accessible to the instructor in front. If there is an AV booth at the back, make sure one does not need to traverse through the large auditorium space to reach the booth. Furthermore, if the space is tiered, there may be difficulties with moving equipment.

From an acoustic and structural standpoint, make auditoriums "rooms within rooms" to isolate them from noise in the rest of the structure. Sound can be blocked by large partitions with mass or smaller, less massive partitions separated by a large air space. When optimizing acoustics, realize that reverberation time and other information depends greatly on how many people are present, since each person is approximately equivalent to 0.8 square meters of acoustically absorbent material. In general, speech sounds best with less reverberation time, and music sounds best with more reverberation time, but this varies greatly depending on the type of music being performed or played through an audio system, and how varied the uses of the space will be.

Aside from the acoustics of sound traveling directly from the source to the listener, the ceiling shape will determine what quality and quantity of reflected sound waves reach the listener. A flat ceiling will reflect sound waves like a mirror will reflect light, whereas a surface with uneven texture and/or with a wavy shape will tend to scatter waves randomly. A convex ceiling will bounce sounds and scatter the sound around evenly. A domed, or concave, ceiling bounces the sound within it, creating places where the sound is focused; these types of ceilings may create problems.

Sight lines are critical in large rooms like auditoriums. To see over the person in front of you, you must either be seated high enough or be staggered to see between two people. A person typically will be uncomfortable if they have to move their head up more than 30 degrees to see the screen. The maximum viewing angle to the screen, based on seat location in the room, should be 45 degrees. Chalkboards and display panels towards the side of the hall should be angled to improve sight lines. Sloped floors and fixed seating can impede the usage of video cameras for taping lectures and bringing in new demonstration equipment.

If the space has auditorium-style chair seating, this may not include enough space for students and their laptops, etc., but if the space instead has stationary chairs and large tables, there will be a large workspace, but the room will feel isolated and inefficient. One difficulty at the University of Washington Physics and Astronomy Building by Cesar Pelli is that the closer spacing of lecture hall seats has forced changes in educational procedures for examinations.

Hallway entries to lecture facilities (as well as classrooms) should be finished with wood or other touches that echo the interior lecture spaces. It would be helpful to include a transitional "lobby" space to enter a lecture hall for handouts, announcements, and traffic. A good scheme for dealing with high volumes of traffic may include aisles through the seats that start at the floor of the lecture hall, with the addition of secondary aisles along the sides that are partially concealed. For example, at the 10-250 lecture hall at MIT, these aisles are concealed by a pair of diagonal, tall, blank walls that are used as a projection surfaces by lecturers.

A lectern is best located to the left of the audience, since people tend to read from left to right and look back at the presenter after each image. This lectern can be a place to incorporate controls and connections for technology, projection, and lighting, but fixtures at the front of the room should not be too bulky as to make the students and lecturer feel separated. If the presenter has hot lights shining down, there should be the possibility of additional air conditioning. It is good

to have a hard surface behind a presenter in a lecture hall if you want to increase the sound directed out to the audience, whereas a soft absorbing surface behind a person will enhance their ability to hear what someone else is saying, such as having a soft surface behind a lobby reception desk.

Especially for smaller facilities, track lighting is good for the front of the room, and recessed lighting is good for "wall wash" light. Overhead lights should be able to be dimmed and return to full brightness quickly for short video clips, etc. during presentations, but exit lights can be a distraction. In lecture facilities, all of the lighting (as well as technology and video inputs) should be controlled at the front by an interface, perhaps integrated into modular presentation tables.

General halls in universities often serve multiple purposes; even in a new facility intended for one department, other campus organizations will want to use the facilities. A facility may need an ability to be functional in settings including humanities discussions and presentations, science lectures, movies, exams, and interactive multimedia presentations. As mentioned above, the teaching styles of professors will likely include several digital and analog technologies within a single lecture, so avoid complex or difficult interfaces for switching media. Within the hall, it may be practical to have larger faculty workstations for colleagues, printer stations for students to produce hard copies of notes, or even coat rooms.

When designing auditorium spaces within the larger context of educational facilities, other adjacent spaces should include considerations for storage, equipment setup, presenter preparation, and catering. Surrounding circulation spaces need to be accessible and efficient to deal with inflow and outflow between classes, but there should be smaller-scaled, more intimate spaces available for informal discussion before and after lectures. One possible way to allow for consolidation of technology spaces while spreading out traffic is to have several lecture halls clustered around centralized AV and storage rooms. Especially with science facilities, the program of lecture and teaching space should be articulated differently from research laboratories, with the interstitial spaces forming a link with the larger community, perhaps as a courtyard or atrium.

Conference room benchmark facilities are often corporate headquarters' meeting rooms. Education institutions often emulate these corporate environments in their meeting and presentation rooms for campus groups and prospective students. Current standards include careful usage of detailed woodwork, wall surfaces, and furnishings, with a wide range of lighting schemes to fit the presentation technologies. "Board-room" configurations of furnishings do not work well for videoconferencing and distance education, so new flexible ways are emerging in which a single wall or truncated corner provides a focal point and location for technology.

Rather than being driven by clunky appliances, new technologies are trying to emulate traditional pen-and-paper or chalkboard/whiteboard usages by providing "smart" surfaces that can read words, drawn images, diagrams, and notes and transmit them to another location for collaboration, client meetings, or a design process. Some of these new technologies are even adopting the name "roomware," alluding to how the entire room is part of the communication and technology apparatus.

As always, though, the delineation between "conference rooms" and other meeting, training, or learning facilities is becoming blurred; the specific use of a conference room is evolving into a general category of smaller presentation spaces that defy classification as merely small auditoriums or technology-rich classrooms. The growing educational trends towards small group project-based learning may cause these small, concentrated spaces to proliferate as a new model for a type of educational and business facility.

Distance learning spaces accommodate education using video technologies that may include 1-way video, 2-way interactive video, satellite-based and/or computer-based teleconferencing, and broadcast and/or cable teleconferencing. Local area networks, fax machines, and other technologies are often needed in these communication-intensive spaces. In distance learning design, parameters are driven by AV requirements, comfort and service to the individual user, and being user-friendly to a technical programmer. Consider the layout of the entire facility in terms of circulation, accessibility, and noise. Keep distance education spaces away from noisy spaces, mechanical rooms, busy streets. Deal with internal room circulation for the instructor to have access into the audience, but minimize disruption of class in session via front and rear doors. Decide how the space will function with respect to local or remote students, and/or whether this facility's characteristics need to match with other existing campus facilities.

Study the customers (faculty and students), and what types of classes, subject matter, and pedagogy are used with existing methods of instruction and/or videoconferencing. Successful

students for distance education are often voluntarily seeking further education and career advances, are highly motivated and self-disciplined, and have post-secondary education goals. It is tempting to make technology the key issue, but it is a tool in support of curriculum; research shows that the instructional format itself has little effect on student achievement as long as the delivery technology is appropriate to the content and all participants have access to the same technology. Good videoconferencing professors need new skills and more specialized presentation strategies, and spaces can help to allow for seamless usages of technology and adapt to new learning strategies.

Distance Learning is not a way for schools to save money outright; technology requires more rather than fewer support staff and more professor involvement. Cost factors include technology hardware, transmission expenses, maintenance, telecommunications infrastructure, production, support, and personnel to staff the aforementioned functions.

Include AV with the architecture scheduling process; consider the impact of special technology such as telecommunication, videoconferencing, and other AV and IT systems on other building trades. Installation of AV and telecommunications requires a clean, dust-free environment, so cooperation and proper sequencing is necessary. Delaying AV work allows purchase of more recent equipment because contractors do most work after the bricks and mortar contractors. Delaying AV work has many disadvantages for smaller jobs because occupancy deadlines may be more important than waiting for new equipment. If the AV contractor can work in the schedule with other contractors, the owner can be aware of all costs up front at bid time and conduit or cabling in can be installed easier.

For cameras and displays, decide on placement, size, whether they will be one- or two-way, and if they need capabilities for one or more camera orientations, sending/receiving graphics, or connecting with a document camera. Address possibilities for horizontal and vertical sight lines of the students to displays and cameras, as well as sight lines from the projectors to the screens. Consider possibilities for tiered floors and angled placement of cameras. White boards can be difficult for video transmission depending on lighting schemes, so consider traditional chalkboards as well.

A built-in demonstration bench/table may require an overhead or wall-mounted camera, and the fixed instructors console should be customized with monitor, keyboard, processor, document camera, control panel, cables, plus any other AV equipment. Students' tables should be built-in with modesty panel or full panel front, with wiring for power, voice, data, and shared microphone. For other built-ins, do all custom casework, and have architects, engineers, and AV consultants create designs to show telecommunications wiring.

Distance educational facilities need extensive telecommunications design to develop a passive cabling infrastructure to facilitate changing technology and instruction. This goal is accomplished when cabling environments enable new systems, capacities, and media in a multi-product, multi-vendor environment. Create structured wiring as a flexible, modular system based on standard cable types deployed through the facility in an organized hierarchy of connection, with standardized interfaces for workstations and connections.

Particularly for distance learning spaces, construction documents need to show the enormous amounts of conduits and cables, with section cuts through corridors to coordinate space for cable trays, conduit, ductwork, piping, and equipment from all disciplines. Remodeling projects will require an exhaustive existing field survey. Require the design team to draw an elevation of the teaching/presentation wall with all screens, vertical writing surfaces (dustless marker boards), equipment, device plates, and switches.

For acoustical privacy, identify noise sources and vibrations from mechanical and electrical systems, adjacent rooms, and outdoors. Distance education spaces need a room sound quality NC (noise criterion) 30 rating. As part of this effort, all drain piping should be outside the distance education room envelope, while keeping supply piping and valves away from the ceiling. HVAC (Heating, Ventilation, and Air Conditioning) design should have unobtrusive duct noise, perhaps incorporating duct silencing equipment. The electrical system should ground to one point to avoid noise in electronic signals, and should be connected to emergency backup generators. Lighting systems should use directional, controlled lighting without glare-producing point sources, keeping light on the walls but off the display; if there are any windows, they must have black-out shading to ensure user control. The color temperatures and reflectivity of interior surfaces should work with the needs of the participants and presenter. For example, interior surfaces should be colored in pale pastels, blue, gray, or mauve, but not in skin tones, warm colors, brown or black, or any busy patterns that would interfere with video transmission.

For videoconferencing design, merely "having" video and audio doesn't mean it is good—it must be useable and effective, so verify that audiovisual and telecommunications consultants have experience in distance learning spaces ranging from programming, design, system interfaces, construction administration, and training.

In contrast to programs that enroll large quantities of students, Drake University has chosen to limit enrollment to keep the quality high in distance learning programs. Besides internet and video correspondence, courses are available through the Iowa Communications Network (ICN) Iowa's statewide fiber optic network. This may be the world's first statewide fiber-optic network to connect all public schools, hospitals, libraries, prisons, and the National Guard in a state. This program puts Iowa "in the middle of everywhere" with web-based courses used worldwide. Educators realize that refined teaching methods and electronically delivered interactive courses take time and money to develop properly.

