CS@CU

NEWSLETTER OF THE DEPARTMENT OF COMPUTER SCIENCE AT COLUMBIA UNIVERSITY VOL.5 NO.2 SPRING 2009

Outgoing Computer Science Department Chair Henning Schulzrinne

An Interview with Computer Science Department Chair **Henning Schulzrinne**

After serving as the Computer Science Department Chair for 5 1/2 years, Professor **Henning Schulzrinne** is stepping down at the end of his term in July.

T.C. Chang Professor Shree Nayar will be the next chair of the department. Professor Schulzrinne gave CS@CU Newsletter Roving Reporter Carlos-Rene Perez some candid insights about his experiences and perspectives as Department Chair.

How did you get into the role of department chair?

I think the way it works in most departments is that some of the senior faculty put their heads together and decide who can we draft, who might be willing, who probably won't do too much damage, etc. For some reason some of the senior faculty picked me, so I was invited to a breakfast at one of the local eateries and it was sprung on me that some people think that "you should be next." Serving as chair was something that I hadn't necessarily planned on at that particular moment, but I decided it was a challenge I was interested in, and I was willing to take it on.

Were there any surprises on your first day of work?

There weren't too many surprises on the first day, but eventually there were surprises in terms of things that you don't really see as a normal faculty member. Usually it's kind of a black box "paper goes in, signature comes out" type of model; what you don't see is all the activities that go on on a daily basis. Or as a student, you probably think of the front office as a group of people that sit there all day and you don't really know what they do—you hope they do something useful! You might only bother them 10 minutes during the year, but of course there are 120 other Ph.D. students, 30 other faculty, and many more Masters students, undergraduates, etc—all who have needs that need to be taken care of.

What I found out is that the department chair tends to be the last stop, at least within the department, for just about everything. One of the things that you notice very quickly is that no issue is too small to make it to the email inbox of the chair—missing toilet paper, somebody ordered the wrong pencils, somebody can't find their mail—you name it, sooner or later you'll get a e-mail with somebody stopping by. That's not a bad thing, because ultimately most organizations only work if somebody takes

responsibility for everything in the end. You want people to feel that there is a person who eventually will listen and hopefully make the right things happen. Of course, as chair you usually don't do it yourself, you end up delegating. So chair management is often what I call management by forwarding; as chair you try to decide who is best qualified to deal with whatever idea or suggestion or complaint or problem you're confronted with. Speaking as a networking person, you act kind of like the central router: you provide a well known address to everybody in the department, and hopefully in many cases you can provide some assistance.

What is your proudest accomplishment as chair?

One thing you learn as a chair, hopefully quickly, is that the Oath of Hippocrates applies: First do no harm. In any department the really important work that the University does—teaching, research, service—gets done by the faculty and the Ph.D. students and to some extent by the other students in the department and the research staff. So what a department chair can do and staff can do is really make sure that those who do this real work of the University have the resources they need so that they can do their work unencumbered by bureaucracy, by silly walls, by resource limitations (to the extent possible).

One of the issues that I discovered relatively early is that the department had grown significantly since I joined in 1996. Back then we were probably roughly half the size that we are now, a little larger, and many of our ways of doing things were "small department" ways. For example, we had very little electronic support, computing support for our administrative functions, everything was done on spreadsheets, word documents in file stores and so on. This worked OK for a small department, but didn't scale very well to a $10 million dollar research, 30+ faculty department. So one of the things that I definitely wanted to accomplish was having better infrastructure that allows the faculty and staff and student volunteers to run things a little bit more efficiently.

What are some goals that you feel still need more work?

I do believe that there are two areas where the department can do better. One is a byproduct of success: I think we have moved up in our external ranking and visibility, and so we need to start "acting our age." I don't mean our biological age as a 25+ year department, but we have now acquired a bit of external reputation. We used to be kind of a "mid-20's" type of department, but now we are easily a top-20 department by our impact. But once you reach that level, you are suddenly in the same league as departments which have a number of activities, in terms of outreach and interaction with the community, that we haven't quite reached. Working with alumni, working with industry, working with the rest of the university—more is expected of you when you're a more visible department.

The second area is a long term challenge. The way we are set up physically, we are in two—if you count the InterChurch group, three—buildings, and so we are very distributed. It's quite possible to graduate, if you were in CEPSR, having never set foot into the CS building, except maybe on graduation day. This means you may not have met many of the people whose offices are in the CS building. So facilitating contact between students is an ongoing goal: between students in CEPSR and students who are in the CS building, between students in, say, the theory group or the graphics group and students who are more on the systems side. This is a challenge.

I think the staff here, particularly the academic staff, has done a much better job than we used to do at integrating things through social activities. I've set up a number of social events, like the family day and the graduation ceremony that we do for Ph.D. students, to help a little bit more with this; also we now have our department-wide coffee hours pretty regularly. But I think we can still do better.

Hopefully this department is a community—a community that has common technical and intellectual interests, but also a social community. Most people spend an inordinate fraction of their week here in these buildings; many people probably spend more time with their colleagues here than they spend with their their children and spouse, besides sleeping. So treating it simply as a place with a bit of concrete and bricks where you get a network connection and a desk doesn't quite cut it; I think we need to treat it as a place where people can connect socially as human beings. So having people that volunteer here, be it for a newsletter or be it for cookies or be it for being a big brother, is very much an important part of that. I very much try to encourage activities along these lines; that's why we have new awards for students who do community service.

How do you balance all the things that a department chair has to do?

Well, I'm not claiming that my schedule is any worse than any other faculty, but certainly the number of hats that a department chair has to wear makes Imelda Marcos's shoe collection seem small. You have to time-slice to an extent that I didn't experience before. Today is a good example: you go from a meeting at the school level with other chairs, to a meeting with an external research sponsor, to a meeting with a department administrator, to meeting with a Ph.D. student, followed by another meeting with the school on some budget issues. Even during a single day, you end up constantly switching between different roles.

Do you have any tips for your successor?

If I were to do this again, I think it's important to set expectations appropriately. I think I have quite a reputation for answering email relatively quickly at all hours, and people have started to expect that. That is not an expectation which you necessarily want to foster! So it's a good idea to set clear expectations: if you do not get an answer to an e-mail until Monday morning, you're going to survive, and you don't have to panic as to why you didn't get an answer to your email on Saturday. The other piece of advice is to take advantage of our great staff. We are lucky now to have a very strong office staff on all sides—business, financial, academic—that can take on many roles. It's a situation where if we can grow the professionalism and responsibility of our staff even further, I think they will step up to the plate. This means many things can be done well that otherwise may not get done at all, if the chair tries to do it all.

How about faculty recruitment? Could you see the department growing by ten more faculty?

Every department wants to grow, and growth is a good way to do new things and do existing things better. But simply relying on faculty growth is not the only way we can be doing things better.

We've reached a size that makes us now one of the largest computer science departments in the Ivy League, certainly, and we are probably mid-sized for general computer science departments. So we are now competitive in terms of size and we have pretty good coverage of areas. I would certainly like us to see growth and I would hope that

the current financial situation improves to such an extent, from a University perspective, that the school allows us to grow to some extent. But I think we also want to make the best use of the faculty assets that we already have. In the last few years we've hired extremely strong junior faculty, and as they grow they become more senior. We have strengths in areas that we never used to have strength in. This means we can now do things that we were never able to do—instead of everybody doing things by themselves, we can operate in larger scales, with larger groups doing large proposals and centers and other efforts that are just more visible than five little efforts. I think this is how good departments and very highly ranked departments operate.

What are you going to miss about being department chair?

Well, one thing I've missed while being chair is the opportunity to do more intense research effort that requires more hands-on work. Right now I've been doing a lot through the Ph.D. student group that I have, but maybe I haven't always been able to provide as much day-to-day supervision or interaction with them as I've had before. I also hope that I can continue to do work on the external professional side; I have a fair amount of ACM involvement now through ACM SIGCOMM and other parties, and that is something that I think is an important part of being a somewhat more senior faculty. You can help shape the general direction of ideas—in my case, a general direction for the network research community.

In closing, do you have any suggestions for how the student community can help out in outreach or other ways?

We have a great student community at both the undergraduate and graduate levels. I hope we can draw on their resources, both reaching out to their alma maters—wherever they came from—and staying in touch once they leave the University. I'm not necessarily talking about a $20 million donation for a new building (though that would be highly welcome), but simply staying in touch, providing us with input, providing us with contacts, providing us with a means of engagement with companies and other entities. This is something where I think alumni can certainly help, and students can help with outreach to wherever they came from earlier. Students are our life blood; we wouldn't get any research done without them. And attracting good students to Columbia is probably the best way to make Columbia Computer Science even better. It helps everybody: students get good colleagues to work with, we all get good research done, and the department over time will become even better. We've already seen this when we have the Ph.D. visit day—many students already help out with this, and it has been tremendously helpful in attracting the best students to come here.

Professor
& Chair
Henning Schulzrinne

APAM Ph.D. student Tianzhi Yang practices a presentation.

Academic Writing & Great Presentations Workshops for CS Grad Students

Starting in 2007, the Department of Computer Science invited Janet Kayfetz to teach writing and presentation skills to our graduate students in short courses taking place during the early weeks of the semester.

These short courses were built on her experience teaching writing and presentation skills at UC Santa Barbara. We asked her to reflect on her experiences.

If you walked through the CS lounge at some point during the first three weeks of the fall semester, you may have been there when it was transformed temporarily into a writers' space. For three weeks in September 2008, 45 engineering doctoral students from Computer Science, Biomedical Engineering, Electrical Engineering, Earth and Environmental Engineering, Applied Physics/Applied Math, IEOR, and Chemical Engineering participated in an Academic Writing Workshop organized by the Department of Computer Science. Students met twice a week for 2 hours each meeting in one of three workshop sections.

The goals of the writing workshop were threefold: to create a discourse community of scholarly student writers who would share a commitment to excellence in writing, to apply basic skills in composing and editing to produce clear, logical, precise, and readable scientific texts, and to practice techniques that would help student writers produce more—and qualitatively better—writing, either for conference and journal papers or for their dissertations. Students analyzed the elements of a scientific paper, including content, form, and language, with particular attention given to reader-oriented writing and clarity. While a three week session goes by quickly, we were able to focus on a number of important genres, including introductions, data commentaries, results and discussions, and abstracts.

