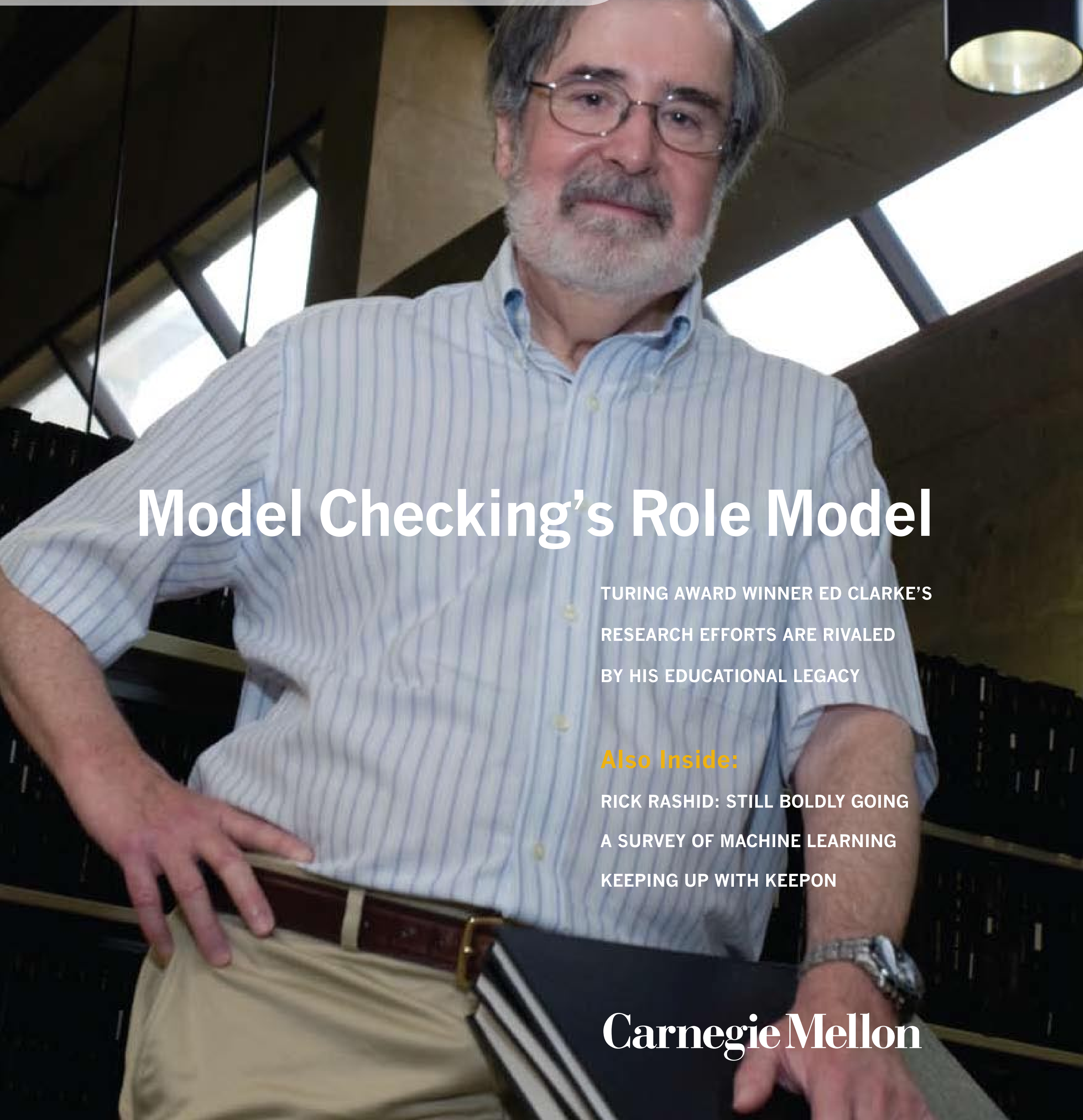
THE LINK

News from the School of Computer Science

Issue 4.1 / Summer 2009



Model Checking's Role Model

TURING AWARD WINNER ED CLARKE'S RESEARCH EFFORTS ARE RIVALED BY HIS EDUCATIONAL LEGACY

Also Inside:

RICK RASHID: STILL BOLDLY GOING

A SURVEY OF MACHINE LEARNING

KEEPING UP WITH KEEPON

Carnegie Mellon

THE LINK

News from the School of Computer Science

The Link provides a mosaic of the School of Computer Science: presenting issues, analyzing problems, offering occasional answers, giving exposure to faculty, students, researchers, staff and interdisciplinary partners. The Link strives to encourage better understanding of, and involvement in, the computer science community.

Editor-in-Chief
Randal E. Bryant

Managing Editor/Writer
Jason Togyer

Contributing Writers
Jennifer Bails, Tina Carr, Mark Dorgan, Meghan Holohan, Philip L. Lehman, Byron Spice, Anne Watzman

Photography
Ken Andreyo, John Barna, EEcue.com, Stan Franzos, Wade H. Massie, Microsoft Corp., Stanford News Service, U.S. Postal Service

Graphic Design
Melissa Stoebe
Communications Design Group

Office of the Dean
4305 Newell-Simon Hall
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh PA 15213

Randal E. Bryant, dean
Tina Carr (HNZ'02), director of alumni relations
Philip L. Lehman (CS'78,'84), associate dean for strategic initiatives
Byron Spice, director of public relations
Jason Togyer (HS'96), writer/editor

Carnegie Mellon University
412-268-7884
TheLink@cs.cmu.edu

Produced for the School of Computer Science by the Communications Design Group, August 2009, 09-670.

CALENDAR OF EVENTS

All events to be held at the Carnegie Mellon University campus in Pittsburgh, unless otherwise noted.

August 24

Fall semester begins

September 2

Student Activities Fair

College of Fine Arts Lawn

(Rain Location: Wiegand Gymnasium)

4:30 to 6:30 p.m.

September 12–13

Career Express Weekend

University Center

September 22

Dedication Ceremony:

Gates Center for Computer Science and Hillman Center for Future-Generation Technologies (see page 3)

October 15

New York Alumni Network Night

October 29-31

Homecoming Weekend

November 25

Andrew Carnegie born 1835, Dunfermline, Scotland

April 16-17

Spring Carnival

FEEDBACK LOOP

Good grief! There's our round-headed managing editor, waiting in vain for all of the cards and letters that never arrived after our last issue.

Whether you're a little red-haired girl or a World War I flying ace, we'd like to hear from you. To comment on a story, make a suggestion, ask a question or submit a news item, call 412-268-8721, email TheLink@cs.cmu.edu, or write to The Link Magazine, Office of the Dean, School of Computer Science, 5000 Forbes Ave., Pittsburgh, PA 15213. You can also leave comments at our website, link.cs.cmu.edu.

Please be his pencil pal. Otherwise, he'll have to see that crabby psychiatrist again, and he's running out of nickels.

—Jason Togyer (HS'96), managing editor

JOHN BARNA PHOTO

C O N T E N T S

Issue 4.1 / Summer 2009

On the Cover:

Inside the newly renovated Engineering and Science Library on the Carnegie Mellon campus, Edmund Clarke poses with some of the theses written by his graduate students. Clarke, a winner of the 2007 Turing Award, has worked with more than 75 graduate students, post-docs and visitors over the past three decades. "They deserve the credit for what I have been able to achieve," says Clarke, who calls them his extended family. See page 8. (Wade H. Massie photo)

From the Dean

Randal E. Bryant

Looking Back, Looking Ahead

I've just passed the five-year mark as dean of the School of Computer Science. Time has gone by very quickly, yet many things have transpired during that period.

In 2004, we were in the earliest stages of thinking about building new space for computer science. Now, a large part of the school is about to move into the Gates Center for Computer Science and the Hillman Center for Future-Generation Technologies. These buildings are both stunning, and they will serve the needs of our students, educators and researchers much better than our existing spaces. We look forward to the official opening ceremony on Sept. 22. (See page 3.)

During the past three years, we also expanded into buildings on neighboring Craig Street, part of the business district of the Oakland section of Pittsburgh. These buildings house more than 200 students pursuing professional master's degrees in software engineering, electronic commerce and human-computer interaction. We published our first edition of The Link in conjunction with CS50, our celebration of 50 years of computer science at Carnegie Mellon (dating back to when the first computer arrived on campus, and when Alan Perlis, Herb Simon and Allen Newell began working together).

Our Center for Automated Learning and Discovery was transformed into the first machine-learning department in the world (see Tom Mitchell's article on page 18). Computational biology became a major focus for SCS, the university and our collaborators at the University of Pittsburgh and elsewhere. We started graduating computer science students at our campus in Doha, Qatar, and in many other ways we have become a more global institution.

We have begun a major focus on data-intensive scalable computing, focusing on the challenges and opportunities for performing computation over terabyte-scale data sets. This form of computing will provide revolutionary capabilities in science, business and society. In the past few years, we received access to very large-scale computing facilities through the generosity of Yahoo!, Google, IBM and Intel, and we procured two major cluster facilities. Still, there is much more work to be done in making these machines more reliable and efficient, finding better ways of programming large-scale computations and solving important, real-world problems.

Jeannette Wing introduced the term "computational thinking" to express the idea that computer science embodies a fundamentally new set of ideas and principles for formulating and solving difficult problems and for designing and implementing complex systems. We have plans to redesign our introductory computer science courses to put more emphasis on computational thinking rather than simply the mechanics of programming, and we want to foster this approach at other universities, in middle and high schools and as a basis for interdisciplinary collaborations in science, humanities and the arts.

I've just been reappointed for another five-year term, and with you, I look forward to the many developments that will take place during this time. Stay tuned as we keep you updated about the many activities by the students, faculty and alumni of the Carnegie Mellon University School of Computer Science.

Randy Bryant

A daylong celebration on Sept. 22 will herald the opening of the striking new Gates Center for Computer Science and Hillman Center for Future-Generation Technologies.

KEN ANDREYO PHOTOS

Gates Center, Hillman Center to be dedicated Sept. 22

The events being planned will include visits from the buildings' namesakes, Microsoft Chairman Bill Gates and Pittsburgh philanthropists Henry and Elsie Hillman, who will help the university celebrate five decades of leadership in computer science education and research.

Although School of Computer Science faculty, staff and students will begin occupying the buildings in August, the September celebration will mark the formal opening and offer a chance for the Carnegie Mellon community to reflect on the history and future of computer science at the university, a spokesman said.

With a lead gift of $20 million, the Bill and Melinda Gates Foundation made possible the Gates Center, which will be home to the school's undergraduate programs. The center will feature classrooms, laboratories, offices, study spaces, conference rooms and collaborative work areas for students, faculty and staff.

The Henry L. Hillman Foundation's gift of $10 million was instrumental in the creation of the Hillman Center for Future-Generation Technologies, which adjoins the Gates Center. The Hillman Center is designed to house some of the university's most promising research in computer science.

Hillman, a successful venture capitalist and investor, and his wife Elsie are prominent regional civic leaders whose philanthropies have focused on ways that science and technology can improve human life, health, work and prosperity. Strong supporters of Pittsburgh-area universities, the Hillmans have called their gift to Carnegie Mellon recognition of its role in the economic growth of western Pennsylvania, and of the powerful role that computer science plays in shaping the future of the region.

A significant number of donors have complemented the generosity of the Gates and Hillman foundations by naming spaces and other features of the building, and additional opportunities remain available.

The event will mark Bill Gates's third visit to a Carnegie Mellon campus in 18 months. In February 2008, Gates visited Pittsburgh to deliver a lecture called "Bill Gates Unplugged: On Software, Innovation, Entrepreneurship and Giving Back."