Since 1993, rural aborigine communities in Australia have used videoconferencing as the primary medium for personal and business communications. These systems (using industry-standard PictureTel equipment) can convey hand gestures; most videoconferences are personal or ceremonial in nature. It is important that the subtleties of human communication are not hindered by technology, because learners will reject systems that are not efficient and realistically communicative.

The North Carolina Public Health Training and Information Network (PHTIN) is a government/educational initiative started in 1995 that provides education, training, and information services to the public health work force in North Carolina. Created in response to challenges of future public health practices, PHTIN uses interactive videoconferencing, connecting the UNC School of Public Health in Chapel Hill, the NC department of health and human services, community colleges, hospitals, local health education centers, and the CDC in Atlanta, enabling best practices to be distributed statewide. On a practical level, during one year in 1998-1999, PHTIN's 211 programs had 8783 participants who saved a total of 834,385 travel miles.

The University of California-Davis Health System includes regional and rural telemedicine programs use a videoconferencing system that allows for community service by exporting clinical expertise to rural Northern California clinics and Community Hospital Network affiliates which have a shortage of local physician specialists. Given the interactive nature of videoconferencing, physicians will be able to receive Continuing Medical Education credit for their participation, and administrative videoconferencing allows healthcare professionals to participate in meetings and presentations without traveling to Sacramento.

Science doesn't occur only in the laboratory; there is less clear demarcation between laboratories, classrooms, offices, and public spaces in science teaching facilities. "Therefore, consider the building as a whole…If the laboratory is in fact a place conducive to experimentation, investigation, and observation, then the whole building can now indeed become the laboratory."[18] Every square foot in a laboratory must be considered for its potential contribution to the laboratory; traditional "net" and gross" area measurements will metamorphose into "learning environments," transitional or contributing spaces," and "utility space."

The location of offices encourages interdisciplinary collaboration and research, and offices need to be in a different space relation to laboratories and classrooms. Especially when the building isn't big enough, avoid a sense of "my office" or "my lab," by bringing together dispersed faculty and/or consolidate disciplines, agencies, etc. Having zones of offices and labs linked by atrium spaces can also bring together people who otherwise wouldn't interact. Site arrangements and using buildings as indoor "streets" faced with visible laboratories and interactive displays can also facilitate interaction between engineering and science students interacting with students from the humanities.

Create a utility and material distribution corridor internal to the labs, separate from public corridors, and consider the need for off-site support and direct exterior access to lab spaces. Since support and repairs are expensive, design labs for both low cost of maintenance and low technological level of maintenance through simplicity and access. Access to utilities can be handled by overhead or underfloor grids rather than attached to fixed furniture, so the labs are not forced into a preset modular bay system or pattern, so some labs can be open while others are subdivided (hot rooms, cold rooms, clean rooms, etc.), with specialized equipment.

As research becomes more team-oriented, industries demand new collaborative skills in workers. Often, expensive equipment must be used continuously by more than one team, so cooperation is necessary for the very existence of the research. Support spaces for preparation rooms and access to major stored equipment

need to reflect this collaboration. With respect to equipment, laboratories include myriad technical concerns relating to physical equipment and utilities as well as information technology networks. Consider ring stands and future equipment when planning hood sizes, and allow equipment to be connected to future computer networks. Indoor Air Quality (IAQ) is especially important within and between laboratory settings for environmental precision as well as human health. The abundance of natural light makes is good for a learning environment and for some biology experiments, but labs need to be completely darkened for video presentations and experiments that require real control over light.

Consider opportunities for multiple seating and options within flexible academic spaces, which can include differentiated spaces for lecture, teaching, research, and other specialized spaces within the larger "laboratory" context. Laboratories for undergraduates fulfilling general requirements need to have larger academic spaces; those for declared majors need smaller, more defined spaces.

Early Cost Estimation for Manufacturing of Tooling in Resin Transfer Molding

Robert W. Lin

Due to the increasing demand for strong and lightweight parts, the aerospace and automobile industries began using composite materials in their parts manufacturing. Of the three main methods of composites processing, manufacturers increasingly use the Resin Transfer Molding (RTM) process because it provides a moderate alternative between the high volume, large capital investment compression molding and low volume, low capital investment vacuum infusion molding. The RTM balance between volume and investment on machines and tooling makes it important to develop an early cost estimation for RTM tools.

Understanding RTM

RTM consists of a thermoset resin injected into a two-part matched mold containing a dry fiber reinforcement. The basic RTM process includes the loading of the preformed reinforcement into the RTM tool, closing the tool, injecting the thermoset resin into the mold, curing the resin, and removing the part from the mold as shown in Figure 11.

The dry fiber reinforcement, often called a preform, consists of a wide range of materials from natural wood fibers to synthetic polyester fibers. The automobile industry most commonly uses glass fibers, whereas the aerospace industry uses mostly carbon fibers.[2] Because of the differing part requirements, industries use various perform fibers and fiber preparation methods to create the preform. In most cases, the toolmaker manually places a shaped perform into a prepared mold cavity. Much like injection molding, the index pins align the two mold halves, while a press closes the mold. Clamps hold the two mold halves in place as the resin injects into the mold.[3] The resin wets the preform, and when resin begins filling the venting ports, the injection stops. An oven, platens, or integral heaters elevate the temperature of the mold, curing the resin. Finally, one removes the cured resin part and if required, finishes the part by trimming or polishing.[1]

Costing

Traditionally, companies assign the cost estimation of these RTM tools to the most experienced mold makers, because the increasing number of manufacturing materials and processes used in RTM makes pricing difficult. Often toolmakers admit ignorance to an exact method for determining cost. Despite the difficulties, cost estimation is an essential part of the design for manufacturing (DFM) of a RTM tool on two different levels.

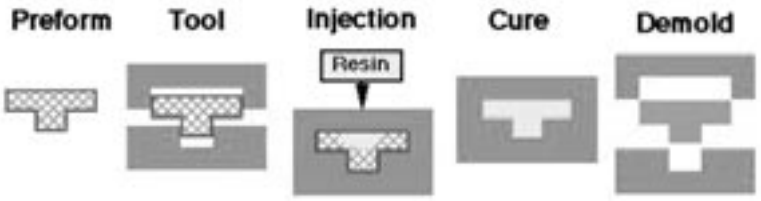


Figure 1

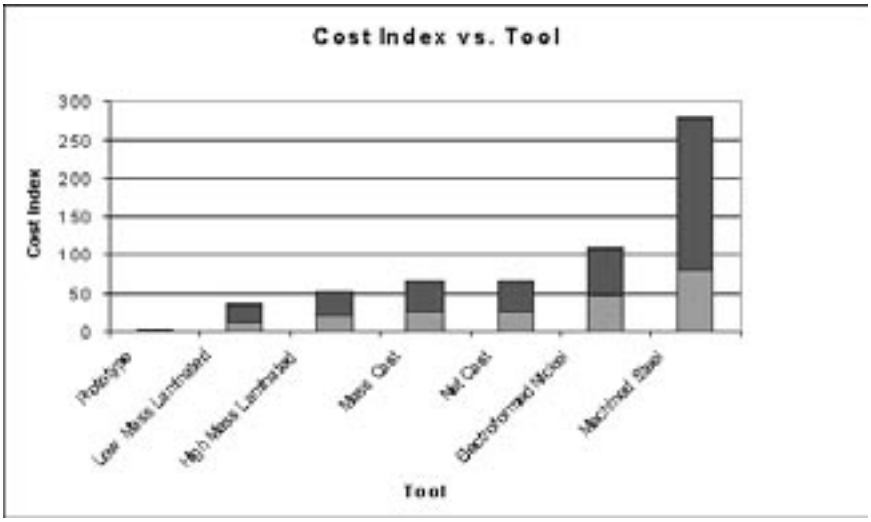


Figure 2

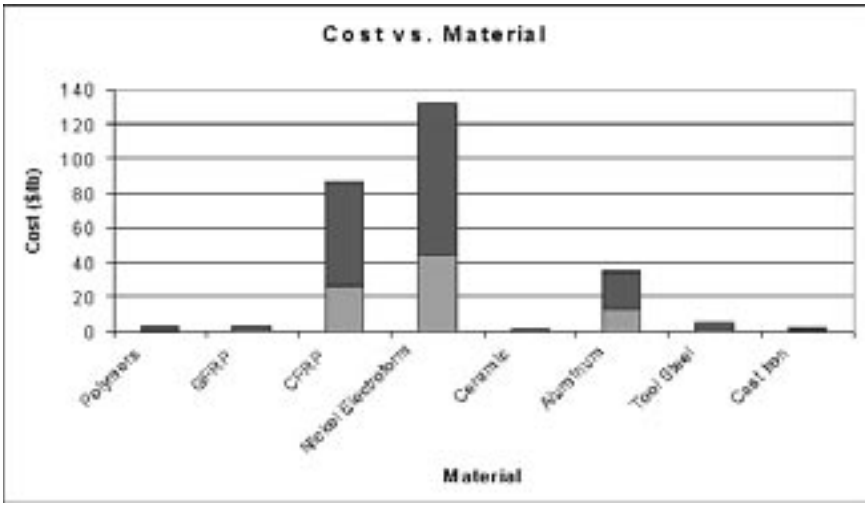


Figure 3

A demand for the part initiates the first level of RTM tooling. The part designer creates a part with strength and dimensional requirements to fulfill a task. At this level, an early estimation method for RTM tooling helps the part designer create an economically viable tool. Therefore, even before the conception of the tooling design, the parts designer considers the cost. For instance, the designer chooses certain features because a long z-axis part adds to the tooling cost or the strength requirement of a slender component in the part call for large manufacturing pressures achieved only with expensive material such as Electroformed Nickel (Interview with Eric Johnson, Everett Pattern and Manufacturing).

On the second level, the tool designer creates a tool to match the part. Due to the relative newness of the composite RTM process, few tool designers know how to choose and design an economically viable tool for a given part. An early cost model helps the tool designer decrease costs on this tooling design level. This is especially important in RTM because so many feasible mold designs exist for a given part. Thus, the use of a cost estimation model can overcome some of the limitations of relying only on design experience in selecting the designs, materials, and production methods for a particular part and tool. This paper presents some guidelines for early cost estimation in RTM tooling.[3]

There are three major steps in tool manufacturing. First, the manufacturer designs the tool for manufacturing (DFM). Then the toolmaker performs analysis on the design. When the tool design passes the analysis, the toolmaker manufactures the tool. Finally, after the final inspection, the toolmaker ships the tool to the customer. Because the methods for estimating shipping costs do not rely solely on the toolmaker, this paper concentrates on the first two steps of manufacturing. In considering an early cost model of RTM tooling, the designer should also split the cost modeling into these major categories to simplify the process.

Design and Analysis

Initially, a designer creates a part, which requires a tool to produce. Next, a toolmaker designs a tool for manufacturing the part (DFM). This is usually done using a computer-aided design to model the tool (CAD model). Unlike injection molding and other manufacturing processes, depending on the performance

requirements, several methods exist for RTM tooling. Two basic categories separate RTM tools. Tools made from composite make up the first category of tooling. The different materials and laminating processes divide the composite tools into low-mass laminate, high-mass laminate, and composite-mass cast tools. Tools made from metal compose the second category of RTM tools. Differing manufacturing processes such as near-net casting, nickel electroforming, and machining divide this category of tooling. Table 1 gives a rough estimation of the appropriate type of tool for the part production volume (Interview with Global Glass, Able Body Corporation, Fiberglas and Plastic).

