CS@CU

NEWSLETTER OF THE DEPARTMENT OF COMPUTER SCIENCE AT COLUMBIA UNIVERSITY VOL.7 NO.1 SPRING 2011

From left to right: Augustin Chaintreau, Michael Collins and Xi Chen

CUCS **Welcomes Three New Faculty**

Michael Collins named Vikram S. Pandit Professor of Computer Science

Vikram S. Pandit Professor of Computer Science
Michael Collins

"**Michael Collins** revolutionized the natural language processing (NLP) field with his work on parsing. He was the first to build a robust, statistical parser, which is now the most downloaded software in NLP and is widely used in applications around the world."

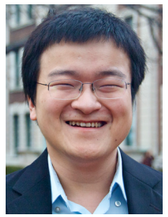
– KATHLEEN MCKEOWN

MICHAEL COLLINS received a Ph.D. in computer science from the University of Pennsylvania in 1998. From 2003 to 2010 he was a professor in computer science at MIT; prior to that he was a researcher at AT&T Labs-Research. Professor Collins's research has focused on topics including statistical parsing, structured prediction problems in machine learning, and NLP applications including machine translation, dialog systems, and speech recognition. His awards include a Sloan fellowship, an NSF career award, and best paper awards at several conferences: EMNLP (2002, 2004, 2010), UAI (2004, 2005), and CoNLL 2008.

CUCS: Tell us about your field of research.

My research is in the area of machine learning for natural language processing (NLP). I work on applications such as machine translation, speech recognition, dialog systems and information extraction. I got into this research area for a few reasons. I first learned about natural language processing while doing my masters, and quickly became fascinated by the idea of formal or computational models of human languages. These models go back to Chomsky's work in linguistics, which were later developed further by computer scientists. In recent years there has been a great deal of interest in statistical models for NLP, the idea being to learn models from large amounts of linguistics data. Statistical NLP poses many interesting challenges, both in terms of the development of

(continued on next page)



"**Xi Chen** has solved a decades-old problem in game theory and market equilibria, showing inherent computational limitations in predicting equilibria even when all the relevant information is readily available."
– MIHALIS YANNAKAKIS

Assistant Professor
of Computer Science
Xi Chen

"**Augustin Chaintreau** is one of the mathematical pioneers of mobile social networking in the networking research community, running the first large scale experiment collecting social contact data. He has been in the trenches before there was an iPhone, Foursquare or Facebook places."
– VISHAL MISRA

Assistant Professor
of Computer Science
Augustin Chaintreau

learning or inference algorithms, and also in terms of the development of statistical models. In addition to the technical challenges, I find the underlying applications in NLP fascinating-machine translation, for example, is a very interesting problem, that in addition has a great deal of practical relevance.

What are your ongoing research projects and what motivated you to focus on them?

One research project I've been working on is models for statistical machine translation that make greater use of linguistic structures. In recent years, great progress has been made in machine translation through statistical models—Google translate being the obvious example of such a system. However these systems make relatively little use of linguistic structure, and because of this sometimes have trouble producing coherent output, or in modeling systematic differences in word order between languages. We've been developing statistical models that make explicit use of linguistic structures, in an effort to address these problems. A second recent research project is more on the algorithmic side, and concerns efficient inference algorithms for several statistical models found in NLP. In many applications, inference (e.g., finding the most likely translation for a sentence) is very costly; because of this, researchers have generally focused on approximate search methods, which have quite weak guarantees. We've been developing exact inference methods, based on ideas from linear programming relaxations, and combinatorial optimization. I'm excited about this work because we've been able to exactly solve several inference problems that people previously thought were intractable, and this has led to improved performance on several problems.

What is your approach to research?

I like working on real-world, practical problems; but I also enjoy developing formal models and algorithms. Ideally I try to develop approaches that give real improvements on problems, but that have a clean theoretical underpinning.

