

click!

University of Illinois at Urbana-Champaign

ILLIAC IV: the first supercomputer

ALSO IN THIS ISSUE:



Inaugural Alumni Awards 4



Parallel Computing at Illinois 16




PPL Celebrates its 20th Anniversary 28



ILLIAC IV 32

College of Engineering





the computer was born
to solve problems
that did not exist before

Bill Gates

click!

University of Illinois at Urbana-Champaign

CS@ILLINOIS ALUMNI MAGAZINE

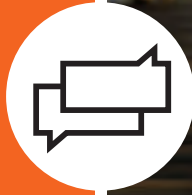
2012, VOLUME I

2	Letter from the Head
4	Inaugural Alumni Awards Banquet
10	Alumni News
13	Did You Know: Blue Waters Stats
14	Q&A With: Bill Gropp
16	Parallel Computing at Illinois
23	In The Lab
28	20 Years of Parallel Programming Lab
32	ILLIAC IV: The First Supercomputer
38	In The Lab
56	In The Classroom
62	Awards
64	In Memory
66	Gifts In Action
70	Faces of CS
75	The #lastword



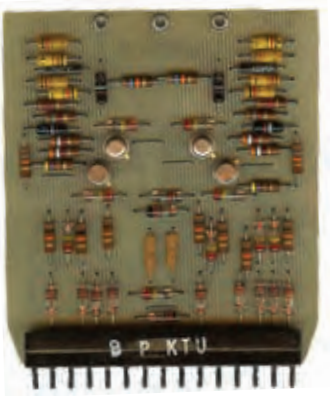
We've been defining the rules of this game for a long time.

Rob A. Rutenbar



Welcome to the new edition of CLICK! -- the alumni/news magazine for the Department of Computer Science at the University of Illinois at Urbana-Champaign. We use CLICK! to give our friends and alumni a regular look at some of the exciting accomplishments of our faculty and students.

I'm particularly excited about this issue, because of the focus on parallelism. This is an area close to the heart of CS, integrating issues across architecture, compilers, operating systems, networks, applications, and theory. It's also an area where Illinois has long been in the lead. Fifty years ago, when a 'serious' mainframe was less powerful than today's cheapest cellphone, we were building Illiac III, a SIMD parallel machine; followed shortly by ILLIAC IV, a 256-node machine with a vector-like architecture. We've been defining the rules of this game for a long time.



And the excitement continues today. New centers in Assured Cloud Computing, and Programmable Extreme-Scale Systems. A new campus-wide effort, the Parallel Computing Institute (PCI), which integrates all our broad, diverse efforts in one place. And lets not forget Blue Waters, a collaboration between the National Center for Supercomputing Applications (NCSA) and Cray, which is today standing up a peta-scale machine with just under 400,000 cores. CS@Illinois is a very good place to be doing parallelism right now.

Those of us working in this area (full disclosure: I'm a hardware guy, and a silicon guy) all have our stories for how we got into this business. It's worth sharing mine. As a new grad student working for Prof. Dan Atkins (CS Ph.D., '69) I was assigned to straighten up a hardware lab. In this process, I came across a box filled with tiny, ancient-looking circuit boards. "Hey Dan!" I asked, "What are these?"

"Illiac III," said Dan. "Want some?"

(I should have grabbed more....)

Rob A. Rutenbar, Head

alumni executive advisory board

Mr. Alex Bratton, BS, 1993, Computer Engineering
Chief Executive Officer and Founder
Lextech Global Services

Mr. Vilas S. Dhar, BS, 2004, Computer Science; BSLAS, 2004, Bioengineering
Co-Founder and Partner
Dhar Law, LLP

Dr. Mary Jane Irwin, Ph.D., 1977; MS, 1975
A Robert Noll Chair in Engineering
Penn State University
Department of CS&E

Ms. Jill C. Magsam, BS, 2000
Senior Product Manager
The TriZetto Group

Mr. Daniel (Dan) L. Peterson, BS, 2005
Product Manager
Google, Inc.

Mr. William (Bill) M. Dunn, MS, 1987; BS, 1986
President
Dunn Solutions Group, Inc.

Mr. Peter (Pete) L. Tannenwald, BSLAS, 1985, Math & Computer Science
Client Partner
Cognizant Technology Solutions

staff

Editor: Jennifer LaMontagne

Writers: Hong Shen, Jennifer LaMontagne,
Kirsten Keller, Tricia Barker (NCSA), Kim
Gudeman (Coordinated Sciences Laboratory),
Nathaniel Lash (ECE), Jenny Applequist
(Information Trust Institute)

Design: Surface 51, Diane Park

Department Head: Rob A. Rutenbar

Associate Department Heads: David Forsyth,
Klara Nahrstedt, Craig Zilles

Associate Director for Communications:
Jennifer LaMontagne

Associate Director for External Relations:
Cynthia Coleman

Contact us at:
alumni@cs.illinois.edu



CS@ILLINOIS

Honors 12 Distinguished Alumni and Faculty

CS@ILLINOIS honored 12 distinguished alumni and former faculty members at the inaugural CS@ILLINOIS Alumni Awards ceremony on November 4. The awards honored computer science graduates and faculty members who have made professional, technical, educational, and service contributions that bring distinction to themselves, the department, and the University of Illinois.

The alumni and faculty honored included:

DISTINGUISHED ACHIEVEMENT

Mary Jane Irwin
Mary McDowell
Max Levchin

DISTINGUISHED EDUCATOR

Baochun Li
C. L. David Liu
Jane W.-S. Liu
Jose Martinez

DISTINGUISHED SERVICE

Dan Peterson
Ira Cohen

MEMORIAL ACHIEVEMENT

Donald Gillies
Erich Hauptmann
Saburo Muroga

distinguished achievement award



MARY JANE Irwin:

Mary Jane Irwin, Evan Pugh Professor and A. Robert Noll Chair, Department of Computer Science and Engineering, The Pennsylvania State University which she joined as a junior faculty member after completing her Ph.D. from UIUC in 1977.

Dr. Irwin's research and teaching interests include computer architecture, energy-aware and reliability-aware design, and emerging technologies. She has supervised more than 25 Penn State Ph.D.'s. She was named a Fellow of IEEE in 1995, a Fellow of ACM in 1996, and was inducted into the NAE in 2003 and into the AAAS in 2009. Awards she has received include the 2003 IEEE/CAS VLSI Transactions Best Paper of the Year Award, the 2004 DAC Marie R. Pistilli Women in EDA Award, the 2005 ACM Distinguished Service Award, the 2007 Anita Borg Technical Leadership Award, and the 2010 ACM Athena Lecturer Award. She received an Honorary Doctorate from Chalmers University, Sweden in 1997.



MAX

Levchin:

Max Levchin is the founder and visionary behind PayPal and Slide. Max is renowned as the co-founder of PayPal, an expert in combating online fraud and one of the hardest working entrepreneurs in Silicon Valley. Before starting Slide, he incubated several other start-ups, including Yelp, where he is currently Chairman of the Board.

Max started PayPal in 1998, took the company public in 2002 and then sold it to eBay for more than \$1.5 billion at the age of 26. In 2004, Levchin founded Slide, a personal media-sharing service for social networking sites such as MySpace and Facebook. Slide was sold to Google in August 2010 for \$182 Million.

Originally from Kiev, Ukraine (then part of the Soviet Union), Max moved to Chicago at the age of 16 and later received his Bachelor of Science degree in Computer Science from the University of Illinois at Urbana-Champaign. Max sits on the board of several other companies and is currently planning his next start-up project.



MARY

McDowell:

Mary McDowell is Executive Vice President in charge of Nokia's Mobile Phones unit, with global P&L responsibility for Nokia's mass market devices and associated services. She is the champion of Nokia's "Next Billion" strategy which brings information and internet access to a new generation of consumers primarily in emerging markets. She has been a member of the Nokia management team since 2004, and was appointed to her current position in July 2010. Mary reports to the CEO. Mary has a successful track record in building new business and as an industry innovator.

Wireless named her one of the "Top Women in Wireless" in July 2008.

Mary joined Nokia in 2004 as Executive Vice President and General Manager of Enterprise Solutions, responsible for the development and marketing of Nokia's range of business-oriented devices and solutions. These included Nokia Eseries devices, mobility software and security and mobile connectivity solutions. Mary served as Executive Vice President and Chief Development Officer from 2008 until assuming her current role.

She serves as a board member of Autodesk, a world leader in 2D and 3D design, engineering and entertainment software.

Before joining Nokia, Mary served 17 years at HP-Compaq, after joining as a systems engineer in 1986. She was Senior Vice President and General Manager of the Industry Standard Server Group at Hewlett Packard and Compaq for five years, leading a multi-billion dollar business and the world's largest server franchise.

She holds a bachelor's degree in computer science from the College of Engineering at the University of Illinois. She serves on the college's Board of Visitors.

distinguished educator award



BAOCHUN

Li:

Baochun Li received his B.Engr. degree from the Department of Computer Science and Technology, Tsinghua University, China, in 1995 and his M.S. and Ph.D. degrees from the Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, in 1997 and 2000.

Since 2000, he has been with the Department of Electrical and Computer Engineering at the University of Toronto, where he is currently a Professor. He holds the Nortel Networks Junior Chair in Network Architecture and Services from October 2003 to June 2005, and the Bell Canada Endowed Chair in Computer Engineering since August 2005. His research interests include large-scale distributed systems, cloud computing, peer-to-peer networks, applications of network coding, and wireless networks.

Dr. Li was the recipient of the IEEE Communications Society Leonard G. Abraham Award in the Field of Communications Systems in 2000. In 2009, he was a recipient of the Multimedia Communications Best Paper Award from the IEEE Communications Society, and a recipient of the University of Toronto McLean Award. He is a member of ACM and a senior member of IEEE.



JANE W.-S.

Liu:

Jane W.-S. Liu received her BSEE degree from Cleveland State University and her Sc.D. from MIT. She is currently a Distinguished Visiting Fellow of Institute of Information Science of Academia Sinica and Bill Benter Honorary Chair Professor of Computer Science at National Tsing Hua University in Taiwan. Before joining Academia Sinica in 2004, she was a software architect in Microsoft Corporation from 2000 to 2004 and a faculty member of Computer Science Department at the University of Illinois at Urbana-Champaign from 1972 to 2000.

Her research interests are in the areas of real-time and embedded systems. In addition to over 160 journal and conference publications, she has also published two text books, one on real-time systems and the other on signals and systems. Her recent research focuses on technologies for building user-centric automation and assistive devices and services and for disaster preparedness and response. While at Illinois, she advised over 30 Ph.D. and over 100 MS students. She continues to supervise graduate students in Computer Science at National Tsing Hua University and recently graduated her first Tsing Hua University Ph.D. student.

She received the Achievement and Leadership Award of IEEE Computer Society, Technical Committee on Real-Time Systems in 2005; Information Science Honorary Medal of Taiwan Institute of Information and Computing Machinery in 2008 and Linux Golden Penguin Award for special contributions of Taiwan Linux Consortium in 2009. She is a fellow of IEEE.

C. L. Liu:

C. L. Liu received his B. Sc. degree at the National Cheng Kung University in Taiwan and his S. M. and Sc. D. degrees from the Massachusetts Institute of Technology. His teaching career spans more than forty years, at MIT, the University of Illinois at Urbana Champaign, and the National Tsing Hua University, where he is now the William Mong Honorary Chair Professor of Computer Science. His academic administrative duties include serving as Associate Provost at the University of Illinois at Urbana Champaign from 1996 to 1998, and as President of the National Tsing Hua University from 1998 to 2002.

His research areas are in computer-aided design of VLSI circuits, computer-aided instruction, real-time systems, combinatorial optimization, and discrete mathematics. He has published over 180 technical papers, and 8 technical books. In addition, he has published three books which are essay collections.

He received the Institute of Electrical and Electronics Engineers (IEEE) Millennium Medal, and the IEEE Circuits and Systems Society Golden Jubilee Medal in 2000. He was also the inaugural winner of the IEEE Computer

Society, Real Time Systems Technical Committee 1999 Technical Achievement Award, received for his contributions in the area of real time scheduling. He received the IEEE Circuits and Systems Society 1998 Technical Achievement Award for his contributions in the area of computer aided design of VLSI circuits, and an Outstanding Talents Foundation Award in 1998. He is the recipient of the 1994 IEEE Education Medal. He also received the Taylor L. Booth Education Award from the IEEE Computer Society in 1992, and the Karl V. Karlstrom Education Award from the Association for Computing Machinery in 1990. In 2004, the University of Macau awarded him an honorary doctorate and the National Chengchi University awarded him an honorary doctorate in 2011. He will receive the 2011 Phil Kaufman Award for Distinguished Contributions to Electronic Design Automation (EDA).

He serves on the Boards of a number of high tech companies and educational and charitable foundations in Taiwan, Hong Kong, and the US. Since 2005, he hosts a weekly radio show on Technology and Humanities in the radio stations IC975 in Hsinchu, Bravo in Taipei, and Mradio in Taichung.

He is a member of Academia Sinica, and a Fellow of IEEE and ACM.



JOSÈ Martínez:

Josè Martínez is associate professor of electrical and computer engineering and graduate field member of computer science at Cornell University. His research area is computer architecture, and his work has earned several awards; among them: two IEEE Micro Top Picks papers; a HPCA Best Paper Award; a NSF CAREER Award; and two IBM Faculty Awards. On the teaching side, he has been recognized with a 2005 Kenneth A. Goldman '71 Excellence in

Teaching Award, as a 2007 Merrill Presidential Teacher, and as the 2011 Tau Beta Pi Professor of the Year in the College of Engineering.

Prof. Martínez earned MS (1999) and Ph.D. (2002) degrees in computer science from the University of Illinois at Urbana-Champaign; his adviser was Prof. Josep Torrellas. He is a senior member of the ACM and the IEEE, Associate Editor in Chief of IEEE Computer Architecture Letters, as well as Associate Editor of ACM Transactions on Computer Architecture and Code Optimization.



distinguished service award



DAN

Peterson:

Dan Peterson is a Product Manager on Google+ with a passion for developer platforms and open source projects. Dan was a member of the Google Wave team, where he grew the developer community and promoted the federation protocol. Prior to that, Dan helped establish the OpenSocial specification, Apache Shindig, co-founded the OpenSocial Foundation, and guided the Google Web Toolkit (GWT) team as it became an open source project. Dan began at Google on the core infrastructure team, contributing to web search and data center management. Dan earned a B.S. in Computer Science from the University of Illinois at Urbana-Champaign, as well as minors in Technology & Management and Philosophy. Dan is also active on the Department of Computer Science's Executive Advisory Council and an occasional angel investor.



IRA

Cohen:

Ira Cohen received his undergraduate degree in Computer Science from the College of Engineering in 1980. He co-founded Advanced Systems Concepts in Schaumburg, Illinois in November of 1982 and helped to author software involving database systems and programmer productivity aids. The company was sold in 2006. Ira has remained active in various charitable pursuits including the American Cancer Society and the Jesse Owens Foundation. He also has remained connected to the University through the Computer Science Department and College of Engineering and is a current member of the University of Illinois Foundation.

memorial achievement award



ERICH WILLIAM

Hauptmann:

Erich William Hauptmann, 25, received a B.S. degree in Computer Science in 2008. His employment with Digital Domain, the award-winning special effects studio in Venice, CA was a dream come true for him.

He worked on *The Curious Case of Benjamin Button*, *GI Joe: The Rise of the Cobra*, *Transformers: Revenge of the Fallen*, *Star Trek*, *Speed Racer*, and *2012*. He is credited as Technical Assistant for *Percy Jackson* and the *Olympians*:

The Lightning Thief. Shortly before his death he was promoted to Assistant Technical Director of the Commercials division and participated in the production of the first 3D cinema spots ever created for a consumer electronic brand, the Samsung LED TV.

One of Erich's managers said: "Erich was able to cross the line between artist and engineer, teacher and student, colleague and friend. ... [He] earned the respect of his colleagues by his sheer motivation, proving that he could tackle any problem thrown his way, no matter how difficult."

DONALD Gilles:



Professor Donald B. Gillies, a native of Canada, did his undergraduate work at the University of Toronto, and received his Ph.D. in mathematics from Princeton University in 1953. While in graduate school he worked as a graduate assistant at the Institute for Advanced Study with John von Neumann in the fields of game theory and computer science. Before coming to the University of Illinois in 1956, he spent two years with the National Research Development Corporation at Cambridge University and London, England. He was among the first mathematicians to become involved in the computer field, helping to program the first Sputnik orbit and later discovering three new prime numbers in the course of checking out Illiac II. Before his death in 1975, he was experimenting with educational uses and networking possibilities of minicomputers.

Professor Gillies was an inspiration to his students, taking an interest in both their professional and personal lives. Long before timesharing terminals, minicomputers and microprocessors made “hands on” computer experience commonplace, he recognized the need for students to have this opportunity and implemented a system to provide it. Throughout his work and teaching he stressed the importance of the ethical use of computing machines in contemporary science. Dedicated to the honest uses of technology, environmentally concerned, a man of wit, vigor and understanding, he challenged and stimulated all who knew him.

The Donald B. Gillies Lectureship in Computer Science was established in the department of computer science in remembrance of his legacy. The lectureship continues to enrich the lives of students and colleagues as an appropriate memorial to a man whose intellectual excellence and moral purpose made him a distinguished teacher and scientist.

SABURO Muroga:



Professor Emeritus Saburo Muroga was one of Japan’s computer pioneers and a globally significant leader in the extensive field of information processing, and he taught and mentored generations of computer science researchers.

Professor Muroga was a pioneer in threshold logic, and was the author of the classic book on the topic “Threshold Logic and Its Applications”, published in 1971. The book enjoyed a renewed interest in recent years, as researchers of neural networks recognized its relevance to their field as well.

Muroga’s research in threshold logic was directed at minimizing the complexity of networks that would still be able to support high-level performance by, for example, minimizing the number of logic gates, interconnections among gates, or number of levels in a logic circuit. His revolutionary thinking led also to the creation of the “transduction method”, representing a new method for simplifying logic circuits based on permissible functions. The transduction method was adopted by major CAD companies and is now considered an industry standard.

Muroga also published widely on improving design automation using mathematical approaches and computer-aided design of VLSI chips.

Muroga joined the computer science faculty at the University of Illinois in 1964. While initially he planned to spend only three years at Illinois before returning to Japan, in fact, he remained at the university teaching and conducting research for 38 years until his retirement in 2002.

In 2004, Muroga was honored by his homeland, receiving the Order of the Sacred Treasure from the Office of the Emperor. His award, the “Gold Rays with Neck Ribbon” was given in recognition of his contributions to the area of information processing, and to the industry of computing in Japan.

RAY

Ozzie:

Alumnus Ray Ozzie (B.S., '79) was inducted into the College of Engineering Hall of Fame, October 14, 2011. Ozzie is a software visionary and entrepreneur who influenced communications and productivity in business, first as creator of LotusNotes and founder of Groove Networks, and most recently as Chief Software Architect of Microsoft.

Raymond E. Ozzie is an independent software entrepreneur and pioneer in social productivity, an area widely known in the field as computer-supported cooperative work. Through late 2010, he was Chief Software Architect of Microsoft, the company's most senior technical strategy and architecture role previously held by Bill Gates.

Ozzie came to Microsoft in 2005 through the acquisition of Groove Networks, a company he founded in 1997 to focus on software and services for small-team dynamic collaboration. Prior to Groove, in 1984 Ozzie founded and led Iris Associates, the creator and developer of Lotus Notes. A decade later, Iris was acquired by Lotus and then by IBM. Under his leadership during that period, Lotus Notes grew to be used for communication and social productivity by hundreds of millions at most major enterprises worldwide. Before creating Notes, he worked on 1-2-3 and Symphony at Lotus, on VisiCalc and TK!Solver at Software Arts, and on operating systems at Data General.

Ozzie studied computer science and engineering at the University of Illinois. While earning his degree, he also worked as a systems developer on the seminal PLATO project. He credits that work with helping him to understand the significance of online community and social interactive systems.

Honored as one of seven Windows Pioneers by Microsoft, Ozzie was named "Person of the Year" in 1995 by PC Magazine, and he has been inducted into the Computer Museum Industry Hall of Fame. He was selected for an Engineering at Illinois Alumni Award for Distinguished Service in 1997. He received the 2000 Institute for Electrical and Electronics Engineers Computer Society's W. Wallace McDowell Award and the 2005 SDForum Visionary Award. In 2001, he was honored as a World Economic Forum Technology Pioneer.

He is a member of the National Academy of Engineering and was named a Fellow of the American Academy of Arts and Sciences last year.

Ozzie has served as a member of the National Research Council's Computer Science and Telecommunications Board, and he was a member of the NRC committee that produced the landmark CRISIS report on the societal impact of cryptography.



Illinois CS Alumnus Ray Ozzie To Receive Honorary Degree

BY JEFF UNGER, UNIVERSITY OF ILLINOIS NEWS BUREAU

Raymond Ozzie, former chief software architect for Microsoft, has been chosen to receive an honorary doctor of engineering degree at the 2 p.m. campus-wide Commencement ceremony May 13 at the Assembly Hall.

In the early 1990s, his creation, Lotus Notes, was the first networked groupware application for the personal computer, revolutionizing business computing. In 1994, he was named one of seven "Windows Pioneers" by Microsoft because of the impact he and Lotus Notes had on the development of the Windows operating system.

In 1995, Ozzie, was named "Person of the Year" by PC Magazine. Ozzie earned a bachelor's degree in computer science at Illinois in 1979.

In 2000, Ozzie was awarded the IEEE Computer Society's W. Wallace McDowell Award for his "vision, determination, and programming skill in the development of Lotus Notes, a program that enables groups of people to collaborate over computer networks."



ROBERT

Bocchino:

Alumnus Robert Bocchino (Ph.D. 2010) was awarded the ACM SIGPLAN Outstanding Doctoral Dissertation Award for 2010. The award honors Bocchino's Ph.D. thesis, entitled "An Effect System and Language for Deterministic-by-Default Parallel Programming." The Outstanding Doctoral Dissertation Award is "presented annually to the author of the outstanding doctoral dissertation in the area of Programming Languages" and includes a prize of \$1,000.