On April 18, Gates visited Carnegie Mellon's campus in Doha, Qatar, to give the keynote address at the Third International Conference on Information and Communications Technologies and Development, the premier conference for innovating technology accessible and relevant to developing economies. (See page 23.)

With multiple entrances and pedestrian bridges, the new centers will literally and figuratively connect disparate areas of campus. Notable connections include the Randy Pausch Memorial Footbridge, which will link the fifth floor of the Gates Center to the Purnell Center for the Arts. The bridge, university officials note, connects computer science with the arts—much as Pausch did in life—while physically tying together the east and west sides of campus.

An enclosed fourth-floor bridge will connect the Gates Center to Newell-Simon Hall, where a new bridge will span the Perlis Atrium. The walkways will provide the first direct pedestrian connection from the Cut to Wean Hall and the rest of the campus mall.

Inside the Gates Center, the Pausch and Newell-Simon bridges are connected via a helix that its designers say echoes the connections of computer science to many related areas of research and education at Carnegie Mellon.

The Gates and Hillman centers were designed by Atlanta-based Mack Scogin Merrill Elam Architects to achieve a silver Leadership in Energy and Environmental Design, or "LEED," rating for sustainable buildings from the U.S. Green Building Council.

— Byron Spice and Philip Lehman

New buildings offer fresh start for fall term

When Jim Skees, director of building facilities for the School of Computer Science, describes the upcoming move into the Gates and Hillman centers, he jokingly describes it as "dollies full of boxes as far as the eyes can see."

Skees is one of the Carnegie Mellon staff members planning the big move, scheduled to take place in five phases beginning in mid-August. The pace is dictated in part by connectivity requirements; SCS has made a promise to faculty, students and staff that no one will lose access to a working computer and network port for more than 24 hours. A few laboratories with more specialized requirements will be moved separately, Skees says.

Some furniture for public spaces may arrive after August, he notes, and some landscaping and façade work will continue after the move as well. But the university is driving to get the buildings open and largely complete before the fall semester begins Aug. 24, Skees says.

After all, classes for the fall term are already scheduled in the Gates Center, he says: "The classrooms have got to be ready to go."

— Jason Togyer (HS'96)

Crews preparing the landscaping at the new Gates and Hillman centers installed these Kentucky Coffeetrees in early July.

Giving Back

Carnegie Mellon's environment makes people better, Takeo Kanade says, and people make the university better through giving.

With a career that extends almost 30 years at Carnegie Mellon, Takeo Kanade has played an integral role in the growth and expansion of the Robotics Institute, founding the master's and doctoral programs and serving as director from 1992 to 2001. His areas of research are broad and diverse, including facial recognition, facial expression analysis and motion analysis. His innovations in robotic vision systems are numerous and have been integrated into a variety of robotic applications.

But Steelers fans—or anyone who watches football—might be most grateful for his research into the extrapolation of three-dimensional models from two-dimensional photos. Takeo pioneered the concept of "Virtualized Reality," a technique that combines several images from an event into a coherent 3-D model that can be viewed from any angle. It's the technique used during TV football broadcasts to provide images of the field from a player's point of view, first demonstrated during CBS's telecast of Super Bowl XXXV in 2001.

In 2006, Takeo was a founding director of the Quality of Life Technology Center. This NSF-funded engineering research center is developing intelligent systems to help older adults and people with disabilities.

Last year, he made a gift as part of Carnegie Mellon's "Inspire Innovation" billion-dollar fundraising campaign to establish the Quality of Life Technology Center Student Research Fund. Supporting students in research related to the center's work provides invaluable experience and is critical to the university's efforts to advance a vital technology for society, Takeo says.

Takeo also made a gift to the Gates Center for Computer Science and Hillman Center for Future-Generation Technologies by "buying" a seat in the Rashid Auditorium. As someone who has benefited greatly from philanthropic support in his work at the university, Takeo believes the environment at Carnegie Mellon makes people better, and that people make Carnegie Mellon better through their giving.

If you have questions about how you can help the School of Computer Science, whether through scholarships and fellowships, faculty support or gifts towards the Gates and Hillman centers, please contact me at mdorgan@cmu.edu or 412-268-8576, or send an email to scsgiving@cs.cmu.edu. You can also learn more about the campaign or make your gift online today at www.cmu.edu/campaign.

—Mark Dorgan is principal giving advisor and development liaison for the School of Computer Science.

Scanning GigaPan's Horizons

Researchers are picturing scientific uses for camera system, software

The GigaPan system makes it easy to generate giant panoramas using everyday digital cameras, but even it has its limits. Laura Tomokiyo, project scientist with the Robotics Institute, learned that the hard way—on the North Slope of Alaska.

It turns out that the batteries powering the robotic camera mount are no match for the bitter cold of Alaska in March. "You can't be outside setting up a camera when it's 50 degrees below zero," Tomokiyo says. The cold sapped power from the batteries before GigaPan could complete its photographic sweep of the horizon.

Nevertheless, Tomokiyo's visit to Barrow, Alaska—America's northernmost community—convinced her that GigaPan has a role to play in documenting the culture and language of the native Inupiat people. This summer, members of the Inupiaq Heritage and Language Center have taken advantage of the warmer weather to gather GigaPan images of the isolated villages that dot the 94,700-square-mile North Slope.

"We hope this project will provide a good opportunity for youth and elders to work together," she says. "The youth may provide leadership in the use of the technology, but they also will be looking to the elders to share their knowledge of history and tradition in the documentation of the panoramas."

Tomokiyo, who earned her Ph.D. in language technologies at Carnegie Mellon in 2001, made the 20-hour trip to Barrow in hopes of finding ways that GigaPan panoramas could be used to study and preserve the endangered Inupiaq language.

Bilingual interpreters are commonly used to elicit lists of important words from native speakers, she says, but the method tends to be biased toward the dominant language. "Even if someone is bilingual, they will think about things differently when talking to English speakers," Tomokiyo says.

The use of GigaPan images, which allows users to zoom in on details, might eliminate this bias by allowing people who speak only Inupiaq to point out objects or activities of interest and name them. An example of such an image would be a GigaPan that Tomokiyo produced of workers preparing seal skins for use on a umiak—a large seafaring canoe used for whaling.

Laura Tomokiyo (left) used GigaPan to document village life on the North Slope of Alaska, including scenes (above) of natives preparing seal skins for use on a whaling canoe or "umiak." See more at www.gigapan.org/viewProfile.php?userid=314.

BOTH PHOTOS COURTESY LAURA TOMOKIYO

Other scientific uses of GigaPan were explored in May during the Fine Outreach for Science, a workshop in Pittsburgh sponsored by the Fine Foundation that brought together 29 paleontologists, geologists, climatologists, primatologists and other researchers from around the world. One had used GigaPan to image rock paintings in Kazakhstan. Another modified the GigaPan mount so that it controlled a scanning electron microscope, making it possible to image an entire fly magnified 500 times.

The GigaPan system—a robotic camera mount that triggers the camera to take numerous photographs of a scene and a software package that electronically stitches the images into a panorama—was created by the Global Connection Project to help people communicate across cultures.

Last year, for instance, the project launched the GigaPan School Exchange, in which students from schools in Pittsburgh, South Africa and Trinidad and Tobago shared GigaPans of their communities with each other. This 21st Century "pen pal" program is expanding this fall, with schools in Brazil and Indonesia joining the group. In the Pittsburgh area, the Falk School in Oakland will be joined by students at Propel charter schools and South Fayette, Burgettstown and Connellsville middle and high schools.

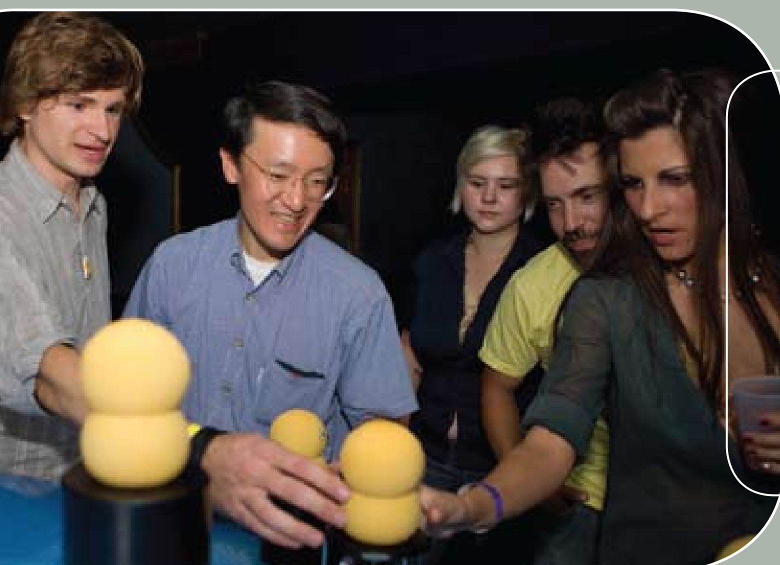
Broadband connections on the North Slope are poor, so the ability of students there to take full advantage of such an exchange would be difficult, Tomokiyo says. Even so, she's thinking about ways to make it happen, such as timing student use to coincide with periods of low Internet traffic. In the meantime, the GigaPan archive will document the culture and activities of the North Slope and reside on computers there, rather than be accessible via the Internet.

The archive is funded through the IHLC by the U.S. Department of Education's Education through Cultural and Historical Organizations, or ECHO, initiative. "We will return to Barrow at the time of the fall whaling season to wrap up, document cultural activities related to this event and explore possibilities for the next phase," Tomokiyo says.

— Byron Spice is director of public relations for the School of Computer Science. To view some of the GigaPans mentioned in this story, visit The Link's Web site at link.cs.cmu.edu.

Keeping Up with Keepon

Bouncy 'bot adds new dimensions to human-robot interaction



Marek Michalowski and Hideki Kozima demonstrate Keepon to visitors at Wired's NextFest in September 2007.

DAVE BULLOCK PHOTO VIA EECUE.COM

It's hard keeping up with a Keepon. Ever since Robotics Institute doctoral student Marek Michalowski taught it to dance, this highly expressive little robot—designed for child therapy—has taken on a new dimension as an entertainer while adding insights to the study of non-verbal communication between robots and people.

Keepon looks like a little yellow snowman with a rubbery body built to withstand hours of use and abuse by young children. Created by Japanese roboticist and developmental psychologist Hideki Kozima, currently at Miyagi University, it was designed to interact with children and provide therapy for developmental disorders such as autism. Its round eyes contain cameras and its charcoal-colored nose hides a microphone, enabling Keepon to record the reactions of the people interacting with it. Four motors in the robot's base allow it to turn left and right, nod its head up and down and rock from side to side.

Michalowski joined Kozima's lab three years ago during a summer internship in Japan, where he began studying rhythm in human-robot social interaction. After working in Yale's Social Robotics Lab and graduating with bachelor's and master's degrees in computer science and a bachelor's degree in psychology, he'd planned to focus on verbal communication at the Robotics Institute. But after working with some of the university's most famous talking bots, he became convinced much of what humans communicate has nothing to do with speech, but rather in non-verbal cues.

"There is a synchronization of rhythmic behavior between interacting people," says Michalowski, who decided he wanted to capture that in robots.

Michalowski felt an immediate rapport with Kozima and Keepon, which can speak volumes without making any sound. After imagining that the robot had a desire to dance, he developed software enabling it to do just that. For fun, he made a video of Keepon dancing to the song "I Turn My Camera On" by the Texas rock band Spoon. It garnered two million hits after being posted on YouTube.

After that posting, Kozima says, "the explosion" came. Soon he and Michalowski were traveling to festivals and contests around the world. In 2007, Wired magazine invited Keepon and Spoon to shoot a follow-up video for the band's hit song, "Don't You Evah." (It featured Kozima and Keepon exploring Tokyo.) Then Keepon won the 10,000-Euro "Robots at Play Prize" in Odense, Denmark, and first place at the International Conference on Robotics and Automation's Human-Robot Interaction Challenge. (It repeated the latter feat in 2009.)

