UMIACS
UNIVERSITY OF MARYLAND INSTITUTE FOR ADVANCED COMPUTER STUDIES

InterConnections

Contents

Former NSF Director Rita Colwell joins UMIACS

Named Distinguished University Professor

Following a six-year term as the first woman director of the National Science Foundation (NSF), Rita Colwell has returned to the University of Maryland faculty as a distinguished professor and is working with UMIACS and Life Sciences faculty. "I had an extremely successful run at NSF and I completed my agenda," says Colwell. "I returned to the university because the contributions right now that can come from science are enormously attractive. I have the opportunity to pursue some very important scientific questions."

During her tenure as NSF director, NSF's budget increased 68 percent and the average annual grant size grew from $90,000 to $142,000. Colwell also was responsible for increases in graduate student stipends, initiatives to support women and minorities in science and engineering, and programs to encourage science and math graduate students to participate in K-12 education.

Colwell was recently named to the National Women's Hall of Fame, joining inductees such as Senator Hillary Rodham Clinton. She is considered a pioneer for her groundbreaking work in the study of global infectious diseases, water and health. Always a scientist at heart, Colwell continued ongoing research projects even while carrying out her NSF administrative duties.

Among her highest priorities now is developing an international network to address infectious diseases and water issues. The network would include research from scientists in countries such as Sweden, Norway, Japan, India and Bangladesh.

Colwell says, "It is imperative that a network be implemented for the prevention and/or mitigation of the global problem of emerging and re-emerging infectious diseases, specifically those caused by waterborne pathogens. We know from satellite observations that when sea surface temperature increases in the Bay of Bengal, there is an increase in the number of cholera cases. This requires a multi-dimensional understanding of

COLLEGE OF COMPUTER, MATHEMATICAL AND PHYSICAL SCIENCES

Director's message

During the past year, it has been a pleasure for me to serve first as interim director, and subsequently director of UMIACS. I have been fortunate to witness spectacular growth in our research program during this period. We have grown in many ways—through the addition of several preeminent members, the addition of exciting new research programs, and the award of numerous national and international honors to our faculty, staff and students. This has been accompanied by a significant leap in external research funding.

We have expanded our faculty by recruiting two distinguished, seasoned, and well-recognized researchers.

The first, Professor Rita Colwell, comes to us from her most recent position as director of the National Science Foundation. Professor Colwell is setting up an ambitious new international center for the study of infectious diseases.

In addition, we have recruited Professor Steven Salzberg to head up the Center for Bioinformatics and Computational Biology (CBCB). CBCB is a joint partnership between UMIACS, the University of Maryland's College of Mathematical, Computer and Physical Sciences and College of Life Sciences. Professor Salzberg will join our faculty on July 1, 2005. He currently serves as senior director of bioinformatics at the Institute for Genomics Research. Professor Salzberg's arrival will be the first of several hires in CBCB. Under his leadership, we hope to see CBCB achieve international stature in bioinformatics. Read more on page 7.

Our senior hires have been accompanied by junior hires—during the past academic year, we hired Amol Deshpande from Berkeley, with expertise in databases; Ramani Duraiswami from Maryland, an expert in computer vision; and Neil Spring from Washington, with expertise in computer systems. Read about them on page 10.

It is impossible to list all the new research programs that took shape during the past year. However, a few significant ones merit mention. We set up a $5 million Joint Institute of Knowledge Discovery to determine how best to extract accurate and valuable knowledge from large text and other media databases. We also set up the Center for Human Enhanced Secure Systems (CHESS) to protect computer systems from malicious attacks using a mix of completely automated methods and sophisticated human interfaces that allow computer security experts to increase their efficiency in targeting malware. You can read about both of these efforts in this newsletter.

Our faculty has had numerous honors showered upon them during the past year. I will just mention two – Professor "Pete" Stewart was elected to the National Academy of Engineering, and Professor Rita Colwell was named to the National Women's Hall of Fame. Our faculty is routinely asked to give invited and keynote lectures at national and international conferences, and are frequently profiled in the news media.

Last, but not least, UMIACS' research awards leapt almost 50 percent from Feb. 1, 2004, to March 31, 2005.

UMIACS Director V.S. Subrahmanian

Rita Colwell joins UMIACS

the complex interaction between biological systems, including humans, and the global physical and chemical environment."

According to Colwell, in addition to classical epidemiology, such diseases are driven by outside forces that include climate, worldwide movement of people and goods, sanitation, personal hygiene and malnutrition. "Understanding these forcing functions as a holistic system can lead to developing models and practical applications for the prediction, prevention, treatment and/or elimination of diseases."

Colwell uses the recent tsunami in South Asia as an example that stresses the importance of such a network. Although the tsunami was a horrible tragedy, had the event taken place just a month earlier it could have been much worse, according to Colwell. "During September, October and November, the sea surface temperature is elevated and results in a large population of plankton that carry the Vibrio cholerae bacteria," she notes. "After the tsunami hit land it would have left behind pockets of water with Vibrio cholerae bacteria, potentially exposing many thousands of individuals to the disease."

A significant role of the network will be to establish an international data repository that will prepare for and address such outbreaks, providing computational analysis and bioinformatics capabilities. Colwell envisions that "the power of information technology will be brought to bear on the complex analyses this program will require, especially in discerning patterns and correlations heretofore unsuspected or unrecognized. Predictive models will be constructed on a global basis, taking into account the data available from other disciplines such as weather, storm and ocean current patterns as well as seasonal vegetation and animal cycles."

Colwell believes that UMIACS is "a terrific place to be working, with its leadership in large database management and operations."

