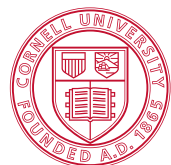




CIS
ANN
UAL
REP
ORT
20
07

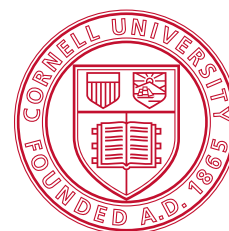
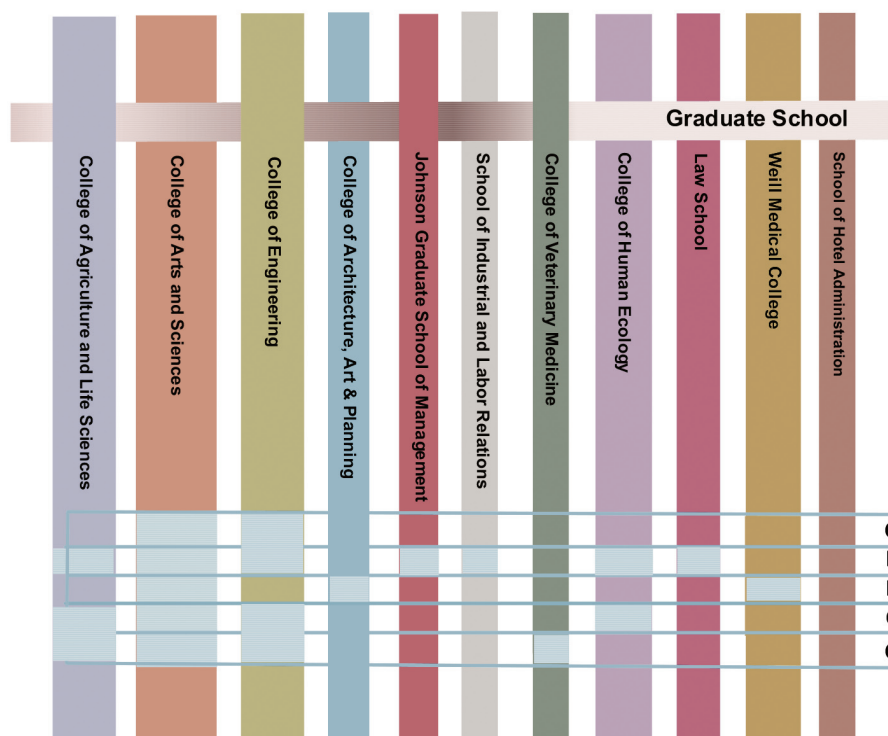




CIS ANNUAL REPORT 2007

TABLE OF CONTENTS

- 03 Message from the Dean
- 05 CIS Units
 - CIS Departments
 - Department of Statistical Science
 - Department of Computer Science
 - CIS Programs
 - Program of Computer Graphics
 - Information Science Program
 - CIS Centers and Institutes
 - The Cornell Theory Center
 - Intelligent Information Systems Institute
 - Information Assurance Institute
 - National Science Digital Library at Cornell
 - Computer Systems Laboratory
- 10 A New Start: Innovative Introductory CS Courses at Cornell
 - Computer Graphics
 - Trustworthiness
 - Scientific Computing
 - Artificial Intelligence
 - Systems
 - Database Systems
 - Theory of Computing
 - Programming Languages
- 13 Research Focus
 - Computational Biology
- 18 BOOM
- 19 Awards
- 21 Corporate and Foundation Interactions
- 23 Alumni Affairs
 - Lecture Series
- 25 2006-2007 CIS Courses
- 27 2006-2007 Research Grants
- 35 Faculty Bios



CIS



A Message from the Dean

OF THE FACULTY OF COMPUTING AND INFORMATION SCIENCE

ROBERT CONSTABLE

One of the key events this year for all Cornellians was the inauguration of David J. Skorton on September 7th as our 12th president. To quote the press, our new president is a "... cardiologist, computer scientist, national leader in research ethics, and jazz musician." Over the past year, I've had the opportunity to travel with our new president and to work closely with him on a number of questions and have learned firsthand that he brings a deep and passionate appreciation for the power of the Information Revolution to his new role.

The information sciences are changing the world and the role of universities are being transformed by that same change. In his first commencement speech President Skorton spoke of the sense in which "... innovation has led to enormous economic growth; the foundation of innovation is research; and the seat of fundamental research is the university. ... As a result of its core missions of education and research, the university reaches out materially and directly to assist and improve the quality of life." He called for American universities to collectively use their influence to improve the condition of the peoples of the world.

It seems to me that this accurately captures the challenge before us. CIS, perhaps more than any unit in the university, is profoundly immersed in the complex realities of this worldwide revolution. Our students come from every corner of the "flat world" and go on to pursue careers that reshape science, technology, and global culture. Some have played key roles in creating entirely new industries; others less visible ones, but I think that Cornell graduates are exceptionally well represented among the ranks of those who are innovating, inventing, making the world smarter, faster, more open. I believe that through these roles, we have a unique opportunity to make sure that the new world we create is a better place. This, to me, is the real CIS opportunity and a major contribution to Cornell's capability to change the world. Here is how we have expressed the opportunities and challenges we in CIS face in the coming years.

Computers and digital information transform how we think, discover, and learn. They change everything they touch. The PC and the Web are symbols of the Information

Revolution, one of the most profound intellectual and technical revolutions in history.

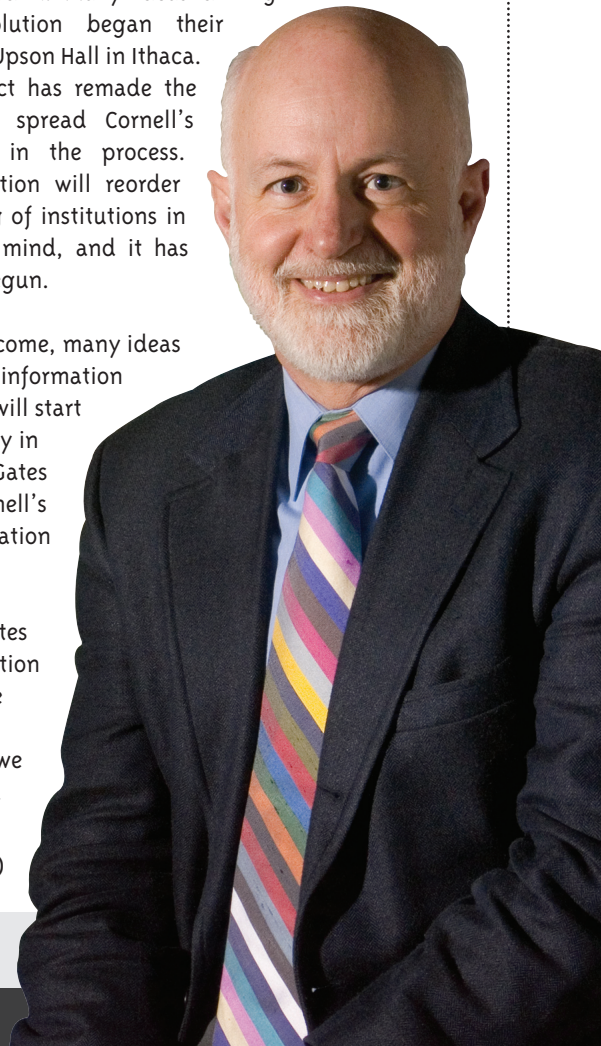
While the Industrial Revolution was about automating mechanical processes using materials, the Information Revolution is about automating intellectual processes using digital information. In CIS, researchers study all information processes—those programmed on networks of computers, those modeling theories of physics, those sustaining living systems from cells to brains, and those designed to capture features of minds and other intelligent systems.

"The path of the Information Revolution is of abiding concern to all peoples of the world."

The path of the Information Revolution is of abiding concern to all peoples of the world. Its course impacts their health and prosperity. It provides new means to monitor the condition of our planet, a "digital skin" for the earth. It gives people new ways to connect and express themselves. It is a democratizing revolution consistent with America's core values, and it is a source of special pride to Cornellians because Cornell is among the handful of universities who have led it. Many ideas driving the revolution began their journey in Upson Hall in Ithaca. Their impact has remade the world and spread Cornell's reputation in the process. This revolution will reorder the ranking of institutions in the public mind, and it has only just begun.

In years to come, many ideas driving the information revolution will start their journey in William H. Gates Hall on Cornell's new information campus. Association with Mr. Gates is a recognition of where we have been and where we are headed.

(continued)



The support of Cornell's alumni will be a major factor in determining when we achieve our goals—goals set by some of the world's leading computer scientists, information scientists, computational biologists, statisticians, and others in the Faculty of Computing and Information Science. Our goals point the way to accelerating scientific advances, innovation to meet human needs, and continuing leadership of the great revolution.

In his commencement speech, President Skorton noted a steady erosion of government support for universities. We live in a difficult period and are increasingly dependent upon our relationship with the private sector, with our alumni and friends with whom we find common cause to sustain our capacity to innovate and educate.

You may be aware that the president and the Board of Trustees have launched a four billion dollar fundraising campaign to provide resources for students, faculty, and facilities. All of the colleges, schools, and faculties have statements on the campaign home page.

Our faculty organized their most pressing needs around constellations of people and spaces, associating endowed chairs and graduate fellowships with laboratories and research groups. We seek funding for several essential clusters such as: a Computer Systems Lab, a Graphics Lab, a Computational Biology Lab, an AI Lab, an HCI Lab, a Scholarly Communications Group, and the oldest cluster, the Theory Group.

The campaign theme is "Far Above," from lyrics to Cornell's alma mater, and the Faculty of CIS is well represented, having goals to complete fundraising for William H. Gates Hall, the signature building on the Information Campus and to raising endowment for professorships, for the deanship, and for graduate fellowships. Our alumni readers of this report might want to look at the campaign home page where the CIS campaign statment is displayed.

In the context of the overall campaign, we are confident that we can raise \$100 million dollars. One reason for this confidence is the exceptional standing of the CIS departments and programs, which are often cited by the university leadership as examples of how Cornell can thrive in a new era. Our standing in computing and information science is not simply a linear ranking by news magazines, which usually put computer science as a department at number five; this year it is rated 5th among all 235 computer science departments by *US News & World Report* and among the top ten in three of the four subareas ranked—theory, programming languages,

and AI—making the Department of Computer Science profile unique among all the Cornell departments rated by that magazine. Another tentative national evaluation this year put the Information Science Program as number one among two dozen or so similar programs. Statistical Science, also located within CIS, was rated number six as a subarea of mathematics.

Rankings by the popular press are flattering, but there are also many concrete ways to measure our standing in the community. For example, this year Eva Tardos was elected to the National Academy of Engineering and was awarded an endowed University chair, the Jacob Gould Schurman chair, named after Cornell's third president and U.S. ambassador to China. (He was also the dean of the Sage School of Philosophy at Cornell before becoming president.) Jon Kleinberg won the highly prestigious Nevanlinna Prize from the International Mathematical Union (IMU), and he was elected to the American Academy of Arts and Sciences. John Hopcroft won the Distinguished Service Award from the Computing Research Association for his outstanding service to computer science while serving on the National Science Board. Fred Schneider was elected to the Computing Research Association Board and to the governing Council of the Computing Community Consortium (CCC Council). Lillian Lee won the Provost's Award for Distinguished Scholarship. The computer science faculty also won three NSF CAREER awards this year, given to Kavita Bala, Uri Keich, and Bobby Kleinberg.

Several information scientists were also honored, including Paul Ginsparg receiving the prestigious Evans Peters Award from Educase, ARL, and CNI, Carla Gomes elected as a fellow of the Association for the Advancement of Artificial Intelligence (AAAI), Geri Gay was appointed to the Kenneth J. Bissett endowed chair in Communication, and Phoebe Sengers won a Cornell Society for the Humanities Faculty Fellowship.

In computational biology, Adam Siepel won a Microsoft Faculty Fellowship, one of only five awarded this year, and an NSF CAREER award. Hod Lipson won a DARPA Young Faculty Award.

I couldn't be more proud, or more excited, about the team we've assembled within CIS. Across the board, I'm seeing recognition of our strength and impact, noteworthy accomplishments of talented faculty, and some of the most dedicated and talented students in the world. This proven combination is a sound basis for pride in Cornell and hope for a better world at a time when it needs more than ever the best of what America has to offer.

CIS DEPARTMENTS

The Department of Statistical Science

Statisticians work at the interface of mathematics, computing, and data analysis. Statistics is, at its core, about representing information. Notions like expectation and variance are meant to provide compact descriptions of large amounts of data. Large amounts of data are the stock in trade of many disciplines, so statisticians often are involved in analyzing data generated by research in other disciplines.

In biology, the role of statistics is so significant—in problems ranging from the design of clinical experiments to the analysis of genetic, genomic, and proteomic data—that a whole new field of biostatistics has emerged. Computer science also has great interest in problems of information representation, there is a great deal of work at the interface of computer science and statistics, particularly on topics such as machine learning (automatic classification of data) and data mining (finding interesting patterns in large datasets). Indeed, we seem to be witnessing the beginning of a field of computational statistics.

The Department of Statistical Science (DSS) offers a Master of Professional Studies (MPS) degree in applied statistics and a Doctor of Philosophy (MS/PhD) degree in statistics. The academic and research programs at DSS take advantage of Cornell's extensive resources, drawing from many colleges and research groups.