In choosing a tooling process, the designer should keep in mind the limitations for each type of tooling as well as the approximate tooling costs. Figure 2 provides a rough costing index useful in the design stage of the tool (Lacovara B (1995), Considering Resin Transfer Molding, CFA Paper).

Manufacturers build RTM tools out of many materials using many processes. Not only do the material costs differ, each material requires a different manufacturing process to build. Thus, the material choice becomes a major cost driver. Some considerations for choosing a material include the strength, coefficient of thermal expansion (CTE), cost, and weight of the material. The required part volume, and the tooling type also limit the list of possible materials.[4]

The geometric complexity of the tool affects both the design and the manufacturing costs of the tool. The main cost-driving features that significantly affects the cost of the tool includes the long z-axis depth of the tool and the amount of material removed in this direction. The small radii also affect the cost. The design and analysis stages of creating the tool account for 30 to 50 percent of all costs (Site tour, Everett Pattern and Manufacturing).

Manufacturing and Verification Costs

Other major contributors to the tooling cost are the manufacturing and inspection of the tool. The material cost, the labor costs, the machine costs, and the detailing costs are buried within the manufacturing costs. The 3-D CAD data fully defines the volume and the dimensions of the final RTM tool. Given the density and cost per unit volume of a material, the manufacturer can calculate the total material cost. Figure 3[5,6] compares the cost of materials for RTM tools. It is important to note that the material selection differs from the tooling selection. After the selection

Table 1

Tools	Part Count
Low Mass Laminated Tool	Less than 1000
High Mass Laminated Tools	Less than 5000
Mass Laminated Tools	Less than 10,0000
Near Net-Cast Tools	Less than 50,000
Nickel Shell	5,000 annual
Metal Fabricated	10,000 Annual

Table 2

Composite Tooling Process (note that inspection occurs after most steps)
Create and Finish the Pattern
Apply Tooling Gel Coat
Lay and Cure Coats of Resin
Lay Core Material
Build and Attach Reinforcement Frame
Glass the Frame
Demold Plug/Finish the Tool
Build Channel
Fix Seal
Make Male Part Similarly
Build Alignment
Build Injector System
Build Venting System
Final Inspection and Shipping

Table 3

Metal Tooling Process
Machine Out Basic Block for Female Part
Inspection
Create Surface Features: i.e. Shipping Holes, Lock Mechanisms, and Demolding Bolts
Inspection
Create Surface: Rough Pass, Finish Pass, and Bench Pass.
Create Secondary Features: Rough Pass, and Finish Pass.
Inspection
Final Finish: Sanding/Buffing, and Hand Finish.
Make Male Part Similarly
Final Inspection and Shipping

References

(1) Michigan State University, Intelligent Systems Laboratory (1999), Tutorial on Polymer Composite Molding.

(2) Schwartz MM (1997), Composites Materials Volume II, Processing, Fabrication, and Applications, Upper Saddle River, NJ:Prentice Hall PRT.

(3) Reinhart TJ (1987), Engineered Materials Handbook: Composites ASM International Handbook Committee Volume 1, Metal Parks, Ohio: ASM International.

(4) Potter K (1997), Resin Transfer Moulding, London; New York: Chapman & Hall, 1997.

(5) Haffner SM, Gutowski TG (1999), Tooling for Manufacturing of Composites-Fabricated Metallic Tooling, Cambridge: MIT Laboratory for Manufacturing and Productivity.

(6) Ashby MF, Jones DRH (1996), Engineering Materials 1, Oxford; Boston: Butterworth Heinemann, 2nd Ed.

(7) Ostwald PF (1998), AM Cost Estimator: Cost Estimating Database, Cleveland, Ohio: Penton Publishing, Inc., 4th Ed.

(8) Polgar KC (1994), Simplified Time Estimation for Basic Machining Operations, Cambridge: MIT Laboratory for Manufacturing and Productivity.

of an appropriate tool, it narrows down the list of possible materials, but the toolmaker still must choose an appropriate material to build the tool. For instance, a toolmaker must select from different materials such as invar, tool steel, cast iron, or aluminum to build a metal fabricated tool.

Next, the designers must estimate the costs of the actual building of the part. Equation 1 shows a good estimation for these costs.

$$Cost_{total} = T_{mfg} Cost_{labor} + T_{mfg} Cost_{machine}$$

Where, the cost of labor per hour differs for each company. However, one can use $100/hr as a quick estimate in early cost approximations. The cost for running the machine every hour also differs for each company. However, $100/hr. is also a reasonable rough approximation. Thus, multiplying the manufacturing time by the labor cost per hour and the machine costs per hour results in a sound estimation of the building cost.[7]

The time for manufacturing (T) comes from a summation of the times for each step in the process plan. The manufacturer should perform several analyses to calculate early estimations for the manufacturing time.

1) Make a rough process plan for the tool. Table 2 (Interview with Falls Fiberglas Products, Inc., A-1 Fiberglas, Accura Industries, Inc.) and Table 3 (Interview with Mark Bellanger, Technician at the Laboratory for Manufacturing and Productivity at MIT) show the rough process plans for the two categories of RTM tools

2) For each step in the process plan, determine the time required to set up the process. Three major steps contributed to the set up time. First, preseting the tool contributes to the manufacturing time. For instance, on a metal tool, each tool requires 15 minutes to preset on a 5-axis CNC machine. Secondly, changing the tool also adds time to the manufacturing. In metal tooling, it requires about 0.5 minutes to change a tool. Lastly, positioning or repositioning the tool to do the process also requires time.[8]

3) Calculate the time required to complete each process in the process plan. For instance, in metal tools, calculate the material volume of the removed material. Then use the material removal rate to find the time require or the actual roughing and finishing machine processes.

4) Find the time required to inspect the tool after each important process and add up the times for each step in order to find the total manufacturing time.

Using these tables above and incorporating more detail according to each particular part, one can estimate the time required for each part. Next, applying Equation 1 results in an early cost estimation for the tool-building steps of RTM tooling.

Conclusion

Adding up many categories of tooling costs gives an early cost estimation for RTM tooling. First, the design costs section reviews the main criteria for determining the costs of the tool. Secondly, the total manufacturing costs come from the material costs as well as the labor and machine costs. Lastly, the shipping costs add minimally to the overall cost of the tool. Using these basic guidelines will help both the part and tool designer in building an economically viable RTM tool.

Image Processing Issues with Fast, Real-Time *In Vivo* Optical Coherence Tomography Systems

Michael Manway Liu

Early detection and diagnosis of diseases are crucial in medicine. Healing and recovery are far easier when the diseases in question have yet to establish a strong foothold. Towards this end, medical practitioners have devised numerous techniques to diagnose diseases before they become a major concern. Such techniques can be grouped into two broad categories: *in vitro* and *in vivo.* In vitro techniques work by examining tissue samples for diseases within a laboratory environment. That is, tissue is removed from a live specimen and tested for suspected pathologies within a lab, usually through microscopic examination (otherwise known as a biopsy). *In vivo* techniques work by testing for diseases within the context of the living system. They attempt to correlate the physical symptoms of the patient with a list of known diseases. Most common diagnoses use only *in vivo* techniques: for example, when a doctor prescribes bed rest and lots of fluids for a case of the common cold, he usually reaches his diagnosis by checking the patient's temperature, physical comfort, and mucus output. All such checks are done in vivo. In extenuating circumstances, however, such as when he cannot determine whether the patient is suffering from a bad cold or the much more serious flu, he might take a sample of the patient's blood and run it through a lab. By doing so, he moves his diagnosis from an in vivo context to an in vitro one. In vivo techniques,

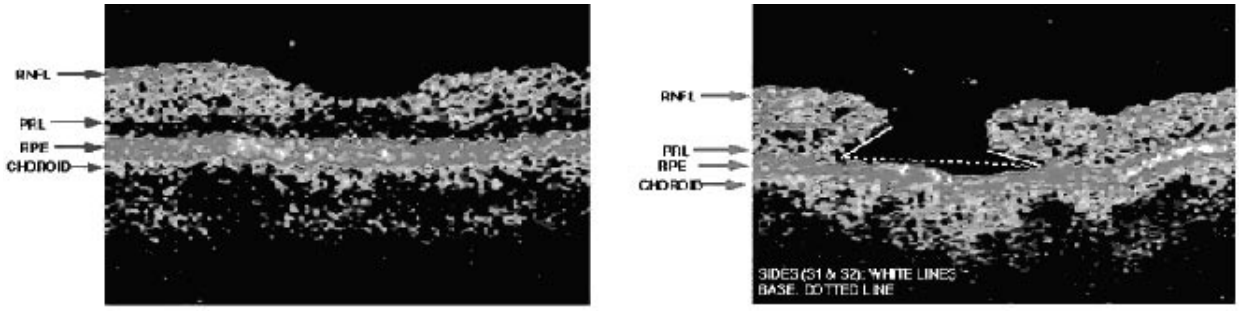


Figure 1: (Left) Image of a normal fovea. (Right) A macular hole, characterized by the absence of retinal tissue. While readily apparent when viewed using OCT, macular holes are difficult to distinguish from normal retinal tears when directly examined. (Images courtesy of the NYEEI Ocular Image Center)[2]

then, are commonly used within a clinical setting to diagnose diseases with in vitro techniques being used whenever the latter fails to reach a definitive diagnosis. This setup, though quite effective in most cases, is unfortunately imperfect. For diseases where few, if any, physical symptoms emerge until the late chronic stages, an in vivo examination would detect nothing until it is too late. Cancer, in all its varieties, comes to mind as one such disease. It may also be true that an in vitro analysis cannot be readily performed because the tissue is sensitive to sampling. Most people, for example, would object to having a piece of their eyeball sliced out for microscopic examination. In both situations, current medical diagnosis techniques remain comparatively poor with a high probability of error.

Optical Coherence Tomography (OCT) is a relatively new technology for scanning and imaging tissue anatomy.[1] It addresses the issues raised above by greatly enhancing the acuity and utility of visual techniques—the technology boasts a resolution of 1 to 10 micrometers and may diagnose both soft (such as skin and eye) and hard (such as teeth) tissue diseases. Beyond its incredibly high resolution, OCT is also completely benign and inherently noninvasive (unless used to image internal tissue). These qualities combined make it extremely attractive as a medical diagnostic tool. The power of OCT is demonstrated in a scanned image of a macular hole, a disease of the eye (Figure 1). Difficult to distinguish from normal retinal wear-and-tear on direct examination, the disease becomes readily apparent when viewed using OCT[3]. While OCT may be used in both in vitro and in vivo contexts, the focus of this paper will deal with the development of an in vivo OCT system, first giving a brief overview of a general OCT system.