She is also a faculty member in the new Center for Bioinformatics and Computational Biology, (CBCB) a partnership between the University of Maryland's

College of Mathematical, Computer and Physical Sciences and College of Life Sciences. CBCB is organized as a center within UMIACS.

Colwell also is a distinguished professor with the Johns Hopkins University Bloomberg School of Public Health and chair of Canon US Life Sciences, Inc., which is working to identify and develop life-science solutions with potential applications in diagnostic and medical instrumentation.

Prior to directing NSF, Colwell was the president of the University of Maryland Biotechnology Institute and before that founding director of the university's Center of Marine Biotechnology. Colwell, who received a bachelor's degree in bacteriology and a master's degree in genetics from Purdue and a Ph.D. in oceanography from the University of Washington, joined the university as a professor of microbiology in 1972. She is a member of the National Academy of Sciences.

During her career, she has held numerous advisory positions in the federal government, nonprofit science policy organizations and private foundations, as well as in the international scientific research community. She has authored and co-authored 16 books and more than 700 scientific publications. In addition, she produced the award-winning film, *Invisible Seas*. Colwell also has held leadership positions with numerous professional organizations, such as chair of the Board of Governors of the American Academy of Microbiology and president of the American Association for the Advancement of Science.

With her return to the University of Maryland, she is entering another phase of her professional life. Colwell says, "For me, there are still questions to be answered, books to be written, and papers to be published."

A delegation from the government of Brazil, including the country's minister of social security, recently visited UMIACS to discuss core IT technology that could improve the management of tax data and identify tax fraud. UMIACS Director V.S. Subrahmanian said, "The meeting went very well and the ministry is currently discussing possible avenues for collaboration."

UMIACS faculty receive NSF grant for CPU-GPU cluster research

UMIACS faculty have received a three-year, $1.1 million grant from the National Science Foundation (NSF) to establish a CPU-GPU cluster for research in high-performance computing and visualization of large scale applications. "We are building a unique computational infrastructure that will leverage the synergistic coupling of CPUs, GPUs, displays and storage," says Amitabh Varshney, associate professor of computer science and principal investigator of the grant.

Varshney says the recent rise in the capabilities and programmability of graphics processing units "has enabled them to become powerful enough to be useful co-processors to the CPUs." He explains that the goal of the project is to study, understand and harness the power of this newly emerging paradigm of a CPU-GPU combination for high-performance computing and scientific applications with visualization-assisted computational steering.

A cluster of CPUs and GPUs offers a rich parallel computing environment—a controlled heterogeneity strengthened with the complementary capabilities of CPUs and GPUs with MIMD and SIMD dataflow paths. A CPU-GPU cluster has the added attractiveness of enabling a tightly coupled interleaving of visualization and computation. Such a coupling permits exploring problems whose solution methodology itself can be profitably changed by using visualization to assist in computational steering and the computation to provide feedback for a richer, more expressive visualization.

The CPU-GPU cluster is enabling the pursuit of several new research directions in computing. The grant has 15 faculty members participating across seven different areas of research, including visual and geometric computing, high-performance querying of large datasets, numerical algorithms for complex systems with uncertainty, computational biology, real-time computer vision, real-time virtual audio, and efficient compilation of signal processing algorithms.

Besides the significant new research underway, Varshney says the project will "greatly assist in education for undergraduate and graduate students through its impact on courses as well as through better research opportunities."

Gordon and Loeb lead new research in economic aspects of information security

The research began with two curious minds and an intriguing topic—the economic aspects of information security. But then it became something much more. "There was a need, and we responded to it," says Professor Lawrence Gordon. Gordon and his colleague Professor Martin Loeb are applying economic concepts to cyber security problems with the goal of ultimately preventing or at least reducing their occurrence.

Professor Martin Loeb

Six years ago the two men stumbled upon the topic and spent the summer looking at it together. "Every issue you read about related to information security has an economic aspect to it," says Gordon.

Gordon and Loeb, who both hold appointments with the Robert H. Smith School of Business and UMIACS, have gone on to address such key issues as how much a firm should invest in information security, the most efficient way to allocate funds invested in information security, the real economic impact of information security breaches on corporations, determining whether or not information sharing actually helps prevent cyber breaches, and whether or not an information security audit enhances the economic value of a firm.

"We saw a real opportunity to have an impact outside, as well as inside the academic community," says Gordon.

For example, Gordon and Loeb were involved with a joint survey conducted by the Computer Security Institute (CSI) and the FBI. They were part of a three-person academic team involved in the survey, considered the most influential in its field. At the request of CSI, they provided advice on improving the survey instrument and helped analyze the data collected. Gordon and Loeb also co-authored the report based on the survey's findings, which confirmed that the threat from computer crime and other information security breaches is a real and continuing problem.

However, the report also found that overall financial losses from cyber crime decreased significantly in the previous year, from almost $202 million to about $141 million; and that organizations are allocating more resources to cyber security and using metrics from economics to evaluate their security decisions.

"Information is a valuable asset in today's world," says Loeb. "In a company it's not just private individuals' information which is being protected, but the details of ongoing mergers, or product development information. The need to protect this asset is a driving force when it comes to allocating resources to information security."