During class, we converted the CS lounge into a digital classroom which allowed us to access the workshop website, choose a student document, project it on the screen in front of the group, edit the text in real time, and then email the edited text to the student for further review. In addition to the details of language use and clarity of written expression, we turned our attention to the verbal articulation of ideas and the appropriate and helpful ways to critique peer writing. Students worked in pairs on the peer editing of their texts, and also on timed, unedited, journal-type writing that builds fluency, speed, and confidence.

The Academic Writing Workshops were first offered to CS Ph.D. students in September 2007. The 2008 sessions were a continuation of our 2007 efforts, with the added feature of opening the workshops to Ph.D. students from other departments in the School of Engineering and Applied Science. As interest in the first semester 2008 sessions seemed to be strong, the Department organized a second round of workshops this past January-February 2009—with a new twist. Instead of a repeat offering of the writing workshops, the most recent sessions offered an Advanced Writing Workshop for students who had attended one of the previous writing sessions, as well as two Great Presentations Workshops, a response to student requests for help with presentation skills.

The advanced writing group worked only on their own writing-in-progress, documents that varied from conference papers and journal articles to technical reports. As students were already familiar with the approaches introduced in the first writing workshop, we had a shared view of the importance of audience and the significance of developing a precise description of the research space and the specific problem to be discussed in the text. We were able to take things to the next level and look at the subtleties of the construction of logical arguments, the idea of consistency of tone, and the notion of rhythm in a text, using the digital classroom approach for group analysis and editing.

The Great Presentations Workshops were a pleasant change of focus, and allowed us to dive into the world of talking about research to an audience of specialists. We used the acronym "STORIES" to highlight the intertwined variables of excellent presentations. We talked about the most important part of a presentation, the development of a clear and logical story (S), as well as the elements of timing (T); organization (O); rehearsal (R) and practicing aloud; interacting (I) with both the audience as well as with the visuals if appropriate; the details of verbal expression (E) including volume, pitch, intonation, choppiness and fluency; and the development of clear and useful visuals and slides (S). Students gave presentations of varying lengths which were videotaped and critiqued by me, and all of the presentation clips were posted on a site where students could review them as often as they liked. In addition to the formal presentation, we also spent time talking about job talks and interviews, and the art of asking and responding to questions.

A huge challenge faced by the workshop participants who are not native speakers of English is that in addition to the central issues of argument development, organization, flow, word choice, precision, and so on, nonnative speakers have the language acquisition challenge and difficulty as well. While the workshops were not conducted with any specific orientation to second language issues, I was able to incorporate work with the nonnative speakers on specific language patterns, pronunciation, and other areas of difficulty.

I want to say that it was a wonderful experience to have the chance to work with Columbia students. From the very first minute, students were free with their observations and comments and were ready to participate. It is no easy thing to create a cooperative group of writers and speakers with a shared commitment to being open and vulnerable about their writing efforts and presentations. All of us know that it is not easy to take in the questions and accept the suggestions that come our way when we throw our text on a screen in front of a group of sharp thinkers or ask "What did you think of my presentation?" Likewise, many students, even those who are skilled at analyzing a text and understand how it might be improved, are not comfortable with the process of offering feedback. But the workshop students did not hold back. They reached a level of analytical ability in three weeks that usually takes much longer to achieve, and I watched them learn how to be willing to receive some very intense and detailed suggestions from group members, myself included.

We hope these workshop experiences take the students a long way to being better collaborators with their professors and their student colleagues, and that their writing and presentations reflect a new level in their communication expertise. I want to extend my thanks to Swapneel Sheth for creating the website for the Great Presentations groups and for coordinating the uploading onto the site of all of the video clips of student presentations over the three-week sessions.

Computer Science Ph.D. students Arezu Moghadam (L) and Ariel Elbaz (R) at work on peer editing.

Computer Science Ph.D. student Ryan Overbeck (L) and APAM Ph.D. student Abby Shaw-Krauss (R) discuss a writing sample.

Interview with New Faculty Member **Martha Kim**

Martha Kim joined the Computer Science Department as an Assistant Professor in the Spring 2009 semester after completing her Ph.D. at the University of Washington.

CS@CU Newsletter Roving Reporter Carlos-Rene Perez caught up with Martha and talked to her about spatial architectures, legacy code, and more.

Tell us about your field of research and why you got into it.

Computer architecture is a broad and cutting-edge field at the moment. It encompasses all manner of technical problem from the device level right up to programming systems. This breadth is what makes it so interesting and extremely rich.

For many years, architects were able to reap performance by scaling processors to higher and higher clock frequencies while leaving the essentials of the design unchanged. For a number of reasons, this old technique will no longer work, so the field is faced with a host of challenges, which, quite frankly, we do not know how to solve. While the problems are quite real and quite difficult, this makes it an exciting time to be in architecture.

This is part of the reason I decided to pursue research in the field. There are many unknowns and unsolved problems.

What type of research interests you?

Problems that are at the boundary of architecture hold a special interest to me. At the software boundary there are many problems relating to how the programming systems interact with the machine on which they are running. How can architects harness the raw computational resources afforded by Moore's Law to help existing and emerging applications run as efficiently as possible? At the device-level boundary, questions arise as to how to adapt architectures and microarchitectures to changing device characteristics. My research has focused on those boundaries.

I have explored algorithms for mapping software to highly distributed or decentralized architectural designs. These architectures consist of many, many processing cores. The challenge lies in mapping code (typically seen as serial) onto a highly parallel computational substrate.

In 2006 you delved into spatial architectures, specifically the Wavescalar architecture. Can you guide us through this?

WaveScalar* is a research project at the University of Washington that I contributed to as a graduate student. The project explored several aspects of spatial architectures and pushed the limits of very small, simple processing cores in a design where communication is a first order citizen.

How do you build this architecture? How would you program such a machine? The design ended up as dataflow machine, which is a different computational paradigm from what is traditionally used. Dataflow has the benefit of exposing a lot of instruction level parallelism.

What is a dataflow machine?

Dataflow is an alternative to von Neumann computation. Von Neumann computation is the more familiar model of computation as a linear set of instructions executed in the order dictated by a program counter. In dataflow computation, the program is represented as a graph, where the nodes of the graph represent instructions, and the edges connecting the nodes represent data dependencies, producer-consumer relationships, between instructions. Under this model, an instruction can be executed at any point once all of its input operands have become available. This is called the dataflow firing rule, and takes the role of the program counter dictating when instructions can execute. Because the dataflow model of computation does not introduce artificial or unnecessary serialization, it exposes all theoretical instruction level parallelism.

Going back to Wavescalar, what did this architecture bring to the field?

One significant contribution of the WaveScalar project was that, unlike previous dataflow machines, it could be programmed in traditional, well-known languages. Previous dataflow machines had to be programmed with special, dataflow languages. This was due to differences between the language and the machines regarding the order in which memory operations were to be applied to memory.

Another significant research contribution is in the micro-architectural space. Traditional processor designs were designed under the assumption that data could be sent anywhere on a chip within a single clock cycle. Over the years, as clocks have become faster and faster, that has ceased to be true. In some sense, other regions of the chip have become more remote because of this. At the microarchitectural level, WaveScalar was focused on these non-uniform communication characteristics from the start. We acknowledged that some parts of the chip were closer than others and designed the microarchitecture accordingly.

Out of this came several innovations in the design and study of distributed micro-architectural components. These results apply to less aggressively parallel chip designs as well. A set of processing cores of any size has this problem. For example, you can consider a quad-core multiprocessor as a distributed architecture with just four locations.

This sounds similar to Intel's Larrabee, so can you make a quick comparison?

Sure. One way to view these architectures is sitting on a continuum. At one end you've got a single-core processor. Then you can apply Moore's law, where you say, in the same sized footprint I'm going to put two half-sized cores, or four quarter-sized cores. Larrabee is out in this direction. You can keep pushing towards more and more, smaller and smaller cores. At the end of the spectrum you run into so-called tiled architectures, of which WaveScalar is one. The cores on these machines are small, lightweight computation engines, but there are many more of them.

Comparing WaveScalar and Larrabee directly, I see them as heading in the same direction but with Larrabee a little closer to what we have today. WaveScalar is unlike anything we have today, but if we take the trend that produced Larrabee to its extreme, we might find something that looks like WaveScalar.

We are unquestionably headed in the direction of multicore. I don't know how to take a processor design and make it twice as fast, but I can very readily shrink it and give you two of them. In doing this we are foisting the problem onto the programmers, leaving it up to them to extract improved performance out of those two parallel cores. I think we are headed in that direction because it's the only thing we know how to do. If we succeed in working with the programming and software system community, we may be able to produce parallel codes that can profit from all these cores. I think also, in the face of power constraints, we really have no other choice that will keep the power consumption of these chips within our budgets. So I do see multi-core as being the future in the next 5-20 years.

What are your thoughts on legacy support? Should we scrap the X86 ISA?

The need to support legacy code was always a funny one to me. People in the hardware community often think of software as being much easier to write and much more changeable than hardware. In software, for example, you can patch your code and you don't need to commit implementations to lithographic masks. But the reality is that software is even more complex. There is much more of it, partially because it's so easy to write. While supporting legacy code does create an enormous burden, the consensus seems to be that re-writing it would be an even greater burden. So we choose to continue to provide that support. Personally I believe that there is too much code out there that we depend on too much. There is no practical way that it can all go away or be rewritten.

Can you tell us more about your recent research on circuits?

My focus for the last three years has been very close to the physical device level. I was motivated by the fact that it is very expensive to fabricate a chip nowadays. In order to design something such as Larrabee, with billions of extremely tiny transistors, there is an enormous engineering cost: a lot of time and money goes into producing that design, validating and verifying that it does what it is supposed to. This takes a lot of time and engineer effort. Then, once you're convinced that it works and is ready to go, you still spend a couple of million dollars for lithographic masks, which are essentially templates that are used to physically fabricate the circuit.

So it's an enormously expensive process, in time and money, before you ever even produce a chip. If you are somebody such as Intel, building Larrabee, you've got a lot of engineers and a lot of capital. You can afford to spend that time and money, because you've got reasonable confidence that you'll sell a fairly large number of these chips. The problem is, if I need only a couple of dozen of my own design, or even a couple thousand, I really can't afford to spend those tens of millions of dollars. So, my research was focused on trying to reduce the initial cost of designing a circuit. The goal was to allow people to actually fabricate hardware in scenarios where today it is not economically feasible to do so.