For more information, see <http://www.stat.cornell.edu/>

The Department of Computer Science

In recent years the department has complemented its strength in the essential core areas (artificial intelligence, programming languages, scientific computing, systems, and theory) with commitments also to databases, graphics, machine learning, and natural language processing. Cornell's Department of Computer Science is widely known for its collegiality and its ability to move rapidly into new interdisciplinary fields such as information science. The department's challenge is to be simultaneously strong in the core areas of computer science, while being a major force

in emerging and interdisciplinary areas that are being developed within CIS.

The high-level research goals are:

- to help science, engineering, and business by making it possible for those communities to (a) interpret unprecedented volumes of data and (b) build models and simulations to cope with exceedingly complex problems,
- to design and build the efficient, scalable, secure, and reliable systems that will form the foundation of our information-based economy and national infrastructure,
- to develop algorithms and languages that extend and automate the human capacity to visualize, reason, and compute,
- and to advance the theoretical foundations of computer science with new mathematics that support the analysis of complex systems and computational processes.



CS faculty members Dexter Kozen, David Gries, Fred Schneider, Bill Arms, and Graeme Bailey at the annual CS faculty retreat.

This broad agenda is made possible by faculty whose research interests radiate in many directions and whose commitment to teaching is deep. The computer science major in the College of Engineering and the computer science major in the College of Arts and Sciences emphasize breadth across the field and encourage students to develop a strong expertise in a second area. The one-year Master of Engineering program gives students the opportunity to take a range of forefront courses and to develop skills that enhance their marketability. The PhD program is highly ranked and over the past forty years has graduated many students who have gone on to influence the course of computer science research in the world.

For more information, see <http://www.cs.cornell.edu/>

CIS PROGRAMS

Program of Computer Graphics

Computer graphics is a rapidly evolving field that has had significant impact on most scientific, artistic, and engineering fields. Graphics is the most common and efficient means of man-machine communication. At Cornell, research and teaching in computer graphics are centered in the computer science department and the closely affiliated Program of Computer Graphics (PCG), one of the world's leading computer graphics laboratories and a dominant force in the international computer graphics community for more than thirty years.

Research in graphics requires a multidisciplinary team with knowledge in algorithms, systems, numerical simulation, machine vision, software and hardware engineering, physics, optics, and perception psychology. The PCG is particularly famous for both its work on realistic rendering—simulating environments that are physically accurate and perceptually indistinguishable from real world scenes—and its influence on the architectural design and computer animation industries. Current graphics research at Cornell includes realistic interactive rendering, advanced material modeling, modeling complex scenes, image-based modeling and rendering, animation, and display technology.

Information Science Program

Information Science (IS) is an interdisciplinary program within CIS. It brings together faculty, researchers, and students who share an interest in the boundary between computer science and the social sciences. The program studies digital information systems in their social, cultural, economic, historical, legal, and political contexts. The IS academic programs are divided into three streams: information systems, which draws from computer science and operations research; human computer interaction, which includes communication and cognitive psychology; and social studies of computing, which draws from science and technology studies, law, sociology, and economics.

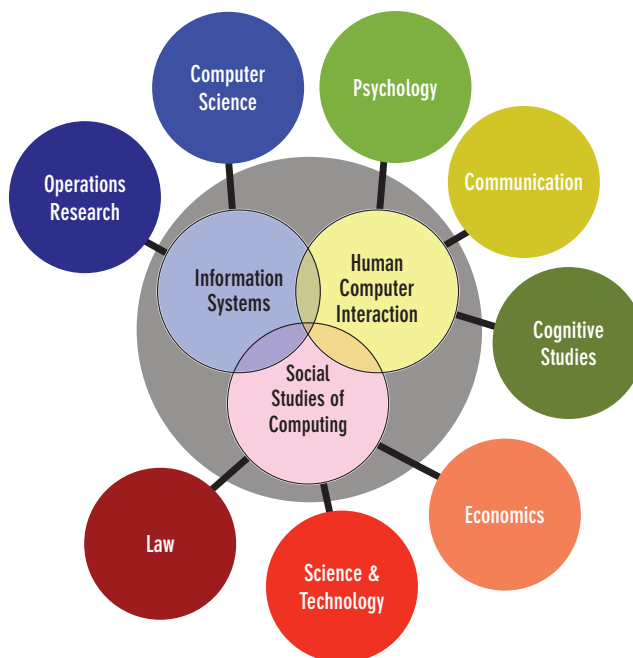
Although the program is quite new, it has already drawn national attention: its faculty were ranked number one in the nation for scholarly productivity in a new index published in the Chronicle of Higher

Education. The small, elite PhD program will have nine students for the 2007–2008 academic year and graduated its first student in 2006.

Our undergraduates—the program offers a major in the College of Arts and Sciences, the College of Agriculture and Life Sciences, and in the College of Engineering—garnered top internship positions at Google and Microsoft for the summer of 2007.

Our interdisciplinary research programs offer a wide range of opportunities for graduate students and undergraduates and are led jointly by faculty from different departments. They include research in Human Computer Interaction, the ePrint arXiv, the National Science Foundation's National Science Digital Library, the Cornell Web Lab, the Cornell eRulemaking Initiative (CeRI), and projects on Deception in Computer-Mediated Communication, Affective Computing, and Social Network Analysis. The 2006 theme project for Cornell's Institute for the Social Sciences, "Getting Connected: Social Science in the Age of Networks," is based on information science research.

For further information see
<http://www.infosci.cornell.edu/>



Information Science examines information systems in their social, cultural, economic, historical, legal, and political contexts. Computer science is an important part of the program, but the emphasis is on systems and their use, rather than on the technologies that underlie them.

The diagram symbolizes the three major areas of Information Science and some of the schools, departments and programs that contribute to each area.

CIS CENTERS AND INSTITUTES

The Cornell Theory Center

The Cornell Theory Center (CTC) is Cornell's high-performance computing and interdisciplinary research center. CTC is focused on providing cyberinfrastructure resources for research and education; these resources include high-performance and data-intensive computing hardware and expertise, visualization, and K-12 outreach. Scientific and engineering projects supported by CTC represent a wide variety of disciplines, including bioinformatics, behavioral and social sciences, computer science, engineering, geosciences, mathematics, physical sciences, and business. The center serves more than 150 faculty research groups across campus and at the Weill Medical College.

CTC is widely recognized for pioneering the use of industry-standard hardware and software to create a productive, large-scale, computing environment. Strategic partnerships with Microsoft and Intel were critical in achieving this vision. CTC's platform expertise includes AIX, Linux, MAC OS X, Windows Server 2003, Windows Compute Cluster Server 2003, and UNIX. Supporting ninety active research projects, CTC runs a cluster computing environment with approximately 1800 processors.

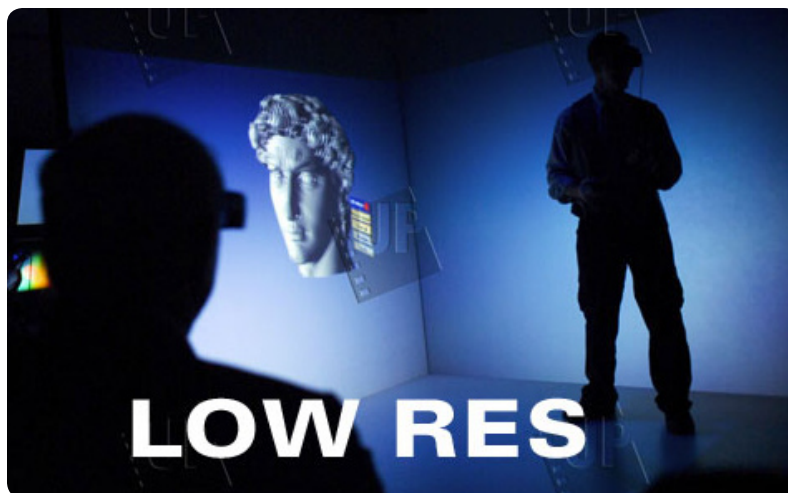
Focused on providing cyberinfrastructure resources for research and education, CTC is one of the founding institutions of NYSGrid. Currently, nineteen New York institutions have joined to create NYSGrid, a 21st century cyberinfrastructure initiative that will provide its constituency with unprecedented resources for research, education, and community outreach. NYSGrid resources will aggregate high-end computing, networking, data storage, visualization, and most importantly, intellectual capital from sites across the State.

CTC also operates a CAVE, providing a three-dimensional, stereo immersive virtual reality environment for

viewing scientific, engineering, architectural, and art applications. The CAVE enhances a variety of research projects and undergraduate courses, including an architectural course and an engineering design course that is taught in conjunction with the NASA/AT&T-sponsored Advanced Interactive Discovery Environment for the Engineering Education Project.

CTC is the home for many large interdisciplinary research projects at Cornell. CTC's Computational Biology Service Unit (CBSU) (<http://cbsu.tc.cornell.edu>) is headed by Andrew Clark, professor of Molecular Biology and Genetics and applies computational resources and expertise to a variety of applications in life sciences, ranging from canine genetics and plant breeding to protein structure modeling. The CBSU is also active in training under-graduates through the CBSU Undergraduate Summer Internship program.

CS professor Johannes Gehrke leads CTC's data-intensive computing initiatives, applying his data mining expertise to a breadth of applications,



including three petabyte-scale data-intensive projects: large-scale astronomical surveys using the Arecibo radiotelescope, physically accurate rendering in computer graphics, and a study of the structure and evolution of the World Wide Web. Professor Gehrke also heads up CTC's new eScience Unit (eSU). eSU provides a breadth of services to researchers with data-intensive applications-database systems and data storage management, database programming and consulting, data curation, and data mining.

CTC also does pioneering work in science communication, outreach, and informal education through the use of multi-user virtual environments.

SciCentr.org, an online multi-user museum, engages youth through the appeal of social gaming. CTC shares research conducted by Cornell scientists and faculty in the fields such as biotechnology, astrophysics, materials science, computer science, communication, fine arts, theatre arts, music, and architecture. The SciCentr museum includes interactive exhibits and laboratories based on computer game technology created by interdisciplinary teams of undergraduate programmers, designers, and content developers. These teams rely on the talents of CS undergraduates.

CS also contributes key members to the pool of undergraduate mentors who support teams of middle and high school students creating their own virtual exhibits at remote locations through SciCentr's SciFair science communication program. CTC began a partnership with Cornell's Public Service Center this year to enhance the community service experience of the SciFair mentors.

For more information, see <http://www.tc.cornell.edu/>

Intelligent Information Systems Institute

The mission of the Intelligent Information Systems Institute (IISI), founded in December of 2000, is to perform and stimulate research in computational and data-intensive methods for intelligent decision-making systems, to foster collaborations within the scientific community, and to play a leadership role in the research and dissemination of the core areas of the institute. The institute is funded by the Air Force Research Laboratory, with Carla Gomes as director.

The IISI supports basic research within Computing and Information Science, promoting a cross-fertilization

of approaches from different disciplines, including computer science, engineering, operations research, economics, mathematics, statistics, and physics. Areas of research within the IISI are search and complexity, planning and scheduling, large-scale distributed networks, data mining and information retrieval, reasoning under uncertainty, natural-language processing, machine learning, multi-agent systems, and combinatorial auctions.

For further information, see
<http://www.cis.cornell.edu/iisi/>

Computer Systems Laboratory

The Computer Systems Laboratory (CSL) in the School of Electrical and Computer Engineering brings together faculty with common interests from the departments of electrical and computer engineering (ECE) and computer science (CS) at Cornell.

The field of computer systems is both experimental and theoretical, having grown out of computer architecture, parallel computer architecture, operating systems and compilers, computer protocols and networks, programming languages and environments, distributed systems, very large scale integration (VLSI) design and fabrication, and system specification and verification.

Graduate students are admitted to either ECE or CS. Usually students with primary interest in computer architecture, multiprocessor design, VLSI, computer-aided design (CAD), and circuit design enroll in ECE, while students with interest in compilers, operating systems, and programming environments enroll in CS. There are no rigid student classifications; ECE students can have a thesis advisor in CS and vice versa. Indeed, the interdisciplinary composition of the research teams is a strength of the Computer Systems Laboratory.

For further information, see
<http://www.csl.cornell.edu/>



IISI director, Carla Gomes.

National Science Digital Library at Cornell

The National Science Digital Library (NSDL) program was established by the National Science Foundation's Division of Undergraduate Education in 2000 as a free, online library that directs users to exemplary resources for science, technology, engineering, and mathematics (STEM) education and research. Since 2001, the NSDL team in CIS at Cornell has been funded to develop the technical infrastructure and run the production facilities of the NSDL. Cornell works together with two other partners, the University Corporation for Atmospheric Research (UCAR) and Columbia University, to provide the core integration for the entire NSDL program. Through the efforts of Cornell and the many other NSF-funded NSDL projects, NSDL is now a production library of well over two million carefully selected STEM resources from over 130 collections.

In January 2007, the Cornell NSDL project released NSDL 2.0, a new kind of digital library that integrates the best of traditional libraries, the powerful tools of Web 2.0, the organizational capabilities of the NSDL partners, and the collective intelligence of all those who use and contribute to the library. The creation of NSDL 2.0 draws on several unique strengths of CIS at Cornell: its long history of research in digital libraries, its development of the leading-edge Fedora repository system, and a tradition of creating production-strength research systems. In the year ahead, NSDL will continue to draw on those strengths, with the planned release of many additional tools to support contribution, collaboration, and context creation within the framework of the high-quality STEM education resources at the core of NSDL.