Bocchino's research, advised by Prof. Vikram Adve, describes the design and evaluation of a new, Java-based object-oriented parallel language called Deterministic Parallel Java (DPJ). DPJ is a safe and modular parallel language that helps developers port (parts of) sequential Java applications to run on multicore systems. The parallel language is the first to guarantee deterministic semantics without run-time checks for general-purpose, object-oriented programs. It is also the first language to use compile-time type checking for parallel operations on arrays of references ("pointers") to objects, and the first language to use regions and effects for flexible, nested data structures.

Bocchino's current research focuses on using and extending the Plaid programming language, together with related verification techniques, to design and verify abstractions that make it easy for programmers to write correct and efficient parallel code.

BARRY

Greenstein:

Ace on the River: Alumnus Named to Poker Hall of Fame

Alumnus Barry Greenstein (B.S., '75), a renowned poker player and prominent fixture of some of the game's biggest events at the highest stakes, has been inducted into the Poker Hall of Fame during the World Series of Poker in November 2011.

A student of the game since childhood - his father taught him to play, and helped him develop his analytical approach to the game - Greenstein played his way through his undergraduate years in computer science, completing his degree in just 3 years. After graduation, Greenstein went on to a Ph.D. program in mathematics.

Greenstein was one of the initial 4 employees at a start-up company now known as Symantec, where he worked on their first product, Q&A. The product was widely hailed at the time of its launch, garnering Product of the Year awards in numerous industry publications. Greenstein still credits his work on Q&A as one of his proudest accomplishments.

In 1991, he left his job at Symantec, to pursue his professional poker career. Since then, Greenstein has won two World Poker Tour titles and 20 WPT cases, and is a three-time World Series of Poker bracelets, which he's won in three different poker disciplines.

Greenstein wrote "Ace on the River," a best-selling poker strategy book. To this day, those who bust him out in a tournament receive an autographed and personalized copy of the book, according to ESPN.

Known as the "Robin Hood" of poker, Greenstein is as well known for his charitable activities as for his skill on the felt.



PHOTO COURTESY OF VIRGINIA TECH

WU-CHEN FENG:

Alumnus Wu-chen Feng (Ph.D., '96) unveiled HokieSpeed, an energy-efficient supercomputer designed for the masses, at Virginia Tech University in December of 2011. HokieSpeed a single-precision peak of 455 teraflops, or 455 trillion operations per second, and a double-precision peak of 240 teraflops, or 240 trillion operations per second.

HokieSpeed's energy-efficiency makes it the highest ranked commodity supercomputer in the United States on the Green500 list, according to Virginia Tech.

"HokieSpeed is a versatile heterogeneous supercomputing instrument, where each compute node consists of energy-efficient central-processing units and high-end graphics-processing units," said Feng. "This instrument will empower faculty, students, and staff across disciplines to tackle problems previously viewed as intractable or that required heroic efforts and significant domain-specific expertise to solve."

Feng is an Associate Professor of Computer Science at Virginia Tech, who's primary research interests are in computer architecture, systems software, and middleware, and applications software. He was named a Person to Watch by HPCWire in 2011 and is the recipient of three R&D 100 Awards in green supercomputing, high-speed networking, and bioinformatics.

TIANYI Wu:

Alumnus Tianyi Wu was named the runner-up for the 2011 ACM SIGKDD Ph.D. Dissertation Award for his thesis "A Framework for Promotion Analysis in Multi-Dimensional Space".

Wu's work involves the use of a novel class of data mining problems, called promotion analysis, for marketing purposes. His thesis discusses how promotion analysis can be used for promoting a given product in a multi-dimensional space by leveraging object ranking information. Wu's key observation is that while most products may not be highly ranked in the

global space, where all products are compared by all aspects, there often exist interesting and meaningful local spaces in which the given product becomes prominent. Therefore, Wu's goal is to break down the data space and discover the most interesting local spaces in an effective and efficient way.

The promotion analysis problem is highly practical and useful in a wide spectrum of decision support applications, says Wu. Typical application examples include merit discovery, product positioning and customer targeting, object profiling and summarization, identification of interesting features, and explorative search of objects.

Wu currently works as an Applied Researcher at Microsoft, focused on the Bing technologies. His Ph.D. advisor was computer science professor Jiawei Han.

did you know?

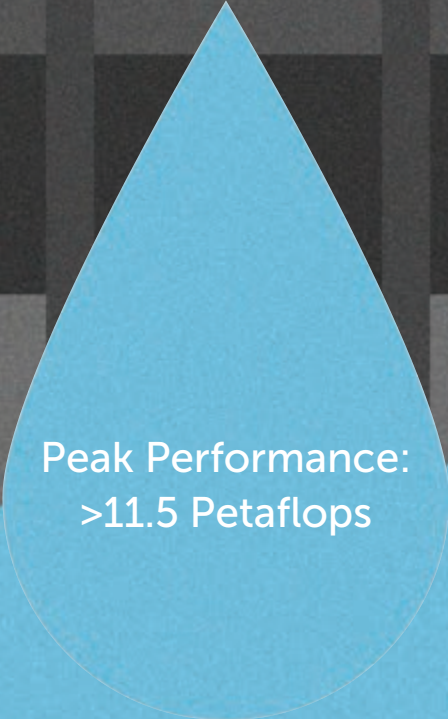
Blue Waters Stats:



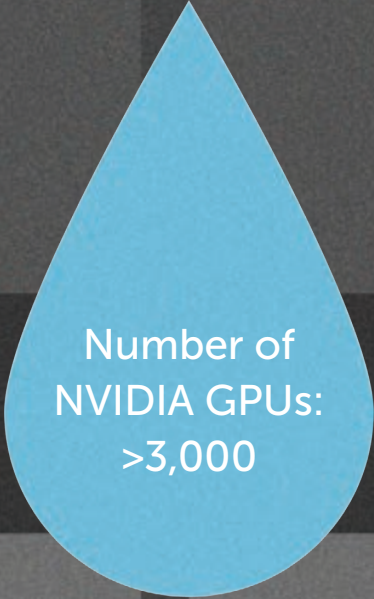
Memory:
>1.5 Petabytes



Number of cores:
>380,000



Peak Performance:
>11.5 Petaflops



Number of
NVIDIA GPUs:
>3,000

Q&A with Bill Gropp



Professor Bill Gropp is the Paul & Cynthia Chair in Computer Science, the Director of the Parallel Computing Institute, and co-PI of Blue Waters

How would you describe the overall vision of parallel computing that Illinois is pursuing with its research? What's the big picture?

Maybe the best way to describe the Illinois vision overall is that we are pursuing parallel computing research that is going to have a lot of impact on our ability to solve problems of all kinds. If you look at the different kinds of research here – we have the supercomputing center (NCSA), which is working on delivering computing resources that are delivering scientific results; we have a broad and deep application community that uses parallel computing every day, both in our department and in the computational science and engineering program. In CS and ECE, you have more fundamental and basic research into the science of parallel computing. This is one of the things we are trying to organize with the Parallel Computing Institute. PCI will provide a community within Illinois to broaden the scope of what individual researchers are doing. It's things like this that are one of the ways Illinois is unique in parallel research.

What does make Illinois such a unique and interesting place to do parallel computing research?

What's unique about Illinois is that it has all of this activity in one place. There is no other place that has such a range of activity in parallel computing, from commodity to extreme scale.

At Illinois, we will have one of the world's fastest machines (Blue Waters) – and in fact for real science it's likely to be the fastest. A major reason Illinois was

selected to build Blue Waters is because of the close connections with the faculty involved on the projects, and the programs that exist to deliver the applications for the system (IACAT and the NCSA Fellows program). There are tremendous opportunities to be working closely at the extreme edge of parallel computing and the applications of it. There are broader applications at mid-range, and all the way down to the important and challenging aspect of parallel computing on the smaller scale, like on consumer devices. Soon, every person on the planet will have experience with parallel computing, whether it's on their phone or their toaster. And Illinois is working on that entire range.

Parallel computing is both getting bigger – from the peta- and exascale – and getting smaller – desktop machines, mobile phones, and toasters. How can we tackle both sides of the problem?

There are a lot of issues that are common. After all, Blue Waters is made up of state-of-the-art processors and state-of-the-art GPUs. So, the issues of how does one program these kinds of systems, what are the algorithms, where are the hardware architectures going when we look at the next generation of extreme scale high performance computing – those questions are likely to have much in common with work that will be done on the smaller scale of parallel computing.

Whether or not exascale machines are built out of commodity components, without question what is built will be very much informed by the direction taken by commodity computing. Extreme scale computing systems would then build on that. But what distinguishes an extreme scale system from a

smaller scale one, beyond the issue of scale – is the interconnect, the ability to move the data between nodes on the system. So, there's a lot of commonality, but there are also a lot of areas to explore that are distinct from extreme scale.

This is why the various efforts on campus, PCI and NCSA, etc., have a lot that they can share, because a lot of pieces are the same, but they also have things that are distinct.

As an example, one question is whether there should be a uniform programming model, or whether it should be built out of a hierarchy of programming models. Part of the challenge is figuring this out.

A programming model that runs on individual chips or GPUs, may be the same model as the one used by people on the extreme scale machines. Extreme scale needs another model to connect everything together, as opposed to having one entirely different model that only used on extreme scale.

We will certainly have a model that deals with issues at scale, which is something that MPI does. We're looking at what's next for MPI, which is something that is unlikely to be of much interest at the small scale, though it will remain critical for midrange clusters.

What are the primary challenges in expanding the scope and applications of parallel computing?

The obvious one is programmability: being able to program these systems, addressing the software issues. This is the one that everyone talks about. I think we might be looking at slightly the wrong question there – I'll get back to this issue in a moment.

However, we also need to look at algorithms, because a lot of the algorithms we have for these problems can be scaled to tens, hundreds, or thousands of cores, but there is a point where they stop scaling. We need algorithms that can scale to 100s of millions or even billions of cores, and we don't have that problem solved. That's another reason Illinois is pretty unique in parallel computing; we have not only strong computer scientists looking at programming models, but also very strong groups looking at numerical and non-numerical algorithms in a parallel context.

With regard to software models, we use parallelism because we can't get enough performance out of our computers. So, the issue really is: how do we get more performance. Maybe that's through parallelism, but

that also might mean better algorithms, better use of the hardware, better ways to get concurrency. Parallelism is an incredibly powerful way to get to that concurrency, but it's really the performance we're after. So rather than focusing on how to get more parallelism into current systems, we should also be looking at how do we get to better performance.

That includes not just parallel programming, but better algorithms, a better understanding of how to get real performance out of real machines. To me, that's a little more fundamental. Parallelism is a fundamental part of that, but we primarily tend to think of parallelism as delivering performance. We need to make it easier for people to make use of the tool of parallelism, but people also need to have the right algorithms and they need to make better use of existing hardware.

Again I think that's something that Illinois can really lead on, because we're so good at applications, and we have so much depth in application areas. We are leaders in parallel computing -we have been historically and we will continue to be – but I think we will continue to be leaders because we understand where parallelism fits, and how to get there. We understand how to use computing to get what you need. It probably means parallelism, but it also means making effective use of the hardware, getting the algorithms right, and more. Understanding and addressing the breadth of the problem is something we probably do better than anybody.

How are applications of parallel computing informing the research at Illinois?

The classic way is that the application areas provide the challenge. If I've got this application and I can't run it, I need some way of getting more performance. So, we can look at the application and say, "OK, for the sort of operations you're performing, the compiler isn't generating the best code, so what are the techniques we can use to fix that. Do we need a better way of expressing what your application is doing?" Or, the perhaps the problem is scale - the work becomes unbalanced – and maybe we need to find better ways of adaptively managing that.


CHARM++ is a wonderful example of collaboration between the needs of the application community and computer science to solve the source of these problems. As you increase the scale, what changes do we need to make? Maybe we need new algorithms to handle how we solve enormous sets of equations.

So, the applications provide lots of challenges in terms of how to solve these problems. Not just how to get this to scale, but what do I need to change in order to reach the performance limits of my machine. Without the applications, it's hard to know what are going to be really important research questions. The applications provide us this continuous challenge to do more, to do better, to keep reaching further for what we can do.




image courtesy of NASA

PARALLEL COMPUTING AT ILLINOIS



With the computing world in the midst of a parallelism revolution, researchers are seeking to take advantage of unprecedented new speed and power potential for applications ranging from astronomy to zoology.

But meeting challenges in new programming technologies, power consumption, and performance for scientific applications will require major interdisciplinary engineering efforts. At Illinois, researchers are investigating every aspect of parallel computing, from peta- and exascale machines to commodity devices like smartphones and more.



enabling
high-impact
collaborations and
providing access to
cutting-edge
resources



PARALLEL COMPUTING INSTITUTE

The Parallel Computing Institute (PCI) is designed to enable Illinois researchers from across campus to come together in new, application-focused research centers and achieve their scientific goals using the latest and most efficient parallel computing technologies.

Led by Illinois computer science professor Bill Gropp, PCI will facilitate their work by providing an incubation and growth mechanism for new interdisciplinary centers and initiatives that will benefit from parallel computing; by expanding access to computing resources and infrastructure; by teaching critical skills; by identifying more opportunities for funding; and by helping to establish key external partnerships. The ultimate goal is to advance Illinois research capabilities by enabling high-impact collaborations and providing access to cutting-edge resources and professional instruction and assistance—all built around the revolutionary new possibilities offered by parallelization.

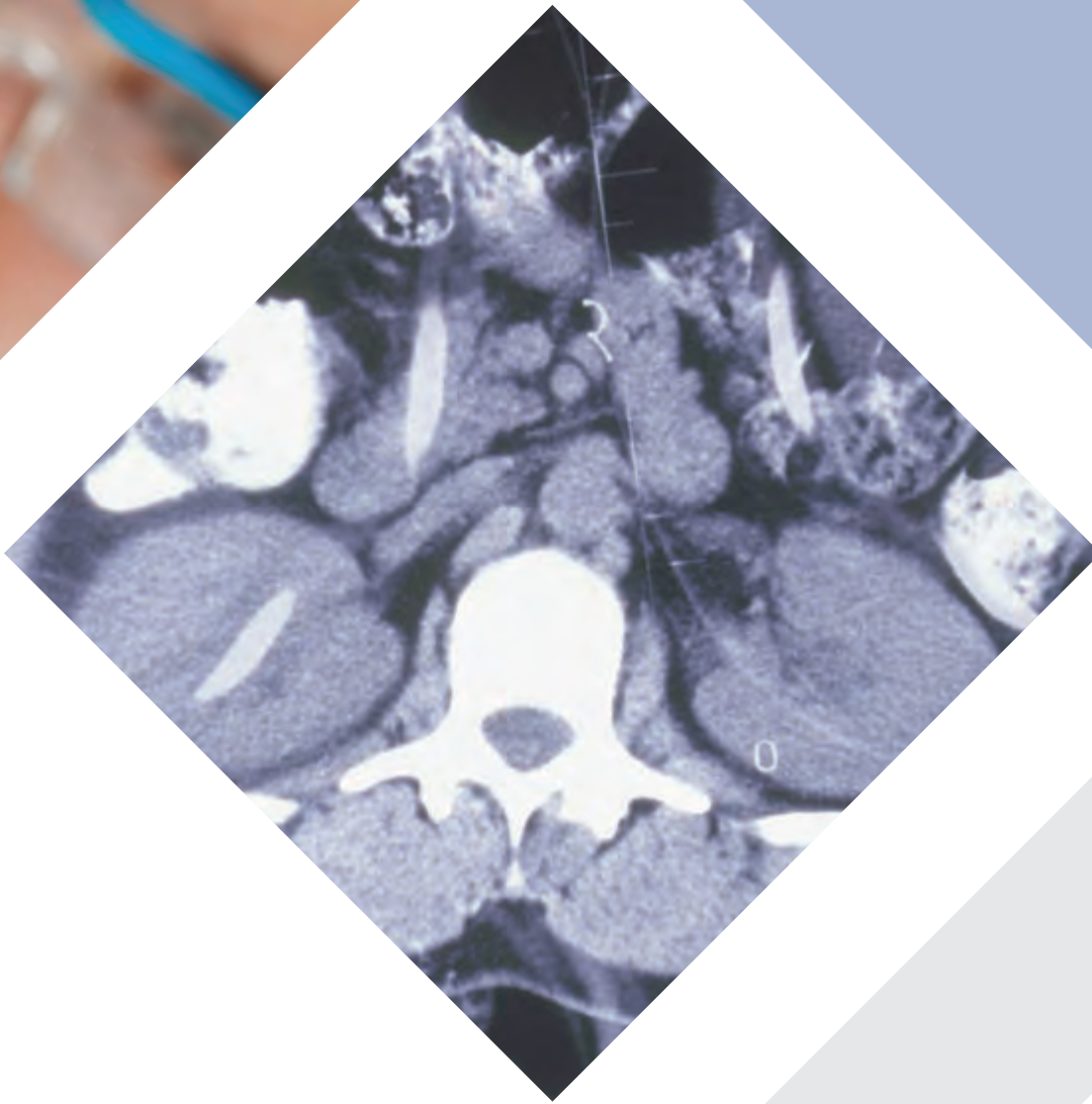
Research thrusts include:

INPATIENT MRI

Through the Illinois Massively Parallel Acceleration Toolkit project, researchers have homed in on a way to speed up the acquisition and reconstruction of MRI images while maintaining accuracy. The project, which draws on Brute Force and Toeplitz technologies, has enabled MRI processing in hours instead of weeks. The research is sponsored in part by the National Institutes of Health.

DNA SEQUENCING

DNA sequencing, an important process in the bioengineering field, is largely used in gene analysis, extraction, and adjustment. However, Illinois researchers believe it could also be a powerful tool in identifying disease. This research thrust explores ways to speed up the de novo genome assembly process using high-performance GPUs. In practice, clinicians could take a sample of DNA and run it through a program of coded diseases to find a match.



PRODUCTIVE HETEROGENEOUS EXASCALE COMPUTING

The “Vancouver Project,” funded by the Department of Energy, is designing a next-generation software infrastructure for productive heterogeneous exascale computing. The project uses GPUs to address issues of realized performance, scalability, and programming productivity in scalable heterogeneous computing (SHC). With attention to issues of energy-efficiency and resiliency of these systems, the goal is to provide tools and infrastructure for the transition to SHC.

DATA LAYOUT TRANSFORMATIONS

As applications require increasing amounts of memory, researchers are exploring new ways to store data. Because data layout is not usually considered during programming, data are frequently saved in a default layout, which is inefficient because the data the programs need at a particular time may not reside nearby in the memory space. This research seeks to change how programs save data, focusing on how it will later be accessed. By matching the code to data layout, data can be accessed more quickly and programs will run faster.

SCALABLE, IMPLICITLY PARALLEL PROGRAMMING

As programming shifts from sequential to parallel modes, researchers want to modify the way coding is accomplished. But instead of requiring programmers to learn an entirely different method of coding, researchers suggest that it is more practical to modify sequential coding techniques. With this approach, programmers can more easily spot problems and don’t need to relearn their trade. Illinois research is focused on creating programs that run on parallel systems in order to fully utilize all of a machine’s resources. This method is termed “implicitly parallel” because the coding presents itself externally as sequential code, though it is actually parallel.



photo: S. Severinghaus

developing
technologies for systems
ranging from terascale portable
devices to peta & exascale
data centers

Enabling programmable, high-performance, and highly power/energy efficient many-cores is a major step toward solving the true grand challenge of popular parallel programming. It is key to our industry's ability to build future-generation computers, keeping our nation competitive in IT and, therefore, ensuring our future quality of life.

Current petascale architectures are hardly scalable. They consume several megawatts (MW) of power and waste about 20% of their capacity because they have typical MTBFs of a few days. Consequently, as we look toward building exascale machines, power/energy-efficiency and resiliency are key challenges. In addition, exascale machines need to support over one billion threads and operate with an even worse ratio of flops to memory bytes than today's already challenging 10/1. This makes concurrency and locality major challenges as well. Finally, providing a programmable environment—one that assists the programmer to express, reveal, and exploit an application's parallelism—is an additional key challenge.

CENTER FOR PROGRAMMABLE EXTREME SCALE COMPUTING

The Center for Programmable Extreme Scale Computing aims to create the next generation of computers that are 1000x more energy efficient, yet deliver the same performance and physical footprint of current computers. The Center aims to create hardware and software technologies that fulfill this mission that are applicable to the whole range of computing systems: from terascale portable devices to petascale departmental servers, and exascale data centers.

“We envision the extreme-scale many-core circa 2020 to be very high performance, and yet very power- and energy efficient,” said Josep Torrellas, director of the Center. “Designing a usable

many-core, however, introduces many challenges. The biggest one is programmability.”

Although programmability is a notoriously difficult metric to define and measure, we suggest it implies two things: (i) the system should be able to attain high efficiency while relieving the programmer from managing low-level tasks, and (ii) the system should help minimize the chance of (parallel) programming errors.

Research at the Center for Programmable Extreme Scale Computing is organized around three major efforts.

UBIQUITOUS HIGH PERFORMANCE COMPUTING: RUNNEMEDE

Funded by: DARPA

As part of the \$49 million, Intel-led effort to develop a new platform for ubiquitous high performance computing, the Center’s researchers are investigating new computing and memory system architectures that aim to dramatically optimize energy efficiency. To this end, researchers will be leading efforts to create new architectures for fast synchronization and communication; innovative new memory hierarchy organizations; aggressive new techniques for power and resiliency management; and in-memory state versioning, among other advances.

In addition, the Center’s researchers are participating in the design of programming systems that facilitate coding and enable automatic program optimization. Very high-level APIs and compiler strategies that support static and dynamic optimization characterize the programming system of extreme-scale computer systems. The multiple optimization dimensions that must be addressed in extreme-scale computing—including execution speed, energy efficiency, and resiliency—complicate tuning and demand the support of novel and sophisticated compiler strategies.

THRIFTY: AN EXASCALE ARCHITECTURE FOR ENERGY PROPORTIONAL COMPUTING

Funded by: Department of Energy (DOE)

Thrifty aims to provide a platform for highly-efficient, energy-proportional computing—a platform where energy is spent proportionally to the work done, and in the resources that truly are the bottleneck. The Thrifty concept and its software stack simultaneously innovate in power/energy efficiency, resiliency, and performance. Each topic is tackled at multiple layers of the computing stack: circuits, architecture, compilation, and runtime/application.