The way I did this was to observe that a lot of chips are built out of the same underlying components. The example I like to use is a camera phone. If I were building your iPhone I would not implement the JPEG encoder from scratch. Other people have implemented the JPEG encoder and I could save time and effort by licensing that encoder for use in my design. The essential idea behind my research, named Brick and Mortar, is reuse—to share not only the design effort (as is done now) but also the physical implementation effort. The proposal is to prefabricate a large volume of small chips containing commonly used hardware modules. These are called bricks. For example, I'd build a JPEG encoder brick, and a micro-processor brick, and and an MPEG codec brick, and so on. I'd prefabricate these system-on-chip-type components, so that all different chip designers need to do is buy these mass-produced components off the shelf at relatively low cost to them and then physically assemble them into the particular hardware system they need. This gets around the high initial cost by sharing those high initial costs across many different custom designs. The questions on the technical side are, how do we make these chips perform reasonably well? How do we size the components? What sort of communication infrastructure should we provide for them to communicate between bricks?

Imagine a multiprocessor system, in which you could go to Intel's website in the same way that you go to Dell's website today and customize your chip. You could request the particular mix of processing cores you want, the particular size and organization of the caches, and so on. You could apply brick and mortar techniques to multi-processors, as an example domain, to make it possible to have made-to-order processors.

* The WaveScalar Architecture, S. Swanson, M. Mercaldi, A. Petersen, A. Putnam, A. Schwerin, K. Michelson, M. Oskin, S. Eggers. ACM Transactions on Computer Systems (TOCS). Volume 25 Issue 2. May 2007.

Application Communities


Yell if you see a buffer overflow coming this way!

Professor **Angelos Keromytis**

Software monocultures have been identified as a major source of problems in today's networked computing environments. Monocultures act as force amplifiers for attackers, allowing them to exploit the same vulnerability across thousands or millions of instances of the same application across the network. Such attacks have the potential to rapidly cause widespread disruption, as evidenced by several incidents over the last few years. The severity of the problem has fueled research behind introducing diversity in software systems. However, creating a large enough number of truly different systems from scratch not only presents practical challenges but can result in systems that are not diverse enough.

As a result, recent research has focused on creating artificial diversity, by introducing "controlled uncertainty" in one of the system parameters that the attacker must know (and control) in order to carry out a successful attack. Such parameters include, but are not limited to, the instruction set, the high-level implementation, the memory layout, the operating system interface and others, with varying levels of success. However, running different systems in a network creates its own set of problems involving configuration, management, and certification of each new platform. In certain cases, running such multiplatform environments can even decrease the overall security of the network.

Given the difficulties associated with artificial diversity and the pervasive nature of homogeneous software systems, the Network Security Laboratory has been working on software self-healing techniques whereby a large, homogeneous software base can be used to improve security and reliability for the whole community of participating users. To that end, we introduced the concept of an Application Community (AC), a collection of identical instances of the same application running autonomously across a network. Members of an AC collaborate in identifying previously unknown (zero day) flaws/attacks and exchange information so that such failures are prevented from re-occurring. Individual members may succumb to new flaws; however, over time the AC should converge to a state of immunity against that specific fault. The system learns new faults and adapts to them, exploiting the size of an AC to achieve both coverage (in detecting faults) and fairness (in minimizing the amount of additional work at each member).

At the core of the system lies our work on selective transaction emulation (STEM). STEM offers a new model for partial emulation of selected segments of legacy applications, in order to detect divergences from expected behavior (e.g., compared to another instance of the application) or fault conditions (e.g., buffer overflow, illegal memory reference). STEM can be used both for data collection/fault monitoring by each member of the AC, and for fault mitigation once the group identifies a new fault. Members of the AC emulate different "slices" of the application, monitoring for low-level failures (such as buffer overflows, illegal memory accesses) as well as deviation from normal application behavior (e.g., unusual function arguments or function call sequences). This same low-level information produced by "normal" runs (those that are not deemed anomalous) is correlated with information from other instances of the application, to build behavior models. When a fault or deviation from normal behavior is detected by a member, the relevant information is broadcast to the rest of the AC. Members may verify the fault and apply STEM on the identified vulnerable code slices, possibly combining this with input filtering or other mitigation techniques. Because of the use of STEM, it is possible to wrap the necessary functionality around existing applications, without requiring source code modifications. In fault-mitigation

mode, STEM operates transactionally at function-call granularity, i.e., it can rollback the current execution, and then force the application to take a different execution path. We identify such paths by using the models of application behavior built from "normal" runs. An alternative mechanism for implementing this rollback functionality is based on the process checkpointing work done at the Network Computing Laboratory.

Application Communities offer a new paradigm in software security and reliability, by exploiting the "wisdom of the masses" to conduct continuous, collaborative fault and attack detection and mitigation. In an environment where attackers routinely control and use thousands of compromised systems to attack users and networks, the natural response is for the defense to pool its resources and information in order to be effective. This work is currently underway under support from the U.S. Air Force and the National Science Foundation, with several patents filed.

New Computer Science Courses

Sebastian Lahaie and **Sergei Vassilvitskii** taught a new course, COMS 6998-3: **Introduction to Algorithmic Game Theory**, in the Fall 2008 semester. Algorithmic game theory is an emerging area at the intersection of computer science and microeconomics. Motivated by the rise of the internet and electronic commerce, computer scientists have turned to models where problem inputs are held by distributed, selfish agents (as opposed to the classical model where the inputs are chosen adversarially). This new perspective leads to a host of fascinating questions on the interplay between computation and incentives.

This course provides a broad survey of topics in algorithmic game theory, such as: algorithmic mechanism design; combinatorial and competitive auctions; congestion and potential games; computation of equilibria; network games and selfish routing; and sponsored search. No prior knowledge of game theory is necessary; the most important prerequisite is mathematical maturity.

Professor **Jason Nieh**

Professor **Jason Nieh** taught a new course, COMS E6998: **Mobile Computing with Iphone and Android**, in the Spring 2009 semester. The course is an intensive study of mobile computing on smartphones with an emphasis on applications. These handheld Internet devices are poised to become the future dominant software platform as a result of the rapid convergence of computers and mobile phones. Topics covered include mobile operating systems and development environments, input modalities and user interfaces for mobile devices, power management issues for mobile devices, wireless mobile networking, thin clients and mobile Web, location-aware and other context-aware services, and virtualization. A course programming project is required. The course was mentioned in a *Forbes.com* piece about mobile computing courses around the country, and had a Spring 2009 enrollment of more than 70 students.

Professor **Kenneth Ross**

Professor **Kenneth Ross** taught a new course, COMS 4112: **Database Systems Implementation**, in the Spring 2009 semester. This new course focuses on how to build a database engine. Topics covered include storage structures, query processing and optimization, concurrency control methods, recovery methods, and parallel and distributed databases. At the same time, the introductory database systems course (4111) has been revamped, focusing on how to effectively use a database system, and including new material on object-relational databases and XML.

Professor **Rocco Servedio**

Professor **Rocco Servedio** taught a new course, COMS 6998: **Advanced Topics in Computational Complexity**, in the Spring 2009 semester. The course is an intensive study of concrete lower bounds for various computational problems in simple models of computation such as decision trees, Boolean formulas, restricted classes of Boolean circuits, and the like. In this line of research unconditional lower bounds are established which rely on no unproven

assumptions such as P vs NP. There has been steady progress made over the years using a range of techniques from combinatorics, algebra, analysis, and other branches of mathematics.

The course gives self-contained proofs of a wide range of unconditional lower bounds for interesting and important models of computation, covering many of the "gems" of the field that have been discovered over the past several decades, right up to results from the last year or two. Students present papers and work on a research project as the main activities in the class.

Professor **Simha Sethumadhavan**

Professor **Simha Sethumadhavan** taught a new course, COMS 6998-1: **Advanced Computer Architecture: Parallel Systems and Programming**, in the Fall 2008 semester.

The predominant question for the Computer Architecture community, and arguably for all of the CS community, is how do we exploit the highly parallel hardware that is available to us today. Researchers have proposed pure software solutions, pure hardware solutions and some hardware/software solutions, but there is no silver bullet. Yet, the future IT successes completely depend on the ability to extract performance from parallel hardware. This class teaches students techniques that may help them tackle this challenge.

The class is both for students and practitioners and covers machine organization and design of parallel systems, the three dominant parallel programming models, performance analysis and optimizations. Students also read and analyze recent research on parallel systems.

Professor **Junfeng Yang**

Professor **Junfeng Yang** taught a new course, COMS 6998-2: **Topics in Computer Science: How to Make Reliable Software**, in the Fall 2008 semester. Despite our increasing reliance on computing platforms, making reliable software systems remains difficult. Software errors have been reported to take lives and cost billions of dollars annually. Making reliable software is one of the most important problems in computer science. In recent years, this problem has drawn huge attention from researchers in systems, software engineering, and programming language communities. A number of automated techniques have been developed to increase software reliability.

This course covers the most practical and most important of these reliability techniques. Specifically, students in the course study:

- **Practical bug-finding techniques.** Students learn effective techniques that have found thousands of serious errors in large systems, as large as an entire Linux kernel.
- **Fault isolation and recovery.** Students learn techniques to prevent a single component failure from crashing an entire system and to completely recover from such component failures.
- **Automated debugging.** Debugging (i.e. finding the root cause of a bug) is usually a painful manual process; students learn techniques to make debugging more automatic and less painful.
- **Concurrency.** CPUs are getting more cores; the implication is that multi-threaded programs will be the mainstream. To better understand the errors in multi-threaded programs (e.g. races and deadlocks), students study their characteristics, and will also study effective techniques to find these concurrency errors.
- **Other reliability techniques.** For example, students learn the underlying mechanisms of a popular dynamic tool Valgrind and a commercial virtual machine VMware.

In addition, students have the opportunity to hear guest speakers who have built large software systems or effective checking tools talking about their first-hand experiences.

2008-9 **Distinguished** Lecture Series

The Computer Science Department enjoyed another successful Distinguished Lecture Series during the 2008-9 academic year. This yearly series brings world-renowned scientists and pioneers in academia and industry from across the country to Columbia's campus.

These distinguished speakers give talks about their research, field questions from the audience, and meet with faculty and students. The lectures are open to the public and announced and documented at our website *(http://www.cs.columbia.edu/lectures)*, and ensure that members of the department and the community at large see first-hand the latest and greatest computer science research emerging outside the Columbia campus. This year, six lecturers visited us and discussed a broad and exciting range of topics.