The main portal of the NSDL is available at <http://nsdl.org/>



Information Assurance Institute

The U.S. Air Force Research Laboratory (AFRL)/Cornell Information Assurance Institute (IAI) supports a broad spectrum of research and education efforts aimed at developing a science and technology base that can enhance information assurance and networked information-systems trustworthiness. IAI is also intended to foster closer collaborations among Cornell and AFRL researchers. Fred B. Schneider is the director.



AFRL researchers participate in Cornell research projects, facilitating technology transfer and exposing Cornell researchers to problems facing the Air Force; Cornell researchers become involved in AFRL projects and have access to unique AFRL facilities. The institute thus makes both Cornell and AFRL more attractive places to work, facilitating recruitment of higher-caliber personnel at each site.

Under the auspices of IAI, Cornell researchers are now involved in the development of the Department of Defense Global Information Grid. Various other technical collaborations are also being explored in the use of "gossip protocols," in language-based security policy-enforcement technology, in the design of trusted computing systems, and in data mining from networks of sensors.

For further information, see <http://www.cis.cornell.edu/iai/>

A New Start

INNOVATIVE INTRODUCTORY CS COURSES AT CORNELL

Cornell has a proud tradition of excelling both at research and at education; for instance, we have produced a number of foundational textbooks, and all our faculty, senior and junior, are actively involved in teaching and advising undergraduates.

In recent years, we have turned a great deal of our attention to the freshman experience. Students entering college bring with them an astounding diversity of interests, skills, goals, and perspectives. Hence, rather than stick with a one-size-fits-all introductory curriculum, Cornell has been taking the initiative in creating an exciting set of radically new courses—almost all taught by full professors—intended to reach out to freshmen, thus catching students early in their careers to excite them about the true depth and breadth of inquiry and applications that computer science represents.

Here, we present a brief overview of some of these new classes. Course webpages can be found at <http://www.cs.cornell.edu/Courses>.

CS 100 Honors—Introduction to Computer Programming

Taught by Graeme Bailey

CS 100 Honors, focused primarily on problem solving, is an example of the smaller, free-form learning environments we are developing to appeal to students from a range of computing and non-computing backgrounds.

Capping this introductory Java course at 35 students allows us to use semi-real, project-based problems to engage students, regardless of expertise. Students work in pairs to build a road-traffic simulation, a basic image manipulation toolkit, a simulation of a small economy, and a basic sound analysis toolkit. In a unit on images, students start by playing with matrices of binary or decimal data representing black-white or greyscale images and then devise schemes to double

the resolution (a project given in the latter part of the course). Since the point is for them to think, rather than learn state-of-the-art image analysis, they attacked this in a naive manner, soon learning that the black-white case was vastly more algorithmic and error sensitive. Yet, they were infected by the challenge and experienced the fun and frustration of discovery and experiment.

Students with no programming experience typically performed at the top of the class, suggesting that this course is an effective route to broadening participation in computing to students not historically engaged by traditional introductory programming classes. The course has extremely high retention rates, regardless of the students' programming experience, and has a female:male ratio of over 1:3.

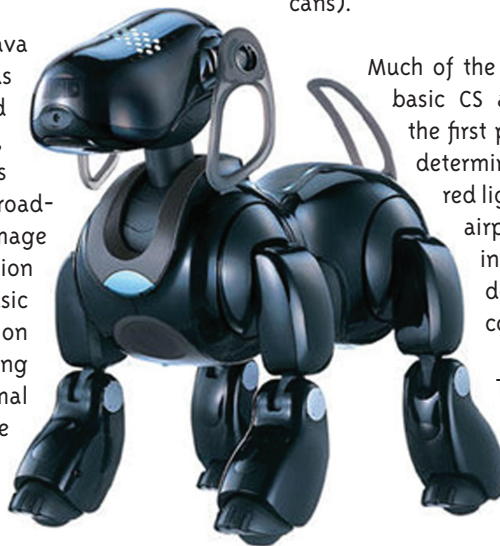
CS 100R—Introduction to Computing Using MATLAB and Robotics

Taught by Ramin Zabih

This new freshman course uses camera-controlled robots to teach fundamental concepts of computer science. The course, taught in Matlab, focuses on common issues that arise in the analysis of real-world data. In one project, the students use least-squares fitting to analyze odometry data from the robots in order to build a robot speedometer and accelerometer. In another project, the students design simple clustering algorithms to enable the robot to distinguish between primarily red objects (e.g., Coke cans) and primarily blue ones (e.g., Pepsi cans).

Much of the course involves the application of basic CS algorithms to analyze images. In the first portion of the course, the goal is to determine the position and orientation of a red light stick (like those used to guide an airplane into its gate). The algorithms include sorting, median finding, depth-first search, connected components, and convex hull.

The robots are a combination of Sony Aibos (robot dogs) and 4-wheeled robots designed for us by the University of Pennsylvania. At the end of the



The robotic dog, AIBO (Artificial Intelligent robot), developed by Sony and used in CS 100R.

course, the students complete a final project of their own choosing, drawing on the robot capabilities that they build up over the semester.



Chris Danis (Computer Science, BS '06, MEng '07), Ramin Zabih (CS faculty), and Brian Rogan (Computer Science, BS '06, MEng '07) developed CS 100R.

CS 101J—Transition to Object-Oriented Programming

Taught by David Gries

Recognizing the importance of computing in engineering, the forward-looking Engineering faculty have changed the computing requirement for engineering students from four to five credits. Students will take a four-credit course that focuses on either Java or Matlab and then a one-credit course in the other language. This summer, we are developing the one-credit Java course, which will focus on object orientation. The challenge is to make this course appealing to all students, and we are introducing new ideas and a novel course format to meet this challenge.

The course will be self-paced, four weeks, and pass/fail. There will be six-seven different alternatives for each programming assignment, so students can find one that piques their interest, but the course will require mastery of the material. Students will continue to take versions of each test until they get 85%.

To promote self-pacing, conventional lectures will be replaced by short, web-based, recorded lectures—not videos of someone lecturing but recordings of a computer monitor and audio as the lecturer speaks and manipulates the mouse and keyboard. Interaction with course personnel is not being sacrificed. Regular discussion sessions will allow students to ask questions on material they have already studied, and help from trained consultants will be available several hours a day.

“An exciting set of radically new courses.”

CS 165—Computing in the Arts

Taught by Graeme Bailey and Carol Krumhansl

Working with poetry, visual art, sculpture, and music, students learn about randomness and stochastic processes, symmetry, structure and group theory, dynamical systems and embedded structures, shape, deformation, and topology—applying all this to actual artistic creations. The initial focus is basic Java programming skills; the remainder of the course is spent learning computing concepts and how to exploit them in various media. Students work in groups of two or three, often across disciplinary divides, sharing their expertise and curiosity both in project development and in artistic critique. Students are also taught fundamental ideas from art (perspective) and music (fugue) to balance the mathematical and computer skills imparted.

Beyond the “technical” and purely artistic elements, the formal perception and cognition content benefited enormously from being co-taught last year by Carol Krumhansl (Department of Psychology), a noted world leader in music perception, and Graeme Bailey, the original developer of the course. Buy-in from departments across the university is reflected in multiple cross-listings, expanded research collaborations, and student projects beyond the Ithaca campus.

This is the core course for a Computing in the Arts minor/concentration, with formal tracks in music, psychology, computer science, film, dance, and art. Five other department-based tracks are under development in three Cornell colleges. Students have praised Computing in the Arts as “a course in creativity.”

CS 167—Visual Imaging in the Electronic Age

Taught by Don Greenberg

The world of computer graphics and imaging is exciting and fast-changing. Many of the ideas in this “concept” course have come from the early Renaissance painters such as Piero della Francesca and Masaccio and the writings of Alberti. Techniques were further refined by architects and stage set designers to provide illusions of three-dimensional spaces. Other concepts originated in the field of perception psychology and photography and are now used in the computer graphics and entertainment industries. But how we can portray and why we can understand 3D environments with 2D representations is still a complex and little understood problem.

An interdisciplinary course cross-listed in engineering, computer science, art, and architecture, CS 167 introduces students to the concepts of digital pictorial representation and display. Many of the assignments are artistic in nature, and no programming skills are required.

Topics include perspective representations, display technology, how television works, bandwidth concepts, digital photography, computer graphics modeling and rendering, matting and compositing, color perception, data acquisition, volumetric imaging, and historical precedents, primarily from the art world.

COM S 172—Computation, Information, and Intelligence

Taught by Lillian Lee

"Computation, Information, and Intelligence" is unique. It is a non-programming yet rigorous freshman-level introduction to computer and information science through the lens of artificial intelligence (AI). Topics include perception-based machine learning, random-graph models of Web evolution, link-based and vector-space-based information retrieval, statistical machine translation, statistical learning in human infants, and Loebner's restricted Turing Test.

Many students fresh out of high school find AI attractive, regardless of their technical background, and are intrigued to discover its applications to search engines and other tools that they actually use on a daily basis. By not teaching or requiring programming, the course can focus on developing algorithmic and modeling ideas at a relatively deep and formal level. For instance, students adapt a complete proof of convergence for a particular learning algorithm to a setting in which the initial assumptions have been altered. (It helps that freshmen still remember high-school geometry, so that they grasp vector mathematics quickly. Probability is also used, but an intuitive understanding suffices.)

As a result, the course effectively conveys the notion that computer science is not "just about programming," and the non-programming nature of the course appeals to a broad range of students. In spring of 2007, 34% of the 47 students were women, and enrollees included a psychology major and a French/English double major. Interestingly, during an internship visit to the offices of the New York Times, this double major challenged an editor's assertion that "on the Internet, all information is equal" with a disquisition on PageRank. This anecdote underscores the importance of "outreach" introductory courses: even those who do not end up pursuing technical disciplines can be empowered by understanding more about computer and information science.

CS 285—Networks

Taught by Jon Kleinberg and David Easley

The study of networks focuses on how the social, technological, and natural worlds are connected and how the structure of the connections affects these worlds. The topic thrives on insights that reach across very different domains. For example, models for the spread of epidemic diseases have been used to analyze the movement of fads and rumors through word-of-mouth communication; principles behind cascading failures in power grids have been applied to cascading extinctions that unravel food webs as well as economic crises that spread through the world-wide financial system; and the notion that certain individuals occupy powerful roles in social networks has helped form the basis for link analysis ranking methods in Web search engines.



Professors Jon Kleinberg and Michael Macy meet with students after class.

Motivated by the growing research interest in this topic, in spring 2007, David Easley and Jon Kleinberg co-taught this new introductory undergraduate course. Cross-listed in computer science, information science, economics, and sociology, the course covered mathematically rigorous applications of networks while relying on minimal prerequisites. Interest in the course came from many parts of campus—the 200+ enrolled students represented 25 different majors from six of Cornell's seven colleges.

Topics included basic properties of social networks; network models of market structure; the design of Web search engines and accompanying economic models; the spread of fads, information, and innovations through human populations; and applications to evolutionary biology, political science, and law. By drawing on current research and adapting it to the introductory undergraduate level, the course managed to combine themes that until recently had existed only in disparate parts of the academic literature.

RESEARCH FOCUS: Computer Science

Computer Graphics

Cornell is a leader in the field of computer graphics, a broad, interdisciplinary field that includes a wide and growing range of applications from science to communication to entertainment.

Research in computer graphics involves algorithms, physics, psychology, computation, computer vision, mathematics, and architecture, among other fields. At Cornell, research and teaching in computer graphics are centered in the computer science department and the closely affiliated Program of Computer Graphics (PCG), one of the world's leading computer graphics laboratories and a dominant force in the international computer graphics community for more than thirty years.

The interests of the computer graphics group span a broad range of areas including realistic materials; scalable high-quality rendering; animation; physical simulation for graphics; haptics and sound rendering; image-based modeling and texturing; and vision and perception.

Trustworthiness

Cornell faculty are leading researchers in computer security and trustworthy computing. Under the aegis of the Information Assurance Institute located within CIS and as a partner in the National Science Foundation TRUST Science and Technology Center, we have a track record of success in tackling fundamental and applied problems associated with ensuring the security and reliability of computing systems.

Current research projects range from system and network security to reliability and assurance, spanning language-based security, trusted computing, reliable and secure online services, static information flow control, and security policy specification and enforcement. Most of the work has

a substantial practical component, including the NEXUS secure operating system for PCs, an electronic voting system that implements strong (and proven) guarantees, Quicksilver live objects where scalable replication is integrated directly with the type system of the .net or jzee environment for use in supporting large data centers, the Fireflies Byzantine-fault-tolerant overlay network for media dissemination, and Night Watch a new probabilistic auditing tool that can monitor applications to detect failures and trigger appropriate corrective action. Some work, however, is primarily theoretical: our new definitions for measuring how much information can be leaked during a given program execution or our abstract characterization for the attacks that program obfuscation resists.

The breadth and depth of the projects undertaken at Cornell are a direct result of the well-integrated, diverse, and collegial environment that the Department of Computer Science provides. Our work draws its strength from the synergy between the groups working on security, programming languages, operating systems, logic, and formal methods.

For more information, see
<http://www.cs.cornell.edu/Research/Security/>



Bala's group is developing scalable algorithms for rendering realistic images of complex scenes; these algorithms eliminate the expensive simulation of light interactions that are not visually important to a viewer. This split-image shows a Roulette wheel (left) that appears blurry (right) because of its fast motion. This motion blurred image can be efficiently rendered by combining novel multiresolution algorithms with perceptual metrics.