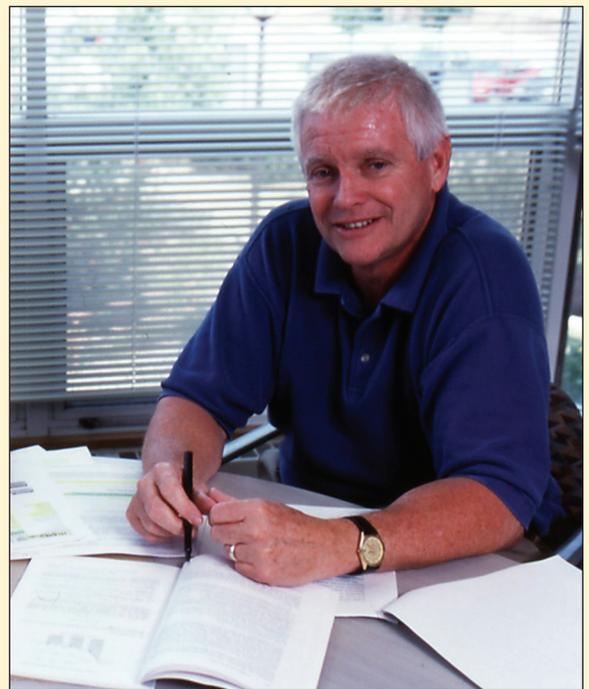"Understanding and properly utilizing economic concepts has real value," says Gordon.

Gordon and Loeb's work has been published in such leading publications as *Communications of the ACM* and *ACM Transactions on Information and System Security*. They have also been frequently written about in the popular media, including recently in *Business Week*. Gordon and Loeb are currently completing a book, *Managing Cyber Security Resources: Myth vs. Realties.* The book pro-

vides a guide for managers dealing with the economic and financial aspects of information security.

Their published papers not only explore relevant issues but offer guidance and solutions. For example,

- "Evaluating Information Security Investments Using the Analytic Hierarchy Process (AHP)" shows how a chief information security officer can apply the AHP to determine the best way to spend a limited information security budget and make a case to the organization's chief financial officer for an increase in funds to further enhance the organization's information security.

- "The Economic Cost of Publicly Announced Information Security Breaches: Empirical Evidence from the Stock Market" examines the economic effect of information security breaches on the stock market value of corporations. "This approach takes into account indirect costs, as well as direct costs, to the firm," says Loeb.

The analysis shows that cyber security

Professor Lawrence Gordon

breaches, in which confidential private information is compromised (e.g., the release of customer credit card numbers, bank account numbers, or medical records to unauthorized parties), have a significant negative effect on the stock market value of the attacked firm. However, security breaches not related to confidentiality (e.g., a temporary shut down of a corporate web site) involve costs that are transitory and unlikely to significantly affect shareholder value.

Gordon says this shows that "market participants appear to discriminate across types of breaches. Economically rational investment strategies should focus on protecting the firm's most valuable information assets."

• "The Economics of Investment in Information Security" presents an economic model that characterizes the optimal monetary investment to protect a given set of information. It is shown that, for a given potential loss, the optimal amount to spend to protect an information set does not always increase with increases in the information set's vulnerability. "Protecting highly vulnerable information sets may be inordinately expensive, and a firm may be better off concentrating its efforts on information sets with midrange vulnerabilities," says Gordon. "Moreover, the paper shows that the amount the firm should spend to protect information sets should generally be only a small fraction of the expected loss."

• "Budgeting Process for Information Security Expenditures: Empirical Evidence" reports the findings of an empirical study of the budgeting process used by U.S. corporations in making decisions concerning expenditures on information security activities. Loeb says the findings show "some senior information security managers use formal discounted cash flow analyses, while many more are initiating a risk management process grounded in basic economic principles to decide on the level of such expenditures." Gordon adds that the findings also indicate "the major impediment to pursuing a rigorous economic cost-benefit approach to information security expenditure decisions is the difficulty in quantifying the potential benefits from security activities."

"When we began this research," says Gordon, "there was essentially nothing out there. We had no idea how much of an opportunity it would turn out to be and how quickly we would have an impact. But, we really have had, and continue to have, a tremendous impact on the field. That's very rewarding and exciting."

The research has energized the two already highly respected and well-established faculty members. "This has been a rebirth for us," says Gordon.

Gordon is the Ernst & Young Alumni Professor of Managerial Accounting and Information Assurance at the Smith School. Loeb is a professor of accounting and information assurance and the Deloitte & Touche LLP Faculty Fellow for the Smith School.
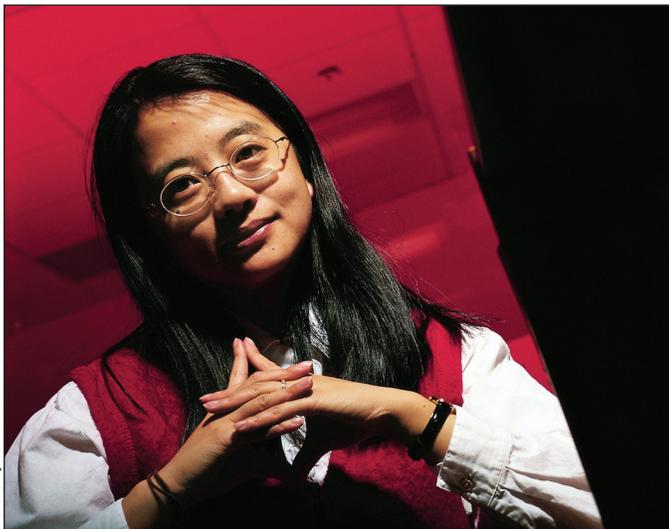
Min Wu among world's top young innovators on 'TR100' list

Photo by John T. Consoli

tal images. Her research has applications from the Department of Defense to the Hollywood film industry.

Wu holds U.S. Patent No. 6,285,775 for technology that adds a digital signature to a black-and-white image by subtly altering individual pixels on the edges of letters and other symbols. If widely adopted, this technology could make it virtually impossible to forge or alter an electronic image or document. Wu also holds three other patents on digital watermarking and multimedia. In 2003 she co-authored *Multimedia Data Hiding* with Bede Liu of Princeton University. She won an NSF CAREER Award for "Signal Processing Approaches for Multimedia Security and Information Protection" in 2001.

