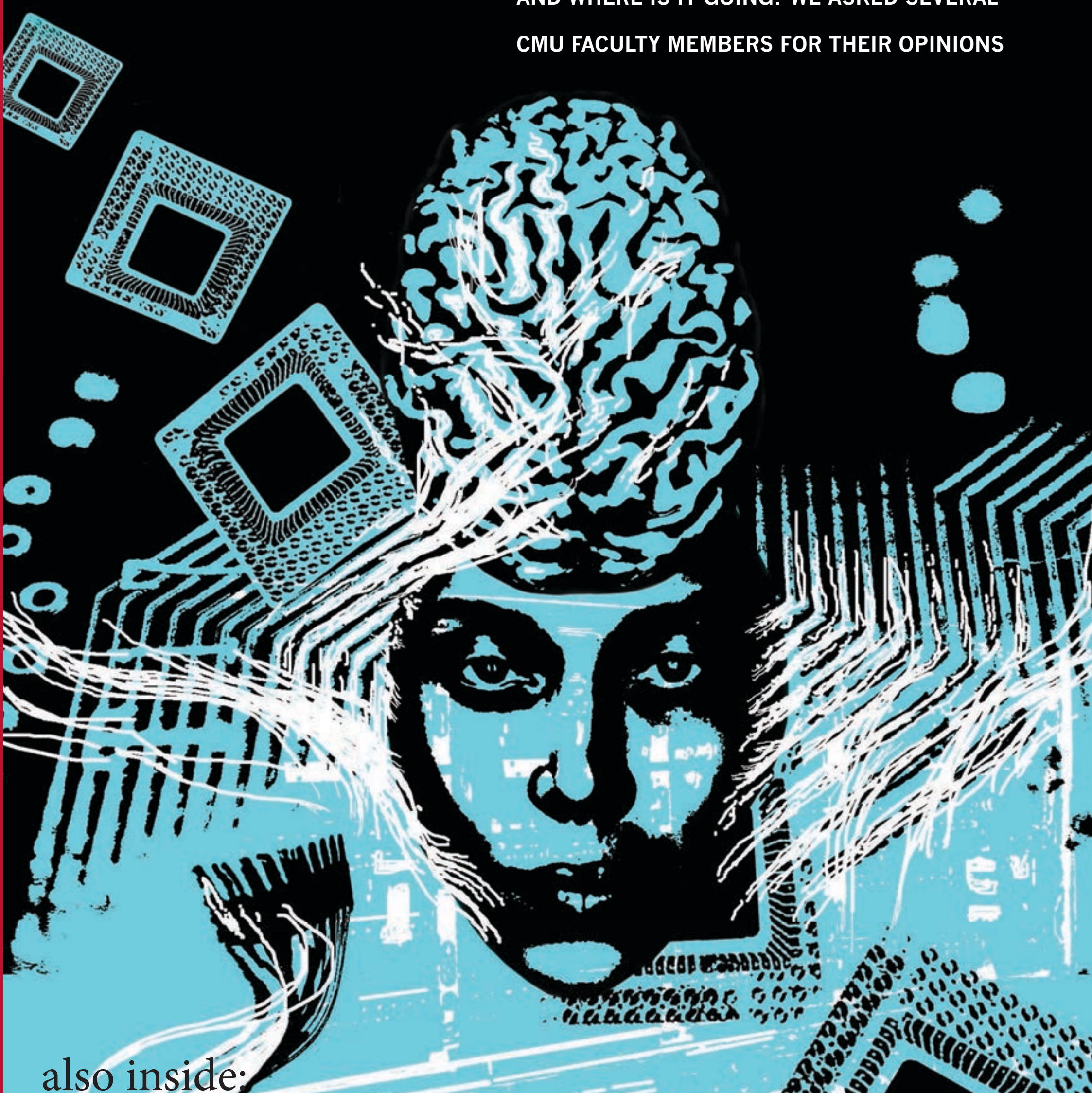
thelink.

The Magazine of the Carnegie Mellon University School of Computer Science

SPRING 2016 ISSUE 10.1

The future of A.I.

WHERE HAS ARTIFICIAL INTELLIGENCE BEEN — AND WHERE IS IT GOING? WE ASKED SEVERAL CMU FACULTY MEMBERS FOR THEIR OPINIONS



also inside:

CMU'S NREC: ALIVE, AUTONOMOUS AND WELL

HCII'S JESSICA HAMMER ISN'T JUST PLAYING GAMES

MOBILITY FOR EVERYONE AND THE INTERNET OF THINGS

Carnegie Mellon University

thelink.

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
Andrew W. Moore

Editor
Jason Togyer

Contributing Writers
Scott Fybush, Meghan Holohan, Nick Keppler, Ashley Patton, Linda K. Schmitmeyer

Photography
All photos by Carnegie Mellon University, except: p. 8, Natural Resources Conservation Service of U.S. Department of Agriculture; pp. 9, 10 (bottom) and p. 11 courtesy National Robotics Engineering Center; p. 10 (top), courtesy Herman Herman; p. 14, Jason Togyer photo; p. 16, Wade H. Massie photo; p. 19 and p. 24 (top) courtesy Carnegie Mellon University Archives; p. 24 (down) and p. 25, courtesy Manuela Veloso; submitted photos: pp. 13, 15, 16, 26, 29, 30, 31, 41, 44 and 45.

Graphic Design
Digital & Creative Services

Office of the Dean
Gates Center 5113
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213

Andrew W. Moore, dean, School of Computer Science

Philip L. Lehman (CS 1978, 1984), associate dean for external relations and strategic initiatives

Mark Stehlik, assistant dean for outreach

Byron Spice, director of media relations

Mark R. Power, director of marketing and communications

Ashley Patton, director of engagement and annual giving, School of Computer Science and Department of Electrical & Computer Engineering

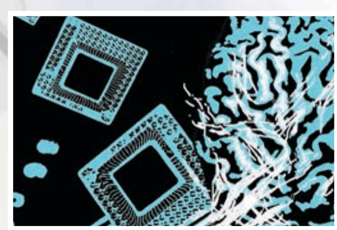
Niccole Atwell, senior associate director of engagement, School of Computer Science and Department of Electrical & Computer Engineering

Jason Togyer (DC 1996), editor

Phone: 412-268-8721
Email: TheLink@cs.cmu.edu
Web: www.cs.cmu.edu/link
Facebook: facebook.com/SCSatCMU
Twitter: twitter.com/SCSatCMU

Produced for the School of Computer Science by Digital & Creative Services, April, 2016, 16-265.

contents



on the cover

Cover illustration by

Jeffrey Katrencik (A 1985)

Dean Andrew W. Moore

Making big plans

Many faculty and alumni have told me they miss the days when Carnegie Mellon was engaged in big, multidisciplinary computing research projects.

Such projects, for instance, as the creation of the Andrew Network, which demonstrated to the world the utility of a GUI-based operating environment that offered seamless, high-speed file transfers from place to place; it anticipated on a campus scale the cloud environment that the Internet now makes possible on a global level.

Large, multidisciplinary projects break people out of their silos. They present challenges and constraints that spark creativity and discovery, and they raise interesting questions that lead to productive future research. For that reason, they are ideal for demonstrating the breadth and quality of Carnegie Mellon research in every facet of computer science and computer engineering.

The more romantic people amongst us might remember a quote attributed to the famed architect Daniel Burnham, who designed some of the world's earliest skyscrapers and shaped the cities of Chicago and Washington, D.C. "Make no little plans," he said, "they have no magic to stir men's blood. Make big plans—aim high in hope and work."

We are making big plans, and we hope to create magic to stir the blood of women and men. With the help and guidance of SCS Council, we are developing large-scale and long-term research projects that we are tentatively calling "moonshots"—ideas and concepts in computing and robotics that both show off the abilities of our talented faculty, staff and students, and help push us into new and exciting frontiers.

In the first round of discussion, more than 30 ideas were presented, and six were selected. We will be examining additional ideas for "moonshots" in the months to come, and you will be hearing about them soon. I predict that some of our goals will exceed our grasp, but that by aiming high, we also will celebrate achievements, and make discoveries, that we hadn't anticipated.

Down here on Earth, I am pleased to announce that Guy Blelloch will become our new associate dean for undergraduate programs at the end of the spring semester. In addition to being an accomplished researcher of data structures, parallel algorithms and programming languages, Guy has always been dedicated to the success of students and fellow faculty members, and was one of the leaders in the creation of the Gates and Hillman Centers.

One of Guy's new jobs will be to lead our undergraduate growth strategy. There are many decisions that we, SCS, must make here. As a consensus-builder among groups of faculty and between departments, he is the right person for this incredibly important duty.

Klaus Sutner, after more than a decade of service as associate dean, is moving on from that role to spend some time working on new course offerings and looking at new ideas around educational technology. Klaus has performed major services for the School of Computer Science. He has tirelessly seen us through tough times and good times and a great deal of change in undergraduate education.

Klaus: from the bottom of my heart, thank you for what you have done.

I am also delighted to announce that Manuela Veloso has been appointed the new head of the Machine Learning Department. The search committee worked very hard, and I believe Manuela is perfect for this role. You can read more about her appointment on page 39 of this issue of The Link.

And thank you to the outgoing and founding head, Tom Mitchell. In a career of astonishing accomplishments (with more to come!) his creation of the world's first Machine Learning Department has been a highlight. He has had an impact on hundreds of careers and has been one of the primary forces pushing the discipline of machine learning to become a major component of the world's economy. He made this happen while following the principles of SCS's founders: cooperation, mutual support and an absolute focus on the most important and impactful research.

Finally, we are making progress in streamlining and bringing order to many internal processes that had frustrated faculty, staff and students, including long queues to get contracts and research approved. I appreciate the frank and honest feedback that we've received; together, we are working to set things right.

Andrew W. Moore
Dean
School of Computer Science

Life, the university and everything

A Google-funded research expedition is allowing CMU faculty and students to live inside the 'Internet of Things'—and understand the promise and perils of unprecedented interconnectedness

By Scott Fybush

If you woke up this morning and asked your Amazon Echo for a weather forecast or relied on your Nest thermostat to get your house toasty, you're already living in the world of the "Internet of Things," or IoT, in a way that would have been unimaginable as recently as the start of this decade.

As the Internet of Things swiftly marches from concept to reality, Carnegie Mellon scientists are leading a Google-funded expedition that will use the campus and the surrounding city as a test bed to build new uses for the technology while protecting users' privacy and safety. There are four broad research areas being explored—machine learning, end-user applications, use of the campus as a "living laboratory," and the privacy and security implications of the technology.

For the first year of research, CMU received $500,000 from Google to be the lead university in the project, code-named "GIoTTo," working alongside Cornell University and the University of Illinois at Urbana-Champaign.

"The fact that [Google] wanted to build an open framework was really compelling to us," says project leader Anind Dey, director of CMU's Human-Computer Interaction Institute. "Any one company that wants to own all of IoT isn't going to succeed, and Google recognizes that nobody's going to own this and you're better off having an open platform that anyone can connect to."

Google's goals for that open platform include ubiquitous inexpensive sensors, better user interfaces for those sensors and the creation of tools to help end users come up with their own uses for all of that new technology. At CMU and its partner schools, those goals are spread across a variety of teams focusing on specific tasks, including one very big one: maintaining privacy and security in a world where it's possible to collect so much data on any individual's behavior.

"We treat privacy as a first-order concern," Dey says. One way to keep sensitive data under control is the specific project of Mahadev Satyanarayanan (CS 1979, 1983), better known as "Satya," CMU's Carnegie Group Professor of Computer Science. Satya is focusing on developing the concept of "cloudlets," miniature data centers that could live in individual homes or offices to manage the data that's passed from local sensors to a broader network.

Ph.D. student Adrian deFreitas (CS 2015) and HCII Director Anind Dey are part of a team of CMU researchers leading an initiative to create a robust platform that will enable Internet-connected sensors, gadgets and buildings to communicate with each other.

"If all of your data leaves your home and goes straight to the cloud, you have no opportunity to control the exposure of that data," he says. "It's quite easy, for instance, to find an IoT sensor that you can attach to your water meter and every time water is used, it's able to report back that this much water was used. Why is it anyone's business how many times I flush the toilet each day?"

Instead, he says, a cloudlet could manage privacy policies locally to determine who gets access to which categories of data. In addition to controlling privacy, a cloudlet can also reduce the bandwidth demands that the rise of IoT could put on the entire Internet.

"Motion sensors, water sensors—the amount of data they generate is small, maybe a few tens of kilobytes every few minutes or seconds. But some of these sensors are video cameras, and if I have 10 cameras in my house or maybe 200 cameras that are watching, you're talking about pretty substantial data rates," Satya says. "Transmitting that to the cloud is not economically feasible."

Instead, a cloudlet could collect all that data and then use locally determined rules to focus only on what its owners find important.

"Suppose you had face-recognition capability running on your cloudlet at home," he says. "If the faces it recognizes are your family members, you'd say, don't bother shipping their faces to the cloud. But any stranger on your property, that's worth shipping"—at which point other pieces of the IoT landscape could alert a security company or a user's mobile device.

Making those facial-recognition capabilities work, as with so many pieces of the emerging IoT ecosystem, requires the sort of test bed for which a curious campus such as CMU is ideal.

"We have people who are technical and people who are non-tech, so that provides us with a really nice mix," says Dey.

In its initial phases, CMU's expedition leaders are exploring some of the problems that are endemic to campus life.

"For this project, we have seven faculty members who come together for a meeting," Dey says. "On any given week I have no idea who's in town and who isn't. So the question I always have is, do I have time to get a cup of coffee before everyone gets there?"

Using a combination of Amazon's Echo devices as well as other sensors, the GIoTTO team has developed an application that figures out where other meeting participants are, whether or not they're planning to attend the meeting, when they'll arrive—and, most critically, how long the line is at the cafeteria where he might or might not be able to sneak in that cup of java before they converge at his office. Along a similar line of inquiry, Satya imagines a situation in which he's in his own office and needs help, say, moving a heavy package. Instead of sending an email or text that would go to an entire office's worth of people, IoT sensors could suppress the sending of that message to anyone who's not actually in the building and able to help.

Projects like that might sound trivial at first, but Dey says they're actually designed to test some bigger ideas. For instance, how many sensors are really needed to get useful information about a home or office? Instead of needing a sensor to determine whether his window is open or closed, for instance, the GIoTTO team is creating applications that could use existing sensors—ones that might be measuring air pressure or temperature—that could be trained to detect the subtle differences between a room with an open window and one where the window is closed.

Dey and his team are already working on the ways in which all of this technology could eventually reach consumers. In mid-November, a CMU team demonstrated for Google a first-generation version of an "IoT app store" that could tie together many of their concepts in a way that average users could easily tap. In the months to come, Dey says, the team will be expanding its reach to test more of its concepts in more of the real world.

"We're really excited to deploy this to a large part of the CMU campus," he says, "and then to Pittsburgh to help make this a smart city." →□

> "Suppose you had face-recognition capability running on your cloudlet at home. If the faces it recognizes are your family members, you'd say, don't bother shipping their faces to the cloud. But any stranger on your property, that's worth shipping."
>
> MAHADEV SATYANARAYANAN

Alive, autonomous and well

When strawberry growers needed a way to sort plants automatically, researchers at CMU's National Robotics Engineering Center moved from idea to prototype in five years.

As robotics technology increasingly moves from the theoretical to practical, clients are still coming to CMU's National Robotics Engineering Center—just not with 'easy projects'

By Nick Keppler

In the strawberry-growing regions of California, nearly every farm displays a "help wanted" sign during harvest season. Over the years, the labor shortage across the state's 40,000 acres of strawberry farms has been growing progressively worse. One farm would raise pay to attract laborers from the farm next door, and another would then follow, increasing labor costs in an ever-escalating wage war. Farms started making greater use of the federal H-2A agricultural guest worker program to import seasonal help from Mexico, but the process was costly and paperwork intensive.

Then, in 2009, a farm owner had an idea: Task a robotics firm with inventing a technology to reduce the labor needed to sort plants for the strawberry nursery. The development cost would be stiff, so the owner of that farm, Lassen Canyon Nursery, convinced some colleagues to form an LLC to fund a future for strawberry processing. They took the project 2,500 miles east, to CMU's National Robotics Engineering Center.

"We didn't even have a mature technology for them [at the time]," recalls Jeff Legault, NREC's director of strategic business development. "They had some seed money and we agreed to take a look at their problems."

Five years later, NREC researchers had a prototype: a conveyer belt monitored by robot eyes that can classify strawberry plants as "good quality," "bad quality" or "questionable" based on appearance. The device uses jets of air to separate the three kinds of strawberry plants, leaving the "questionable" ones for human inspection. By eliminating this one task, the machine could greatly reduce costs for large farms.

This is typical of NREC's method: Someone has a complex problem and the engineers craft a solution using robotics. Whereas NREC's projects once came almost entirely from U.S. government contracts—particularly branches of the military—increasingly, NREC's clients are from the business world. Herman Herman (CS 1993, 1996), who was named director of the nearly 20-year-old institution in February 2015, says NREC's work "is now roughly 50/50 split between the public and private sector."

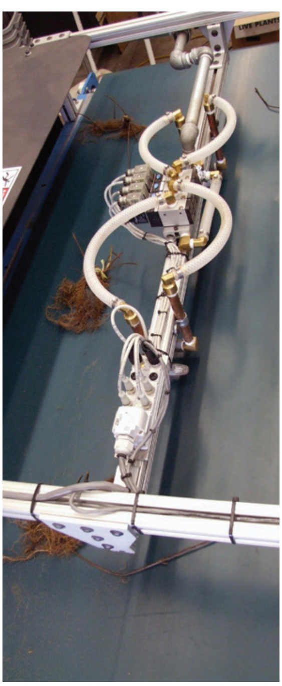
"If this was a place where everyone was working on a Ph.D., the focus would be solving (purely) technical problems," Herman says. Instead, he says, NREC is unique in academia in that it tackles real-world challenges in ways that have immediate, practical benefits. "We have to not just solve technical problems, but resolve those problems in a cost-efficient manner," Herman says.