Workings of an OCT System

OCT operates on the same physical principles as sonar, except that broadband light rather than sound is used.[4] OCT systems employ broadband light because of their low coherence length (coherence is a term describing light waves that are in phase in both time and space domains). A low coherence length is desirable: The lower the length, the greater the spatial resolution an OCT system may achieve. Any OCT system consists of a light source, a beam splitter, a sample arm, a reference arm, an optical gate, and a detector. The light source in an OCT system emits near-infrared light that is split by the beam splitter into two separate light beams. One beam is sent towards the sample arm while the other is sent towards the reference arm. The sample arm beam travels until it hits the tissue sample, where it then penetrates several micrometers deep before scattering back.[5] When the scattered sample beam returns, it passes through an optical gate, such as a Michelson white-light interferometer[6], that filters away all but the minimally scattered photons. Only minimally scattered photons are useful in constructing an image; highly scattered photons cannot effect a clean, detailed image. The reference arm beam travels until it hits and is reflected off the adjustable reference mirror. The filtered sample beam and the reference beam are recombined at the detector after they both return (Figure 2). The detector traces interference between the two light beams: The two beams interfere only when the pathlength difference between them is within the coherence length of the source.[7] When interference does occur, a depth-resolved reflectivity profile is generated for that part of the sample by measuring the light intensities of the backscattered photons as a function of their axial and transverse positions in the tissue. A series of adjacent scans as the sample arm travels transversely across the sample then yields a complete reflectivity profile.[8] This entire process is known as the acquisition phase of a scan.

A constructed image of the tissue sample is then generated from the reflectivity profile through a series of image processing steps. First, the reflectivity profile is scaled and mapped to a smaller, more manageable subset of itself. Secondly, the light intensities of the profile are mapped to values in either a false color domain or a gray-scale one. Now formatted as image data, the profile undergoes additional image processing, such as distortion removal, before being displayed on a host computer. Depending on the application, additional image processing may occur to tailor the output to the application's needs. Collectively, this set of steps is referred to as the processing phase of a scan. Unlike the acquisition phase, the processing phase is dependent on the implementation of each OCT system and may vary.

Issues with *In Vivo* OCT Systems

An *in vivo* OCT system must deal with additional issues not addressed in the basic setup. First is the issue of image contamination. Patients, being live biological organisms, are always in motion. Even when supposedly sitting still, their eyes blink, their hearts beat, and minute vibrations course throughout their bodies. Consequently, the tissue sample in an *in vivo* OCT system is never at rest, which poses problems for any OCT system. If during a scan-

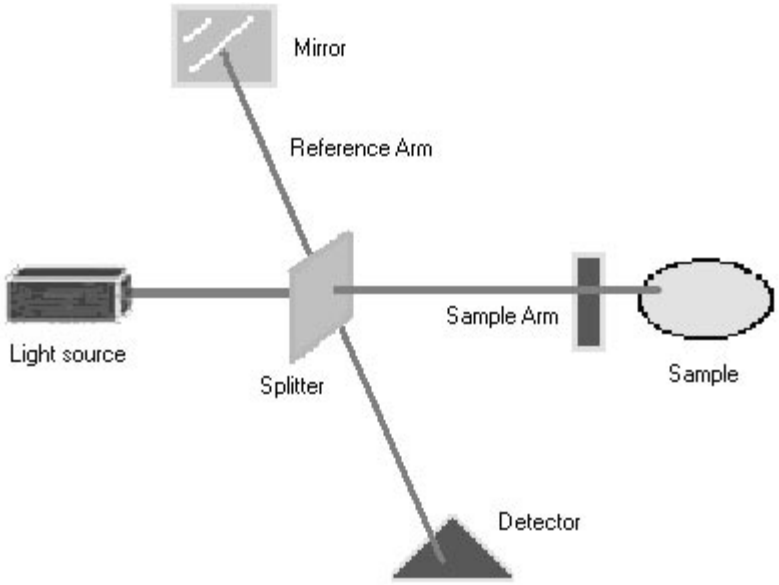


Figure 2: A typical OCT setup. The light source emits a laser beam, which is splitted into two by the spillter. One beam goes towards the sample arm and is backscattered by the sample. The other heads to the reference arm and is reflected back by the mirror. The two beams combine at the detector, which checks to see whether they interfere. Interference leads to a reflectivity profile found by measuring the backscattered photon intensities as a function of their position in the tissue.[7]

ning process the tissue under the scanning region shifts and moves, the resulting image will lack clear boundaries; it would be blurry. For large degrees of movement, the image becomes useless. The term motion artifacts describes the set of image contamination caused by motion. Motion artifacts generally arise from the patient moving about; however, it may also arise when the sample arm moves in a manner undefined in its specification. For example, scans of internal tissue are performed via a human-controlled catheter inserted into an anesthetized patient. Wiggling or shaking the catheter during a scan will result in motion artifacts.

The second issue is speed. *In vivo* systems lack the luxury of time given to their in vitro counterparts. Given a time consumption of x seconds from acquisition to display, and n desired number of images, the total time spent in scanning would equal nx seconds. For either a small x or n, the time commitment is minimal. However, the number of desired images is typically large (in the hundreds), so x must be small. A small x is also advantageous in terms of patient comfort—especially for internal scans where patients are anesthetized beforehand. The less time required to perform a scan, the less time the patient needs to be put under, and the less chance of complications or discomfort. Moreover, a large x also worsens the motion artifact problem: Motion artifacts alone will only contaminate an image of the current tissue sample (bad enough as it is). Motion artifacts coupled with a slow process time may result in a contaminated image of a tissue region that is no longer even the current region being scanned.

Designing an *In Vivo* OCT System

One possible *in vivo* design addresses the motion artifact and speed issues by optimizing the processing phase of an OCT scan. The rationale is that a faster processing phase would accelerate system performance and consequently alleviate the more extreme motion artifacts. At the Optics Group in the Research Laboratory of Engineering, this is the design we chose to implement.

As previously mentioned, the processing phase of an OCT scan consists of the conversion of a reflectivity profile of the tissue to a reconstructed image of the tissue. Optimizing the speed of the processing phase would therefore require optimizing the conversion algorithm that transforms a reflectivity profile to a displayable image. Under our design, we resolved the transformation into four key steps: light intensity homogenization, color domain mapping, removing scanning distortions, and converting between coordinate-spaces. Optimizing the conversion algorithm would require optimizing these four steps.

ABCDEFGHIJKLMNOPQRSTUVWXYZ
ZYXWVUTSRQPONMLKJIHGFEDCBA
ABCDEFGHIJKLMNOPQRSTUVWXYZ
ZYXWVUTSRQPONMLKJIHGFEDCBA

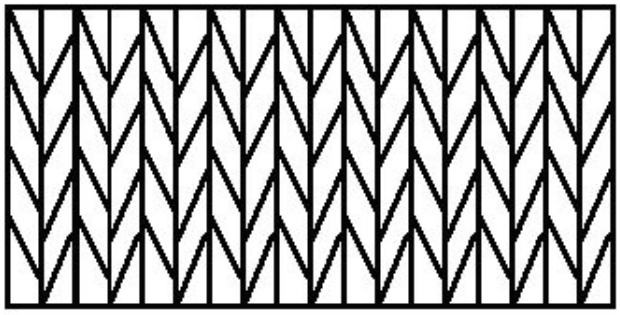


Figure 3: (Top) Writing a continuous sentence on a word processor that outputs both left to right and right to left reverses the orientation of every other line (in gray). In this example, the word processor is assumed to hold only 26 letters per line. (Bottom) A similar situation occurs in a reflectivity profile: because the sample arm traverses the sample crosswise, the orientation of every downward scan is opposite thatof every upward scan. This effect is known as bi-column inversion.

Light Intensity Homogenization

The first step involves homogenizing the light intensity values, making relatively equal intensities exactly equal. The point of OCT is to image tissue structure such that important features are easily discernible. Relatively minute differences in light intensities should therefore reflect minute differences in the tissue, while relatively large differences in light intensities should reflect large differences in the tissue. For example, a light intensity of 100 and a light intensity of 1000 differ significantly from each other and probably represent different physical layers in the tissue; these two intensities should not be homogenized. On the contrary, a light intensity of 100 and a light intensity of 101 are nearly equal and probably represent the same physical layer in the tissue; these two intensities should be homogenized. Any conversion algorithm must perform accordingly.

Color Domain Conversion

The second step concerns the conversions of the light intensity domain and the color domain. Colors are described on a computer according to a Red, Green, and Blue (RGB) specification that allocates x bits to Red, y bits to G, and z bits to Blue. The exact number of bits to allocate per color depends on the total number of bits currently assigned per color value. For example, under a typical 16-bit specification, the bit allocation may be 5-6-5: 5 bits to Red, 6 bits to Green, and 5 bits to Blue. However, this is a conventional allocation—there is nothing to prevent a 16-bit allocation scheme that assigns all 16 bits to, say, Blue. The RGB specification itself is also not strictly enforced. Colors may be described in terms of RBG or any other permutations of the RGB schemata. When mapping the light intensity values of the reflectivity profile against the color values of a color domain, the conversion algorithm divides, by bits, the light intensity values into the appropriate color representation of the host computer.

Bi-Column Inversion

The third process involves ameliorating distortions introduced into the reflectivity profile. As noted before, a reflectivity profile is formed from consecutive adjacent scans as the sample arm travels transversely across the tissue. The transverse motion of the sample arm, while efficient, leads to a negative side effect: The orientation of every downward transverse scan is opposite to that of an upward one. To grasp the idea, think of typing a continuous sentence (without carriage returns) on a word processor that outputs both from left to right and from right to left. Every other line of text would read in the opposite direction from the previous line. The same is true of a scanned constructed image, except that every column rather than line is inverted. This phenomenon is known as bi-column inversion (Figure 3). Converting a reflectivity profile to a displayable image requires reorienting every other scan column in the profile.

Coordinate Space Mapping

The final step entails coordinate space conversion. In some situations it makes sense to display an image in other coordinate systems other than Cartesian. Catheter scans of internal tissue and retinal scans are instances where a polar representation is preferable. A conversion scheme from Cartesian to polar is unfortunately problematic. Not all integer coordinates in Cartesian space map directly to integer coordinates in polar space. However, images are always internally represented as a collection of color values with integer coordinates. Any conversion scheme must therefore ensure that noninteger coordinates are mapped to integer ones. Such a mapping, though, introduces other complications: Depending on its implementation, data overload and loss can occur. Two distinct points in the Cartesian representation may be mapped to the same point in the polar representation, leading to data overloading and subsequently to data loss at that point.

References

1. Huang D, Swanson EA, Lin CP et al. "Optical Coherence Tomography," Science 1991, V254 1178-1181.

2. The New York Eye and Ear Infirmary. "Macular Hole and Normal Fovea," www.nyee.edu/glaucoma/oct-data.htm

3. Swanson, EA, Izatt JA, Hee MR et al. "In-vivo Retinal Imaging by Optical Coherence Tomography," Opt Lett 1993, V18 1860-1866.

4. The New England Eye Center. "Optical Coherence Tomography," www.neec.com/medical/OCT.html

5. Cottrell GW, Morhenn V, Chen Z. "Early detection of Basal Cell Carcinoma using Optical Coherence Tomography and Pattern Recognition Techniques," Technical Report, Mar 2000, UC San Diego.