These speakers helped make this year an exciting and inspirational one for computer science at Columbia. Stay tuned next year, when more of the world's greatest researchers will visit our department and tell us about their work.

Dr. Peter G. Neumann of SRI International visited on October 6 and spoke about **Integrity of Elections.**

Dr. Neumann is Principal Scientist at SRI, where he has been since 1971; his research deals with computer systems and networks, trustworthiness with respect to security, reliability, survivability, and safety, and risks-related issues such as voting-system integrity, crypto policy, social implications, and privacy.

Elections demand end-to-end integrity of voting processes, with additional trustworthiness requirements such as system security, privacy, usability, and accessibility. The overall system aspects present a paradigmatic hard problem. In today's systems and procedures, essentially everything is a potential weak link. The pervasive nature of the risks is astounding, with a serious lack of system architecture, good software engineering practice, and understanding of security problems. Dr. Neumann's talk discussed limitations in existing systems, processes, standards, and evaluations. It also considered some possible alternatives—including nontechnological approaches, computer-based systems, and possible roles for cryptography.

Professor **Herbert Edelsbrunner** of Duke University visited on November 10 and spoke about **Measuring Shape Before Simplification.**

Professor Edelsbrunner is also Founder and Principal at Geomagic, a software company in the field of Digital Shape Sampling and Processing. He is a recipient of the Alan T. Waterman Award from the National Science Foundation, a member of the American Academy of Arts and Sciences, and a member of the German Academy of Sciences, the Leopoldina. Professor Edelsbrunner specializes in the combination of computing and advanced mathematics to solve problems in applications. His methodology is to search out the mathematical roots of application problems and to combine mathematical with computational structure to get working solutions.

Nature is inherently multi-scalar; Professor Edelsbrunner's talk presented an attempt to measure this aspect mathematically. Rooted in algebraic topology, this idea has ramifications inside and outside mathematics. A particularly important application is coping with noise in scientific data. As suggested by the title, the talk advocated measuring noise, but not necessarily removing it from the data, because doing so has side-effects.

Professor **Alberto Sangiovanni-Vincentelli** of the University of California, Berkeley, visited on November 19 and spoke about **Quo Vadis System Design?**

Professor Sangiovanni-Vincentelli holds the Buttner Chair of Electrical Engineering and Computer Sciences at Berkeley. He was a cofounder of Cadence and Synopsys, the two leading companies in the area of electronic design automation. He is the recipient of numerous awards, an author of more than 800 papers and 15 books in the area of design tools and methodologies, a Fellow of the IEEE, and a member of the National Academy of Engineering. In 1995 Professor Sangiovanni-Vincentelli received the worldwide Graduate Teaching Award of the IEEE for "inspirational teaching of graduate students."

The electronics industry ecosystem is undergoing a radical change driven by an emerging three-layered architecture characterized by:

- Computing and communication infrastructure that will offer increasingly faster data transfer and manipulation via powerful data centers, compute farms and wired interconnection;
- Access devices such as PDAs, cell phones, and laptops, which allow leveraging the immense capabilities of the infrastructure to users that can be humans or intelligent physical systems;
- A swarm of sensors, actuators and local computing capabilities "immersed in all kinds of physical systems that offer a wide variety of personal or broad-use services."

Most refer to these swarms as embedded systems. Recently there has been a growing interest in Cyber Physical Systems (CPS) where the interaction between the computing and electronic elements with the physical systems they are immersed into is emphasized. CPS will allow developing a wide span of applications because of the availability of a new generation of sensors, actuators, and local computing that leverage novel interconnect capabilities and centralized computation.

Dealing with system-level problems requires more than simply developing new tools, although of course they are essential to advancing the state of the art in design. Rather, the focus must be on understanding the principles of system design, the necessary changes to design methodologies, and the dynamics of the supply chain. Developing this understanding is necessary to define a sound approach that meets the needs of the system and component industries as they try to serve their customers better and develop their products more quickly and with higher quality.

Dr. Sangiovanni-Vincentelli's talk presented directions, challenges, and potential solutions to the design of future systems, for which heterogeneous subsystems such as mechanical and electrical components must be designed concurrently. The possible scenarios pose fundamental questions to the engineering and scientific worlds regarding how to deal with the design and management of global systems with such huge complexity. A unified design methodology that can extend from cyber physical systems (CPS) all the way down to chips, boards, and mechanical components with general environments capable of hosting specific design flows for the industry segments is the ultimate enabling technology. The talk presented a potential approach to such a unified design methodology, called platform-based design (PBD), and some examples of its use.

Professor **Bonnie John** of Carnegie Mellon University visited on December 1 and spoke about **Cognitive Crash Test Dummies: Where We Are and Where We're Going.**

Professor John has more than 25 years experience in usability analysis and design. She heads the Masters Program in Human-Computer Interaction at Carnegie Mellon, where she researches both human performance modeling and software engineering and consults regularly in government and industry. Professor John is an ACM CHI Academy member, recognized for her contributions to HCI through her work in cognitive modeling and the implications of usability concerns on the design of software architecture.

Crash dummies in the auto industry save lives by testing the physical safety of automobiles before they are brought to market. "Cognitive crash dummies" save time, money, and potentially even lives, by allowing computer-based system designers to test their design ideas before implementing those ideas in products and processes. In her talk, Professor John reviewed the uses of cognitive models in system design and the current state of research and practice. She also presented some exciting new research directions that promise to make predictive human performance modeling even more useful. Along the way, her talk discussed the role of applications in science and validity versus useful approximation.

Professor **Gail Murphy** of the University of British Columbia visited on March 2 and spoke about **Attacking Information Overload in Software Development.**

Professor Murphy works primarily on building simpler and more effective tools to help developers manage software evolution tasks. She is a recipient of the UBC Killam Research Fellowship, the AITO Dahl-Nygaard Junior Prize, an NSERC Steacie Fellowship, the CRA-W Anita Borg Early Career Award, and the University of Washington College of Engineering Diamond Early Career Award. Professor Murphy is also a co-founder and CEO of Tasktop Technologies Inc.

The productivity of software developers is continually degrading due to an inundation of information: source code is easier and easier to traverse and to find, email inboxes are stuffed to capacity, RSS feeds provide a continual stream of technology updates, and so on. To enable software developers to work more effectively, software engineering researchers often deliver tools which produce even more information. The effect of more and more tools producing more and more information is placing developers into overload.

To combat this overload, Professor Murphy's research group has been building approaches rooted in structure and inspired by memory models. As an example, the Mylyn project packages and makes available the structure that emerges from how a programmer works in an episodic-memory inspired interface. Programmers working with Mylyn see only the information they need for a task and can recall past task information with a simple click. Professor Murphy's group has shown in a field study that Mylyn makes programmers more productive; the half a million programmers now using Mylyn seem to agree. In her talk, Professor Murphy described the overload faced by programmers today and discussed several approaches her group has developed to attack the problem, some of which may also pertain beyond the domain of software development.

Finally, Professor **Johanna Moore** of the University of Edinburgh visited on April 22 and spoke about **Language Generation for Spoken Dialogue Systems.**

Professor Moore is Director of the Human Communication Research Centre, and Deputy Head of the School of Informatics at the University of Edinburgh. She is a Fellow of the Royal Society of Edinburgh and of the British Computing Society, and has served as President of the Association for Computational Linguistics and Chair of the Cognitive Science Society. Prof. Moore's research brings together theory and practice from computer science, linguistics and cognitive science to inform the development of natural language technologies for use in a wide variety of applications, including spoken dialogue systems, recommender systems, and information retrieval from multimodal archives.

The goal of spoken dialogue systems (SDS) is to offer efficient and natural access to applications and services. A common task for SDS is to help users select a suitable option (e.g., flight, hotel, restaurant) from the set of options available. When the number of options is small, they can simply be presented sequentially. However, as the number of options increases, the system must have strategies for summarizing the options to enable the user to browse the option space.

In Professor Moore's talk, she evaluated two recent approaches to information presentation in SDS: (1) the Refiner approach (Polifroni et al., 2003) which generates summaries by clustering the options to maximize coverage of the domain, and (2) the user-model based summarize and refine (UMSR) approach (Demberg & Moore, 2006) which clusters options to maximize utility with respect to a user model, and uses linguistic devices (e.g., discourse cues, adverbials) to highlight the trade-offs among the presented items.

To evaluate these strategies, Professor Moore has gone beyond the typical "overhearer" evaluation methodology, in which participants read or listen to pre-prepared dialogues, which limits the evaluation criteria to users' perceptions (e.g., informativeness, ease of comprehension). Using a Wizard-of-Oz methodology to evaluate the approaches in an interactive setting, Professor Moore has shown that in addition to being preferred by users, the UMSR approach is superior to the Refiner approach in terms of both task success and dialogue efficiency, even when the user is performing a demanding secondary task. Finally, Professor Moore hypothesized that UMSR is more effective because it uses linguistic devices to highlight relations (e.g., trade-offs) between options and attributes, and described experimental results on users which supported this hypothesis.

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

Ph.D. student **Miklos Bergou** received an Autodesk Research Fellowship Award. Miklos is a Ph.D. candidate in the Columbia Computer Graphics Group, advised by Professor **Eitan Grinspun**. Miklos investigates the simulation and direction of physical systems, numerically modeling real-world objects and the interactions between them. His research seeks out principled and efficient discrete models that mirror the key geometric properties of the physical system. Miklos is also interested in developing intuitive tools that can be used to control the behavior of these systems, with applications in engineering and entertainment. His work on thin shell simulations is currently used by special-effects studios.

Ph.D. student **Rebecca Collins** received an IBM Ph.D. Fellowship. Rebecca began her Ph.D. studies at Columbia in the fall of 2005 and is advised by Professor **Luca Carloni**. Rebecca's research explores ways to harness the potential power of multi-core processor systems for general purpose programming. Rebecca has developed a tool that automatically generates parallel code which integrates distributed scheduling and adaptive memory management into SPMD-like threads running on the cores. Rebecca is also developing a method that combines static task deployment with dynamic runtime schedules to flexibly balance the computational load among unbalanced stream tasks for stream programs, and increase the overall processing throughput.