Another trend is that clients are coming to NREC, as opposed to NREC searching them out, Legault says. Part of the reason is that robotics technology is becoming more affordable, and thus more mainstream. "In the past, there often was viable technology we could develop, but it would have been too expensive to implement commercially," he says. "That's now changing."

NREC is located about three miles from CMU's Pittsburgh campus in a warehouse-like building in the Lawrenceville neighborhood. About 100 people work at the center at any given time, including a mix of full-time NREC staffers, school faculty and a handful of students.

On the inside, the facility looks like something from a comic book movie—maybe a secret subsidiary of one of Tony Stark's companies. The open floor contains a prototype Humvee programmed to create digital maps and change detection as it roams—sort of a version of the cars that take Google Street View pictures, but for dangerous or war-torn areas. Nearby, there's a self-driving all-terrain vehicle called "REC-Tamer," which can be lowered from a helicopter into a disaster zone. And there's CHIMP, a versatile robot with orangutan-like arms designed for search and rescue missions. Built at NREC, CHIMP won $500,000 in the June 2015 DARPA Robotics Challenge, a two-day event that tested 24 of the world's most advanced robots against one another in a series of disaster-relief tasks. A looping video on display shows CHIMP picking up and using power tools to enter a mockup of a disaster area.

Whether they're from the private sector or the military, NREC clients these days want robots that are autonomous — able to act and maneuver with little or no human input. REC-Tamer, for instance, was developed for the U.S. Defense Department by NREC in cooperation with Sikorsky Aircraft. In a demonstration on Oct. 27 in West Palm Beach, Fla., a modified Black Hawk helicopter autonomously flew itself to a remote location and



On the field-tested strawberry sorter, air jets (left) blow the classified plants into bins based on quality.

> “The NREC is unique in academia in that it tackles real-world challenges in ways that have immediate, practical benefits. We have to not just solve technical problems, but resolve those problems in a cost-efficient manner.”
>
> HERMAN HERMAN

lowered the REC-Tamer to the ground, where it autonomously navigated over six miles of terrain, surveying environmental conditions as it went. The project also won the “Distinguished Engineering Project Achievement Award” from The Engineers’ Council.

The project was in part spurred by the disaster at the Fukushima nuclear reactor in Japan, Herman says. “After Fukushima, the U.S. Army realized it didn’t have an autonomous vehicle that could go into such an environment to survey damages or contamination. People often forget that the army is in charge of a lot of rescue and humanitarian work.”

The NREC building also has a prototype autonomous tractor, commissioned by Deere & Company, which can perform agricultural tasks, such as spraying or mowing, on its own. The prototype has accumulated thousands of miles of autonomous operation during extensive testing in Florida. There’s even a boat inside the Lawrenceville facility, propped up on a hitch trailer, which NREC engineers are enhancing to be able to navigate by itself to and from docks, picking up and discharging passengers—a self-driving water taxi. To facilitate all the fabrication of the robots, the building includes an in-house full machine shop with 3-D prototyping capability and a temperature-controlled test chamber where robots meant to function in severe weather conditions, such as the arctic or the desert, can be tested.

It’s hard to generalize about the process of commercializing technology developed at NREC, because it varies greatly, depending on the scope of the work and the market for which it’s intended—Herman says it can take anywhere from one to five years and cost between “hundred of thousands” and “tens of millions.” But once NREC develops and extensively tests a solution for a client, those clients can have their finished product fabricated or commercialized by one of several different specialty manufacturing companies— of which are spinoffs from CMU, founded or run by alumni or former NREC employees, and which often are located in the Pittsburgh region. CMU retains a stake in technology created at the university, which creates revenue streams to support its programs.

In a test in October 2015, NREC teamed with Sikorsky to use an autonomous helicopter to deploy a CMU-developed Land Tamer unmanned ground vehicle. The helicopter picked up the vehicle, flew a 12-mile route, delivered it and released it.

Past and current NREC team members pose with the CMU Highly Intelligent Mobile Platform robot, which won $500,000 in 2015 at the DARPA Robotics Challenge.

Four new NREC projects total $11 million in research

NREC has been selected as a prime contractor or subcontractor on four major new federal research projects totaling more than $11 million over the next three years. The projects range from research on a wheel that can transform into a track to automated stress testing for critical software.

The new research initiatives include:

Robustness Inside-Out Testing

This $4 million project for the Defense Department's Test Resource Management Center project will dramatically increase the power of robustness testing used to find defects in complex unmanned autonomous system software. Michael Wagner (E 1998, 2002), NREC senior commercialization specialist, and Philip Koopman (E 1989), CMU associate professor of electrical and computer engineering, are applying active-learning algorithms to develop a capability called "backward chaining" that will better identify inputs that can cause failures deep within a software system.

Reconfigurable Wheel-Track for All-Terrain Mobility

This $4.2 million project for the Defense Advanced Research Projects Agency is developing technology that would enable a wheel to transform into a track, and vice-versa, so vehicles could tackle a variety of terrains.

The technology would allow for an increase in versatility and access across a variety of terrains for both manned and unmanned ground vehicles. Dimi Apostolopoulos (E 1991, CS 1998), senior systems scientist, heads the project.

Automated Phenotyping System for Genetic Improvement of Energy Crops

As part of a $1 million U.S. Department of Energy project being done in conjunction with Texas A&M University's AgriLife Center, researchers are developing robotic vehicles that can monitor sorghum plants being bred to enhance their use as energy feedstocks.

A mechanical arm will capture images and make measurements of the crops; machine vision and learning algorithms will predict plant growth. David Wettergreen (S 1987, E 1989, CS 1995), research professor of robotics, is the principal investigator for the CMU portion of the project, with Apostolopoulos and Herman serving as co-investigators at the NREC.

Aircrew Labor In-Cockpit Automation System

This DARPA program, dubbed ALIAS for short, will develop automation technology that could be added to existing aircraft to enable operation with a reduced onboard crew. Sikorsky is leading the research, with a team that includes NREC, United Technologies Research Center and Veloxiti Inc., on a contract valued at $2.4 million.

For the past 20 years, NREC has been an important national resource, combining unique technical skills and testing capabilities to solve problems that other groups can't," says Martial Hebert, director of CMU's Robotics Institute, which includes the NREC. "These new projects are a reminder that NREC continues to advance the art and science of robotics and that it remains a vital part of Carnegie Mellon's Robotics Institute."

—Byron Spice

Roughly the size of a human, CHIMP can move on all fours or can stand upright, and operate for 90 minutes or more using batteries. The robot was the only competitor in the DRC that was able to right itself after falling down.

To facilitate the commercialization of its technology, part of the NREC facility is dedicated for incubating spinoff companies and related activities. One entity that is currently a resident at NREC is The Robotics Hub, an independent startup accelerator meant to fund robotics high-tech companies in Pittsburgh, including additional NREC spinoffs. The Robotics Hub was started with a seed funding from GE Ventures, an investment firm associated with General Electric, to pursue various commercial applications for robotics technology created at CMU.

Last spring, NREC made headlines, but not for its world-changing research. In February 2015, the ride-sharing company Uber established a Pittsburgh research lab to develop driverless vehicles, and hired about 40 NREC employees. Media reports used words like "gutted" and "completely raided" to describe the hiring process; Herman says it was nothing of the sort, and most projects continue as usual at NREC, where staffing levels have been "healthy and growing again," he says.

"I think what is lost in the media reports is the reason why Uber came to Pittsburgh to establish their research lab," Herman says. It's because, in his words, Pittsburgh has an "enviable wealth of talent" with top researchers being educated and employed at centers such as NREC. "I've always said, if you work at NREC for three or four years, you can get a job anywhere," Herman says, adding that it's never been unusual for NREC staffers to leave for private enterprise. In fact, the center prides itself on the number of its alumni who have started robotics spin-off companies or taken important positions in developing autonomous robotics technology at companies such as Google.

"We're accustomed to a natural flow of technical and research talent back and forth between academia and industry," says Andrew Moore, dean of SCS and a professor of robotics and computer science. Moore himself took a leave of absence from CMU in 2006 to launch Google's Pittsburgh office. Although four Robotics Institute faculty are now on leave working at Uber, Moore says, CMU has hired four new faculty in robotics as well as another 13 in machine learning, systems and algorithms.

In addition, NREC has hired 10 new technical staff members in the past six months, Herman says, and plans to hire another five to 10 in the coming months to augment its existing staff. "Our staffing has always reflected the number of projects we've had, and we have the advantage of being part of the Robotics Institute, which has a deep pool of technical talent to draw upon in terms of faculty, staff and students," he says.

If the future research direction of NREC isn't easily predictable, that's the nature of working on advanced technology, Herman says. The autonomous vehicle technology that currently underpins many of NREC's current projects, for instance, didn't even exist in a cost-feasible form six years ago. What will NREC be working on six years from today? No one is certain, but NREC researchers say they're sure they'll be called upon. The "sky is the limit" in terms of new robotics applications, Herman says. "We're encouraging companies to dream big, and we'll make it happen for them," he says.

Most NREC clients—such as mining companies, aircraft manufacturers and the U.S. military—already have top-notch engineers working for them, Legault says. They only outsource the projects that are the most perplexing to their qualified staff. "They come to us with difficult problems that require a great level of expertise to come up with the unique solution," he says. "People don't come to us with the easy projects."

—Nick Keppler is a Pittsburgh-based freelance writer who frequently contributes to magazines and websites such as Vice, Slate, Nerve and The Village Voice.

More mobility for everyone

Smartphones have revolutionized life for most people, but many of their apps exclude those with limited or no vision. Researchers are working on ways to make mobile devices an "all-access" pass for the sight-impaired

By Nick Keppler

Here is a uniquely 21st-century problem that blind people face: They hear someone say, "Hello" and retort with a "hello." Then the first speaker trails off into what seems like a non sequitur, and the blind pedestrian awkwardly realizes he or she was answering a cell phone.

"It's a very real scenario," says Chieko Asakawa, a researcher at IBM, who has been blind since age 14. "Often a person says 'hi' and I say 'hi' back and a colleague tells me he was talking on the phone. Socialization is a very big challenge."

But Asakawa, a veteran IBM researcher whose past projects have included a Braille word processor and a talking Web browser, has a solution—one that relies on the same technology that begat the dilemma.

Asakawa, an IBM research fellow, is currently the IBM Distinguished Service Professor in CMU's Robotics Institute. She and her collaborators, Kris Kitani of the Robotics Institute and Jeffrey Bigham of the Human-Computer Interaction Institute, imagine a time when a blind person will plug into a smartphone app that connects to a sensor signal, which will guide them step by step, literally, through a public space. The same Siri-like voice that will someday be able to tell them to "walk three steps and open the door on the left" also might then identify passersby via facial recognition software. It will even say if the acquaintance seems happy or sad, and whether or not they're holding a phone to their ear.

Called NavCog, it's one of several projects under development by CMU and IBM researchers that seeks to use personal technology to help people with sight impairments. From apps to helper robots, it's the kind of technology that has the potential to remold almost every aspect of the lives of the blind.

Chieko Asakawa

M. Bernardine Dias

Asakawa navigates CMU's Pittsburgh campus using NavCog.

"Blind people hope to gain much from well-designed and robustly implemented technology," says M. Bernardine Dias, an associate research professor at the Robotics Institute. "In many situations, technology can play a critical role in enhancing independence, safety and access to new opportunities for blind people and people with visual impairments and other disabilities."

Asakawa says that her current work represents a leap into the "real world." "My [past] research focus has been about information on the net," she says. "Now, I am thinking, 'How can I get to the classroom, the post office, the museum?'"

The NavCog project on which she is working has spread 250 Bluetooth signal emitters called beacons—white squares a bit smaller than the average smoke detector—through three School of Computer Science buildings: Newell-Simon Hall, the Gates Center for Computer Science and Wean Hall. The signal currently offers directions around campus. You can plug in, choose start and end points, and a voice guides you. "It's like a navigation system that makes both outdoor and indoor navigation seamless, for humans," Asakawa says. Her team hopes that someday, major foot-traffic areas—like airports, bus stations, malls and concert venues—are dotted with beacons to make areas more accessible for people with limited or no vision. They hope to add other facets, such as facial recognition, to an overall smartphone package for the sight impaired, allowing an easier pedestrian experience.

Kris Kitani, a collaborator on NavCog, has developed another smartphone app called EdgeSonic, which he describes as "audio Braille for images." EdgeSonic can take an image captured by a smartphone and turn it into a crude representation of the most pronounced lines in the photo. The user can then trace her finger along the phone's image and it will create differing clicks and bleeps as she moves along the edges of the objects and outside of it. It will be an audio language that reads shapes, essentially. The blind person could photograph a table or shelf and get an audio map of the objects on it. EdgeSonic also has less utilitarian uses, Kitani says. The user could snap a photo of a Christmas card and feel along the outline of a tree, "hearing" the parameters of the object.

Dias and her research group, TechBridgeWorld, are exploring ways to bring the power of smartphones and other mobile devices to populations underserved by technology—including the blind and visually impaired. Many are from developing countries, where almost 90 percent of visually impaired people reside, according to the World Health Organization. (The disparity is due to lack of access to medical care, particularly preventative care, according to the WHO.)

One of the sensors used by NavCog.

TechBridgeWorld's first project was developing computerized Braille tutors for the Mathru School for the Blind in India. The first version of the tutor connected with a computer, while the second was battery-powered and had its own on-board computing. Easy to use and transportable, the tutors have now been provided to organizations serving the visually impaired in six other countries.

More recent TechBridgeWorld projects include NavPal (not to be confused with NavCog), a smartphone app that provides navigational assistance to visually impaired adults as they move around unfamiliar indoor and outdoor environments, as well as Assistive Robots for Blind Travelers, which explores how robots may be able to help users with limited vision to safely move around a busy urban environment. The latter project is deploying a Baxter Research robot named "Rathu" (which means "red" in Sinhalese) that's been specially programmed to assist blind travelers. With its long, multi-jointed arms and complex sensors, Rathu can differentiate between objects such as bus passes and credit cards and hand them to the blind person. Dias foresees a time when robots such as Rathu will be able to trace a map of a room along a person's hand. She and her team are currently working to recognize users and incorporate past experiences into how it interacts with them. Dias says she "envision[s] assistive robots being available at key locations, such as transit locations, in future smart cities."

Assistive technology also will offer social ease to those with sight impairments, according to Aaron Steinfeld, an associate research professor at the Robotics Institute involved with TechBridgeWorld. People with disabilities are sometimes reluctant to ask for help, wanting to stay independent and avoid feeling like burdens. "There's less social resistance when asking a robot to repeat a gesture," Steinfeld says, "because robots are inherently patient and compliant." →

screenshot

REACHING OUT TO MAKE LIFE EASIER

Kevin (not his real name) was born missing part of his left arm below the elbow. An aspiring cellist, Kevin had a prosthetic that could hold a bow—but he was outgrowing it, making it difficult for him to play, and a new, professionally made prosthetic might cost $6,000 to $10,000.

Do-it-yourself technologies—especially 3D printers—promise to bring down the cost of simple prosthetics. And since 2013, there's been a network of volunteer "makers" called e-NABLE who collaborate to create prosthetics whose parts can be output to 3D printers.

Unfortunately, the e-NABLE makers are having a hard time keeping up with the demand.

Enter Jennifer Mankoff and Scott Hudson of CMU's Human-Computer Interaction Institute. They've teamed up with the founder of e-NABLE, Jon Schull of the Rochester Institute of Technology, to find ways to make customization, production and distribution of do-it-yourself assistive technologies more streamlined and effective.

Working together with other researchers, they've created a project called "Revolutionizing Assistive Devices via Innovation in Distributed Teamwork and Production." It will use crowdsourcing and distributed, online collaboration to create "virtual service teams" that can solve problems; work to encourage participation from clinical specialists; and adapt machine learning techniques to improve the design of assistive technologies.

In Kevin's case, making sure that the prosthetic was the right length was problematic. The team was able to rapidly test and iterate designs using Lego building bricks, then create a bow holder. (See inset photo.) Kevin proudly used his new prosthetic in a recent recital.

Mankoff says the project is about creating "real-world objects that serve real-world needs," but "without requiring the level of expertise that's currently required." That, she says, will make it easier "to make things that matter" and help people with disabilities.

"For now it's prosthetics, but I hope that eventually we can move beyond that," Mankoff says.

You can follow the team's progress, and read additional case studies, at makeabilities.org.

—Jason Togyer (DC 1996)

in the loop:

What were your earliest gaming experiences?

I started designing games when I was 8 years old, planning birthday parties for everyone in the class. So I went to the library and got a book called "101 Party Games for Kids." This book was my secret sauce. I would take the games in the book, and then modify them for each party.

Later on, I started babysitting and ran a youth group, and I designed games for those, but I didn't know it was a career path. I just thought it was just something you did when you were a babysitter.

Jessica Hammer

Jessica Hammer is an assistant professor at Carnegie Mellon University, with a joint appointment between the Human-Computer Interaction Institute and the Entertainment Technology Center. She is a graduate of Harvard, with a B.A. in computer science, and earned a master's of professional studies degree in interactive telecommunications at New York University and a Ph.D. in cognitive studies in education at Columbia University. A member of SIGCHI, the American Educational Research Association, the International Academy of Digital Arts and Sciences, the International Game Developers Association and Women in Technology International, Hammer joined the CMU faculty in 2014.