Implementing an *In Vivo* OCT System

The implemented algorithm optimized three of the four steps in the processing phase. The light intensity homogenization step was optimized by mapping the reflectivity profile onto a logarithmic scale. In the interest of increasing processing speed, a less accurate bit-masking algorithm replaced a real logarithm function. This implementation reduces the number of required mathematical operations by a linear factor. Color domain conversions were handled by stipulating an RGB specification of 5-6-5 and by mapping the light intensities into a color lookup table. The light intensity values become the indices to a predefined table of color values. Passing the reflectivity profile through the color lookup table resulted in formatted color image data. Coordinate space conversions were handled in a similar manner using a coordinate-space lookup table. Having only integer coordinates in the lookup table enforced integer-to-integer coordinate mapping. Linear interpolation then assigned values to any unmapped points in the polar space. Performance was faster using this implementation than using pure mathematical transformations by an exponential factor in time, but more costly by a linear factor in space. The savings in time occur only with pointer arithmetic (to reference the lookup tables); the loss in space resulted from the needed memory to hold the color and coordinate-space lookup tables. Bi-column inversions were treated by memory swapping data. This step of the processing phrase saw no improvement.

Overall, the design improved from earlier ones by several factors in processing speed. The rate of process and display measured at between 7~10 fps, which is an improvement on earlier systems running at under 1 fps. The improved speed, however, still falls quite short of video frame rate (typically ~ 30 fps). It is highly likely that hardware changes to the OCT setup (i.e., improved sample arm speed) would be necessary in conjunction with optimizing algorithms to achieve a real-time system. Improvements in motion artifacts were not clocked under the new system due to a lack of live test subjects at the time.

Further Research and Applications

In the *in vivo* context, there is still much work to be done to remove motion artifacts and achieve a real-time system. OCT in general, however, is also wide open to further research. One of the most exciting research areas today is in extending OCT imaging to the three-dimensional realm. By stacking consecutive, cross-sectional OCT images along the z-axis, a three-dimensional representation of the scanned tissue sample may be obtained. Depending on the number of cross-sectional images used, the 3D images can become incredibly detailed. The possibilities for coupling 3D images with a real-time *in vivo* system are profound: An exact, virtual replica of the tissue sample would allow doctors to rotate, manipulate, and examine the tissue from every angle without ever touching it.

Another area attracting much research is the combination of OCT imaging with artificial intelligence. At its heart, *in vivo* diagnosis techniques do nothing more than match a set of symptoms with a set of possible diseases. For example, a strep throat diagnosis involves following a checklist of symptoms:

* Does the patient have a sore throat?
* Is the patient feverish?
* Does the patient have white pus on the tonsils?
* Does the patient have tender lymph nodes?
* Is the patient coughing?

From the answers to these questions alone, a physician can give an accurate diagnosis. The same holds true for diagnosing diseases based on OCT images—a list of physical characteristics in the images corresponds to a diagnosis for the disease. This paradigm of matching a set of patterns (the symptoms) to a set of outcomes (diseases) is essentially a rule-based system; the symptoms are antecedents to rules whose consequences are the diagnoses. This is a well-known area of research in AI. It may one day be possible to fully automate diagnosis using OCT: A computer generates scanned images of the tissue sample, matches the images to a database linking sets of images to diseases, and outputs a diagnosis.

OCT may one day expand to other fields outside of medicine. For example, the technology is ideally suited to developmental and molecular biology, as well as industrial uses. Because of its high resolution and nondestructive qualities, OCT may have uses in materials production as an imaging tool for polymers and composites. In summation, it must be noted that the potential for OCT is still expanding, with its promise yet to be fully realized.

Acknowledgements

This paper is based on undergraduate work done between 1999 and 2000 under the supervision of Dr. Xingde Li and Prof. James Fujimoto of the Optics and Quantum Electronics Group in the Research Laboratory of Engineering at MIT.

6. University of Florida Medical College. "Optical Coherence Tomography," www.medinfo.ufl.edu/other/profmed/slides/luttge/slide20.html

7. Everett M, "Optical Coherence Tomography," http://lasers.llnl.gov/lasers/mtp/oct.html

8. Lankenau IE, Welzel J, Birngruber R et al. "In Vivo Tissue Measurements with Optical low Coherence Tomography," SPIE 1997, Paper 2981-11 78-84.

Suggested Readings

1. Rollins A, Yazdanfar S, Kulkarni M et al. "In Vivo Video Rate Optical Coherence Tomography," Optics Express 1998, V3.6 219.

2. Yazdanfar S, Kulkarni M, Izatt J. "High resolution imaging of in vivo cardiac dynamics using color Doppler optical coherence tomography," Optics Express 1997, V1.13 424.

3. Brezinski ME, Pitris C, Boppart SA et al. "Micron scale optical biopsy with optical coherence tomography," Journal of Clinical Oncology 1997, V15.3 972.

4. Ghanta RK, Fujimoto JG. "Ultrahigh Resolution Ophthalmic Structural Imaging Using Optical Coherence Tomography," HST Forum 2002 http://hst.mit.edu/forum/abstracts.

The Science and Politics of Global Warming:

The Climate of Political Change at MIT

Patsy Sampson

"Climate Experts Clash on Global Warming." This could have been the headline from The Washington Post or The New York Times or any major publication any time over the past 12 years. Why it was the headline of the Massachusetts Institute of Technology's Tech Talk provides an ongoing story about science, research, and politics at MIT and elsewhere.

In 1988 explosive hearings in the Senate were making major news out of global warming. Dr. James Hansen, a climatologist from the NASA Goddard Institute for Space Studies (in New York City) had revealed in the public forum of the Senate hearing floor that there was a 99 percent chance that the earth's temperature was rising and that there was some greenhouse warming. Seizing on his statements, the environmental advocacy movement adopted the issue. It was soon implied that mankind had no choice but to take measures now to prevent climactic disaster. Overzealous human burning of fossil fuels, filling the atmosphere with carbon dioxide (CO_2), was to be the culprit for the warming greenhouse effect. Among the more hysterical reactions were news stories claiming that "the resultant rise in the [Pacific Ocean] could wipe out San Francisco and Oakland airports with damage in the billions."

At MIT in June of 1990, the "clashing climate experts" were the focus of a Tech Talk story describing a major forum that had been held as a year-end alumni event attracting, as it turned out, more than alumni to a packed Kresge Auditorium. In fact, the audience was filled with students, MIT faculty, and members of the press. Although the general public and even this audience were not sophisticated enough to understand or fully appreciate the complex scientific discussion of atmospheric conditions involved in debate, they did catch the urgency and depth of conviction presented by two of the panelists, MIT climatologist Richard Lindzen and Dr. Stephen Schneider

of the National Center for Atmospheric Research (NCAR) in Boulder, Colorado.

As Dr. Nazri Choucri, professor of political science at MIT, concludes in retrospect, "The early stages of the global warming debate were focused on the science; the politics of the issue was still unspecified, without yet assigning winners and losers." She and Henry D. Jacoby, professor at the MIT Sloan School of Management, were also on the panel. It was clear at the forum, however, that the real "gladiators" in this "spectacle" were the scientists Lindzen and Schneider. Then as now, an open collegial debate looked the media and the lay public as a sensational "clash."

Likewise, the complex science issues of the global warming debate continue to perplex climatologists, leaving prediction open to wide uncertainty. Professor Lindzen pointed out that if the large models that were predicting (in 1990) a 4-degree-Celsius (1 degree Celsius is equal to 1.8 degrees Fahrenheit) temperature rise within the next century for a doubling of carbon dioxide were true, then a 2-degree-rise should already have occurred. As he stated, "This [the 2-degree rise] is something that has not happened." Although, the two climatologists were far from agreement on most issues, they both acknowledged that the lack of scientific understanding of the role of clouds and water vapor in the atmosphere was cause for the greatest uncertainty in the models. The situation, 10 years later, is still cloudy, and the pursuit of the science of this and other global warming issues at MIT is pervasive.

Dr. Chris Forest, researcher with the MIT Joint Program on the Science and Policy of Global Change, describes the overall picture of climate with reassured ease. (He is a climate modeler, individuals who generally are—or appear to be—a more relaxed breed, somehow removed from the climate hype. (See insert.) Forest explains that the atmosphere changes quickly, as anyone who experiences New England weather might attest. The effect over time, known as climate, or the average of these cumulative atmospheric variations, is to mix things up. Other influences on the atmosphere include the earth's land masses and vegetation; human uses of these land areas including agriculture and deforestation; the tremendous effect of the ocean, which accounts for nearly two-thirds of the earth's surface; and of course the input of energy from the sun.

Although climate modeling is his present research area, Forest admits that he is somewhere between being a modeler and a theoretician (see insert), as he continues to avidly describe more fundamentals of the climate question. His early training was in physics, and he is completely at home discussing any aspect of the question.

The atmosphere, he continues, is the medium connecting all these factors, and complicated physics concerning winds, convection, and energy transfer processes must be understood to describe its behavior. Over the long term, the "forcings," or additions, to the atmosphere of added gases by human input or otherwise (such as volcanic eruptions) can bring about, in geologic terms, a fairly "quick" atmospheric response. In contrast, the ocean absorbs greenhouse gases very slowly, including CO2, showing the effect only after long time periods, retaining these dissolved gases for as long as 1,000 years.

How sensitive the ocean is to perturbations, such as additions of large amounts of CO2, is the focus of graduate student Veronique Bugnion in the Department for Earth, Atmospheric and Planetary Sciences (EAPS). Her computer screen on the 17th floor of the Green Building at MIT is revealing: models are blazing blue, showing various ocean configurations in 3D. She explains, with a directness and enthusiasm reminiscent of Chris Forest's, that she looks at the sensitivity of the ocean surface for disturbances that will change the large circulation(s) occurring in its depths. These circulations could be very significant in influencing climate. She uses a coupled model. In other words, the ocean model is coupled with a simpler atmospheric model to study the influence of winds, rain, and atmospheric gases on the surface waters. In her atmospheric model she can increase the CO^2 concentration and get an idea what the effect might be on her 3D ocean model.

The mathematical equations, admits Veronique, are "semifancy and applied on a pretty big and sophisticated model, so it does end up being fairly complicated. But after a while it becomes second nature to you. It feels like home. All these codes and numbers on my machine [computer] mean something to me, but they won't mean anything to most others," she tries to reassure a layperson who happens to ask about her work.

The North Atlantic is Veronique's territory. The major circulation of the North Atlantic, she explains, is the prevailing thermohaline, or heat-and-salt-sensitive conveyor belt, circulating ocean water between the surface and the depth of the ocean over tremendous areas and depths up to nearly three miles. Very cold, salty waters will sink in the far north and travel southward toward the equator. Eventually, these waters, after warming and becoming less saline, will rise and flow north again, joining the Gulf Stream. In a global warming, or increased CO^2 scenario, the warm-

ing of the oceans will bring about a weakening of the North Atlantic thermohaline circulation by decreasing the sinking of the cooler waters in the northern seas and by lowered salt concentration because of increased rain. In simplistic terms, this would ultimately lead to a lessening of the surface current flow, such as the Gulf Stream, whose warming influence is the lifeblood of European civilization. In the 1300s this phenomenon actually led to "the Little Ice Age," nearly bringing life in that area to a standstill.

"Just as the sun's energy drives the atmosphere," Veronique explains, "the atmosphere is apparently the driver of the ocean." Perhaps with a slight bias, she emphasizes that a great many more uncertainties and unknowns about this vast mass compared to the atmosphere, so the ocean is the wild card in the global warming debate. Little is known to date about ocean behavior with respect to the atmosphere, the prevailing winds, the heat and energy exchange between the interface of the ocean and the atmosphere, and the internal large-scale ocean circulations. In her words, "over time, the ocean may be the slower respondent to the warming scenario, but it may be the more important key."