Keepon and the researchers have twice been invited by Wired to exhibit at NextFest, and they're slated for their third appearance later this year at the magazine's conference, which is probably the premier showcase of the latest global innovations in computers and information technology.

And Keepon's reputation continues to grow. In April, Keepon was featured in an NBC "Today Show" series on autism and in a Popular Science magazine article on robots and autism. In May, Pennsylvania's Technology Collaborative, which focuses on economic development of robotics and other digital tech type companies, invited Michalowski and Kozima to demonstrate Keepon at a briefing of the Congessional Robotics Caucus in Washington, D.C.

Michalowski and Kozima's collaboration has changed both of their lives. They've formed a company, called BeatBots, to commercialize the robot for research, therapy and entertainment. Right now there are only about a dozen Keepons in the world, each loaded with motors and cameras and costing about as much as a small car; Michalowski is developing a less expensive version of the robot. Meanwhile, as he finishes research for his thesis, he's conducting studies at Carnegie Mellon's Children's School and with children in Japan, observing the way they play and dance with Keepon. He is refining the robot's perception algorithms, designing interactive scenarios and developing methods to evaluate rhythmic interactions.

"We see Keepon as a research platform, something to be used by hobbyists, researchers or therapists," Michalowski says. "And we're enabling this technology to mesh with people's natural rhythms, so that robots can be more useful and accepted in society."

Robotics Institute Research Professor Reid Simmons, who's working on a game-playing robot to be housed in the Gates Center for Computer Science, is one of Michalowski's advisors. "There are things Marek's working on that definitely will be important in this next generation of robots that we're developing," he says.

—Anne Watzman recently retired as director of public relations for the School of Computer Science. For more information about Keepon, visit www.beatbots.net.

As the Morris Worm Turned

Twenty years later, cybersecurity remains a challenging set of problems

"The worm had access to every file on every infected computer. It could have violated the data on every infected computer ... we were daunted by the unrealized potential for vast damage."

Bill Scherlis

At half past midnight on November 3, 1988, subscribers to an email list for TCP-IP developers received an ominous message: "There may be a virus loose on the Internet," it warned.

A worm written by a Cornell grad student, Robert Tappan Morris, had been released from Massachusetts Institute of Technology. Designed to test the size of the Internet, it got access to computers by exploiting a variety of vulnerabilities, including easy-to-guess passwords and weaknesses in the sendmail and finger server processes that ran on most of the hosts.

There were only about 88,000 computers on the Internet, then used mainly by computer scientists and the Defense Department. But as the "Morris worm" wiggled around, it copied itself over and over on some machines, executing multiple simultaneous processes and causing crashes.

Bill Scherlis was an assistant professor of computer science at Carnegie Mellon temporarily working at the Defense Advanced Research Projects Agency. He and others at DARPA suddenly began receiving reports about problems from all over the country, but because panicked users were pulling sites offline to avoid infection and the remaining email queues were saturated with the worm, Scherlis wound up glued to his phone for the next 72 hours.

"I'm glad I had a headset," says Scherlis, now a professor in SCS and director of the Institute of Software Research and its Ph.D. program in software engineering.

Although the Morris worm wreaked a high toll in terms of lost productivity, it actually helped the Internet's development, Scherlis says: "It got us thinking about a systematic approach to protection." The approach he and his colleagues suggested—a permanent organization to respond to cybersecurity problems—became the Computer Emergency Response Team. Today, the rapid response unit (known as the CERT Coordination Center) is one part of a broader effort by CERT at assessing risks, analyzing network security, devising procedures and "best practices" and training IT professionals.

CERT celebrated its 20th anniversary March 10 and 11 in Pittsburgh with a two-day technical symposium designed to bring together people and groups that have taken the lead in cybersecurity. The invitation-only gathering was an intentionally low-key affair, says Archie Andrews, senior member of CERT and the event's coordinator. "Rather than widely advertising the event, we wanted to bring the movers and shakers together in one place at one time to encourage frank and honest exchange," he says.

A generation ago, the Internet was making the transition from what Scherlis calls a "small-town" culture to that of "a big city where you need to keep your eyes open all the time." When the Morris worm hit, DARPA researchers had to quickly learn how it worked, what damage it was causing and who could rapidly engineer fixes. Scherlis worked with Stephen Squires, then software technology program manager at DARPA, to study the aftermath and make recommendations to the Pentagon.

Most of the havoc created by the Morris worm was caused by denial-of-service problems, but if it had been malicious, it could have corrupted or compromised data on every machine it touched. "The worm had access to every file on every infected computer," Scherlis says. "It could have violated the data on every infected computer … we were daunted by the unrealized potential for vast damage." One of DARPA's recommendations was creation of a neutral team "to collaborate with the community, mobilize resources and coordinate response from multiple organizations," Scherlis says. After their memo reached then-DARPA Director Craig Fields, he asked Larry Druffel, director of Carnegie Mellon's Software Engineering Institute from 1986 to 1996, to oversee CERT's creation.

Those gathered for the anniversary symposium looked at the past but spent more time examining the frontiers of Internet security. Keynote speakers included Vinton Cerf, vice president and chief Internet evangelist for Google; U.S. Navy Rear Adm. Elizabeth Hight, vice director of the Defense Information Systems Agency; John Gilligan, former chief information officer for the U.S. Air Force and now president of the Gilligan Group, a Virginia-based consultancy specializing in defense information networks; Scott Charney, Microsoft's corporate vice president for trustworthy computing; and Jonathan Zittrain, professor of law and co-founder of the Berkman Center for Internet and Society at Harvard University.

A panel led by David Farber, professor of computer science and public policy at Carnegie Mellon, discussed emerging security challenges, while SCS Dean Randy Bryant led a forum on security and privacy issues in cloud-based computing systems. A government panel focused on areas of concern within the U.S. government and means of sharing those concerns with the private sector.

Privacy issues were another topic of discussion. Knowing the identity of other users can prevent scams and predatory behavior, Andrews says, but if you track everyone's identity, participants noted, are you violating their right to privacy?

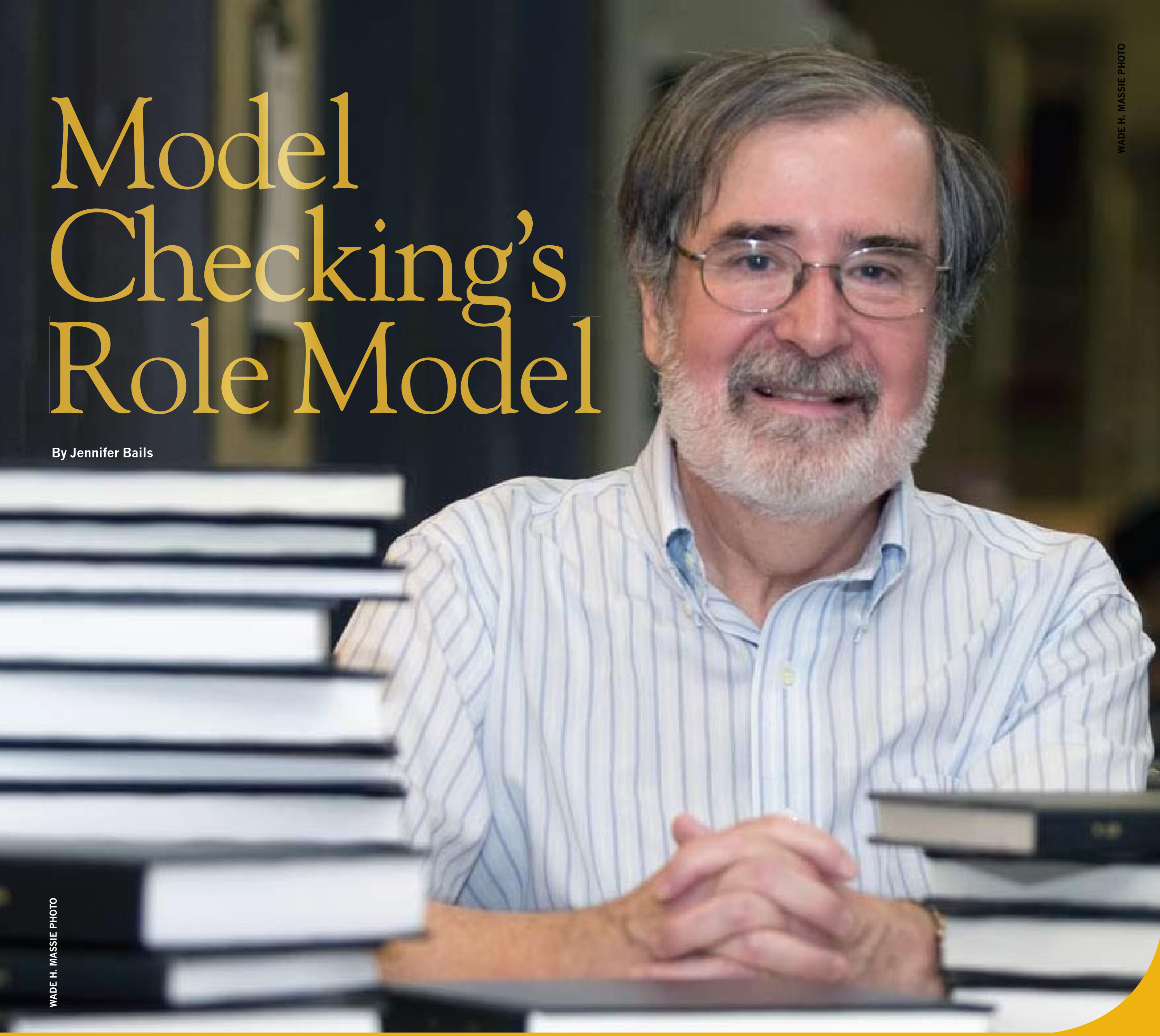
Twenty years after the Morris worm, cybersecurity researchers are still wrestling with the legacy of the Internet's "small-town" culture, Scherlis says. "Perhaps we continue to wrestle because we know it is such a shame to have to abandon trust just because there are few bad actors," he says. "That's our challenge—how to build technology and institutions so we can have it both ways. We need to remember that despite its scale and pervasiveness the Internet is still quite malleable."

CERT's presence in Pittsburgh since 1988 has helped make Carnegie Mellon a "powerhouse" in cybersecurity, Scherlis says. "When we took on this public service role, there were amazing things that happened," in terms of attracting top faculty, students and projects, he says. "We have real opportunities ahead of us."

—Meghan Holohan is a Pittsburgh-based freelance writer whose work has also appeared in Mental Floss, Geek Monthly and ComputerWorld.

Model Checking's Role Model

By Jennifer Bails



WADE H. MASSIE PHOTO

WADE H. MASSIE PHOTO

Ed Clarke's graduate students stay in close contact with the Turing Award winner, whose educational legacy rivals his research output

It was the biggest speech that Edmund Clarke had ever delivered. Last Oct. 20, before the International Conference on Embedded Software in Atlanta, Clarke rose to give his official lecture as a winner of the 2007 A.M. Turing Award.

Recognized as the highest honor in computer science, the $250,000 prize is bestowed annually by the Association for Computing Machinery and is widely considered the Nobel Prize of computing. Clarke shared the honor with a former student, E. Allen Emerson, now at the University of Texas at Austin, and Joseph Sifakis of France's National Centre for Scientific Research.

The trio was honored for its efforts to develop an automated method—called model checking—to find design errors in computer hardware and software.

Model checking has transformed how the computer industry tests chips, systems and networks, and is increasingly being used to de-bug applications as well, and no one has been at the forefront of its development over the past three decades like Clarke, the FORE Systems university professor of computer science and professor of electrical and computer engineering at Carnegie Mellon.