She spoke to Link Editor Jason Togyer.

How did you get into computer science?

My dad was a professor of computer science at MIT, so the two things I was convinced I was never going to do were become a professor and study computer science. Instead, I was going to be a poet. My father, who was very tolerant, said, "Well, you should do whatever you want to do in life, but you should take one computer science class. Computer science is a powerful tool that will help you to do your work better."

I almost failed the course, and that made me mad. I said, "No! Computer science is not going to beat me!" I went back and took the next course in the sequence, and at some point, I fell in love. I realized that solving problems and making things was really satisfying.

From there, did you go into game design?

No—I went to work for a nonprofit that was teaching STEM skills to girls. In the office next door, there was a guy named Scot Osterweil who was developing the Zoombinis games. I kept going over to peek at what they were doing until finally Scot said, "Would you like to just come and work for me one day a week?" All of a sudden, I was doing meaningful things with computer science and code that would bring people so much pleasure, and in the case of Zoombinis, help them learn mathematical skills. It was bringing back all of these feelings I had from childhood, and I felt like I'd hit the jackpot.

After that, I started work as a game designer, but it wasn't quite the same. Even though I was doing work I loved, I wasn't making progress on the questions I wanted to answer, such as—how can you use games to change people's lives?

What brought you to CMU?

The Human-Computer Interaction Institute was looking for someone who did game design, and when I saw what kind of work they were doing I said, "Yes, me!"

I actually have a joint appointment between the HCII and the Entertainment Technology Center. At the HCII, I run a research lab where I have Ph.D. students, while at the ETC, I work primarily with the master's students who are doing deep project work on designing games and other interactive experiences. It's a really rich opportunity to combine research into learning with research into the game design process, and it allows me to work on some projects that would be much harder to do in a conventional academic research lab.

What sort of research questions can you probe with a game?

You can make people do ridiculous things in games—jump up and down, croak like a frog, dance in front of a screen—because people participating in games have to take the context of the games seriously. The ability to do that is an amazing opportunity for research, because you can basically put people into whatever context you can imagine and then see what they do.

Right now, I'm working with an ETC team to develop a game that will change people's beliefs about what happens after a natural disaster. In popular fiction, including news media, portrayals of what happens after a major disaster depict life as Hobbesian—nasty, brutal and short. But that's basically a lie. Communities are very resilient after a disaster, and the majority of people are incredibly resourceful and capable. Natural disasters often bring out the best in people.

The problem is that these narratives are not just popular in fiction—they're narratives that policymakers use to make decisions. We think that we can make a game that can be played by young researchers and aides who put together policy proposals that are read by people who are more powerful. We are hoping we can make a game that has some power in the popular consciousness to change people's minds, and change disaster policy.

Do people really change their beliefs based on an experience they have in a game?

The experiences you have in a game don't disappear when you stop playing the game. Those experiences are vicarious to you in the same way that reading a novel is vicarious to you, but the experiences still matter. You still had the emotional reaction that you had. You can take advantage of this to design really interesting, powerful game experiences, so yes—people do.

Some politicians and public figures have been caught playing ultra-violent video games. Should we worry that their gaming activities are a sign of something they want to do in the off-line world?

The life someone has inside a game doesn't necessarily reflect what a person literally wants to do. Perhaps it's something that they need and aren't getting—but that doesn't have to be a literal need. When you tell me about a politician playing a super-violent game, I would say maybe what they need isn't violence. Maybe what they need is simplicity. If you have to compromise and negotiate all day in your daily life, maybe when you get home and play a game, you just want to solve some problems by blowing things up.

Why is playing computer games still held in low esteem by some people?

I think there are certain game activities that are more acceptable than others. The idea of a kid spending 20 to 30 hours a week playing sports is perfectly acceptable, even though it's a game. Digital games, though? It requires a certain amount of persuasion for people to see gaming as not just a ridiculous waste of time.

But think about a 14-year-old kid who's leading a guild of 200 players in "World of Warcraft," and organizing events where they're overcoming major challenges and everyone has to be coordinated. They're basically managing a small business.

What do you tell parents who are worried about the amount of time their kids are spending playing games?

I give them three pieces of advice. First, encourage your kids to play games socially. Playing games with other people is a great way to build relationships, to create opportunities to collaborate, to try on different identities and different roles. If you're going to limit their game time, limit their solo game time and encourage their social game time. Second, worry less about game content and more about their behavior. So they want to play "Call of Duty" and blow things up? That's less of a big deal than if they're bullying other people in the game.

Finally, encourage them to look for games that allow players to do modding or level-building—creative ways to contribute to the game. And that includes non-digital games. Computers are one technology used to play games, but so are dice and pens and paper. Get them playing games that are hackable, because at the end of the day, that will have them thinking analytically about why their design is better. If you're writing rules for games, that's not that much different than writing code—the only difference is that the processor you're writing for is the human brain.

Much more than the

sum of the parts

Modern artificial intelligence research integrates different approaches and disciplines—and so does a writer trying to understand the state of the field in the 21st century

By Linda K. Schmitmeyer
Illustrations by Jeffrey Katrencik (A 1985)

I first heard the Indian legend of "The Blind Men and the Elephant" in a junior high school English class, via John Godfrey Saxe's famous poem:

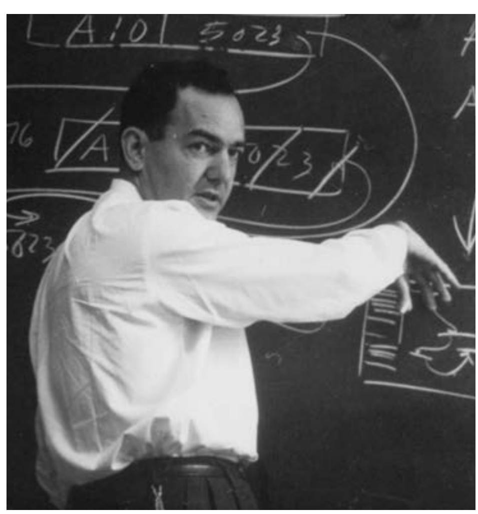
It was six men of Indostan
To learning much inclined,
Who went to see the Elephant
(Though all of them were blind),
That each by observation
Might satisfy his mind.

Each man encounters a different part of the elephant and imagines what it might look like, based on what he touches—the first, feeling the elephant's side, believes an elephant must be like a wall; the second, touching its tusk, thinks it like a spear. The third, grabbing the elephant's trunk, decides an elephant is like a snake, while the fourth, feeling one of the legs, thinks it like a tree. And so on.

They begin arguing about what an elephant is. Because each has felt only one part of the elephant, none really understands what an elephant is. As Saxe concludes his poem, each "was partly in the right, and all were in the wrong!"

I recently spoke about artificial intelligence research with seven different Carnegie Mellon University professors—Emma Brunskill, Eric Nyberg, Ariel Procaccia, Tuomas Sandholm, Aarti Singh, Manuela Veloso and Eric P. Xing. After each conversation, I felt like those blind men of Indostan; I walked away with a fragmented sense of the breadth and depth of the field of AI. Only when I was able to integrate what each of the faculty members had to say did I begin to develop an understanding.

In computer science terms, AI can be considered a very old discipline. The idea of a "thinking machine" that could synthesize concepts and ideas like a human being has fascinated scientists and futurists for generations. The field of "artificial intelligence" itself was formalized and named in the summer of 1956, when a group of researchers that included Carnegie Tech's Allen Newell (TPR 1957) and Herb Simon (H 1990) met for two months at Dartmouth College to talk about the work being done with machines that displayed intelligent behaviors. The four organizers of the Dartmouth Summer Research Project on Artificial Intelligence, John McCarthy (then at Dartmouth), MIT's Marvin Minsky, IBM's Nathaniel Rochester and Bell Labs' Claude Shannon, proposed "every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves." [continued]



Artificial intelligence pioneer and longtime CMU faculty member Herb Simon (H 1990)

In short, says Ariel Procaccia, an assistant professor in CMU's Computer Science Department, "the Dartmouth conference had a very broad vision of trying to understand how people think." Different approaches to simulating human intelligence emerged from the conference. Some researchers favored rule-based systems; Newell and Simon presented their version of a "thinking machine" called the Logic Theorist, which attempted to solve problems using the formal rules a human would use. McCarthy preferred a more abstract approach that came to be known as "circumscription," which excluded any variables that weren't explicitly known in an attempt to solve problems more efficiently.

Many approaches to AI were tried in the 1960s and 1970s, including early attempts to construct artificial neural networks. Much of the research focused on rule-based systems, in which programmers would imagine some state or condition, and then devise the rules that a computer would have to follow to move from that state or condition to another state or condition. "Rule-based approaches enabled the first use of computers to perform true symbolic reasoning," says Manuela Veloso, CMU's Herbert A. Simon University Professor of Computer Science. She calls them a "major contribution" to the field of AI; for the first time, she says, scientists were attempting to emulate human problemsolving in a structured, deliberate way. And as rule-based systems evolved, the research evolved as well. Soon, Veloso says, AI developers were focused on ways to most "efficiently search (the) enormous space of possible rules to find a sequence of actions that would transform some current state into a desired final state."

But there were limitations to rule-based approaches. In structured domains, such as playing chess, it was hard enough to write a rule for every possible situation; achieving artificial intelligence that could handle wider, more abstract problems through purely rule-based methods was nearly impossible. Other attempts at AI, such as artificial neural networks, were struggling as well. By the early 1990s, organizations that had been funding research into pure AI, such as the Defense Advanced Research Projects Agency, began questioning whether their investments would ever yield practical results. "It was like, 'Where are the success stories? What are the big AI systems that have been fielded?'" says Tuomas Sandholm, a professor in the Computer Science Department.

The advent of quicker processors, parallel processing and better methods for storing and retrieving massive quantities of information made it possible to use statistical methods to analyze gigabytes and terabytes of data to look for patterns, and then use algorithms to draw inferences and make predictions. The resulting field has become known as machine learning, and as a result, computer science "is not what it used to be, many years ago," says Aarti Singh, the A. Nico Habermann Associate Professor in CMU's Machine Learning Department.

Statistical machine learning as applied to big data—the catchall term that describes voluminous amounts of both structured and unstructured digitized information—has allowed researchers to make major breakthroughs in fields ranging from speech recognition, language translation

"The Dartmouth conference had a very broad vision of trying to understand how people think."
— *Ariel Procaccia*

> "Rule-based approaches enabled the first use of computers to perform true symbolic reasoning."
> *— Manuela Veloso*

and image processing to economics and computational biology. Singh says machine learning has become so prominent in computer science in part because it is making an impact in so many different areas. Her own research includes designing algorithms that, when confronted with a massive quantity of what she refers to as "big and dirty" data, balance the need for accurate statistical approximations with the desire for computational efficiency. The work has broad possible applications in physics, psychology, economics, epidemiology, medicine and analysis of social networks.

"Artificial intelligence research has a history of being very interdisciplinary, especially at CMU, and it is going to continue to be in the future," Singh says.

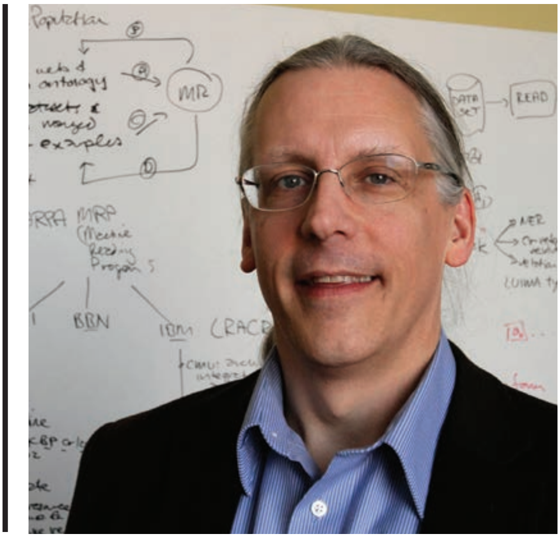
But modern methods of achieving artificial intelligence look "totally different" than they used to, says Sandholm, founder and director of CMU's Electronic Marketplaces Laboratory. In 2014, he led the team that won the championship at the Association for the Advancement of Artificial Intelligence's annual computer poker competition, and in 2015, his team's AI, now called "Claudico," took on four of the world's best human poker players in 80,000 hands of no-limit Texas Hold'em.

"Researchers today have a much more analytical approach, and AI has become more statistical and more rigorous, both theoretically and empirically," says Sandholm, who holds patents on many methods of optimizing online marketplaces such as auctions. "Instead of building a system and saying, 'Look, it kind of works,' there is real evaluation today on the empirical side, proving that things work."

Sandholm, for instance, has extensively deployed and demonstrated electronic marketplaces. In 1997, Sandholm founded CombineNet, which enabled major companies such as General Mills and Procter & Gamble to move their sourcing of supplies from manual processes and simple reverse auctions to highly sophisticated combinatorial auctions; CombineNet's sourcing platform, powered by AI, allowed purchasing agents to place bids on extremely diverse combinations of goods and services. CombineNet powered 800 auctions worth $60 billion before being acquired in 2010. Sandholm's newest startup, Optimized Markets Inc., is again applying sophisticated AI optimization techniques to change how advertising inventory is allocated and sold.

Although statistical machine learning has reshaped the study of artificial intelligence, not everyone was thrilled when those approaches first emerged, says Eric Nyberg, a professor in CMU's Language Technologies Institute, who recalls the reaction from what he calls "traditionalists" in the machine translation field when they were first confronted with statistical machine learning. "When statistical methods first came on the scene, there was a (negative) reaction because the traditionalists, who had been working on rule-based systems for years, got frustrated when they found out that by simply observing data, a model could automatically be trained much more quickly than having a human observe and then write the rules," Nyberg says.

He remembers attending a conference on machine translation in 1992 with several other CMU colleagues where, during one session, a debate ensued between rationalists and empiricists. The rationalists championed machine translation systems that used rules to describe a grammar that formally dictated correct English as it's taught to humans learning to read and write. The empiricists supported the idea of doing statistical analysis of texts, and creating rules based on probabilities that would allow computer systems to understand language at something close to human level, even if they weren't using formal rules of grammar.



"At the conference, the people from Carnegie Mellon decided to fly over this controversy and say, 'You're missing the point. The point is that we have to combine these techniques because they have complementary strengths,'" Nyberg says. Rule-based systems "are really good at helping encode exceptions," he says, but statistical methods "are good at teasing out the general rules of a domain."

"They don't work equally well on all cases, and until you learn what your domain is all about, you can't presuppose which approach will be better," he says.

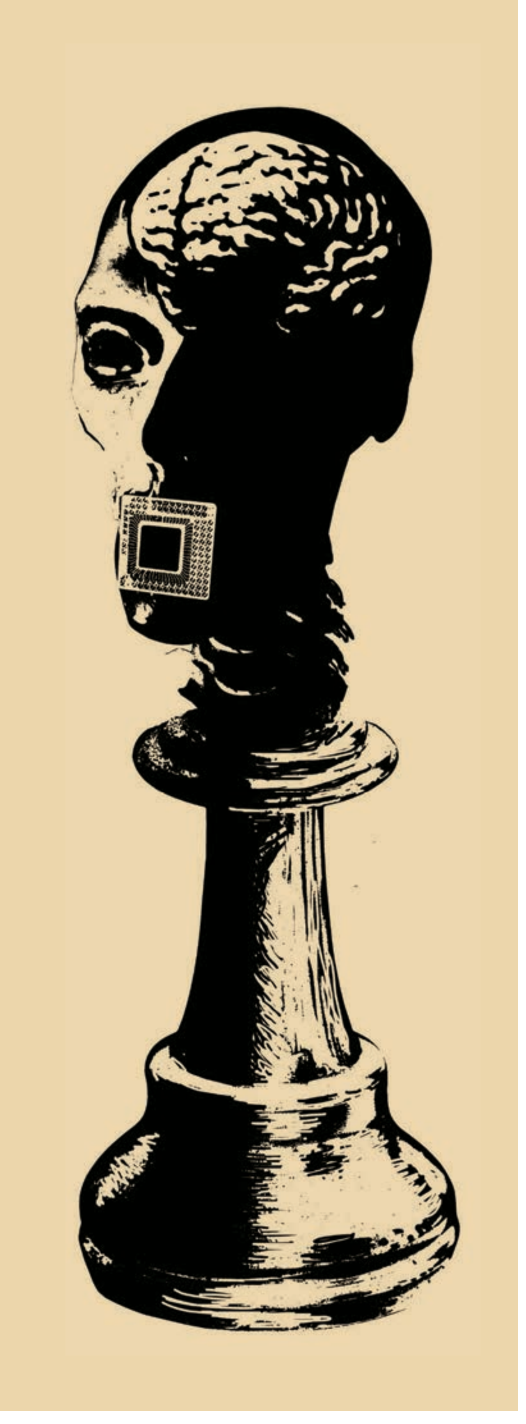
In the early 1990s, Nyberg says, CMU was a pioneer of Multi-Engine Machine Translation, or MEMT, a method for processing data that uses multiple machine translation approaches, including both statistical and rule-based methods, within a single machine-translation system. This multi-strategy approach has since been leveraged by other language systems, most notably in IBM's Watson question-answering system, which attracted international attention when it defeated two human champions on the TV game show "Jeopardy!" in 2011. Nyberg's own research involves the open advancement of question-answering, or QA, systems; he and other CMU researchers collaborated with IBM and other universities to create the algorithms that powered Watson to its "Jeopardy!" victory.