Database Systems

The digitization of our daily lives has led to an explosion in the collection of data. Experts estimate that the amount of data doubles every three years. Cornell is a leader in addressing data management problems that require scalability and usability. Ongoing research includes Web 2.0 techniques for building, extending, and personalizing hosted data-driven web applications; techniques for data privacy that enable sharing and mining of data while limiting the disclosure of private information; and design and implementation of a system that supports real-time discovery of complex patterns in event streams and which scales to large numbers of pattern queries and high stream rates. We also have recently begun work on scalability and extensibility of computer games and virtual worlds, with the goal of creating "Games 2.0" environments that enable players to participate in the creation of the virtual world and still scale to millions of participants. We also are collaborating with physicists, mechanical engineers, astronomers, and ornithologists across campus on several eScience projects that concentrate on novel data management and analysis services for the sciences. For example, we are collaborating with the Cornell Lab of Ornithology on mining citizen science data.

For more information see
<http://www.cs.cornell.edu/database>

Artificial Intelligence

Artificial intelligence (AI) research at Cornell is both theoretical and experimental and covers a range of topics, including decision theory, information retrieval, knowledge representation, machine learning and data mining, natural language processing, planning, reasoning under uncertainty, search, and computer vision. Natural language processing, the ability to talk to machines in plain English, is an important research area in our Information Science program. Research in AI often involves collaboration with other disciplines such as biology, economics, linguistics, medicine, operations research, physics, and psychology. A particular strength at Cornell

is computer-intensive approaches to AI problems. Professors who work in AI are also key participants in the university-wide Cognitive Studies program and in CIS's Intelligent Information Systems Institute.

For more information, see
<http://www.cs.cornell.edu/Research/ai/> and
<http://www.cis.cornell.edu/iisi/>.

Systems

The systems group at Cornell is concerned with the design and implementation of the software systems that constitute the global computing infrastructure. Our interests span operating systems, networking, mobile computing, peer-to-peer systems, programming languages, and distributed computing. We build high-performance, reliable, and working systems based on a principled, scientific approach to practical, pragmatic problems. Past research by the group in fault-tolerance, peer-to-peer systems, distributed communication, extensible systems, and Internet networking is widely cited and used. Concrete software artifacts developed by Cornell researchers are used by the New York Stock Exchange and the French air traffic control system, form a part of all firewalls, and are deployed in commercial operating systems.



Cornell hosted the North-East Student Colloquium on Artificial Intelligence (NESCAI) again in April 2007. NESCAI seeks to foster discussion, provide networking opportunities, and a forum for feedback among graduate students from around the region. For more information, see <http://www.cs.cornell.edu/Conferences/nescail/>

Joseph Halpern (left), Nat Hyafil (Toronto), Pradeep Ravikumar (CMU), Eric Breck, Jure Leskovec (CMU), Lucian Leahu, Phoebe Sengers, Thorsten Joachims, Carla Gomes

Scholarly Communication and Digital Libraries

The explosion of the Web and the manner in which it digitizes our intellectual, political, and cultural activities, and in fact our everyday lives, has led to revolutionary changes. Not only have we come to expect a world where everything is searchable and accessible from our desktops, laptops, and palmtops, but we increasingly demand a level of customization and interaction with the information that we use daily. This new form of online information environments is exemplified by the so-called social networking websites, such as FaceBook, MySpace, and Wikipedia. Researchers in Digital Libraries, Web Information, and Scholarly Communication in CIS are collaborating with colleagues in theory, machine learning, sociology, and communication to build the infrastructure, systems, and tools to manage and analyze these new social information environments.

Research in information infrastructure in CIS has had a major impact on the protocols, standards, and middleware that support interoperation and persistence among digital information systems. Fedora, developed in CIS from NSF-funded research and now an international open source project based at Cornell,

is middleware that is the foundation of advanced Web, library, archive, and scholarly communication applications worldwide. The Open Archives Initiative develops widely-deployed protocols and standards for the interchange of information resources among heterogeneous systems in eScholarship and library applications. The Web Library is a joint project of Cornell University and the Internet Archive to provide data and computing facilities for research about the structure and evolution of the Web and the information on the Web.

Research on information systems in CIS is widely recognized for its impact on how scholars, teachers, and students use and share information. The arXiv, joint work between researchers in CIS and the Cornell Library, is the premier source for information among physicists and mathematicians, dramatically changing the communication culture in these communities. The NSF-funded National Science Digital Library (NSDL) is building an online learning environment that combines traditional library notions with so-called Web 2.0 social overlays.

Both of these application-oriented research areas provide the context for investigations into new methods to analyze and infer information and knowledge from these advanced information environments. CIS

researchers are using a mixture of data mining, bibliographic analysis, machine learning techniques, and social network analysis to explore phenomena such as the evolution of ideas, the nature of online collaboration, and the composition of compound information from basic information units distributed over the network.

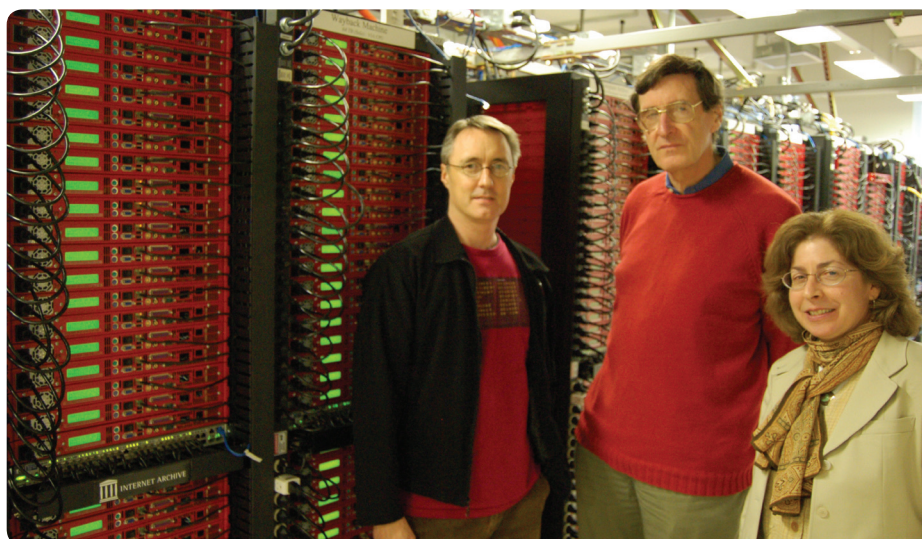


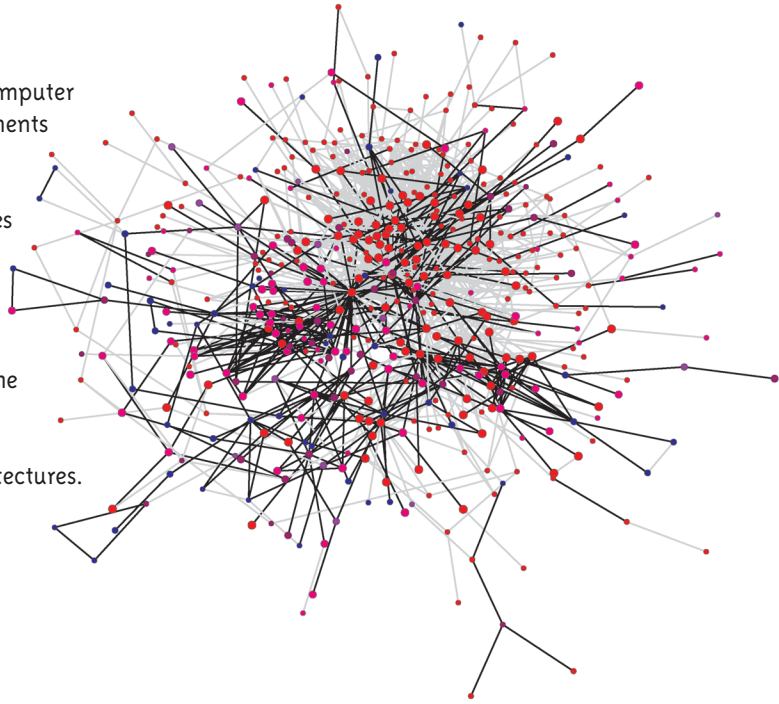
Photo by Felix Weigel.

Bill Arms, computer science faculty (middle), and Ruth Mitchell, Theory Center, tour the Internet Archive with John Berry.

Scientific Computing

Scientists and engineers increasingly rely on computer models and simulation to buttress their experiments and designs. From the prevention of metal fatigue in engines to the design of new medicines, scientific and technological advances rely on powerful computers and the algorithms that run on them. The scientific computing group develops the algorithms that underlie simulation and optimization. Matrix computations are a recurring theme in the research. The focus is on efficient and robust algorithms with an eye toward modern, high-performance parallel and multi-threaded architectures.

For more information, see
<http://www.cs.cornell.edu/Research/scientif/>



Theory of Computing

Cornell is a world leader in the theory of computing with fundamental research that spans the breadth of the field. Our faculty and students have contributed to such diverse areas as algorithms, combinatorial optimization, computational complexity, cryptography, learning theory, program logic and semantics, and automated deduction. Cornell is also at the forefront in applying theoretical ideas to problems in areas such as machine learning, data mining, software and hardware verification, reliable systems, security, networks, computer vision and medical imaging, computer graphics, programming languages and compilers, information science, and the computational sciences. Yet, while the theory of computing is now recognized as a deep and foundational scientific discipline with broad impact, many of its central questions remain unanswered. As we enter the information age, continued progress on the fundamental theoretical questions will be essential for heightened understanding of complex systems and software, the nation's network and information infrastructure, and the physical, biological, and social sciences.

For further information, see
<http://www.cs.cornell.edu/Research/theory/>

Programming Languages

Cornell has a long history of leadership in many programming languages, with strength in many topics, including logics and semantics of programming languages, language design and implementation, program analysis and verification, and language-based security. The programming languages area draws on Cornell's strengths in both the theoretical and practical sides of computer science.

Active research includes our work on Jif, a language for analyzing secure information flow and for automatically generating secure distributed systems, on theorem-proving and logic-based programming tools in the NuPRL system, on the J&E language, which supports the development of extensible systems such as the Polyglot compiler framework, and on program analysis techniques for the automatic detection of software errors and for memory management optimizations, in the Crystal and Frex systems.

Strong synergies exist with researchers in other areas. For instance, NuPRL-based theorem-proving technology has been used by the systems group to verify communication protocols, and the work on language-based security spans the areas of computer security, systems, and programming languages.

Computational Biology

New computational tools are being developed to analyze, understand, and manipulate newly available, richly-detailed information on the basic structures of life. Without advanced computing, biological databases with gigabytes of information would defy attempts to recognize patterns in them. Cornell has a university-wide initiative in the science of genomics, in which the Department of Computer Science and others are taking part. CIS oversees interdisciplinary teaching and research in many aspects of computational biology. The Department of Computer Science initiated a novel graduate program in computational biology and medicine (CBM) in collaboration with Rockefeller University and Sloan Kettering Research Institute. The Tri-institutional program in CBM, now part of the new field of computational biology, builds on the strengths in the medical sciences in the New York City campuses and on the strengths in the computer science at the Ithaca campus.

For more information, see
<http://www.triiprograms.org/cbm/>

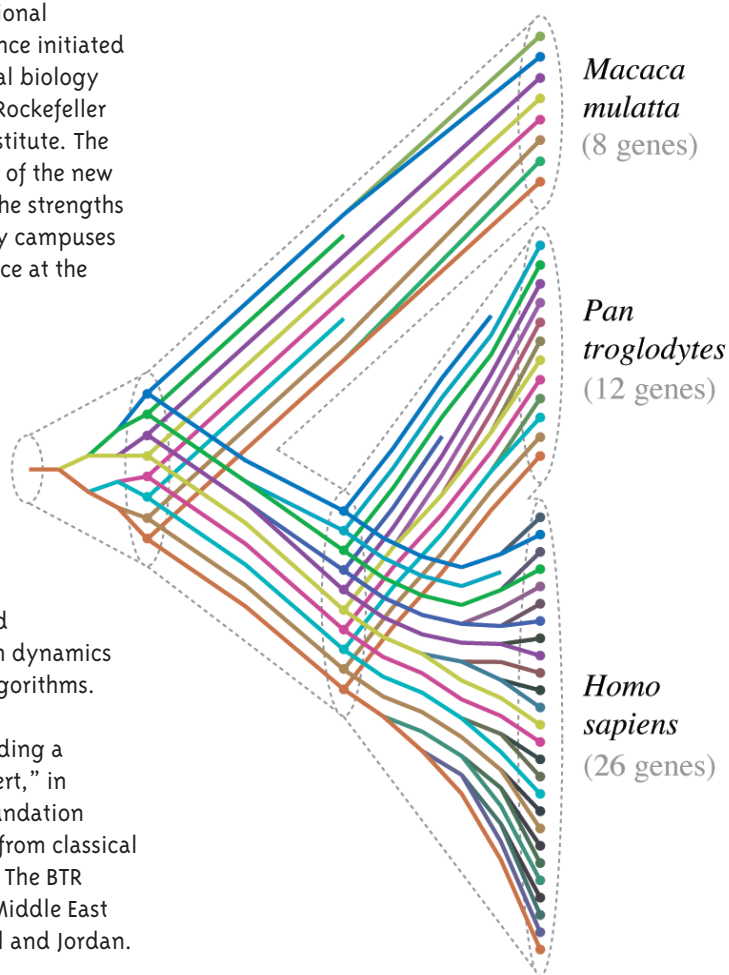
Within the Department of Computer Science, the focus is on computational molecular biology. This includes computational statistic and algorithmic problems that arise in biological sequence analysis, such as motif finding; protein fold prediction; long-time simulation of protein dynamics and function; and structure comparison algorithms.