XI CHEN received his B.S. degree in Physics Mathematics in 2003 and his Ph.D. in Computer Science in 2007, both from Tsinghua University, China. Before joining Columbia, he was a postdoctoral researcher at the Institute for Advanced Study, Princeton University and USC. His research interests include Algorithmic Game Theory and Theoretical Computer Science in general. He works on natural and fundamental computational problems that arise from a game-theoretic study of the Internet, e-commerce and other decentralized systems. He won best paper awards at FOCS 2006 and at ISAAC 2009.

What area of computer science do you focus on?

My research is in theoretical computer science in general, and I am particularly interested in algorithmic game theory and economics (AGTE). In the past decade, concepts and approaches from game theory and economics have been widely applied in computer science. Many of the applications are motivated by the rise of the Internet and e-commerce. Because of the absence of central authority, it is not surprising that being able to understand and predict the behavior of selfish agents is critical to the success of many Internet-based applications. The goal of this field is then to help the design of such decentralized systems that are both computationally efficient and robust against selfish and strategic manipulations.

What are some of your ongoing research projects?

Recently I have been working on the computation of market equilibria. In a classical market, there are buyers and sellers exchanging money and goods. An equilibrium is then a stable state in which everybody feels satisfied about the outcome of the exchange. The model has been studied in economics for a long time, and has been recognized by many as the crown jewel of mathematical economics. While an equilibrium always exists under mild conditions, in general we do not know how to find one efficiently. Computational problems related to this model have been studied intensively in the computer science community in the past decade, from either the algorithmic or the complexity-theoretic perspective.

What do you consider the most exciting challenges of your current research?

One of the greatest challenges I see is to bridge the gap between theory and practice. This is by no means easy and coming up with a good research problem that is interesting in both theory and practice is probably the hardest part in many situations. I believe that working in AGTE would offer me a lot of opportunities to collaborate with researchers in other areas of CS. Therefore, the diversity and excellence of the Columbia CS faculty is definitely the factor that attracted me the most for coming here.

AUGUSTIN CHAINTREAU graduated in 2006 from École Normale Supérieure, Paris. He joined Technicolor (previously Thomson) to work on advanced communication platforms. Prior to that, he worked at Intel Research Cambridge where he conducted the first measurement campaign of opportunistic mobile dissemination. During his Ph.D, he worked at Alcatel Bell and IBM Research on characterizing scalable resource sharing systems in the presence of fairness and reliability constraints. His research deals with mobile and social dynamics in information systems.

Tell us about your area of specialization and what attracted you to it?

My research background is in probabilistic modeling and networked algorithms for mobile and social computing. This is an area packed with both exciting applications and fundamental problems that we are only starting to discover. I first got into this field by chance when, as an undergraduate I completed a research internship with Sprint in the bay area. We used the algebraic property of matrices to explain why several commercial products would not scale, and that was a lot of fun.

What attracted you most to a faculty position at Columbia?

Where you work is incredibly important! I have always looked for academic rigor and energetic support from my colleagues. Visiting Columbia last year and seeing how people work here just blew me away. Another important fact is that Columbia's "footprint" spans the social sciences, economics, business and journalism. One reason I'm moving to academia after five years in industry is that research in mobile and social networks today impacts all these fields.

What are some of your ongoing research directions?

As mobile networks are now dealing with an abundance of local information, we are working on a distributed recommendation engine. We try to predict what information would be relevant to you the same way Netflix recommends movies.

Do you have a method or approach that you like to use in your research?

I learned at least one important thing from my mentors: theory and system building truly enrich each other. This is interesting because sometimes they might seem to have very different goals. When a set of simple assumptions, inspired by empirical tests and practical insights, allows us to prove or disprove an intuition, this is when I feel passion for my work.

What are the most interesting aspects of your research?

Certainly one thing that keeps me awake is how we will exploit the vast wealth of social and mobile information that we all produce everyday. We are starting to understand which tasks could be solved more efficiently using social networking tools, as well as the limits of these tools, and how they can be misused. This in turn leads to a series of computational challenges, and I think solving any of them could have a large impact.