The Thrifty research involves the development of a detailed architecture simulator of a 1K-core Thrifty system and its evaluation; the implementation and evaluation of a power-aware optimizing compiler based on ROSE; the development of libraries for application characterization and tuning; and the fabrication of a large test chip in 22nm technology by Intel to test novel circuit technologies.

The goal of this project is to attain exascale performance at about 20MW by year 2018–2020 at 8nm technology.

DESIGNING PROGRAMMABLE MANY-CORE FOR EXTREME SCALE COMPUTING

Funded by: National Science Foundation

This project aims to address the challenge of programmable extreme-scale computing through a multidisciplinary and ambitious effort that cuts across the architecture, compiler, operating system, and correctness/performance tools areas. The Center’s goal is to make fundamental advances toward designing a programmable many-core for extreme-scale computing in mobile platforms (netbooks and smartphones) by the year 2020. Researchers focus on a key concept in this work: that cores and all of the software continuously operate in chunks (i.e., atomic blocks) of instructions at a time—eliminating the need for in-order, single-instruction-at-a-time commit. To implement this scheme, researchers are working on innovations in architecture, compiler support, operating systems, and correctness/performance tools.



enabling
easy-to-program,
energy & power-efficient,
parallel portable
devices

ILLINOIS-INTEL PARALLELISM CENTER: MAKING PARALLELISM EASY

The Illinois-Intel Parallelism Center (I2PC) is advancing the state of the art in parallel applications, software, and architectures to enable easy-to-program, energy/power-efficient, parallel portable devices. I2PC is a joint research center funded by the Illinois Departments of Computer Science and Electrical and Computer Engineering, and the Coordinated Science Lab, with support from corporate sponsor – Intel Corporation.

Research is organized around three projects:

AVASCHOLAR

AvaScholar (a portmanteau of Avatar and Scholar) leverages parallelism to make internet-based education work better with new tools for the instructor and the learner. The AvaScholar Instructor module applies parallel 3-D capture and reconstruction software on the instructor and any visual aids in 3-D so they can be viewed from any angle remotely by the students. The AvaScholar Learner module applies new parallel machine learning “soft biometrics” techniques that use each student’s webcam to help the instructor gauge how well large classes of remote students understand the material (through expression recognition and shrug detection) across various automatically recognized demographic groups such as age or gender.

ACROBATICS

Acrobatics aims to develop basic technologies in programming, translation, architecture, and tools to enable the development and execution of high-performance, power-efficient parallel code. It relies on high-level abstractions for parallel programming, together with compiler optimization and program synthesis techniques, to generate code for architectures designed for programmability. The system is supported by a set of parallel programming tools for debugging, testing, refactoring, and performance/power monitoring.

SAFESPEED

The goal of SafeSpeed is to ensure safe and efficient parallel computing by utilizing explicit annotations. The SafeSpeed components utilize explicit annotations for disciplined shared memory: the programmer specifies the annotations that the software testing techniques and hardware design can exploit. These annotations are used by SafeSpeed components at all levels from specification to run-time. DPJ, DPJizer, Accord, and AMP focus on the specification and checking, Penelope is used for testing, and DeNovo supports parallelization at run-time.



How fast is a supercomputer?

BILL GROPP, PCI DIRECTOR

Cray has begun the installation of the Blue Waters supercomputer at the National Center for Supercomputing Applications at Illinois (see <http://timelapse.ncsa.illinois.edu/pcf/inside2/index.php> for live images of the installation).

Blue Waters will be the “Track 1” system for the National Science Foundation and will provide unprecedented computing power for the nation’s scientists and engineers. To provide such computational power, the system makes use of multiple levels of parallelism — large numbers of state-of-the-art multicore processors, connected with a high-performance interconnect, and including a modest fraction (a little over 10%) of nodes that include a graphics processing unit (thus most of the peak performance is in conventional multicore processors, while the GPUs can provide a significant performance boost). Also part of the system is a high performance I/O system, ensuring that moving data into and out of the system will not be a major bottleneck.

So how fast is Blue Waters? That depends on what you mean by fast. Perhaps the easiest measure is the peak performance in floating point operations per second (FLOPS). However, everyone in high performance computing knows that this number is at best a “guaranteed not to exceed” number and does not reflect the performance that is seen on applications. The next choice is the benchmark used for the Top500 list. However, this benchmark, which uses a direct algorithm to solve a large, dense linear system, is also widely understood to not reflect the performance that users of the system will see. Other benchmarks, including the HPC Challenge and Graph500 benchmarks, attempt to address these limitations by picking “better” benchmarks, but still do not typically represent the performance seen by applications.

A better measure of performance, though much more difficult to define and measure, is the sustained performance seen by the applications that run on the system. And even this is somewhat artificial (we all know how to increase the FLOPS achieved by an application by using algorithms and computations designed to increase compute intensity at the cost of computational efficiency). The real measure will be in what the system accomplishes — what new knowledge results from using Blue Waters.

So how fast will Blue Waters be? The Blue Waters team (of which I’m a member) is confident that the system will exceed a PetaFLOPS on a representative set of computationally efficient, extremely parallel applications, and we expect groundbreaking science results from the teams using the system. And that’s really all that counts.

briefs

See <http://timelapse.ncsa.illinois.edu/pcf/inside2/index.php> for live images of the installation



parallel computing



Encyclopedia of Parallel Computing

Professor David Padua served as Editor-in-Chief for a new encyclopedic effort to survey the concepts behind the significant shift towards parallel computing in today's computer industry. The 4 volume Encyclopedia of Parallel Computing contains more than 300 entries on topics related to the critical demand for continued advances in parallel programming. Contributors to the effort were leading international experts in the field, including Illinois computer science professors William Gropp and Marc Snir, professor emeritus David J. Kuck, and Jack Dongarra, Michael Flynn, David E. Shaw, and Guy L. Steele Jr. among others.

"Parallel computing has already impacted or will soon impact everyone who uses a computing device, from supercomputers to laptops, tablets, and smart phones," said Padua.

Today's supercomputers are massively parallel machines with thousands of processors. The fastest supercomputer today uses 705,024 processing cores, capable of 10.51 quadrillion calculations per second.

Ten years ago the world's fastest supercomputer used a total of 8,192 processing cores and was only capable of 12.3 trillion calculations per second, almost one thousand times less powerful. This type of accelerated parallelism is critical to science and engineering, enabling discoveries and designs that

would not be possible otherwise. For consumer and mobile devices, parallelism is the only viable strategy for continued performance gains, while also allowing chipmakers to optimize for energy efficiency.

With the need for parallelism at an all-time high, the Encyclopedia of Parallel Computing provides researchers and developers with an authoritative reference that pulls together the tools necessary to take advantage of these pioneering concepts.

"This monumental work will be an invaluable resource for practitioners and students in all areas of computer science," said Alex Nicolau of UC Irvine. "In today's world where parallel computing is ubiquitous--from desktops to cell phones and game consoles--this reference work will be indispensable."

Key concepts in the Encyclopedia for professionals, researchers and students of Parallel Computing include:

- Programming models and programming languages
- Debugging and race detection
- Laws of parallel computing
- Theoretical models of computation
- Supercomputer/high-performance computing machines
- Interconnection networks

Learn more about
Deterministic Parallel
Java at
<http://dpj.cs.uiuc.edu>



Rethinking Memory Hierarchy

Illinois CS researchers have won the Best Paper Award at the Parallel Architectures and Compilation Techniques (PACT 2011) conference for their paper, “DeNovo: Rethinking the Memory Hierarchy for Disciplined Parallelism.” The research team, led by Illinois computer science professor, Sarita Adve, is working on the DeNovo project that takes a new approach to building multicore hardware. DeNovo exploits emerging software trends in disciplined parallel programming to make hardware simpler, higher performance, and lower energy, all at the same time.

Most multicore programs use a shared memory programming model. Shared-memory programs have many advantages, but they are known to be difficult to program, debug, and maintain. At the same time, shared-memory hardware is complex and inefficient, leading to unnecessary energy consumption and performance bottlenecks. After decades of trying, researchers have found it difficult to even develop satisfactory shared-memory semantics for common languages such as C++ and Java. A recent article, co-authored by the DeNovo lead, calls the research community to rethink how we design both parallel languages and parallel hardware.

At the root of these problems is what the Illinois team refers to as “wild shared memory” behavior. Shared-memory programs tend to exhibit unstructured parallelism with implicit communication and side effects, leading to hard-to-debug data races and ubiquitous non-determinism. The Illinois team believes general-purpose languages must enforce more discipline and eliminate such wild behavior by design if parallel computing is to become tractable for mass programmers. Such a discipline would enforce more structured parallelism and make side effects of a parallel task be more explicit. Many software researchers today are working on such an approach, including pioneering work by an Illinois team on the

Deterministic Parallel Java (DPJ) language, led by Vikram Adve. The DeNovo team, working closely with the DPJ team, has shown that the same disciplined parallel programming features that simplify software can also enable more performance-, energy-, and complexity-scalable hardware. As their first step, they have developed a cache coherence protocol and consistency model that takes an order of magnitude less time to verify and runs some applications in less than half the time with less than half the network traffic and cache misses than the state-of-the-art. The simplicity and low network and cache traffic means that the performance increases come with significant power and energy benefits. It is rare in computer architecture that a hardware design improves complexity, performance, and power all at once.

According to Sarita Adve, “this paper is a first step towards an ambitious vision. While it presents significant new technical contributions that we hope will eventually be adopted, it also opens up many new questions. We hope that the largest impact of this paper will be in inspiring a broad research agenda anchored in a more disciplined approach to parallel systems. The paper motivates hardware research driven by disciplined programming models and also seeks to inspire architects to extend their influence on the development of such models. We believe this is an opportune time for such a co-designed evolution of the current hardware-software ecosystem.”

The paper was authored by a team of Illinois computer science graduate students and faculty, including Byn Choi, Rakesh Komuravelli, Hyojin Sung, Robert Smolinski, Nima Honarmand, Sarita Adve, and Vikram Adve, in collaboration with Nicholas Carter and Ching-Tsun Chou from Intel. The work was supported in part by Microsoft and Intel through the Universal Parallel Computing Research Center and the Illinois-Intel Parallelism Center, and by the National Science Foundation.





CS Course Leads to Groundbreaking Health Informatics Text

Bruce Schatz, an affiliate faculty member in Computer Science, has co-authored a groundbreaking book on Health Informatics based on his popular computer science course. The text is the first book combining the solutions of modern computer science with the problems of modern medical science. The book is expected to be a key reference for professionals working in health management, from information to healthcare executive, health information technologist to computer scientist, and physician to patient.

Healthcare Infrastructure: Health Systems for Individuals and Populations describes the new healthcare infrastructure that will gather personal health records from every individual and correlate each longitudinal record across whole populations. The book explains the problems of personal medicine and public health, then the solutions possible with information technology.

"What is needed [in health care] is a vast backbone, a health care infrastructure consisting broadly of health and deeply of medical information, which is recorded through personal sensors, analyzed on supercomputers, communicated through networks, and accessed through computers," says Schatz. His new text is about that infrastructure: who will use it, what problems it solves, where it will be used, why it chooses its designs, and how it works.

we wrote the books

Hoiem Authors Computer Vision Text

A new text in computer vision co-authored by professor Derek Hoiem introduces the subject of 3D object recognition and scene interpretation in depth. The text, *Representations and Techniques for 3D Object Recognition & Scene Interpretation* has a primary focus on recent efforts to fuse models of geometry and perspective with statistical machine learning.

The text aims to make the latest developments in 3D scene understanding and object recognition accessible to newcomers to the field. With existing research scattered across journal and conference papers, the subject was left to the purview of experts. Hoiem's text organizes this research and provides an historical and technical background so that newcomers to the field can learn about this emerging area.

"In recent years, the rigid, algebraic view of 3D geometry has given way to a more statistical, probabilistic view. In consequence, we've seen amazing new abilities to reconstruct 3D scenes and recognize 3D objects from a photograph," said Hoiem. "These technologies could have far-ranging impact, from robotics, to vehicle safety, to content creation and photo enhancement."

Har-Peled Authors Text on Geometric Approximation Algorithms

A new text on geometric approximation algorithms authored by professor Sarel Har-Peled is the first to cover the subject in detail. *Geometric Approximation Algorithms* describes key techniques in geometric approximation algorithms and also surveys some of the more traditional computational geometry techniques, such as sampling and linear programming.

The field of geometric approximation algorithms is a subfield of both computational geometry and approximation algorithms, explains Har-Peled.

"Exact algorithms for dealing with geometric objects are complicated, hard to implement in practice, and slow," says Har-Peled. "Over the last 20 years, a theory of geometric approximation algorithms has emerged. These algorithms tend to be simple, fast, and more robust than their exact counterparts."

What began as a collection of class notes on the subject was expanded by Har-Peled into his full length text.



**PPL
Celebrates
its 20th
Anniversary**
Parallel
Programming
Laboratory
(PPL) proudly
celebrated
20 years of
inspiring
computing
innovation
and enabling
inter-
disciplinary
collaboration.

by Hong Shen

For computer science professor Laxmikant (Sanjay) Kale, the wonderful thing about being a professor at Illinois is to “have the freedom to do what you want.”

In 1990, the year he started the Parallel Programming Laboratory (PPL), Kale wanted to do two things: first, to develop a parallel programming system that allows humans and computers to do what they can do best; second, to perform useful computer science by working with applications.

Over the past twenty years, the group has achieved both objectives with many accomplishments: a machine-independent parallel programming system called Charm++ and a variety of successful parallel applications.

This year, PPL is going to celebrate its 20th anniversary.

“Charm++ is our core technique,” said Kale. “It is one of those simple ideas that turns out to be a very powerful one.”

Charm++ is an object-oriented portable parallel programming language based on

C++ that includes an adaptive runtime system. Programs written using this system will run unchanged on MIMD machines with or without a shared memory. It provides high-level mechanisms and strategies to facilitate the task of developing even highly complex parallel applications.

“It simplifies parallel programming by separating the share of resource management,” Kale said. “So programmers don’t have to worry about when and which processor does what work, they only talk about what work is to be done in parallel.”

In some sense, the principle of Charm++ is to let the programmers do what they can do best and the system do what it can do best, Kale added.

“I was attracted to Illinois by Charm++,” said Eric Bohm, a senior research programmer at PPL. “I like the degree of freedom it gives to the application’s author. When you’re writing a Charm++ application, you’re not constrained in the design of the physical machine. You can focus on the fundamental algorithm expression of parallelism. So the machine and the programmer can have a nice separation of concerns.”

Over the years, Charm++ has become one of the most popular parallel programming languages in the supercomputing world. According to a 2007 survey, two of the major supercomputing centers in the U.S.—the National Center for Supercomputing Applications at Illinois (NCSA) and the Pittsburgh Supercomputing Center (PSC)—spent 15 percent and 10 percent of their time running Charm++ applications, respectively.

“It’s such a big and interesting parallel runtime system,” said Orion Lawlor, a PPL alumnus now working as an assistant professor at the University of Alaska Fairbanks (UAF). “Literally hundreds of students have worked on the system directly, over the last twenty years. We’ve also worked with dozens of outside groups to speed up parallel programs to record-breaking sizes.”

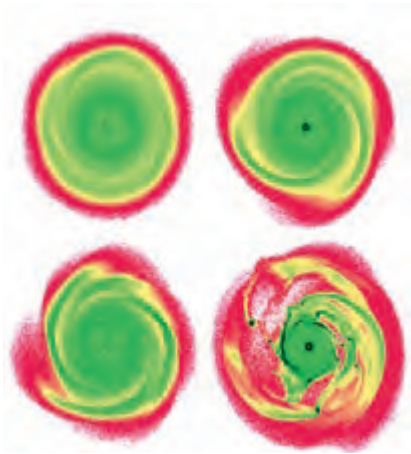
One of the “outside groups” that has close collaboration with PPL is the Theoretical and Computational Biophysics Group, led by Illinois physics professor Klaus Schulten.

“Our collaboration started in the 1990s,” Schulten said. “Sanjay and I worked together to develop software for biomedicine and one of the ideas was to use parallel computers for better computing in the field. The most successful application is called NAMD.”

NAMD is a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems.

“The idea is to do atom-by-atom simulation of large biomolecules and biomolecular systems to see how they interact and function. But it’s very challenging computationally because it involves a large number of steps in a very tiny amount of time,” said Kale. “It’s also challenging because every step has to be done completely before you can go to the next





“I am most impressed with PPL’s vision of enhancing performance on real applications.”

— Professor Tom Quinn, University of Washington

one—which means you have to parallelize each step efficiently.”

Based on Charm++ parallel objects, NAMD scales to hundreds of processors on high-end parallel platforms and tens of processors on commodity clusters using gigabit Ethernet.

In 2002, NAMD received the Gordon Bell award—one of the most prestigious awards to recognize outstanding achievement in high-performance computing applications.

“We worked very closely for 20 years and it’s a wonderful experience,” Schulten said. “We are actually an application scientist and a computer scientist working together for a very long time and having really excellent results. I think it’s a famous relationship and it’s well known beyond Illinois.”

To ensure relevance and long-term impact, PPL also works with other groups in the context of real applications. Besides NAMD, the collaborative research projects involve rocket simulation, cosmology (ChaNGa), quantum chemistry (OpenAtom), space-time meshes, and structural dynamics.

ChaNGa is a collaborative project between PPL and the N-Body Shop, led by professor Tom Quinn at the University of Washington.

“The overall project is computational cosmology and cosmogony: how galaxies form via gravitational collapse, and how planets form in the protoplanetary nebula,” said Quinn. “And both of these problems involve a complex interaction of gravitational collapse, hydrodynamic, and radiation processes.”

During the collaboration, the N-Body Shop provides computational cosmology problems and algorithms, such as the formation of structure in galaxy clusters and solar systems, and tree-based gravity and gas dynamics algorithms, while PPL uses the Charm++ system to help make that code run efficiently on large parallel machines with thousands of processors.

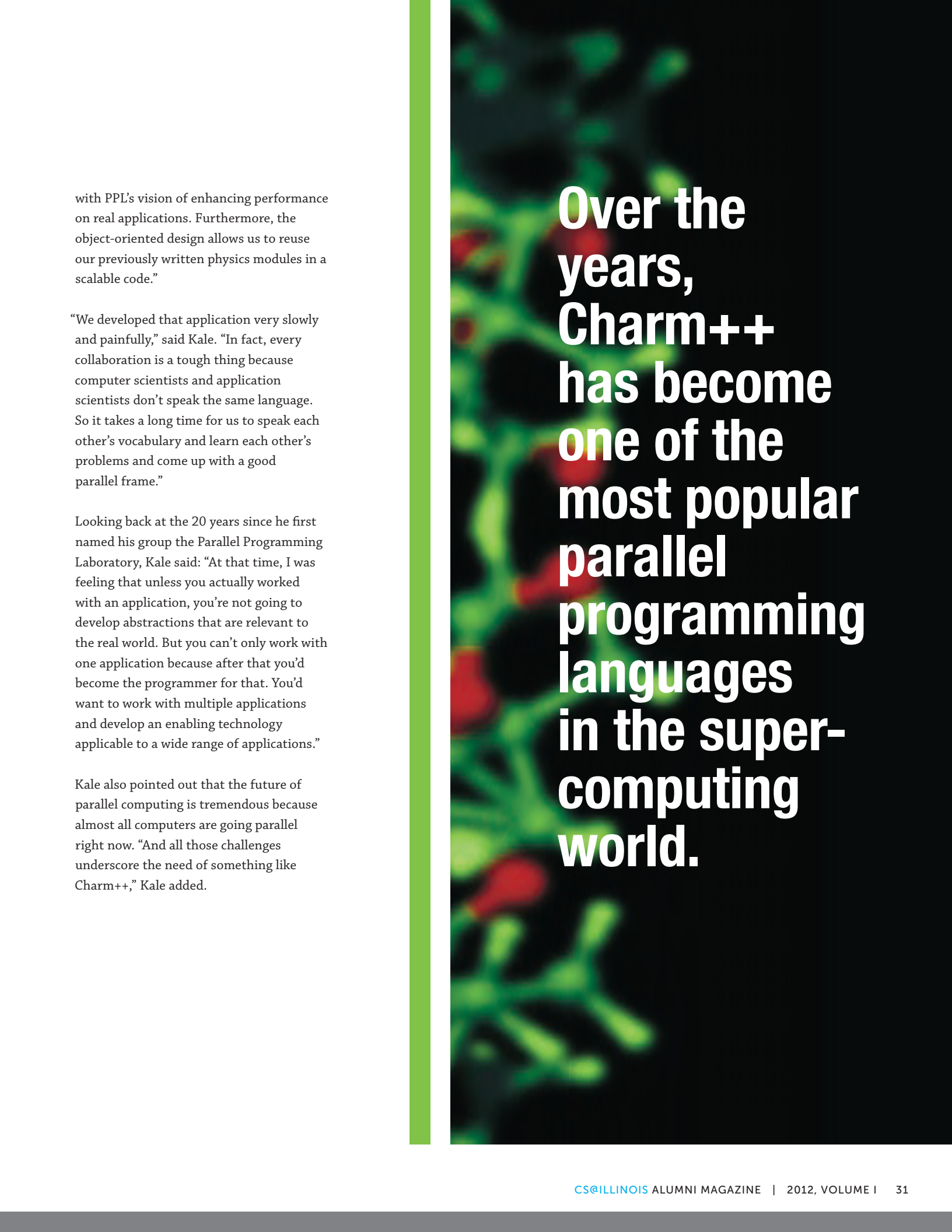
The project started in 2001. “At that time, my group was looking for ways to extend our work in large cosmological computations to scale better and be more flexible,” said Quinn. “I am most impressed

with PPL's vision of enhancing performance on real applications. Furthermore, the object-oriented design allows us to reuse our previously written physics modules in a scalable code."

"We developed that application very slowly and painfully," said Kale. "In fact, every collaboration is a tough thing because computer scientists and application scientists don't speak the same language. So it takes a long time for us to speak each other's vocabulary and learn each other's problems and come up with a good parallel frame."