The Department of Computer Science is leading a new project, "The Library of Life of the Desert," in collaboration with the Bridging the Rift Foundation (BTR) that will integrate (and create) data from classical and molecular biology of life in the desert. The BTR project also aims to promote peace in the Middle East and is located on the border between Israel and Jordan.

For more details see
<http://www.news.cornell.edu/features/BTR/>

The Computational Biology Service Unit, overseen academically by the Department of Computer Science, was established as a high performance center for bioinformatic applications by Microsoft, with annual funding of \$400K.

For more information on computational biology at Cornell, see
<http://www.cs.cornell.edu/Research/compbio/> and
<http://www.cis.cornell.edu/cb/>



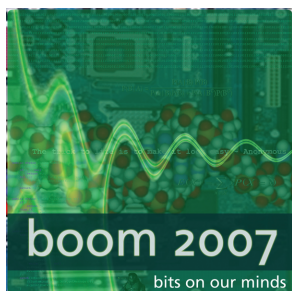
Researchers in computational biology at Cornell work to unravel complex evolutionary histories using the macaque genome. This image illustrates how PRAME genes (a gene family associated with cancer) expanded greatly in humans and chimpanzees between four and eight million years ago (Ma), but remained stable in macaques.

Image: Brian Raney, UC Santa Cruz; Tomas Vinar and Adam Siepel, Cornell University.

BOOM 2007

IS A HIT!

18



BOOM (Bits on our Minds) is the premiere showcase event for students across campus to show off their digital projects (<http://www.cis.cornell.edu/boom>). Not limited to computer science majors, BOOM increasingly dem-

onstrates the broad impact that computer technology has on intellectual and artistic pursuits across Cornell. Each year, 40 to 50 student projects, from a dozen different departments, jostle for four prestigious prizes. Two of the prizes are judged and awarded by the BOOM sponsors, which this year were Cisco and Credit Suisse. The winning projects were an amorphous robot and a tool for creating "sound sculptures." The "People's Choice" award is judged by the roughly 500 attendees themselves. This year's winner had an unfair advantage because it kept escaping from its booth and getting in people's way. The fourth award is judged by computer science faculty and focuses on the digital science behind the curtain. Called the "Where's the BOOM" award (where BOOM here means "Brains on our Machines"), it looks behind the glitzy GUIs and rewards those projects with a real intellectual core. This year's "Where's the BOOM" award went to a project that coupled RFID tags with the Internet to provide sound-based navigation for the visually impaired.

Each year, a faculty committee selects a student—or in this case, students—whose academic star is on the rise and acknowledges them in the BOOM Student Spotlight.



Joseph Nelson poses with his "Where's the BOOM" award. He won for his project, Subni Rfid, a web service for visually impaired users that gives physical objects, places, people, and things a presence on the internet. For more information, see <http://www.subni.com/>.

This year's Student Spotlight honors Chris Danis (Computer Science, BS '06, MEng '07) and Brian Rogan (Computer Science, BS '06, MEng '07).

Together with Professor Ramin Zabih, Chris and Brian developed the CS100R curriculum and technology platform, including the robot hardware, software, and lab infrastructure for the class. Chris has also served as a lead programmer on Cubesat and as the lead consultant at the ACCEL lab for four years. Brian is a founding member and group leader of CUAir. He also worked on projects related to energy minimization techniques. Among the many other projects they were involved with at Cornell, both were founding members and leaders, with Professor Dan Huttenlocher, of a project that enables mobile autonomous robots to map and classify the environment around them.



Spencer Topel (left) and faculty Kevin Ernste explore artistic renderings of timbres within a 3-dimensional space. For more information, see <http://digitalmusic.cornell.edu/>.

Walking across the showroom floor, it is impossible not to be infected by the students' enthusiasm for showing off the results of long hours of work. Part of the challenge for them is explaining what their projects do and how they work in the brief time they get with roaming attendees. The students particularly benefit from the high-quality exposure they get with recruiting companies. The sponsors get a unique opportunity to sell themselves to top students. Cornell also uses BOOM for community outreach by inviting the public. This year, an essay contest gave an iPod to the high-school or junior high-school student who could best describe their favorite BOOM project.

David Albonesi

IEEE Micro Top Picks in Microarchitecture (2007)

Kavita Bala

NSF CAREER Award (2007)

James and Mary Tien Excellence in Teaching Award,
College of Engineering (2006)

Paul Ginsparg

Paul Evans Peters Award from Educause, the Association
of Research Libraries, and the Center for Networked
Information (2006)

Carla Gomes

Fellow, Association for the Advancement of Artificial
Intelligence (2007)

Hans Bethe House Fellow, Cornell University (2007)
AAAI Best Paper Award (With B. Selman and A.
Sabharwal) (2006)

John Hopcroft

CRA Distinguished Service Award (2007)



Jon Kleinberg, professor of computer science, received the 2006 Rolf Nevanlinna Prize at the International Congress of Mathematicians in Madrid, Spain, for his "deep, creative and insightful contributions to the mathematical theory of the global information environment."

The prize has been awarded every four years since 1982 by the International Mathematics Union, in recognition of the most notable advances made in mathematics in the "information society". It is considered to be a computer-science oriented parallel to the prestigious Fields Medal, regarded as the equivalent of a Nobel Prize in mathematics.

Kleinberg was also among 203 new fellows elected to the American Academy of Arts and Sciences for 2007 in honor of their distinguished contributions to their professions.

Eva Tardos, professor of computer science and chair of the Department of Computer Science, was among 64 new members and nine foreign associates elected to the National Academy of Engineering (NAE). Election to the academy is among the highest professional distinctions accorded to engineering faculty members.

Tardos was chosen for "contributions to the design and analysis of efficient algorithms for network problems." Her research focuses on "optimization," in which a computer is asked to find the most efficient way to organize a large number of elements.

Doug James

Fellow, Alfred P. Sloan Research (2006–2008)

Thorsten Joachims

ACM SIGKDD Best Paper Award (2006)

Uri Keich

NSF CAREER Award (2007)

Jon Kleinberg

American Academy of Arts and Sciences (2007)
Rolf Nevanlinna Prize, International Mathematical
Union (2006)

Robert Kleinberg

NSF CAREER Award (2007)

Lillian Lee

Cornell Provost's Award Distinguished Scholarship
(2007)

Hod Lipson

DARPA MTO Young Faculty Award (2007)
NSF CAREER Award (2006)

Rajit Manohar

Fellow of the World Technology Network (2006)
Globus Indus Technovators Award (2006)

Jose Martinez

IEEE Micro's Top Picks in Computer Architecture (2007)
IBM Faculty Award (2006)
NSF CAREER Award (2006)

Sally McKee

IBM Faculty Award (2006–2007)

Fred Schneider

Elected to CRA Board of Directors (2007)
Invited to CCC Council (2007)

Phoebe Sengers

Cornell Society for the Humanities Faculty Fellowship
(2007–2008)

Adam Siepel

Microsoft New Faculty Fellowship (2007)
NSF CAREER Award (2007)

Eva Tardos

Awarded Jacob Gould Schurman Endowed Chair (2007)
Elected to the National Academy of Engineering (2007)
Dantzig Prize (2006)

Getting CONNECTED

20

The Institute for the Social Sciences at Cornell was established in 2004 to promote cross-disciplinary social science research; as part of its efforts in this direction, it funds a campus-wide "theme project" each year. The theme project for the 2006–2007 academic year was focused on "Social Science in the Age of Networks," with a core team of faculty consisting of John Abowd, Larry Blume, David Easley, Geri Gay, Dan Huttenlocher, Jon Kleinberg, Michael Macy (project leader), Kathleen O'Connor, Jeffrey Prince, and David Strang. The intellectual contributions of CIS to the project were clear both in the subject matter and in the fact that eight of the ten core team members have affiliations with CIS.

The project team pursued a range of research directions, extending from theoretical work on network structure to empirical studies of on-line social networks including university e-mail communication and socially constructed knowledge through sites like Wikipedia. These projects reinforced the

interdisciplinary nature of networks research, as they involved collaboration among faculty specializing in computing and information science, economics, sociology, and social psychology, among other disciplines.

The project team also organized a series of events to engage the broader campus community in the topic of networks. These included a weekly seminar series featuring talks by Cornell faculty and external visitors; a set of courses, including an introductory undergraduate course on networks that drew over 200 students; and two conferences that brought well-known researchers from around the country to campus to discuss dynamic processes on networks and the structure of on-line communities.

"A powerful way to catalyze research..."

Overall, the project served as a powerful way to catalyze research and educational initiative related to the study of networks, with many of the new activities spanning departmental boundaries and likely to continue for many years to come.



"Social Science in the Age of Networks" team members. Back row (Left to Right): David Easley, Jeffrey Prince, John Abowd, Dan Huttenlocher, Larry Blume, Jon Kleinberg, Michael Macy, Geri Gay, and Mary Still. Front row sitting (Left to Right): Kathleen O'Connor, H. Ted. Welser, and David Strang.

CORPORATE AND FOUNDATION INTERACTIONS

Computing and Information Science (CIS) provides excellent recruiting opportunities for students skilled in core computer science and information science and in the intersections of these and other disciplines, such as the physical sciences, the life sciences, social sciences, digital arts, and engineering.

Each spring the student research fair, "Bits On Our Minds" (or BOOM), provides a showcase of student technology projects. This showcase is funded by corporations, and in 2006 Cisco Systems, Credit Suisse, and FAST Research provided \$30,000 to support the fair. Representatives from other companies commonly attend in order to learn about developing technologies and to talk to the students who have developed these projects.

A new facility is being developed to house CIS and related disciplines at Cornell. This \$50+ million building will enable major programmatic and research advances that are currently in critical need of space. A lead gift by the Bill and Melinda Gates Foundation of \$25 million has helped this effort enormously. Additional funding is being sought to complete Gates Hall and to establish at least three new named professorships in CIS.

Microsoft has awarded \$35,000 to Johannes Gehrke to fund work on data privacy for medical data. Filip Radlinski is supported on a PhD Fellowship. Support for Paul

Ginsparg continues through the Technical Computing Initiative (TCI) to investigate and implement new methods for scholarly research communication. The project builds on Ginsparg's arXiv.org, which holds more than 340,000 open-access articles in physics, mathematics, computer science, and related fields.

Microsoft expanded its support for research in systems this year with total awards of \$325,000 to facilitate the CS department's efforts in operating systems, languages, and security. The funds will help maintain the cohesive and collegial atmosphere within the systems group and will also facilitate greater interactions between Cornell's systems researchers and the exploratory development groups at Microsoft (providing an unprecedented opportunity for Cornell research ideas to have a broad impact).

Cisco contributed \$92,000 for Paul Francis's research on Internet Protocol (IP) anycast deployment and peer-to-peer models for establishing connections through firewalls over the Internet.

Google Research is supporting Rich Caruana's work on model compression.

Intel provided \$3,150 to support Ramin Zabih's new course using camera-controlled robots to teach fundamental concepts of computer science. PhD



Representatives from Cisco attended BOOM 2007. Cisco was a sponsor of BOOM again this year, as well as a supporter of Paul Francis's research on Internet Protocol anycast deployment.

student Kevin Walsh, whose advisor is Gun Sirer, was awarded a \$54,000 PhD fellowship.

Morgan Stanley sponsored the CS undergraduate team's participation in the ACM programming competition.

Xerox Foundation continued its funding of Claire Cardie's research on machine learning and natural language processing and Johannes Gehrke's work on trend and sequence mining with awards of \$20,000.

Yahoo awards were given to several faculty in Computing and Information Science. Johannes Gehrke received \$75,000 in support of his research on anonymizing search data. A \$100,000 award was provided to Jon Kleinberg, Dan Huttenlocher, Lillian Lee, and Eva Tardos to support their work on the diffusion and temporal dynamics in web information. \$95,000 was awarded to Thorsten Joachims to support his research on large margin methods for predicting structured outputs.

Emerging partnerships are also developing with game companies and large animation studios. Pixar funded Doug James's computer graphics research. Electronic Arts, the world's largest games company, is one of several companies that has chosen Cornell as a core recruiting school. EA has continued to provide expertise and funding for critical programs like the Game Design Initiative at Cornell.

Please address questions relating to corporate interactions with CIS to:

Justine Schaffner

CIS Director of Development & Corporate Relations
js536@cornell.edu
607.227.8721

Aimee Potts

Asst to the CIS Director of Development & Corporate Relations
ab366@cornell.edu
607.254.8327

corprel@cs.cornell.edu
Fax: 607.254.8888
<http://www.cis.cornell.edu/>

Among the various gifts from corporations this year were totals of \$415,000 from Microsoft; \$270,000 from Yahoo; \$105,000 from Cisco; \$88,000 from Intel; \$65,000 from Google; and \$40,000 from Xerox.

Microsoft®

YAHOO!

CISCO™

intel®

Google™

XEROX®

A Message from the Director

OF ALUMNI AFFAIRS AND CORPORATE RELATIONS

JUSTINE SCHAFFNER

The greatest aspect about living in the midst of a digital revolution is the knowledge that change produces tremendous opportunities for engagement and the exchange of ideas. Rapid advancements in technology, campus research, and business models are all making this possible. Through increased channels of communication from the CIS faculty and Alumni Development, you will no longer be last to hear about the exciting ideas being generated within the halls of Cornell University.