Yet rather than kicking off his speech by recapping his ground-breaking career, Clarke instead displayed a PowerPoint slide listing over 75 graduate students, post-docs and visitors who worked with him over the past three decades. "They deserve the credit for what I have been able to achieve," he said, with genuine humility. "I attribute my success to all of these people. I often just steered them in the right direction and then tried to keep up."

The sentiment wasn't out of character for a professor as devoted to teaching the next generation of computer scientists as he is to advancing model checking. "Ed has played an important role in model checking, but for me and his other students, he has also 'checked out' to be an important role model," says Bud Mishra (CS'82,'85), professor of computer science, mathematics and cell biology at New York University and one of Clarke's first graduate students at Carnegie Mellon.

Computer scientists used to spend countless hours manually poring over lines of code to hunt for bugs, often in vain. Instead, they now rely on model checking to catch errors in hardware for everything from cell phones to digital memory chips. Fast and automatic, model checking can verify highly sophisticated circuits with as many as 10^{120} configurations. If a design contains an error, model checking will produce a counterexample to help pinpoint the mistake.

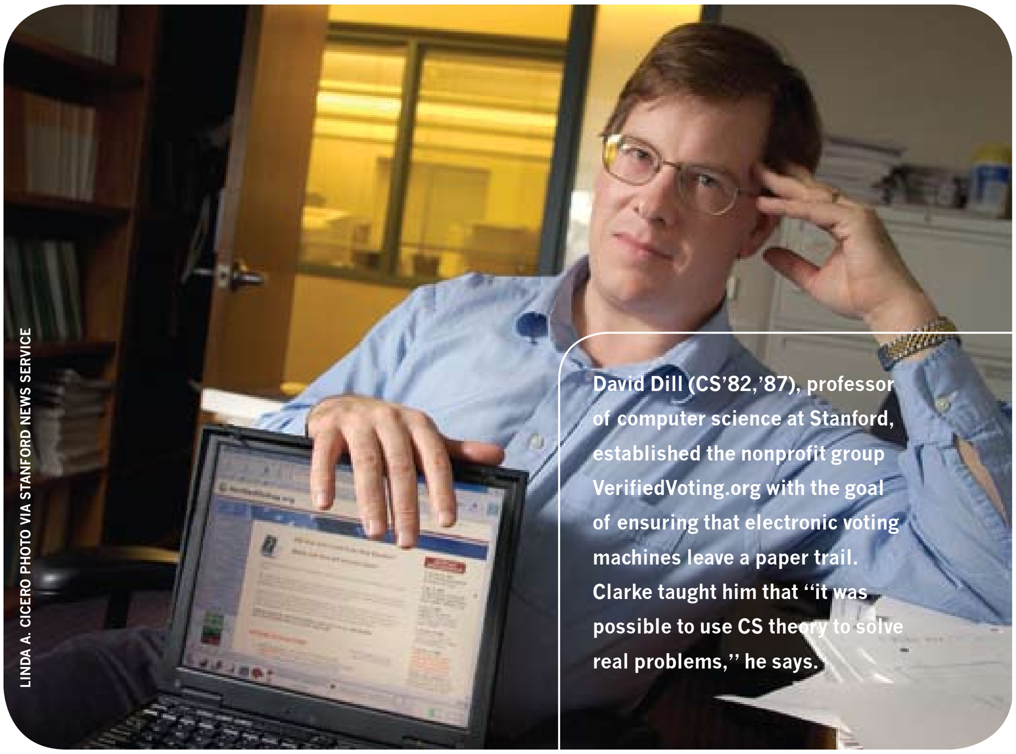
The method also is being used to test the reliability of Internet security and privacy protocols, and more recently, to verify that safety-critical software in cars, medical devices and nuclear control systems behave as intended.

Clarke's Turing Award lecture naturally documented his 27-year effort to transform model checking from a theoretical algorithm into an indispensible technology. But his legacy of education is as rich as his legacy of research. His dozens of former graduate students and post-docs have made key contributions in both academia and industry in fields as diverse as systems biology, Internet security and even public policy.

No matter their current interests, they attribute much of their success to the intellectual freedom he afforded them to pursue their own ideas, as they developed the knowledge and skills to become top-notch scientists.

"He gives his students a lot of freedom to pick their problems," says Somesh Jha (CS'96), associate professor in the department of computer science at University of Wisconsin at Madison. "That was excellent training for when I got out of graduate school and really had to do things on my own."

Clarke says he looks upon his students as equals. "At the same time, I feel like they are my children, and I am sincerely interested in their welfare," he says.



LINDA A. CICERO PHOTO VIA STANFORD NEWS SERVICE

David Dill (CS'82,'87), professor of computer science at Stanford, established the nonprofit group VerifiedVoting.org with the goal of ensuring that electronic voting machines leave a paper trail. Clarke taught him that "it was possible to use CS theory to solve real problems," he says.

The spirit of collaboration he fosters in this "family" remains very much alive among Clarke's extensive network of graduates, who call each other to share ideas, seek career advice or just to chat. A dozen or so "Ed Clarke alumni" living in the San Francisco Bay area have even organized themselves more formally and gather often. "Every time I am out there, they take me to dinner and talk about their work," Clarke says. "They still seem to really enjoy teaming up with each other." And they all enjoyed the privilege of learning under a legend in computer science.

One member of Ed Clarke's extended family is David L. Dill (CS'82,'87). When he first heard about a nationwide push to replace traditional election ballots with paperless electronic voting machines, he grew uneasy. That's in part because of the lessons he learned as one of Clarke's graduate students at Carnegie Mellon in the early 1980s. There, he helped write a key paper showing that model checking was useful in very-large-scale integration, or VLSI, design, the process of building integrated circuits by combining thousands of transistors on a microprocessor chip.

"I developed a profound respect for the complexity of systems and the incredible difficulty of making them correct—and the even greater difficulty of making them secure," says Dill, professor of computer science at Stanford University since 1987. "Designers are getting much better at putting bugs in systems than we are at finding them, so that's why my suspicions about electronic voting were immediately raised."

Dill's research has concentrated on the theory and application of model checking and other "formal verification" techniques to system designs. But after the 2000 presidential election debacle, he put some of his science on hold and made a foray into public policy, believing the integrity of American democracy was in jeopardy.

"I got involved in late 2002, early 2003 because states and counties were gearing up for massive purchases of touch-screen machines, and I started wondering why we could trust them," Dill says.

"I developed a profound respect for the complexity of systems and the incredible difficulty of making them correct—and the even greater difficulty of making them secure."

David L. Dill (CS'82,'87)

He established the lobbying group VerifiedVoting.org and a nonprofit foundation, with the goal of ensuring every vote has a paper record and that voters can verify the accuracy of that record before casting their ballots. So far, the organization and its partners have persuaded 31 states to establish regulations requiring voter-verified paper records—not just bug-prone computer systems. "Watching the admittedly slow national shift toward better election systems has been really satisfying," Dill says.

Dill views his involvement in political affairs as a natural product of the most lasting insight he gleaned from Clarke. "He taught me that it was possible to use CS theory to solve real problems, and, surprisingly, that I was capable of doing it," Dill says. >>>

Another of Clarke's former collaborators who remains in close contact is Orna Grumberg, who first came to the School of Computer Science from the Technion—or Israel Institute of Technology—as a post-doc under Clarke in 1985. Now a full professor in the Technion's computer science department, she returns to Carnegie Mellon each summer as a visiting researcher with Clarke's group.

"Ed always welcomes his visitors warmly and teaches them whatever he knows," says Grumberg, a renowned expert in computer-aided verification of hardware and software. "The atmosphere is one of wonderful collaboration."

Grumberg and Clarke have co-authored dozens of journal articles together, and in 1999, they published the first comprehensive textbook on the theory and practice of model checking. And like Clarke, most of Grumberg's career has been dedicated to solving what is called the state explosion problem in model checking.

Model checking works by expressing the specifications of an abstract model of the hardware or software being analyzed in a special type of notation called temporal logic, which describes how a system will behave over time. Then the model checker reviews every possible state of the model to determine if it is consistent with these specifications.

But as the complexity of a system increases, the number of states to analyze multiplies exponentially, overwhelming the model checker. Finding ways around this conundrum is the key to unlocking all of the inherent power of the method.

"It's a problem we are still fighting, and there is a lot of work left to do," says Grumberg, who in December was named vice dean of the Technion's graduate school. "But there are new inventions and new horizons all the time so I hope that we haven't seen the limits of model checking yet."

Internet security is one of the new frontiers of model checking. Take the security protocols that ensure that credit card information and other personal data are kept confidential when users shop online.

These complex authentication protocols are tough to get right, partly because it is difficult to predict the behavior of potential attackers, says Jha of UW-Madison. Flaws have been identified years after security protocols are released—much too late, for even small bugs easily can be exploited to cause serious harm, he says.

THE TECHNION – ISRAEL INSTITUTE OF TECHNOLOGY PHOTO

In May, Clarke visited the Computer Science Department at the Technion – Israel Institute of Technology in Haifa to deliver a series of lectures. Shown at left is one of Clarke's former post-docs, Orna Grumberg, now a full professor of computer science at the Technion.

"Ed always welcomes his visitors warmly and teaches them whatever he knows. The atmosphere is one of wonderful collaboration."

Orna Grumberg

Jha joined Clarke's group in 1991, earning his Ph.D. and then staying at Carnegie Mellon until 2000 to complete a postdoctoral fellowship. There, he helped develop several improvements to model checking, including one advance that exploited symmetry in hardware design to increase the speed of the technique and another that provided a systematic way to refine the abstract models used.

Now he's drawing on his model checking expertise to make computer networks safer. The method is valuable in the context of security because it exhaustively searches for all possible vulnerabilities. After all, an attacker only needs one door to enter a system and wreak havoc, says Jha, who won the prestigious CAREER award from the National Science Foundation in 2005 to support his work in computer security and privacy.

Recently, Jha co-founded the company NovaShield, which develops technology to detect and eliminate the next generation of malware on personal computers including viruses, worms, trojans, spywares and other threats. Jha says he learned a great deal about "how to be a good scientist" from Clarke and still depends on him for counsel.

Another student who calls Ed Clarke for advice is Bud Mishra, who came to Carnegie Mellon in 1982, the year Clarke left Harvard University to come to Pittsburgh. Along with Dill, he helped write the first paper showing model checking had practical use in hardware verification and wasn't just an academic exercise.

After completing his degree, Mishra searched for applications of his training outside computer science. Today he works at the interface of computing, mathematics and biology as principal investigator for the NYU Bioinformatics Group at the university's Courant Institute of Mathematical Sciences. Founded by Mishra, the group focuses on applying algorithmic, statistical and mathematical models to genomes, cells and populations of cells.

Mishra has developed cutting-edge technologies and statistical analysis tools to attack biological problems that range from deciphering the structure of a genome to understanding chromosomal aberrations and their relation to cancer, malaria, autism and other diseases.

One of Mishra's innovations—called optical mapping—uses fluorescence microscopy to image individual DNA molecules and then produce a map of an organism's genome at low cost compared to conventional sequencing methods. Another can tease apart the genetic contributions people receive from each of their parents, which could help scientists read the genome more accurately and understand evolutionary patterns.

Interestingly, Mishra says, many of his colleagues are beginning to apply the tools of model checking to computational models of biological systems—a daunting task. "With hardware or software, the engineer gives you the model," he says. "With biology, there is no documentation. If there is an intelligent designer, we didn't get the design."

Mishra said a key lesson he took away from Clarke is that the most beautiful problems to work on are not necessarily the most technically difficult.

Ed Clarke's Scrapbook

The Link asked Ed Clarke to page through some old photos—including some of his former graduate students and post-doctoral researchers—and tell us a little bit about them.