Looking back at the 20 years since he first named his group the Parallel Programming Laboratory, Kale said: "At that time, I was feeling that unless you actually worked with an application, you're not going to develop abstractions that are relevant to the real world. But you can't only work with one application because after that you'd become the programmer for that. You'd want to work with multiple applications and develop an enabling technology applicable to a wide range of applications."

Kale also pointed out that the future of parallel computing is tremendous because almost all computers are going parallel right now. "And all those challenges underscore the need of something like Charm++," Kale added.



Over the years, Charm++ has become one of the most popular parallel programming languages in the super-computing world.



the first super



A computer science historian could add numerous tags to ILLIAC IV: the world's first supercomputer, the world's fastest computer in the 1970s, the first large-scale SIMD machine, the first general purpose parallel computer ... and many more. However, talk to most ILLIAC IV alumni and they can tell you: it's a computer that changed their lives forever.

ILLIAC IV: computer

by Hong Shen

Until today, Ahmed Sameh, Ph.D. '68 Civil Engineering, still felt he owed a great deal to a computer built more than 40 years ago.

"I owe the ILLIAC IV project a great deal of gratitude for introducing me to the field of parallel computing," said Sameh, now a professor of computer science at Purdue University. "After that, I left my Ph.D. research in structural mechanics/dynamics in favor of developing numerical linear algorithms for various parallel architectures."

That was in 1967. Sameh had finished most of the codes for his Ph.D. thesis in civil engineering. Running numerical simulations on the old IBM 360 was taking too much time and eventually, he got bored.

"One day, I saw a flyer in the Engineering Library seeking students' help in programming a new kind of computer," Sameh said. "I was intrigued by its concept of reducing the time of numerical simulations in many science and engineering applications and decided to apply for an hourly position."

of the art in digital system technology, architecture, and performance.

As a pioneer in "massively parallel" computing, ILLIAC IV had 64 processing elements working together, and it was designed to have up to 256. All processing elements executed the same instruction simultaneously, and each one had its own local memory.

The computer was the fourth in a line of computers built at the University of Illinois starting in 1948. It was moved to the NASA Ames Research Center in 1974, where its applications included fluid dynamics, digital image processing, and weather forecasting.

Only one was ever built, and it was finally retired in 1982. Nevertheless, from 1972 to 1982, ILLIAC IV was the fastest and most powerful computer on the face of the earth at that time, achieving record speeds for various numerical calculations at a peak of 200 million instructions per second and a one gigabit per second I/O transfer rate.

Yet, perhaps more importantly, it pioneered the new concept of parallel computation, stimulated the design of subsequent supercomputers including the Cray machine in the '70s, and created the school of parallel computing at the University of Illinois that has lasted for more than 40 years.

"From 1965 on, as a result of the ILLIAC IV program, first the Urbana laboratory and subsequently many other university, government, and industrial labs have undertaken the analysis of the relations between computer architecture, algorithms, both numerical and non-numerical, and their expression in programming languages," Slotnick said in a memoir written for the Annals of the History of Computing in 1982.

Talk to most ILLIAC IV alumni and they can give you a quick summary of how the world's first supercomputer was conceived at the University of Illinois, usually as a short personal story.

... from 1972 to 1982,
ILLIAC IV was the fastest
and most powerful computer
on the face of the earth ...

What Sameh discovered was ILLIAC IV—the world's first supercomputer, conceived at the University of Illinois in the mid-1960s. The aim of the project was to build a large parallel processing computer to cut down the time taken by computers in completing many of the computational science and engineering tasks. It was led by late computer science professor Daniel Slotnick as a research vehicle for advancing the state

David Kuck, an emeritus computer science professor, came to Illinois from MIT in 1965. He said there was some theoretical work going on in parallel computing at MIT at that time, but ILLIAC IV was the first opportunity he had to work on a real machine.

“ILLIAC IV was the most ambitious floating point parallel computer of its time,” said Kuck. “It was an experiment, but it still showed that something could be done. We were able to get a lot of parallel programming applications running on it and demonstrated that they actually worked well.”

As the sole software person on the ILLIAC IV project in contrast to all the other hardware-oriented members, Kuck was responsible for developing many of the initial ideas of how to restructure computer source code for parallelism.

“I spent the first year roughly thinking about architecture and how we could produce some applications and software that would run well on that kind of machine.” Kuck said.

For Kuck, one of the most influential features of ILLIAC IV is its architecture: ILLIAC IV was one of the first experiments in a large-scale SIMD (single instruction, multiple data) machine, and today SIMD is still in use—now Intel machines have small SIMD vector units inside them.

“The basic idea is very important,” said Kuck. “I’d say developing that idea included the control of the machine and the way the data was stored in the memory and flowed through the machine—that was probably the most important architectural feature of the machine and we demonstrated that software could actually be written for it.”

Kuck worked on the project from 1965 to 1968. After that, he created the Center for Supercomputing Research and Development (CSRD) at the University of Illinois and became one of the most influential figures in parallel computing, especially in productivity tools for parallel programming. Over the past four decades, he not only influenced a wide range of areas including architecture design and evaluation, compiler technology, programming languages and

algorithms, he also trained many of that field’s major players around the world.

Duncan Lawrie, one of the “major players” from Kuck’s group, also an emeritus computer science professor at the University of Illinois, said there were many difficulties when they first started.

“The biggest difficulties I saw were that we knew so little about parallel algorithms and applications. It was difficult to design languages and compilers when you didn’t have a good set of applications to guide you.” Lawrie said.

Lawrie joined the ILLIAC IV project as a graduate assistant, yet he designed and implemented the first high-level language for that machine: GLYPNIR.

“GLYPNIR is one of the earliest existing languages designed for programming the ILLIAC IV computer,” said Lawrie. “Several attempts were made to produce a higher-level language/compiler, but in the end we



translated Fortran into GLYPNIR (I think we may have called it ‘Fourtran’). The syntax of the language is based on ALGOL 60, but has been extended to allow the programmer explicitly to specify the parallelism of his algorithm in terms of 64-word vectors.”

For 10 years, GLYPNIR was the only working ILLIAC IV higher-level language, according to Slotnick.

“[ILLIAC IV] has inspired many minds and created the school at Illinois

— David Padua, Ph.D.

Computer science professor Mark Snir pointed out two major contributions of ILLIAC IV.

“Although parallelism is not a new idea in computer science, ILLIAC IV was definitely one of the first to explore massive parallelism,” said Snir. “First, it showed this kind of architecture is feasible—you can really leverage this level of parallelism and achieve good performance; second, it motivated a lot of scientists and engineers (many of them Illinois alumni) to do research on subsequent supercomputers, for example, the Cray machine in the ’70s and even today’s Blue Waters.”

Steve Chen, Ph.D. ’75 Computer Science, was one of those alumni that became a key player in the field of supercomputing.

“Steve was one of the first five students in my group,” said Kuck. “Just after ILLIAC IV, in 1971 I started a research program to work on parallel architectures that were driven by software needs—there were many things of the ILLIAC IV project done very intuitively without enough deep thought.”

After graduation, Chen joined Cray Research as its chief designer, where he led the development of the

world’s most commercially successful parallel vector supercomputers, the Cray X-MP and its successor, the Cray Y-MP. Chen began by making some architectural changes to the Cray-1, which was introduced in 1971. In the Cray X-MP, Chen introduced shared-memory multiprocessing to vector supercomputing. The machine contained two pipelined processors compatible with the Cray-1 and shared memory. The X-MP series was expanded to include 1- and 4-processor machines. The X-MP4 was the first supercomputer installed at the National Center for Supercomputing Applications (NCSA) at Illinois.

“Chen became very famous later on because Cray was the world’s fastest supercomputer for almost a decade,” said Kuck.

The New York Times claimed in an article that Chen was “considered one of the nation’s most brilliant supercomputer designers while working in this country for the technology pioneer Seymour Cray in the 1980s.”

For other people, another important contribution of the ILLIAC IV project was its influence on the University.

“... ILLIAC IV was definitely one of the first to

to work on parallel computing that has lasted for 40 years.”

photo: Steve Jurvetson

“At Illinois, we have one of the strongest groups in parallel computing in the U.S. and very active programs,” said computer science professor David Padua, Ph.D. ’80 Computer Science. “And the reason for that is not just an accident. It is ILLIAC IV. It has inspired many minds to work on parallel computing and created the school at Illinois that has lasted for 40 years.”

For Padua, ILLIAC IV was also the reason he came here. “At the beginning of the ’70s, I was very interested in computer system and machine design, and at that time, the most interesting computer system project was ILLIAC IV.”

Padua pointed out that parallel computing has become more and more important in today’s world.

“Parallel computing used to be the concern of a small function of the computer industry; now it has become the concern of the entire industry,” said Padua.

According to Padua, through all the 20th century, computers have gained an incredible speed increase every couple of years by simply increasing the clock speeds of its microprocessors. But the industry hit a wall as chips reached the melting point.

As a consequence, the computer industry has generally stopped relying on regular increases in the processing speed of chips. In recent years it has bet instead that future advances in speed and energy efficiency will come from putting multiple processors on a single silicon chip. Numbers of computer functions can then be done in parallel rather than sequentially.

“It seems like the era of sequential computing has to give way to a new era in which parallelism is at the forefront,” said Padua. “And so far it looks like the only way in the foreseeable future.”

“Parallelism has been part of the computing arena almost since the beginning of computers. It takes different forms, and the form it took in ILLIAC IV was particularly appealing. It showed people that things can be done in massive parallelism and it has profoundly influenced the way people saw supercomputing.” Padua said.

explore massive parallelism.” – Professor Mark Snir



Science of Security “Lablet” Established at the University of Illinois

The University of Illinois at Urbana-Champaign will receive an initial \$1 million in grant funds from the U.S. National Security Agency (NSA) to stimulate the creation of a more scientific basis for the design and analysis of trusted systems.

It is widely understood that critical cyber systems must inspire trust and confidence, protect the privacy and integrity of data resources, and perform reliably. To tackle the ongoing challenges of securing tomorrow’s systems, the NSA concluded that a collaborative community of researchers from government, industry, and academia is a must.

To that end, the NSA grant will seed an academic “Lablet” focused on the development of a Science of Security (SoS) and a broad, self-sustaining community effort to advance it. A major goal is the creation of a unified body of knowledge that can serve as the basis of a trust engineering discipline, curriculum, and rigorous design methodologies. The results of SoS Lablet research are to be extensively documented and widely distributed through the use of a new, network-based collaboration environment. The intention is for that environment to be the primary resource for learning about ongoing work in security science, and to be a place to participate with others in advancing the state of the art.

The Illinois Lablet, which will be housed in the Information Trust Institute at Illinois, will contribute broadly to the development of security science while leveraging Illinois expertise in resiliency, which in this

context means a system’s demonstrable ability to maintain security properties even during ongoing cyber attacks. David M. Nicol, the Illinois Lablet’s principal investigator, explains, “The complexity of software systems guarantees that there will almost always be errors that can be exploited by attackers. We have a critical need for foundational design principles that anticipate penetrations, contain them, and limit their effects, even if the penetration isn’t detected.”

The Lablet’s work will draw on several fundamental areas of computing research. Some ideas from fault-tolerant computing can be adapted to the context of security. Strategies from control theory will be extended to account for the high variation and uncertainty that may be present in systems when they are under attack. Game theory and decision theory principles will be used to explore the interplay between attack and defense. Formal methods will be applied to develop formal notions of resiliency. End-to-end system analysis will be employed to investigate resiliency of large systems against cyber attack. The Lablet’s work will draw upon ideas from other areas of mathematics and engineering as well.

Nicol, the project’s principal investigator, is a professor of Electrical and Computer Engineering (ECE) at Illinois and the director of the Information Trust Institute. The Lablet’s leadership is shared with co-principal investigators William H. Sanders, who is an ECE professor and director of the Coordinated Science Laboratory at Illinois, and José Meseguer, a professor of Computer Science.

“ It is widely understood that critical cyber systems must inspire trust and confidence, protect the privacy and integrity of data resources, and perform reliably. ”

Real March Madness is Relying on Seedings to Determine Final Four

BY LIZ AHLBERG, ILLINOIS NEWS BUREAU

As you've probably already learned the hard way, if you think picking all the top-seeded teams as the Final Four in your March Madness bracket is your best bet for winning the office pool, you should think again. According to an operations research analysis model developed by professor Sheldon H. Jacobson, you're better off picking a combination of two top-seeded teams, a No. 2 seed and a No. 3 seed.

"There are patterns that exist in the seeds," Jacobson says. "As much as we like to believe otherwise, the fact of the matter is that we've uncovered a model that captures this pattern. As a result of that, in spite of what we emotionally feel about teams or who's going to win, the reality is that the numbers trump all of these things," Jacobson said. "It's more likely to be 1, 1, 2, 3 in the Final Four than four No. 1's."

Jacobson's model is unique in that it prognosticates not based on who the teams are, but on the seeds they hold. He describes his model in a forthcoming paper in the journal *Omega* with co-authors Alex Nikolaev, of the University of Buffalo; Adrian Lee, of CITERI (Central Illinois Technology and Education Research Institute); and Douglas King, a graduate student at Illinois.

Jacobson has also integrated the model into a user-friendly website to help March Madness fans determine the relative probability of their chosen team combinations appearing in the final rounds of the NCAA men's basketball tournament.

A number of websites offer assistance to budding bracketologists, such as game-by-game probabilities of certain match-ups or determining the spread on a given team reaching a particular point in the tournament. Jacobson's website is the only one to look at collective groups of seeds within the brackets.

"What we do is use the power of analytics to uncover trends in 'bracketology.' It really is a

mathematical science," he said. "What our model enables us to do is look at the likelihood or probability that a certain set of seed combinations will occur as we advance deeper into the tournament."

Jacobson's team applied a statistical method called goodness-of-fit testing to NCAA tournament data from 1985 to 2011, identifying patterns in seed distribution in the Elite Eight, Final Four and national championship rounds. They found that the seeds themselves exhibit certain statistical patterns, independent of the team. They then fit the pattern to a stochastic model they can use to assess probabilities and odds.

Two computer science undergraduates, Ammar Rizwan and Emon Dai, built the initial website bracketodds.cs.illinois.edu based on Jacobson's model.

For example, the probability of the Final Four comprising the four top-seeded teams is 0.026, or once every 39 years. Meanwhile, the probability of a Final Four of all No. 16 seeds — the lowest-seeded teams in the tournament — is so small that it has a frequency of happening once every eight hundred trillion years. (The Milky Way contains about an estimated one hundred billion stars.)

"Basically, if every star was given a year, the years it would take for this to occur is 8,000 times all the stars in the galaxy," Jacobson said. "It gives you perspective."

However, sets with long odds do happen. The most unlikely combination in the 26 years studied occurred in 2000, with a Final Four seed combination of 1, 5, 8 and 8. But such a bracket is only predicted to happen once every 32,000 years, so those filling out brackets at home shouldn't hope for a repeat.

What amateur bracketologists can be confident of is upsets. For even the most probable Final Four combination of 1,1,2,3 to occur, two top-seeded schools have to lose.

"In fact, upsets occur with great frequency and great predictability. If you look statistically, there's a certain number of upsets that occur in each round. We just don't know which team they're going to be or when they're going to occur," Jacobson said.

After the 2012 tournament, and in years to come, Jacobson will integrate the new data into the model to continually refine its prediction power.

Until then, users can find out how likely their picks really are, and compare them against friends' picks — or even sports commentators'.

"We're not here specifically to say 'Syracuse is going to beat Kentucky in the Elite Eight.' What we're saying is that the seed numbers have patterns," Jacobson said. "A 1, 1, 2, 3 is the most likely Final Four. I don't know which two 1's, I don't know which No. 2 and I don't know which No. 3. But I can tell you that if you want to go purely with the odds, choose a Final Four with seeds 1, 1, 2, 3."

Networking Data Centers Randomly

While computer network designers have traditionally adhered to carefully planned, symmetrical structures, a team of computer science researchers at the University of Illinois is proposing the exact opposite: complete randomization of the network topology.

Ph.D. students Ankit Singla and Chi-Yao Hong, computer science professor Brighten Godfrey, and Lucian Popa of HP Labs are working together working together on the project, called “Jellyfish”. The project harnesses randomized networks to increase flexibility in network growth over time, and to achieve high network capacity at low cost.

“Intuitively, Jellyfish makes network capacity less like a structured solid and more like a fluid,” Godfrey said. “Coincidentally, it also looks like a jellyfish.”

Hong presented the team’s paper at the Networked Systems Design and Implementation Symposium on April 26 in San Jose, CA.

The project has two goals. The first is to provide high network capacity to facilitate data-intensive computation across many servers, an increasingly important use of compute clouds.

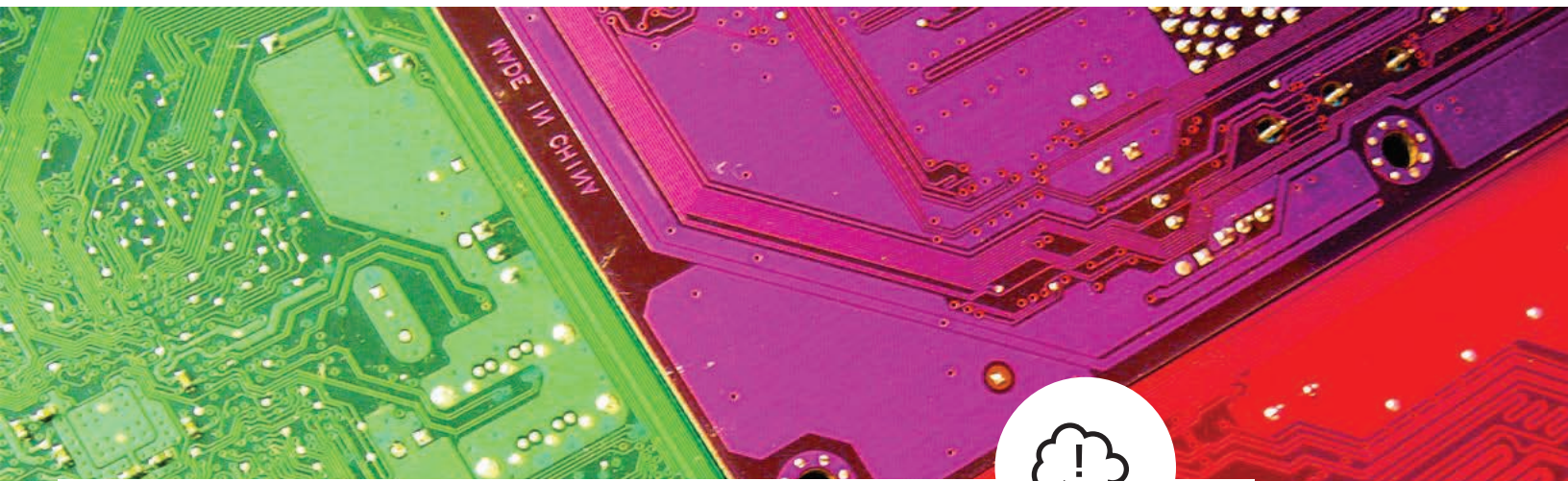
The second goal is to make incremental network growth easier. Current high-capacity networks like the fat-tree, hypercube, and butterfly have very constrained structures that restrict networks to certain sizes, and are hard to modify incrementally.

A Jellyfish network’s unstructured design allows it to expand smoothly while preserving efficiency, meaning cloud providers can meet increasing operational demands without incurring huge up-front capital expenditures. And surprisingly, the team has found that Jellyfish supports about 25% higher network capacity than an equally-priced fat-tree: Jellyfish’s diverse random connections result in short paths between servers, so that the network has less work to do for each packet of information and can deliver more data per unit time.

Hong’s presentation generated positive feedback and numerous questions from the symposium audience. “The project is rightly seen as an interesting, fresh look at network topology design that tackles an important practical problem for data centers,” Singla said. And, Hong added, “Many people think it’s a really cool idea to build a random network.”

The team is now developing techniques to build Jellyfish networks in practice. One problem is choosing the routes that packets should follow through the network – a process that traditionally takes advantage of simple network structure. “I feel routing is the main challenge,” Hong said. “Because of the lack of structure in Jellyfish, today’s standard equal-cost multipath routing may not work well.” The team’s results published in NSDI explored several promising solutions for routing, as well as other problems such as how to physically lay out the network cables.

“Jellyfish is sufficiently unlike past designs that implementation challenges will certainly arise,” Godfrey said. “But so far it seems like a completely random network just might work.”



Machine Learning and Perception

Three University of Illinois engineering faculty members will be investigating machine learning and perception via stochastic computation as part of a new Intel-funded center on embedded computing. Computer science professors Rob Rutenbar and Paris Smaragdis join electrical and computer engineering professor Naresh Shanbhag as part of the Intel Science and Technology Center on Embedded Computing headquartered at Carnegie Mellon University.

A key area of research for the center is to make it easier for these everyday devices to continuously collect, analyze and act on useful data from both sensors and online databases in a way that is timely, scalable and reliable. For example, in cars, this data could be used to customize in-vehicle entertainment options when specific passengers are recognized, and provide them better routing, retail, dining, and entertainment recommendations while on-the-road.

Tapping into the expertise of leading researchers from across the country, the ISTC for embedded computing forms a new collaborative community to

drive research to transform experiences in the home, car and retail environment of the future.

The team's research will be informed by new applications of machine listening in embedded computing environments for cars, retail environments, and homes. For example, the Illinois team's efforts may lead to cars that can listen for driver activity such as sleepiness, mechanical issues, or external traffic cues such as accidents. In a retail environment, the Illinois team envisions applications that can listen for customer movement and patterns linked to in-store advertising or product placement, or for emergency situations. Home applications may include listening for accidents and emergencies for the elderly or those needing assistance, or for building systems malfunctions like broken pipes.

"These new ISTCs are expected to open amazing possibilities," said Justin Rattner, Intel Chief Technology Officer. "Imagine, for example, future cars equipped with embedded sensors and microprocessors to constantly collect and analyze traffic and weather data. That information could be shared and analyzed in the cloud so that drivers could be provided with suggestions for quicker and safer routes."