Ph.D. student **Charles Han** was awarded an ATI 2008-2009 Fellowship. ATI's highly selective panel awards between four and six fellowships each year to outstanding doctoral students studying a broad range of topics spanning computer graphics, multimedia, chip or system design, or related research. Charles is a Ph.D. student in the Computer Graphics Group, co-advised by Professors **Eitan Grinspun** and **Ravi Ramamoorthi**. His research focuses on finding principled representations and efficient algorithms that operate well across a wide range of visual scales. In work presented at SIGGRAPH 2007, Charles presented a solution to the long-standing problem of normal map filtering. By reinterpreting normal mapping in the frequency-domain as a convolution of geometry and BRDF, this work has enabled accurate multiscale rendering of normal maps at speeds orders of magnitude faster than previously possible.

Sahar Hasan, a junior in the Department of Computer Science, and **Ian Vo**, a senior, were selected for Honorable Mention in the 2008 Outstanding Undergraduate Award competition held by the Computing Research Association (CRA). The CRA honored a total of 22 female and 44 male students in this year's competition. Sahar Hasan worked with Professor **Gail Kaiser** on a project which was accepted for presentation at SIGCSE 2009, the annual conference of the ACM Special Interest Group on Computer Science Education. The paper is entitled *Retina: Helping Students and Instructors Based on Observed Programming Activities*. Ian Vo wrote a paper titled *Quality Assurance of Software Applications Using the In Vivo Testing Approach*, which was accepted for presentation at 2009 ICST, the 2nd IEEE International Conference on Software Testing, Verification and Validation.

Ph.D. student **Steve Henderson** received the Best Paper Award at ACM VRST 2008, held in Bordeaux, France. Steve is advised by Professor **Steven Feiner**. The award-winning paper, *Opportunistic controls: Leveraging natural affordances as tangible user interfaces for augmented reality*, was coauthored by Steve Henderson and Professor Feiner. It presents a class of interaction techniques, called opportunistic controls, in which naturally occurring physical artifacts in a task domain are used to provide input to a user interface through simple vision-based processing. Tactile feedback from an opportunistic control can make possible eyes-free interaction. For example, a ridged surface can be used as a slider or a spinning washer as a rotary pot.

Ph.D. student **Snehit Prabhu** was awarded a Microsoft Research and Live Labs Ph.D. Fellowship. Snehit is advised by Professor **Itsik Pe'er**; he transitioned from the MS to the Ph.D. track in January 2008. Snehit's research involves the application of computational models to high throughput DNA sequencing. He developed a method to analyze DNA from pooled sets of individuals, using error-correcting codes to identify each person. This work has been accepted to RECOMB 09 and selected as one of four papers considered for the Genome Research special issue for the conference. Snehit has also been involved in applied analysis of sequenced organisms, and is a joint-first-author on a publication describing the first animal mutant whose genome was assembled—a worm with two right "brains" (Nature Methods 2008).

Ph.D. student **Sean White** will speak on behalf of his fellow Ph.D. candidates at the convocation for doctoral candidates in the schools of Architecture, Business, Engineering, Journalism, Law, Nursing, Physicians and Surgeons, Public Health, and Teachers College. The convocation is held on Monday, May 18, 2009, at 3:30 PM in the Chapel. Congratulations to Sean on this honor!

Professors **Peter Belhumeur** and **Shree Nayar** have received a 5 year Multidisciplinary University Research Initiative (MURI) from the Office of Naval Research. The total award is for $7.5 million and is for a joint project on *Remote Multimodal Biometrics for Maritime Domain*, between Columbia University, University of Maryland, University of Colorado at Colorado Springs, and University of California at San Diego.

Professor **Steven Feiner** served as General Chair of the 2008 International Conference on Intelligent Technologies for Interactive Entertainment. He also served as General Co-chair for the 2008 ACM Symposium on Virtual Reality Software and Technology, and as chair of the best paper awards committee for IEEE Virtual Reality 2008. Professor Feiner gave invited talks on his research at the Universidade Nova de Lisboa, the University of Florida, the International Symposium on Ubiquitous Virtual Reality in Gwangju, Korea, the VTT Research Technical Research Center in Espoo, Finland, and the opening of the Experimental Media and Performing Arts Center at Rensselaer Polytechnic Institute. Professor Feiner was also invited speaker at the 2008 IBM Research Symposium on Human-Computer Interaction at TJ Watson Research Center and at Visual Computing Trends 2009 in Vienna, Austria, and spoke at invited panels at the 2008 IEEE International Symposium on Wearable Computers and the 2008 IEEE and ACM International Symposium on Mixed and Augmented Reality.

Professor **Steven Feiner** is a member of the new Games For Learning Center at NYU, a multi-disciplinary, multi-institutional games research alliance with support from Microsoft Research that is working to provide fundamental scientific evidence to support games as learning tools for math and science subjects among middle-school students.

Professor **Tony Jebara** was featured in a Business Week article on *Mapping a New Mobile Internet*.

Professors **Angelos Keromytis** and **Sal Stolfo** received a research gift from Symantec to aid their studies and investigations on techniques for scalable program whitelisting, for common software vulnerability discovery across large software installations, and for vulnerability-oriented analytics for risk assessment.

Professors **Angelos Keromytis** and **Sal Stolfo** received a 30-month grant titled *Automated Creation of Network and Content Traffic For the National Cyber Range* from DARPA, as part of a team led by BAE. A team of researchers led by BAE won a DARPA contract to design and build a National Cyber Test Range, a facility to test and evaluate new security technologies to defend the Internet from cyber attack.

Columbia University spin-off company, **StackSafe, Inc.**, is the winner of the 2008 National University Start-Up Competition. "Beating out more than 400 entrants from across the country, StackSafe was awarded first prize after a rigorous assessment by an online panel of over 300 venture capitalists, angel investors, and university judges." The company employs patent-pending network security and intrusion detection technologies developed by Professor **Angelos Keromytis** and **Salvatore Stolfo**. StackSafe, based in Vienna, Virginia, leverages virtualization technology to reduce costly IT production problems and downtime. StackSafe's flagship product, Test Center, enables IT departments to stage their existing software infrastructure, conduct pre-deployment tests, and predict IT complications in a secure environment.

The Department of Homeland Security (DHS) is supporting a new three-year project titled *Privacy Preserving Sharing of Network Trace Data*, in cooperation with BAE Systems National Security Solutions Inc. The project will be lead by Professors **Tal Malkin**, **Sal Stolfo**, **Tony Jebara**, **Steve Bellovin**, **Vishal Misra**, and **Dan Rubenstein**. The team will develop a next-generation network-trace anonymization tool that preserves individual and organizational privacy while still allowing cross-trace correlation for detection, understanding, and prevention of complex attacks and other network behavior. The tool will rely on three techniques that will be developed at Columbia University: (1) Hidden Markov Model (HMM)-based clustering will be used to divide raw network traces into groups for which statistical and other properties can be preserved across the anonymized equivalents; (2) more aggressive application- and definition-specific anonymization will prevent recovery and attribution of private topology, flow, and content information under attack; and (3) efficient and application-specific secure computation will allow this clustering and anonymization without centralizing or revealing the contents of individual raw traces during the clustering stage.

Professors **Vishal Misra** and **Dan Rubenstein** co-authored the paper *Opportunistic Use of Client Repeaters to Improve Performance of WLANs*, which received the Best Paper Award at ACM CoNext 2008 in Madrid. The paper was co-authored with Victor Bahl (Microsoft Research), Ranveer Chandra (Microsoft Research), Patrick P. C. Lee (Columbia U.), Jitendra Padhye (Microsoft Research), and Yan Yu (Google).

Professor **Steve Nowick** was named an IEEE Fellow for his contributions to asynchronous and mixed-timing integrated circuits and systems.

Professor **Itsik Pe'er** received a three-year grant from the National Science Foundation within the Program for Emerging Models and Technologies to reconstruct a genealogy of the human species. The project is driven by a novel approach for detecting relatedness of individuals from high throughput data of DNA variation across millions of genetic markers and tens of thousands of individuals. The approach is based on a new linear-time algorithm that stores words of marker data in a dictionary of genetic variants. This framework will handle the diversity of human genetic data in terms of populations and experimental platforms culminating in a complete map of human genetic genealogy.

Professor **Itsik Pe'er** received a four-year National Institute of Health (NIH) R01 award titled *Genomewide survey of allele-specific somatic copy number changes*, together with Tom LaFramboise of Case Western University and Matthew Freedman (Harvard) as a sub-contractor. Cancer is a genetic disease in two levels: First, like many diseases inherited mutations at the individual level may contribute to disease risk and susceptibility in families. Second, and more specific to cancer, it is a genetic disease of cells, that acquire mutations during the lifetime of the patient, transforming the cells from normal tissue into tumors. Professor Pe'er proposes a novel computational method to process signals from tumor DNA to detect interaction between these two kinds of mutations.

Professor **Itsik Pe'er** has been participating in the Severe Adverse Effects Consortium, an industry-academia partnership to investigate the genetic basis to fatal or near-fatal adverse reaction to commonly used medications. The consortium brings together major pharmaceuticals (Abbott, Daiichi Sankyo, GlaxoSmithKline, Johnson & Johnson, Novartis, Pfizer, Roche, Sanofi Aventis, Takeda, Wellcome Trust and Wyeth) with joint interest in safely prescribing drugs by means of personalized medicine. Columbia is serving as the informatics center, with Aris Floratos (C2B2) and Professor Pe'er.

Professor **Itsik Pe'er** received a grant from the National Institute of Health (NIH) to support a project investigating the genetic basis of schizophrenia with computational analysis. The study is led by Pablo Gejman (Northwestern) and Douglas Levinson (Stanford). The PIs are conducting a genomewide study for the association between genetic variants and schizophrenia, a highly heritable disease, but without clear common genetic factors. The research project started in May 2008.

Professor **Itsik Pe'er** won second prize, among 200 entries, in a genome sequencing award competition from Applied Biosystems. The company will sequence 30 billion "bases", i.e., letters of DNA. Professor Pe'er and his collaborators will generate the genomic sequence for individuals from an isolated Pacific population, aiming to provide insights on the genetics of obesity, lipid levels and diabetes.

Professor **Ravi Ramamoorthi** won a 2007 Presidential Early Career Award for Scientists and Engineers. Professor Ramamoorthi was cited for "investigating foundations of the visual appearance of objects and developing mathematical and computational models for recreating complex scenes for automated image understanding, and for mentoring undergraduate and graduate students." The awards are the nation's highest honor for faculty members that are beginning their independent research careers.