Computer Science Meets Journalism at Columbia University

This year, Columbia's Department of Computer Science will begin offering a novel Dual M.S. Degree Program in collaboration with the Graduate School of Journalism. The new program is designed to prepare students for careers in the digital-media world of the 21st century.

The new program is the first of its kind, as it aims at bridging the gap between the evolving discipline of journalism and new information technologies, while leveraging the unique environment of New York City. Home to the three traditional major American television networks, four of the ten largest newspapers, several large cable television channels and radio stations, and most of the book and music publishing industry, New York has been called the "Media Capital of the World." And along with traditional media, the city has seen a continuous growth of its media-oriented information-technology industry, which includes software and systems development for interactive media, networking, and Internet services.

In the fall of 2011, the first class of students will start the five-semester graduate program. By taking courses in both the School of Journalism and the School of Engineering and Applied Science, as well as working on inter-disciplinary projects under the joint supervision of professors from both schools, the students will simultaneously earn M.S. degrees in Computer Science and Journalism.

To better prepare students for careers in the digital-media world, the new program will teach them about the impact of digital techniques on journalism, the emerging role of citizens in the news process, the influence of social media, and the changing business models that will support news gathering, reporting, and delivery.

The following is a sample of the program course offering from the two disciplines:

Journalism

- Reporting and Writing
- Journalism essentials: law, business, ethics, and history
- Skills classes in photo, audio, video, and other digitally based media
- Digital Media Workshop
- Business and Economic Reporting
- Science Reporting
- Interactive Media Storytelling

Computer Science

- Introduction to Databases
- Advanced Software Engineering
- User Interface Design
- Analysis of Algorithms
- Network Security
- Natural Language Processing
- Web-Enhanced Information Management
- Intrusion Detection
- Seminar on anonymity and privacy
- Machine Learning
- Search Engine Technology

At the same time, students will have the opportunity to apply some of the new media-related technologies that have been developed by professors and researchers in the Department of Computer Science, including: speech understanding and speech synthesis, automatic text summarization, 3D user interface design and augmented reality, computer vision and digital imaging, statistical machine learning for the analysis of natural languages, and machine learning for the analysis of massive data sets generated by social networks.

Possible Career Paths

Online Editor/Manager of Information Technology at a large news organization.
SKILL SET: knowledge of software creation, databases, user interfaces, security, large-scale text processing to engage readers on the web and mobile devices while ensuring content and design meet the standards for quality journalism.

Data-mining expert for journalistic applications and investigative journalism.
SKILL SET: knowledge of statistics, machine learning, and advanced language processing techniques, Web search for the purpose of gathering and synthesizing data.

Entrepreneur/founder of media startup.
SKILL SET: strong editorial background and knowledge about the business of journalism; basic knowledge of computing with specialization in social media, data mining, and other application areas

For more information: http://www.cs.columbia.edu/education/ms/journalism

WHAT PEOPLE ARE SAYING

The concept makes sense, the problem it addresses is real, and Columbia is capable of taking on the challenge. But we were most fascinated by the technologies these professors hope their graduates will contribute. –WIRED MAGAZINE

One goal of the Columbia program, according to Bill Grueskin, the dean of academic affairs, is to produce journalists who will "take it several steps beyond—to where they're creating a lot of their own new tools." That means learning enough computer science and software engineering to be able to design tools for information gathering, synthesis, analysis and circulation—or enough, at least, to see what technology can do for journalism. –THE NEW YORK TIMES

Many of today's programmer-journalists got to where they are thanks to self-directed education and hacking together courses and other educational opportunities to build their skills. But the new Columbia program, along with other initiatives, suggests that the next wave of programmer-journalists could be trained in specialized education programs that combine a traditional engineering/computer science degree with a traditional journalism education. Universities are working to either alter existing journalism programs or create new joint degrees to formalize the training of these workers. – PBS MEDIASHIFT

Professor
Mihalis Yannakakis

"As one of the leading academics working to define the laws that govern the computational world, Mihalis is both an exceptional educator and a ground-breaking scholar." – FENIOSKY PEÑA-MORA

NAE Elects Professor Yannakakis a Member

Professor **Mihalis Yannakakis** has been elected a member of the prestigious National Academy of Engineering (NAE), one of the highest professional distinctions accorded to an engineer.