Illinois-Developed Data Creation Tool a Boon for Genetic Biologists

BY: LIZ AHLBERG, ILLINOIS NEWS BUREAU

With the BeeSpace Navigator, University of Illinois researchers have created both a curation tool for genetic biologists and a new approach to searching for information.

The project was a collaboration between researchers at the Institute for Genomic Biology and the department of computer science. Led by Bruce Schatz, professor and head of medical information science at the U. of I. and affiliate professor in computer science, the team described the software and its applications in the web server issue of the journal *Nucleic Acids Research*. The research team also included CS faculty members Chengxiang Zhai and Saurabh Sinha.

When biologists need information about a gene or its function, they turn to curators, who keep and organize vast quantities of information from academic papers and scientific studies. A curator will extract as much information as possible from the papers in his or her collection and provide the biologist with a detailed summary of what's known about the gene – its location, function, sequence, regulation and more – by placing this information into an online database such as FlyBase.

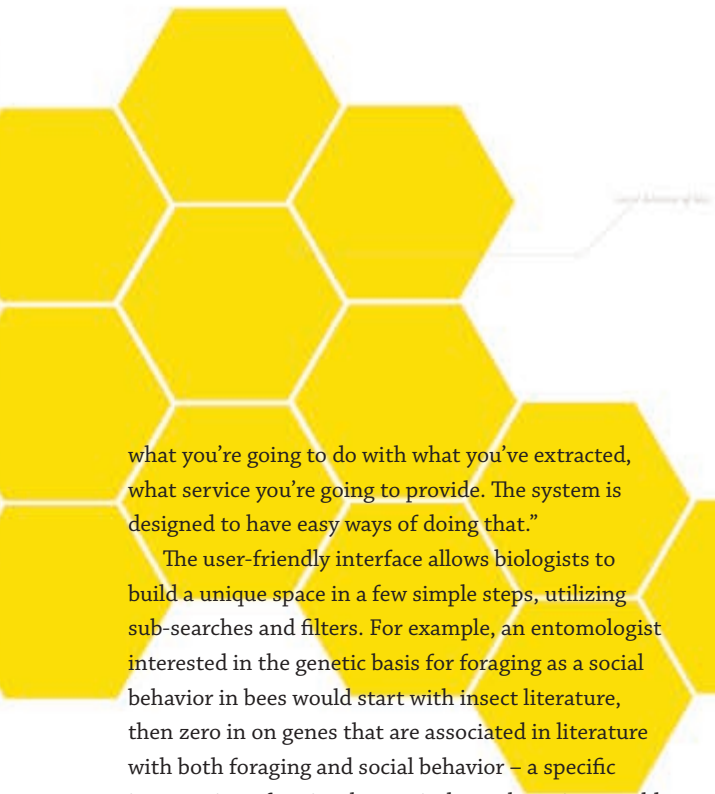
“The question was, could you make an automatic version of that, which is accurate enough to be helpful?” Schatz said.

Schatz and his team developed BeeSpace Navigator, a free online software that draws upon databases of scholarly publications. The semantic indexing to support the automatic curation used the Cloud Computing Testbed, a national computing datacenter hosted at U. of I.

While BeeSpace originally was built around literature about the bee genome, it has since been expanded to the entire Medline database and has been used to study a number of insects as well as mice, pigs and fish.

The efficiency of BeeSpace Navigator is in its specific searches. A broad, general search of all known data would yield a chaotic myriad of results – the millions of hits generated by a Google search, for example. But with BeeSpace, users create “spaces,” or special collections of literature to search. It also can take a large collection of articles on a topic and automatically partition it into subsets based on which words occur together, a function called clustering.

“The first thing you have to do if you have something that’s simulating a curator is to decide what papers it’s going to look at,” Schatz said. “Then you have to decide what to extract from the text, and then



what you're going to do with what you've extracted, what service you're going to provide. The system is designed to have easy ways of doing that."

The user-friendly interface allows biologists to build a unique space in a few simple steps, utilizing sub-searches and filters. For example, an entomologist interested in the genetic basis for foraging as a social behavior in bees would start with insect literature, then zero in on genes that are associated in literature with both foraging and social behavior – a specific intersection of topics that typical search engines could not handle.

This type of directed data navigation has several advantages. It is much more directed than a simple search, but able to process much more data than a human curator. It can also be used in fields where there are no human curators, since only the most-studied animals like mice and flies have their own professional curators.

Schatz and his team equipped the navigator to perform several tasks that biologists often perform when trying to interpret gene function. Not only does the program summarize a gene, as a curator would, but it also can perform analysis to extrapolate functions from literature.

For example, a study will show that a gene controls a particular chemical, and another study will show

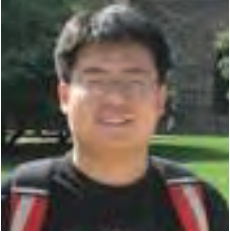
that chemical plays a role in a certain behavior, so the software makes the link that the gene could, in part, control that behavior.

BeeSpace can also perform vocabulary switching, an automatic translation across species or behaviors. For example, if it is known that a specific gene in a honeybee is analogous to another gene in a fruit fly, but the function of that gene has been documented in much more detail in a fruit fly, the navigator can make the connection and show a bee scientist information on the fly gene that may be helpful.

"The main point of the project is automatically finding out what genes do that don't have known function," Schatz said. "If a biologist is trying to figure out what these genes do, they're happy with anything. They want to get as much information as possible."

The project included the work of several CS students, now alumni, including Qiaozhu Mei, Jing Jiang, Xu Ling, Xin He, and David Arcoleo.

The BeeSpace Navigator, now in its fourth version, is available free online. Overview documentation is available as well. The National Science Foundation supported this work as the bioinformatics flagship of the Frontiers in Integrative Biological Research program.



Search Engine Optimazation

Search for the term “computer virus”, and it’s likely that some of your search results will contain only computer, or only virus. Why do these results show up instead of other more relevant results?

Ph.D. student Yuanhua Lv sought the answer to this question as he began exploring new methods to improve search algorithms. The performance of a search engine is mainly determined by its retrieval model which formally specifies how to score and rank documents optimally for a user’s query. Optimization of retrieval models is a fundamentally important research problem because an improved retrieval model would lead to improved performance for all search engines.

Lv first analyzed the deficiencies in current models, and revealed a previously unknown common deficiency in all current retrieval methods, namely that the component of term frequency normalization by document length is not properly lower-bounded. As a result of this deficiency, long documents which do match the query term can often be scored by search engines unfairly as having a lower relevancy than shorter documents that do not contain the query term at all. For example, for a query such as “computer virus”, a long document matching both “computer” and “virus” can easily be ranked lower than a short document matching only “computer”.

Lv’s discovery led him to create a novel methodology for introducing a sufficiently large lower bound for term frequency normalization, which can be used as a plug-and-play patch to multiple current retrieval models to eliminate the problem. Lab results indicate that Lv’s methodology incurs almost no additional computational costs while delivering more precise search results, particularly in cases where the queries are verbose.

Until Lv’s work, it has proven difficult to improve the current retrieval models. The well-known BM25 retrieval function was proposed in 1994, but efforts to improve its performance since then have been fruitless.

“[Lv’s work] makes a breakthrough contribution in improving the general search algorithms used in all search engines,” said Illinois computer science professor ChengXiang Zhai, Lv’s advisor and the co-author of a paper explaining Lv’s discovery and new methods. “The new models proposed can improve over many strong baseline methods that have been the state of the art for more than a decade. These new models can be immediately used to replace the old models in all search engines to improve search accuracy.”

The paper describing Lv’s work, “Lower-bounding term frequency normalization” received the Best Student Paper Award at the 20th ACM Conference on Information and Knowledge Management (CIKM 2011).



“Until Lv’s work, it has proven difficult to improve the current retrieval models.”

A stylized illustration of a network of people. Several colorful human figures (blue, orange, pink, green, teal) are arranged in a circular pattern, connected by dashed lines. Each figure has a speech bubble of a different color floating above them, suggesting communication and interaction. The background is white with a light blue gradient at the top.

Building a Better Internet

University of Illinois computer science professor Brighten Godfrey was among a select group of academic researchers and Internet visionaries chosen to participate in Verisign's "Building a Better Internet Symposium". Godfrey's project was one of four chosen internationally to receive a \$75,000 infrastructure grant that Verisign awarded as part of its 25 Years of .Com commemorations.

The University of Illinois project, a collaboration with Ph.D. students Wenxuan Zhou and Qingxi Li and Professors Matthew Caesar and Brighten Godfrey, developed methods to accelerate the Web and other interactive networked applications, via secure, deployable extensions to the domain name system (DNS) and transport control protocol (TCP). The team created the Accelerated Secure Association Protocol, or ASAP, which establishes a connection between client and server quickly and securely. The protocol enables the server to verify the key security property that the client's source address is not forged, yet avoids the delay of TCP's "handshake" method of verification.

"What I'm really excited about is how do we make the other side of the world feel like it's right at our fingertips," said Godfrey. "The exciting thing is that this work can have broad impact. If ASAP is widely deployed, it would make every connection on the web faster."

Pre-Social Networks

A technology that can tell where users are going to be, how long will be there, and who they will meet.

Sound like a sci-fi movie?

At Professor Klara Nahrstedt's lab, it's a reality. Nahrstedt and computer science graduate student Long Vuis new technology Jyotish draws up maps of people's movements by monitoring the connections their smart phones make to WiFi and Bluetooth networks. Over time, the system is able to determine the patterns in users activities and movements, and can make predictions on where people will be in the future, and what other people might be nearby during the same time frame.

The project began as an effort by Boeing to find better ways to track and predict the movements of work crews in its aircraft manufacturing facilities.

"It is well known that people movement exhibits a high degree of repetition since people visit regular places and make regular contacts for their daily activities," says Vu. "Our work constructs a predictive model by exploiting the regular pattern of people movement found in real joint Wifi/Bluetooth trace."

The model constructed by Jyotish is able to answer three fundamental questions: (1) where the person will stay, (2) how long she will stay at the location, and (3) who she will meet.

In order to construct the predictive model, Jyotish includes an efficient clustering algorithm to exploit regularity of people movement and cluster Wifi access point information in Wifi trace into locations. Then, a Naive Bayesian classifier assigns these locations to records in Bluetooth trace. Next, the Bluetooth trace with assigned locations is used to construct predictive model including location predictor, stay duration predictor, and contact predictor.

Assured Cloud Computing

JENNY APPLEQUIST, INFORMATION TRUST INSTITUTE

The U.S. Air Force Research Laboratory Technology Directorate (AFRL) has announced plans to create a new \$6 million "University Center of Excellence in Assured Cloud Computing," which will be a combined effort of AFRL, the Air Force Office of Scientific Research (AFOSR), and the University of Illinois at Urbana-Champaign. The center will be led by University of Illinois computer science professor Roy Campbell.

A new \$6 million research center funded by the Air Force Research Laboratory Technology Directorate will focus on assured uses of cloud computing infrastructures. Led by Prof. Roy Campbell, the Assured Cloud Computing (ACC) Center will perform research, provide technical exchange, and educate professionals and students in the secure cloud computing sciences and technologies that are needed to allow the Air Force to succeed in air, space, and cyberspace missions. ACC's research activities will focus on developing technology for assured, mission-critical cloud computing across "blue" and/or "gray" networks that ensures the confidentiality and integrity of data and communications, job completion in the presence of cyber attacks and failures, and timely completion of jobs to meet mission requirements.

A computational cloud used in military applications may include both blue and gray networks, where "blue" networks are U.S. military networks, which are considered secure, and "gray" networks are those in private hands, or perhaps belonging to other nations, which are considered insecure. In order to reach mission goals, it will sometimes be necessary to coordinate computation across a mixture of these blue and gray resources. Thus, cloud computing in a military context presents special security challenges. Specifically, assured mission-critical cloud computing across blue and/or gray networks requires the

realization of "end-to-end" and "cross-layered" security, dependability, and timeliness.

"We're trying to offer the military extended functionality, to implement a vision of global vigilance, global reach, and global power giving us the ability to meet overseas commitments," explains Campbell. "If we can use a variety of secure networks plus insecure networks, it gives us lots of agility and mobility and the ability to manage situations where normally we wouldn't be able to reach."

He points to humanitarian missions in potentially unfriendly territory as an example application. "Suppose you have a rescue mission, say a disaster in Pakistan, like an earthquake or a river flooding. If their government requests help, do we have the capabilities to safely assist in their aid? Not all the people in Pakistan might agree with the U.S. providing assistance. Staging such an operation would be risky without a cloud infrastructure that has secure properties. So how do you assure a successful mission in a possibly hostile environment? How do you benefit from the cloud, its communication, computations, and data in missions to help people in need?"

Ilesanmi Adesida, the Dean of the College of Engineering at Illinois, observed that the planned research has broad implications. "Although the new project will primarily benefit the Air Force over the short term, assured use of the cloud will be a tremendous benefit to humanity. Today, you can't really trust cloud computing because of the security issues that remain to be addressed," he explains. "No one has been able to use cloud computing in a task-oriented way before. The work of the Assured Cloud Computing Center will make it possible to deploy cloud computing in task-based, mission-oriented human activity. For that reason, this work will be groundbreaking."

In addition to Roy Campbell, the Center's research team includes computer science faculty members Gul Agha, Indranil Gupta, and JosÈ Meseguer among other Illinois faculty. The center is part of the Information Trust Institute at Illinois.



NCSA, Cray partner on sustained-petascale Blue Waters supercomputer

BY NCSA COMMUNICATIONS

The National Center for Supercomputing Applications at the UI has finalized a contract with Cray Inc. to provide the supercomputer for the National Science Foundation's Blue Waters project. Cray replaces IBM, which terminated its contract in August because, IBM said, the technology required more financial and technical support than originally anticipated.

Cray will begin installing hardware in the UI's National Petascale Computing Facility within the next few weeks, with an early science system expected to be available in early 2012. Blue Waters is expected to be fully deployed by the end of 2012.

This new Cray supercomputer will support significant research advances in a broad range of science and engineering domains, meeting the needs of the most compute-intensive, memory-intensive and data-intensive applications. Blue Waters is expected to deliver sustained performance, on average, of more than one petaflops on a set of benchmark codes that represent those applications and domains.

"We are extremely pleased to have forged a strong partnership with Cray. This configuration will be the most balanced, powerful and useable system available when it comes online. By incorporating a future version of the XK6 system, Blue Waters will also provide a bridge to the future of scientific computing," said Thom Dunning, NCSA director.

Physically Asynchronous Logically Synchronous Architecture

In networked cyber-physical systems, real time global computations require consistent views, consistent actions, and synchronized state transitions across network nodes. However, the convergence of sensing, control, communication and coordination in cyber-physical systems (CPS) such as modern airplanes, power grid, train and medical device networks poses enormous challenges.

Computer science researchers at the University of Illinois, led by Professors Lui Sha and Jose Meseguer, are developing formal complexity reduction architecture patterns aimed at addressing these challenges. The team has developed the Physically Asynchronous Logically Synchronous (PALS) architecture pattern that supports real time global computation. The objective of the PALS protocol is to provide the optimal real time logical (virtual) synchronization protocol.

The groups are working in collaboration with industry and research partners Rockwell Collins Inc., Lockheed Martin Corporation, Software Engineering Institute, and the University of Oslo. Researchers and engineers of Rockwell Collins Inc., in cooperation with Illinois computer science researchers, led the study on the implementation of PALS for an Integrated Modular Avionics application. The results showed that PALS reduced verification time from 35 hours to less than 30 seconds as compared with traditional design in a dual redundant flight guidance system prototype. This PALS study, authored by Steven P. Miller, Darren D. Cofer, Lui Sha, Jose Meseguer and Abdullah Al-Nayeem, won the David Lubkowsky Award for the Advancement Digital Avionics.

AADL is an open standard architecture analysis and description language supported by the avionics community. Because of this remarkable success in complexity reduction, Illinois researchers, in cooperation with Software Engineering Institute and University of Oslo, are now working on a special AADL Annex called Synchronous AADL to support PALS based designs. In addition, Illinois researchers, led by Dr. Cheolgi Kim, in cooperation with Lockheed Martin, are developing a production quality implementation of PALS library.

International Coalition Aims to Enable Climate Simulation at Extreme Scale

Policy decisions for mitigating climate change or adapting to it are subjects of great discussion throughout the world. Uninformed decisions will impose a heavy cost on future generations, both financial and human. Therefore, it is essential to reduce the current uncertainties about future climate changes and their impact by running climate simulations at 1,000 times larger scales than today. Exascale supercomputers (those capable of 10¹⁸ operations per second) may appear in 2018-2020, featuring a hierarchical design, gathering hundreds of millions of computing cores. The numerical models of the physics, chemistry, and biology affecting the climate system need to be improved to run efficiently on these extreme systems. Without improvement, these codes will not produce simulation results required to respond to the societal and economical challenges of climate change.

The objective of the G8 ECS project is to investigate how to efficiently run climate simulations on future exascale systems and get correct results. This project gathers top researchers in climate and computer science to focus on three main topics:

- how to complete simulations with correct results despite frequent system failures
- how to exploit hierarchical computers with hardware accelerators close to their peak performance
- how to run efficient simulations with 1 billion threads.

This project also aims to educate new generations of climate and computer scientists about techniques for high-performance computing at extreme scale.

This project is a direct result of the collaboration between the University of Illinois at Urbana-Champaign, the National Center for Supercomputing Applications and INRIA, the French National Institute for Research in Computer Science and Control, through their Joint Laboratory for Petascale Computing. In addition to these three institutions, other partners on the project come from Canada's University of Victoria, the German Research School, Japan's Tokyo Tech and University of Tsukuba, Spain's Barcelona Supercomputing Center, the University of Tennessee and the National Center for Atmospheric Research. The project will employ the top supercomputers to experiment with new techniques in the previously described three topics.

This three-year project receives G8 coordinated funding from the Natural Sciences and Engineering Research Council of Canada, French National Research Agency, German Research Foundation, Japan Society for the Promotion of Science (JSPS) and U.S. National Science Foundation. This project, together with five other projects, was funded as part of the G8 Research Councils Initiative on Multilateral Research, Interdisciplinary Program on Application Software towards Exascale Computing for Global Scale Issues. This is the first initiative of its kind to foster broad international collaboration on the research needed to enable effective use of future exascale platforms.

New Illinois center to develop smarter infrastructures, smarter publics

From smart utilities like the smart grid and intelligent transportation systems to social networks on sites like Facebook and YouTube, the infrastructures of tomorrow will heavily utilize information technology. While these “smart” infrastructures promise many benefits, they often require new kinds of interaction between people and the machines meant to serve them. Yet the social, cultural, economic and political side of these relationships often receives little attention.

The new Center for People and Infrastructures at the University of Illinois at Urbana-Champaign seeks to address these issues by better understanding social norms, market structures, public policies and human capabilities that shape and are affected by the development of smart infrastructures. The center, part of the Coordinated Science Laboratory, brings together experts in engineering, design, the social sciences and computer science.

The Center will initially focus on research about broadband telecommunications and energy. Researchers will work with fiber networks like UC2B (Urbana-Champaign Big Broadband), which will deliver high-quality broadband connections to several thousand households in Champaign and Urbana, to understand the consequences of fiber infrastructure for education, the economy, health and community participation. In addition, the Center will work to help identify the next generation of broadband Internet applications enabled by fiber infrastructure.

“Infrastructures are about computers, wires and pipes but they are also about human relationships, economics and justice,” said Co-Director Karrie Karahalios, associate professor of computer science. “We want to see infrastructures that not only work, but that help humans to flourish.”

National Petascale LEED

The University of Illinois’ National Petascale Computing Facility has been certified LEED® Gold in the U.S. Green Building Council’s Leadership in Energy and Environmental Design (LEED®) rating system, which is the recognized standard for measuring sustainability in construction.

The building, which opened in summer 2010, is home to supercomputers and other high-performance systems operated by the National Center for Supercomputing Applications and used by scientists and engineers across the country.

The LEED system awards points in a half-dozen categories, including energy and water efficiency, indoor environmental quality, and innovation in design. NPCF’s energy-saving features include:

- A highly efficient power distribution system that is based on 480 V power for the computational equipment.
- Focusing on water-cooled computational and storage equipment. Liquid cooling is two times more efficient than air cooling.
- External cooling towers that let Mother Nature chill the water needed for cooling the building and the supercomputers a large part of the year. This is expected to cut the chilled water costs for the facility by about \$1 million per year.
- Low-impact landscaping with native prairie plants that thrive without frequent watering or mowing.
- Using best practice construction methods to improve the air quality environment within the facility.



Regenerating Codes for Reliable Data Center Storage

A paper co-authored by Professor Brighten Godfrey received the 2010 IEEE Communications Society Technical Committee on Data Storage Best Paper Award.

The award is given by the Signal Processing for Storage Technical Committee of the IEEE Communications Society to a conference or journal paper published in 2010 that presents novel research on data storage.

The paper, "Network Coding for Distributed Storage Systems", was authored by Alex Dimakis, Brighten Godfrey, Yunnan Wu, Martin Wainwright, and Kannan Ramchandran, and appeared in the IEEE Transactions on Information Theory.

The paper created a new sub-area of coding theory called regenerating codes, which can make reliable storage in data centers more efficient.

Large data stores, which often appear in today's data centers and cloud services, use erasure coding to provide redundancy across multiple disks so if part of the file is lost it can be reconstructed. But the erasure-coded fragments must also be replenished after a failure, to maintain the same level of reliability. Traditionally a lost fragment is replaced by first reconstructing the entire file, and then distilling it into a small additional fragment to replace the lost fragment. The paper introduced a technique called regenerating codes which can produce new fragments much efficiently, without reconstructing the entire file.

The award was presented at the IEEE Global Communications Conference (Globecom) 2011 in Houston, Texas, in December.

Data-Driven Graphics Models

Professor Yizhou Yu received a Best Paper Award for his work to recreate physical imperfections in scanned models of a human hand. The work, "Controllable Hand Deformation from Sparse Examples with Rich Details" received the Best Paper award at the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA 2011).