Originally from Beijing, Wu received her Ph.D. in electrical engineering from Princeton in 2001. She joined Maryland in 2001 with a joint appointment in UMIACS and the Department of Electrical and Computer Engineering. She is also affiliated with the Institute for Systems Research.

Technology Review began the TR100 list five years ago to recognize the world's top innovators under age 35. Inclusion among the TR100 has become one of the most prestigious awards for young innovators around the world.

Assistant professor Min Wu is among the world's 100 top young innovators according to Massachusetts Institute of Technology's *Technology Review* magazine, which named her to its "TR100" list. The list names 100 individuals under age 35 whose innovative work has made a profound impact on the world.

The nominees were recognized for transforming biotechnology, medicine, computing and nanotechnology.

Wu is developing novel solutions to multimedia security and content protection problems such as fraud prevention of binary documents and forensic tracking of classified multimedia content. She is a signal processing expert on security in multimedia applications and on hiding information in digi-

With guidance from UMIACS professor, high school student is Intel finalist

Based on the research he did at UMIACS last summer, Michael Forbes was recently selected as one of 40 finalists in the 2005 Intel Science Talent Search, one of most prestigious science competitions for high school seniors in America. The prize was the capstone of a most interesting journey.

UMIACS Professor Samir Khuller clearly recalls the day last spring when Forbes, a shy student from Montgomery Blair High School in Silver Spring, Md., knocked hesitantly on his door. He had been doing some research for his senior project and wanted Khuller to look over it. Reading the abstract of his paper, Khuller realized that the main result in Forbes' paper was a few decades old and showed him a book that confirmed it.

While Forbes was disappointed to learn that his findings were not new, Khuller was impressed with him all the same. "For a high school student to discover this theorem on his own seemed quite an achievement from my point of view," he says. Khuller suggested that Forbes read some related papers. But he did not think he would hear from Forbes again "because the papers were kind of technical and hard to read, even for a Ph.D student."

A few weeks later, Forbes sent him an e-mail and asked to schedule a meeting. Khuller recalls, "I met with him and this time he came with a long list of questions on the papers I had given him. He had carefully made his way through them and marked the parts that he wanted to discuss. I was most impressed. It was clear that he understood the papers very well."

Forbes then asked Khuller if he could work with him that summer. For Khuller, who had never worked with a high school student, this would be a new experience. He arranged for a desk for Forbes and during the summer of 2004, Forbes came to campus almost every day for eight weeks. Khuller was initially unsure about how best to mentor Forbes. He wondered whether he should give him a programming project (as he would with his undergraduate students), or throw some real research problems at him.
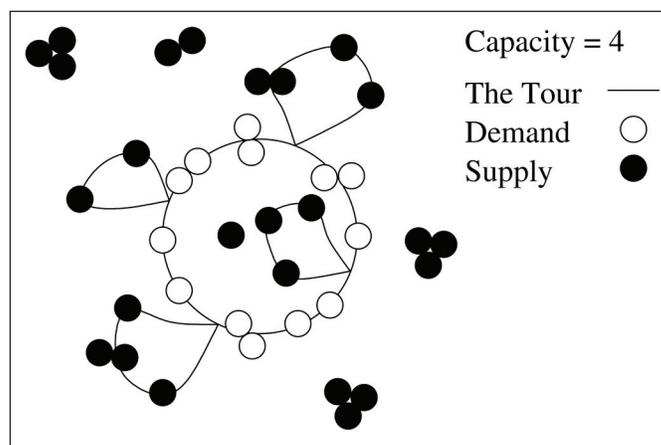
Eventually, "since Forbes seemed so sharp, and quick to understand things, I decided that it would not hurt to throw him a real research problem," says Khuller. "I formulated two related questions for him based on vehicle routing work I had done with Moses Charikar (Princeton) and Balaji Raghavachari (UT Dallas) in 1997. I thought that this would be a problem Forbes could start working on after reading a small number of papers."

Each week Forbes met with Khuller individually and also participated in group meetings with Khuller's graduate students. It soon became apparent that he could hold his own among them. As Khuller says, "Michael is a deep thinker and very quick to learn new things. He has a sharp mind and strong analytical skills."

Forbes took the opportunity and ran with it. "I felt right from the start a strong connection with Dr. Khuller," he says. "I can't begin to describe how much I learned in one summer. Dr. Khuller and the graduate students were very helpful in teaching me many aspects of computer science. I gained a lot from discovering how to find papers online and reading them and gaining insight into the many different ways to approach a problem."

Over the summer, Forbes read Khuller's course notes from his graduate class on algorithms. He also worked on research that formed the basis for his Intel entry, "Capacitated Vehicle-Routing and the k-delivery n-Traveling Salesman Problems." The basic Traveling Salesman problem asks for the shortest route visiting a given set of locations. This problem is computationally challenging and the best algorithms for it run in exponential time. The Vehicle Routing problem is a significantly harder problem where one has a vehicle

A tour for a vehicle that can only hold four units of supply.

with a fixed capacity that has to perform deliveries. Previous work had studied the case when the number of objects to be delivered were the same as the available supply. Forbes focused on the case where the supply and demand are not necessarily equal. In this case, in addition to finding the shortest route, one has to decide which deliveries to actually perform. Forbes worked on developing algorithms that ran quickly, but did not necessarily find the shortest route. In this case, he wanted to show that the obtained solution is *guaranteed* to be close to optimal.

"This area is extremely hard," says Khuller. "One has to be very creative and mathematically agile. One has to be thorough and have tremendous intuition. Getting nice results is not easy."