Nyberg predicts that, going forward, artificial intelligence researchers will more and more be blending the two different approaches. "You can have a rule-based component that does a task and a statistically trained component doing the same task and constantly monitor how they're doing," he says. "We can combine the components on a task until such time when we realize one has surpassed the other."

Over and over again, CMU faculty engaged in artificial intelligence research told me about the importance of blending different approaches. Sandholm says the "silos of learning" inherent in academia have to come down everywhere, but particularly in artificial intelligence. In the 2004 edition of her classic book, "Machines Who

"Researchers today have a much more analytical approach, and AI has become more statistical and more rigorous, both theoretically and empirically. Instead of building a system and saying, 'Look, it kind of works,' there is real evaluation today on the empirical side, proving that things work."
— Tuomas Sandholm

"When statistical methods first came on the scene, there was a (negative) reaction because the traditionalists, who had been working on rule-based systems for years, got frustrated when they found out that by simply observing data, a model could automatically be trained much more quickly than having a human observe and then write the rules."
— Eric Nyberg

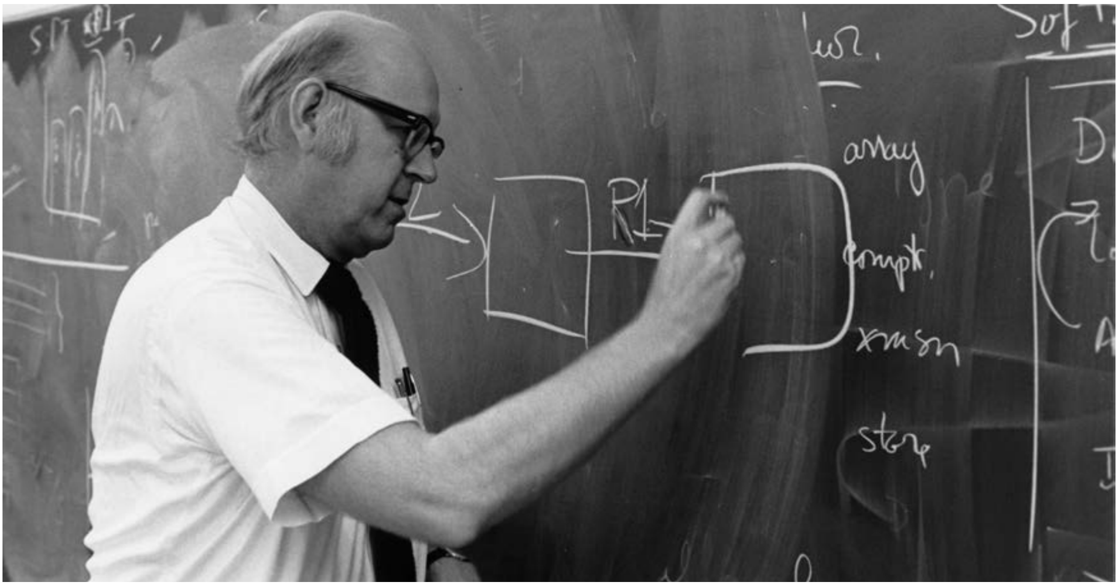


Think," Pamela McCorduck laments that AI researchers have often broken into "subfields" such as "vision, natural language, decision theory, genetic algorithms, robotics" that "hardly have anything to say to each other."

"Cross-fertilization of ideas is important within AI research, but also between AI and other disciplines," Sandholm says. "From AI to operations research. From AI to economics. From AI to astrophysics. I tell my students never to write in a paper, 'a computer scientist says this, but an economist says that.' You should never think like that."

Newell and Simon embodied that interdisciplinary approach. Unlike many early computer scientists, neither one came to artificial intelligence research directly from electrical engineering or mathematics; Simon was an economist and political scientist, while Newell was studying group dynamics and decision-making.

Manuela Veloso (CS 1989, 1992), who today holds the university professorship named in Simon's honor, came to CMU as a student in 1986, when both Simon and Newell were still active faculty members. At a time when there was a lot of excitement about different divergent projects in artificial intelligence, "such as chess, and tools like neural networks in planning and problem-solving techniques, and the different architectures for cognition, and trying to figure out how machines learn," Veloso says Newell was already arguing that these different research areas had to be integrated. "I was there, in 5409 Wean Hall, when Allen Newell went to the blackboard and picked up the chalk and wrote perception, cognition and action and drew boxes around the words. Underneath, he wrote the word agents. He said that the field of AI was fragmented and that it was time to put these things together." Veloso, the immediate



Allen Newell (TPR 1957) argued that A.I. research was too fragmented and that "it was time to put these things together," says Manuela Veloso (CS 1989, 1992). Her soccer robots are examples of A.I.'s that can act autonomously and collaborate on tasks.

past president of the Association for the Advancement of Artificial Intelligence, said she personally took his comments "as the research direction for my life."

In the early 1990s, Veloso founded the research lab called CORAL—for Collaborate, Reason, Act and Learn—which she still directs. With her students, she created robots that could play soccer together as a team, performing coordinated activities against other teams of robots. The first "RoboCup" soccer games were held in 1997 at the 15th International Joint Conference on Artificial Intelligence in Nagoya, Japan, with 10 teams competing in the real robot league, and 29 teams competing in computer simulations. Teams from CMU, led by Veloso, have competed in every "RoboCup" since. Her teams—competing in the "Small-Size League"—have won the competition five times and placed second four times; in 2015, they won all of the games in which they competed, finishing with a combined score of 48-0. "The robot soccer teams are a remarkable example of artificial intelligence, robotics and some machine learning," Veloso says.

Roaming the halls of Carnegie Mellon's Gates and Hillman Centers today are additional manifestations of that integration of AI research that Newell called for a quarter century ago. They are the CoBots, created by Veloso and her students, which deliver packages, lead visitors to various offices for scheduled meetings, and, when necessary, seek assistance from humans to push elevator buttons so they can carry out their tasks. To date, the CoBots have logged more than 1,000 kilometers navigating the GHC buildings.

They would never have reached that milestone—would not be able to demonstrate any artificial intelligence—without integration of their various capabilities. Veloso's vision algorithms, for instance, must be connected to task-planning algorithms that in turn control actuators. For the CoBots to be successful, Veloso says, she realized that they would have to be aware of their own perceptual, physical and reasoning limitations. "We introduced the concept of symbiotic autonomy, in which the robots proactively ask for help from humans, or resort to searching the Web when they realize they lack the capabilities or understanding to perform parts of their service tasks," Veloso says. The research and development of the robot soccer teams, of the CoBot robots and of other robots in the CORAL lab requires integration of robotics, engineering and many other disciplines. "It's a science itself, of determining how we are going to make all these (systems) work together," she says.

Real-world successes, such as combinatorial sourcing auctions and the CoBots, have validated the integrative approach to artificial intelligence. Another of the real-world success stories in AI is the Kidney Paired Donation Exchange, which is powered by algorithms and software developed in Sandholm's research group. There are two types of kidney transplants—a kidney swap, where a donor gives a kidney directly to someone (often a blood relative) who is medically compatible, or the much more complicated kidney chain, where a donor gives a kidney to a patient who they don't know. Most kidney transplants worldwide now happen through such chains. Doing the work manually is a "very tough task," Sandholm says.

The kidney exchange, launched in 2010, medically matches donors who are incompatible with their intended recipients to others in the network who need a transplant. AI algorithms make the transplantation plan autonomously for the entire United States twice a week. Sandholm and Procaccia are collaborating on the kidney exchange research, along with CMU computer science professor Avrim Blum, who developed an early version of the fielded algorithm.

"AI is doing a much better job than any human could possibly do," Sandholm says. "It's not because the doctors are dumb. It's just that they're sifting through more alternatives than there are atoms in the universe, and that's not something humans can do." But computers can.

Sandholm predicts that within the next five to 10 years, similar artificial intelligence techniques will be helping humans make smarter choices in many other critical areas, both in high-level planning and low-level decision-making. That's because of the progress in AI algorithms over the past decade, as well as in the amount of data now available, he says.

Yet there are concerns that with their inherent lack of structure, current statistics-based methods are hitting their own limitations, just as purely rule-based methods did in the past. "A large fraction of the contemporary machine learning systems in action is using very simple logical or mathematical rules as of now," says Eric Xing, professor of machine learning, language technologies and computer science and director of CMU's Center for Machine Learning and Health. Those rules, he says, are mostly propositional rules—relatively simple inferences such as "if x, then y"—or measurements of dispersion of data. Future approaches to statistical analysis of big data will increasingly apply higher-order logic, Xing predicts, such as "relational or probabilistic rules," that enable the resulting intelligence mechanisms to interface with data in ways that are both "stochastic and elastic."

"The marriage of the two approaches will create richer models that are more likely to work in the real world," Xing says.

Today's hardware designs also affect the capacity to learn from big data. Because of the heavy computational power required to take advantage of the data, operating systems need to evolve. "In the past, I don't feel that either the machine learning or artificial intelligence fields have been paying strong attention to the hardware and to the operating systems," Xing says. "And vice versa. The hardware and system people don't pay enough attention to the other side."

He says the merging of hardware engineering and operating system design "is imminent."

One of the next-generation platforms that brings machine-learning principles and artificial intelligence needs directly into the design of both an operating system and its hardware is called Petuum. Developed by a research group within Xing's CMU lab, called "Statistical Artificial InteLligence & INtegrative Genomics," or "SAILING" for short, Petuum is designed specifically for machine learning algorithms. It provides essential distributed programming tools to tackle the challenges of running machine-learning algorithms at large scale.

> As the field evolves, one aspect of artificial intelligence on which all seven Carnegie Mellon professors concur is that it will reach into almost every aspect of human existence. Another thing that became clear from my conversations is that from the days of Newell and Simon's participation in the Dartmouth conference, CMU has played a key role—perhaps the most integral role—in moving artificial intelligence research forward.

"The Petuum system ties together a cluster of machines with communication procedures, scheduling procedures and resource allocation and management procedures so that they turn up as a single machine interface to the user," Xing says. One of those users is the recently established Pittsburgh Health Data Alliance, a joint effort between the Center for Machine Learning and Health, the University of Pittsburgh and Pittsburgh's UPMC health care system. The Health Data Alliance collects massive amounts of data from sources as varied as electronic health records, genomic profiles and wearable devices. But instead of passively recording that data for humans to sort through later, the Health Data Alliance is using Petuum to analyze the data in real-time, find noteworthy patterns, and generate notifications or warnings, if necessary.

Brunskill doesn't dismiss the possibility of a computer some day achieving human-level intelligence, but she says there are aspects of the human brain that remain elusive.

"We want to use our strength in artificial intelligence and computing to amplify the value of that data," Xing says. The Center for Machine Learning and Health is also planning to develop new technologies—including a series of increasingly data-driven apps—that will change the way diseases are prevented and patients are diagnosed and treated.

When the Pittsburgh Health Data Alliance was announced in early 2015, CMU President Subra Suresh predicted it would help caregivers and patients to "move from reactive care to immediate, proactive prevention and remediation, from experience-based medicine to evidence-based medicine, and to augmenting disease-centered models with patient-centered models." Xing has a catchier phrase for the effort. By mining UPMC's health data, he says, the alliance hopes to "'smart-ify' the entire health care system."

But are we any closer to developing a computer with human-level intelligence? At the Dartmouth conference in 1956, the organizers proposed that a "significant advance" in simulating human thought could be made as the result of their summer-long brainstorming session. More than 50 years later, we don't seem yet to have created a machine that can demonstrate both human-level intelligence as well as creativity.

"We have systems that can mimic how we think human intelligence works, but how humans can spontaneously arise with ideas of identity, of self-will, of control, of consciousness, to me those concepts are very, very new," says Emma Brunskill, an assistant professor in CMU's Computer Science Department who is also affiliated with the Machine Learning Department. How a machine could replicate those attributes "is incredibly unclear at this point," she says.

Advances in both processors and storage have allowed researchers to develop artificial neural networks that do simulate the way human brains engage in activities such as solving problems and processing language, Brunskill says. "With the increased focus on data and neural networks, we're starting to tackle how agents and autonomous systems learn representations of the world from scratch," she says. The challenge is in developing systems that take in raw

"In the past, I don't feel that either the machine learning or artificial intelligence fields have been paying strong attention to the hardware and to the operating systems. And vice versa. The hardware and system people don't pay enough attention to the other side."

— Eric Xing

data of many different kinds, build up a representation of what the data means, and then act on that representation in abstract ways—not just seeing, for example, a phone on a table, but understanding one of the concepts represented by a phone on a table: communication.

"Those types of systems are really exciting because they are able to go directly from sensory input of the world to real decision-making," Brunskill says.

One of Brunskill's research interests is developing online tutoring programs that continually improve their teaching methods as they interact with students. A major focus of concern is making sure it doesn't take too long for those systems to self-optimize. "We don't want to teach a million students before we figure out what pedagogical activities are most effective for teaching fractions," Brunskill says.

In an ongoing research project, Brunskill and her colleagues have developed an intelligent tutoring system for teaching students how to use histograms—graphical representations of statistical data. Many students, Brunskill says, have trouble understanding histograms, even after taking a statistics class. The tutor automatically found a strategy for teaching histograms by iteratively changing how it instructed students. When the tutor was evaluated, the researchers found that within the first 30 students who used it, the quality of instruction was comparable in results to a good, hand-designed tutorial.

Brunskill doesn't dismiss the possibility of a computer some day achieving human-level intelligence, but she says there are aspects of the human brain that remain elusive. It may be possible that consciousness will evolve from a system once it reaches a certain level of mathematical or computational sophistication, she acknowledges, but that would raise another question. "How would we test if something was conscious?" Brunskill says.

In a 1950 article called "Computing Machinery and Intelligence," Alan Turing suggested the hypothesis that we now call the "Turing test"—the idea that a computer can be considered "intelligent" if human users can conduct a conversation with the computer and not realize they're talking to a machine. If computers aren't yet there, question-answering systems—such as those being developed by Nyberg and his colleagues—are getting close.

That, in turn, has sparked fears of a computer uprising. "I, for one, welcome our new computer overlords," joked Ken Jennings, one of the two human "Jeopardy!" champions defeated by Watson.

The pop-culture trope of rogue computers enslaving humans grew out of the classic definitions of artificial intelligence, which envisioned AIs as autonomous agents that didn't need human input. But Nyberg says Watson actually represents a different approach, one that IBM calls "cognitive computing"—humans and machines collaborating to accomplish tasks better than either humans or machines can do on their own.

A team of CMU researchers led by Ariel Procaccia is exploring one such collaboration. They've applied artificial intelligence to social choice and game theory to help human beings allocate resources and make collective decisions, such as who gets a favorite family heirloom when a loved one dies, or how much an individual owes when sharing a cab with friends.

“Artificial intelligence research has a history of being very interdisciplinary, especially at CMU, and it is going to continue to be in the future.”
— Aarti Singh

Understanding the “fair division” problem, Procaccia says, has until recently been the domain of economists and political scientists, but not so much computer scientists. Yet artificial intelligence is uniquely suited to providing unbiased advice, unaffected by emotion. Once a group of individuals faced with a decision establish what “fairness” means to each of them—whether it’s a social good or a personal goal—“then we can feed that into a computer and develop algorithms that achieve those notions of fairness,” Procaccia says. In 2014, he and other researchers launched the website Spliddit.org, a not-for-profit academic endeavor that utilizes economics, mathematics and AI research to provide people with provably fair methods to resolve everyday dilemmas, such as how to split rent, divide tasks or apportion credit for a project.

They’re now applying their research to more complex scenarios, says Procaccia, who’s working with California’s largest public school system, the Los Angeles Unified School District, to develop a computer program that will make suggestions for the fair allocation of students in classrooms in the district’s 241 charter schools.

“I want to leverage economic theory to develop computer programs that make intelligent suggestions for interactions between multiple people,” he says.

Unlike the blind men of Indostan, who remained steadfast in their conclusions about the nature of an elephant, I came away from my conversations with all seven Carnegie Mellon professors with a better appreciation of the breadth and depth of artificial intelligence—and the magnitude of its impact on our lives.

As the field evolves, one aspect of artificial intelligence on which all seven Carnegie Mellon professors concur is that it will reach into almost every aspect of human existence. “It’s already pretty broadly in our lives today,” says Sandholm, “but in the next 10 to 20 years, it’s going to be running almost everything.”

Another thing that became clear from my conversations is that from the days of Newell and Simon’s participation in the Dartmouth conference, CMU has played a key role—perhaps the most integral role—in moving artificial intelligence research forward.

In October 2000, at CMU’s Earthware Symposium, Herb Simon ranked the computer as one of the top three human inventions, along with language and organization. “It has already become, and will continue to become increasingly, a constant companion and partner of the human mind,” Simon said, arguing that the task of computer scientists is “to design a future for a sustainable and acceptable world, and then devote our efforts to bringing that future about.”

Carnegie Mellon today is arguably living out Simon’s vision. “CMU is not just one of the players in AI research,” Xing says. “It is unique and has the advantageous position of leading the research.”

—Linda K. Schmitmeyer is a freelance writer and editor and teaches non-fiction writing at Pittsburgh’s Point Park University. Link Editor Jason Togyer contributed to this story.

Growing in amazing ways

Ashley Patton
awpatton@andrew.cmu.edu

Niccole Atwell
atwell@andrew.cmu.edu

It's an exciting time to be a part of the School of Computer Science.

We're growing in amazing ways. We're about to launch what SCS Dean Andrew Moore calls the Moonshots Initiative: big, ambitious projects that attempt to change the world in dramatic, positive ways. We're working on a partnership in Pittsburgh to expand the number of high school students who are interested in computer science.