Just down the hall, with a spectacular view of the Charles River, framed by its now greening banks and the city of Boston beyond, is the very well-used office, full of papers and shelves of scholarly volumes on climate, of Richard Lindzen. He has been a student of clouds and their behavior for most of his career. It is apparent, in the face of global warming, that his scholarly observations have also included the political behavior of people, sociological groups, and governments. What amazes Lindzen is how this issue, from a journalistic perspective, has dragged on over eleven years and still gets coverage. "Although the coverage may be going down," he observes, "politically, the issue has a life of its own." Who would know better than Professor Lindzen, who has been assigned by the media the title of "climate contrarian."

He recalls a 1983 in Stockholm of meteorologists who would eventually form, under the wing of the United Nations, the Intergovernmental Panel on Climate Change (IPCC). At this meeting, the IPCC founder, Swedish meteorologist Bert Boline, suggested that a mechanism was needed to launch global warming as an issue. This mechanism became, as Lindzen pointed out, the famous Senate hearing in 1988. Taking action has become very difficult—where all involved are judged on environmental correctness. At MIT, he points out, "the situation is not bad. The level at MIT, where everyone has a reputation as a scientist, is much higher than at the IPCC, where the reason many contributors have a reputation at all is solely due to their association with this group. Very little of the true science of climate research reaches the political public. And most politicians [bureaucrats] would rather have a month of eating gruel than read science and would accept three weeks of gruel rather than read economics!"

Professor Lindzen's work on tropical meteorology, and more recently on the mechanisms of positive and negative feedbacks in terms of cloud physics, is lofty, and so it is helpful to get a grounding in the basics first. Chris Forest, again eager to share in the theories behind the models, provides some facts concerning the greenhouse principle and how water vapor plays such a key role. On the second floor of the Energy Lab, in a nondescript but clean and efficient brick structure, the mysteries of CO_2 and water vapor, unfold.

Everybody talks about the greenhouse effect but most know very little about it. It may as well be the environmental equivalent of the Internet—it's everywhere but most people don't know how it got here or how it works. Forest compares the global greenhouse to a huge blanket surrounding the earth. The atmosphere naturally contains a large amount of water vapor (1.5—3%, a "large" number in comparison with other greenhouse gases). Each water vapor molecule can reside in the atmosphere for as long as a week before precipitating back to earth. Water vapor is the chief greenhouse gas, although CO_2 and several other gases, including methane and nitrous oxide, act with it. This atmospheric mix behaves like a blanket preventing the escape to space of longwave radiation that the earth sends back after being "warmed" by the sun's shortwave radiation.

As long as the earth has existed, it has been surrounded by gases, though the concentration of CO_2 has changed drastically over the earth's roughly 4 billion years. Presently, the amount of CO_2 in the atmosphere is 27 percent higher than it was at the beginning of the industrial revolution, 150 years ago, when the increased combustion of fossil fuels by humankind initiated the release of added greenhouse gases, including CO_2. In addition, CO_2 molecules in the atmosphere don't go away easily. The average CO_2 molecule has an atmospheric lifetime of roughly 100-120 years.

In terms of the greenhouse mechanism, says Forest, "Think of the earth as a red-hot-poker. It will give its heat off to the air around it when removed from the fire." The significance of the

"greenhouse gases," including water vapor and CO_2, is that as molecules they have unique abilities to "accept" or be activated by this outgoing longwave radiation. Water vapor has a particularly wide band or ability to accept this energy, whereas CO_2's acceptance of energy is much less (narrower). Nevertheless, because CO_2 accepts the energy that the water vapor cannot, the overall outcome is that the combined gases take on more energy, resulting in warming of the atmosphere and eventually the ocean. Warmer air means greater evaporation, which leads to more water vapor. Thus, there is a cumulative effect known as a "positive feedback."

But the situation is not so simple. It turns out that in terms of feedbacks there are competing effects. As the atmosphere warms and clouds increase, the resulting cloud cover can act as a cooling shield, reflecting the initial incoming shortwave radiation from the sun; i.e., negative feedback. Whether the clouds are high or low and over what region of the globe they cover also influence the feedback character—positive or negative.

Ten years ago, Professor Lindzen had suggested that in a warmer world taller clouds, which detrain or spread out and dissipate at higher levels (about 10—15 km), would leave less water, reducing humidity. This would allow more of the longwave radiation to escape from the earth's atmosphere, and ultimately produce a negative feedback. More recently, he and his colleagues have concluded that the upper atmospheric water vapor is actually detrained ice is precipitated from very high cirrus clouds producing moisture in the environment. Ultimately, in a warmer world, it is possible that more rain would reach the surface and leave less water available to reach the upper atmosphere. This would imply less detrained ice and less cloud cover, allowing far more effective cooling. Using observational data, Lindzen's group has concluded that this effect is approximately four to five times bigger than most models' water vapor feedback and of the opposite sign, i.e., a negative feedback, resulting in cooling. The effect of Lindzen's finding would knock the water vapor factor (the positive feedback) in these models down to a much lower significance. Historically, no climate model shows anything but positive feedback for water vapor. Presently, Professor Lindzen's paper on this work is in review, the process in the scientific world which allows for peer appraisal prior to publication.

"It is fascinating how the whole notion of water vapor feedback arose in the absence of data," Professor Lindzen muses. "There was no theoretical foundation. It was supposition but it held sway because it was the way the models behaved."

"Philosophically," he concludes, "this will take yet many more years to sort out, but the science, in spite of the politics, will evolve."

Despite the uncertainties, or because of them, the ongoing global warming climate change research at MIT is extensive, including theoretical and practical approaches. In the Parsons Lab, a block away but still under the gaze of the towering EAPS structure, a group of hydroclimatologists looks patiently and persistently at the effects of climate change from a regional perspective. These are more down-to-earth researchers, looking at real-world data to track regional flood and drought conditions.

Graduate student and aspiring hydroclimatologist Luis Prado-Perez is studying the 30-year drought of West Africa. A native of Puerto Rico, Prado-Perez has had direct experience being at the mercy of the weather. In his good-natured but determined and dedicated way, he explains that the African drought problem might be affected by the overall global warming situation, but his research is much more specific to the region. He is concerned with studying the mechanisms and reasons for prolonged drought—whether it is caused by global warming or not.

The geography of West Africa is influenced more than any other land mass by the actions of two large air circulations, the Hadley Cell and the Monsoon Circulation. Luis carefully describes the action of the air rising along the equator to 10—15 kilometers vertically. During this rising, "not the subsequent falling as you might think," he cautions, there is the most likelihood of rainfall. Thus the tropics have a climate that is always warm and moist. Eventually, the air will stop rising, on average, and flow in a large arc, swooping down toward the earth at a higher latitude. Lying along the equator, the Hadley Cell occurs all around the globe like a very large inner tube, covering up to 40 degrees lattitude with rising air in the summer hemisphere and sinking air in the winter hemisphere.

On the other hand, West Africa is also exposed to the effects of the Monsoon Circulation, which is entirely a product of the temperature gradient created by the warm sea and land proximity. In the best of situations, the action of both the Monsoon Circulation and the Hadley Cell in the summer would mean added moisture for West Africa. But in a persistent drought, Prado-Perez explains, the regional vegetation and resident soil moisture play a significant feedback role. He looks at the energy in the boundary layer, the vertical parcel of air influenced by the land below it.

With less moisture given off by decreased vegetation due to overgrazing and forest clearing, the air circulation will contain less moisture, equivalent to less energy, and result in less precipitation. Why the process is self-perpetuating is the focus of Prado-Perez's research.

West Africa is an ideal regional model to study because it has tropical forest, savannah, and desert. But there are difficulties as well. For all kinds of reasons—lack of universities and data collecting centers and even wars—available and reliable data other than basic temperature and precipitation measurements, are scarce. In addition, the drought started 30 years ago, and reliable measurements are only available for the period starting after the drought. So Prado-Perez and his colleagues must rely on input from satellites and weather balloons, collected in central places such as weather centers in the UK and in Germany. If he and his colleagues wish to make studies involving CO_2, data for CO_2 measurements must be input globally to study the regional effects of varying vegetation and surface moisture. Studying these effects using 2D computer models is a significant part of his work.

Climate Modelers: Are they the same as Climate Theorists?

At the Massachusetts Institute of Technology, a respected member of the faculty of science in the Department of Earth, Atmospheric, and Planetary Sciences, has changed his title from "Professor of Meteorology" to "Professor of Climate Dynamics." As one of his research scientists, Dr. Chris Forest, notes, "This is the closest I've heard of anyone trying to be called a 'climate theorist.'"

Yet Professor Stone has had more than 20 years' experience in developing climate models, including the first general circulation model (GCM) at the Goddard Institute for Space Studies (GISS). GCMs are generally two- or three-dimensional models that are used to study global climate, covering areas in grids of up to 250 square kilometers. The models are based on intricate mathematical equations describing the multiple components for each grid cell that are needed to get an accurate picture of the system.

For the past 10 years, Professor Stone has been the director of MIT's Climate Modeling Initiative. To get an idea of how this affects his approach to global warming, one might ask him, as a layperson, about his perspective. His first statement will concern the high degree of uncertainty in predicting the future of global climate and therefore the use of a modeling system which will allow for a wide range of possible climate inputs, or "forcings," including greenhouse gases balanced with atmospheric, biospheric, and ocean reactions to them. He will show how the range of predicted temperature rise goes from 1 to 5 degrees Celsius within the next century. This is not necessarily troublesome to someone like Peter Stone. His abiding interest in climate is driven by his fascination with what can be accomplished with the present computer models and even more by his faith in the future capabilities of these models. (He does reveal that he thinks that the reality of temperature rise lies in the 2 degree range.)

Like Professor Stone, a number of his students and research workers at MIT consider themselves climate theorists who may have started out as physicists or mathematicians, but developed interests in climate as well as computational science. As Dr. Forest observes, "When computational science came into the game, the term modeler was used to separate those theorists who did things on paper with elegant mathematics from those who tried to compute answers to practical questions, like tomorrow's weather." Climate modelers often find, as they load their computer models with the complicated simulations needed to imitate global climate, that they must turn to climate theory to explain the results when these do not make sense with real-world data or phenomena.

Unfortunately, there is no accessible way to experiment with the weather in a laboratory. A computer model becomes the next best thing. Forest expresses some frustration saying, "If we had our own toy world to experiment with, we might not need to simulate it in a computer." In his opinion, the frontier of climate modeling and the point of greatest learning is when a climate model is "pushed and experimented with even to the point that it might break or blow up (not literally, but just crash). It is like throwing a bunch of chemicals together to see how they react," he explains.