Yannakakis, the Percy K. and Vida L. W. Hudson Professor of Computer Science, was honored for his contributions to algorithms and computational complexity. He is one of 77 new members elected this year and joins a total membership of nearly 2,500 worldwide.

NAE membership honors those who have made outstanding contributions to engineering research, practice, or education, including, where appropriate, significant contributions to the engineering literature, and to the pioneering of new and developing fields of technology, making major advancements in traditional fields of engineering, or developing/implementing innovative approaches to engineering education.

"It is a great honor for me to join such a distinguished body," Yannakakis says. "I am very happy to be elected."

The honor has garnered him praise from Columbia Engineering leaders.

"We are thrilled that the work of Mihalis Yannakakis has been recognized by his election to the National Academy of Engineering," said Feniosky Peña-Mora, Dean of Columbia Engineering School. "As one of the leading academics working to define the laws that govern the computational world, Mihalis is both an exceptional educator and a ground-breaking scholar. This is a wonderful tribute to him and, by extension, to our world-renowned faculty at Columbia Engineering."

Computer Science Department Chairman Shree Nayar, himself an NAE member, calls Yannakakis one of the world's leading theoretical computer scientists.

"His contributions to computer science are remarkable in terms of their depth and impact," says Nayar, the T.C. Chang Professor of Computer Science. "His foundational work on algorithms and computational complexity has led to new approaches to important optimization, approximation, and testing problems in broad areas of computer science. Mihalis's election to the National Academy of Engineering is a richly deserved honor. We are proud to have him as a colleague."

Yannakakis studied at the Varvakeio High school, at the National Technical University of Athens (Diploma in Electrical Engineering, 1975), and at Princeton University (Ph.D. in Computer Science, 1979).

He worked at Bell Labs Research from 1978 until 2001, as a member of the technical staff and then as Head of the Computing Principles Research Department. He was later Director of Computing Principles Research at Avaya Labs and a Professor of Computer Science at Stanford University, before joining Columbia University in January 2004.

2010-11 **Distinguished** Lecture Series

The Computer Science had an exciting lineup of speakers during the 2010-11 academic year. The speakers include leading figures in science in both academia and industry with a particular emphasis on the internet's origins, internet advertising, legal aspects of the digital world and user interaction.

The lectures were open to all members of the Columbia University campus as well as the general public. Details on the talks are available via the department's website *(http://www.cs.columbia.edu/lectures)*. Visitors also had opportunities to meet with students and faculty in one-on-one settings. Such visits frequently inspire new research ideas at the department and help forge collaborations between our faculty and leading colleagues. We hope you will attend future lectures in our subsequent series.

Professor **Leonard Kleinrock** of the University of California, Los Angeles visited on October 25 and spoke about **A Brief History of the Internet and its Dynamic Future.**

ABSTRACT: In this presentation we discuss the history and future of the Internet. The early work on packet switching is traced and then a brief description of the critical events in the growth of the Internet is given. We then present a vision of where the Internet is heading with a focus on the edge where user participation, flexible applications and services, and innovation are appearing. We foresee a network with extreme mobility, ubiquity, personalization, adaptivity, video addiction and surprising applications as yet unimagined.

BIOGRAPHY: Professor Leonard Kleinrock is Distinguished Professor of Computer Science at UCLA. Known as a "Father of the Internet", he developed the mathematical theory of packet networks, the technology underpinning the Internet. His Host computer at UCLA became the first node of the Internet in September 1969. He wrote the first paper and published the first book on the subject; he also directed the transmission of the first message ever to pass over the Internet. Kleinrock's work was recently recognized when he received the 2007 National Medal of Science, the highest honor for achievement in science bestowed by the President of the United States.