And even beyond our campuses, the need for computer science education is gaining prominence in the national spotlight, with President Obama's announcement in late January of the new White House initiative called Computer Science for All.

To support the exciting initiatives happening in SCS, we've launched a new, joint Office of Engagement and Annual Giving in partnership with the Department of Electrical and Computer Engineering in CIT. The goal of this office is to identify and implement new ways for our SCS and ECE alumni to connect to the work being done on campus.

We've already begun to roll out new programs, including the launch of our new alumni mentoring platform at cmu.firsthand.co. Many of you have told us that the way you're most interested in giving back is to support current students, and we want to make it easy for you to volunteer in this extremely valuable way. The platform allows you to connect directly to students electronically and on your own schedule, while still maintaining the privacy of your information. If you're interested in learning more, or becoming a mentor yourself, please visit cmu.firsthand.co.

Our next major project is an overhaul of the alumni page on the SCS website. We're hoping to make it easier for you to find the information you need about events, volunteer opportunities and the ways you can make a direct philanthropic impact on the college projects that matter most to you.

So many of you support us with gifts of both time and resources, and we're incredibly grateful for that—you make much of the good work that's happening here possible! In addition to launching new programs and tools, we're also rethinking how we say "thank you" to those of you who have supported us in many different ways. We want to make sure you know how important you are, so we're working to provide more detailed information about the impact your gifts have made.

If you have questions, suggestions or comments about what's going on at SCS, we'd love to hear from you! Feel free to reach out to either my colleague Niccole Atwell or me—our contact information is at left.

Ashley Patton
Director of Engagement and Annual Giving
School of Computer Science
and Department of Electrical and Computer Science

Jennifer Cerully

B.S., computer science, Carnegie Mellon University, 2004
M.S., psychology, University of Pittsburgh, 2008
Ph.D., psychology, University of Pittsburgh, 2011

For those of us who grew up during the Cold War, the RAND Corporation might conjure up classic movie stereotypes of sterile, well-lit corridors lined with giant mainframe computers. "A lot of people ask me about the movie 'Dr. Strangelove,'" says Jennifer Cerully (CS 2004), a behavioral and social scientist at RAND's Pittsburgh office, located on Craig Street, next to CMU's Software Engineering Institute.

But she hadn't seen the movie until after she began working at RAND—and what Cerully has found at RAND isn't so strange. In fact, it's very human-focused.

"There's real potential here for your work to make an impact," she says. "We do research and policy analysis, in whatever form that takes. We work in many different sectors—health, education, defense, labor, population studies—and our independence and objectivity is very much valued. So when I'm doing research, I'm never pressured to make the data tell a certain story. That was one factor that was very important for me before coming to RAND."

Cerully is currently studying mental health care as it's provided to Americans in the armed forces, and looking for areas where the Department of Defense can improve the delivery and timeliness of those services. "There's a lot of stigma surrounding mental illness," she says. "We've come in to evaluate programs and mental health awareness campaigns the department has in place, and make recommendations for reducing stigma and other barriers to care."

"I'm really hoping my work will make it a little easier for military service members to get good quality mental health care if they need it," Cerully says.

It's not the path on which Cerully expected to find herself when—inspired by high school programs designed to introduce women to science and technology careers—she arrived at CMU as a computer science undergraduate. "I still remember the privilege in my freshman immigration course of hearing Herb Simon speak," she says. But by the end of her degree, Cerully says, "I found I didn't really like programming, so the idea of doing that as a career wasn't very appealing to me." Working in information technology wasn't attractive, either.

What she did enjoy was studying how humans interact with one another—and that led her, while at CMU, to earn a double major in social and decision sciences, working with CMU professor Jennifer Lerner, who's now at Harvard's Kennedy School of Government. Upon graduation, Cerully moved across Panther Hollow to the University of Pittsburgh, where she earned her master's degree in psychology, and then her Ph.D. in 2011. She joined RAND the same year.

"I thought I was making a complete split from computer science by moving into social science, but I've found that my computer science degree helps me a lot," Cerully says. "Having learned very early on how to code helps me in very specific ways—doing syntax-based analysis, for instance—and where I really started to see the difference was in graduate school. I had a different approach to problem-solving than some of my peers, and it was because I'd been taught to think in a different way."

And computer science intersects her work in many areas. For example, on one project, she's working with a doctoral student at Pardee RAND Graduate School to use machine learning to examine the discourse about mental health and mental illness via discussions that take place on Twitter.

A native of Altoona in central Pennsylvania, Cerully says she didn't expect to stay in Pittsburgh forever. But now she's married to a fellow CMU grad, Jonathan Chu (CS 2004), who works at the SEI (they met outside of Wean 7500) and says, together with their daughter, they're a "real CMU family." She still enjoys coming back to campus to talk to students.

"In some ways, I feel like I struggled as an undergraduate about what direction I was going to go," Cerully says. "That's actually very common. People take all kinds of paths, and I love to go back to CMU and talk to students about that. You can kind of see the relief on their faces when they realize it's not strange at all to be wrestling with those issues." —*Jason Togyer (DC 1996)*

Brendan Meeder

B.S., computer science, Carnegie Mellon University, 2007
Ph.D., computer science, Carnegie Mellon University, 2015

Andrew's Leap was "an eye-opening experience," says Brendan Meeder (CS 2007, 2015) of the long-running summer enrichment program for middle-school and high-school students which was recently renamed "Leap@CMU."

Until he'd participated in Andrew's Leap as a high school student, Meeder's computer science and robotics experience consisted of "casually programming websites" and making simple games. But in Andrew's Leap, Meeder says, he got a "multifaceted" overview of computer science as well as its connections to the physical world. "Steven Rudich does a great job teaching," Meeder says, "and on the robotics side, Matt Mason got me really excited about that field."

After his experience with Leap, there was little doubt that Meeder wanted to go into computer science and attend Carnegie Mellon. In fact, he enjoyed the experience so much, that after two years with Leap, he came back the following year to serve as a teaching assistant in the program; and the summer before his first undergraduate year at CMU, Meeder worked in Mason's manipulation lab.

Young people today have a wealth of programming tools and robotics toys to choose from that didn't exist a few years ago, he says. That—along with programs such as Leap@CMU—exposes them to the usefulness of computer science much earlier than that used to happen, Meeder says.

"It used to be that things were so hard, and all you could do was make simple 'guess-a-number' games in Visual Basic, or program a graphic calculator, and it was like, 'Why should I bother?'" he says. "Now, the frameworks are there so that you can start building interesting things much faster. Yes, the barrier to entry is lower, and yes, the satisfaction of learning to build utilities is going down, but when you're doing something now, in the physical world with something like an Arduino, it's so much more rewarding."

Meeder completed his bachelor's degree in computer science in 2007, then left to work for Microsoft Research in Redmond, Wash., on speech recognition devices. But his heart was back in Pittsburgh, and he returned to CMU after 14 months to begin his Ph.D. under CMU's Luis von Ahn and Manuel Blum. "One of the courses that Luis was teaching at the time was on the mathematical modeling of the Internet and social networks, and I was really attracted to that because it included theoretical math—which I love—plus, you could derive a lot of understanding of the Internet from it," Meeder says, "and that became my thesis topic."

In 2011, Meeder was recruited by von Ahn to be one of the first employees to work for Duolingo. A language-learning app developed at CMU by von Ahn (CS 2003, 2005) and Severin Hacker (CS 2009, 2014), Duolingo now provides free online language education in 23 different languages to 100 million registered users. Meeder is one of those users; when we talked, he'd logged time with Duolingo for 840 consecutive days.

Although Duolingo is still growing, Meeder felt the urge to keep "learning as fast as I can," and in April 2015, he joined Uber's growing research center in Pittsburgh, where he works with many CMU alumni and former employees on vehicle technology, including autonomy, mapping and safety. "It's a magical place," Meeder says. "There's a lot of variety in the kind of research that needs to be done and the kind of software that needs to be written. It's almost like a symphony, where everyone has a role."

Meeder has had his own role in a real symphony; until recently, he played bass trombone in CMU's All-University Orchestra. In his spare time, he and his wife, fellow CMU alum Ariel Levavi (S 2007), enjoy cooking, relaxing and walking around Pittsburgh's Squirrel Hill neighborhood.

"It's really good to still be in Pittsburgh," Meeder says. "With the number of companies like Uber and Google growing their research presence here, it's really attractive to stick around after graduating."

—Jason Togyer (DC 1996)

The next 50 years of data

Carnegie Mellon University
Computer Science Department
Fiftieth Anniversary

By Andy Pavlo

Editor's Note: This article is adapted from an article that originally appeared in the commemorative book printed in 2015 to celebrate the 50th anniversary of the Carnegie Mellon University Computer Science Department. Additional portions are adapted from the author's blog. We are grateful to Andy Pavlo for allowing us to reprint these portions here.

The first database management system, or DBMS, came online in 1968. IBM's Information Management System was built to keep track of the supplies and parts inventory for the Saturn V and Apollo space exploration projects. It introduced the idea that an application's code should be separate from the data that it operates on. This allows developers to write applications that only focus on the access and manipulation of data, and not the complications and overhead associated with performing these operations and ensuring that data is safe.

IMS was later followed by the pioneering work in the early 1970s on the first relational DBMSs, IBM's System R and the University of California's INGRES.

The database workloads for these first systems were less complex and diverse than they are today. In these earlier applications, a human operator started a transaction through a terminal and then entered new data into the system manually. Back then, the expected peak throughput of a DBMS was only from tens to hundreds of transactions per second, and the response times were measured in seconds.

The architecture of these early DBMSs was based on the computing hardware that was prevalent at the time they were invented. They were typically deployed on computers that had a single CPU core and a small amount of main memory. For these systems, disks were the primary storage location of databases because they were able to store more data than could fit in memory and were less expensive.

The Present

Although a lot has changed 50 years later in terms of how we use databases, the relational model and Structured Query Language are still the predominant ways to organize a database and interact with it. Many Internet applications need to support hundreds of thousands or even millions of transactions per second, and with each transaction's processing, latency must be kept on the order of milliseconds. That's because the applications are connected to millions of users and other computer systems at the same time.

And now that businesses and organization are able to collect a large amount of data from these applications, they want to analyze it to extrapolate new information to guide their decision-making. For this reason, in recent years we have seen the rise of specialized systems that target specific application scenarios that are able to perform much better than general purpose DBMSs based on 1970s architectures. There are now DBMSs that are designed to ingest new information quickly for online transaction processing, or OLTP, applications and other DBMSs that are designed to store large amount of data for complex online analytical processing, or OLAP, programs.

These newer DBMSs also take advantage of the three major hardware trends that have emerged in recent years.

Large-Memory Computers: The first is the advent of large-memory computers, which make it now affordable to deploy a small number of machines that have enough dynamic random-access memory (DRAM) to store all but the largest OLTP databases. Storing data in memory ensures that the DBMS can process many transactions simultaneously with low latencies. In our experience, the databases for modern OLTP applications are typically several hundred gigabytes in size. Contrast this with an OLAP data warehouse, where the DBMS could be managing databases that are several petabytes in size. While an OLTP database bases

stores the current state of an application (for example, orders from the last 90 days), an OLAP database stores all of the historical information for an organization (for example, all of the orders that have ever been placed). Thus, OLAP DBMSs are still primarily stored on disks and employ several optimizations like compression or columnar storage to overcome their slower access times.

Multi-Core CPUs: The second hardware trend is the shift away from increasing the clock speeds of single-core CPUs and toward multi-core CPUs. Clock frequencies have increased for decades, but now the growth has stopped because of hard power constraints and complexity issues. Aggressive, out-of-order, super-scalar processors are being replaced with simple, in-order, single-issue cores. Exploiting this increased parallelism is difficult in a DBMS because of the complexity of coordinating competing accesses to shared data for hundreds of threads. Modern DBMSs are employing low-overhead concurrency control and other lock-free techniques to improve the scalability of the system.

Low-Cost Commodity Hardware: The third trend is the lower cost for commodity hardware. This is especially pronounced in cloud computing platforms. It is now possible to deploy a large-scale cluster with a large processing and storage capacity for a fraction of what it was ten years ago.

Despite these advancements, there are still significant barriers that impede the deployment of data-intensive applications. One overarching theme is that databases are still a human-intensive component of computing systems (e.g., deployment, configuration and administration). Using two independent DBMSs separates the OLTP and OLAP workloads to avoid one slowing down the other, but it requires additional processes to transfer data from one system to the other.

In addition, tuning a DBMS to get the best performance for a particular application is notoriously difficult. Many organizations resort to hiring experts to configure the system for the expected workload. But as databases grow in both size and complexity, optimizing a DBMS to meet the needs of these applications has surpassed the abilities of humans.

Main Memory Database Systems

Way back in 2007, the experimental database management system H-Store, developed by a team that included researchers from CMU, Brown, Massachusetts Institute of Technology and Yale, was the vanguard for a new era of main memory-oriented database systems. H-Store was a forerunner of VoltDB, an in-memory database that is now widely used in many mission-critical applications.

But it wasn't obvious right away to some that a memory-oriented architecture was the way to go for scalable OLTP applications. In the early days of VoltDB, customers were uncomfortable with the idea of storing your entire database in volatile DRAM. I visited PayPal with VoltDB co-founder and database researcher Michael Stonebraker in 2009 when he went to talk to them about VoltDB, and I remember that the senior management was unnerved by the idea of a DBMS that did not store all physical changes to tuples immediately on a disk.

The prevailing conventional wisdom has obviously changed. Since then, several other memory-oriented systems are now available, including MemSQL, SAP HANA and Microsoft Hekaton. Other notable in-memory research systems that came along after H-Store include Shore-MT, HyPer and Silo. In particular, the HyPer team has amassed an impressive corpus of research publications.

> The first database management system, or DBMS, came online in 1968. IBM's Information Management System was built to keep track of the supplies and parts inventory for the Saturn V and Apollo space exploration projects. It introduced the idea that an application's code should be separate from the data that it operates on.

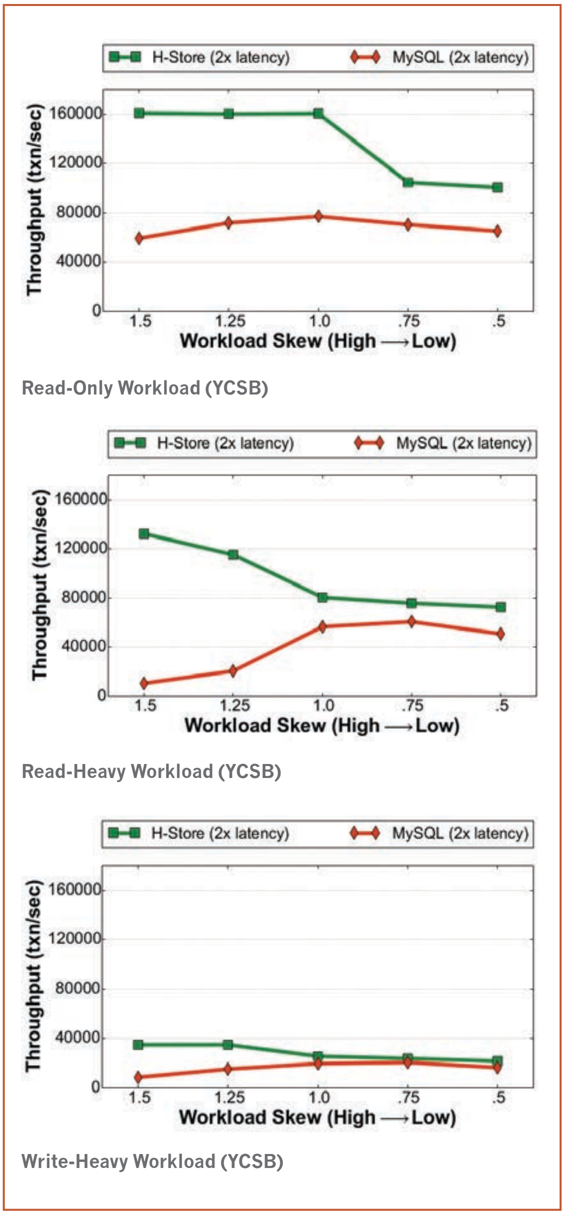


Figure 1: H-Store and MySQL running on a NVM+DRAM system with Intel Lab's hardware emulator.

Main memory DBMS research is mostly a solved problem from a research point-of-view. A more interesting topic is how the advent of non-volatile memory, or NVM, such as flash memory, will overturn the traditional storage hierarchy in computing systems. DBMSs are uniquely positioned to utilize this technology for a wide variety of application domains.

NVM+DRAM Systems

Based on my discussions with hardware vendors, database management systems that deploy block-addressable, non-volatile memory on PCI-e cards are less than five years away. I refer to this as an NVM+DRAM storage hierarchy. Memory-oriented DBMSs will still be the best performing architecture for this hierarchy because they do not use legacy architectural components from the 1970s (e.g., heavyweight concurrency control) that are designed to mask the latency of slow disks.

Since a DBMS on the NVM+DRAM hierarchy still uses DRAM for ephemeral storage, it will need to flush out changes to stable storage for recovery and durability. This logging will be the major bottleneck for all DBMSs, even if they have fast NVM.

At the Intel Science & Technology Center for Big Data, based at MIT's Computer Science and Artificial Intelligence Laboratory, we did some initial experiments testing MySQL and H-Store (with anti-caching) on Intel Lab's NVM emulator. The emulator was configured so that all reads/writes to NVM were approximately 180 nanoseconds, compared to a 90-nanosecond read/write to DRAM. We benchmarked our tests using the Yahoo! Cloud Serving Benchmark released in 2010.