In the Spring of 2007, CIS gave me the opportunity to serve the incredibly bright and creative alumni population that CIS has graduated over the years. By leveraging my past three years as the Director of the Cornell Entrepreneur Network (CEN) and prior ten years in the technology industry, I plan on developing a strong business network of technology-orientated alumni to facilitate greater opportunities for engagement and the exchange of ideas.

Computing and Information Science is one of the best kept secrets at Cornell and it is my responsibility to strengthen our presence and reputation, both on and off campus. In the coming years, you will have the opportunity to develop stronger relationships with

other alumni and gain insights into the numerous advancements and commercial ventures that are taking place at the impressive organizations where our alumni work.

It is the goal of this office to supplement that alumni contact with communications about the latest advancements and research occurring on campus. To aid that exchange, CIS faculty and alumni will host events in major metropolitan areas and tech centers across the country. We also plan on evangelizing the accomplishments of CIS via podcasts, webinars, and targeted distribution lists to highlight key initiatives at Upson Hall as well as feature notable alumni, faculty and technologies. Through the Alumni Development Office, we will "eat our own dog food" and utilize technology that is created and shaped on campus and in your organizations to facilitate an exchange of communication and ideas.

This is truly one of the most exciting times to be actively engaged with alumni, faculty and friends of Computing and Information Sciences at Cornell University. I look forward to meeting and serving you in the years to come.

THE GERARD SALTON LECTURE SERIES

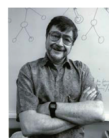
2006–2007



SEPTEMBER 21, 2006

HECTOR GARCIA-MOLINA
STANFORD UNIVERSITY

"GENERIC ENTITY RESOLUTION"



NOVEMBER 14, 2006

JUDEA PEARL
UCLA

"THE MATHEMATICS OF CAUSAL INFERENCE"

BROUGHT TO YOU WITH THE SUPPORT OF AMIT SINGHAL, PHD '97

PRESENTED BY THE DEPARTMENT OF COMPUTER SCIENCE
CORNELL UNIVERSITY
4130 UPSON HALL, ITHACA, NY 14853
WWW.CS.CORNELL.EDU/EVENTS/SALTONSERIES

Distinguished Career Lecture Series

In 2006, Deborah Estrin visited Cornell to share a "personal perspective on research collaborations and transitions" as part of the CS Distinguished Career Lecture Series. A pioneer in the development of embedded networked sensing (ENS), Estrin has wide experience to draw on with her research ranging from the design of network and routing protocols for very large, global, networks, to applications to environmental monitoring for ENS systems.

Deborah Estrin is a Professor of Computer Science with a joint appointment in Electrical Engineering at UCLA, holds the Jon Postel Chair in Computer Networks, and is Founding Director of the NSF-funded Center for Embedded Networked Sensing (CENS). Estrin received her Ph.D. (1985) in Computer Science from the Massachusetts Institute of Technology, and her B.S. (1980) from U.C. Berkeley. Before joining UCLA she was a member of the University of Southern California, Computer Science Department.

Games in Computer Science

24

The undeniable popularity of games draws the attention of academia, industry, and even the government on areas of design, development, and social impact. The game industry, like the film industry, is an unmistakable force in entertainment. And like filmmaking, game design can only thrive and evolve with the support of a strong academic foundation. In Fall 2003, the Game Design Initiative at Cornell (GDIAC) was formed by a group of students, faculty, and alumni who wished to study game design in an academic setting.

GDIAC supports undergraduate courses in game design and development, game-oriented graduate student projects, and game-related faculty research. All undergraduate students from colleges across Cornell are eligible to complete the Minor in Game Design. The game design courses culminate with students presenting their games, discussing their research, and explaining their shared design visions in the semi-annual Game Design Showcase.

Since fall 2003, GDIAC has offered Ithaca-area middle- and high-school students the opportunity to participate in the computer science department's game design courses. As "interns" these K-12 students work closely with undergraduate students and faculty as team members integrated into the design process. In May 2007, NSF formalized this outreach program with three years of funding to refine and enhance a game design curriculum specifically for 7th-12th graders, and to develop a replicable undergraduate mentor training program. This outreach program will serve to broaden participation in computing by testing game design as a gateway mechanism for study and career interest in computer science. Collaborating organizations include: Ithaca Youth Bureau, Tompkins County Action, Learning Web, and the Finger Lakes Unschooler's Network.

In addition to the educational opportunities, games also provide

many interesting research opportunities. As game design involves both software and artistic content, modern games are designed in a data-driven fashion. This means that game-specific content, such as artwork, music, or character behavior, is kept in data files that are separate from the game software. Many games now permit the game players to modify these external files (so-called "modding"), blurring the boundary between game designer and game developer and enabling the game players to express their creativity through the design of their own part of the game.

“...blurring
the boundary
between game
designer
and game
developer.”

The Cornell database group is developing ways for players to better express their creativity in modding games. A team led by Walker White, Alan Demers, and Johannes Gehrke have recently received funding from the NSF, the Air Force, and Microsoft in order to research scalability and creativity in computer games. The initial result of their work is a novel approach for scaling the behavior of non-player characters for data-driven games using techniques from database systems. In a nutshell, the research uses the same techniques that allow databases and search engines to scale to petabytes to enable an order of magnitude more non-player characters in games compared to existing techniques. These first results were recently published in the paper "Scaling Games to Epic Proportions" that appeared in the ACM SIGMOD 2007 Conference. Ongoing research focuses on the design and implementation of a full-fledged game engine.



Dean Cirielli (BA '88, MENG '91), center, tours the Spring 2007 Game Design Showcase.

2006-2007 CIS COURSES

CIS 121 INTRODUCTION TO MATLAB
CIS 122 FORTRAN APPLICATIONS IN EARTH SCIENCE
CIS 165 COMPUTING IN THE ARTS
CIS 167 VISUAL IMAGING IN THE ELECTRONIC AGE
CIS 300 INTRODUCTION TO COMPUTER GAME DESIGN
CIS 372 SPTP: STUDIO IN SPACE AND TIME
CIS 400 ADVANCED PROJECTS IN GAME DESIGN
CIS 490 INDEPENDENT READING AND RESEARCH
CIS 504 APPLIED SYSTEMS ENGINEERING I
CIS 505 SYSTEMS ARCHITECTURE, BEHAVIOR, AND OPTIMIZATION
CIS 576 DECISION THEORY I
CIS 577 DECISION THEORY II
CIS 629 COMPUTATIONAL METHODS FOR NONLINEAR SYSTEMS
CIS 750 EVOLUTIONARY COMPUTATION AND DESIGN AUTOMATION
CIS 797 TOPICS IN CIS/IGERT SEMINARS

2006-07 CS UGRAD COURSES

COM S 100J INTRODUCTION TO COMPUTER PROGRAMMING—JAVA
COM S 100M INTRODUCTION TO COMPUTER PROGRAMMING—MATLAB
COM S 100R INTRODUCTION TO COMPUTER PROGRAMMING—ROBOTS
COM S 101 INTRODUCTION TO COGNITIVE SCIENCE
COM S 113 INTRODUCTION TO C
COM S 114 UNIX TOOLS
COM S 130 INTRODUCTORY DESIGN AND PROGRAMMING FOR THE WEB
COM S 165 COMPUTING IN THE ARTS
COM S 167 VISUAL IMAGING IN THE ELECTRONIC AGE
COM S 172 COMPUTATION, INFORMATION, AND INTELLIGENCE
COM S 211 COMPUTERS AND PROGRAMMING
COM S 212 JAVA PRACTICUM
COM S 214 ADVANCED UNIX
COM S 215 INTRODUCTION TO C#
COM S 230 INTERMEDIATE DESIGN AND PROGRAMMING FOR WEB
COM S 280 DISCRETE STRUCTURES
COM S 285 NETWORKS
COM S 305 CREATIVE PROBLEM SOLVING
COM S 312 DATA STRUCTURES AND FUNCTIONAL PROGRAMMING
COM S 314 COMPUTER ORGANIZATION
COM S 316 SYSTEMS PROGRAMMING

COM S 322 INTRODUCTION TO SCIENTIFIC COMPUTATION
COM S 330 DATA-DRIVEN WEB APPLICATIONS
COM S 381 INTRODUCTION TO THEORY OF COMPUTING
COM S 411 PROGRAMMING LANGUAGES AND LOGICS
COM S 412 INTRODUCTION TO COMPILERS
COM S 413 PRACTICUM IN COMPILERS
COM S 414 OPERATING SYSTEMS
COM S 415 PRACTICUM IN OPERATING SYSTEMS
COM S 416 COMPUTER ARCHITECTURE
COM S 419 COMPUTER NETWORKS
COM S 421 NUMERICAL ANALYSIS
COM S 426 INTRODUCTION TO BIOINFORMATICS
COM S 428 INTRODUCTION TO COMPUTATIONAL BIOPHYSICS
COM S 430 INFORMATION RETRIEVAL
COM S 431 ARCHITECTURE OF WEB INFORMATION SYSTEMS
COM S 432 INTRODUCTION TO DATABASE SYSTEMS
COM S 433 PRACTICUM IN DATABASE SYSTEMS
COM S 465 COMPUTER GRAPHICS I
COM S 467 COMPUTER GRAPHICS II
COM S 468 PRACTICUM IN COMPUTER GRAPHICS
COM S 472 FOUNDATIONS OF ARTIFICIAL INTELLIGENCE
COM S 473 PRACTICUM IN AI
COM S 474 INTRODUCTION TO NATURAL LANGUAGE PROCESSING
COM S 478 MACHINE LEARNING
COM S 480 INTRODUCTION TO CRYPTOLOGY
COM S 482 INTRODUCTION TO ANALYSIS OF ALGORITHMS
COM S 485 MATHEMATICAL FOUNDATIONS FOR THE INFORMATION AGE
COM S 490 INDEPENDENT READING AND RESEARCH

2006-07 CS GRAD COURSES

COM S 501 SOFTWARE ENGINEERING
COM S 513 SYSTEMS SECURITY
COM S 514 INTERMEDIATE COMPUTER SYSTEMS
COM S 516 PARALLEL COMPUTER ARCHITECTURE
COM S 519 ADVANCED COMPUTER NETWORKS
COM S 530 THE ARCHITECTURE OF LARGE-SCALE INFORMATION SYSTEMS
COM S 567 PHYSICALLY BASED ANIMATION
COM S 578 EMPIRICAL METHODS IN MACHINE LEARNING AND DATA MINING

COM S 611 ADVANCED PROGRAMMING LANGUAGES
 COM S 614 ADVANCED SYSTEMS
 COM S 615 PEER-TO-PEER SYSTEMS
 COM S 619 RESEARCH IN COMPUTER NETWORKS
 COM S 621 MATRIX COMPUTATIONS
 COM S 626 COMPUTATIONAL MOLECULAR BIOLOGY
 COM S 632 DATABASE SYSTEMS
 COM S 665 ADVANCED INTERACTIVE GRAPHICS
 COM S 672 ADVANCED ARTIFICIAL INTELLIGENCE
 COM S 674 NATURAL LANGUAGE PROCESSING
 COM S 678 ADVANCED TOPICS IN MACHINE LEARNING
 COM S 681 ANALYSIS OF ALGORITHMS
 COM S 683 ADVANCED DESIGN AND ANALYSIS OF ALGORITHMS
 COM S 687 CRYPTOGRAPHY
 COM S 709 COMPUTER SCIENCE COLLOQUIUM
 COM S 715 SEMINAR ON PRL
 COM S 718 SEMINAR IN COMPUTER GRAPHICS
 COM S 719 SEMINAR IN PROGRAMMING LANGUAGES AND COMPILERS
 COM S 726 PROBLEMS AND PERSPECTIVES IN COMPUTATIONAL MOLECULAR BIOLOGY
 COM S 733 SEMINAR IN DATABASE SYSTEMS
 COM S 750 EVOLUTIONARY COMPUTATION AND DESIGN AUTOMATION
 COM S 754 SYSTEMS RESEARCH SEMINAR
 COM S 764 VISUAL OBJECT RECOGNITION
 COM S 772 SEMINAR IN ARTIFICIAL INTELLIGENCE
 COM S 775 SEMINAR IN NATURAL LANGUAGE UNDERSTANDING
 COM S 778 TOPICS IN MACHINE LEARNING
 COM S 785 SEMINAR ON INFORMATION NETWORKS
 COM S 787 TOPICS IN CRYPTOGRAPHY
 COM S 789 SEMINAR IN THEORY OF ALGORITHMS AND COMPUTING
 COM S 790 INDEPENDENT RESEARCH
 COM S 990 DOCTORAL RESEARCH