Allan Fisher, former associate dean for undergraduate education, and Clarke help Ken McMillan (back to camera) celebrate the latter's successful thesis defense in 1992. McMillan, now at Cadence Berkeley Labs in California, was "one of my most memorable students ever," Clarke says. <<<

Clarke, wife Martha and son Jeffrey pose with their hosts during a 1993 visit to the Institute of Software of the Academia Sinica (now the Chinese Academy of Sciences) in Zhongguancun, Beijing, China. <<<

"This was one of my largest groups of students, though I've probably had groups even larger than that," he says of this 2002 snapshot.

Shown are Muralidhar Talupar (CS'02,'06), Flavio Lerda (CS'05), Michael Theobald (post-doc), Pankaj Chauhan (CS'01,'07), Sagar Chaki (CS'05), Daniel Kröning (post-doc), Clarke, Nishant Sinha (E'03,'07), Alex Groce (CS'05), Ofer Strichman (post-doc), Karen Yorav (post-doc), Joel Ouaknine (post-doc) and Natalya Sharygina, currently adjunct assistant professor of computer science at Carnegie Mellon and senior researcher at Software Engineering Institute.

Marius Minea (CS'95,'99) toasts his successful thesis defense with Clarke in 1999. Minea is now an associate professor of computer and software engineering at Politehnica University of Timisoara, Romania.

Three of Carnegie Mellon's 10 Turing Award winners gather after hearing news that Clarke had won the prize in 2008: Raj Reddy, Mozah Bint Nasser University Professor of Computer Science and Robotics and former dean of SCS; Clarke; and Manuel Blum, Bruce Nelson Professor of Computer Science.

ALL PHOTOS COURTESY EDMUND CLARKE

"I should have realized that when I abandoned model checking and moved on to other things," he says. "But it feels good to be in the wrong." Rather than debugging software, Mishra says, "now I'm interested in debugging humans."

Whether Clarke's students continue to work directly in the field of model checking or move further afield into other disciplines, one thing is certain—his legacy as an extraordinary teacher and role model will carry on in perpetuity. Last year, a student of a student of a student (who Clarke refers to with amusement as his "great-grand-student") came to host a seminar for one of his classes. And future generations of top computer scientists and leading thinkers will also be able to trace their intellectual roots on a family tree whose lineage begins in his laboratory.

For Clarke, that is perhaps the best aspect of his success, as he imparts his most important advice to those following in his footsteps. "You've got to pick something that you think is an important project and then devote your life to it," he says.

DAN HART ILLUSTRATION

Still Boldly Going

For Rick Rashid, the voyage from Iowa to building the corporate world's largest computer science research organization went through Carnegie Mellon

By Jason Togyer

It was a familiar sensation for graduate students working on the Mach kernel in the 1980s—grinding out code under the fluorescent lights of Wean Hall, they would suddenly feel the presence of Rick Rashid behind them. Rashid, then a professor in the computer science department, hated meetings. Instead, he engaged in what his protégés jokingly remember as "managing by wandering the halls."

"Rick would tailor the frequency of his visits based on how he thought you were doing," Rich Draves (CS'94) says. "If you were seeing a lot of him, it meant either that he was excited, or worried. And that's still the case."

Draves is still working for Rashid, now senior vice president of research for Microsoft Corp. When Rashid left Carnegie Mellon in 1991 to set up the software giant's initial research effort, Draves was among his first 20 hires. Today, Microsoft Research includes 850 people—including Turing Award winners Tony Hoare and Butler Lampson, and, until his disappearance and presumed death, Jim Gray—and laboratories in Bangalore, Beijing, both Cambridges (Massachusetts and England), and Mountain View and San Francisco, Calif. Under Rashid, Microsoft is investing heavily—one trade magazine estimates about $6 billion annually—into research into computer vision, e-Commerce applications and Web services, including search engine technology. Work done at Microsoft Research has led directly to creation of:

- Microsoft Virtual Earth and TerraServer;
- Data mining and machine-learning tools on the company's SQL Servers;
- Compression and encoding algorithms of Windows Media Player; and
- Rendering technology and artificial intelligence inside Xbox.

"Rick Rashid, to his credit, has been able to structure Microsoft Research in a way that's incredibly relevant to software research," says another of Rashid's former students, Jeff Eppinger (CS'88), one of the founders of Transarc Corp. and currently professor of the practice in Carnegie Mellon's Institute for Software Research.

"There's a saying in the industry, 'Don't bet against Steve Jobs,' and I'd say, 'Don't bet against Rick Rashid,' " Eppinger says.

Despite Microsoft Research's extraordinary growth, Draves, now area manager for systems and networking, says the organization retains a Carnegie Mellon flavor. Rashid still tells researchers about the "reasonable person principle"—the SCS tradition favoring initiative over bureaucracy, as long as the results don't infringe on other people's rights. "The working atmosphere here feels very academic," Draves says.

That's no accident, says Rashid, who modeled Microsoft Research after Carnegie Mellon's computer science culture. His approach to building the organization has focused on finding researchers who can make an impact in their fields; collaborate with others from different backgrounds; and thrive when given freedom to pursue ideas.

"And I look at their curiosity—do they pull in knowledge from other areas?" Rashid says. "Are they excited not just about the specific things they're doing, but is computer science as a discipline exciting to them? … If you don't have that attitude that 'I'm going to keep moving and stake out new territory intellectually,' you're not going to have a long career."

Maybe it's no surprise that Rashid—a proponent of boldly exploring new frontiers—is a huge fan of the "Star Trek" franchise. He identifies with the thoughtful management style of Captain Jean-Luc Picard of "Star Trek: The Next Generation," who delegated authority rather than barking commands, and even has a Starfleet uniform like the one Patrick Stewart wore on the TV series. During Rashid's tenure at Carnegie Mellon, he started taking his researchers to the premiere of each new "Star Trek" movie—at his own expense. (For the debut of the latest film, the tradition continued, though it required renting a theater for more than 500 people.)

And in the best "Star Trek" tradition, Rashid seeks out researchers who are invested in the organization, not just their personal goals. "It's a Carnegie Mellon way of doing things," he says. Eighteen years after leaving Pittsburgh, Rashid remains close to the university. His son, Daniel, is a doctoral candidate in the Language Technologies Institute, and the new Gates Center for Computer Science includes the 250-seat Rashid Auditorium, created with help from a gift to Carnegie Mellon from the family.

"I owe a lot of my personal success to the time I was at CMU, and I thought it was important to give back to the university, and the students and faculty," Rashid says. "When I go back to Carnegie Mellon, it's the closest thing I have to going home."

Home originally was Fort Madison, Iowa, population 14,000, where his parents, Farris and Ramona Rashid, ran a wholesale grocery business. Located on the Mississippi River two hours south of the "Quad Cities," Fort Madison's biggest employers were the Iowa State Penitentiary and the Sheaffer pen company—not exactly hotbeds for scientific research.

Nevertheless, Rashid's sister, Norma, a former television newscaster who now teaches journalism at the University of Cincinnati, says her older brother was always fascinated by science. As president of the science club at Fort Madison High School, Rick Rashid once locked himself in a laboratory so no one could interfere with an experiment, she says. >>>

Christian Borgs, deputy managing director, Microsoft Research New England, Jennifer Chayes, managing director, Microsoft Research New England, and Rashid celebrate the opening of the corporation's newest lab in Cambridge, Mass., in September.
MICROSOFT CORP. PHOTO

"Whatever he did, he was intense about it," Norma Rashid says. "When he was in high school, he put together an artificial heart for the science fair—most kids grow bean plants, but that was the kind of thing he did." She remembers avidly watching television coverage of America's space race with her brother, and his interest in science fiction—from writing his own stories to falling in love with "Star Trek."

"Rick was always interested in learning more about problems, and asking questions," Norma Rashid says, "and he's always loved science fiction—so if you're going from fantasy to real science, the logical pathway is through computers."

Rashid didn't see his first computer (a Hewlett-Packard 2116) until his sophomore year at Stanford University, where he was a mathematics and comparative literature major. Fellow student Dan Ling—later a Microsoft Research employee—encouraged Rashid to take a basic course in Stanford's Algol-W programming language.

His first opportunity to program a computer was "a rush," Rashid says, and instead of pursuing a higher degree in math, he decided to go to the University of Rochester to earn his master's and doctorate in computer science. When his mother found out, she cried. "For someone from Iowa, I might as well have said 'I'm going to go get a graduate degree in underwater basket-weaving,' " Rashid says, adding his parents, both now deceased, were "never anything but supportive of me."

> "Some people think research is all about technology. It isn't. Research is all about people who have the ability to think and innovate. I think it's because we built that kind of organization with those kinds of values that we've been successful, both in the early days and in the long run."
>
> *Rick Rashid*

At Rochester, Rashid worked on research into the emerging field of personal computing over local-area networks, which attracted the attention of Carnegie Mellon's Raj Reddy. Reddy was spearheading what became the Scientific Personal Integrated Computing Environment project and in 1979 asked Rashid to join the faculty. One of SPICE's goals was a so-called "three-M" workstation—a personal computer with a megabyte of memory, a megapixel display and a processor capable of handling a million instructions per second.

SPICE spawned the Accent operating kernel, whose novel use of memory mapping saved processor time by copying only the data needed for specific operations rather than complete files; and allowed processors to share information via secured ports. But it wasn't compatible with UNIX and was tightly woven around the Three Rivers PERQ workstation, limiting its usefulness. Rashid and others began work on the Mach kernel, which

Rashid through the years—at left as a newly installed Carnegie Mellon faculty member circa 1979, and below during development of the Mach kernel.

BOTH PHOTOS COURTESY RICK RASHID

incorporated Accent's ideas but could run UNIX applications.

In a building full of computer scientists used to late hours, Rashid, lead developer on Mach, stood out in part because he arrived "bright and early" every day, remembers one of his graduate students, David Black (CS'88,'90). He also remembers Rashid's persistent optimism and encouragement to students. "He was an adviser in the best sense of the word," Black says.

"Tireless," adds Dave Hornig (CS'84), one of Rashid's advisees and now the vice president of product development at Pittsburgh-based Touchtown Inc. Rashid "clearly thought about computers from the moment he woke up in the morning until the time he went to bed," Hornig says.

Unlike other scientists consumed with finding elegant solutions, Rashid was most interested in effective ones. "I tended to have some instincts toward formalisms in theory, where he had the instinct to try something, see if it works, and if it doesn't, try something else," Hornig says.

Black says Rashid stressed the value of learning by doing. "One of the goals of the Mach project was to show that these concepts would work in practice, not just in principle," Black says. As a result, faculty and staff were soon running Mach-based operating systems on their own computers—allowing the Mach developers to "eat their own dog food." They didn't always like the taste, especially when they found bugs, but that didn't discourage Rashid.

> "There were no software companies that had basic research groups in those days. Moreover, Microsoft had a relatively small number of products at that point. I had to be convinced that Microsoft would still exist in a few years, that it was willing to make the investment, and that we had enough 'intellectual runway' to get off the ground."
>
> *Rick Rashid*

"I never saw him looking depressed—maybe sometimes a little frustrated," Hornig says. "He had a can-do, positive attitude that rubbed off on a lot of people. In a place that was often rather hectic and sometimes negative, he could make you feel like it was all going somewhere."

Though compatible with UNIX, Mach wasn't hampered by some of the older operating system's restrictions. While UNIX allowed only a single thread of execution in a process, Mach supported multiple threads, sharing memory but running independently. That made it easy to expand Mach to run on multiple processors or across clusters of machines. Mach's modular design also kept the kernel size small while allowing it to be ported to any number of platforms.