We used three variants of the YCSB workload: (1) read-only with 100 percent reads, (2) read-heavy with 90 percent reads plus 10 percent updates, and (3) write-heavy with 50 percent reads plus 50 percent updates.

Our experiments (see Figure 1) showed that there was a significant decrease in throughput as the number of update transactions in the workload increased. H-Store exhibited an approximately 77 percent drop in its peak performance for the read-only workload compared to its best performance in the write-heavy workload. For MySQL, this difference was nearly 75 percent. This is due to the overhead of preparing and writing the log records out to durable storage to overcome DRAM's volatility.

In Figure 1, we also see that for the read-heavy workload, H-Store achieves 13 times better throughput over MySQL when skew is high, but only a 1.3 times improvement when skew is low. This is because H-Store performs best when there is high skew since it needs to fetch fewer blocks from the NVM anti-cache and restart fewer transactions. In contrast, the disk-oriented system

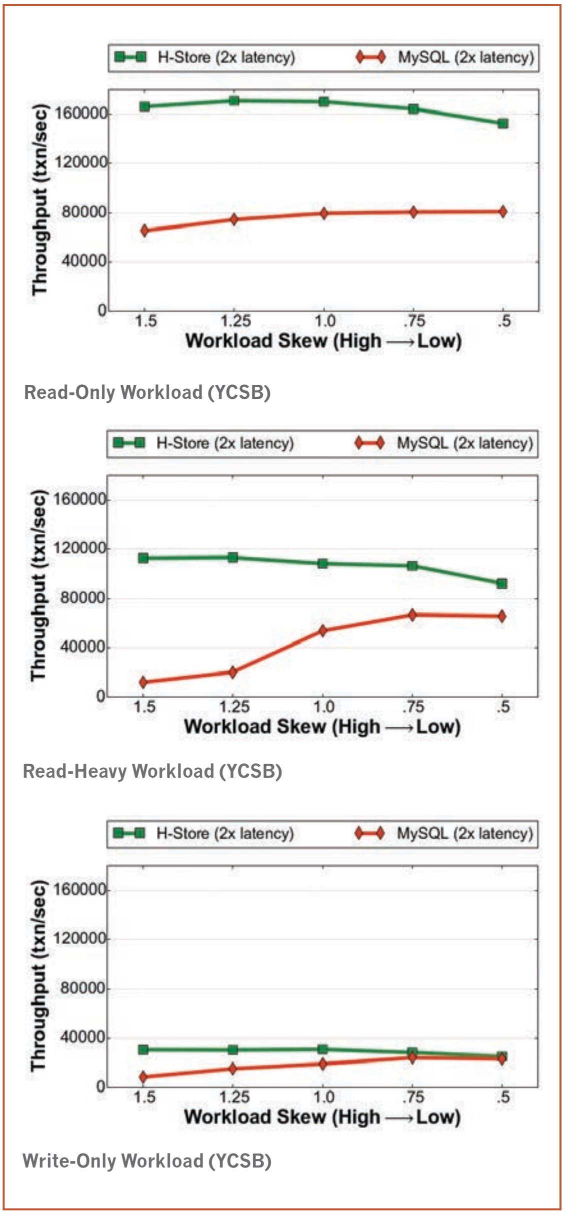


Figure 2: H-Store and MySQL running on an NVM-only system with Intel Lab's hardware emulator.

performs worse on the highly skewed workloads due to lock contention. But this performance difference is nearly non-existent for the write-heavy workload. We attribute this to the overhead of logging.

NVM-only Systems

A more interesting scenario involves a system that does not have any volatile DRAM and only has byte-addressable NVM. Using an NVM-only hierarchy has implications for how a DBMS manages data, because now all memory writes are potentially persistent.

One approach is to designate half of the NVM's address space for temporary data that is deleted after a restart and then use the other half for durable storage. For example, a disk-oriented DBMS using this hierarchy is not "aware" that modifications to data stored in its buffer pool are persistent. As such, many aspects of their architecture are unnecessary. This would require minimum source code changes to existing systems, but I believe that this does not exploit the full potential of NVM and could more quickly wear out the memory device; flash memory has a finite number of program-erase cycles and can be prematurely degraded if it's been written too often and unnecessarily.

MySQL uses a doublewrite mechanism to flush data to durable storage by first writing out the pages to a contiguous buffer on disk before writing them out to the data file. This means that on an NVM-only system, every modification to a tuple in MySQL will cause four writes to the storage device: (1) to the record in the buffer pool, (2) to the write-ahead log, (3) to the doublewrite buffer, and (4) to the primary storage.

Likewise, a system such as H-Store does not have any way to ensure that a change to the database has been completely flushed from the CPU's last-level cache. Thus, when a transaction commits, any changes that it made might still be in volatile caches, and they could be lost before the processor writes them out to the NVM. This means that it also still needs to write to a log for recovery.

Using the same workload described above, we tested MySQL and H-Store again using Intel Lab's emulator as an NVM-only system. All memory allocations were slowed down using a special CPU mode, and all permanent data was stored on the emulator's NVM-optimized file system.

In Figure 2, we see that just as with the NVM+DRAM configuration, the DBMSs' throughputs are greatly affected by the number of update transactions in the workload. This is because each system still writes log entries for each transaction to durable storage, even though all updates to NVM-resident data are potentially persistent. When comparing the systems' peak performance in the read-only workload with the write-heavy workload, we see that H-Store's throughput drops by approximately 81 percent while MySQL's drops by about 72 percent. Since the write-heavy workload has more transactions that update the database, each DBMS has to write more log entries, and thus there is more contention on the logging manager.

The Rise of Multi-Core

Beginning in the mid-2000s, as the trend in hardware shifted to multi-core CPUs, much of the research and development in DBMSs focused on adapting existing architectures that assumed a single CPU core to now utilize additional cores. The 2009 dissertation by former CMU graduate student Ryan Johnson (E 2006, 2010) on optimizing Shore-MT storage manager is probably the best example of this. Although this work was important, timely and necessary, the number of cores that it targeted was relatively small compared to what people expect future processors to be able to support. More recent research papers on shared-memory OLTP DBMSs also evaluate a relatively modest number of cores. For example, I consider the current state-of-the-art in high-performance transaction processing to be the Silo system developed by Harvard's Eddie Kohler and other researchers. In their 2013 paper for the Symposium on Operating Systems Principles, they used a machine with 32 cores. Similarly, a paper on HyPer presented at SIGMOD in 2015 used a 32-core machine. One of the Shore-MT experiments used an 80-core machine, but even 80 cores is still small compared to how I think things will be in the future.

Why Many-Core is Different

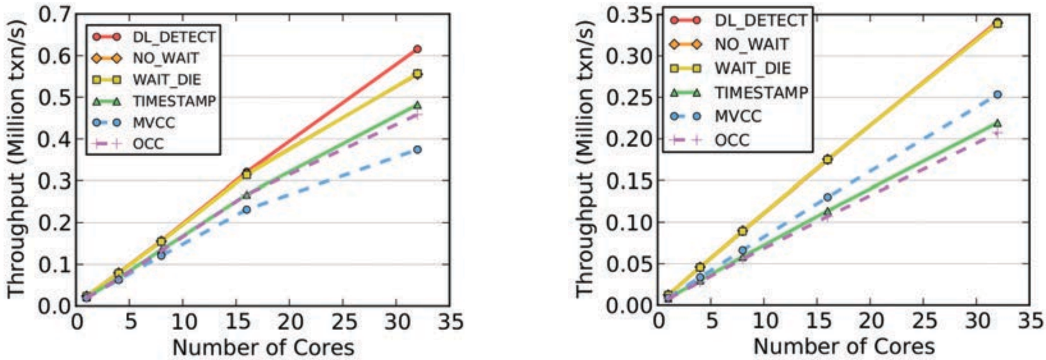
We will soon enter the era of "many-core" machines that are powered by potentially hundreds or even thousands of smaller, low-power cores on a single chip.

You may be wondering how future DBMS research for many-core CPUs is going to be different than what was previously done with multi-cores. I believe the previous systems were all about software approaches that used hardware in a smarter way. But the difference is that with many-cores there seems to be a fundamental barrier into what gains can be achieved with software. Think of it like a "speed of light" limit. As such, it is my hypothesis that the advancements from the previous decade for running DBMSs on multi-core CPUs will not be enough when the core counts are much greater.

To prove this point, I am working with Srini Devadas at MIT and MIT Ph.D. student Xiangyao Yu to investigate how OLTP DBMSs perform on a CPU with 1,000 cores. Such chips obviously do not exist yet, so we have been using a CPU simulator developed by Devadas' group called Graphite, and using a new DBMS written by Yu called DBx1000, to test different aspects of transaction processing. Graphite's simulated architecture is a tiled chip multi-processor where each tile has a small processor core, two levels of cache and a 2D-mesh network-on-chip for communication between the cores. This is similar to other commercial CPUs, such as Tilera's Tile64 (64 cores), Intel's SCC (48 cores) and Intel's Knights Landing (72 cores).

We have conducted initial experiments using DBx1000 to test what happens when executing transactions with high core counts. We started with concurrency control schemes, as this is always the main bottleneck that needs

> Over the next 50 years, we will see significant changes to the database landscape. Beyond obvious things like the volume of data increasing and the velocity at which it is stored being much greater, there will be major changes in how databases are used in applications and the type of hardware on which they will be deployed.



Intel Xeon E7-4830 (Dual Socket)

Graphic Simulator

Figure 3: Comparison of the concurrency control schemes with DBx1000 running in Graphite and a real multi-core CPU using the YCSB workload with medium contention.

to be addressed first in an OLTP DBMS. DBx1000 supports a pluggable lock manager that allows us to swap in different two-phase locking and timestamp-ordering concurrency control schemes. We adapted these schemes from algorithms in Phil Bernstein's seminal 1981 paper on concurrency control and recovery in database systems, but implemented state-of-the-art variants. For example, our MVCC implementation is similar to Microsoft's Hekaton and our OCC implementation is based on Silo. In the graphs below, we compared six of these concurrency control schemes using YCSB. We first execute DBx1000 on a Intel Xeon CPU and then again on our simulator (see Figure 3.)

These results show that all of the algorithms achieve better performance up to 32 cores with the same relative trends. This corroborates previous findings from the papers discussed previously. But now when we crank up the number of cores to 1,000 in the Graphite simulator, things look strikingly different when going past 64 cores:

The results in Figure 4 show that the two-phase locking variant NO_WAIT is the only scheme that scales past 512 cores, but even it fails to scale up to 1,000 cores. There are too many concurrent threads trying to access the same set of records. Clearly, no scheme is ideal here.

The Future

Over the next 50 years, we will see significant changes to the database landscape. Beyond obvious things like the volume of data increasing and the velocity at which it is stored being much greater, there will be major changes in how databases are used in applications and the type of hardware on which they will be deployed.

It is difficult to predict what the major paradigm shift will be in the field. Nor is it realistic to predict which database companies and products will still be available.

The relational model will likely still dominate for most applications, but developers will no longer need to explicitly worry about which data model their application uses. There will be a tighter coupling of programming frameworks and DBMSs such that all database interactions will be transparent (and optimal).

Likewise, SQL (or some dialect of it) will remain the de facto language for interacting with a DBMS, but humans will never actually write SQL. They will instead ask questions about data in natural language. Such changes will cause a major shift in how we write programs; the developer will model their data in a way that is best understood by

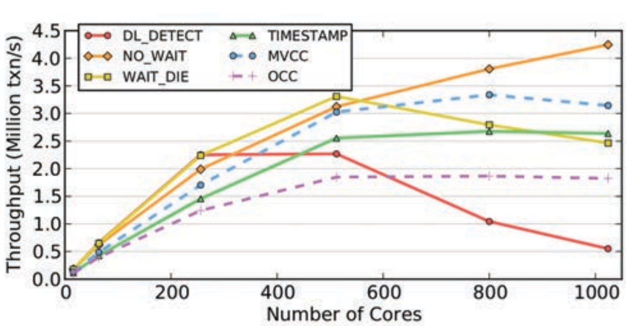


Figure 4: Comparison of the concurrency control schemes with DBx1000 running in Graphite using the YCSB workload with medium contention.

The omnipresent "Internet of Things" will mean that every device is able to collect data about its environment. This will range from small, embedded sensors to larger, autonomous robots. Smaller devices will use an on-chip DBMS in the same way that cellphones now contain on-chip video decoders.

humans and then the framework (in conjunction with the DBMS) will automatically generate the optimal storage scheme for it.

All programs will execute using strongly consistent ACID transactions. That is, the eventual consistency methods that are used in today's Web-based applications will be avoided due to the complexity of managing them. There will be major advancements in network communication, concurrency control and resource management that make using ACID transactions preferable and scalable.

There will be an increasing number of applications where it is more natural to store data in arrays or matrices. This is because organizations will want to analyze large corpora of unstructured information, especially video. We will have mastered the ability to convert all unstructured data into semi-structured formats that are more easily organized and indexed in a DBMS. As part of this, temporality will become important as well, because it matters how information changes over time. Current systems are not able to account for this because of the large overhead of storing extracted information about each video frame in a time series.

The omnipresent "Internet of Things" will mean that every device is able to collect data about its environment. This will range from small, embedded sensors to larger, autonomous robots. Smaller devices will use an on-chip DBMS in the same way that cellphones now contain on-chip video decoders. The databases for all of these systems will be completely composable and easily federated through some standard API (possibly SQL).

This means that DBMSs will communicate with each other with no configuration necessary. You will point two DBMSs at one another, and they will immediately transfer their information and ensure that they are synchronized. Humans will not need to manually configure extract-transform-load utilities or other tools to keep the data on disparate systems consistent.

It will be a significant engineering job to make all of the various DBMSs composable and interoperable in this manner. As such, there will be a toolkit that uses artificial intelligence and machine learning to automatically map the different variations of the DBMSs to each other for the same operation.

In new hardware, more flexible and programmable processing fabrics will become prevalent. DBMSs will compile the critical sections of their program (e.g., the lock manager) into a hardware accelerator. We also will see the disappearance of the dichotomy between volatile and non-volatile memory. The DBMSs will assume that all memory is fast and durable; the need to maintain caches in volatile memory will be unnecessary. This new memory will be orders of magnitude larger than what is available today. Thus, the DBMS will store multiple copies of its data in pre-computed materialized views in order to quickly respond to any possible query.

The role of humans as database administrators will cease to exist. These future systems will be too complex for a human to reason about. DBMSs will finally be completely autonomous and self-healing. Again, the tighter coupling between programming frameworks and DBMSs will allow the system to make better decisions on how to organize data, provision resources and optimize execution than human-generated planning.

And we will likely see the rise of database transactions for inter-planetary devices (e.g., space probes). In this scenario the DBMSs running on these vessels will be at greater distances from each other than Earth-bound systems and incur significantly longer latencies (i.e., minutes or hours). This means that the weak consistency techniques and practices that are used in today's Web-based applications will then be applied to these inter-planetary systems.

—Andy Pavlo (www.cs.cmu.edu/~pavlo/) is an assistant professor in CMU's Computer Science Department. His research interest is in database management systems, including main memory systems, non-relational systems (NoSQL), transaction processing systems (NewSQL), and large-scale data analytics. At CMU, he is a member of the Database Group and the Parallel Data Laboratory. His work is also done in collaboration with the Intel Science and Technology Center for Big Data at MIT.

VELOSO NAMED HEAD OF MACHINE LEARNING DEPARTMENT

Manuela Veloso (CS 1989, 1992) has been appointed the new head of CMU's Machine Learning Department.

She replaces Tom Mitchell, the E. Fredkin University Professor of Computer Science and founding head of the department, who has stepped down to focus on his research. Geoff Gordon (CS 1999), associate professor of machine learning, served as interim department head during the final stages of the search.

In an email announcing Veloso's appointment, Andrew Moore, dean of SCS, called her the "modern-day embodiment" of the legacy of artificial intelligence research that dates to the earliest days of computer science at CMU.

"The founders of computer science at CMU were also two of the four founders of the entire field of AI," Moore said. "One of the main things the search committee saw in Manuela was her dedication to the continued success of this half-century legacy, and someone who was excited about all levels of the machine learning 'stack,' from sensors and kernel-level improvements, to algorithms and fundamental statistics research, to big decision-making systems which improve through experience."

Moore thanked Mitchell for his role in founding the Machine Learning Department—the first stand-alone department of its kind in the world—and in advancing the field through his teaching and research.

Mitchell, he said, has had an impact on "hundreds of careers" and has been one of the "primary forces pushing the discipline of machine learning to become a major component of the world's economy."

Moore also thanked Gordon for serving as interim department head, as well as members of the search committee, including chair Bob Murphy and Emma Brunskill, Aarti Singh, Alex Smola and Larry Wasserman.

Manuela Veloso (CS 1989, 1992) greets Democratic presidential candidate Hillary Clinton during her April 6 visit to CMU's Pittsburgh campus. The former U.S. secretary of state toured the Planetary Robotics Lab before making a speech in the Skibo Gymnasium to a crowd estimated at more than 3,000.

Veloso, CMU's Herbert A. Simon University Professor of Computer Science, has been a faculty member since earning her Ph.D. in computer science at Carnegie Mellon in 1992. Her thesis work involved automated planning and learning by analogy. That led to her long-standing research objective of achieving robots with full autonomy, including agents capable of planning, executing, learning and cooperating, particularly in complex, uncertain and adversarial environments.

With her students, Veloso conducts research on a variety of autonomous robots, including pioneering work on robot soccer. Last year, her CMU robot soccer team won its fifth world championship in the RoboCup small-size league.