"This type of research is what I envision myself doing in life," says Forbes, who graduated this spring. "This opportunity allowed me to preview what it would be like to be a professor and do research for a living. I am enthusiastic about this topic more than ever and look forward to pursuing it in college."

Forbes' college choices are Maryland, Cornell and MIT. Khuller would like the chance to work with Forbes again. "I really look forward to working more with Michael," he says. "He is going to be an extraordinarily gifted scientist."

Leading researcher in bioinformatics and genomics to head new center

Steven Salzburg, a leading researcher in bioinformatics and genomics, is joining UMIACS. He will assume the Phillip H. and Catherine C. Horvitz Professorship in Computer Science and will head up the new Center for Bioinformatics and Computational Biology, (CBCB) a partnership between the University of Maryland's College of Mathematical, Computer and Physical Sciences and College of Life Sciences. CBCB is organized as a center within UMIACS.

"We anticipate that Dr. Salzberg will be able to make the university's bioinformatics program one of the finest in the world," says V. S. Subrahmanian, UMIACS director.

Salzburg certainly brings with him the credentials to do so. He comes to UMIACS as the senior director of Bioinformatics at the Institute for Genomic Research (TIGR) in Rockville, Md. TIGR is a nonprofit, basic research institute devoted to the sequencing and analysis of humans, animals, plants and microbes. Salzberg was also a research professor of computer science and a research professor of biology at Johns Hopkins University.

Early in his career, Salzberg focused on machine learning and its applications to fields ranging from astronomy to molecular biology. Later, his interest in the human genome project motivated him to develop one of the first computational gene-finding systems for the human genome in the mid-1990s. "There really has been a genomic revolution," he says. "I knew that it was going to be one of the most exciting projects in the world and would be for some time to come."

Salzberg joined TIGR as an investigator in 1997 and became its director of bioinformatics in 1998. His initial collaborations with TIGR led to the development of a microbial gene-finding system that was subsequently used in the analysis of the bacterial genomes of Borrelia burgdorferi, which causes Lyme disease; T. pallidum, which causes syphilis; M tuberculosis; V. Cholerae; and more than 50 other bacterial, viral and other species that have been sequenced since then.

"I work closely with life scientists and biologists to understand problems and where someone with my skills can contribute to solving interesting problems," says Salzberg, who earned his Ph.D. in computer science at Harvard. He also earned a bachelor's degree with honors in English and a master's degree in computer science, both from Yale.
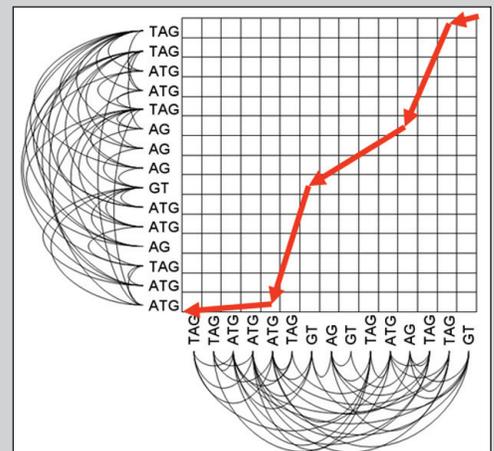
Salzberg and his research team developed a eukaryotic gene finder, first used for Plasmodium falciparum or malaria parasite, and later adapted to many other eukryoes, including plants and animals. The team also developed systems for large-scale genome sequence alignment, operon discovery and genome assembly. "These open-source software systems have been distributed to thousands of scientific laboratories around the globe," says Salzberg. In addition, his group developed and applied computation methods to analyze genome duplications, rearrangement and other evolutionary phenomena in a wide range of organisms.

"You have to ask the right questions," says Salzberg of his research. "I look at the questions that will have the biggest impact on science and society."

He finds much satisfaction in the human element of that research. "I do a lot of work on diseases, and I think it's rather exciting to think about computer science having a hand in helping cure some of these diseases," he says.

Salzberg says the new center at the University of Maryland will be a computational facility that will include an integrated pathogen database with state-of-the-art algorithms designed to facilitate the rapid identification of microbial pathogens.

Salzberg is excited about the challenge of the new center and the research it will produce. "The University of Maryland is very committed to making this a world-class research center," he says. "This is a wonderful time in biology."

The increased availability of genome sequences of closely related organisms has generated much interest in utilizing homology to improve the accuracy of gene prediction programs. Generalized Pair Hidden Markov Models (GPHMMs) have been proposed as one means to address this goal. Dr. Salzberg's team developed an open-source GPHMM gene finder, TWAIN, and applied it to a pair of related Aspergillus species, A. fumigatus and A. nidulans. On these fungal genomes TWAIN performs very well, predicting 89% of the exons correctly and 74% of the gene models correctly in a held-out test set of 147 conserved gene pairs.

This chart illustrates how TWAIN obtains a pair of predictions. Tracing back through the optimal trellis path highlights two corresponding paths in the parse graphs. These paths outline the selected gene predictions in the two genomes.

Victor Basili receives honorary doctorate; symposium held in his honor

Dr. Basili, center, receives the honorary doctorate from (left) the Chairman of the Computer Science Department Professor Dr. Hans Hagen, and (right) the President of the Technical University of Kaiserslautern Professor Dr. Helmut J. Schmidt. Dr. Schmidt holds "The Gateway to Knowledge," the symbol of the university, which was given as a gift to Dr. Basili.

On January 28, 2005, Professor Victor Basili was awarded an honorary Ph.D., Doctor Honoris Causa in Natural Science, from the Technical University of Kaiserslautern, Germany. The event included technical talks by Laszlo (Les) Belady, CEO of Eutecus Inc., USA (the former vice president of Mitsubishi Research) and Professor Lionel Briand, of Carleton University, Canada. There were also presentations by the University President Professor Dr. Helmut J. Schmidt, and the Chairman of the Computer Science Department, Professor Dr. Hans Hagen. The laudation speech was given by Professor Dr. Dieter Rombach. Professor Basili gave a presentation on the future directions of empirical software engineering.