"Recent advances in laser scanning technology have made it possible to faithfully scan a real object with tiny geometric details, such as pores and wrinkles," said Yu in the paper. "However, a faithful digital model should not only capture static details of the real counterpart, but be able to reproduce the deformed versions of such details."

The team used models of a human hand to test their approach. Hand models tend to be among the most challenging, because human hands have such a large degree of freedom of motion, and highly deformable wrinkles.

The team developed data-driven methods to reproduce such deformities both on a large-scale, and in high resolution. The team's framework was capable of synthesizing high-resolution hand mesh animations with rich and varying details from as few as 14 training examples. The team's approach is able to be applied both to keyframe animation systems as well as performance-driven animation.

The papers coauthors included Haoda Huang and Xin Tong (Microsoft Research Asia), Ling Zhao and Yue Qi (Beihang University), and KangKang Yin (National University of Singapore).

New Approach to Open-Domain Language Processing

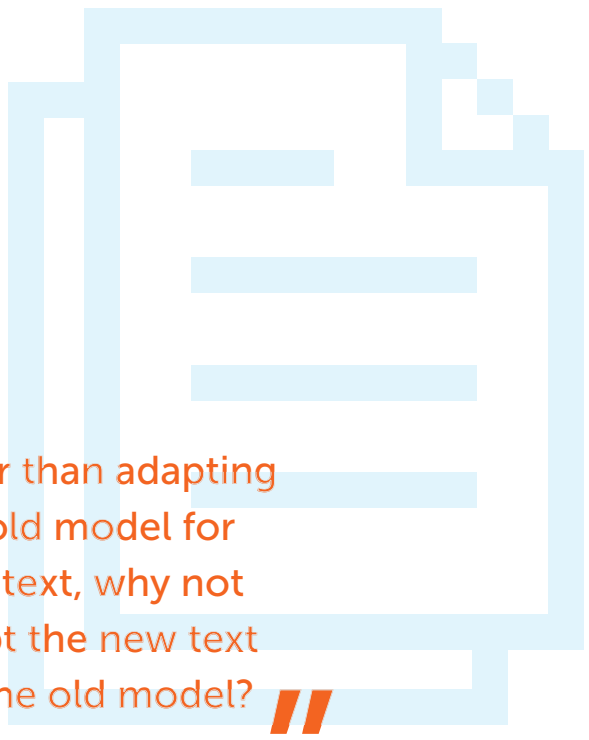
Graduate student Gourab Kundu received the Best Student Paper Award at the 2011 Conference on Computational Natural Language Learning (CoNLL 2011) for his work on open domain natural language processing. His paper, "Adapting Text Instead of the Model: An Open Domain Approach" introduces a new training approach that allows Natural Language Processing (NLP) researchers to adapt existing systems to new domains without the need for retraining their system on labeled data from the new domain.

Given the sentence "Gourab wrote two excellent papers and submitted them to 2012 conferences," a Semantic Role Labeling program would analyze the sentence at the level of "who does what to whom, when, where", and determine, for example, that Gourab is the Writer and also the Submitter, that two papers were submitted, that the submission happened in 2012, etc. While this task is trivial to humans, automating this process is very difficult due to the inherent ambiguity and variability of natural language.

Modern NLP approaches therefore use machine learning and inference approaches to accomplish this task. Consequently, these systems must be "trained" or fed a lot of sample text, along with the corresponding analysis, in order to learn how to analyze new, previously unobserved text. However, while natural language systems trained on labeled data perform well on the domain in which they were trained, they adapt poorly to other domains due to changes in vocabulary, sentence structure, etc. For example, a system that was trained to parse natural language sentences on the Wall Street Journal may perform very poorly when parsing language from fiction novels.

Current approaches to this problem require researchers to conduct the time-consuming task of training a new model for the new domain using data from that domain. This is costly especially since there is a need to first annotate the data from the new domain with the correct analysis.

Kundu proposes looking at the problem from another angle. Rather than adapting the old model for new text, why not adapt the new text for the old model?



“ Rather than adapting the old model for new text, why not adapt the new text for the old model? ”

The result is ADUT (ADaption Using label-preserving Transformation), an approach that avoids the need for retraining and does not require any knowledge of the new domain. ADUT applies multiple label-preserving transformations to the text it is attempting to analyze in order to make the target text "more similar" (in terms of vocabulary and structure) to the text on which it was initially trained. The ADUT system then applies the existing model on the transformed sentences, and combines the predictions using a constrained optimization inference approach to produce the desired prediction on the target text. In this manner, ADUT is able to work on the fly to analyze new text.

Kundu's ADUT approach can use a single trained model to analyze text from multiple domains, even when the notion of a domain is not well defined (e.g., streaming data from the web). The approach was shown to yield a 13% reduction in error for the task of Semantic Role Labeling when adapting from news wire text to text from fiction.

Kundu conducts his research in computer science professor Dan Roth's research group. Roth was co-author of the paper. One of Professor Roth's former students, alumnus Scott Wen-tau Yih (Ph.D., '05), also won the Best Paper Award (non-student) at the same CoNLL conference for his paper, "Learning Discriminative Projections for Text Similarity Measures."

ADSC Working to Analyze Biomedical Data without Risking Individuals' Privacy

BY: JENNY APPLEQUIST, CSL

Professor Marianne Winslett is leading a new project that will eliminate significant privacy risks that currently impede researchers' ability to analyze biomedical data about individuals. Immense amounts of valuable data now exist that are unusable by the research community due to the lack of an effective method for concealing individuals' identities.

The project is taking place at the Advanced Digital Sciences Center (ADSC) and is funded by A*STAR. The new ADSC work will generate new publication schemes for the results of data analyses, thus making detailed summaries of health data available that can offer unprecedented insight into a vast range of medical conditions and provide useful input for urban planners, public health officials, and researchers.

The \$2 million project, "Enabling Mining of Medical Records through Differential Privacy," is led by Winslett. Her co-principal investigators include Prof. Xiaokui Xiao of Nanyang Technological University, Prof. Jiawei Han of the Department of Computer Science at the University of Illinois at Urbana-Champaign, Dr. See Kiong Ng of the Institute for Infocomm Research, and Prof. Nikita Borisov of the Department of Electrical & Computer Engineering at Illinois. The team also includes biomedical researchers Dr. E Shyong Tai from the National University of Singapore and Dr. Edison Liu from the Genome Institute of Singapore.

The widespread availability of biomedical data, ranging from reports of the locations of new cases of dengue fever to individuals' genomic variations, appears to offer researchers a tremendous opportunity. Statistical analysis of such data can help researchers and public health officials better understand a disease and its transmission patterns, gain new insights into the human body, and develop new treatments and services that can improve the quality of life of millions of people.

Unfortunately, privacy concerns make it infeasible to provide researchers with unlimited access to biomedical information. Previous attempts to solve this problem have tried to anonymize data by removing personally identifiable information from medical records, but this does not provide sufficient protection. The main problem is that external knowledge can be used to re-identify individuals whose data appear in supposedly anonymized data sets. Many ideas for mitigating the problem have been proposed, but all of them have made the unrealistic assumption that adversaries had limited prior knowledge.

"In fact, this has been shown to be a fundamental barrier," explains Winslett. "An anonymized database will either reveal private information, given certain external knowledge -- or will be useless for answering some questions."

To the extent that databases of patient information have already been made available, they have made many lifesaving discoveries possible. For example, a University of San Antonio study involving data collected from over 9,000 breast cancer patients showed that amplification of the HER-2 oncogene was a significant predictor of both overall survival and time to relapse in patients with breast cancer. This information subsequently led to the development of Herceptin (trastuzumab), a targeted therapy that is effective for many women with HER-2-positive breast cancer. Likewise, it was medical records research that led to the discovery that supplementing folic acid during pregnancy can prevent neural tube birth defects (NTDs), and population-based surveillance systems later showed that the number of NTDs decreased 31 percent after mandatory fortification of cereal grain food products. No one doubts that additional valuable findings would follow if a way to tackle the privacy limitations can be found, so that far more patient data can be made available to researchers.

To that end, medical studies funded by the National Institutes of Health (NIH) in the U.S. are required to make the data they collect, as well as summaries of analysis results, available to other

Differential privacy offers us the tantalizing possibility of being able to do privacy-preserving data analysis that is both useful and secure

Marianne Winslett



researchers. Originally, the statistical summaries were freely available to other researchers via NIH's dbGaP database (<http://www.ncbi.nlm.nih.gov/gap>), while access to the detailed patient records required researchers to undergo a rigorous and fairly arduous approval process with their Institutional Review Boards (IRBs). Privacy concerns subsequently led NIH to restrict dbGaP access, so that today many of the statistical summaries cannot be viewed without IRB approval. The need for IRB approval is a significant hurdle for researchers who want to access the summary statistics from old studies to help them plan their future work.

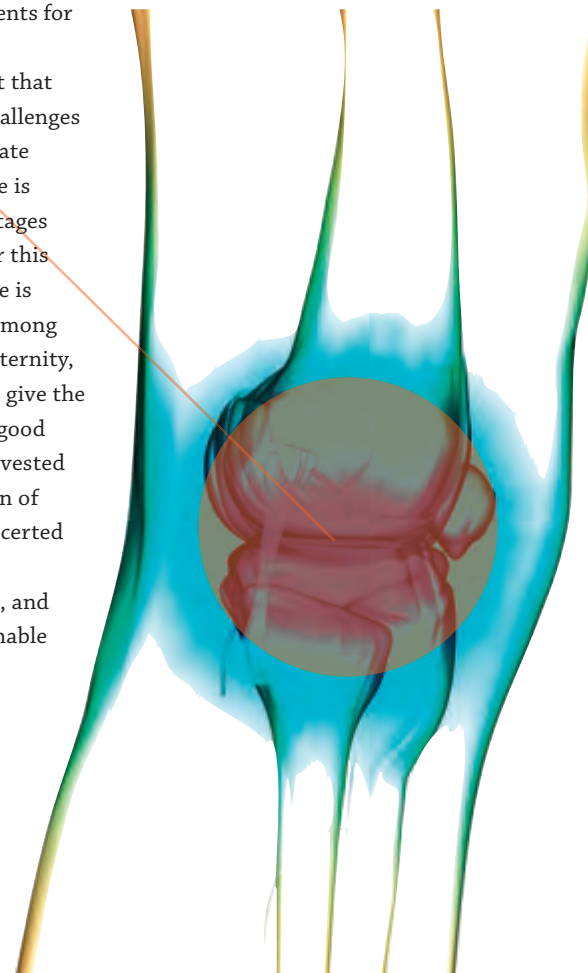
To find a practical solution, the ADSC team is using the recently developed concept of "differential privacy." Differential privacy works by adding a small amount of noise to the results of statistical analyses of sensitive data sets. Under differential privacy, the contributions of any one individual's data towards the outcome of an analysis are negligible; analysis results will be essentially identical regardless of whether a particular person's data are included. This should not limit the usefulness of the results, since in a large and well-designed medical study, the history of a single individual should not have a significant impact on overall results. When analysis of a data set begins, its owners decide on a total "privacy budget" for the entire data set. Each published analysis result uses up a little bit of the privacy budget, and once the budget has been exhausted, no more results can be published, as they could open the possibility of at least one individual's data having a non-negligible impact on overall results.

"Differential privacy offers us the tantalizing possibility of being able to do privacy-preserving data analysis that is both useful and secure," says Winslett. "It's such a new concept, but the implications are immense. Whoever comes up with a practical approach to differentially private access to biomedical data -- which is what we aim to develop with this new project -- will set off a free-for-all. It will open up so many new opportunities to revolutionize treatments and reduce health care costs."

The new project is starting by analyzing privacy issues in the statistics released by Singapore's Ministry of Health (MOH). Due to the potential for privacy breaches, MOH currently publishes detailed statistics only for highly dangerous infectious diseases, and only very sketchy information, such as the total numbers of male and female patients in Singapore, for other types of diseases. For instance, one report says that there was exactly one male patient aged 30-39 with relapsed tuberculosis in 2010. The team's goal is to make it possible to publish detailed statistics for all diseases, but with strong privacy-preservation guarantees.

The project's next step will be to investigate ways to re-enable open access to the summary information in dbGaP by making the summary tables differentially private. The researchers will also target other custodians and users of health-related statistics in Singapore. That work is projected to include applications in pharmacoeconomics and in analysis of hospital records to reveal the effectiveness of different treatments for a disease.

Winslett is quick to point out that several fundamental research challenges remain before differentially private analyses will be practical, but she is optimistic that ADSC has advantages that make it an ideal location for this research. In particular, Singapore is unique in its close cooperation among the government, the medical fraternity, and research institutes. This will give the ADSC researchers exceptionally good access to the parties who have a vested interest in broader dissemination of health data summaries. This concerted effort to bring together medical researchers, computer scientists, and medical records could one day enable Singapore to be a world leader in technologies for analyzing sensitive data.





Picture in Picture: Adding 3D Objects to 2D Images

Inserting objects into existing photographs or videos is extremely difficult because the object's appearance depends on the complex lighting and geometry of the environment. Researchers from University of Illinois show how to realistically render 3D objects into an existing photograph, requiring no special equipment and much less effort from the artist than previous approaches.

Researchers at Illinois have shown how to get the lighting right, by combining methods from computer vision, human computer interaction, and computer graphics. A user who wanted to insert an object or an animation into a picture would sketch the main layout of the space (or use a computer vision program to estimate it) and the lights using an interface, and use a standard modeling tool to make the object or the animation.

"Detecting photo fraud is often easy, because it is very difficult to ensure that the new object has the right shadows, highlights, and shading; casts shadows in the right places; and is reflected in mirrors," said computer science professor David Forsyth. "In contrast, experimental work conducted at Illinois suggests that this new system (a) produces accurate estimates of shading, tested by comparison with real photographs and (b) fools human observers into believing that an edited image is real."

The technical breakthrough is getting the lighting right by taking advantage of small amounts of annotation to recover a simplistic model of geometry and the position, shape, and intensity of light sources. The team's approach applies a computer vision

program to estimate strong directional sources of light like the sun through windows; applies a computer vision program to estimate how much light the surfaces in the picture reflect, and to correct the user sketch of light; and then uses a standard computer graphics program to generate the picture from detailed physical models of light behavior.

In addition to the overall system, the research team also developed a semiautomatic algorithm for estimating a physical lighting model from a single image. The Illinois method is able to generate a full lighting model that is demonstrated to be physically meaningful through a ground truth evaluation. As part of their work, the team also introduced a state-of-the-art image decomposition algorithm for single image reflectance estimation.

"With a single image and a small amount of annotation, our method creates a physical model of the scene that is suitable for realistically rendering synthetic objects with diffuse, specular, and even glowing materials, while accounting for lighting interactions between the objects and the scene," said computer science Ph.D. student Kevin Karsch, lead author of the approach. "Our approach has applications in the movie and gaming industry, as well as home decorating and user content creation, among others. Imagine being able to slay dragons right in your own living room."

Kevin's work, joint with Varsha Hedau, Derek Hoiem, and David Forsyth, will appear in this year's SIGGRAPH Asia conference. They plan to produce a web-server version of this application in the near future.



Kevin Karsch Wins \$30k Lemelson-MIT Illinois Prize

STEPHANIE LARSON, TECHNOLOGY ENTREPRENEUR CENTER

Kevin Karsch, a passionate and diligent graduate student in computer science (CS) at Illinois, is the sixth annual winner of the \$30,000 Lemelson-MIT Illinois Student Prize, funded through a partnership with the Lemelson-MIT Program.

“Like the Post-It® Note, Kevin’s invention succeeds because it solves a common problem in an easy-to-use package,” said his letter of recommendation writer, Alyosha Efros, Finmeccanica Associate Professor of Robotics and Computer Science at Carnegie Mellon University. “The key to the method’s success is a well-designed collaboration between the human and machine.”

Karsch’s software is able to relight inserted objects and even works for complex lighting patterns. Note the strongly directed pools of light on the dragon’s back (in the image above), giving the appearance that the sculpture was really in the photo.

In his research, Karsch noted that currently adding digital actors or props to media is a painstakingly long process that requires artistry, expertise, and physical measurements of the scene. To simplify this process, Karsch developed a new technique for inserting objects

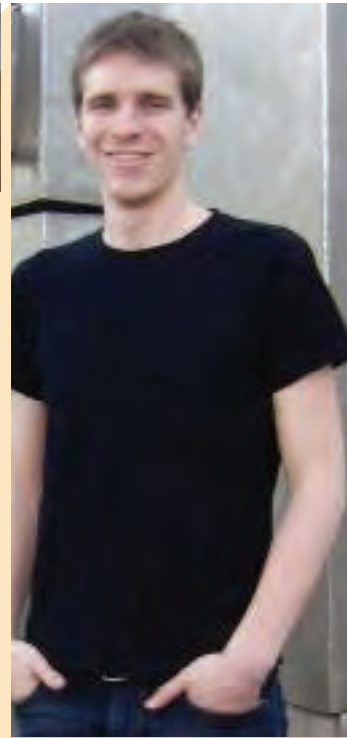
and special effects into photographs and videos that requires no scene measurements and can be performed by novice users in only a few minutes.

With over 850,000 views on his online demonstration video and interest from a variety of companies, the tool’s possibilities are versatile and ubiquitous. Thus far, proposals for uses of Karsch’s technology span from 3D modeling to picture editing as well as virtual home and property redecoration.

The University is working with him to file for a patent and has already begun licensing agreements. The future looks promising for the advancement of Karsch’s technology.

“Kevin’s system takes several complex technical methods (lighting estimation, perspective modeling, interactive matting, etc.) and puts them into an intuitive interface that outputs a 3D model that works with standard rendering packages,” said Efros. “Kevin has the most important skills of an inventor: ability to incorporate and innovate complex technical ideas; a grasp of the human element in product design; and an understanding of the end-to-end use of a product.”

In addition to his advisor, David Forsyth, Karsch works closely with CS faculty members Derek Hoiem and John Hart. Karsch has high hopes that his invention will revolutionize current image editing and significantly reduce production cost and time.



KEVIN KARSCH

View a video of this work in action at: vimeo.com/28962540





student apps

BenchStats Helps Coaches Improve Their Game

NATHANIEL LASH, ELECTRICAL AND
COMPUTER ENGINEERING

Illinois engineering students are developing a program that could change the way tennis is played, or at least coached. ECE student Rod Serna teamed up with Computer Science students Marcell Vazquez-Chanlatte and Eric Mills to create BenchStats, a new program that records and analyzes tennis players' performance.

Serna said this software will be able to help coaches and players alike.

"It will help the coach by simplifying some of his tasks such as scheduling practices and help him develop his players by showing him their strengths and weaknesses," Serna said. "Similarly for the player, allowing them the ability to see their stats will help show them what area of their game they need to work on."

Serna said the goal of the program is to provide tennis coaches a direct way to keep track of their players' statistics. Among these stats are first serve percentages, including aces and double faults, which players would be able to view as soon as they are updated.

The tool can help coaches streamline some of their tasks as well.

"It will also be managing software, allowing coaches to maintain their schedule online for their players to view," Serna said.

BenchStats started while Serna worked with the Illinois tennis team. Serna, who played tennis in high school and already attended most of the team's tennis matches, volunteered to help with statistics for the University's team, first with only pen and paper, but soon evolving into a simple PHP script.

"This evolved into one of the coaches suggesting I might try and make this into a platform where coaches could calculate statistics on their own easily," Serna said.

Later, he was joined in the project by Vazquez-Chanlatte and Mills, and they used Ruby on Rails, an open-source web development tool, and enhanced the project further by incorporating C++.

The collaboration started in a physics discussion section, where Vazquez-Chanlatte, who has developed playbook simulators for sports before, jumped on the opportunity when Serna pitched his idea, and brought on Mills and his Web-designing experience.

Vazquez-Chanlatte hopes to develop the software for a mobile platform, and he is currently investigating whether it would be a viable option for the Android Tablet platform.

"I very much am interested in pushing it into the mobile space due to portability and the potential to eliminate the score card medium, allowing direct input into the cloud," Vazquez-Chanlatte said.



CS Students Create "Flops Fever" iPhone App

Scientists and engineers keep supercomputers very busy simulating severe storms, galaxies, molecules, and more. Scheduling those jobs is like putting together a complex puzzle. Computer science students Ari Morgan and Harry Hsiao have created an iPhone/iPad game, Flops Fever, that emulates the activity.

Created at the National Center for Supercomputing Applications (NCSA), Flops Fever can be downloaded free from the Apple app store. An Android version of the game will be coming soon.

To play the game, place the different types and sizes of science and engineering jobs on the available nodes of the supercomputer. If you're too slow, jobs will pile up in your queue and your supercomputer will be idle! Flops failure!

Can you juggle jobs and keep the flops count climbing, or will your queue explode?



Barnyard Break

KIRSTEN KELLER, COMPUTER SCIENCE

The recent surge of apps for smart phones has made people crazy about games like Angry Birds, Plants vs. Zombies and Temple Run. Now, thanks to computer science students Jay Prombo and Neal Whittington, another addicting app has been added to the mix: Barnyard Break.

Barnyard Break is an interactive game where animals break out of their pens and it is the user's job to round them back up. With 90 levels, over 1,000 animals to round up and four tools available to catch them, Barnyard Break is a user-friendly yet challenging app.

"We wanted to create something that someone could play in class or on the bus, where you'd just rather not pay attention to what's going on around you, and want something more fun to do," Prombo said.

"Barnyard Break starts off easy but gets tough fast!" Whittington added. "We made the game like this so it frustrates the user only enough to keep them playing over and over again."

Prombo and Whittington, both juniors, met through a mutual class their sophomore year. As their friendship strengthened, they bounced ideas off each other, one of which was Barnyard Break.

"The idea of Barnyard Break came to me when I was contemplating "Why is Angry Birds so popular?" Whittington said. "I took those reasons, applied them and the result was Barnyard Break."

The pair used Flex and ActionScript 3 to develop the app. After a year of work, Barnyard Break was released on Mar. 18.

"We're very happy to actually have it out on the market," Prombo said. "It lifts a big weight off of our shoulders."

Barnyard Break is available on the iPhone and on Android phones.