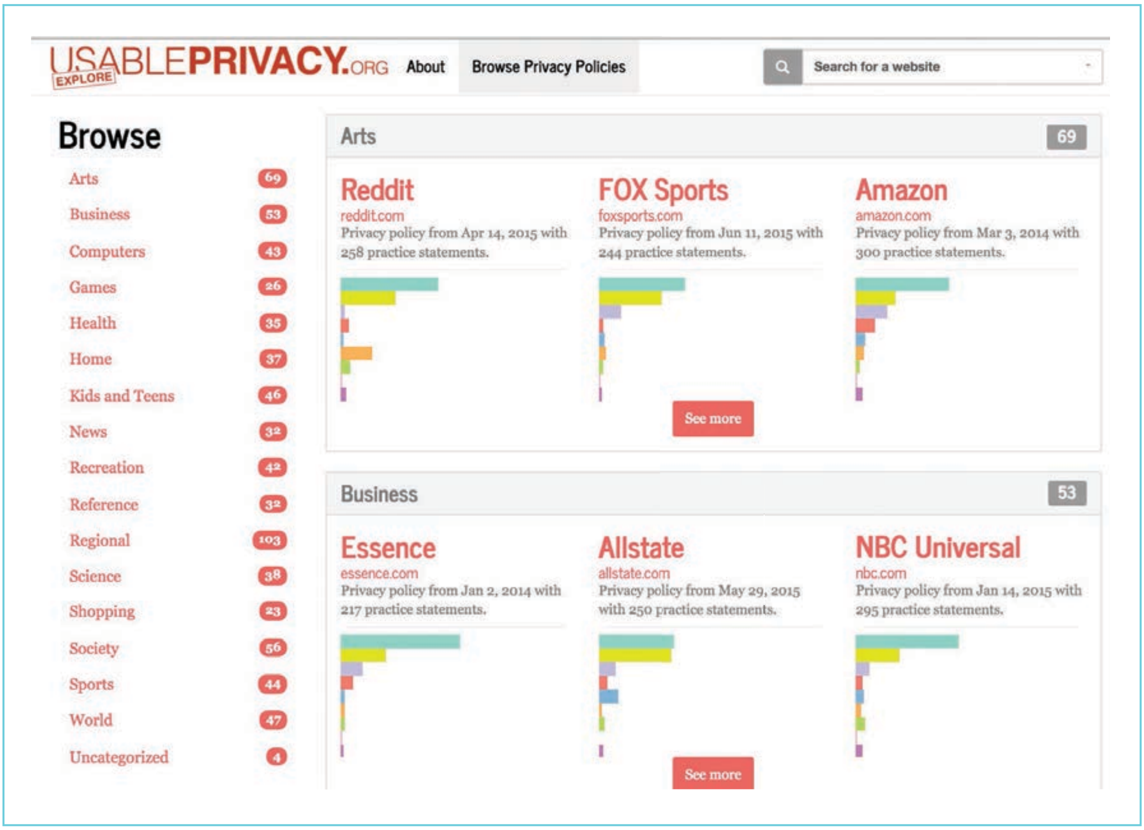
Using her CoBot service robots, Veloso has developed the concept of symbiotic autonomy, in which intelligent mobile robots are autonomous, but also aware of their physical, cognitive and perceptual limitations and able to ask for help when necessary. Using this approach, her CoBot robots have been running errands and doing other tasks in CMU's Gates and Hillman centers since 2011.

A mentor and advocate for increasing the number of women in the discipline of computer science, Veloso presented a keynote address last fall to the annual Grace Hopper Celebration of Women in Computing in Houston, Texas.

In 2014, she was named a University Professor, the highest academic accolade bestowed by CMU, and in 2012 she was honored as an Einstein Chair Professor by the Chinese Academy of Sciences.

Veloso is the past president of the Association for the Advancement of Artificial Intelligence, as well as co-founder and past president of the International RoboCup Federation. She is a fellow of AAAI, IEEE and AAAS.

WEBSITE STRIPS CONFUSION FROM DENSE PRIVACY POLICIES



Few people read privacy policies. But research conducted over the past two years by researchers at CMU, Fordham and Stanford is paving the way to a day when technology may provide users with short summaries of privacy policies.

The Usable Privacy Policy project has introduced a free, online application that enables visitors to navigate more than 23,000 privacy policy annotations covering 193 websites.

The project leverages crowdsourcing, machine learning and natural language processing to semi-automatically annotate privacy policies, extracting relevant statements from the long and convoluted policies found on many websites and mobile apps today.

"This is the first site to provide analysis of privacy policies at this scale," says Norman Sadeh (CS 1991), professor in the Institute for Software Research, lead principal investigator on the study and a researcher in Carnegie Mellon's CyLab security and privacy institute.

The project's objective is to provide what Sadeh calls "succinct yet informative summaries" of privacy policies that can be included in browser plug-ins or incorporated into privacy assistants.

Color codes help users select from a menu of privacy practices that might interest them. The interactive tool covers a comprehensive number of different practices, including whether the site provides opt-out or opt-in choices for users, discloses its retention policy, complies with applicable laws, and much more.

The tool also gives each privacy policy a grade on reading level based on its language. Google's privacy policy, for example, is written on the level of a first-year college student. But according to the tool, the privacy policy for Playstation.com—a site used by many children and teen visitors—is written on the level of someone who's already graduated from a four-year college.

While the annotations on the website were crowdsourced from law students at Fordham University, Sadeh says the researchers are working toward fully automating the service using machine learning and natural language processing.

The website design team includes Institute for Software Research post-doctoral fellows Mads Schaarup Andersen, Florian Schaub and Shomir Wilson; Language Technologies Institute graduate student Aswarth Dara; and undergrad computer science freshman Sushain Cherivirala.

It's available at https://explore.usableprivacy.org/

PROJECT PINPOINTS PESKY POTHOLES AND PAVEMENT PROBLEMS

Pavement riddled with cracks. Graffiti on stop signs. Icy surfaces that need rock salt. Municipalities must respond to road infrastructure problems that are changing constantly.

"It's essential to get eyes on every road, every year, to stay ahead of what could become costly repairs," says Jason Dailey, director of public works in Cranberry Township, Butler County, north of Pittsburgh. "Expensive services are available that have on-board tools and sensors, but these are typically out of the price range of the average community."

Instead of those more costly solutions, Christoph Mertz, principal project scientist at the Robotics Institute, is researching whether it might be possible to harness the sensors on-board any commercially available smartphone.

Mertz is experimenting with a system that uses smartphone cameras to grab high-resolution images of roadways, which are then analyzed using computer vision algorithms. When looking for road damage, for example, the system can detect the ratio between cracked and un-cracked pavement. It also can look for signs that are missing or damaged, and detect snow or slush on the road.

Areas that need care are flagged for further investigation; software displays the data using easy-to-read maps and visuals.

Mertz's technology has been tested with the City of Pittsburgh, Marshall Township, Cranberry Township and the Pennsylvania Department of Transportation.

In a fully complete system, Mertz says, data collection and analysis "could take a matter of days" instead of months or years.

A big appeal of Mertz's system is the simplicity of integrating it with existing procedures. For example, Mertz suggests mounting smartphones onto garbage trucks so that they can routinely assess the roadways as they make their rounds. Similarly, snow plows with smartphones could provide real-time road conditions in winter.

"Mertz demonstrated a viable advancement that may bring inspection technology into the everyday operations, making it not only affordable, but practical," Dailey says.

The project was funded by CMU's Traffic21 institute, launched in 2009 with the help of Pittsburgh's Hillman Foundation. Traffic21's goal is to design, test, deploy and evaluate technology-based solutions to address the problems facing transportation systems.

ANALYSIS: PRESIDENTIAL CANDIDATES SPEAKING ON SIXTH-GRADE LEVEL

If you think politicians seem to be acting awfully juvenile nowadays, you might be right.

A readability analysis of speeches made by presidential candidates finds most are using words and grammar typical of students in grades six through eight.

In terms of grammar, none of the presidential candidates could compare with Abraham Lincoln's Gettysburg Address—an admittedly high standard, with grammar well above the 10th grade level.

The study, done by researchers in CMU's Language Technologies Institute, suggests that all five candidates in the analysis—Republicans Donald Trump, Ted Cruz and Marco Rubio, and Democrats Hillary Clinton and Bernie Sanders—have been using simpler language as the campaigns have progressed. Trump has tended to lag the other candidates, the analysis indicates; his use of grammar and choice words was highest during the Iowa caucus, then plummeted during a later speech at the Nevada caucus.

A comparison of the candidates with previous presidents show Lincoln outpacing them all, boasting grammar at the 11th grade level, while President George W. Bush's fifth-grade grammar was below that currently being used by Trump.

Maxine Eskenazi (DC 1973), LTI principal systems scientist, performed the analysis with Elliot Schumacher, a graduate student in language technologies. Eskenazi says assessing the readability of campaign speeches in tricky, because most of the language analysis formulas are geared to the written word. But the written word is very different from the spoken word, she says.

"When we speak, we usually use less structured language with shorter sentences," Eskenazi says.

Analyzing campaign speeches is also difficult because it is often hard to obtain transcripts of speeches, Schumacher says.

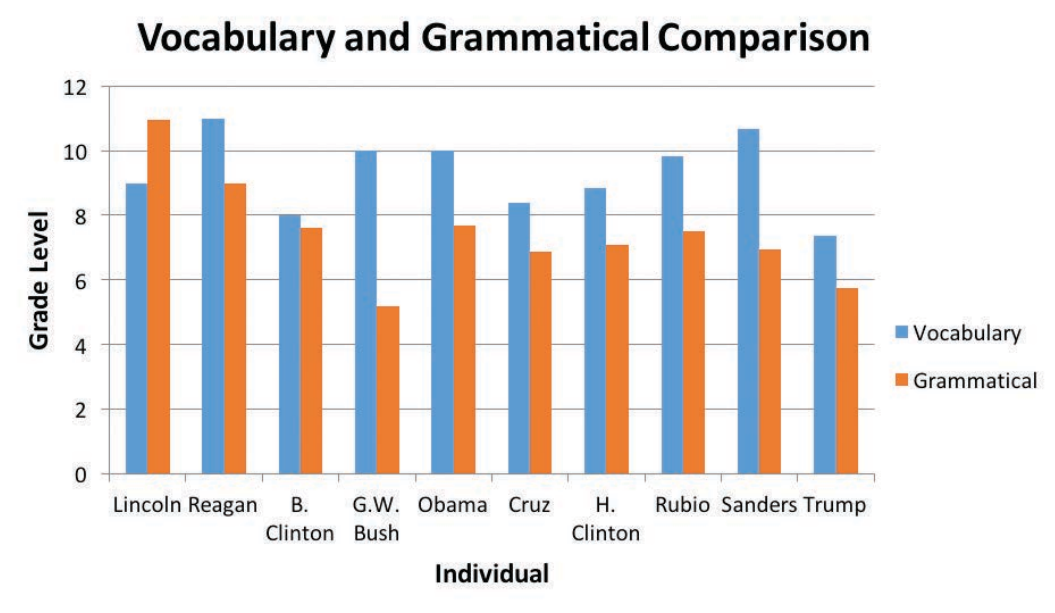
It is possible to generate reliable transcripts from video using automatic speech recognition systems such as those developed at the LTI, but he and Eskenazi opted not to use today's automated methods because they were likely to introduce errors in the noisy environment of campaign rallies.

Once they obtained or created transcripts, the researchers used a readability model called REAP, which looks at how often words and grammatical constructs are used at each grade level.

Based on vocabulary, campaign speeches by past and present presidents—Lincoln, Reagan, Bill Clinton, George W. Bush and Obama—were at least on the eighth grade level, while the current candidates ranged from Trump's seventh grade level to Sanders' 10th grade level.

Trump and Hillary Clinton's speeches showed the greatest variation, suggesting they may work harder than the others in tailoring their speeches to particular audiences, Schumacher says.

The complete study is available online via the LTI website at www.lti.cs.cmu.edu.



UPMC FUNDS FIRST RESEARCH UNDER HEALTH DATA ALLIANCE PROGRAM

A system for helping physicians make smarter treatment decisions is the first CMU project to be funded by the Pittsburgh Health Data Alliance.

UPMC Enterprises made the award to the Clinical Genomics Modeling Platform as part of an initial, $3 million round of funding to projects both at CMU and the University of Pittsburgh.

Announced last March, the Pittsburgh Health Data Alliance is a collaboration among CMU, the $12 billion UPMC healthcare system and Pitt. It will focus on building new companies that create data-intensive software and services, with the potential to revolutionize health care and wellness.

The Clinical Genomics Modeling Platform project is being led by Carl Kingsford and Christopher Langmead, both associate professors of computational biology. It is part of CMU's Center for Machine Learning and Health, which is spearheaded by Eric Xing, professor of machine learning.

The platform is an engine for easily building precision-medicine models for various diseases and populations. Triage algorithms, for instance, might help to determine if patients with a certain disease should be sent home with monitoring or sent to the intensive care unit.

"We are excited to move forward with the first of many exceptional ideas in the Health Data Alliance pipeline," said Tal Heppenstall (TPR 1985), president of UPMC Enterprises, an investment arm of UPMC created to provide early funding to healthcare technologies that have commercial potential. "This promising start bodes well for the Alliance's goal of transforming health care by unleashing the creativity and entrepreneurialism of leading scientists and clinicians in Pittsburgh."

NREC TEAM WINS $10K

A team from CMU's National Robotics Engineering Center has been named one of four semi-finalists and awarded $10,000 in the Naturipe Blue Challenge, a contest to develop innovative technologies for harvesting blueberries.

Dimi Apostolopoulos and Gabriel Goldman were instrumental in developing the concept for the NREC entry.

ENERGY DEPARTMENT FUNDS SPECIALIZED TRAINING FOR CLEANING UP NUKE SITES

A new robotics "traineeship" program will provide specialized education for graduate students interested in development of robots that can remediate nuclear sites.

Under an agreement with the federal Department of Energy, up to $3 million in funding will be provided over five years to provide full or partial support for as many as 20 master's and Ph.D. students in robotics.

The selection of CMU for the program by the federal Office of Environmental Management was announced March 16 by Elizabeth Sherwood-Randall, U.S. deputy energy secretary.

To carry out the program, CMU will team with two DOE laboratories, Savannah River National Laboratory in Aiken, S.C., and Pacific Northwest National Laboratory in Richland, Wash.

Nathan Michael, assistant research professor of robotics, said the program is expected to begin in the fall.

Like a fellowship, the traineeship will provide financial support for the students' education, Michael said. But it also will support the Robotics Institute's development of specialized courses and will provide research opportunities in association with the partner labs that will help extend the use of autonomous systems in remediation efforts.

The DOE's Office of Environmental Management was created to address the safe cleanup of the environmental legacy created by five decades of nuclear weapons development and government-sponsored nuclear energy research.

A DOE spokesperson said the new program answers a critical need for workers trained in retrieval, treatment, processing, storage, transportation and disposal of radioactive waste, as well as safeguarding of spent nuclear fuel, operations and maintenance of nuclear facilities, worker and facility safety, deactivation of former nuclear sites and other activities.

The new traineeship at CMU is available to students who have been admitted to an existing robotics graduate program and have expressed an interest in environmental remediation.

SCS ALUMNA HELPS NYC 'RIDE THE WAVE' OF CS EDUCATION IN PUBLIC SCHOOLS

Leigh Ann DeLyser (CS 2010, 2014) with teacher Brian Schott and students at the Bronx Academy for Software Engineering.

An SCS alumna and former high school teacher is helping New York City ride a "tidal wave" of computer science education.

Leigh Ann DeLyser (CS 2010, 2014) is now director of education and research for the New York City Foundation for Computer Science Education, or CSNYC for short.

Along with CMU's Mark Stehlik and Chris Stephenson of the Computer Science Teachers' Association, DeLyser was co-author of "Running on Empty," a 2011 report for the Association for Computing Machinery that warned about the lack of computing education in America's K-12 schools. (See "Closing the Educational Gap," The Link, Winter 2011.)

Stehlik, now SCS associate dean for outreach, first became acquainted with DeLyser in 2000 when she was a teacher for an advanced placement program in computer science for high school students.

"Leigh Ann stood out immediately as someone with a learned and passionate voice, especially about teaching computer science at the high school level," he says.

New York City, Chicago and San Francisco are among the few American cities that have taken steps to mandate computer science education. At CSNYC, DeLyser helps coordinate programs across the system to ensure strong implementation of computer science education at the school level.

In an announcement this past fall, New York Mayor Bill de Blasio said the city's public schools will now offer computer science to all students; more than 20,000 students in the New York City school system have had CS education in the past three years, and the administration has ambitious plans to reach all 1.1 million students in the next 10 years.

CSNYC may indeed be at the leading edge of a trend. In February, the White House asked for $4 billion in funding so that states can expand computer science education, and directed the National Science Foundation to spend $125 million to advance research in computer science education and train teachers.

While at CMU, DeLyser pursued an interdisciplinary, self-directed doctoral degree through the university's Program in Interdisciplinary Education Research. She explored topics in psychology, computer science, human computer interaction, statistics, physics, economics and public policy, and applied them to education research. Stehlik was one of her faculty advisors.

DeLyser was the first PIER student to bridge the two disciplines of computer science and psychology, says Sharon Carver (DC 1987), associate training director in PIER and a teaching professor in the Department of Psychology.

A large obstacle to getting computer science taught in K-12 schools is a lack of teacher certification requirements in the subject. In New York City, DeLyser is working directly with school leaders, teachers and policymakers to discuss curriculum development and professional education for teachers.