Leonard Kleinrock received his Ph.D. from MIT in 1963. He has served as a Professor of Computer Science at the University of California, Los Angeles since then, serving as Chairman of the department from 1991-1995. He received his B.E.E. degree from CCNY in 1957, and his M.S. degree from MIT in 1959. He has published over 250 papers and authored six books on a wide array of subjects, including packet switching networks, packet radio networks, local area networks, broadband networks, gigabit networks, nomadic computing, performance evaluation, and peer-to-peer networks. During his tenure at UCLA, Dr. Kleinrock has supervised the research for 47 Ph.D. students and numerous M.S. students.

Dr. Kleinrock is a member of the National Academy of Engineering, a member of the American Academy of Arts and Sciences, an IEEE fellow, an ACM fellow, an INFORMS fellow, an IEC fellow, a Guggenheim fellow, and a founding member of the Computer Science and Telecommunications Board of the National Research Council. Among his many honors, he is the recipient of the National Medal of Science, the L.M. Ericsson Prize, the NAE Charles Stark Draper Prize, the Marconi International Fellowship Award, the Dan David Prize, the Okawa Prize, the IEEE Internet Millennium Award, the ORSA Lanchester Prize, the ACM SIGCOMM Award, the NEC Computer and Communications Award, the Sigma Xi Monie A. Ferst Award, the CCNY Townsend Harris Medal, the CCNY Electrical Engineering Award, the UCLA Outstanding Faculty Member Award, the UCLA Distinguished Teaching Award, the UCLA Faculty Research Lecturer, the INFORMS President's Award, the ICC Prize Paper Award, the IEEE Leonard G. Abraham Prize Paper Award, and the IEEE Harry M. Goode Award.

Doctor **Andrei Broder** of Yahoo! Research visited on November 8 and spoke about **Computational Advertising**.

ABSTRACT: Computational advertising is an emerging scientific sub-discipline, at the intersection of large scale search and text analysis, information retrieval, statistical modeling, machine learning, classification, optimization, and microeconomics. The central challenge of computational advertising is to find the "best match" between a given user in a given context and a suitable advertisement. The context could be a user entering a query in a search engine ("sponsored search"), a user reading a web page ("content match" and "display ads"), a user watching a movie on a portable device, and so on. The information about the user can vary from scarily detailed to practically nil. The number of potential advertisements might be in the billions. Thus, depending on the definition of "best match" this challenge leads to a variety of massive optimization and search problems, with complicated constraints and difficult data representation issues. This talk will give an introduction to computational advertising focusing on the interplay between science, engineering, and marketplace.

BIOGRAPHY: Andrei Broder is a Yahoo! Fellow and Vice President for Computational Advertising. Previously he was an IBM Distinguished Engineer and the CTO of the Institute for Search and Text Analysis in IBM Research. From 1999 until 2002 he was Vice President for Research and Chief Scientist at the AltaVista Company. He was graduated Summa cum Laude from Technion, the Israeli Institute of Technology, and obtained his M.S. and Ph.D. in Computer Science at Stanford University under Don Knuth. His current research interests are centered on computational advertising, web search, context-driven information supply, and randomized algorithms. Broder is co-winner of the Best Paper award at WWW6 (for his work on near-duplicate page detection) and at WWW9 (for his work on mapping the web). He has authored more than a hundred papers and was awarded thirty patents. He is a member of the US National Academy of Engineering, a fellow of ACM and of IEEE, and past chair of the IEEE Technical Committee on Mathematical Foundations of Computing.

Professor **Danielle Citron** of the University of Maryland visited on November 15 and spoke about **Digital Discrimination.**

ABSTRACT: Social network sites and blogs have increasingly become breeding grounds for anonymous online groups that attack women and minorities. The attacks include rape threats, privacy invasions, defamation, and technological attacks that silence victims. Victims go offline or assume pseudonyms to prevent future attacks, impoverishing online dialogue and depriving victims of the social and economic opportunities associated with a vibrant online presence.