"Rick's work was really seminal," says Eppinger, because it successfully demonstrated ideas that became widely discussed among OS developers. And industry took notice—Apple co-founder Steve Jobs visited Pittsburgh, left impressed, and >>>

recruited Rashid's students, including researcher Avie Tevanian (CS'85,'88) to his new company, NeXT, where the Mach kernel became the heart of the NeXTSTEP OS.

Microsoft's chief technology officer, Nathan Myhrvold, also came calling—to recruit Rashid to set up a research unit, much to Rashid's surprise. "There were no software companies that had basic research groups in those days," Rashid says. "Moreover, Microsoft had a relatively small number of products at that point. I had to be convinced that Microsoft would still exist in a few years, that it was willing to make the investment, and that we had enough 'intellectual runway' to get off the ground."

Before leaving Carnegie Mellon, Rashid set two key conditions. First, Microsoft's research unit had to advance the state of the art in any area it explored. Second, it had to turn that research into useful products. "It's not really that much different from Carnegie Mellon," Rashid says. "It's about developing ideas, and then getting those ideas to corporations, and the general public." Assured by Myhrvold that he'd have the freedom to pursue those goals, Rashid says, he made the leap.

> Microsoft Research is "not really that much different from Carnegie Mellon. It's about developing ideas, and then getting those ideas to corporations, and the general public."
>
> *Rick Rashid*

Although industrial research organizations aren't unusual—IBM established its first lab in 1945—Eppinger says Microsoft stands out as the only software company funding basic scientific inquiry. Microsoft scientists working under Rashid have an unusual amount of latitude to pursue research not directly tied to products.

Take computational biology: Microsoft researchers are experimenting with programming languages to alter living cells for use in medical therapies. Another effort is adapting algorithms used for digital compression and email filters to model the ways that immune systems fight off intruders. Such research could eventually lead to a vaccine that adapts to the many mutations of the virus that causes AIDS. Other teams are pursuing research into fundamental theories of information retrieval, natural language processing, organizational behavior and human-computer interaction.

"The great contribution to mankind is when they let you go off and focus on something fundamental, but the great benefit for the company when you create this level of expertise is that the research group becomes a resource for the product centers of the company," Eppinger says.

Microsoft Research also has an unusually strong record—at least among for-profit corporations—of collaborating with outside scientists. Carnegie Mellon DNA is apparent in the collaboration process, which is governed by the RPP, Draves says. "In a lot of companies, when you have a paper that you want to publish, you have to go through the legal process and get various people to sign off on it," he says. "We don't have that process here. The onus is on the researcher to do the right thing. Go ahead and give talks, but be responsible about what you talk about." Microsoft Research investigators also publish frequently—more than 5,100 peer-reviewed papers so far, including well over 10 percent of those accepted annually by SIGGRAPH since 2002.

After nearly four decades in computer science, Rashid says he remains excited about the frontiers—especially cloud computing and what he calls "disembodied computation," where an end user's device delivers a service rather than processing a lot of data locally. Data-intensive super computing is another major focus for Microsoft Research, he says. "It's refined our notions of what our specifications for programs should be," Rashid says. "It's given us a better understanding of how we have to describe and talk about structuring data, so that we do a better job of sharing it and mapping it."

Whatever Microsoft Research pursues, Rashid says its main focus has to remain on people. "Some people think research is all about technology," he says. "It isn't. Research is all about people who have the ability to think and innovate. I think it's because we built that kind of organization with those kinds of values that we've been successful, both in the early days and in the long run."

STAN FRANZOS PHOTO

"Betting on the Future" describes Rashid's outlook on life—and also was the title of the lecture he delivered to a packed room at Newell-Simon Hall on April 19, 2006, during the CS50 celebration at Carnegie Mellon.

JOHN BARNA PHOTO

Mahadev Satyanarayanan—better known as Satya—is the Carnegie Group Professor of Computer Science. Founding director of Intel Research Pittsburgh, he also was a principal architect and implementer of the Andrew File System (AFS).

Mahadev Satyanarayanan

Much of Satya's research focuses on mobile and pervasive computing—the ability to move from place to place and platform to platform seamlessly. One outcome of this work was Coda, an open-source system to support distributed file access on intermittent networks. Key ideas from Coda were incorporated into Windows 2000 and Outlook 2003.

Satya's most recent research in this area has been the Internet Suspend/Resume system, or ISR, which allows users to leave one computer and pick up their work at the exact same place on another computer. He spoke with Managing Editor Jason Togyer:

How did Coda evolve from AFS?

As you become more dependent on a distributed file system such as AFS, your inconvenience level becomes higher when one of the servers goes down. We wanted to preserve the positive features of AFS, but make it more resistant to failure. The answer was the disconnected operation mechanism in Coda—the forerunner of today's hot-sync mechanisms—which caches server data on local machines so users can continue working even when their network connections are lost. At the time, laptops were just starting to appear, and I realized that Coda was a perfect fit for the world of wireless computing—even today, wireless networks are spotty.

After we published a paper describing the evolution of Coda from AFS, a Japanese magazine, Nikkei Electronics, asked if they could translate the article. Of course, I can't read Japanese, so how could I know if the translation was accurate?

Well, they added cartoons by a Japanese artist, Gaich Muramatsu. One half of the first cartoon shows unhappy AFS users dealing with a broken network connection to the servers. The other half shows happy Coda users still working. I could see just from the cartoons that everyone understood it—they captured the idea perfectly.

How does ISR build on Coda?

If you run Windows XP at home, and move to a machine using Windows Vista, you don't get the same user experience. Why not? At all times, you should have one world you're dealing with, not different worlds on different machines.

Most people think of cloud computing as data being stored in the cloud. With ISR, the entire machine is stored in the cloud, and if the right business model can be constructed, people would treat computers like furniture. If you check into a hotel, you don't have to bring your own furniture. Why should you lug around your computer? Just have your world magically appear as you access parts of it at any loaner computer.

Are there issues with platform portability?

Yes, but just about all of today's laptops and desktops use Intel processors, so that can be handled through virtual machine technology. I'd like to see ISR deployed at campus scale. All it takes are the resources to make that happen—I see no technical challenges, but we need more funding.

Does Diamond build on Coda and AFS?

Diamond is completely unrelated! When I started the Intel lab, I asked myself: "Data caching is a powerful technique in hardware and software and on the Web, but is there any circumstance in which caching data is not useful?"

It turns out that caching and indexing work well when you're searching text, but not when you're searching rich content, such as images. Creating an index ahead of time for all conceivable queries you might pose data is almost impossible.

But caching the results of searches is useful. So rather than building an index, Diamond does discard-based search. For example, you highlight a portion of an image and tell Diamond to find similar images. You discard the results that don't fit your query, refine your query, then search again. Since the refined query typically has significant overlap with the previous queries, Diamond can reuse cached results from those queries.

Can it handle data other than images?

Diamond knows nothing about the underlying data content. We've recently extended it to live data sets from webcams, and you could also use it to sort audio. You couldn't just throw sound at it and say "search"—you'd have to write some sound-specific plug-in for that—but once you did, it would do the search very efficiently.

What are some practical applications?

The medical and pharmaceutical communities have shown a lot of interest. Let's say you have a mammogram, and a tissue mass causes you concern. Should you call the patient in for a biopsy? With Diamond, you could highlight a portion of the image, ask it to show you other patients who had the same kinds of masses, and find out what their biopsy results were. Our goal here isn't to replace doctors, but to provide a Google-like tool that helps doctors with decision-making.

One application we've built on Diamond is called FatFind—it automatically counts fat cells, which is useful in automating drug discovery. Another is called PathFind, which allows pathologists to rapidly search through stained tissue samples.

Tom Mitchell

The Discipline of Machine Learning

A scientific field is best defined by the central question it studies. In machine learning, that question is: "How can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes?"

This covers a broad range of learning tasks, such as how to design autonomous mobile robots that learn to navigate from their own experience; how to data mine medical records to learn which future patients will respond best to which treatments; and how to build search engines that automatically customize themselves to their user's interests.

More precisely, we say that a machine "learns" with respect to a particular task T, performance metric P and type of experience E, if the system reliably improves its performance P at task T following experience E. Depending on how we define T, P and E, the learning task might also be called "data mining," "autonomous discovery," "database updating" or "programming by example."

Machine learning is a natural outgrowth of computer science and statistics. Where computer science has focused primarily on how to program computers, machine learning asks how we can get computers to program themselves, from experience plus some initial structure. Where statistics has focused primarily on what conclusions can be inferred from data, machine learning incorporates additional questions about what computational architectures and algorithms can be used to most effectively capture, store, index, retrieve and merge these data; how multiple learning subtasks can be orchestrated in a larger system; and questions of computational tractability.

Closely related to machine learning are psychology, neuroscience and related fields. The questions of "how do animals learn?" and "how can computers learn?" are highly intertwined. So far, the insights machine learning has gained from studies of human learning are much weaker than those it has gained from statistics and computer science, due primarily to the weak state of our understanding of human learning. Nevertheless, the synergy between studies of machine and human learning is growing, with machine learning algorithms now being suggested as explanations for neural signals observed in learning animals. It's reasonable to expect the synergy between studies of human learning and machine learning to grow substantially, as they are close neighbors in the landscape of core scientific questions.

Other fields—from biology to economics to control theory—also have a core interest in the question of how systems can automatically adapt or optimize to their environment, and machine learning will have an increasing exchange of ideas with these fields in the years ahead.

Application Successes

It's worth noting that as late as 1985, there were almost no commercial applications of machine learning. One measure of progress in the field is the growth of significant real-world applications.

Speech Recognition: All currently available commercial systems for speech recognition use machine learning in one fashion or another to train the system to recognize speech. The reason is simple—speech recognition accuracy is greater if one trains the system than if one tries to program it by hand. In fact, many commercial speech recognition systems involve two distinct learning phases—one before the software is shipped (training the general system in a speaker-independent fashion), and a second phase after the user purchases the software (to achieve greater accuracy by training in a speaker-dependent fashion).

Computer Vision: Many current vision systems, from face recognition systems to systems that automatically classify microscope images of cells, are developed using machine learning, again because the resulting systems are more accurate than hand-crafted programs. One massive-scale application

Machine learning techniques have improved the optical-character recognition software that sorts mail for the U.S. Postal Service. Over 85 percent of handwritten mail in the United States is now sorted automatically.

is the system used by the U.S. Postal Service to automatically sort letters containing handwritten addresses. Over 85 percent of handwritten mail in the United States is sorted automatically, using handwriting analysis software trained to very high accuracy using machine learning over a very large data set.

Bio-Surveillance: A variety of government efforts to detect and track disease outbreaks now use machine learning. For example, the Real-time Outbreak and Disease Surveillance, or RODS, project launched by the University of Pittsburgh in 1999 involves collection of admissions reports by emergency rooms and the detection of anomalous patterns of symptoms and their geographical distribution. Current work involves adding in a rich set of additional data, such as retail purchases of over-the-counter medicines, to increase the information flow into the system, further increasing the need for automated learning methods given this even more complex data set.

Robot Control: Machine learning methods have been successfully used in a number of robot systems. For example, several researchers have demonstrated the use of machine learning to acquire control strategies for stable helicopter flight and helicopter aerobatics.

Accelerating Empirical Sciences: Many data-intensive sciences now make use of machine learning methods to aid in the scientific discovery process. Machine learning is being used to learn models of gene expression in the cell from high-throughput data; to discover unusual astronomical objects from massive data collected by the Sloan Digital Sky Survey; and to characterize the complex patterns of brain activation that indicate different cognitive states of people in fMRI scanners. Machine learning methods are reshaping the practice of many data-intensive empirical sciences, and many of these sciences now hold workshops on machine learning as part of their field's conferences.

While there will remain software applications where machine learning may never be useful, the niche where it will be used is growing as applications grow in complexity; as the demand grows for self-customizing software; as computers gain access to more data; and as we develop increasingly effective machine learning algorithms.