Professor Rombach noted that this was the first honorary Ph.D. degree awarded by the department since its beginning in 1975. Professor Rombach said, "We give Professor Basili this award today due to his outstanding scientific achievements, his contributions to the establishment of the sub-discipline of empirical software engineering, his outstanding ability to teach and motivate students and younger scientists, his personal leadership, and finally his direct and indirect contributions to our department here in Kaiserslautern."

Belady also said Professor Basili "found a way to gather valuable data whose analysis can then lead to generally usable formulae to guide research as well as development."

Briand noted Professor Basili's accomplishments: his emphasis on the scientific method as a means to achieve quality and productivity improvement in software engineering (Quality Improvement Paradigm), the organizational model to implement it in a typical software development setting (Experience Factory), the link of corporate measurement programs to the scientific method and quality/productivity improvement (Goal Question Metric Paradigm), the large body of experimental works, and his creation of a school of thought that has helped spread the application of the scientific method throughout software engineering research.

This was the second honorary degree for Professor Basili. He also received a Laurea Honoris Causa in Informatic Engineering from the University of Sannio, Italy, in April 2004.

On May 16, 2005, a Symposium on the Foundations of Empirical Software Engineering: The Legacy of Victor Basili, was held at the International Conference on Software Engineering in St. Louis, Mo., to celebrate Basili's 65th birthday. A book of several of his papers, *Foundations of Empirical Software Engineering: The Legacy of Victor R. Basili,* was edited by Barry Boehm, Dieter Rombach, and Marvin Zelkowitz. It was published in June 2005 by Springer.

Software defect detection tool exceeds 100,000 downloads

FindBugs, a software defect detection tool developed at the University of Maryland, has been downloaded more than 100,000 times, with more than 10,000 downloads in each of the past three months. FindBugs analyzes Java programs to find various types of coding defects, such as errors that lead to null pointer exceptions, incorrect use of Java libraries, vulnerabilities to malicious code or SQL injection attacks, and incorrect synchronization.

FindBugs has found hundreds of serious errors in each of several large production applications, such as Sun's core Java run-time library and IBM's Websphere application server. Compared with many similar commercial tools for Java, FindBugs does deeper analysis aimed at finding actual software defects, rather than just violations of coding standards. In several studies, more than half of the high-confidence warnings generated by FindBugs corresponded to serious coding defects.

FindBugs is an open-source project, and other developers have made contributions such as translations to Japanese, French, and German, and plug-ins for tools such Eclipse and Maven. The primary developers are David Hovemeyer, a graduate student in the Department of Computer Science (CS), and William Pugh, a professor in CS. The development of FindBugs has been supported by National Science Foundation grant CCR-0098162.

For more information or to download, visit *findbugs.sourceforge.net.*

UMIACS faculty, postdocs and students win awards for their research

Neil Spring wins William R. Bennett Prize

The paper, "Measuring ISP Topologies with Rocketfuel," co-authored by Neil Spring and published in the *IEEE/ACM Transactions on Networking* in 2004, has been awarded the IEEE Computer Society's William R. Bennett Prize.

In it, Spring and his colleagues describe methods for measuring detailed maps of the connections between routers in 10 of the ISP networks that make up the Internet. They found that by using snapshots of global Internet routing information, they could focus their efforts on a few ISP networks at a time, discovering detail in those networks not found by prior efforts. They used these topologies to better understand how the Internet's constituent networks are designed. The Rocketfuel topologies that were measured have been widely used to evaluate new research in networking.

The William R. Bennett Prize is given annually to the best original paper published in the *IEEE/ACM Transactions on Networking.* Papers are judged on quality, originality, utility, timeliness and clarity of presentation. The web site *www.comsoc. org/socstr/org/operation/awards/bennett.html* lists some very well-respected papers from prior years.

UMIACS postdoc wins best paper award at ACL-2005

David Chiang, a UMIACS postdoctoral researcher working with Philip Resnik, won the Best Paper Award at ACL-2005, the 43rd Annual Meeting of the Association for Computational Linguistics. This is the top conference in computational linguistics, with an acceptance rate of 18 percent this year. David's paper was on a hierarchical phrase-based model for statistical machine translation. His system learns a synchronous context-free grammar from texts in parallel translation without any prior syntactic information.

David's work served as a key component of the University of Maryland's entry at the National Institute of Standards and Technology's (NIST) 2005 Machine Translation Evaluation (MTEval) this year. The University of Maryland was among a small handful of top performers. The NIST evaluations are the primary community-wide evaluation forum for research in machine translation.
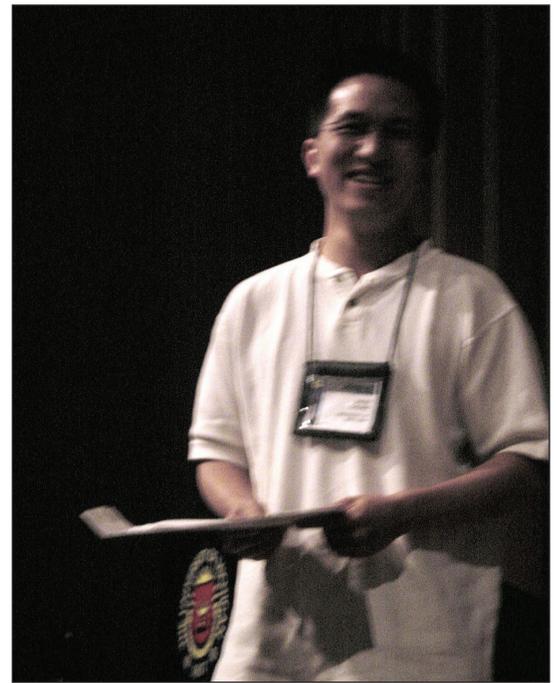
Grad student wins best student paper award at digital libraries conference