Nyan Cat

Think the world already has enough Nyan Cat? Think again, says computer science senior Kevin Lange. Lange developed a Nyan Cat telnet server to serve up the ubiquitous poptart cat to anyone, from anywhere, on any system, at any time. The animated, color ASCII-text telnet server renders a loop of the classic Nyan Cat animation from telnet miku.acm.uiuc.edu.

“The idea came while some colleagues and I were playing around with alibi,” said Lange. “We all decided it would be fun to get the classic Internet meme running in a terminal, and after a bit of work I had a prototype written that ran on a toy operating system I have been working on since January. When one of my friends from the original group saw the prototype, he said I should make a telnet server so people can remotely connect and see NyanCat.”

One of the challenges that Lange faced was making the

animation work on different terminals — 13 officially supported ones in all, ranging from Gnome to VT220.

For Lange, the best part of the project thus far has been how far it has spread.

“I’ve been posted about by Google engineers on Google+, Modmyi (a popular iPhone and iPod modding/jailbreaking community), the head of security at Twitter (who has over 9000 followers!), and hundreds of other people.” Shortly after Lange launched Nyan Cat, there were 820 results on Twitter when searching for “miku.acm.uiuc.edu”, and the Github project had 88 watchers.

“On reddit, my post in r/linux has 467 votes, and the one r/programming has 140. It’s definitely the most popular anything I have made has ever been,” continued Lange.

As with most of Lange’s projects, Nyan Cat is open source (NCSA license) and available on github (kevinlange).

Watch NyanCat on a VT 220 at <http://www.youtube.com/watch?v=D4c5xpLdm-A>



Classwhole

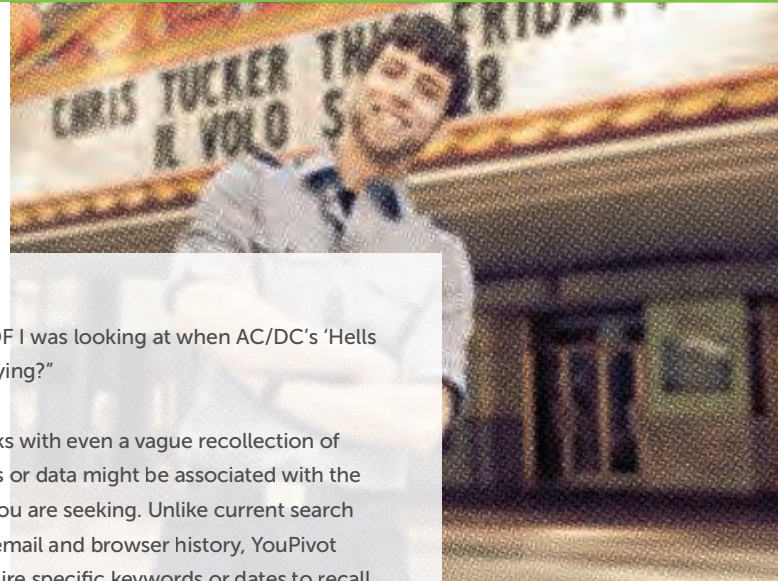
For CS seniors Kiran Ryali and Jon Hughes, and ECE senior Scott Wilson, scheduling can sometimes be a pain in the class. As part of their Technology & Entrepreneurship Senior Design course, they decided to do something about it.

Result: Classwhole, a socially interactive class scheduling system. With Classwhole, students planning their next semester’s schedules can enter their desired courses and drag and drop potential class meeting times around an interactive calendar. While shifting times for an optimal schedule, users can also see which other Classwhole users have opted for the same courses.

“The project started out just being a pure schedule, but as we started working on it, we realized we could improve some things,” said the team. Considering additional functionality, the team realized that they could introduce the ability to drag and drop classes around a calendar view, building new schedules on the fly, rather than presenting the user with potentially dozens of permutations of their potential schedule to choose from. The team credits this functionality as what differentiates it from Scedule - and also as their most difficult technical challenge to date.

This fall, the team plans to release a mobile app version that will show students their schedule on a calendar as well as provide directions to campus buildings and maps.

View YouPivot at:
[http://youtube.com/
watch?v=v1mnEf63W2c](http://youtube.com/watch?v=v1mnEf63W2c)



JOSH HAILPERN

Finding Lost Data

Misplace your keys and the advice you'll be given is to "try and remember where you were the last time you had them." Why shouldn't finding your lost digital data work the same way?

Thanks to new technology developed by University of Illinois computer science Ph.D. student Joshua Hailpern, it can. Hailpern's software, YouPivot, helps you retrace your steps back to your lost files, websites, or documents by tapping into how the human memory works.

YouPivot keeps track of your activities on your computer at any given point in time - music you're playing, websites you're browsing, files you're working with. Using this data as a reference, YouPivot allows you to browse, search and "pivot" to see what else was happening on your computer at the time and rediscover what you are looking for. As Hailpern told IEEE Spectrum in a recent article, "What was that

website or PDF I was looking at when AC/DC's 'Hells Bells' was playing?"

YouPivot works with even a vague recollection of what activities or data might be associated with the information you are seeking. Unlike current search methods for email and browser history, YouPivot does not require specific keywords or dates to recall information. What's more, pivoting is only one of the many new features Hailpern created to allow users to find their "digital car keys."

YouPivot is currently available in alpha as a free extension to the Google Chrome browser, though Hailpern and his team are working on a beta release by end of 2011, and full release in spring of 2012.

Hailpern's work on YouPivot was recently profiled in an IEEE Spectrum article "A Prototype of Pivot Searching." Learn more about Hailpern's technology by reading the article, or by viewing a YouTube video on the software.

"That will be especially helpful for freshmen," Scott points out.

"Well, we're even seniors and a lot of times we don't know where some buildings are," Jon adds.

After a soft launch on the Illinois campus, the team had more than 1000 students visit the site. Word of mouth spread through Facebook and Reddit posts. Since then, classwhole.com has seen more than 4000 unique visitors.

"It felt really good to have it pay off. We didn't really do any advertising, so it was nice to see people start using it," says Kiran.

"A lot of our traffic is through Facebook, so it's like a view into the university's social graph, watching it go viral through students liking it," adds Scott.

Still, the team would like to see "pretty much everyone at the University using it."

The team recently won a regional coding contest sponsored by Google, beating out teams from Purdue (with a Wikipedia crawler) and University of Illinois at Chicago (with an interactive lecture slide product). The team credits their product's polish, successful launch, and user base with giving them the edge over their competition.

Students Take on World at Programming Finals

KIRSTEN KELLER, COMPUTER SCIENCE

For Hassan Eslami, computer science Ph.D. student, no two problems in computer science can be solved using the same approach.

“Usually problems in theoretical computer science have a feature that there is no routine mechanism to approaching them,” Eslami said. “When I try to solve problems, I start with lots of examples, and then I try to figure out what the pattern is in all the examples.”

Using these techniques, Eslami and two other computer science students, seniors Justin Kopinsky and Krishan Chockalingam, will advance to the world finals of the ACM Intercollegiate Programming Competition in Warsaw, Poland from May 14–18. 112 teams from all over the world will participate, with only 18 universities from the United States competing.

The three students, under the team name “ILLIAC,” qualified for the world finals when they placed first at the Mid-Central Regional Programming Contest in November of 2011. They beat teams from other schools including Northwestern, Illinois Institute of Technology and the University of Chicago.

“We have been fortunate to have the cream of the crop to work with,” said Marsha Woodbury, faculty advisor to Team ILLIAC.

According to Darko Marinov, associate professor in computer science and the coach of Team ILLIAC, a competitive local competition was held to pick students for the regional contest. Out of 30 participants, the top 15 were selected to form five three-member teams for the regional contest.

Team ILLIAC, which was the one team out of the five to advance to the world finals, is the ninth team in ten years from Illinois to compete in the world finals.

Kopinsky also served on the ACM ICPC World Finals team in 2010.

The world finals will consist of a set of eight problems, and each team has five hours to complete as many as possible. Using a single computer, teams will rank the level of difficulty of the problems and proceed to figure out what needs to be done for each separate problem, design tests and ultimately build software systems that solve the problems.

“The very beginning of the competition is quite stressful. Since, at the beginning there are several problems which are quite easy, and thus you are trying to code as fast as possible. Towards the later part of the competition it’s a lot more relaxed, as time is no longer as important. The focus of the later part is a lot more intensive thinking, rather than a race to code as fast as possible,” said team member Krishnan Chockalingam.

Past problems include mapping the prime placement of a channel for farmland irrigation and improving travel times in mass transit systems.

“[W]e challenge the finalists to look past the codes they manipulate and engage in active problem solving for some of the most urgent issues facing our planet,” said Salvatore Vella, vice president of architecture and technology at IBM Software Group.

Last year, Illinois received honorable mention in the world finals.

“We have not explicitly specified [a] goal this year,” Eslami said. “The best ranking [by Illinois] was 17, by solving five problems. We will try to solve at least four to five problems and be in the ranking this year.”

Advancing to the world finals is something Eslami has wanted to do since studying as an undergraduate in Iran.

“It will be a great experience to compete with capable students in programming and algorithm,” Eslami said.

2 Students Win CRA Research Honors

KIRSTEN KELLER, COMPUTER SCIENCE

Every year, the Computing Research Association (CRA) honors undergraduate students in North American universities who demonstrate extraordinary research potential in an area of computing research with the Outstanding Researcher Award. This continental award is highly selective and competitive, with only two to three winners each year. Therefore, it is a great honor that this year, the University of Illinois' Forrest landola and Rohan Sharma were finalists for the award.

To apply for the award, landola and Sharma had to submit a resume, a one-page research summary and letters of recommendation.

landola's research lies in many areas, including real-time systems, parallel particle transport, computational medical physics, GPU cluster computing and facial deformation simulation.

"Following my graduation, I plan to attend graduate school and to focus on scheduling theory for real-time distributed systems," landola said. landola's research into real-time systems began as part of a class project with Professor Tarek Abdelzaher. During this project, landola and Prof. Abdelzaher attempted to discover and prove theoretical boundaries on the worst-case execution times of tasks in real-time distributed systems. landola's work on the project led to a proof on the behavior of systems in merging workflows, and developed a new system that demonstrated tighter bonds than those previously implemented.

In the long run, landola plans to become a professor in computer science.

"Forrest is a remarkable student who has been conducting research since high school," said Rob Rutenbar, professor and head of the Computer Science department. "[He] is organized, energetic, and passionate about computer science."

Sharma's research is based in software testing to find techniques for finding bugs in code and to automatically generate data structures to use as test inputs.

"Array-based and linked structures such as trees and hash tables are an integral part of programming language libraries, and hence they are vital to real-world code," Sharma said. "It is crucial to make sure their implementations are bug-free."

In his research, Sharma demonstrated a new semi-automated approach to predicate coverage. In tests of his new approach, he has shown that using randomization to generate test suites was better than a more systematic state space exploration.

Sharma's research has resulted in two papers, with a likely third on the way.

"As juniors, many undergraduates are only just beginning to get involved with research, yet Rohan has already made valuable contributions to software testing," Rutenbar said.



awards

Sarita Adva

IEEE Fellow

ACM Fellow

Anita Borg Institute Women of Vision
Innovation Award

Lawrence Angrave

Illinois Student Senate Teaching Excellence
Award

Campus Award for Excellence in
Undergraduate Teaching

Robert Bocchino

ACM SIGPLAN Dissertation Award

Matthew Caesar

NSF Career Award

Brighten Godfrey

NSF Career Award

Bill Gropp

IEEE Fellow

SIAM Fellow

National Academy of Engineering

Jiawei Han

Campus Award for Excellence in Graduate and
Professional Teaching

Julia Hockenmaier

NSF Career Award

Derek Hoiem

NSF Career Award

Forrest Iandola & Rohan Sharma

CRA Outstanding Undergraduate Researcher
Award, Finalists

Brett Jones & Rajinder Sodhi

Qualcomm \$100,000 Innovation Fellowship

Laxmikant Kale and Parallel Programming Laboratory

HPC Challenge, First Place, Class 2

Karrie Karahalios

Kavli Frontiers of Science Fellow

Kevin Karsch

Lemelson-MIT Illinois Prize Winner

David Liu, emeritus

Phil Kaufman Award for Distinguished
Contributions to Electronic Design Automation

Klara Nahrstedt

IEEE Technical Achievement Award

Bryan Plummer

Barry M. Goldwater Scholarship

Dan Roth

ACM Fellow

Josep Torrellas

ACM Fellow

Tianyi Wu

ACM SIGKDD Dissertation Award, Runner-Up

Best Paper Awards

2011 ACM Multimedia Conference (MM 11): Best
Student Paper Award, November 2011

"Color-plus-Depth Level-of-Detail in 3D
Teleimmersive Video: A psychophysical Approach"
W. Wu, A. Arefin, G. Kurillo, P. Agarwal, K.
Nahrstedt, R. Bajcsy

2011 ACM SIGGRAPH/Eurographics Symposium
on Computer Animation (SCA 2011): Best Paper
Award, August 2011

"Controllable Hand Deformation from Sparse
Examples with Rich Details" H. Huang, L. Zhao, K.
Yin, Y. Qi, Y. Yu, X. Tong

2011 IEEE Symposium on Reliable Distributed
Systems (SRDS 2011): Best Paper Award, October
2011

"Modeling the Fault Tolerance Consequences of
Deduplication" E. Rozier, W. H. Sanders, P. Zhou, N.
Mandagere, S. M. Uttamchandani, M. L. Yakushev

2011 ACM Conference on Information and
Knowledge Management (CIKM 2011): Best
Student Paper Award, October 2011

"Lower Bounding Term Frequency Normalization"
Y. Lv, C. Zhai

2011 International Conference on Advances in
Social Networks Analysis and Mining (ASONAM
2011), August 2011

"Evolutionary Clustering and Analysis of
Bibliographic Networks" M. Gupta, Y. Sun, C.
Aggarwal, J. Han

2011 Parallel Architectures and Compilation
Techniques (PACT 2011): Best Paper Award,
October 2011

"DeNovo: Rethinking the Memory Hierarchy for

Disciplined Parallelism" B. Choi, R. Komuravelli, H. Sung, R. Smolinski, N. Honarmand, S. Adve, V. Adve, N. Carter, C-T. Chou

2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2011): Best Student Paper Award, September 2011

"Simplicial Dijkstra and A* Algorithms for Optimal Feedback Planning" D. Yershov, S. M. LaValle

2011 IEEE Computer Vision and Pattern Recognition (CVPR 2011): Best Student Paper, June 2011

"Recognition Using Visual Phrases" A. Farhadi, M. A. Sadeghi

2011 Conference on Computational Natural Language Learning (CoNLL 2011): Best Student Paper, June 2011

"Adapting Text Instead of the Model: An Open Domain Approach" G. Kundu, D. Roth

2011 IEEE Transactions on Computer-Aided Design: Donald O. Pederson Best Paper Award, April 2011

"Statistical Blockade: Very Fast Statistical Simulation and Modeling of Rare Circuit Events and Its Application to Memory Design" R. Rutenbar, A. Singhee

2011 IEEE International Conference on Data Engineering (ICDE): Best Student Paper Award, April 2011

"How Schema Independent Are Schema Free Query Interfaces?" A. Termehchy, M. Winslett, Y. Chodpathumwan

2011 IEEE Pervasive Computing and Communications Conference (PERCOM 2011): Mark Weiser Best Paper Award, March 2011

"Jyotish: A Novel Framework for Constructing Predictive Model of People Movement from Joint Wifi/Bluetooth Trace" L. Vu, Q. Do, K. Nahrstedt

David Liu



Professor Emeritus David Liu Receives Phil Kaufman Award

Professor emeritus Dr. C. L. David Liu was the winner of the 2011 Phil Kaufman Award for Distinguished Contributions to Electronic Design Automation (EDA). The Phil Kaufman Award honors individuals who have made an impact on the field of EDA and pays tribute to Phil Kaufman, the late industry pioneer who turned innovative technologies into commercial businesses that have benefited electronic designers.

Early in his career, Dr. Liu led the transformation from ad hoc EDA to algorithmic EDA. He was an early advocate for more rigorous design automation, arguing that powerful, formal algorithmic techniques were essential to the effective solution of complex design automation problems. His technical contributions are at the foundation of a multitude of current EDA tools within several disciplines, including behavioral synthesis, logic synthesis and physical design.

His technical impact includes the first floorplanning algorithms and scheduling algorithms for hard real-time tasks. His research on floorplanning received DAC's Best Paper Award in 1986 and has been widely regarded as seminal. Dr. Liu's work on Rate Monotonic Scheduling (RMS) is a cornerstone of modern scheduling theory, applicable in the design of real-time operating systems. As of today, his 1973 paper on the subject has over 7,000 citations.

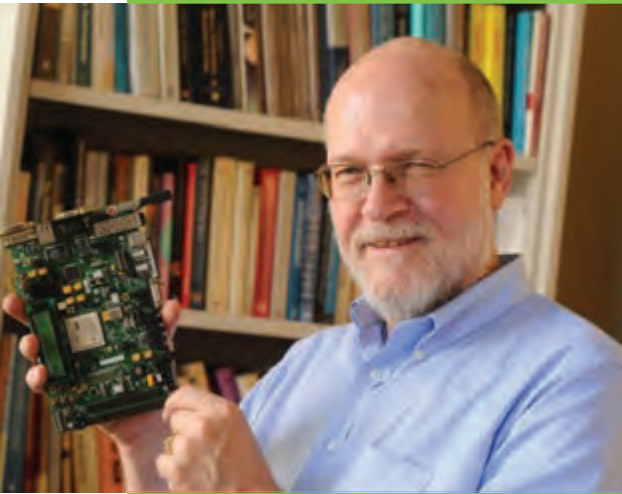
"We should not be surprised when one of our community's leading technical contributors has remarkable leadership and business skills as well," observes Rob A Rutenbar, Abel Bliss professor and head the Department of Computer Science at

the University of Illinois at Urbana-Champaign, a colleague of Dr. Liu's for more than 25 years. "But neither should we fail to notice when a colleague like David Liu manifests in such an impressive sphere of activities, visible worldwide."

Over the past 12 years, Liu's contribution to Taiwan's semiconductor industry has been broad and significant. He serves as chairman of the board of TrendForce, a market intelligence provider in the DRAM, LED, LCD and Solar Energy technical segments. He is a member of the board of Powerchip Semiconductor Corp., United Microelectronics Corp., MediaTek and Macronix International Co., Ltd. Additionally, he is a member of the board of Anpec Electronics Corporation, Andes Corporation, and Cadence Methodology Service Company.

For the last six years, he has hosted a weekly radio show on technology, natural science, social science and literature in Taiwan. He has published three essay collections based on the presentations in the show. One of them is a 2011 winner of a book award in the category of popular science.

in memory



COURTESY OF EILEEN BARROSO FOR COLUMBIA ENGINEERING

David L. Waltz

a former Illinois electrical and computer engineering faculty member, and a computer scientist whose early research in information retrieval provided the foundation for today's Internet search engines, has died at the age of 68.

David Waltz spearheaded advances in artificial intelligence. Many computer science alumni will remember taking courses in artificial intelligence from him.

During his career as a teacher and a technologist at start-up companies as well as large corporate laboratories, Dr. Waltz made fundamental contributions to computer science in areas ranging from computer vision to machine learning.

One signal achievement was the development of a basic technique that makes it possible for computers to render three-dimensional scenes accurately. The 3-D research was seminal in the fields of computer vision and artificial intelligence. Known as "constraint propagation," the technique is now used in industry for solving problems like route scheduling, package routing and construction scheduling.

At M.I.T., Dr. Waltz was taught by Marvin Minsky, a pioneer in artificial intelligence. Dr. Waltz graduated in 1972, then taught computer science at the University of Illinois at Urbana-Champaign and, later, at Brandeis University in Massachusetts, then Columbia University.

While at the University of Illinois, Dr. Waltz turned to the field of natural language understanding, a component of artificial intelligence involving the interpretation of language. With support from the Office of Naval Research, he built a question-answering system called Planes and explored the use of neural networks in language processing.

Dr. Waltz was also instrumental in establishing interdisciplinary research centers: the Beckman Institute at the University of Illinois, and the Volen National Center for Complex Systems at Brandeis.

Ryan Andrew Beavers

It is with great sadness that we note the sudden death of computer science student Ryan Beavers on Saturday, April 21. Ryan was a junior in CS and was conducting research into mobile computing with Prof. Robin Kravets, working to create a vehicular test-bed running on AMD Geode-based systems. His work on the project aimed to design and experiment with new opportunistic protocols that could enable efficient data uploading and data aggregation for vehicle-to-vehicle connectivity.

Ryan was a course assistant for CS 125 in past semesters and was a member of the ACM Gamebuilders organization.

"Ryan was one of those students you noticed because he was friendly and curious, and he cared about the concepts he was learning. He came to my office more than once to clarify his understanding of the material, and also just to chat about what courses he should take next," said CS lecturer Cinda Heeren. "I feel fortunate to have bumped into him in the stairwell just last week. It was only a quick hello, but I'm glad to have it as a memory."

Ryan was a loyal friend to many and will be greatly missed by all of those lives whom he touched.



gifts in action



Alex Chiang

CS@ILLNOIS is pleased to present Alexander Chiang as the inaugural winner of the Frima Lukatskaya Scholarship in Computer Science.

The Frima Lukatskaya Scholarship in Computer Science honors an incoming freshman student in the College of Engineering majoring in Computer Science who demonstrates remarkable passion about programming. The scholarship provides a full year of in-state tuition and housing assistance, and is renewable for up to three additional years with strong academic performance. In addition, the winner of the Frima Lukatskaya Scholarship will have an opportunity to meet with Max Levchin.

The Frima Lukatskaya Scholarship was established by alumnus Max Levchin in honor of his grandmother, who, through her remarkable life and career in science, taught him that talent and knowledge must be catalyzed by relentless drive to achieve excellence.


Alex joins the Engineering Computer Science Class of 2015 from North Allegheny Senior High School in Wexford, PA, where he received the school's President's Award for Educational Excellence, and the Distinguished Achievement Award. Alex was also

recognized as a North Allegheny Top Scholar and an AP Scholar with Distinction.

Initially attracted to Illinois because of its reputation in computer science, Alex said that after he applied, "Illinois just kept calling to me. Illinois was one of the first to offer me admission, the first to offer me honors admission, and the first to offer me a scholarship."