Professors **Simha Sethumadhavan**, **Angelos Keromytis**, and **Sal Stolfo** were awarded a $650,000 equipment grant from the Air Force Office for Scientific Research (AFOSR), titled *SCOPS: Secure Cyber Operations and Parallelization Studies Cluster*. The funds will be used to deploy a new compute cluster capable of continuous

line-speed capture and near-online analysis of network traffic and for storage and analysis of large traces generated from run-time profiling of (legacy) applications. The computational capabilities provided by this cluster will allow detailed modeling and in depth analysis of real world scenarios.

Professor **Joseph Traub**, Edwin Howard Armstrong Professor of Computer Science, gave a College of Computing Distinguished Lecture at Georgia Tech. The title of his lecture was *Exponential Improvement in Qubit Complexity*.

Together with Corinna Cortes, Professor **Vladimir Vapnik** has been awarded the Paris Kanellakis Theory and Practice Award for their revolutionary development of a highly effective algorithm known as Support Vector Machines (SVM). The Kanellakis Award honors specific theoretical accomplishments that significantly affect the practice of computing. Support Vector Machines are a set of related supervised learning methods used for data classification and regression common in the field of artificial intelligence. As a result of this work, SVM is one of the most frequently used algorithms in machine learning, and is used in medical diagnosis, weather forecasting, and intrusion detection among many other practical applications.

The Department at **Play**

Computer Science department students and faculty participate in a range of organized social activities. Here are some photos from recent hiking and ski trips.

Hike on the Hudson

Ski Trip to Hunter Mountain

Hrvoje Benko (Ph.D. '07) was pictured in a feature article on the front page of the New York Times Business section. The article, *Microsoft Mapping Course to a Jetsons-Style Future*, is about technology that is being developed at Microsoft, including Hrvoje's gesture recognition system.

Matthew Blaschko (SEAS '01) writes, "I've just finished my Ph.D. at the Max Planck Institute for Biological Cybernetics in Tuebingen, Germany, and am now at the University of Oxford as a postdoctoral research fellow where I work on computer vision and machine learning. If anyone wants to get in touch, send an email to *blaschko@caa.columbia.edu*."

Casey Callendrello (SEAS '07) writes, "I'm currently a Senior Network Operations Engineer at Akamai, and I've been having a great time there. It certainly justified the long hours I spent in the windowless confines of the Columbia NOC and the Network Lab. I'm also volunteering at a Boston-based charity called Bikes not Bombs, working on their website and other technology. Bikes not Bombs takes donated bicycles, repairs them, and sends them to needy people in developing countries. Our most recent shipment went to a group called Maya Pedal, a Guatemalan organization that turns bicycles into pedal-powered farm machinery."

Sebastian Enrique (M.S. '05) writes, "When I left Columbia I moved from New York to Vancouver to fulfill one of my dreams: to work in the game industry. It's been more than 3 years and I can say that I've successfully accomplished it by working on the FIFA video game franchise at Electronic Arts Canada. In all that period I have worked as a Software Engineer and Game Designer on four released titles: FIFA 06 and FIFA 07 for Xbox, PS2, Nintendo GameCube, and PC; and FIFA 08 and FIFA 09 for Xbox 360 and PS3.

It's been a rollercoaster ride where learning and contributions are at a maximum point in every development cycle. First working on the demos to gather my first video game and console development experience, then leading an area, and lately being responsible for the design, implementation, and execution of innovative features like Be A Pro (conceived in FIFA and adopted by all EA Sports titles) and Player Reactions, among others.

Columbia University has been invaluable for preparing me for the challenges I faced and the ones to come. I'm proud of my work and I encourage everyone to pursue your inner dreams and accomplish those goals.

You can see more about what I've been doing at *www.sebastianenrique.com.ar*."

Eleazar Eskin (Ph.D. '02) has been named a 2009 Sloan Foundation Research Fellow. The Sloan Research Fellowships seek to stimulate fundamental research by early-career scientists and scholars of outstanding promise. These two-year fellowships are awarded yearly to 118 researchers in recognition of distinguished performance and a unique potential to make substantial contributions to their field. Eleazar works in the area of molecular biology and is an assistant professor at UCLA. He was advised by Professor Sal Stolfo.

Tarun Kapoor (M.S. '01) will be joining Colibria, a Norwegian-based company specializing in IM & Presence Services for Mobile Operators and Service Providers, as Product Director of e2e solutions for SIP-based solutions, where he will focus on helping with the SIP strategy & solutions for the operators and service providers market. This position will be based out of Oslo, Norway.

Han Liang (M.S. '06) writes, "After I graduated, I worked for IBM for a year and half as a software engineer in Massachusetts. Then I joined TripAdvisor (*www.tripadvisor.com*) as a senior software engineer in 2008. Last April, TripAdvisor began to build a Chinese version of their site, and I was working on a project team to help them launch their site. Currently I am permanently relocated to the newly established Beijing office, and switched my role to manage the search engine marketing (SEM), search engine optimization (SEO) efforts here, and some product management. This office will expand to roughly 15 headcounts this year, and I am enjoying this startup feeling very much."

Ramiro Ordonez (SEAS '04) writes, "I am working for Charles River in Boston (after the fall of Lehman Brothers, which forced me to leave NYC). Things are really good for me here considering the economy, and I am looking forward to being more involved with the faculty and alumni."

Peter Ottomanelli (SEAS '03, M.S. '05) and his wife Lisa (CC '05) welcomed their first son on Christmas morning (12/25/2008) this past year.

Kundan Singh (Ph.D. '06) was married to Mamta Singh, also a computer professional, on November 29th in Kanpur.

Akira Tsukamoto (M.S. '03) writes, "I moved to Sony Computer Entertainment in 2006 and have been working on Cell Broadband Engine (Cell/B.E.) processor projects jointly with IBM and Toshiba since then. Currently I am leading an open source project called MARS to make programming easier and more efficient on Cell/B.E., which is used on PLAYSTATION 3 and IBM Cell Blades (also used on the first 1 petaflops super computer, IBM Roadrunner). MARS is a multi-tasking programming framework for Cell/B.E. processors which simplifies maximizing the usage of SPE cores on the Cell/B.E. MARS is SPE centric, in that each SPE runs a micro kernel that provides light weight context switching with synchronization with other SPEs and generally runs independent of the PPE core. Detailed documentation is at *http://ftp.uk.linux.org/pub/linux/SonyPS3/mars/1.0.1/mars-docs-1.0.1/html/* and a download is available at *http://ftp.uk.linux.org/pub/linux/Sony-PS3/mars/*."

Recent & Upcoming **PhD Defenses**

Paul Blaer

Advisor: Professor Peter Allen

View Planning for Automated Site Modeling

Abstract:
Constructing highly detailed 3-D models of large complex sites using range scanners can be a time consuming manual process. One of the main drawbacks is determining where to place the scanner to obtain complete coverage of a site. We have developed a system for automatic view planning called VuePlan. The system proceeds in two distinct stages. In the initial phase, the system is given a 2-D site footprint with which it plans a minimal set of sufficient and properly constrained covering views. We then use a 3-D laser scanner to take scans at each of these views. When this planning system is combined with our mobile robot, it automatically computes and executes a tour of these viewing locations and acquires them with the robot's onboard laser scanner. These initial scans serve as an approximate 3-D model of the site. The planning software then enters a second phase in which it updates this model by using a voxel-based occupancy procedure to plan the next best view (NBV). This next best view is acquired, and further NBVs are sequentially computed and acquired until an accurate and complete 3-D model is obtained. A simulator tool which we developed has allowed us to test our entire view planning algorithm on simulated sites. We have also successfully used our two-phase system to construct precise 3-D models of real world sites located in New York City: Uris Hall on the campus of Columbia University and Fort Jay on Governors Island.

John Cieslewicz

Advisor: Professor Kenneth Ross

Architecture-Sensitive Database Query Processing on Chip Multiprocessors

Abstract: Microprocessor design is currently in the midst of a sea change, which will have a significant impact on all computer applications including database management systems, the focus of this thesis. For the past several decades, microprocessor performance has been driven primarily by higher clock rates and the subsequent faster execution of a single stream of instructions (a single thread). Faster chip speeds require more power and produce more heat. A single thread also has difficulty providing enough work to keep a chip busy because applications often have to wait for high latency operations such as loading data from memory, an operation that becomes more expensive in terms of processor cycles the faster a chip is. For these reasons, processor architects have recently begun to improve processor performance by increasing the degree of on-chip thread-level parallelism rather than the chip speed. The processor designs featuring on-chip parallelism are termed chip multiprocessors and include multi-core and simultaneous multi-threading processors, both of which are examined in this thesis.

The move toward on-chip parallelism has important implications for computer applications. Database management systems are a core computer application used across disciplines within computer science and in almost every sector of the economy. This means that database performance has a big impact on most data related activities. This thesis examines the performance of several core database operations on different chip multiprocessor architectures. Performance challenges are identified and architecture-sensitive algorithms designed specifically for chip multiprocessors are proposed and verified. Three core challenges impact database performance on chip multiprocessors: (1) identifying sufficient parallelism with database operations to take advantage of on-chip parallelism, (2) managing shared on-chip resources including memory cache, memory bandwidth, and execution units, and (3) balancing the aforementioned resource sharing with inter-thread communication overhead.

The first part of this thesis examines the use of simultaneous multithreading technology to help overcome performance bottlenecks due to high memory latency in common database join algorithms. Then a highly parallel supercomputer is used to demonstrate the benefits of massively parallel architectures for data processing tasks. Finally, the last sections examine core database operations on a commodity chip multiprocessor with the highest currently available degree of on-chip parallelism. Chip multiprocessors are shown to have significant benefits for database performance, but achieving optimal performance requires database algorithms to be crafted with on-chip parallelism as a design parameter from the start.