UMIACS graduate student Jack Kustanowitz was the lead author for the paper that won the Best Student Paper award at the ACM/IEEE Joint Conference on Digital Libraries 2005. Jack will receive a cash award, courtesy of the IEEE Technical Committee on Digital Libraries, TCDL. Jack collaborated with Professor Ben Shneiderman on "Meaningful Presentations of Photo Libraries: Rationale and Applications of Bi-Level Radial Quantum Layouts." View a PDF version of this paper at *www.cs.umd.edu/hcil/brqlayer/papers/BiLevl HierarchicalLayouts.pdf.*

The project on novel layouts for photo libraries is described with examples at: *www.cs.umd.edu/hcil/brqlayer/.* The software is available for download; the lively dynamic effects can be seen with your own photo collection.

Gang Qu receives Microsoft Trustworthy Computing Research Award

Assistant Professor Gang Qu has won a Trustworthy Computing (TWC) Curriculum RFP award from Microsoft Research. Dr. Qu received a $50,000 cash award for his winning project titled "A Multidisciplinary and Integrated

Professor Bonnie Dorr recorded the moment when David Chiang received his best paper award at ACL-2005.

Approach to Raise the Global Awareness of Trustworthy Computing." This award will enable Qu to conduct research and curriculum development that integrates security, privacy, reliability, and business integrity.

Qu will offer a new upper-level undergraduate course titled "Introduction to Trustworthy Computing" in Spring 2006. The new course will have deep computer science and engineering content, but will also include business case studies and analysis, guest speakers from social science and industry, a seminar series associated with the course, and a column in the school paper on trustworthy computing topics. The class will conduct campus-wide surveys to test the success of the project and raise the publicity and visibility of trustworthy computing. The course materials will be made available to the public both at the Microsoft Developer Network Academic Alliance Curriculum Repository and the National Information Assurance Training and Education Center.

Dr. Qu's was one of 15 Trustworthy Computing Curriculum RFP awards presented by Microsoft Research.

For additional information, visit: *research.microsoft.com/ur/us/twc/.*

Deshpande, Duraiswami, Spring join UMIACS faculty

Amol Deshpande, Ramani Duraiswami, and Neil Spring have joined UMIACS as new faculty members. "We are excited about these new additions to the UMIACS faculty," says UMIACS Director V.S. Subrahmanian. "We look forward with great anticipation to what these talented individuals will bring to our program."

Amol Deshpande joined the University of Maryland in spring 2005 after completing his Ph.D. at the University of California, Berkeley. His thesis work was in database management systems. "Specifically, my research was on handling estimation errors during query optimization by improved estimation procedures through the use of adaptive query processing techniques," Deshpande says. He obtained his B.Tech. degree in 1998 from the Indian Institute of Technology in Bombay.

Deshpande's current research interests include exploring data management issues in emerging domains such as wireless sensor networks, data streams and wide-area data sources. Deshpande believes "wireless sensor networks, comprising tiny, battery-powered sensing devices capable of communicating over wireless radios, have virtually unlimited potential in improved sensing, automation and control." He is interested in applying probabilistic modeling techniques to reason about the data collected from sensor networks. He is also interested in query optimization and query processing in data streams, wide-area data sources and XML data sources.

Ramani Duraiswami received his B.Tech Degree in 1985 from the Indian Institute of Technology in Bombay and his Ph.D. in 1991 from Johns Hopkins University. For the next seven years he worked for private companies in research and development. Ramani joined UMIACS as an assistant research scientist in 1998 and was promoted to associate research scientist in 2003. He joined the Department of Computer Science in July 2004 as an assistant professor.

In the last three years, Duraiswami has published more than 60 articles and a book and has mentored 10 Ph.D. students, four of whom have since graduated. He has developed a successful research program and was named among the top 100 faculty at the university for securing outside funding. Duraiswami also has formed strong scientific collaborations within and external to the university.

Duraiswami has interdisciplinary research interests. He says, "My work emphasizes use of scientific computation, mathematical and physical modeling, and systems development for solving problems in varied fields." His current interests are in computational auditory perception, computer vision, visual perception, virtual reality and wave scattering. Of note is a recently developed ingenious technique to measure a quantity called the "head-related transfer function" important in spatial audio. "I look forward to forging further collaborations and more exciting research in the years ahead," Duraiswami says.

Neil Spring received his bachelor's degree in computer engineering from the University of California, San Diego, in 1997, and his Ph.D. in computer science from the University of Washington in 2004. His dissertation focused on tools for finding accurate topologies of networks run by large providers in the Internet and characterizing the routing policies that determine which paths these providers choose. The network topologies measured as part of his dissertation research have been widely used to evaluate new network models and protocols.

His current work focuses on the tools and technologies for accurate, repeatable and scalable Internet measurement. "Surprisingly, few data sources track how the design of the Internet has changed over time, despite the dramatic changes in users and applications over the last decade," Spring says.

Spring designed and implemented a safe framework, called scriptroute, for executing often-dangerous network measurement code. He says, "Scriptroute has become a model service supported by the Planet Lab network testbed and allows students and researchers everywhere to build distributed tools to understand the Internet." Building upon the success of scriptroute, Spring is currently constructing systems to support a collaborative "reverse-engineering" of the Internet.