2006-07 INFO COURSES

INFO 130 INTRODUCTION TO DESIGN AND PROGRAMMING FOR THE WEB
 INFO 172 COMPUTATION, INFORMATION AND INTELLIGENCE
 INFO 204 NETWORKS
 INFO 214 COGNITIVE PSYCHOLOGY
 INFO 230 INTERMEDIATE DESIGN AND PROGRAMMING FOR THE WEB
 INFO 245 PSYCHOLOGY OF SOCIAL COMPUTING
 INFO 292 INVENTING AN INFORMATION SOCIETY
 INFO 295 MATHEMATICAL METHODS FOR INFORMATION SCIENCE
 INFO 330 DATA-DRIVEN WEB APPLICATIONS
 INFO 345 HUMAN-COMPUTER INTERACTION
 INFO 349 MEDIA TECHNOLOGIES
 INFO 355 COMPUTERS: FROM THE 17TH CENTURY TO THE TO DOT COM BOOM
 INFO 372 EXPLORATIONS IN AI
 INFO 430 INFORMATION RETRIEVAL
 INFO 431 ARCHITECTURE OF WEB INFORMATION SYSTEMS
 INFO 435 SEMINAR ON APPLICATIONS OF INFORMATION SCIENCE
 INFO 440 ADVANCED HUMAN-COMPUTER INTERACTION DESIGN
 INFO 445 SEMINAR IN COMPUTER-MEDIATED COMMUNICATION
 INFO 447 SOCIAL AND ECONOMIC DATA
 INFO 450 LANGUAGE AND TECHNOLOGY
 INFO 490 INDEPENDENT READING AND RESEARCH
 INFO 491 TEACHING IN INFORMATION SCIENCE, SYSTEMS, AND TECHNOLOGY
 INFO 515 CULTURE, LAW, AND POLITICS OF THE INTERNET
 INFO 530 THE ARCHITECTURE OF LARGE-SCALE INFORMATION SYSTEMS
 INFO 614 COGNITIVE PSYCHOLOGY
 INFO 631 ARCHITECTURE OF WEB INFORMATION SYSTEMS
 INFO 634 INFORMATION TECHNOLOGY IN SOCIOCULTURAL CONTEXT
 INFO 635 SEMINAR ON APPLICATIONS OF INFORMATION SCIENCE
 INFO 640 ADVANCED HUMAN-COMPUTER INTERACTION DESIGN
 INFO 645 COMPUTER-MEDIATED COMMUNICATION GRADUATE SEMINAR
 INFO 648 SPEECH SYNTHESIS
 INFO 651 CRITICAL TECHNICAL PRACTICES
 INFO 709 INFORMATION SCIENCE COLLOQUIUM
 INFO 747 SOCIAL AND ECONOMIC DATA
 INFO 790 INDEPENDENT RESEARCH
 INFO 990 THESIS RESEARCH

2006–2007 RESEARCH GRANTS

Funded Research—Computing & Information Science / Computer Science

INVESTIGATOR	SPONSOR	AWARD	PERIOD OF AWARD	TITLE
Arms	NSF	120,000	10/1/06–9/30/07	Using the Cyberinfrastructure to build a Full Text Index to the Web
Bala	NSF	450,000	2/1/07–1/31/12	CAREER: Scalable Rendering for Visual Realism in Scale-Complex Scenes
Bala	NSF	75,000	8/15/05–7/31/06	Feature-based Rendering
Bala	NSF sub of U TX	535,880	1/1/07–7/31/09	CSR–AES: Parallelizing Irregular Applications for Multicore Processor
Bala	Intel	90,000	7/1/06–6/30/09	Efficiently Parallelizing Graphics Applications for Multicore Processors
Bala	NIST	7,350	5/1/07–9/1/07	NIST Summer Undergraduate Research Fellowship Program
Birman	Intel	180,000	7/1/05–6/30/08	Using QuickSilver to Support Autonomic Services in Clusters
Birman/Gehrke/ Demers	AFOSR	4,000,000	5/1/02–4/30/07	Scalable Technology for a New Generation of Collaboration Applications
Birman	DARPA sub of Telcordia	306,734	10/1/05–9/30/06	Adaptive Cognition-Enhanced Radio Teams (ACERT)
Birman	AF sub of UC Berkeley	1,200,000	4/1/06–1/31/09	Partnership for Research Excellence and Transition (PRET) in Software and Security
Birman	AFRL	1,100,000	2/28/06–8/31/08	Widely Distributed Scalable Infospaces: CASTOR
Cardie	NSF	500,000	9/15/02–8/31/06	Reducing the Corpus Annotation Bottleneck for Natural Language Learning
Cardie	ARDA sub of U Pittsburgh	463,368	5/1/04–9/30/06	Opinions in Question Answering
Cardie	DHS sub of U Pitt	842,145	7/1/06–6/30/09	Information Extraction of Events and Beliefs from Text
Cardie	Xerox	20,000	7/1/06	Machine Learning in Natural Language Processing
Caruana	NSF	507,000	6/15/04–5/31/09	CAREER: Meta Clustering: Improving User Efficiency on Real-World Clustering Applications
Caruana	NSF	14,000	6/1/06–5/31/07	Student Poster Program and Travel Scholarships for the 23rd International Conference on Machine Learning (ICML 2006)
Caruana/Joachims	NSF	276,000	8/15/04–7/31/07	Optimizing Classification Models to Application-Specific Performance Metrics
Caruana	Google	65,000	12/1/06	Model Compression
Constable	NSF	180,000	8/1/06–7/31/08	Developing a Theory of Events to Improve Distributed Systems
Demers/ Shanmugasundaram	NSF	500,000	2/15/06–1/31/09	Hilda: A High-Level Language for Data-Driven Web Applications
Elber	NIH	1,339,390	3/1/04–2/29/08	Long Time Dynamics of Biomolecules
Elber	NIH sub of U Pitt	58,305	1/1/05–12/31/07	Computational Prediction of Biomolecular Dynamics
Elber	HFSP	44,500	4/10/07–4/10/10	Biophysical and Bioinformatic Methods for Protein Structure Refinements
Francis	NSF	496,421	9/1/03–8/31/07	Very Fine-grained Proximity Addressing
Francis	ARDA sub of Telcordia	190,166	10/1/05–5/30/07	RapidTrace: Rapid Traceback of Cyber Attacks
Francis	NSF	300,000	9/1/06–8/31/09	NeTS-FIND: Collaborative Research: Towards Complexity Oblivious Network Management
Francis	Cisco	92,246	10/20/06	End-Middle-End Internet Connection Establishment

INVESTIGATOR	SPONSOR	AWARD	PERIOD OF AWARD	TITLE
Gehrke	NSF	340,000	9/1/02–8/31/07	CAREER: Towards Sensor Database Systems
Gehrke	NSF	500,000	9/15/03–8/31/06	SENSORS: Data-Driven Sensor Networks
Gehrke	Sloan	40,000	9/16/03–9/15/07	Sloan Research Fellowship
Gehrke	AFOSR	600,000	3/1/06–11/30/08	Stateful Publish-Subscribe for XML Data Streams
Gehrke	Microsoft	200,000	6/1/07–5/31/08	Scalability in Computer Games and Simulations
Gehrke/Birman	I3P	300,000	4/1/07–3/31/09	Distributed Monitoring for Human Behavior, Insider Threat and Awareness
Gehrke/Demers	AFOSR	608,333	5/1/07–11/30/09	Scalable Simulations of Dynamics of Relationships
Gehrke/Demers	KD-D thru NSF	450,000	7/1/06–9/30/07	Mining and Monitoring the Blogosphere
Gehrke/Sirer/ Shanmugasun- daram/Demers/ Birman	NSF	913,320	9/1/02–8/31/06	ITR: Massively Convergent Distributed Computing
Gehrke/White	NSF	200,000	7/1/06–6/30/08	A Formal Approach to Data Stream Processing
Gehrke	Microsoft	35,000	11/6/06	Data Privacy for Medical Data
Gehrke	Xerox	20,000	5/10/07	Trend and Sequence Mining
Gehrke	Yahoo!	75,000	12/4/06	Anonymizing Search Data
Ginsparg	NSF	796,395	10/1/04–9/30/07	Classification, Analysis and Navigation Tools for Physics Research Communication
Ginsparg	Microsoft	1,259,869	7/1/06–6/30/09	TCI Proposal for Scholarly Communication Collaboration Between Cornell University and Microsoft Corporation
Gomes	AFRL	3,870,000	3/1/04–2/28/08	Intelligent Information Systems Institute
Gomes	NYSTAR CEIS	30,000	7/1/07–6/30/08	Computational Intelligence for Print Shop Workflows
Gomes	Eastman Kodak	170,000	7/1/07–6/30/08	Computational Intelligence for Print Shop Workflows
Halpern	ONR	526,058	5/1/01–12/31/06	Software Quality and Infrastructure Protection for Diffuse Computing
Halpern	ONR sub of U Penn	189,000	6/1/04–12/31/06	Trustworthy Infrastructure, Mechanisms, and Experimentation for Diffuse Computing
Halpern	AFOSR	429,816	1/1/05–12/31/07	Reasoning About Authorization and Security
Halpern	ONR sub of ITT	300,000	10/10/05–12/29/06	Designing a Policy Language
Halpern	NSF	394,733	10/15/05–9/30/08	Taking Awareness, Language, and Novelty into Account in Decision-Making and Game Theory
Halpern	NSF	8,000	4/1/07–3/31/08	2nd Northeast Student Colloquium on Artificial Intelligence
Hopcroft	AFOSR	374,974	2/1/07–11/30/09	Information Forensics
Hopcroft/Selman	NSF	300,000	8/1/03–7/31/06	ITR: Emerging Communities in Large Linked Networks: Theory Meets Practice
Hopcroft/Selman	NSF	475,000	4/1/06–3/31/09	The Analysis and Modeling of Large Linked Networks
Huttenlocher	NSF	100,000	5/1/06–10/31/07	SGER: Recognizing Objects by Simultaneously Combining Appearance and Geometry
James	NSF	406,000	2/15/04–1/31/09	CAREER: Precomputing Data-Driven Deformable Systems for Multimodal Interactive Simulation
James	NSF sub of UMD	119,000	9/1/06–8/31/09	Reality-based Data-driven Computer Models for Surgical Simulation
James	NIH sub of UMD	375,000	9/1/06–8/31/10	Data-Driven Real-Time Surgical Simulation from Reality-Based Soft-Tissue Models
James	Boeing Phantom Works	75,000	12/18/06–12/31/07	Interactive Deformable Object Simulation for Aerospace Applications
James	Sloan	45,000	9/16/06–9/15/08	Sloan Research Fellowship
James	Pixar	27,776	8/17/06	
Joachims	NSF	400,000	9/1/03–8/31/08	CAREER: Improving Information Access by Learning from User Interactions

INVESTIGATOR	SPONSOR	AWARD	PERIOD OF AWARD	TITLE
Joachims	BSF	22,500	10/1/05–9/30/07	Efficient Non-Parametric Revelation of Ordinal Preferences
Joachims/Caruana	NSF	270,000	8/15/04–7/31/07	Discriminative Methods for Learning with Dependent Outputs
Joachims	Yahoo!	95,000	10/1/06	Large Margin Methods for Predicting Structured Outputs
Keich	NSF	644,880	7/1/07–6/30/12	CAREER: Improving the Performance of Motif Finding Tools
Kleinberg, J/Lee/ Huttenlocher/Tardos	Yahoo!	100,000	10/1/06	Diffusion and Temporal Dynamics in Web Information
Kleinberg, R.	NSF	400,000	9/1/07–8/31/12	CAREER: Algorithms for Environments with Incomplete Information
Kozen	AF Sub of ATC-NY	30,000	8/15/06–5/1/07	AppMon: Application Monitors for Not-Yet-Trusted Software
Kozen	NSF	250,000	10/15/06–9/30/09	Specialized Logics for Applications in Computer Science
Krafft	NSF sub of UCAR	284,756	1/1/05–2/28/07	National Science Digital Library Subcontract
Krafft/Lagoze	NSF sub of UCAR	1,049,516	1/1/07–12/31/07	Partial Supplemental Request 2007 (NSDL)
Lagoze	Mellon sub of U VA	794,553	10/1/04–9/30/07	The Open Source FEDORA Repository Development Project: Phase II
Lagoze	Mellon	524,000	10/1/06–9/30/08	Pathways Fabric: Augmenting Interoperability Across Scholarly Repositories
Lagoze/Warner	NSF	1,297,550	9/15/04–8/31/07	II: Pathways
Lee/Kleinberg	NSF	449,897	9/1/03–8/31/07	Graph-Based Approaches to Text Processing
Lee	Sloan	40,000	9/16/02–9/15/06	Sloan Research Fellowship
Marschner	NSF	412,000	2/15/04–1/31/09	CAREER: Modeling the Properties and Appearance of Materials
Marschner	Unilever	81,000	4/1/06–3/31/07	Scattering and Computer Visualisation of Human Hair
Marschner	NSF	300,000	2/1/06–1/31/09	Unifying Geometric and Volumetric Light Scattering for Accurate Rendering of Dense Geometry
Marschner	Sloan	45,000	9/16/06–9/15/08	Sloan Research Fellowship
Myers	NSF	349,999	1/1/02–12/31/06	CAREER: Practical Language-Based End-to-End Security
Myers	NSF	350,000	9/1/06–8/31/09	CT-ISC: Diaspora: A Secure, Reliable Federated Execution Platform and Object Store
Myers	Sloan	40,000	9/16/02–9/15/06	Sloan Research Fellowship
Myers	Microsoft	83,382	6/1/07–5/31/08	Language-Based Security Methods to make Web Applications Secure by Construction
Myers/Birman/ Schneider	NSF	1,600,000	9/1/04–8/31/07	Integrating Security and Fault Tolerance in Distributed Systems
Payette	Mellon	514,000	3/1/07–3/31/09	Fedora Service Framework: Filling out the Blueprint for Next-Generation Scholarly Communication Systems
Pingali/Bala/ Chew	NSF	98,061	8/1/06–12/31/06	CSR-AES: Parallelizing Irregular Applications for Multicore Processor
Pingali/Rugina	NSF	364,182	8/15/04–12/31/06	NGS: A System for Semi-Automatic Application-Level Checkpointing of Parallel Programs
Rugina	NSF	374,915	2/1/06–1/31/09	Compiler and Run-Time Support for Memory Management Using Explicit Memory Reclamation
Rugina	NSF Sub of U TX	136,088	1/1/07–7/31/08	NGS: A System for Semi-Automatic Application-Level Checkpointing of Parallel Programs
Schneider	AFOSR	4,638,325	6/1/02–3/31/07	AFRL/Cornell Information Assurance Institute