Frank Enos

Advisor: Professor Julia Hirschberg

Detecting Deception in Speech

This dissertation describes work on the detection of deception in speech using the techniques of spoken language processing. The accurate detection of deception in human interactions has long been of interest across a broad array of contexts and has been studied in a number of fields, including psychology, communication, and law enforcement. The detection of deception is well-known to be a challenging problem: people are notoriously bad lie detectors, and no verified approach yet exists that can reliably and consistently catch liars. To date, the speech signal itself has been largely neglected by researchers as a source of cues to deception. Prior to the work presented here, no comprehensive attempt has been made by speech scientists to apply state-of-the-art speech processing techniques to the study of deception. This work uses a set of features new to the deception domain in classification experiments, statistical analyses, and speaker- and group-dependent modeling approaches, all designed to identify and employ potential cues to deception in speech. This dissertation shows that speech processing techniques are relevant to the deception domain by demonstrating significant statistical effects for deception on a number of features, both in corpus-wide and subject-dependent analyses. Results also show that deceptive speech can be automatically classified with some success: accuracy is better than chance and considerably better than human hearers performing an analogous task.

The work also examines speaker and group differences with respect to deceptive speech, and we report a number of findings in this regard. We provide a context for our work via a perception study in which human hearers attempted to identify deception in our corpus. Through this perception study we identify a number of previously unreported effects that relate the personality of the hearer to deception detection ability. An additional product of this work is the CSC Corpus, a new corpus of deceptive speech.

Agustin Gravano

Advisor: Professor Julia Hirschberg

Turn-Taking and Affirmative Cue Words in Task-Oriented Dialogue

Abstract: As interactive voice response systems spread at a rapid pace, providing an increasingly more complex functionality, it is becoming clear that the challenges of such systems are not solely associated to their synthesis and recognition capabilities. Rather, issues such as the coordination of turn exchanges between system and user, or the correct generation and understanding of words that may convey multiple meanings, appear to play an important role in system usability. This thesis explores those two issues in the Columbia Games Corpus, a collection of spontaneous task-oriented dialogues in Standard American English.

We provide evidence of the existence of seven turn-yielding cues—prosodic, acoustic and syntactic events strongly associated with conversational turn endings—and show that the likelihood of a turn-taking attempt from the interlocutor increases linearly with the number of cues conjointly displayed by the speaker. We present similar results related to six backchannel-inviting cues—events that invite the interlocutor to produce a short utterance conveying continued attention.

Additionally, we describe a series of studies of affirmative cue words— a family of cue words such as 'okay' or 'alright' that speakers use frequently in conversation for several purposes: for acknowledging what the interlocutor has said, or for cueing the start of a new topic, among others. We find differences in the acoustic/prosodic realization of such functions, but observe that contextual information figures prominently in human disambiguation of these words. We also conduct machine learning experiments to explore the automatic classification of affirmative cue words. Finally, we examine a novel measure of speaker entrainment related to the usage of these words, showing its association with task success and dialogue coordination.

Rean Griffith

Advisor: Professor Gail Kaiser

The 7U Evaluation Method: Evaluating Software Systems via Runtime Fault-Injection and Reliability, Availability and Serviceability (RAS) Metrics and Models

Abstract: Renewed interest in developing computing systems that meet additional non-functional requirements such as reliability, high availability and ease-of-management/self-management (serviceability) has fueled research into developing systems that exhibit enhanced reliability, availability and serviceability (RAS) capabilities. This research focus on enhancing the RAS capabilities of computing systems impacts not only the legacy/existing systems we have today, but also has implications for the design and development of next generation (self-managing/self-*) systems, which are expected to meet these non-functional requirements with minimal human intervention.

To reason about the RAS capabilities of the systems of today or the self-* systems of tomorrow, there are three evaluation-related challenges to address. First, developing (or identifying) practical fault-injection tools that can be used to study the failure behavior of computing systems and exercise any (remediation) mechanisms the system has available for mitigating or resolving problems. Second, identifying techniques that can be used to quantify RAS deficiencies in computing systems and reason about the efficacy of individual or combined RAS-enhancing mechanisms (at design-time or after system deployment). Third, developing an evaluation methodology that can be used to objectively compare systems based on the (expected or actual) benefits of RAS-enhancing mechanisms.

This thesis addresses these three challenges by introducing the 7U Evaluation Methodology, a complementary approach to traditional performance-centric evaluations that identifies criteria for comparing and analyzing existing (or yet-to-be-added) RAS-enhancing mechanisms, is able to evaluate and reason about combinations of mechanisms, exposes under-performing mechanisms and highlights the lack of mechanisms in a rigorous, objective and quantitative manner.

The development of the 7U Evaluation Methodology is based on the following three hypotheses. First, that runtime adaptation provides a platform for implementing efficient and flexible fault-injection tools capable of in-situ and in-vivo interactions with computing systems. Second, that mathematical models such as Markov chains, Markov reward networks and Control theory models can successfully be used to create simple, reusable templates for describing specific failure scenarios and scoring the system's responses, i.e., studying the failure-behavior of systems, and the various facets of its remediation mechanisms and their impact on system operation. Third, that combining practical fault-injection tools with mathematical modeling techniques based on Markov Chains, Markov Reward Networks and Control Theory can be used to develop a benchmarking methodology for evaluating and comparing the reliability, availability and serviceability (RAS) characteristics of computing systems.

This thesis demonstrates how the 7U Evaluation Method can be used to evaluate the RAS capabilities of real-world computing systems and in so doing makes three contributions. First, a suite of runtime fault-injection tools (Kheiron tools) able to work in a variety of execution environments is developed. Second, analytical tools that can be used to construct mathematical models (RAS models) to evaluate and quantify RAS capabilities using appropriate metrics are discussed. Finally, the results and insights gained from conducting fault-injection experiments on real-world systems and modeling the system responses (or lack thereof) using RAS models are presented. In conducting 7U Evaluations of real-world systems, this thesis highlights the similarities and differences between traditional performance-oriented evaluations and RAS-oriented evaluations and outlines a general framework for conducting RAS evaluations.

Andrew Howard

Advisor: Professor Tony Jebara

Large Margin Transformation Learning

Abstract: With the current explosion of data coming from many scientific fields and industry, machine learning algorithms are more important than ever to help make sense of this data in an automated manner. The support vector machine (SVM) has been a very successful learning algorithm for many applied settings. However, the support

vector machine only finds linear classifiers so data often needs to be preprocessed with appropriately chosen nonlinear mappings in order to find a model with good predictive properties. These mappings can either take the form of an explicit transformation or be defined implicitly with a kernel function.

Automatically choosing these mappings has been studied under the name of kernel learning. These methods typically optimize a cost function to find a kernel made up of a combination of base kernels, thus implicitly learning mappings. This dissertation investigates methods for choosing explicit transformations automatically. This setting differs from the kernel learning framework by learning a combination of base transformations rather than base kernels. This allows prior knowledge to be exploited in the functional form of the transformations which may not be easily encoded as kernels. Additionally, kernel based SVMs are often hard to interpret because they lead to complex decision boundaries which are only linear in the implicitly defined space. However, by working with explicit mappings, models with an intuitive meaning can be learned. The learned transformations can be visualized to lend insight into the problem, and the hyperplane weights indicate the importance of transformed features.

The two basic models that will be studied are the mixture of transformations and the matrix mixture of transformations. The matrix mixture reduces to mixture of transformation learning when the matrix M is rank 1 and mixture of kernel learning when M is a diagonal matrix. First, greedy algorithms are proposed to simultaneously learn a mixture of transformations and a large margin hyperplane classifier. Then, a convex semidefinite algorithm is derived to find a matrix mixture of transformations and large margin hyperplane. More efficient algorithms based on the extragradient method are introduced to solve larger problems and extend the basic framework to a multitask setting. Another cost function based on kernel alignment is explored to learn matrix mixture of transformations. Maximizing the alignment with a cardinality constraint on the mixture weights gives rise to approximation algorithms with constant factor approximations similar to the Max-Cut problem. These methods are then applied to the task of learning monotonic transformations which are built from a mixture of truncated ramp functions. Experimental results for synthetic data, image histogram classification, text classification and gender recognition demonstrate the utility of these learned transformations.

Alpa Jain

Advisor: Professor Luis Gravano

Query Processing over Relations Extracted from Text Databases

Abstract: Text documents often embed data that is structured in nature, and this structured data is increasingly exposed using information extraction systems. Information extraction systems generate structured relations from documents, thus enabling expressive, structured queries over text databases. This dissertation studies the problem of processing structured queries over relations extracted from text databases.

To process structured queries over text databases, we face multiple challenges. One key challenge is efficiency: information extraction is a time-consuming process, so query processing strategies should minimize the number of documents that they process. Another key challenge is output quality: information extraction systems are often far from perfect, and might output erroneous information or miss information that they should capture, hence hurting output accuracy and completeness. At the same time, query processing decisions, such as the choice of information extraction systems or document retrieval strategies, also impact the output quality. Finally, depending on the nature of the information need, users may have varying preferences regarding the execution efficiency and quality expected from the querying process. This dissertation builds on the critical observation that, in addition to efficiency, which is important just as in traditional relational query optimization, the output quality of an execution is critical.

In our extraction-based scenario, query processing can be decomposed into a sequence of basic steps: retrieving relevant text documents, extracting relations from the documents, and joining extracted relations for queries involving multiple relations. Each of these steps presents different alternatives and together they form a space of possible query execution strategies. Our goal is to consider the user-specified requirements for execution efficiency and quality, and choose an execution strategy for each query based on a principled, cost-based comparison of the alternative execution strategies. We first introduce a simple, integrative optimization approach for processing queries involving single as well as multiple extracted relations. This approach considers each execution strategy as a whole and exploits database-specific statistics to predict the execution strategy characteristics. We then move towards an in-depth understanding of the impact of each component of an execution strategy on the overall execution. With this in mind, we rigorously analyze the critical components of an execution strategy and build statistically robust representations for information extraction systems, as well as statistical models for document retrieval strategies and join processing algorithms. These models help predict the efficiency and output quality of a variety of query execution strategies. Finally, we also consider the common scenario where information extraction systems report the extracted tuple together with scores that reflect the confidence in the correctness of the extracted tuples. Specifically, we present query processing algorithms that leverage these confidence scores and efficiently produce the high-confidence tuples, in turn discarding extracted tuples that are likely to be incorrect.

In summary, this thesis presents a principled query optimization approach for processing structured queries over text databases, taking into consideration both the efficiency and the output quality of the query execution strategies. Our hope is that the contributions of this work will help shrink the gap between structured databases and text databases, by enabling the seamless and expressive querying of all available information, regardless of whether it is in structured databases or embedded in natural language text.