To an outside, nonscientific observer, the day-to-day work of a climate modeler would appear very dull. Indeed, Dr. Andrei Sokolov, research scientist under Professor Stone at the Joint Program on the Science and Policy of Global Change, describes his work in a thick Russian accent with some humor, "My work is very boring." As he points out, there is active climate modeling in his native Russia, but "they have a lot more things to worry about"; the climate modeling there is not well supported. He is obviously enthusiastic about his recent work on building a coupled

atmospheric-ocean model to show how the ocean will act as the world's most significant absorber of CO_2 and to show calculated findings for thermal expansion of the ocean as it absorbs the heat from global warming. Nevertheless, Dr. Sokolov critically admits that, at most, 10 percent of the time spent modeling is "useful" (leading to new creative approaches or thinking), "the rest is in the trash." Although climate models created by other climate centers are often available for use, Andrei Sokolov finds that the process of developing a model is more interesting in itself.

Physically, the kind of computer that can handle the large computing task required for the Integrated Global System Model at MIT is now run by the Lab for Computer Science. There, a mainframe called the Pleiades Cluster Project (Pleiades for short) is used at least half-time for the work of the Climate Modeling Initiative. The Pleiades cluster can process a 3-D atmospheric-ocean GCM model in as little as 2 hours of computer time. On a relative basis, the cost of such a computer is now about $100,000. Fortunately, the funding for such equipment is readily available to the MIT Climate Modeling Initiative through government and private sources.

Other laboratories at MIT, such as those relying on computer modeling at the Parsons Lab in the Department of Civil and Environmental Engineering, must rely on supercomputers at outside centers such as the National Center for Atmospheric Research (NCAR) in Boulder, Colorado. To achieve a decent regional model output for a three-month prediction, for example, Jeremy Pal estimates that up to 36 hours of computer time would be required. As Luis Perez-Prado in the same laboratory admits, it is better to use a two-dimensional model, which requires less run time and less money, with more results, than using a three-dimensional model. Funding for these modeling runs comes mainly through National Science Foundation research grants.

Ideally, the theorist and the modeler will be one and the same or at least collaborate as colleagues. Professor Richard Lindzen's latest work involving predictions of cloud behavior and feedback response in tropical environments will soon be tested by putting a parameterization of his and his colleagues' predictions into a model to see if the effect of their theoretical estimate is in agreement with the model's prediction. Professor Lindzen urged in the spring of 1990, "it is a pressing obligation of meteorologists [climatologists] and oceanographers to find out why the models are wrong." There appears to be some progress in this direction.

As long as computing power continues to expand, there will be more opportunities to equally improve the climate modeling process. The best scenario is that this progress will correspond to increased climate research and theoretical knowledge, and that climate modelers will be, by training, climate theorists.

A Modernized Medicine for Our Times:

Reforming the American Health Care System

Jasper James Chen

America's health care system is unparalleled in terms of its physicians, hospitals, and prompt and extensive treatments. People from all over the world come to our hospitals for the most advanced treatments known to man and to receive care from renowned physicians. As the largest service industry in the United States, health care is also a major contributor to our nation's robust economy. Health care spending constitutes approximately 20% of our GDP and is a "business beholden to, and a force in, the economy."[1]

However, our health care system is in need of an overhaul. For one, an increasing number of Americans are unsatisfied with our health care system. The elderly have expressed a need for Medicare to defray the costs of their prescription drugs, a current issue for the two presidential candidates. Working-class Americans are also concerned about the possibility of losing their insurance coverage and spiraling health insurance premiums. Additionally, by pulling out of Medicare and other markets, HMOs are not helping the situation. All of this results in diminished access to health care.

Failing to provide a minimum level of health care for all citizens has negative consequences for America as a whole. One chief weakness of our health care system is that one-third of Americans—the uninsured and underinsured—lack access to essential health services. A population that is not healthy cannot be fully productive or reach its potential. In our increasingly competitive world, human resources are precious assets, so the entire nation gains if all its citizens can attain optimal health.

In this report, I hope to achieve the following: 1) describe the problems resulting from lack of health care access, 2) delineate recent health care reform attempts, 3) explain how health care is inextricably bound with the politics of our nation, and 4) suggest some possibilities for modernizing our health care delivery and financing system to meet the needs of the 21st century.

Problems Resulting from a Lack of Health Care Access

More than 43 million underprivileged Americans lack adequate access to health care services. The Health Insurance Association of America projects this number to reach 50 million by the year 2007. Governmental and societal efforts to establish reasonable equity for members of all socioeconomic classes in accessing health care constitute a valuable national objective. A nation with self-respect and decency will do everything in its power to provide good health for all its citizens. One can hardly expect equity in a health care system that is fragmented, professionally controlled, hospital centered, middle-class in its orientation, and unresponsive to the social, physical, and biological conditions in which the urban and rural poor live.

The medically uninsured and underinsured often resort to more costly emergency rooms; the delay in receiving medical care may exacerbate their existing condition as well. Underserved urban populations have turned to the emergency rooms and the clinics of neighborhood hospitals in order to receive episodic health care treatment. The problem with this is that they are deprived of any real continuity in the patient-physician relationship-a serious shortcoming in the quality of care. An individual who receives medical care as a charitable service by a voluntary hospital or free clinic is not on a par with the vast majority of citizens who can obtain necessary medical care because they have insurance.

Overview of Recent Attempts at Health Care Reform

The Clinton Health Care Proposal

Only recently has health care arrived at center-stage in American politics. Issues that have traditionally dominated US politics include the following: economic stability, foreign affairs, crime, defense, and education. In order to ameliorate the plight of the medically underserved, numerous reform efforts have been attempted. Perhaps the most ambitious of these was the Clinton Administration's attempt for universal health care, which occurred in 1994.

What President Clinton did—and which set him apart from other presidents who have introduced proposals for some form of universal health coverage, including Truman, Nixon, and Carter—was to base his proposal expressly on the premise that access to comprehensive health care is a fundamental right for all Americans.[2] In this manner, Clinton attempted to help those underprivileged and uninsured individuals who needed access to health care the most.

However, why did the Clinton program fail? For one, his health care initiative could not find the support of the health industry and was too complicated to make it through congress.[3] Second, the political will was nonexistent at the time. Jeffrey C. Merrill, former principle health economist for the U.S. Congressional Budget Office and former director of legislation and policy for the Health Care and Financing Administration, asserts that in order for health care reform to occur, certain long-standing beliefs, practices, and vested interests built up over the years must be overcome.

Unless citizens perceive the current crisis has serious effects on them personally, significant change will not occur because of historical inertia and fear of the unknown. A critical mass of the American population needs to feel personally affected by any health care crisis in order to create the political will necessary to address this crisis.[4]

Not only was the Clinton health reform program overly ambitious, it was unclear in specifying how universal access to health care would be financed. Clearly there is a need to overhaul the structure of our health care delivery system if investing more health care dollars into the system will not result in a feasible solution to addressing the health care needs of the underprivileged. Rather, the money would eventually run out and more and money would be required to continue delivery of care.

The Idea of Managed Competition

What can be done with the uninsured and the underprivileged? How can there be a market to sell to these unfortunate individuals, who do not have deep pockets? A number of scholars have put forth theories. One of the most significant is the idea of managed care and managed competition. In the 1970s, Alain Enthoven, Professor of Public and Private Management and Professor of Health Research and Policy at Stanford, envisioned that:

"in a competitive environment of multiple and responsible consumer choice of health plan, groups of doctors, or networks of independent doctors, would find it in their interest to commit to work together to manage quality and cost of care for their voluntarily enrolled populations. They would accept total capitation payment for all the costs of care (with appropriate reinsurance for high cost cases). Among the many strategies these medical groups would pursue to improve quality and cut cost, would be:

1 Merrill, Jeffrey C. The Road to Health Care Reform: Designing a System That Works. New York: Plenum Press, 1994: 28.

2 Merrill, Jeffrey C. The Road to Health Care Reform: Designing a System That Works. New York: Plenum Press, 1994: 7.

3 Langdon, Steve. Congressional Quarterly, 27 JUL 1996.

Langdon, Steve. Congressional Quarterly, 3 AUG 1996.

4 Merrill, Jeffrey C. The Road to Health Care Reform: Designing a System That Works. New York: Plenum Press, 1994: 15.

5 Enthoven, Alain C. "Managed Care: What Went Wrong? Can It Be Fixed?" The Donald C. Ozmun and Donald B. Ozmun and Family Lecture in Management, The Mayo Clinic, Rochester, MN. 29 NOV 1999: Introduction

6 Herzlinger, Regina. Market-Driven Health Care: Who Wins, Who Loses In The Transformation of America's Largest Service Industry. Reading, Massachusetts: Perseus Books. 1997: x.

7 Merrill, Jeffrey C. The Road to Health Care Reform: Designing a System That Works. New York: Plenum Press, 1994: 49.

8 Vinluan, Frank. "Scholar hopes to help fix health care." The Seattle Times, 4 Aug 2000: Local News.

9 Mitchell, Alison. "Bush Spells Out Major Overhaul in Medicare Plan." The New York Times, 6 SEP 2000.

10 Ibid.

11 Ibid.

12 Ibid.

13 Ibid.

14 Ibid.

15 Ibid.

16 Ibid.

17 Alliance for Health Reform. Cost, Access, Quality: Managed Care and the American Health System. Washington, D.C., 1998: Section 7.4.

- Match the resources used to the needs of the populations they serve (thus correcting the specialty imbalances and surplus of beds and high tech facilities);
- Concentrate costly complex services in regional centers that would realize economies of scale and experience;
- Adopt what we now call Continuous Quality Improvement, including study of practice variations and adoption of best practices, process improvement or reengineering, benchmarking, analysis of errors and development of methods to prevent them, etc.;
- Follow rational personnel substitution and referral processes among specialists, generalist physicians and other health professionals;
- Evidence-based medicine and outcomes measurement and management, avoiding treatments that were inappropriate or that conferred very low marginal health benefits on patients relative to cost;
- And development of formularies based on the judgments of the physicians treating the patients in order to achieve a high degree of compliance, and the bargaining power that comes with it.

These groups or networks would team up with one or a few carriers as partners who would market their services, including pricing, advising on plan design, contracting with employment groups, and the many other functions they perform. Examples would include Med Center's Heath Plan at the St. Louis Park Medical Center, the arrangement the Kelsey Seybold Clinic had with one carrier as its main marketing partner, or the mutually exclusive arrangements at Kaiser Permanente, Group Health Cooperative of Puget Sound and Harvard Community Health Plan. The doctors would bear the risk of costs and make all of the medical decisions, including quality and utilization management, and they would determine the premium by mutual agreement with the carrier in recognition of the demands of a competitive market. I emphasize partnership because, in the long run, health plans cannot succeed without the loyalty, commitment and responsible participation of the doctors. Also we envisioned a gradual transformation so that people would have time to develop the underlying delivery systems."[5]

The idea of managed competition can only succeed if patients become consumers in the sense that they can read consumer reports on medical treatments, procedures, and providers, and decide intelligently which is best for them-similar to buying a car. As of yet, there is a lack of

18 Wiebe, Christine. "First HIPAA Rules Published." Medscape: Money & Medicine, 2000.

19 Epstein, Richard A. Mortal Peril: Our Inalienable Right to Health Care?" Cambridge, MA: Perseus Books, 1999.

20 Herzlinger, Regina. Market-Driven Health Care: Who Wins, Who Loses In The Transformation of America's Largest Service Industry. Reading, Massachusetts: Perseus Books, 1997: 282.