Drawn to computing at a young age, Alex began to teach himself programming principles in middle school. His explorations led him to begin teaching himself Java and ultimately led to his passion in programming for interactive entertainment.

Alex considers video game development and interactive entertainment to be the perfect blend of literature, art, and science. He enjoys the challenge of not only building the game story and creating its art, but also integrating a wide variety of concepts in computer science.



On Fire Emblem Zero, the game development project Alex has been working on for the last two years as a tribute to a favorite series, Alex serves as the sole designer, programmer, and artist. He plans to release his project to the open source community before the new school year begins to get critiques on his work so far.

Alex counts as his biggest challenge to date integrating classroom knowledge with his hands-on projects and real life coding. He considers his personal projects to be integral in helping to advance his knowledge of computing and programming.

In addition to game development, Alex also was a founding member of a semi-finalist team in the Cyber Patriot III Defense Team competition. The competition pitted Alex's team of white hat hackers against a team of experienced black hat hackers to test the team's skills at defending vulnerabilities and securing systems. Alex's team advanced from the opening round of 160 teams to the semi-final round of 36.

Alex is looking forward to gaining a solid foundation in computer science at Illinois, and hopes to continue to apply the computing concepts he learns in class to his own personal projects. During a visit to campus recently, Alex was struck by the atmosphere on campus and particularly in the Siebel Center. He hopes to participate in undergraduate research early on in his student career, and is looking forward to being a part of ACM and Gamebuilders.

"I am so grateful and thrilled to have received the Frima Lukatskaya Scholarship. What a great inspiration for me to become as successful as Max Levchin!"



The Frima Lukatskaya Scholarship

The Frima Lukatskaya Scholarship was established by CS@ILLINOIS alumnus Max Levchin (founder of PayPal, Slide, VP of Engineering at Google) in honor of his grandmother, who, through her remarkable life and career in science, taught him that talent and knowledge must be catalyzed by relentless drive to achieve excellence.

Siebel Scholars



top row (left to right): Swapnil Ghike,
Nipun Sehrawat, Akhil Langer

bottom row: Katrina Gossman,
Harshitha Menon

Swapnil Ghike volunteered to teach illiterate villagers and participated in NGO projects in his native India, and now is inspired to create sustainable technological solutions for the developing world. Katrina Gossman's simple idea to create one-way gates for randomly moving robots kickstarted a research program so innovative that it is generating an unusual amount of discussion and controversy in the field of robotics. Akhil Langer's SMS texting solution to help people find answers to medical questions put forth in natural language was awarded for its high impact factor in developing countries. Harshitha Menon's observation of parallelism in bee colonies inspired her to seek solutions at the intersection of high performance computing and distributed systems. Nipun Sehrawat's belief in the transformative potential of cloud computing and virtualization is resulting in solutions that reduce development time and improve developer productivity in large software development environments.

These five accomplished Computer Science graduate students have been named the 2012 Siebel Scholars. They join an elite group chosen on the basis of outstanding academic performance and demonstrated qualities of leadership. Each receives a \$35,000 prize award established by the Siebel Foundation to recognize the most talented students at the world's leading schools of business, computer science, and bioengineering.

"The Siebel Scholars Program recognizes students who have demonstrated academic and leadership excellence at the world's leading graduate schools of business, bioengineering, and computer science, and confirms the excellence of our institutions," stated Ilesanmi Adesida, dean of the College of Engineering. "We are very proud to be part of the Siebel Scholars Program in our efforts to create informed scholars and leaders, and to be considered among the top institutions in providing this interdisciplinary training."

SWAPNIL GHIKE

As Ghike learned during his time in India working on projects that addressed illiteracy, rural business models, and more, a social problem is generally harder than it appears, and many times, simple solutions are best though they may be the most difficult to find. His experiences in India gave him an appreciation for simple engineering solutions to the challenges of the developing world, and fostered his aspiration to create sustainable technological solutions that will serve developing nations. Ghike's research has focused on writing dynamic load balancing algorithms for parallel computation on thousands of processors and cross kernel optimizations for OpenCL programs. His current work with Professor David Padua aims to extract high performance from applications by applying compiler transformations and harnessing the parallelism offered by graphics processing units, without impacting programmer productivity.

KATRINA GOSSMAN

As an undergraduate, Gossman began exploring minimalist manipulation of simple ergodic bodies, where she was considered one of the most influential undergraduates ever to work in the robotics lab. Her design of simple gates to control randomly moving robots enforced a desired behavior at virtually no cost in terms of expense or energy. Continuing her research as a graduate student, Gossman is exploring minimalist solutions that can allow simple vehicles (in this case, a \$4 weasel ball) to achieve tasks that are typically performed by much more expensive and complex robotics. Her advisor, Professor Steve LaValle characterizes her research work as "so innovative that it has led to an unusual amount of discussion and controversy in the robotics community."

AKHIL LANGER

Langer's work as an undergraduate student in India has already garnered him recognition for creating technology solutions with a high impact factor for developing countries. His system to analyze

biomedical questions posed via noisy SMS text delivers answers to medical questions and helps the general population locate doctors, and also resulted in a national level award. His current work with Professor Kale focuses on collaborative, cross-disciplinary applications of high performance computing at peta- and exa-scale. His projects range from parallelizing the decision-making process for allocation of military and civilian aircraft to simulations using adaptive mesh refinement with applications in numerical cosmology, global atmospheric modeling and mantle convection modeling. He also works on scalable algorithms for optimizing communication in parallel programs.

HARSHITHA MENON

Observing the intrinsic parallelism found in nature, Menon was inspired to make the beauty of parallelism pervasive in computing. Towards her goal, Menon's research spans the areas of parallel computing and distributed systems. Working with Professor Kale, Menon's research aims to harness the power of multiprocessors and distributed systems to sustain the ever-increasing performance requirement of today's applications. Menon's passion for teaching has led her to create innovative teaching labs and collaborative problem solving methods in the computer science courses she currently teaches.

NIPUN SEHRAWAT

Sehrawat has been involved in various projects relating to virtualization and cloud computing since his undergraduate years. His current work in the area focuses on scalability and performance aspects of the emergent fields, with particular emphasis on highly available and replicated cloud databases. His work with Professor Indranil Gupta to develop a distributed build solution led to the development of an algorithm that outperforms the next best solution, and may lead to shorter development cycles and increased productivity for large software development firms.



Kolla and Lazebnik Join CS@ILLINOIS



Two new faculty will be welcomed at Illinois beginning in 2012! Alexandra Kolla (Microsoft Research) and Svetlana Lazebnik (University of North Carolina – Chapel Hill) have joined the faculty beginning in January 2012.

Tackling some of the hardest problems in their field, these accomplished young researchers and educators are exploring innovative new avenues in complexity theory and computer vision.

Alexandra Kolla research lies at the intersection of theoretical computer science and discrete mathematics, with a particular interest in the mathematical tools used to solve problems in graph theory, convex optimization, approximation algorithms, and complexity theory. Inspired by spectral graph theory and semidefinite programming (sDP), Kolla aims to understand how these tools can be used to solve outstanding questions in the field of algorithms and complexity theory.



Kolla adds terrific depth to the College's expertise in complexity theory. She is working on some of the most difficult and most high-impact problems, trying to answer fundamental questions about what we can and cannot compute efficiently.

Svetlana Lazebnik's research interests in computer vision have led her to explore topics ranging from modeling and organizing large-scale Internet image collections to developing effective image representations for recognition and comprehensive 2D and 3D scene descriptions. Her work has implications for assistive technologies, security and surveillance, navigational systems, field biology, and more.

Lazebnik joins the world-leading computer vision group at Illinois. Her innovative work to integrate statistical and geometric techniques to describe images, and her use of information theoretic models to model famous landmarks and entire cities from very large-scale image collections, will open up exciting new collaborations at Illinois.

Gropp Co-Founds ACM High Performance Computing Group

University of Illinois computer science professor Bill Gropp was one of the driving forces behind the launch of a new ACM Special Interest Group on High Performance Computing. Gropp and several others founded the group to address the needs of students, faculty, and practitioners of high performance computing. SIGHPC is the first such group within a major professional society to address this emerging and important topic in computing.

SIGHPC will help address this challenge for professionals and students alike by:

- disseminating research and experiences by those using computing resources to tackle our society's toughest problems;
- promoting the mentoring and education of the next generation of HPC professionals; and
- serving as a source of information about the field to the larger computing and scientific communities.

Snir Named HPC Person to Watch in 2012

Professor Marc Snir has been named one of HPCWire's People to Watch in 2012.

Each year, HPCwire announces its annual "People to Watch" list, comprised of an elite group of community leaders selected from academia, government, business, and industry who we predict will be impacting the world in 2012 and beyond.

"The recent flurry of news last fall surrounding IBM's departure from the 10 petaflop NCSA petaflop "Blue Waters" Project and Snir's role as a co-PI for the project will share the focus of a lot of attention this year as the supercomputer nears its delivery date, slated for fall 2012. Combined with his role as a one of the original co-founders and as a co-chair for the relatively new Graph 500, all but guarantees that Snir will be a person of interest to follow this year," writes HPCWire in announcing Snir's selection.

Snir has been a leader in shaping high performance computing (HPC) architectures and parallel programming, including contributions to IBM's SP and Blue Gene systems and to MPI, the standard communications library used in HPC.



Snir to Lead Mathematics and Computer Science Division of Argonne

Illinois computer science professor Marc Snir has been chosen to direct the Mathematics and Computer Science Division (MCS) at Argonne National Laboratory (ANL). Snir, the Michael Faiman and Saburo Muroga Professor in computer science, has been a leader in shaping high performance computing (HPC) architectures and parallel programming, including contributions to IBM's SP and Blue Gene systems and to MPI, the standard communications library used in HPC. At MCS, he will be directing over 200 researchers and staff members, who are working on projects ranging from algorithm development and software design in key areas like optimization, to exploration of new technologies such as distributed computing and bioinformatics, to numerical simulations in challenging areas like climate modeling. Snir will continue to hold his appointment as professor of computer science. He will divide his time between MCS and the University of Illinois and will continue to be associated with the Blue Waters project.

"Argonne is one of the most prestigious national labs, conducting significant research across many key disciplines. This is a rare and outstanding opportunity, not only for Marc, but also for the Department of Computer Science and the College of Engineering. The potential for new collaborations between our two world-class institutions is exciting," said Rob A. Rutenbar, the Abel Bliss Professor of Engineering and computer science department head.

A distinguished researcher and scholar, Snir chaired the Department of Computer Science at the University of Illinois, Urbana-Champaign from 2001 to 2007. While at Illinois, he has also co-directed the Intel and Microsoft Universal Parallel Computing Research Center, was the first director of the Illinois Informatics Institute, is the Associate Director for Extreme Scale Computing at NCSA, and is co-PI of the Blue Waters petascale computing project. In addition, Snir co-chaired the National Research Council's Committee to Study the Future of Supercomputing, and he is a co-author of its influential 2004 report, "Getting Up to Speed: The Future of Supercomputing."





Han named Bliss Professor

Professor Jiawei Han was among six College of Engineering faculty members to be recognized with Bliss Professorships.

The generous Bliss bequest, established by Helen Eva Bliss in memory of Abel Bliss Jr., is used to advance scholarly activities in the College of Engineering. Holders of college professorships are nominated by the dean upon recommendation of the College Advisory Committee on Endowed Appointments and approval of the Provost. Faculty members with named professorships are bestowed to recognize leaders who are among the most talented and accomplished on our faculty.

Han's groundbreaking and highly influential research has made him one of the top computer scientists in the world. With a focus on knowledge discovery and data mining, data warehousing, and database systems, he is recognized as a pioneer in the field. Han was the first to introduce a pattern-growth methodology for mining frequent, sequential, and structured patterns, as well as the first to develop a set of important algorithms for mining such patterns. These methodologies have been extremely influential in subsequent research and are widely used. Google Scholar lists Han as the second most cited author in the field of data mining with his FP growth algorithm cited more than 3,700 times. This algorithm or its variations have been introduced in most data mining textbooks and has been used at various companies such as Google and Microsoft.

Torrellas Named Program Chair for ISCA 2012



Professor Josep Torrellas has been named the Program Chair for the 39th Annual International Symposium on Computer Architecture (ISCA) in 2012. ISCA is the premier forum for new ideas and experimental results in computer architecture, covering topics ranging from parallel and multi-core systems, interconnection networks, dependable architectures, power and energy efficient architectures, and more.

As Program Chair, Torrellas will be charged with the technical contents of the conference: the papers presented, keynote speeches, and panels. Torrellas will select the program committee members and external reviewers that will review the papers, prepare the final program, and choose the keynote speakers and the panels in the conference, among other responsibilities.

"I am honored at this appointment because ISCA is the premier conference in computer architecture. It allows me to contribute with my efforts to the continued excellence of our research field," said Torrellas. "I am

particularly excited at the fact that our field is now bursting with new concepts that range from data centers to hand-held devices, and from novel memory and interconnect technologies to DNA-based and brain computing. I hope the final program captures the most exciting aspects of this research."

Torrellas and his students have been a major contributor to the ISCA conference for many years. At ISCA 2011, he and his ex-Ph.D. students authored 6 of the 40 papers presented. The papers were authored by Jose Renau from University of California Santa Cruz, Milos Prvulovic from Georgia Tech, Yan Solihin from North Carolina State University, Michael Huang from University of Rochester, and by Josep Torrellas.

Other Illinois researchers and alumni will be making contributions to ISCA 2012 as part of the organizing and program committee, including Sarita Adve, Wonsun Ahn, Pradip Bose, Luis Ceze, Tom Conte, Xin Fu, Scott Mahlke, Jose Martinez, Trevor Mudge, Jose Renau, Yan Solihin, and James Tuck.

Roth Named ACM Fellow

Professor Dan Roth was named a 2011 ACM Fellow for his “contributions to machine learning and natural language processing.” The ACM Fellows Program celebrates the exceptional contributions of the leading members in the computing field. These individuals have helped to enlighten researchers, developers, practitioners and end-users of information technology throughout the world.

Dr. Roth’s research interests focus on the computational foundations of intelligent behavior. He has made seminal contributions to several aspects of this problem — from theoretical questions in Learning and Reasoning, to new models and algorithms that have contributed significantly to the use of machine learning and inference techniques in natural language processing. Advances made by Roth have changed how computer scientists develop algorithms and programs for natural language understanding, and how they think about computational modeling of learning and reasoning.

In his research, Roth has pursued several interrelated lines of work that span multiple aspects of this problem — from fundamental questions in learning and inference and how they interact, to the study of a range of natural language processing (NLP) problems, including multiple disambiguation problems, shallow parsing, semantic role labeling, co-reference, question answering and textual entailment, to large scale Natural Language Processing and Information Extraction system development - resulting in a number of software packages for NLP tools.

He has published broadly in machine learning, natural language processing, knowledge representation and reasoning and learning theory, and has developed advanced machine learning based tools for natural language applications that are being used widely by the research community, including an award winning Semantic Parser.

Dr. Roth is the Director of the DHS Institute of Discrete Science Center for Multimodal Information Access & Synthesis Research (MIAS). He is a fellow of the Association for the Advancement of Artificial Intelligence, the premier AI professional society and was an NSF CAREER Award winner. Roth received his Ph.D. in Computer Science from Harvard University in 1995 and is a Willett Faculty Scholar of the University of Illinois College of Engineering.



Adve Named IEEE Fellow

Professor Sarita Adve has been named an IEEE Fellow for the class of 2012. Adve was selected for her “contributions to shared memory semantics and parallel computing.” Adve’s research in computer architecture and systems, parallel computing, and power and reliability-aware systems focuses on a full-systems view and is notable for its multidisciplinary collaborations.

Adve’s broadest impact has been in hardware and software memory consistency models. She received the 2008 SIGARCH Maurice Wilkes award for this work, specifically “for formalization of memory consistency models, especially data-race free models, and their influence on both hardware and high-level languages.”

The memory consistency model lies at the heart of the semantics of any threaded software or hardware. Arguably, it has been one of the most challenging and contentious areas in concurrent hardware and software specification for many years. There is finally now a convergence in both the hardware and software communities. Adve has been a common thread and a leader in the multiple community-scale efforts that have driven this convergence.

The memory consistency model is a hardware/software interface, affecting programmability and performance. Unfortunately, designing a model satisfying all desirable properties has proven difficult. Sequential consistency is simplest to program, but most systems do not provide it for performance reasons. Instead, we had divergent models (often ambiguously specified) for different hardware.

Adve’s early work departed from the prevalent hardware-centric approaches to use a combined hardware/software view more appropriate for an interface. She observed that for well-synchronized programs (which she formalized as data-race-free), both sequential consistency and high performance can be provided. She developed a comprehensive framework to specify memory models as providing “sequential consistency for data-race-free programs.” Adve’s data-race-free model forms the foundation of the memory models for Java and C++.

Adve has been a leader in power- and reliability-aware architectures. Her group was among the first to recognize that significant power reductions required breaking traditional system boundaries in favor of a collaborative, cross-layer system-wide power management framework. The GRACE project she led was the first to demonstrate a prototype system where the hardware, operating system, network, and applications all adapted collaboratively to minimize energy while still meeting real-time quality of service requirements.

Most recently, Adve’s work has challenged the research community to rethink how both parallel languages and parallel hardware are designed. Adve recently won honors for her DeNovo system, which exploits emerging software trends in disciplined parallel programming to make hardware simpler, higher performance, and lower energy, all at the same time. As their first step, Adve and her research team have developed a cache coherence protocol and consistency model that takes an order of magnitude less time to verify and runs some applications in less than half the time with less than half the network traffic and cache misses than the state-of-the-art. The simplicity and low network and cache traffic means that the performance increases come with significant power and energy benefits, a rarity in hardware design.

Professor Adve received the Maurice Wilkes Award in 2008, an IBM Faculty Award in 2005, was named a UIUC University Scholar in 2004, and received an Alfred P. Sloan Research Fellowship in 1998, an IBM University Partnership award in 1997 and 1998, and a NSF CAREER award in 1995.

She served on the National Science Foundation’s CISE directorate’s advisory committee from 2003 to 2005 and on the expert group to revise the Java memory model from 2001 to 2005. She co-led the Intel/Microsoft funded Universal Parallel Computing Research Center (UPCRC) at Illinois as its director of research in its founding year (2008-09). She currently serves on the board of directors for ACM SIGARCH and the Computing Research Association (CRA).



Learn more about Adve’s recent work on page 25

Adve Named IEEE Fellow



Professor Sarita Adve has been named an IEEE Fellow for the class of 2012. Adve was selected for her “contributions to shared memory semantics and parallel computing.” Adve’s research in computer architecture and systems, parallel computing, and power and reliability-aware systems focuses on a full-systems view and is notable for its multidisciplinary collaborations.

Adve’s broadest impact has been in hardware and software memory consistency models. She received the 2008 SIGARCH Maurice Wilkes award for this work, specifically “for formalization of memory consistency models, especially data-race free models, and their influence on both hardware and high-level languages.”

The memory consistency model lies at the heart of the semantics of any threaded software or hardware. Arguably, it has been one of the most challenging and contentious areas in concurrent hardware and software specification for many years. There is finally now a convergence in both the hardware and software communities. Adve has been a common thread and a leader in the multiple community-scale efforts that have driven this convergence.

The memory consistency model is a hardware/software interface, affecting programmability and performance. Unfortunately, designing a model satisfying all desirable properties has proven difficult. Sequential consistency is simplest to program, but most systems do not provide it for performance reasons. Instead, we had divergent models (often ambiguously specified) for different hardware.

Adve’s early work departed from the prevalent hardware-centric approaches to use a combined hardware/software view more appropriate for an interface. She observed that for well-synchronized programs (which she formalized as data-race-free), both sequential consistency and high performance can be provided. She developed a comprehensive framework to specify memory models as providing “sequential consistency for data-race-free programs.” Adve’s data-race-free model forms the foundation of the memory models for Java and C++.

Adve has been a leader in power- and reliability-aware architectures. Her group was among the first to recognize that significant power reductions required breaking traditional system boundaries in favor of a collaborative, cross-layer system-wide power management framework. The GRACE project she led was the first to demonstrate a prototype system where the hardware, operating system, network, and applications all adapted collaboratively to minimize energy while still meeting real-time quality of service requirements.

Most recently, Adve’s work has challenged the research community to rethink how both parallel languages and parallel hardware are designed. Adve recently won honors for her DeNovo system, which exploits emerging software trends in disciplined parallel programming to make hardware simpler, higher performance, and lower energy, all at the same time. As their first step, Adve and her research team have developed a cache coherence protocol and consistency model that takes an order of magnitude less time to verify and runs some applications in less than half the time with less than half the network traffic and cache misses than the state-of-the-art. The simplicity and low network and cache traffic means that the performance increases come with significant power and energy benefits, a rarity in hardware design.

Professor Adve received the Maurice Wilkes Award in 2008, an IBM Faculty Award in 2005, was named a UIUC University Scholar in 2004, and received an Alfred P. Sloan Research Fellowship in 1998, an IBM University Partnership award in 1997 and 1998, and a NSF CAREER award in 1995.

She served on the National Science Foundation’s CISE directorate’s advisory committee from 2003 to 2005 and on the expert group to revise the Java memory model from 2001 to 2005. She co-led the Intel/Microsoft funded Universal Parallel Computing Research Center (UPCRC) at Illinois as its director of research in its founding year (2008-09). She currently serves on the board of directors for ACM SIGARCH and the Computing Research Association (CRA).

Learn more about Adve’s recent work on page 25

Department of Computer Science
University of Illinois at Urbana-Champaign
Thomas M. Siebel Center for Computer Science
201 N Goodwin Ave
Urbana, IL 61801

All gifts make a difference in upholding our past legacy,
preserving our present excellence, and reaching out to future
generations. If you would like to make a contribution, please
visit <http://cs.illinois.edu/alumni/giving> or scan this QR code



click!

The word "click!" is written in a large, white, sans-serif font. The letter "i" is highlighted in orange. A white mouse cursor arrow is positioned to the right of the exclamation mark, pointing towards it.

cs.illinois.edu