UMIACS
University of Maryland Institute for Advanced Computer Studies

InterConnections

Winter 2009

Contents

InterConnections is published by the University of Maryland Institute for Advanced Computer Studies (UMIACS).

Visit UMIACS on the web at: *www.umiacs.umd.edu*.

Director: V.S. Subrahmanian

Contact information
UMIACS
2119 A.V. Williams Bldg.
University of Maryland
College Park, MD 20742-3251
Phone: 301.405.6722

New WiMAX Forum Lab to Develop Next-Generation Wireless Applications

With the widespread prevalence of handheld devices like PDAs, Internet Tablets, BlackBerries, and iPhones, people increasingly want and expect to be able to browse the internet at high speed anytime, anywhere, and not just at wireless hot spots where broadband internet access is available over a few hundred feet. Ashok Agrawala, a professor in the department of computer science, calls WiMAX, which stands for Worldwide Interoperability for Microwave Access, the next generation of wireless technology, providing some of the solutions for mobile broadband communications needs. This fall, Sprint launched wireless internet service using WiMAX over the city of Baltimore. Washington D.C. is expected to have its own WiMAX network up and running this winter.

Since the fall, Agrawala has led a team that is deploying a WiMAX network on the University of Maryland campus. An international body called the WiMAX Forum, which controls WiMAX technology and promotes conformance and interoperability of this standard, selected the university as only the second laboratory in the world approved as a WiMAX Forum Applications Lab. (The first applications lab was established in Taiwan.) Agrawala named the lab he directs MAXWell Lab. By giving some 45,000 users at the university access to the wireless resource, developers can test WiMAX applications in a realistic environment.

The MAXWell Lab is a joint laboratory of UMIACS and the Office of Information Technology (OIT).

Masoud Olfat of Sprint Nextel co-chaired the committee that selected UM as a WiMAX Forum Applications Lab. In addition to meeting all the requirements for a lab, UM had the advantage of a "strategic location," he says, in its proximity to government resources and to the first commercial WiMAX network in Baltimore. Among the forum's requirements was that a lab have existing infrastructure and experience developing applications. "We would have hesitated with others with no such experience," Olfat says. As the director of the Maryland Information and Network Dynamics, or MIND, laboratory, Agrawala has collaborated extensively with private industry as well as government agencies to create applications in the areas of wireless networking, networking infrastructure, information services, and security. For example, he's helped develop rapidly deployable systems to track emergency personnel and to allow them to communicate effectively. He's

continued on page 4.

Director's message

UMIACS Director V. S. Subrahmanian

2008 has seen incredible growth in UMIACS' research portfolio, as well as expanding impact of UMIACS faculty activities in the research, national, and international arena.

First and foremost, the WiMAX Consortium selected UMIACS to host the second WiMAX Forum laboratory in the world - this is also the first WiMAX Lab in the western hemisphere. As developing countries increasingly turn to wireless communications, largely bypassing the expensive investment in wired connections, WiMAX is emerging as the standard of choice for wireless communications in developing economies. The selection of UMIACS faculty member Ashok Agrawala to create the first WiMAX Forum Lab in UMIACS demonstrates the degree of excellence achieved by UMIACS in this field.

UMIACS research has had far reaching impact in other areas as well. Professor Bill Pugh's pioneering open source FindBugs system to find bugs in Java code has been downloaded over 500,000 times and was the topic of a day-long global Google "fixit" session in which over 700 engineers used FindBugs that fixed over 1000 problems.

More recently, NVIDIA named UMIACS an NVIDIA Center of Excellence - one of a handful of such centers in the world. Led by Professor Amitabh Varshney, and supported generously by NVIDIA, this center focuses on the development of core CPU-GPU technology and its applications to a wide variety of fields including bioinformatics, fluid dynamics, and scientific computing.

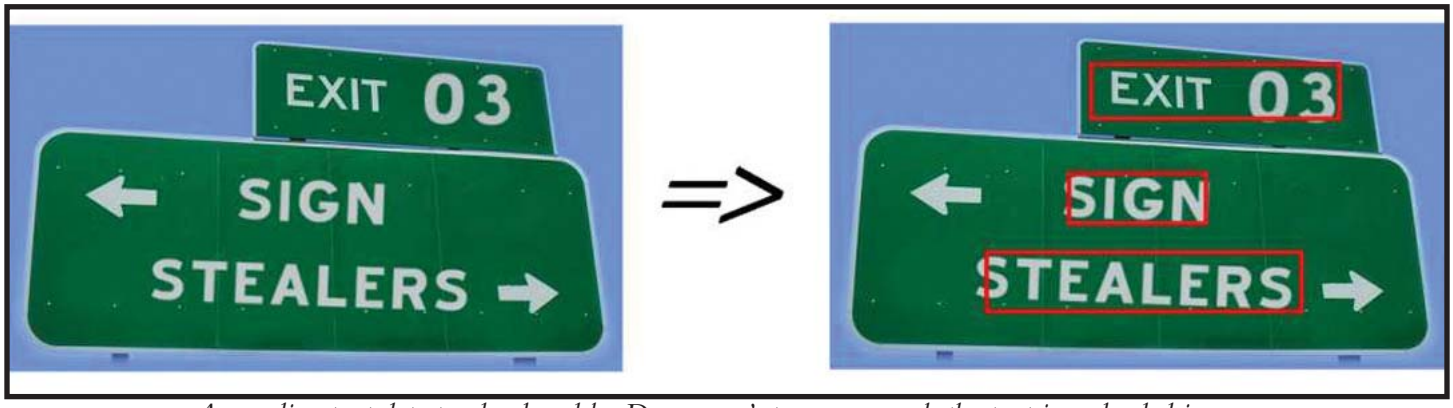
In addition, UMIACS faculty received numerous honors. These included two honorary doctorates from other universities, election to Fellow Status in various national and international scientific associations, numerous best paper awards and best student paper awards at leading conferences. We see incredible additional upcoming growth in these fields in 2009 and I look forward to further reports of this kind.