Anshul Kundaje

Advisor: Professor Itsik Pe'er

Predictive Models of Gene Regulation

Abstract: Inferring gene regulatory networks and signal transduction pathways from high-throughput genomic and proteomic data is one of the central problems in computational biology. There has been an explosion in the quantity and types of high-throughput biological data. These datasets tend to be noisy, incomplete, high-dimensional and undersampled. Due to the complexity of the biological systems and a constant evolution of our understanding of these systems, building flexible and robust models that aid automatic hypothesis generation remains a challenge.

We present a new predictive modeling framework for deciphering gene regulation. The core computational approach is based on the GeneClass and MEDUSA algorithms. MEDUSA integrates promoter sequence, mRNA expression, and transcription factor occupancy (ChIP chip) data to learn regulatory programs that predict the differential expression of target genes. MEDUSA is able to learn context-specific regulatory networks and discover regulatory sequence motifs ab-initio. We design new techniques for extracting biologically and statistically significant information from the learned regulatory models. We show the utility of this approach in analyzing several interesting biological contexts (environmental stress responses, DNA-damage, hypoxia) in the budding yeast Saccharomyces cerevisiae. We present wet-lab experiments to validate hypotheses generated by our models. Experimental results demonstrate that MEDUSA is able to achieve high prediction accuracy and reveal context-specific regulatory mechanisms even on small datasets.

We present extensions to the model for integrating interactome data such as signaling and protein-protein interaction data. We use a graph-based representation of the interactome data for reconstructing active signaling pathways. We demonstrate that our methods scale well for predictive modeling in higher eukaryotes such as C. elegans and humans. We also present an off-the-shelf software tool for learning, interpreting and visualizing predictive models of gene regulation. In this thesis, we introduce algorithms and software that make powerful new techniques for modeling gene regulatory networks broadly accessible to the biological community.

Sujit Kuthirummal

Advisor: Professor Shree Nayar

Flexible Imaging for Capturing Depth and Controlling Field of View and Depth of Field

Abstract: Over the past few centuries cameras have greatly evolved to better capture our visual world. However, the fundamental principle has remained the same—the camera obscura. Consequently, though cameras today can capture incredible photographs, they still have certain limitations. Recent years have seen several efforts to overcome these limitations and extend the capabilities of cameras through the paradigm of computational imaging— capture the scene in a coded fashion, which is then decoded computationally in software. This thesis subscribes to this philosophy. In particular, we present several imaging systems that enable us to overcome limitations of conventional cameras and provide us with flexibility in how we capture scenes.

First, we present a family of imaging systems called radial imaging systems that capture a scene from a large number of viewpoints, instantly, in a single image. These systems consist of a conventional camera looking through a hollow conical mirror whose reflective side is the inside. By varying the parameters of the cone we get a continuous family of imaging systems. We demonstrate the flexibility of this family—different members of this family can be used for different applications. One member is well suited for reconstructing objects with fine geometry such as 3D textures, while another is apt for reconstructing larger objects such as faces. Other members of this family can be used to capture texture maps and estimate the BRDFs of isotropic materials.

We then present an imaging system with a flexible field of view—the size and shape of the field of view can be varied to achieve a desired scene composition in a single image. The proposed system consists of a conventional camera that images the scene reflected in a flexible mirror sheet. By deforming the mirror we can generate a wide and continuous range of smoothly curved mirror shapes, each of which results in a new field of view. This system enables us to realize a wide range of scene-to-image mappings, in contrast to conventional imaging systems that yield a fixed or a fixed set of scene-to-image mappings.

All imaging systems that use curved mirrors (including the ones above) suffer from the problem of defocus due to mirror curvature; due to local curvature effects the entire image is usually not in focus. We use the known mirror shape and camera and lens parameters to numerically compute the spatially varying defocus blur kernel and then explore how we can use spatially varying deconvolution techniques to computationally "stop-up" the lens—capture all scene elements with sharpness while using larger apertures than what is usually required in curved mirror imaging systems.

Finally, we present an imaging system with flexible depth of field. We propose to translate the image detector along the optical axis during the integration of a single image. We show that by controlling the motion of the detector—its starting position, speed, and acceleration—we can manipulate the depth of field in new and interesting ways. We demonstrate capturing scenes with large depths of field, while using large apertures to maintain high signal-to-noise ratio. We also show how we can capture scenes with discontinuous, tilted or non-planar depths of field.

All the imaging systems presented here subscribe to the philosophy of computational imaging. This approach is particularly attractive as with Moore's law computations become increasingly cheaper, enabling us to push the limits of how cameras can capture scenes.

Patrick Lee

Advisor: Professor Vishal Misra

Achieving Network Robustness with Diversity Overlay Routing

Abstract: Network systems, in both wired and wireless domains, traditionally forward data along a single best routing path for communication. Given that network communication has become a major part of today's economy, it is important to guarantee the robustness of network communication toward general failures or malicious attacks. However, there are fundamental constraints on re-engineering existing network systems to use robustness mechanisms. First, network systems are not amenable to centralized solutions because they usually span multiple administrative domains. Also, it is non-trivial to modify low-layer hardware and protocols of legacy network systems to support new robustness mechanisms.

To overcome the fundamental robustness challenges, we propose to adapt existing network systems into robust architectures using the concept of overlays. Layered atop the physical network, overlays provide a flexible platform for building robustness mechanisms. In particular, we seek to design distributed solutions that are amenable to nodes that can communicate with one another, but are independently operated and prefer to make local decisions on link costs and route choices.

This thesis proposes to use diversity overlay routing to address two robustness problems: (1) resilient multipath routing, in which distributed algorithms are proposed to route data along multiple available network paths in order to minimize the loss of throughput due to node or link failures, and (2) network fault correction, in which an end-to-end inference approach is proposed to aggregate information

from multiple paths in order to minimize the cost of diagnosing and repairing network failures.

First, we develop a distributed resilient multipath solution that routes data across available network paths to improve routing resilience in wired overlay networks. Our solution allows routing decisions to be made locally by network nodes without requiring centralized information, such as the entire network topology. We devise two distributed algorithms termed the Bound-Control algorithm and the Lex-Control algorithm, with an objective of minimizing the loss of throughput due to a single-link attack. We formally prove that both algorithms converge to their respective optimal solutions. Using simulation, we show that our resilient multipath solution effectively mitigates the loss of throughput due to single-link attacks as well as multi-link attacks.

We next consider the use of opportunistic routing to boost the throughput gain of routing in multi-hop wireless networks. Our goals are to overlay the 802.11 MAC layer, exploit the broadcast nature of wireless communication, and utilize the spatial diversity of a network. We consider a novel routing scheme called Stable Opportunistic Routing (SOR), which integrates the techniques of forwarder scheduling and network coding to improve the throughput gain. Using nsclick simulation, we show that SOR improves throughput when compared with traditional single-path routing and previous opportunistic routing schemes under various settings, including the scenarios where link-level measurements are inaccurate. We also verified the benefits of SOR using testbed experiments.

Finally, we consider an end-to-end approach of inferring network faults in an overlay system, with an optimization goal of minimizing the expected cost of correcting (i.e., diagnosing and repairing) all faulty nodes. Conventional fault localization problems first check the most likely fault node, i.e., the node with the highest conditional failure probability given a network with faulty nodes. By taking into account the cost of correcting faulty nodes, we prove that an optimal strategy should start with checking one of the candidate nodes, which are identified based on a potential function that we develop. We propose several efficient heuristics for inferring the best node to check in large-scale networks. By extensive simulation, we show that checking first a candidate node decreases the cost of correcting all faulty nodes when compared with checking first the most likely faulty node.

This thesis seeks to provide insights into the solutions of overcoming the fundamental robustness challenges of existing network systems. We consider one possible solution, i.e., using diversity routing on top of an overlay architecture, and show that how this solution improves the robustness of network communication toward general failures or malicious attacks.

Sean White

Advisor: Professor Steven Feiner

Interaction and Presentation Techniques for Situated Visualization

Abstract: Every day we move through a sea of invisible information. As computation, sensing, and display become more mobile and distributed, interaction shifts from our desktop to our environment. This shift changes how we interact with our surroundings, creating the opportunity for situated visualization—visual representation of data presented in its spatial and semantic context. This dissertation contributes novel interaction and presentation techniques for situated visualization in three research areas: mobile visualization, objects as context, and scenes as context. In support of this research, we make further contributions developing algorithms, architectures, and working artifacts.

For mobile visualization, we present results from a field study and task analysis of botanists performing species identification in the field. We develop an iteratively designed, extensible Electronic Field Guide (EFG) system and architecture, a conceptual data model, and interface (LeafView) for mobile visualization. Field experiments show improved identification speed and interaction efficacy.

Next, we explore spatially- and semantically-driven situated visualization using objects as context in head-worn augmented reality (AR). We develop and evaluate tangible AR and head-movement controlled AR interaction techniques. In interviews following lab experiments, participants reported improved speed for inspection and comparison and a preference for tangible AR. Building on this work, we design, develop, and evaluate visualization and activation techniques for discovering and learning gestures for these user interfaces (Visual Hints). Lab experiments show preference for visual hints that combine overlaid graphics and animation. In addition, we investigate menu techniques, activated by shaking an object, for interacting with visualizations (Shake Menus). We compare display-, object-, and world-referenced coordinate systems for presentation of menu options. Lab experiments show increased speed and accuracy when using the display-referenced coordinate system.

Finally, we present techniques for spatially-driven visualization in the context of physical scenes. Based on our field study of site visits by urban designers, a hand-held AR visualization tool (SiteLens) embodying these techniques enables interaction with invisible aspects of urban sites, such as georeferenced sensor data. Field experiments provide evidence for new insights derived from situated visualization, preference for specific representations, and improved interaction with data using a novel stabilization algorithm.

Stay in Touch!

Visit the CUCS Alumni Portal at **https://mice.cs.columbia.edu/alum** to:

- Update your contact information
- Look at recent job postings
- Get departmental news

If you know another alumni who may not be receiving the newsletter, please forward the Alumni Portal link to them.

In the spirit of environmental responsibility, we are moving towards more electronic (and less hard-copy) distribution of the newsletter.

If you'd like to continue receiving paper copies of the newsletter, please visit the Alumni Portal to sign up.

You can subscribe to the CUCS news mailing list at **http://lists.cs.columbia.edu/mailman/listinfo/cucs-news**

CUCS colloquium information is available at **http://www.cs.columbia.edu/lectures**

Read the CUCS Newsletter online at **http://www.cs.columbia.edu/resources/newsletters**

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