Placing Machine Learning Within Computer Science

Given this sample of applications, what can we infer about the future role of machine learning in the field of computer applications? The above applications suggest a niche where machine learning has a special role to play. In particular, machine learning methods are already the best methods available for developing two types of applications:

1.) Applications too complex for people to manually design the algorithms: Sensor-based perception tasks, such as speech recognition and computer vision, fall into this category. Machine learning is the software development method of choice because it's relatively easy to collect labeled training data.

2.) Applications that require software to customize itself to fit its users: This machine learning niche is growing rapidly. Examples include speech recognition programs that learn to recognize the user's voice patterns, online merchants that learn your purchasing preferences and email readers that evolve to block your particular definition of spam.

While there will remain software applications where machine learning may never be useful, the niche where it will be used is growing as applications grow in complexity; as the demand grows for self-customizing software; as computers gain access to more data; and as we develop increasingly effective machine learning algorithms.

Beyond its obvious role as a method for software development, machine learning is also likely to help reshape our view of computer science.

By shifting the question from "how to program computers" to "how to allow them to program *themselves*," machine learning emphasizes the design of self-monitoring systems that self-diagnose and self-repair, and on approaches that model their users, and then take advantage of the steady stream of data flowing through the program rather than simply processing it.

Some Current Research Questions

Substantial progress has already been made in the development of machine learning algorithms and their underlying theory. The field is moving forward in many directions, exploring a variety of types of learning tasks, and developing a variety of underlying theory. Here is a sample of current research questions:

- Can unlabeled data be helpful for supervised learning?
- Are there situations where unlabeled data can be guaranteed to improve the expected learning accuracy?
- How can we transfer what is learned for one task to improve learning in other related tasks?
- What is the relationship between different learning algorithms, and which should be used when?
- For learners that actively collect their own training data, what is the best strategy?
- What is the most efficient training strategy for actively collecting new data as its learning proceeds?
- To what degree can we have both data privacy and the benefits of data mining? >>>

It's also interesting to consider longer-term research questions.

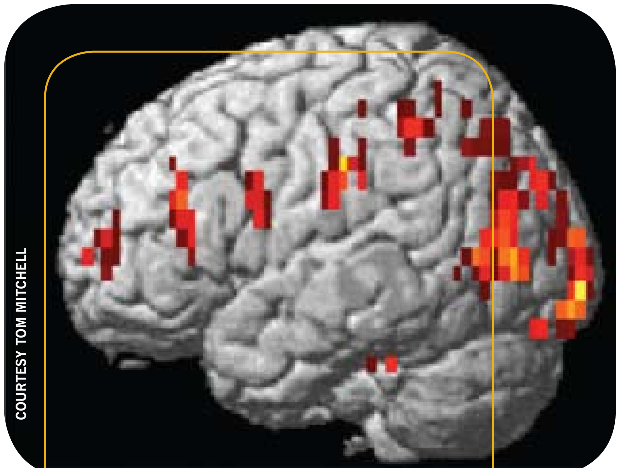
For instance, can we build never-ending learners? The vast majority of machine learning work to date involves running programs on particular data sets, then putting the learner aside and using the result. In contrast, learning in humans and other animals is an ongoing process in which the agent learns many different capabilities, often in a sequenced curriculum, and uses these different learned facts and capabilities in a highly synergistic fashion.

Why not build machine learners that learn in this same cumulative way, becoming increasingly competent rather than halting at some plateau?

Can machine learning theories and algorithms help explain human learning? Recently, theories and algorithms from machine learning have been found relevant to understanding aspects of human and animal learning. And machine learning algorithms for discovering sparse representations of naturally occurring images predict surprisingly well the types of visual features found in the early visual cortex of animals. However, theories of animal learning involve considerations that have not yet been considered in machine learning, such as the role of motivation, fear, urgency, forgetting and learning over multiple time scales. There is a rich opportunity for cross fertilization here—an opportunity to develop a general theory of learning processes covering animals as well as machines, and potential implications for improved strategies for teaching students.

Can a new generation of computer programming languages directly support writing programs that learn? In many current machine-learning applications, standard machine-learning algorithms are integrated with hand-coded software into a final application program. Why not design a new computer programming language that supports writing programs in which some subroutines are hand-coded while others are specified as "to be learned"? Such a programming language could allow the programmer to declare the inputs and outputs of each "to be learned" subroutine, then select a learning algorithm from the primitives provided by the programming language. Interesting new research issues arise here, such as designing programming language constructs for declaring what training experience should be given to each "to be learned" subroutine, when, and with what safeguards against arbitrary changes to program behavior.

Can a new generation of computer programming languages directly support writing programs that learn? In many current machine-learning applications, standard machine-learning algorithms are integrated with hand-coded software into a final application program. Why not design a new computer programming language that supports writing programs in which some subroutines are hand-coded while others are specified as "to be learned"?



COURTESY TOM MITCHELL

Machine learning is allowing researchers to characterize complex patterns of brain activation that indicate different states of people being examined by fMRI scanners.

Will computer perception merge with machine learning? Given the increasing use of machine learning for state-of-the-art computer vision, computer speech recognition and other forms of computer perception, can we develop a general theory of perception grounded in learning processes? One intriguing opportunity here is the incorporation of multiple senses (sight, sound, touch, etc.) to provide a setting in which self-supervised learning could be applied to predict one sensory experience from the others. Already researchers in developmental psychology and education have observed that learning can be more effective when people are provided multiple input modalities, and work on co-training methods from machine learning suggests the same.

Considering Ethical Questions

While it's impossible to predict the future, further research in machine learning will almost certainly produce more powerful computer capabilities. This, in turn, will lead on occasion to ethical questions about where and when to apply the resulting technology.

Consider that today's technology could enable discovering unanticipated side effects of new drugs, if it were applied to data describing all doctor visits and medical records in the country along with all purchases of drugs. Recent cases in which new drugs were recalled following a number of unanticipated patient deaths might well have been ameliorated by already available machine-learning methods. However, applying this machine-learning technology also would have impacted our personal privacy, as our medical records and drug purchases would have had to be captured and analyzed. Is this something we wish as a society to do?

Related questions occur about collecting data for security and law enforcement or for marketing purposes. Like any powerful technology, machine learning will raise its share of questions about whether it should be used for particular purposes. Although the answer to each of these questions will have a technical component, in some cases the question will also have a social policy component requiring all of us to become engaged in deciding its answer.

—Tom Mitchell is the Fredkin Professor of Artificial Intelligence and Machine Learning and chair of the Machine Learning Department—the first of its kind when it was established in 2006. In May 2009, he was named a University Professor, the highest distinction for a member of the Carnegie Mellon Faculty. A graduate of Massachusetts Institute of Technology, Mitchell earned his Ph.D. in electrical engineering with a minor in computer science at Stanford University. A former president of the Association for the Advancement of Artificial Intelligence, he is a fellow of both the AAAI and the American Association for the Advancement of Science and winner of the 2007 AAAI Distinguished Service Award.

Tina M. Carr (HNZ'02)
SCS Director of Alumni Relations
tcarr@cs.cmu.edu

Thoughts on Giving Back

In this edition, SCS alumnus and Associate Dean Philip Lehman offers a perspective on giving to Carnegie Mellon. — Tina

In previous columns, we've described the "Gift of Time" given by alumni and friends of SCS who participate and volunteer for events, talks, advisory organizations and other initiatives. Quite a few alumni and friends "roll up their sleeves" to collaborate with the university on a broad range of activities and events. We've also described the importance of communication with Carnegie Mellon, and advocacy—helping spread the word about the amazing things we do.

In addition, many of you (alumni and friends, organizations as well as individuals) support the university financially. We're very grateful for your help, and we have, from time to time, profiled some generous individuals in The Link (see page 4 of this issue, for example).

What motivates us to "give back" in this way? There are as many answers as there are benefactors. Some donors are alumni who want to acknowledge the benefits they received here; others are friends who have come to know, admire and love Carnegie Mellon; still others are corporations and foundations who have seen the profound effect the university has on the world.

Some people support education and research, by funding fellowships and scholarships. Others support physical facilities (such as the new Gates and Hillman centers) or creative endeavors, student groups and various activities. Some target specialized funds such as the Dean's Innovation Fund, which invests in the future via new research; others offer unrestricted gifts. Some of us enjoy having donations acknowledged publicly, but others prefer to remain anonymous.

The only common theme is that we all want to help advance the great work that Carnegie Mellon continues to do, and the inherent satisfaction—regardless of the size of the gift—we receive in being collaborators in the university's mission. In short—to coin a phrase—we all want to "inspire innovation," the theme of Carnegie Mellon's ongoing $1 billion development campaign!

We sincerely thank those of you who have given gifts of your time as well as financial gifts. And we welcome inquiries about ways in which your generosity makes a real difference to our faculty, students and staff, who move us forward every day—and whose hearts, like yours, are truly in the work!

—Philip L. Lehman (CS'78,'84)
Associate Dean for Strategic Initiatives
School of Computer Science
Philip.Lehman@cs.cmu.edu

SUBMITTED PHOTO

The 2009 SCS Alumni Award for Undergraduate Excellence was presented May 17 during commencement exercises in Pittsburgh to two of our newest graduates, Andrew Maas (CS'09) and Hatem Alismail (CS'09, Qatar). Boris Sofman (E'05, CS'05,'07), a member of the SCS alumni award committee, made the presentation. Shown from left are Andrew, Boris and Hatem.

Andrew's thesis was entitled "Learning about Related Tasks with Hierarchical Models." His advisors were Andrew Bagnell (CS'01,'04), assistant research professor of robotics, and Charles Kemp, assistant professor of psychology. Hatem's thesis was entitled "Exploring Visual Odometry for Mobile Robots." His advisor was Brett Browning, systems scientist in the Robotics Institute.

Established in 2003, the SCS Alumni Award recognizes technical excellence in research and development by an undergraduate participating in the senior thesis program. The awards committee evaluates both the written thesis as well as an oral and poster presentation, judging each entry on factors such as originality, technical excellence, potential societal impact, accessibility and general excitement level.

David M. Steier

B.S., Computer Science, Purdue University, 1982
Ph.D., Computer Science, Carnegie Mellon University, 1989

Hollywood doesn't make flashy TV crime shows about people who develop accounting software. But as David Steier (CS'89) describes the work he and his colleagues do at PricewaterhouseCoopers (PwC), it's clearly got many of the same elements—it's fast-paced, involves high technology and unravels mysteries.

With offices in 150 countries and 155,000 employees, PwC audits thousands of publicly traded companies, private corporations and non-profits. That makes reliable data-mining tools potentially one of the most important pieces of an auditor's workbench, says Steier, director of PwC's Center for Advanced Research in San Jose, Calif. "One big area we work on is fraud risk detection. There are millions and millions of data points—there's no way that a manual review can find some of these things."

Tools developed by Steier's team look for anomalies—for example, is a certain department paying prices inconsistent with industry norms? And they spot coincidences that are—as a TV detective might say—"a bit too convenient." (One clue: Do vendors share addresses or other relationships with employees?)

Ideas for some applications come directly from the field. "We deliberately look for high-risk, high-reward problems," Steier says. One internal social-networking application developed by the center can match any one of thousands of partners and staff with hard-to-solve problems by looking at resumes, past experience, reports they've authored and other data.

While much of their work is confidential, PwC researchers do publish and collaborate with peers, including those at Carnegie Mellon. PwC scientists recently worked with Christos Faloutsos, professor of computer science at SCS, and Mary McGlohon, a doctoral candidate in the Machine Learning Department, on applying algorithms that find patterns in social networks to looking for links in financial data.

Steier finds his work rewarding because of its broad reach. "The way that the financial statements of companies you invest in are audited are likely to be impacted by the work we do here," he says.