Although social and legal norms have dampened offline discrimination, the internet's Wild West culture and architecture invites bigots to move their hatred to cyberspace. The Internet facilitates anonymity, loosening social norms that constrain noxious behavior. It brings people together—a benefit when used for good purposes but especially dangerous when cyber mobs band together. Moreover, unlike offline hate that loses its impact with time, online hatred can be permanent.

These are serious problems. Nonetheless, the public pays little attention to cyber hate. Law enforcement routinely trivializes cyber harassment of women and minorities, deeming it ranting victims can, and should, ignore. Police officers often refuse to pursue cyber harassment complaints on the grounds that the conduct is legally insignificant. Victims often do not pursue legal claims for fear that they won't be taken seriously or that no law protects them. That is an unfortunate result of false information and mistaken assumptions—the law criminalizes cyber harassment and it may provide compensation for discrimination online.

Digital Discrimination aims to chart a way forward. Leaving well enough alone is not an option. The more that we ignore cyber hate, the more prevalent in mainstream social media it will become. Although law has an important role to play, the existing legal framework can only take us so far. Other institutions can, and should, tackle cyber discrimination, including online intermediaries, educators, and interest groups.

BIOGRAPHY: Danielle Citron is a Professor of Law at the University of Maryland School of Law, where she focuses on information privacy law, cyber law, and administrative law. Her scholarship appears in the Boston University Law Review, California Law Review, George Washington Law Review, Hastings Law Journal, Michigan Law Review, Southern California Law Review, Washington University Law Review, U.C. Davis Law Review, and University of Chicago Legal Forum. Professor Citron is an Affiliate Fellow of the Yale Information Society Project and an Advisory Board member of The Future of Privacy. She presents her work widely, including at Yale Law School, Harvard Law School, University of Chicago Law School, Stanford Law School, New York University School of Law, University of Michigan Law School, and Princeton University. The Denver University Law Review devoted a symposium to her work on cyber gender harassment entitling it Cyber Civil Rights: New Challenges to Civil Rights and Civil Liberties in our Networked Age. Danielle blogs at Concurring Opinions. In 2005, the University of Maryland School of Law's students voted her the "Teacher of the Year." Before teaching, she clerked for the late Honorable Mary Johnson Lowe of the U.S. District Court for the Southern District of New York and practiced at Willkie, Farr & Gallagher.

Further details of the lecture series are available online at: http://www.cs.columbia.edu/lectures

Salman Abdul Baset and Professor **Henning Schulzrinne** win best paper award at IPTCOMM 2010 for their paper titled "Reliability and Relay Selection in Peer-to-Peer Communication Systems".

About Face: NPR interviews Professor **Peter Belhumeur** on Face Recognition. He discusses his work on face recognition on National Public Radio. The full interview is available at *http://www.studio360.org/episodes/2010/12/17*

Matei Ciocarlie has been named the winner of the 2010 Robot-Dalen Scientific award for his Ph.D. Thesis work on "Low-Dimensional Robotic Grasping: Eigengrasp Subspaces and Optimized Underactuation".

Scientific American features research of Ph.D. student **Oliver Cossairt** and Professor **Shree Nayar**, "Draw the Curtains: Gigapixel Cameras Create Highly Revealing Snapshots".

Oliver Cossairt

Professor **Shree Nayar**

Today's high-resolution cameras capture images with pixel counts in the tens of millions. In the future digital cameras may be able to capture images with billions of pixels. Columbia's CAVE lab has shown that by using a large ball lens, an array of planar sensors, and deconvolution as a post processing step, we can capture gigapixel images with a very compact camera.

Ang Cui

Ang Cui and Professor **Salvatore Stolfo** won best paper award at the 2010 Annual Computer Security Applications Conference for their paper "A Quantitative Analysis Of The Insecurity Of Embedded Network Devices: Results Of A Wide-area Scan."