INVESTIGATOR	SPONSOR	AWARD	PERIOD OF AWARD	TITLE
Schneider	AFOSR	773,378	12/1/05–11/30/08	From Fault-tolerance to Attack Tolerance
Schneider	AFOSR	1,000,000	1/1/07–12/31/07	AFRL/Cornell Information Assurance Institute
Schneider	Microsoft	125,000	11/6/06	Research Gift
Schneider	Microsoft	200,000	6/29/07	Unrestricted Gift
Schneider	Microsoft	83,316	6/1/07–5/31/08	Trustworthy Distributed Services
Schneider	Microsoft	83,302	6/1/07–5/31/08	Operating Systems that leverage Trusted Co-Processors
Schneider/ Morrisett	AFOSR/PECASE	820,000	4/15/01–12/31/06	Next Generation Systems Languages
Schneider/ Morrisett/Kozen/ Myers	ONR	4,247,977	7/9/01–9/30/07	CIP - Language-Based Security for Malicious Mobile Code
Schneider/Sirer	AFRL	1,200,000	6/4/07–12/3/08	Nexus Operating System for Trustworthy Computing
Selman	DARPA sub of SRI	20,000	6/1/06–8/30/06	Analyzing Heuristic Approaches to Real-time, Distributed Coordination Problems
Selman/Gomes/ Kreitz	DARPA/AFRL	1,606,256	6/8/04–6/7/09	Boosting Reasoning Technology Through Randomization, Structure Discovery, and Hybrid Strategies
Selman/Gomes	NSF	405,000	6/1/07–5/31/10	Extending the Reach of SAT Technology - Quantification, Counting, and Sampling
Sengers	NSF	500,000	2/15/03–1/31/08	CAREER: Using Cultural Theory to Design Everyday Computing
Sengers	NSF	300,000	1/1/06–12/31/08	Collaborative Research: Closing the Affective Gap
Sengers	NSF	11,963	2/1/06–1/31/08	The Open World - The World of Open Systems
Shanmugasundaram	NSF	406,750	9/1/03–8/31/08	CAREER: Towards Unifying Database Systems and Information Retrieval Systems
Sirer	NSF	400,000	2/1/06–1/31/11	CAREER: Building Robust, High-Performance Infrastructure Services through Informed Resource Tradeoffs
Sirer	I3P	150,000	6/12/06–5/30/07	I3P Research Fellowship
Sirer	NSF	400,000	9/15/06–8/31/09	CT-ISC: Nexus: A New Operating System for Trusted Computing
Sirer/Birman/Pass	I3P	400,000	4/1/07–3/31/09	Secure Identity Management Infrastructure
Tardos	NSF	150,000	8/1/03–7/31/06	Approximation Algorithms and Applications in Network Games
Tardos	ONR	375,000	1/1/05–9/30/07	Network Games and Approximation Algorithms
Van Renesse	AFOSR	127,174	4/15/06–4/14/07	An Improved Testbed for Highly-Scalable Mission-Critical Information Systems
Van Renesse	NSF	25,000	9/15/06–8/31/07	Student Travel Support for OSDI 2006

TOTAL EXPENDITURES FOR FISCAL YEAR 2006-2007: \$16,799,076

Collaborative Research at Cornell

INVESTIGATOR	SPONSOR	AWARD	PERIOD OF AWARD	TITLE
Birman/Sirer/ECE	NSF	1,250,000	9/1/04–8/31/08	NETS-NOSS: Ultra Low-Power Self-Configuring Wireless
Cardie/Law/Hotel	NSF	750,000	11/15/05–10/31/08	Natural Language Processing Support for eRulemaking
Cardie/Ling/ Comm	NSF	679,991	11/15/06–10/31/09	DHB -Dynamics of Deception in Computer-Mediated Environments
Caruana/ Riedewald/ Lab of Ornithology	NSF	2,701,622	9/15/04–8/31/08	ITR-(ASE+EVS)- (dmc+sim): Tracking Environmental Change through the Data Resources of the Bird-monitoring Community

INVESTIGATOR	SPONSOR	AWARD	PERIOD OF AWARD	TITLE
Chew/Vavasis/ MAE	NSF	1,393,410	9/1/04–8/31/08	ITR-(ASE)-(sim+dmc):Collaborative Research: Algorithms for Large-Scale Simulations of Turbulent Combustion
Demers/ Huttenlocher/ Kleinberg, J./ Marschner/ Astronomy	NSF	1,799,183	9/15/04–8/31/09	Petabyte Storage Services For Data-Driven Science
Elber/Joachims/ CTC	NIH	980,042	2/1/04–1/31/08	Optimization of Folding and Threading Potential
Elber/Weill	CU seed grant	50,000	11/15/05–12/31/06	Allosteric Mechanisms and the Function of G Protein-Coupled Receptors
Elber/CTC	Microsoft	400,000	1/1/06–12/31/06	High Performance Computing for Bioinformatics on Windows
Francis/ECE	DARPA sub of BAE	500,000	3/1/06–12/14/08	CBMANET
Gehrke/EAS	NASA	239,188	2/15/06–2/14/07	Predicting Right Whale Distributions from Space: An Operational System for Marine Ecosystem Modeling
Gehrke/CISER	NSF	488,950	9/15/06–8/31/10	CT-T: Collaborative Research: Preserving Utility While Ensuring Privacy for Linked Data
Gehrke/Physics	NSF	149,380	6/15/07–5/31/08	Scientific Workflow Management in Energy Frontier Physics
Gries/CTC	NSF	597,638	1/1/07–12/31/09	BPC Worlds for Information Technology and Science (WITS)
Guckenheimer/ Math/Physics/ CCB/MBG/ Biomedical/ Mole Med/Eng/Soc	NSF	3,436,000	10/1/03–9/30/08	IGERT: Program in Nonlinear Systems
Huttenlocher/ Arms/ Kleinberg, J./ Soc / CTC	NSF	1,999,990	1/1/06–12/31/08	Very Large Semi-Structured Datasets for Social Science Research
Huttenlocher/ Selman/MAE/ CCMR	DARPA	1,000,000	9/27/06–12/31/07	Team Cornell: Autonomous Vehicle for Operations in Urban Environments
Krafft/Arms/ Lagoze/ Eng Library/ Comm	NSF	8,845,453	1/1/03–12/30/06	Collaborative Project: Core Integration - Leading NSDL Toward Long-Term Success
Marschner/PCG	Eastman Kodak	129,665	11/1/06–10/31/07	Digital Imaging Media: Measurement, Modeling and Perception
Marschner/ James/FSE&AD	NSF	300,000	7/1/07–6/30/10	Accurate and Efficient Visual Simulation of Fiber-based Mechanical Structures
Riedewald/ Caruana/ Lab of Ornithology	NSF	987,334	7/1/06–6/30/10	SEI+II: Ecological Discovery & Inference: Tools for Data-driven Exploration and Testing of Observational Data
Schneider/ Birman/Sirer/ ECE	NSF	3,040,000	7/1/05–6/30/10	Team for Research in Ubiquitous Secure Technology (TRUST)

INVESTIGATOR	SPONSOR	AWARD	PERIOD OF AWARD	TITLE
Sirer/Gehrke/ Demers/ECE	NSF	410,000	9/15/02-8/31/06	The Ad Hoc Classroom: Integrating Emerging Wireless Communications and Networking Technologies into Mainstream Computer Science and Electrical Engineering Curricula
Tardos/ Kleinberg/ Huttenlocher/ Halpern/ OR&IE/Econ	NSF	2,468,677	9/1/03-8/31/08	ITR: Networks of Strategic Agents: Theory and Algorithms
Vavasis/CEE	NSF	500,000	8/1/04-7/31/07	MSPA-MCS: Automatic Geometric Simplification

Submitted Proposals

INVESTIGATOR	SPONSOR	AWARD	PERIOD OF AWARD	TITLE
Birman	I3P	150,000	9/2/11-9/1/12	I3P Fellowship: Offloading Framework: Exploiting Computational Capacity on Peripheral Devices for Security
Birman / Gehrke	NSF	1,600,000	7/2/11-7/1/15	CT-T: Collaborative Research: FISC: The Financial Infrastructure Security Consortium
Birman / Gehrke Computing	US Treasury	149,422	1/2/11-1/1/12	SAFE: Secure Architecture for Financial Enterprise
Constable	AF sub of ATC-NY	30,000	8/2/11-5/1/12	Scalable Formal Methods for Distributed Systems
Gehrke	NYSTAR	748,380	11/2/11-11/1/14	From Web 2.0 to Games 2.0
Gries / Van Loan	NSF	149,626	10/2/11-10/1/13	Individualizing Instruction in Computer Science
Halpern / Gehrke	ARI	697,799	10/2/11-10/1/14	A Calculus of Evidence and a Database Implementation of It
Huttenlocher	NSF	448,819	6/2/11-6/1/14	RI: A Context-Based Approach to the Recognition and Localization of Visual Object Categories
Huttenlocher	NYSTAR CEIS	30,000	7/2/11-7/1/12	Detection and Localization of Visual Object Categories in Photographic Images
Huttenlocher	Eastman Kodak	75,837	7/1/07-6/30/08	Detection and Localization of Visual Object Categories in Photographic Images
Joachims	NSF	405,000	6/1/07-5/31/10	RI: Learning Structure to Structure Mappings
Joachims	AFOSR	754,078	10/1/07-9/30/10	Content-Aware Models of Social Network Dynamics
Kleinberg, R.	BSF	53,000	10/1/07-9/30/11	Impossibilities and Alternative Solution Concepts for Algorithmic Mechanism Design
Kleinberg, R.	NSF	250,000	8/1/07-7/31/10	Combinatorial and Algorithmic Aspects of Network Coding
Krafft	UCAR sub of NSF	403,152	8/1/07-7/31/09	Beyond Penguins and Polar Bears: Literacy and IPY
Krafft	NSF sub of UC Berkeley	90,000	11/1/07-10/31/10	Science and Math informal Learning Educators (SMILE)
Krafft	NYHS	58,788	10/1/05-9/30/08	VolTS Partnership with NSDL
Lagoze	SGER	100,000	9/1/07-8/31/08	SGER: Advancing the State of eChemistry: Workshop and Pilot Study
Myers	AFOSR	596,077	2/1/07-7/31/08	Accountable Information Flow for Java-Based Web Applications
Pass	BSF	92,000	10/1/07-9/30/11	Composition of Cryptographic Protocols
Pass	AFOSR	396,000	12/1/07-11/30/10	Concurrent Security of Cryptographic Protocols: From Foundations to Practice
Schneider/ Gehrke	Fast	950,000	1/1/07-12/31/11	Information Access Disruptions Enabling Technology and Emerging Services in the Information Age
Schneider/ Birman	AFOSR	75,000	7/1/07-6/30/08	Homogeneous Enclave Software vs Controlled Heterogeneous Enclave Software
Van Renesse	NSF	500,000	7/1/07-6/30/09	CT-ISC: Trustworthiness Through Offloading: Exploiting Computational Capacity on Peripheral Devices for Security

INVESTIGATOR	SPONSOR	AWARD	PERIOD OF AWARD	TITLE
Van Renesse	DARPA sub of ATC-NY	369,802	10/1/07–12/31/08	ATC-NY IAMANET ZENO Byzantine
White/Gehrke/ Demers	NSF	800,000	8/1/07–7/31/10	Design Methodologies for Scalability in Computer Games

Submitted Collaborative Research at Cornell

INVESTIGATOR	SPONSOR	AWARD	PERIOD OF AWARD	TITLE
Arms/ Huttenlocher/ Kleinberg, J./ Soc/ CTC	NSF	897,657	10/1/07–9/30/09	Computer Science Research using the Cornell Web Lab to Study Social and Informational Processes on the Web
Caruana/ECE	NSF	784,317	7/1/07–6/30/10	Computer Architecture Optimization: A Machine Learning Approach
Elber/AEP	NIH	377,701	12/1/07–11/30/09	Local Initiation of Global Molecular Events: Experiment and Theory
Elber/Biotech/ PBG/MBG	NSF	975,673	8/1/07–7/31/08	Acquisition of Massively Parallel Sequencing Instruments to be used as a University-Wide Core Facility Shared Research Resource
Gehrke/Physics	DoE	6,799,738	7/1/06–6/30/11	Development and Deployment of a Dynamic, Grid-Award Analysis-Development and Interactive Data Analysis Site for CMS
Gehrke/EAS	NASA	1,032,883	5/1/06–4/30/10	Predicting Right Whale Distributions from Space: An Operational System for Marine Ecosystem Modeling
Huttenlocher/ECE	Army	6,378,000	10/1/07–9/30/12	Center for Networked Autonomous Systems
Kozen/M&I/ Vet/Lab of Ornithology	NSF	1,585,765	8/1/07–7/31/12	Adaptive Modeling of the Transmission of H5N1 Avian Influenza by Migratory and Domesticated Birds in North America
Kleinberg, J./ Sociology	McDonnell	444,742	1/1/08–12/31/10	Social Network Structure and Dynamics of Self-Organization in Online Communities
Marschner/PCG	Unilever	785,204	4/1/07–3/31/10	Optics and Computer Visualization of Hair and Textiles

The background of the page features a grey field filled with a pattern of overlapping, semi-transparent arrows pointing in various directions. A solid blue horizontal bar spans the top of the page, containing the title text in white.

FACULTY BIOGRAPHIES