—Jason Togyer (HS'96)

Andrew Dubois

B.S., Computer Science, Carnegie Mellon University, 2003

He's gone from checking source code to checking vendor performance. And while they wouldn't seem that similar, Andrew Dubois (CS '03) says the two fields share some common virtues. For one thing, you need to troubleshoot and uncover hidden problems in both code and contracts.

"You have to understand the big picture but be willing to step through each element," says Dubois, a senior project manager at Stamford, Conn., based UBS Investment Bank. "You're much better off if you can identify and resolve an issue early before the impact multiplies. It's the same thing with vendor management—people are counting on you to be the quality control."

An effective audit report, he says, should be as elegant as a concise piece of code: "It's got to be clean and it can't leave any ambiguity. If someone looks at your findings, they need to know exactly what was discovered and the level of severity."

Dubois joined UBS as an application developer and regularly found himself negotiating with vendors. Soon he was crafting the terms of the deals. From there, it was a short hop to Dubois's current position, where he leads a team of five people who enforce the terms and conditions of contracts for the U.S. investment-banking arm of the Swiss-based global financial services company, UBS.

An SCS education left Dubois well prepared to make the transition, he says, which is one reason why he likes to hire computer science students at UBS.

"You learn attention to detail, and how to manage your time," says Dubois, who balanced his own coursework with serving as a student senator and student government treasurer and playing trombone in the Kiltie Band. He also was a founder of Carnegie Mellon's Online Gaming Society and an SCS teaching assistant.

Dubois currently serves on the SCS alumni advisory board, and he and his wife, fellow grad Jennifer Li (CS'04), made it back for Spring Carnival this year. Their Oakland hotel was filled with "floor after floor of CMU people," Dubois says, laughing. "It was incredible." —JT

Watzman, PR director since 1991, retires

Anne Watzman is one of a handful of people who can say their career spanned both coal mining and data mining.

Working out of Pittsburgh as a staff writer and free-lance correspondent, Watzman spent more than two decades covering the biggest industrial stories in the nation for McGraw-Hill's magazines, including Business Week, Engineering News Record, Chemical Week and—yes—Coal Age.

But as Pittsburgh changed from the "Steel City" into a center for science, research and education, Watzman changed direction, too. In 1987, the Squirrel Hill native and Penn State alumna joined Carnegie Mellon as a member of the university's public relations staff. Four years later, she became director of media relations for the recently created School of Computer Science.

Then-Dean Raj Reddy put her in the "cave"—a windowless room of cubicles in the 4600 wing of Wean Hall—where her first assignment was to cull articles from The Wall Street Journal. "Pretty soon," says Watzman, laughing, "it got busier and busier, and Raj had to read the Wall Street Journal himself." (She got a real office, too.)

For nearly a generation, Watzman, who retired April 1, had VIP access to pioneers in artificial intelligence, computer networking, machine learning, software engineering and robotics—and a seemingly endless supply of great stories to tell to media outlets around the world.

Still, she admits, it was sometimes a challenge to get reporters to understand that computer science is more than just robots. Pressed to recall the most important story of her tenure, Watzman rattles off a half dozen, from model checking and the DARPA Grand Challenge to Randy Pausch's "Last Lecture."

Since 2006, Byron Spice, former science editor of the Pittsburgh Post-Gazette, has served as "co-director" of public relations for SCS. He now soldiers on alone as Watzman looks forward to traveling, gardening and spending more time with her four grandchildren, aged 20 months to 9 years old. "I also feel obligated to sort through the thousands of family photos that have accumulated in my house," she says.

Watzman may even write a novel. "I guess I have to wait and see if I have anything to say that's worthwhile," she says. Her record over the past 18 years suggests that's the least of Watzman's worries. — Jason Togyer


FILE PHOTO

Watzman: Looking forward to traveling, gardening and spending more time with her grandchildren

Gates delivers keynote as Qatar campus hosts ICTD 2009

Microsoft Chairman Bill Gates was the keynote speaker when the university's campus at Education City in Doha, Qatar, hosted the third International Conference on Information and Communication Technologies and Development from April 17–19.

More than 350 technologists, social scientists and others attended ICTD, considered the premier multidisciplinary forum for researchers designing information and communications technologies for developing countries. The conference—which was significantly expanded this year—also gave the university a chance to show off its new building, which opened two months earlier.

Workshops and panel discussions focused on topics such as technology-based approaches to education, literacy and healthcare; the challenges of deploying technology in rural settings; and implications of new eServices for policy, government and non-governmental organizations.

CARNEGIE MELLON IN QATAR PHOTO

Gates with Carnegie Mellon Qatar Dean Chuck Thorpe

Gates spoke April 18 to more than 800 people about the work of the Bill and Melinda Gates Foundation, which has a strong focus on bringing technology to the developing world. Poverty "wastes human potential," he said. "Helping the poorest people is the most important thing we can do. Technology is a tool for achieving that goal."

Assistant Research Professor M. Bernardine Dias (CS'00,'04) served as conference chair for ICTD 2009, while Rahul Tongia (E'98), senior systems scientist in the Institute for Software Research, served as program committee co-chair. — Philip Lehman

New faculty member joins CS department

Venkatesan Guruswami, a leading researcher in theoretical computer science whose work has provided new insights on digital communications, became an associate professor in the Computer Science Department on July 1.

A graduate of the Indian Institute of Technology at Madras and Massachusetts Institute of Technology, Guruswami had been an associate professor of computer science and engineering at the University of Washington since 2002.

His research has focused on error-correcting codes in digital media, enabling the more robust detection of weak signals in noisy environments. Guruswami has also explored the degree to which approximation provides answers for computationally difficult problems.

Guruswami was a visiting professor at Carnegie Mellon during the past academic year.

Habermann named 'influential educator' by SIGSOFT

The late A. Nico Habermann has been honored with the inaugural Influential Educator Award from the Association for Computing Machinery's Special Interest Group on Software Engineering.

Habermann, founding dean of the School of Computer Science, was honored "for significant and lasting contributions to the field of software engineering as a teacher and mentor," according to the citation. His widow, Marta Habermann, accepted the award May 21 at the International Conference on Software Engineering. Two of Habermann's children, Frits and Marianne, also attended. "This is the most appropriate award Nico ever received because he was always a teacher," Marta Habermann said.

Laurie Williams, associate professor of computer science at North Carolina State University, also received the award this year.

Habermann, who died in 1993, joined the Carnegie Mellon faculty in 1969. He was head of the Computer Science Department from 1980 to 1988, when he became dean of the new School of Computer Science; and helped establish the Software Engineering Institute, which he served as acting director from 1984 to 1985.

CARNEGIE MELLON PHOTO

Marta Habermann, widow of Nico, accepts the Influential Educator Award plaque while her son Frits (left), SIGSOFT chair William Griswold and daughter Marianne look on.

Report: Gene regulators similar to cloud computing networks

Gene regulatory networks in cell nuclei are similar to cloud computing networks, reports an international research team led by Ziv Bar-Joseph, assistant professor of machine learning and computer science.

The finding—reported June 16 in the online peer-reviewed journal Molecular Systems Biology—helps explain not only the robustness of cells, but also some experimental results that have puzzled biologists.

"Similarities in the sequences of certain master genes allow them to back up each other to a degree we hadn't appreciated," said Bar-Joseph, a member of Carnegie Mellon's Ray and Stephanie Lane Center for Computational Biology. That's analogous to a cloud-computing network, which keeps working even when individual processors fail.

The research team said this explains the curious results of experiments in which biologists removed one master gene from a cell at a time, only to find out that most of the genes controlled by the master gene remained activated.

Anthony Gitter, a graduate student in the Computer Science Department, was lead author of the article. Other authors included Itamar Simon, Zehava Siegfried and Michael Klutstein of Hebrew University Medical School in Jerusalem, Oriol Fornes of the Municipal Institute for Medical Research in Barcelona and Baldo Oliva of Pompeu Fabra University, also in Barcelona.

— All stories on this page compiled from staff reports

Yahoo! makes research grants to four doctoral students

Four SCS Ph.D. students are among 20 young researchers selected as winners of Yahoo!'s first Key Scientific Challenges program.

Honored were Pinar Donmez and Jaime Arguello of the Language Technologies Institute and Polo Chau and Yi Zhang of the Machine Learning Department. Arguello and Chau won recognition for their work in search technology, while Donmez and Zhang were cited in the category of machine learning and statistics.

Each recipient received $5,000 in unrestricted seed funding for their research, exclusive access to certain Yahoo! datasets and the opportunity to collaborate directly with the corporation's scientists.

No other university had as many winners in the program as Carnegie Mellon. "We clearly were impressed by the quality of the applicants from Carnegie Mellon and believe Pinar, Yi, Jaime and Polo each hold great potential for making significant contributions as researchers," said Ken Schmidt, Yahoo! director of academic relations.

RENEE ROSENSTEEL PHOTO

SCREEN SHOT

We thought dragons were the only animals that flew over Wean Hall, but photographer Renee Rosensteel spotted this peregrine falcon high above the Pittsburgh campus.

Although these birds of prey occasionally startle unsuspecting pedestrians along Frew Street, they're no danger—unless you happen to be a tasty pigeon or squirrel.

Listed as a Pennsylvania endangered species, peregrine falcons can reach speeds of 40 to 55 mph in level flight and dive at up to 200 mph, making them the fastest birds in the world. There are only two known peregrine falcon nesting sites in the western half of the state—atop the Gulf Tower in downtown Pittsburgh and at the University of Pittsburgh's Cathedral of Learning in Oakland.

According to the Western Pennsylvania Conservancy, which tracked the falcons locally from the 1990s through 2008, those buildings are ideal nesting sites because they offer tall, secluded ledges, perfect for raising chicks—and teaching them to fly. Pittsburgh's National Aviary now manages the local falcon recovery project.

Renee's photo also shows the new Gates Center for Computer Science nearing completion at lower left, and its location relative to Newell-Simon Hall, Wean Hall and the rest of campus.

Find us on Facebook

Now you can find The Link magazine on Facebook! Become a fan of The Link to connect with other SCS friends and alumni and get updates between issues.

Plus, The Link's website has been redesigned and is now accepting your comments. You can download copies of past issues and see additional material that doesn't make it into print at link.cs.cmu.edu.

Carnegie Mellon

Office of the Dean
5000 Forbes Avenue
Pittsburgh, PA 15213

Then and Now

CARNEGIE MELLON UNIVERSITY ARCHIVES

One of the most ambitious building projects in Carnegie-Mellon University's short history was announced March 14, 1968. The $13 million "Research-Computer Building" would include the University Computation Center, Engineering and Science Library and a 200-seat auditorium, along with offices, laboratories and classrooms for biotechnology, computer science, materials science, mathematics, metallurgy, physics and statistics.

Shown above right making the announcement are Carnegie-Mellon President H. Guyford Stever; Aiken Fisher, chairman of the board of trustees; Frederick Bloom, chairman of the trustees' building and grounds committee; Fred Foy, chairman of the university's Fund for Distinction; and Walton Forstall, professor of mechanical engineering and faculty representative on the planning committee.

The building was designed by the Pittsburgh firm Deeter Ritchey Sippel, better known as architects of Mellon Arena and the former Three Rivers Stadium. The photo above left shows construction progress as of May 1969.

Renamed "Science Hall," the building opened in May 1971. A decade later, it was rededicated in honor of Raymond J. Wean Sr. (E'17), a pioneering engineer, distinguished alumnus and emeritus life trustee of the university.

The university has scheduled a Sept. 22 dedication for its latest ambitious building project—the Gates Center for Computer Science and Hillman Center for Future-Generation Technologies. For details see page 3, and for more background on Raymond Wean and Wean Hall, visit The Link website at link.cs.cmu.edu.