DeLyser has a strong belief that computer science education in K-12 schools must be grounded in sound theory, so she is also arranging collaborations between schools and industry leaders to bring real-world needs into the classroom. "We want every child to have a meaningful experience in computer science," she says.

"She was definitely one of those people who had a vision of what she wanted to do, and CMU enabled it to become a reality," Stehlik says. "We are extremely proud of her as an alumna."

LIBRARY USERS CHECK THEIR AIR WHILE THEY CHECK OUT A BOOK

Indoor air-quality monitors developed at CMU will now be available for free use at 100 public libraries across the United States.

Speck sensors, invented in CMU's CREATE Lab, detect fine particles of air pollution in homes and offices. Hundreds of patrons have already borrowed them from the Carnegie Library of Pittsburgh.

Now, the CREATE Lab and spinoff company Airviz, which makes and markets the device, are offering three free Specks, informational materials and training to public libraries that agree to make them available to their patrons.

"We have too many communities where the air is hazardous from time to time, yet people can't readily find out what they are breathing in their own homes," says Illah Nourbakhsh, professor of robotics and head of the CREATE Lab. "This is the air quality you can actually do something about—if you know that a hazard even exists. That's why it is so important that people of all income levels have access to a sensor such as Speck."

Speck is designed for use indoors, helping users realize when polluted outside air is coming into the house, or to recognize pollution sources within the home. It is also Wi-Fi-connected, so air quality data can be uploaded for analysis and shared, if desired, via the Internet.

Interested libraries can apply for the National Speck Library Program at www.specksensor.com. In addition to three free Specks, participating libraries also receive a 15 percent discount on purchases of additional Specks. Carnegie Library of Pittsburgh, which loans Specks through 16 of its 19 branches, is providing advice and support for the national campaign.

The CREATE Lab and Airviz introduced the Speck personal air quality monitor a year ago at the SXSW Interactive Festival in Austin, Texas. To keep the monitors affordable—they retail at $149—Speck uses a low-cost infrared sensor to detect pollutants.

Developers employed machine learning algorithms that learn to recognize and compensate for spurious "noise" in each detector, boosting accuracy.

Speck was placed in the Pittsburgh library branches with support from the Heinz Endowments, Fine Foundation and Pittsburgh Foundation. The national library campaign thus far is being supported by CREATE Lab and Airviz in a bid to "pay it forward," says Mary Beatrice Dias (E 2009, 2011), project director in the Robotics Institute.

Adds Nourbakhsh: "Providing equitable access to monitoring technology is too important for us not to do this."

—Compiled from staff reports

NAMES IN THE NEWS

A CMU student team won the $10,000 grand prize at the 2015 Facebook Global Hackathon, besting 20 other teams from 11 countries. The team included **Tiffany Jiang**, a sophomore design and human-computer interaction major; **Avi Romanoff**, a sophomore psychology and human-computer interaction major; and **Nikhil Choudhary**, a junior electrical and computer engineering major. **Sumanth Pandugula**, a graduate student at the University of Illinois at Chicago, also was a member of the team.

Award-winning filmmaker Werner Herzog's latest film, "Lo and Behold: Reveries of the Connected World," prominently features several CMU scientists, including **Marcel Just**, the D.O. Hebb University Professor of Psychology, and **Joydeep Biswas** (CS 2009, 2014), a former post-doctoral researcher in the Computer Science Department, discussing CMU's championship robot soccer team. The film delves into society's dependence on the Internet for just about everything. It premiered at the 2016 Sundance Film Festival.

The National Academy of Sciences will award **John R. Anderson** the 2016 Atkinson Prize in Psychological and Cognitive Sciences for his "foundational contributions to systematic theory and optimality analysis in cognitive and psychological science and for developing effective, theory-based cognitive tutors for education." Anderson will receive the prize—a gold-plated bronze medal and $100,000—at the academy's annual meeting May 1 in Washington, D.C.

Ed Felten, the deputy U.S. chief technology officer in the White House Office of Science and Technology Policy, was this year's keynote speaker at CMU's Privacy Day celebration held Jan. 28.

The Association for the Advancement of Artificial Intelligence has elected **Eric P. Xing**, professor of machine learning and director of the Center for Machine Learning and Health, one of six 2016 AAAI fellows. The new class of fellows was recognized Feb. 14, during the annual AAAI conference in Phoenix. Xing joins 17 current or former CMU faculty members previously named AAAI fellows.

Robert Kraut is the recipient of the 2016 SIGCHI Lifetime Achievement in Research Award. Kraut, CMU's Herbert A. Simon Professor of Human-Computer Interaction, joined the Carnegie Mellon community in 1993. His research has focused on the design and impact of social computing. Kraut will receive the award at the 2016 ACM Conference on Human Factors in Computing Systems, May 7–12 in San Jose, Calif.

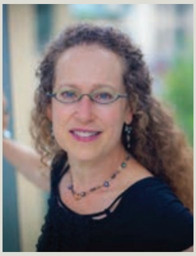
CMU awarded its 2015 Dickson Prize in Science to **Judea Pearl**, a computer scientist at UCLA internationally known for his contributions to artificial intelligence, human reasoning, causality and the philosophy of science. Pearl accepted the award and presented the Dickson Prize Lecture, "Science, Counterfactuals and Free Will," on Feb. 29 in the Cohon University Center.

Abhinav Gupta, an assistant professor of robotics who specializes in computer vision and large-scale visual learning, is among 126 outstanding U.S. and Canadian researchers chosen as recipients of the 2016 Sloan Research Fellowships. **Wesley Pegden**, assistant professor of mathematical sciences in the Mellon College of Science, was also honored. They each received $55,000 to further their research.

Alumnus **Ivan Sutherland** (S 1959), known as the father of computer graphics, will be inducted into the National Inventors Hall of Fame on May 5 in Washington, D.C. Sutherland will be among 16 new inductees into the hall.

Jaime Carbonell, director of the Language Technologies Institute, was awarded the 2015 Okawa Prize for "outstanding contributions to research in language technologies, machine learning and computational biology in the field of artificial intelligence." The prize includes a certificate, gold medal and an $81,000 cash award. Previous SCS winners of the Okawa Prize include **Raj Reddy** (2004) and **Takeo Kanade** (2007).

A new book explores ways to encourage women to pursue computer science education, using strategies developed at CMU. “Kicking Butt in Computer Science: Women in Computing at Carnegie Mellon University” was written by **Carol Frieze** (DC 1989, CS 2007), director of Women@SCS, and **Jeria Quesenberry**, associate teaching professor of information systems in the Dietrich College of Humanities and Social Sciences.



Justine Cassell, SCS associate dean for technology strategy and impact, was named a corresponding fellow of the Royal Society of Edinburgh. The society cited Cassell for her contributions to computer science and to human-computer interaction in particular. She and 55 other new fellows will be inducted at a May 16 ceremony.

A computer poker program called Baby Tartanian8 continued CMU’s hot streak at the Annual Computer Poker Competition, taking first place in the total bankroll category and third place in the bankroll instant run-off category in the Heads-Up, No-Limit Texas Hold’em game. **Tuomas Sandholm**, professor of computer science, and **Noam Brown** (CS 2014), a Ph.D. student in the Computer Science Department, created the pokerbot. The results were announced Feb. 13 at the Association for the Advancement of Artificial Intelligence meeting in Phoenix. Ten teams competed.

calendar of events

All events to be held on the Carnegie Mellon University campus in Pittsburgh, unless otherwise noted. Dates and locations subject to change without notice. Visit www.cs.cmu.edu/calendar for a more up-to-date and current listing of events.

May 4
Meeting of the Minds
Undergraduate Student
Research Symposium
10 a.m. to 5 p.m.,
Cohon University Center

May 13
Carnegie Mellon Alumni Awards
6 p.m., Choskey Theatre,
Purnell Center for the Arts

May 14 and 15
2016 Commencement

May 16
Summer semester begins

May 30
Memorial Day, no classes

June 8
Bay Area alumni happy hour, hosted by LinkedIn, Sunnyvale, Calif.

Summer 2016
Regional alumni events
Seattle, Boston, Washington, D.C.
Locations TBA

July 4
Independence Day, no classes

Aug. 14–18
Freshman Orientation Week (CMU Qatar)

Aug. 21
Fall semester begins (CMU Qatar)

Aug. 29
Fall semester begins

Sept. 5
Labor Day, no classes

Sept. 11–15
Eid al-Adha Break, no classes (CMU Qatar)

Sept. 12
Add-drop deadline

Oct. 21
Mid-semester break, no classes

Nov. 23–25
Thanksgiving holiday, no classes

Dec. 1
Fall semester, last day of classes (CMU Qatar)

Dec. 4–8
Final exams (CMU Qatar)

Dec. 12–18
Final exams

Tracking changes, 1965 to today

R. K. Mellon Gives Tech $5 Million For Computer Studies

Carnegie Alumni News, September 1965, p. 3

The intellectual preeminence of Carnegie Tech in the computer and information sciences will be of lasting value to the community and the nation, said General Richard K. Mellon in giving Tech $5 million dollars for expanded activity in those fields.

> Jason Togyer 3/28/2016
> **Comment [1]:** The merger between Carnegie Institute of Technology and Mellon Institute of Industrial Research was announced in 1966 and in 1967, Carnegie-Mellon University was created.

The grant from General Mellon, and the charitable trusts established by him is for a new department of computer and information sciences in the College of Engineering and Science.

In making the gift, General Mellon added, "Mrs. Mellon and I hope this grant will provide the basis for an expanded program leading to the practical and intellectual preeminence of Carnegie Tech in computer and information sciences."

Dr. H. Guyford Stever, president of Carnegie Tech, said the grant will be used for the following projects:

> Jason Togyer 3/28/2016
> **Comment [2]:** Horton Guyford Stever (1916-2010) was a physicist with a bachelor's degree from Colgate and a Ph.D. from Caltech. He served as director of the National Science Foundation from 1972 to 1976 and was science advisor to Vice President (later President) Gerald Ford from 1973 to 1977.

Construction of a $2 million 60,000 square foot building located on the west side of Doherty (Engineering) Hall to serve as the focal point of research and education in computer science.

> Jason Togyer 3/28/2016
> **Comment [3]:** Work on this building—at first called "Science Hall" and later renamed "Wean Hall"—began in 1968. The total cost was $13 million.

Expenditure of $1,250,000 over a five-year period for research in computer development and applications, including existing projects in computer languages and systems, thought processes, design, management information systems and other areas of potential benefit.

One million dollars towards the purchase of a new advanced computer and $750,000 to endow a named professorship.

> Jason Togyer 3/28/2016
> **Comment [4]:** The first recipient of the Richard King Mellon Professorship of Computer Science and Psychology was Herbert A. Simon. John R. Anderson currently holds the chair.

In making the announcement, Dr. Stever said, "One of the important factors in my decision to accept the presidency of Carnegie Tech, was that I found it one of the leaders in the field of computer sciences in terms of faculty, students and intellectual accomplishments. This magnificent gift will enable us to expand our program and help coordinate, centralize and increase the many areas of computer research.

> Jason Togyer 3/28/2016
> **Comment [5]:** Stever became the fifth president of Carnegie Tech on Feb. 1, 1965. Stever House, CMU's newest residence hall, is named in his honor.

"It will also make possible an increase in the number of outstanding faculty members in this area and enable us to maintain a strong graduate program in many fields. While it will enroll only graduate students, initially, it is expected that the new department will be expanded to include undergraduate programs in the next few years.

> Jason Togyer 3/28/2016
> **Comment [6]:** This took longer than anyone expected; CMU didn't offer its first undergraduate degree in computer science until 1989, when seven sophomores were admitted as CS majors.

"In addition, Tech will be able to offer facilities in computer and information sciences which will make Pittsburgh more attractive for industrial and governmental operations. With the acquisition of the newest IBM computer (system 360-Model 67), our facilities will be the equal in quality of those of any college or university in the country."

> Jason Togyer 3/28/2016
> **Comment [7]:** Stever's prediction certainly came true; CMU's prowess in computer science has attracted Apple, Google, Intel, Uber and many other modern industrial giants to Pittsburgh.

An important area of research to be intensified is the development of computer languages and systems. One objective of research in computer language is to communicate directly with the computer in English instead of a variety of symbolic languages. This program will be administered by

Dr. Alan Perlis, international authority and pioneer in computer languages. He is currently director of Carnegie Tech's Computation Center.

Work in the field of simulation of thought processes begun by Professors Herbert A. Simon and Allen Newell will also be increased.

Other intensified research activities include problems of engineering design, and the development of complete information systems for business firms.

Established in 1956, the Computation Center has had a remarkable growth. There are almost 100 fulltime employees and an annual operating budget of nearly $1,500,000. Facilities include a paired computer with the second largest memory storage unit in education or industry in the country.

Tech recently was chosen as one of three major computer facilities, along with MIT and the System Development Corp., for study of Information Processing established by the Advanced Research Projects Agency of the Department of Defense.

Current research programs related to computer sciences at Tech involve an expenditure of $2.5 million yearly. They include projects in behavioral sciences, management sciences, fine arts and in all the departments of the College of Engineering and Science.

Computer courses are available for undergraduates in the departments of electrical engineering and mathematics. At present some 1300 students and faculty regularly make use of Tech's computers.

Jason Togyer 3/28/2016

Comment [8]: Perlis (S 1942), a Pittsburgh native, became chair of the computer science department at Yale University in 1971. His article "Epigrams on Programming," written in 1982, remains widely quoted.

Jason Togyer 3/28/2016

Comment [9]: This was CMU's innovative G-21, which consisted of two Bendix G-20 computers running in parallel, with a shared memory. It was supplanted by a Univac 1108 and an IBM System 360/67.

Jason Togyer 3/28/2016

Comment [10]: Following the merger with Mellon Institute, the College of Engineering and Science was split into the Carnegie Institute of Technology (engineering) and the Mellon Institute of Science.

CARNEGIE ALUMNI NEWS

published for the Alumni of Carnegie Institute of Technology, Pittsburgh, Pennsylvania September 1965

R. K. Mellon Gives Tech $5 Million For Computer Studies

The intellectual preeminence of Carnegie Tech in the computer and information sciences will be of lasting value to the community and the nation, said General Richard K. Mellon in giving Tech $5 million dollars for expanded activity in those fields.

The grant from General Mellon, and the charitable trusts established by him is for a new department of computer and information sciences in the College of Engineering and Science.

In making the gift, General Mellon added, "Mrs. Mellon and I hope this grant will provide the basis for an expanded program leading to the practical and intellectual preeminence of Carnegie Tech in computer and information sciences."

Dr. H. Guyford Stever, president of Carnegie Tech, said the grant will be used for the following projects:

Construction of a $2 million 60,000 square foot building located on the west side of Doherty (Engineering) Hall to serve as the focal point of research and education in computer science.

Expenditure of $1,250,000 over a five-year period for research in computer development and applications, including existing projects in computer languages and systems, thought processes, design, management information systems and other areas of potential benefit.

One million dollars towards the purchase of a new advanced computer and $750,000 to endow a named professorship.

In making the announcement, Dr. Stever said, "One of the important factors in my decision to accept the presidency of Carnegie Tech, was that I found it one of the leaders in the field of computer sciences in terms of faculty, students and intellectual accomplishments. This magnificent gift will enable us to expand our program and help coordinate, centralize and increase the many areas of computer research.

"It will also make possible an increase in the number of outstanding faculty members in this area and enable us to maintain a strong graduate program in many fields. While it will enroll only graduate students, initially, it is expected that the new department will be expanded to include undergraduate programs in the next few years.

"In addition, Tech will be able to offer facilities in computer and information sciences which will make Pittsburgh more attractive for industrial and governmental operations. With the acquisition of the newest IBM computer (system 360-Model 67), our facilities will be the equal in quality of those of any college or university in the country."

An important area of research to be intensified is the development of computer languages and systems. One objective of research in computer language is to communicate directly with the computer in English instead of a variety of symbolic languages. This program will be administered by Dr. Alan Perlis, international authority and pioneer in computer languages. He is currently director of Carnegie Tech's Computation Center.

Work in the field of simulation of thought processes begun by Professors Herbert A. Simon and Allen Newell will also be increased.

Other intensified research activities include problems of engineering design, and the development of complete information systems for business firms.

Established in 1956, the Computation Center has had a remarkable growth. There are almost 100 full-time employees and an annual operating budget of nearly $1,500,000. Facilities include a paired computer with the second largest memory storage unit in education or industry in the country.

Tech recently was chosen as one of three major computer facilities, along with MIT and the System Development Corp., for study of Information Processing established by the Advanced Research Projects Agency of the Department of Defense.

Current research programs related to computer sciences at Tech involve an expenditure of $2.5 million yearly. They include projects in behavioral sciences, management sciences, fine arts and in all the departments of the College of Engineering and Science.

Computer courses are available for undergraduates in the departments of electrical engineering and mathematics. At present some 1300 students and faculty regularly make use of Tech's computers.

Carnegie Mellon University
School of Computer Science

Office of the Dean
5000 Forbes Avenue
Pittsburgh, PA 15213

Follow us!

SCSatCMU

THANKS FOR HELPING US CELEBRATE THE <SOURCE> OF IT ALL.

FIFTY
Carnegie Mellon University
Computer Science Department
Fiftieth Anniversary

Thank you to everyone who helped us celebrate CSD50—the 50th anniversary of the Computer Science Department, Oct. 23–24, 2015. We'd also like to thank **Google**, **Infosys**, **Microsoft**, **crowdpoll** and **GoDaddy** for their support.

For a replay, visit **www.cs.cmu.edu/csd50/video**

Google Infosys Microsoft Go Daddy crowdpoll