Professor **Salvatore Stolfo**

David Elson

Ph.D. Student **David Elson** and Professor **Kathleen McKeown** won the IBM Best Student Paper award at the 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010) in Uppsala, Sweden along with co-author Nicholas Dames, Professor in the Department of English and Comparative Literature. Elson, McKeown and Dames studied the properties of social networks based on conversation in a corpus of 60 British novels from the Victorian era. They examined the correlation of these properties with literary theories about interaction in the 19th century novel. The interdisciplinary paper was titled "Extracting Social Networks from Literary Fiction."

Professor **Kathleen McKeown**

Professor **Steve Feiner** has been elected to the CHI Academy, which is an honorary group of individuals who have made extensive contributions to the study of HCI and who have led the shaping of the field. For the citation, please see: *http://www.sigchi.org/about/awards/2011-sigchi-awards*

Congratulations to Professor **Steve Feiner** and his co-authors B. MacIntyre, M. Haupt, and E. Solomon, for winning the UIST'10: Lasting Impact Award for their 1993 paper "Windows on the world: 2D windows for 3D augmented reality". The conference byline is "UIST (ACM Symposium on User Interface Software and Technology) is the premier forum for innovations in the software and technology of human-computer interfaces." The conference has been held yearly for the past 22 years. This is the eighth year of Lasting Impact awards.

Professor **Luis Gravano** Awarded NSF Grant for "Detection and Presentation of Event Content in Social Media".

Social media sites such as Twitter, Facebook, YouTube, and Flickr host an ever-increasing amount of user content captured or produced in association with real-world events, from presidential inaugurations to community-specific events. Unfortunately, the existing tools to find, organize, and present the social media content associated with events are extremely limited. This project will develop critical end-to-end information processing and presentation methods that will transform public access to real-world event information from social media sources.

More information can be found at *http://www.nsf.gov/awardsearch/showAward.do?AwardNumber=1017389*

David Harmon received a CRA/CCC Computing Innovation Fellowship, which will fund Harmon for up to two years at a postdoctoral position to "advance the field of computing and its positive impact on society."

C.I. Fellow Harmon will be working at New York University, where he will investigate contact algorithms for geometric modeling. Geometric modeling is concerned with the description and manipulation of shapes for many purposes, such as computer-aided design, mechanical engineering, and image processing. The developed contact algorithms will take contacts and collisions into consideration when modeling complex geometry, such as subdivision surfaces. The impact of such algorithms can be applied to many practical scenarios including computer-aided manufacturing and medical simulation systems, where contact sensitive models are a physical requirement.

Harmon, who was an NSF Graduate Research Fellow during his doctoral studies at the Columbia University School of Engineering & Applied Science, completed his Ph.D. thesis in 2010, as a member of the Columbia Computer Graphics Group directed by Professor **Eitan Grinspun**. He has worked at both Walt Disney Animation Studios (makers of *Snow White* through *Tangled*) and Weta Digital (makers of *The Lord of the Rings* through *Avatar*), applying research technologies to problems in digital special effects. His work on contact algorithms for the motion of fabric is used in films such as Disney's *Tangled*.

Professor **Julia Hirschberg** was interviewed on *The Leonard Lopate Show* on WNYC on 8/25/10 discussing speech technologies in general and emotional speech in particular. The interview is available at *http://beta.wnyc.org/shows/lopate/2010/aug/25/computers-and-language/*

Professor **Kenneth Ross** was awarded a NSF Grant for "Databases on Multicore Processors".

Applications such as traffic monitoring, mobile user management, and sensor networks need to process large volumes of updates while supporting on-line analytic queries. With large amounts of RAM, single machines are potentially able to manage hundreds of millions of items. With multiple hardware threads, as many as 64 on modern commodity multicore chips, many operations can be processed concurrently. Processing queries and updates concurrently can cause interference. Queries need to see a consistent database state, meaning that at least some of the time, updates will need to wait for queries to complete. To address this problem, a RAM-resident snapshot of the database is taken at various points in time. Analytic queries operate over the snapshot, eliminating interference, but allowing answers to be slightly out of date. Several different snapshot creation methods are being developed and studied, with the goal of being able to create snapshots rapidly (e.g., in fractions of a second) while minimizing the overhead on update processing. These problems are studied both for traditional server machines, as well as for multicore mobile devices. By keeping personalized, up to date data on a user's mobile device, a wide range of potential new applications could be supported while avoiding the privacy concerns of widely distributing one's location. The research focus is on how to efficiently utilize the many processing cores available on modern machines, both traditional and mobile devices. A primary goal is to allow performance to scale as additional cores become available in newer generations of hardware. More information can be found at *http://www.nsf.gov/awardsearch/showAward.do?AwardNumber=1049898*