21 Chen, Jasper. "Bridge to The 'Real' World: Internships help students turn classroom theory into careers." TIME/The Princeton Review: The Best College for You, 2000: 104.

accessible medical information for patients to be educated enough so that they can be responsible for their own health in the same manner that they may be responsible for the upkeep of their car.

Market-Driven Health Care

Partly, this is due to the medical establishment's traditions and fear of change—which together may result in a lack of innovation regarding new health care ventures. Luckily, however, a number of health care professionals have a knack for business. Regina Herzlinger, Professor of Business Administration at Harvard, praises the advent of "focused health care systems," such as "Salick Health Care that deliver convenient, integrated, state-of-the-art cancer care; Dr. Denton Cooley's open-heart center that provides high-quality surgeries at a 40% lower price than the average; and Railin, an organization that effectively helps the victims of chronic diseases to manage their health."[6]

Market-driven approaches to health care can be extremely difficult, if not impossible, to implement. There may be little validity to the logic of market-based approaches to health care. That is, the market principles governing our economy may not be relevant to health care. None of us behaves as a wise cost-conscious consumer when we, or a loved one, are sick. Cost plays little or no role in our decision to find the best care possible and our reliance on our physicians as to what we need becomes great. Getting health care can hardly be equated with buying a car. Therefore, we do not search for the cheapest provider; rather we ignore price and look for the best[7].

Worse yet, the underprivileged are unlikely to possess the education and know-how to be cost-conscious consumers. Thus, if marketed towards educated consumers, market-driven approaches will not benefit the underprivileged who need health care the most. The underprivileged will have to continue relying on compassion, charity, and social responsibility.[8]

Health Care and Politics: The 2000 Presidential Race

In the past, the principal policy issue in the debate over health insurance focused on the poor and unemployed. Nowadays, this problem affects a much broader segment of society since the uninsured encompasses those who are employed and not necessarily poor. A growing number of Americans may be affected by the health care crisis.

What this has done is to move the debate from the political left, where concerns centered on

obtaining health insurance for the poor, to include a much broader ideological spectrum. No longer is the focus only the minority of the population who are generally disenfranchised, but now the problem affects our neighbors, families, and even ourselves (Merrill 36).

As a growing number of Americans lose their insurance or find it impossible to pay for it, the politics of the middle class may create a political will for real reform (Merrill 91).

Health care has recently been placed on center stage in the arena of American politics. Health care issues are among the most important ones on the campaign trail this year, in particular how to handle the spiraling cost of prescription drugs for the elderly. It remains to be seen how the American public will perceive the health care agendas of the two primary presidential candidates.

Republican Proposals

"Medicare is an enduring commitment of our country. It must be modernized for our times," says George W. Bush. He wishes to overhaul Medicare, the federal health insurance program for the elderly. In addition to pledging that the government would cover the full cost of drug coverage for the low-income elderly, Bush would create a system in which private insurers compete with the government to provide coverage for America's 39 million Medicare beneficiaries. 9

Bush's plan builds on the expansion of managed care in Medicare required by the balanced budget act of 1997. In a system modeled after the health plan now offered to federal employees, the government would solicit various health insurance plans including prescription drug benefits from private companies and then allow people to choose from a multitude of plans it approved or else remain in the current Medicare system. Citizens would be free to buy the more generous plan at more cost to themselves.[10]

To assist the elderly with drug costs, Bush would furnish a sliding scale of subsidies, starting with the full premium coverage for low-income elderly to obtain a standard benefit package including drug coverage. Elderly with incomes between $11,300 and $14,600, would be entitled to an unspecified partial subsidy for the cost of their prescription drug coverage. For all other elderly, the government would pay 25% of the premium costs for prescription drug coverage. The government is additionally responsible for all catastrophic medical expenses in excess of $6,000 a year. These subsidies would cost approximately $110 billion across a decade.[11]

Since such a reform of Medicare could take years to enact, Bush also proposed a temporary four-year $48 billion plan in state-grants to subsidize prescription drug coverage for low-income elderly. Bush's aides say that 23 states—Texas not included—offer versions of such plans. However, as critics point out, such plans provide for only roughly one million people. "Past experience has shown the states have only been able to cover small portions of those who need prescription drug coverage," said Ron Pollack, the executive director of Families USA, a nonpartisan, nonprofit group for health care consumers. Finally, Mr. Bush also called for $40 billion to be restored to Medicare for payments to doctors, hospitals and health care providers that were reduced under the budget act in 1997.[12]

Mr. Bush's entry into the ensuing debate on health care reform spurs a sharp competition for the votes of the elderly regarding health care. Intending to uphold their traditional advantage on this issue, the Democratic National Committee ran a commercial capitalizing on a federal district judge ruling—which is now undergoing appeal—that Texas violated a 1996 agreement to improve children's Medicaid programs.[13]

Democratic Proposals

Gore and the Democrats attack Bush's plan as one that does not provide meaningful enough subsidies to help many of the elderly purchase drug coverage. They argue that it would leave half those who need it without drug coverage. Moreover, Bush's proposal would force more people into health maintenance organizations at a time when many HMOs have been pulling out of the Medicare market.[14]

Gore's prescription plan is more expensive than Bush's in that it costs at least $253 billion over 10 years and creates the prescription drug benefit inside the traditional Medicare program. The plan would pay all drug costs for the roughly 13 million uncovered Medicare recipients with incomes making up to about $11,300 a year. For those with higher incomes, his plan would pay half a Medicare recipient's drug costs, up to $5,000 a year. Once that cap is met, beneficiaries would pay out-of-pocket up to an additional $1,500. After that, catastrophic coverage would kick in.[15] The appeal of Gore's plan is it's cohesive nature—it would give Medicare extensive purchasing power and the ability to dig up discounts from pharmaceutical companies.

In comparison with Bush's plan, which clearly favors drug companies and investors, Gore's plan can be critiqued as a step towards price controls. "By making government agents the largest purchaser of prescription drugs in America," Bush says, "by making Washington the nation's pharmacist, the Gore plan puts us well on the way to price control for drugs."[16] However, is establishing price controls an avenue for controlling health care costs?

Avenues for Reform: Learning from the Health Insurance Portability and Accountability Act of 1996 (HIPAA)

The biggest mistake in reforming health care would be to increase government's role in the regulation of health care delivery and financing. As it stands, the health care industry is over-regulated.

All too often when a policy determines to cut costs in health care by regulating the practices of HMOs or health insurance companies, these corporations will thwart these attempts at reform by finding creative ways to recoup the money they are no longer making. So in effect some health policies result not in diminishing the costs of health care but in shifting them elsewhere.

Take the Health Insurance Portability and Accountability Act of 1996 (HIPAA), the purpose of which was to protect workers from losing their insurance if they changed jobs. What HIPAA did not address is how much insurance companies can increase their premiums in order to adopt the requirements to insure high-risk individuals. This example shows how there could be any number of unintended consequences to any legislative attempt.

But HIPAA was not completely useless: it was also designed to lower other health care costs due to regulations standardizing administrative and financial transactions in the health care industry.[17] When the federal government finally published the first set of HIPAA rules this August, it set in motion a sweeping overhaul of the nation's health care information network. We must keep in mind, though, that government efforts take a very long time. HIPAA became a statute in 1996, it is now the year 2000, and not until the year 2002 finally rolls around will compliance be required.[18]

Health Care For Our Times

From HIPAA and numerous other pieces of legislation, policymakers have realized that even more important than coming up with sound and efficient health care policy is the willingness not to produce half-way policies that result in unintended consequences. Richard Epstein cautions us against "the hidden consequences of reform which can backfire against its stated objective."[19] Perhaps the best way to reform health care is to encourage competition among providers of health care so that there will be a true incentive to cut costs. Alain Enthoven's landmark theory of managed competition provides a useful illustration for this approach.

In order for health care to become a real marketplace, however, several things need to be accomplished. As is already becoming the trend, patients are no longer as patient as they once used to be. Nowadays, they are becoming more and more like consumers: they want to make their own decisions about health and want to be active participants in their health. And they want to do this in the least amount of time possible. This is assumed to be true of the educated patient at least. Thus, health care must first be convenient and oriented towards consumerism.

Second, doctors' roles are also undergoing a change. They are no longer merely the health care practitioners they used to be, having to deal with the medical needs of their patients and not being attentive to how much such treatment would cost. Now they must be cognizant of both the cost and management of their care—particularly if the HMOs with which they contract will only reimburse them for certain procedures performed. In this way, doctors must rely more and more on their managerial skills. A modernized health care delivery system will require efficient and innovative management on the part of clinicians.

More importantly, however, doctors must be aware that the doctor-patient relationship will not be the same as before. Traditionally, patients have looked up to doctors as the authority on all matters pertaining to health. Patients rarely questioned the decision-making abilities of their doctors, if at all. Today, as more and more patients are having access to more medical information and newer technologies, they tend to question their doctors. Thus, doctors are no longer seen as the omniscient individuals as they were previously. The doctor-patient relationship-which used to be one in which the doctor's order is an order-is evolving into one in which the patient has the final say. Doctors must be patient with patients and allow them to make choices regarding their own health.

In order for this to occur, a breakthrough must be achieved in the dissemination of medical information so that the general population can become intelligent consumer-patients. At present there may be any number of legal and political hurdles to cross before this can happen. In an ideal marketplace for health care, however, patients will have access to health care information at their fingertips—much like they can obtain information on any number of products on-line or in advertisements and consumers' reports—so they can have mastery over their own health.

More generally, our health care system anxiously awaits entrepreneurs and innovators to revolutionize the way in which health care can profit off of people while serving them more effectively. Entrepreneurial ventures in health care would be feasible under a consumer-controlled and market-driven health care system. Indeed, this new health care system is emerging: one that is "created by the people-by entrepre-

neurial people-for the people-for efficiency-, convenience-, mastery-, and medically technology-loving people."[20]

Furthermore, there should be generous tax breaks or other financial incentives for individuals investing in the health care needs of the poor. Government financing alone cannot be relied upon to address the manifold health needs of the disadvantaged. The power of charitable institutions in ameliorating the health care needs of the underprivileged are immense and should not be underestimated. Every effort should be made to encourage almsgiving for medically underserved populations.

Above all, our American ideology—individualism, minimal government interference, willingness to conquer new frontiers, and firm commitment to free-market enterprise—cannot be ignored in any attempt to move health care in a forward direction. Where can this revolution begin? It starts with us here at MIT, for we are the doctors, politicians, entrepreneurs, and scientists of tomorrow. As the health-care system is an endless power struggle among the needs of corporations, the medical profession, and patient-consumers[21], leadership from all directions is necessary for any viable reform effort to occur.

Not an alien life form or the latest video game.

It's what you become when you've made the right career choice.

Science & engineering opportunities for students & grads that are out of this world.

Don't let your career opportunities drift away.

Online shoppers want more than merchandise. They want advice. And when they're making an important decision, the immediacy of the response becomes the difference between a completed transaction and an abandoned sale.

LivePerson is continuously developing and advancing the software and infrastructure that allows millions of online consumers to interact directly with the sites they visit. At LivePerson, you'll help shape our future as a leader in the ASP/CRM arena.

Looking for a challenge? Looking for a fun and rewarding work experience? Visit us at www.liveperson.com/jobs or call 1.877.228.7984.

www.liveperson.com/jobs

LivePerson℠
Give your site a pulse.™