University of Maryland adds new Joint Institute for Knowledge Discovery to UMIACS

The University of Maryland has created the Joint Institute for Knowledge Discovery (JIKD), which will be a part of UMIACS. Funded by a combination of federal government and industrial affiliate funds, JIKD will serve as a center of excellence in the development of new and innovative approaches to finding relevant knowledge in large-scale data stores and semi-structured or unstructured data and text repositories.

JIKD will explore both research and technology application opportunities in a wide range of knowledge technologies including database research, social network analysis, human language technology, semantic web applications, human computer interaction, multimedia databases, and probabilistic predictive reasoning.

JIKD will bring together leading researchers in databases, artificial intelligence, speech analysis, computer vision, digital libraries, human computer interaction, web technology, and secure communications.

More than 20 faculty members are currently involved in JIKD. The institute was established with the help of a $5M contract whose principal investigator is V.S. Subrahmanian. Co-PIs are Rama Chellappa, Carol Espy-Wilson and Jim Hendler. Professor Hendler will serve as the first director of JIKD.
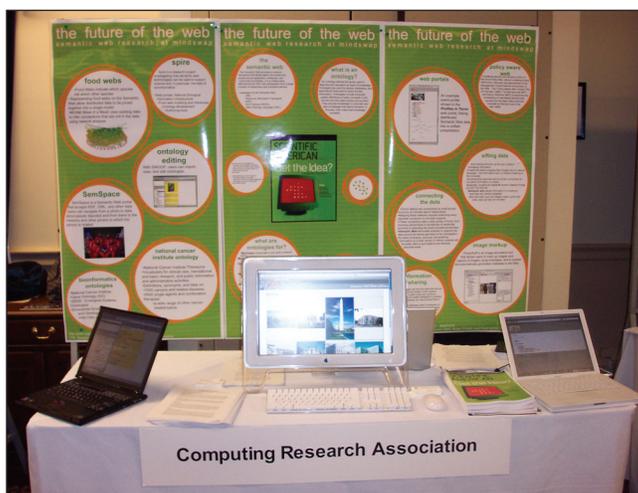
JIKD will initially focus on the following technical areas:

- Exploring knowledge discovery in large e-mail repositories: Headed by Professor Doug Oard (College of Information Studies and UMIACS), this project focuses on the language, search and modeling, and interface issues involved in exploring and finding relevant information in large, semi-structured, text-based knowledge repositories.

- Integrating probabilistic and symbolic information for data integration: Headed by UMIACS Director V.S. Subrahmanian (Computer Science and UMIACS), this project focuses on exploring how probabilistic analysis can be coupled with ontologies and schemas for structured data to provide highly correct and complete integration, search, and presentation of the information in extremely large data repositories.

- Annotating and evaluating knowledge products: Headed by Professor David Doermann (UMIACS), this project will explore the "ground truthing" of data sources that are used in the other projects. This will allow researchers to explore the efficacy of their approaches against known criteria and look at new approaches to ground truthing tools based on the results of research in the earlier projects.

In addition to these projects, an exploration will be made of semantic technologies for the integration of the many different types of data sources explored at JIKD and an infrastructure for exploring the distributed and dynamic processing of large-scale and possibly uncertain data resources. This will not only create an infrastructure for integrating the results of the projects described above, but also will allow research into the support of advanced knowledge-oriented repositories and the meta data-based integration of multiple text and multimedia data repositories.

Jim Hendler talks about the semantic web at the 11th annual Coalition for National Science Funding science exhibition and reception on Capitol Hill, held in June 2005. The event brings together 31 universities and scientific associations to highlight for members of Congress and their staff some of the interesting and important research supported by the National Science Foundation.

The Mind Lab/Computing Research Association exhibit featured examples of semantic web applications in science and anti-terrorism efforts. The semantic web application has helped group members explore links between terrorists operating in Iraq and elsewhere around the globe—including an eye-catching web of links between the leader of al Qaeda in Iraq, Abu Musab Al-Zarqawi, and Osama bin Laden. As new information was entered, the demonstration dynamically created new linkages between persons in the database.

Center for Human Enhanced Secure Systems established
Will focus on human element of computer security

A new UMIACS center, the Center for Human Enhanced Secure Systems (CHESS) was established last fall to foster interdisciplinary research in computer security. CHESS Director Jeff Hollingsworth says the center's unique work emphasizes "the human aspects of computer security: programmers, system administrators, CIOs, and elected officials."

It's important to address more than the theoretical issues of computer security, Hollingsworth says. "We believe that technology alone can't solve the problems of computer security. We're looking at how to make this technology more usable, more reliable and more understandable. No matter how well it works on paper, if you can't reliability build it or understand it, you're not going to use it."

CHESS researchers are involved with investigating security at all levels, such as developing new ways to write and test software, analyze network and wireless security, and investigate the public policy and economic implications of secure systems.

According to Hollingsworth, the center concentrates on four primary areas: code analysis and development; database security and information management; network and wireless security; and policy and management. As these research areas suggest, the center's faculty bring with them varied expertise backgrounds from not only the Department of Computer Science, but also the Department of Electrical and Computer Engineering, the Smith School of Business, the School of Public Policy and the College of Information Studies.

"A lot of good security-related research is already being conducted within UMIACS," says Hollingsworth. "This center will help foster collaborations among the faculty."

Among the researchers is the young, talented Min Wu *(see related story on page 5)*. Well-known business school professors Lawrence Gordon and Martin Loeb, who conduct groundbreaking research in the economic aspects of information security *(see related story on page 4),* are also among the faculty.

CHESS is working in partnership with IBM, Microsoft and Sun Microsystems. "This is an important topic," says Hollingsworth. "The focus of UMIACS on interdisciplinary research makes it the perfect place for this center."