Malek Ben Salem and Professor **Salvatore J. Stolfo**, won best paper award at The 2nd International Workshop on Managing Insider Security Threats (MIST 2010). The paper was entitled "Detecting Masqueraders: A Comparison of One-Class Bag-of-Words User Behavior Modeling Techniques". The workshop was conducted in conjunction with the 4th IFIP International Conference on Trust Management in Morioka, Japan, June 2010.

Professor **Henning Schulzrinne** is the recipient of the 2010 William Terry Lifetime Distinguished Service Award for his contribution and service to the IEEE and Region 1. This award is intended to recognize those whose personal efforts have provided leadership, creativity, guidance, hard work, and inspiration in a wide range of IEEE activities over a long period of time.

Professor **Rocco Servedio** is the 2010 recipient of the Distinguished Teaching Award from the Department of Computer Science. This award is in recognition of his dedication to teaching and his efforts to make Computer Science accessible to all students.

Professor **Junfeng Yang** was chosen for a NSF Career Award for "Making Threads More Deterministic by Memoizing Schedules".

Multithreaded programs are becoming increasingly critical driven by the rise of multicore hardware and the coming storm of cloud computing. Unfortunately, these programs remain difficult to write, test, and debug. A key reason for this difficulty is nondeterminism: different runs of a multithreaded program may show different behaviors depending on how the threads interleave. Nondeterminism complicates almost every development step of multithreaded programs. For instance, it weakens testing because the schedules tested may not be the ones run in the field; it complicates debugging because reproducing a buggy schedule is hard.

In the past three decades, researchers have developed many techniques to address nondeterminism. Despite these efforts, it remains an open challenge to achieve both efficiency and determinism for general multithreaded programs on commodity multiprocessors.

This project aims to address this fundamental challenge. Its key insight is that one can reuse a small number of schedules to process a large number of inputs. Based on this insight, it takes an approach called schedule memoization that memoizes past schedules and, when possible, reuses them for future runs. This approach amortizes the high overhead of making one schedule deterministic over many reuses and makes a program repeat familiar behaviors whenever possible. A real-world analogy to this approach is animals' natural tendencies to follow familiar routes to avoid hazards and discovery overhead of unknown routes.

The greatest impact of this project will be a novel approach and new, effective systems and technologies to improving software reliability, thus benefiting every business, government, and individual.

Professor **Junfeng Yang** and his students got two papers accepted to the 2010 Symposium on Operating Systems Design and Implementation (OSDI 2010), which is one of the top two operating systems conferences. Out of a few hundred submissions, only 32 papers were accepted! The two Columbia papers will be presented next month in Vancouver.

Columbia University in the City of New York
Department of Computer Science
1214 Amsterdam Avenue
Mailcode: 0401
New York, NY 10027-7003

ADDRESS SERVICE REQUESTED

NON-PROFIT ORG.
U.S. POSTAGE
PAID
NEW YORK, NY
PERMIT 3593

Extra News

Kasparov and Thiel Visit CUCS

Chess grandmaster **Garry Kasparov** and PayPal co-founder **Peter Thiel** visited the Columbia Robotics lab on November 4. Kasparov and Thiel were filming an episode of the award winning German-French television series "Into the Night with..." In each episode, two artists meet and spend one night exploring a city. Kasparov and Thiel are both very interested in technology, and began their exploration of New York City at night with a tour of the Robotics Lab where they discussed man vs. machines and Artificial Intelligence.