Visual Aspects of Document Analysis, continued.

the work we do is up-front analysis," Doermann says, such as identifying directly from the image and prior to obtaining converted text the type of document that is in hand and the language it's written in. Anyone can buy optical character recognition (OCR) software that can convert crisp, laser-printed office documents into electronic text, but Doermann's challenge is to examine documents that are not standardized and perfect. "What the government has is often old and highly degraded. It may be hardcopy that came out of a decade-old file cabinet, or it may be something that came from someone's pocket," he explains. Doermann also focuses on developing OCR for handwritten text. The post office is known for such OCR, but addresses are very constrained text that allow for built-in feedback. When such context isn't available, most OCR for handwriting is still very weak, Doermann notes.

Besides working to identify the type of document in hand and analyzing its content, Doermann and his colleagues work to isolate the clutter in documents by segmenting out characteristic noise, for example, automatically eliminating the frame around a page created by photocopying or removing the speckle from a scanned document. Two students who work with him are developing such enhancements.

"Our group has stayed away from character recognition that commercial applications do well. Our focus has been in identifying the difficult problems and doing the front end, the triage end," Doermann says. For example, his group removes some of the noise from analysis by segmenting zones and text lines and by normalizing content to produce a uniform page.



An on-line text detector developed by Doermann's team can mark the text in uploaded images.

"Once we do those things, we can pass along more standardized documents to commercial OCR."

Another focus of Doermann's work is to develop methods to analyze less commonly taught languages for which commercial OCR software is not likely to be developed. For example, Doermann's team has worked to develop methods to analyze Hindi, Burmese, and major languages in Afghanistan, Pashto and Dari.

Human in the loop

One of the principles that has emerged from researchers who try to automatically process language is the benefit of using human input in limited ways to bootstrap and refine automated systems. For example, Doermann and his coworkers developed a way to digitize bilingual dictionaries that requires a human to first teach the program the conventions used in a given dictionary. A human defines how entries are laid out. In general, a linguist can analyze a small amount of data, and programs can be built to be adaptive, able to incorporate iterative cycles of human feedback to establish better methods of analysis.

At this point, Doermann says that his team can develop OCR for any true-type font. To help make the programs his team develops usable in the real world, Doermann set up a core set of algorithms to debug and test software to help convert programs "from a graduate student project … into a bullet-proof program." Having programs in use helps the lab as well as users. "Once customers are engaged, they tell you what they really need. Students get motivated when there's a real problem to solve," Doermann says.

Not all the applications Doermann and his students develop are for government use or even focus on language. One of his students, Xu Liu, who defended his Ph.D. last fall, has worked to extend the use of cameras in cell phones and other mobile devices by developing a number of applications. One application uses cameras in mobile devices to read currency and is being deployed by the National Federation for the Blind. The project won a 2008 ACM Student Research Competition award. "We don't start with a theory. We start with a problem," Liu says of work in the Doermann group. Because the group is flexible and Doermann encourages students to be actively involved in every step of a project from writing proposals to reports, Liu says, "Working in the lab is like working in a small startup. We got to work on the details of the whole project. We have the chance to know everything."

Dave Doermann

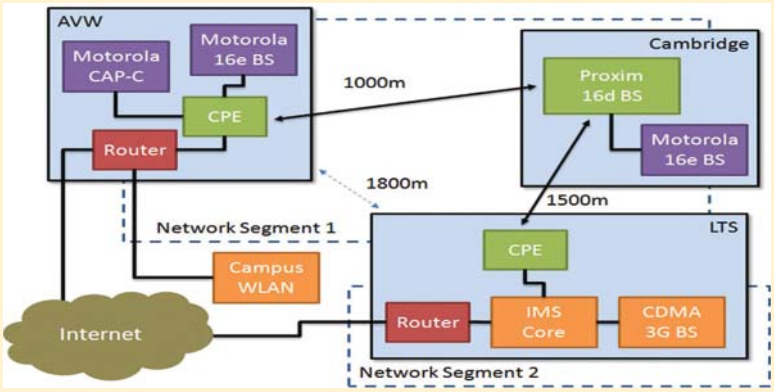
Ashok Agrawala in his office.

also developed systems for helping local law enforcement agencies to communicate with each other.

Communication speed on the campus WiMAX network is comparable to cable and ISDN services. WiMAX networks are established by erecting towers like those used for cell phones. "We are putting up two antennas," Agrawala said in October, "one on top of A.V. Williams and the other on Centreville Hall, one of the taller dorms on campus." These two antennas will establish a WiMAX network for a large part of the campus by Spring 2009.

Campus applications for the wireless network

"In order to assess the performance of applications, we want to put as much load as possible on the network and get as many students to use it as possible," Agrawala said. The network is expected to be used largely through hand-held devices although laptops can access it as well. "The speed will be better than the wireless I have at home," Agrawala said. He envisioned that students would use the network for such things as looking up class schedules, grades and assignments, dining hall menus, and movie schedules at the Hoff Theater. Students can use the high-speed network to communicate with friends via audio and video.



A map of the WiMAX network on the university campus.

Agrawala expects to develop new kinds of "context-aware, location-aware" applications for the WiMAX network. His research group has developed technologies that can locate people in buildings within a couple of feet and outdoors on campus within 10 feet. In MAXWell, Agrawala and his students plan to develop and test a social networking application for cell phones and PDAs that they call MyeVyu, which uses the ability to locate users on campus and to give them relevant information. For example, users could be alerted when they are close to friends, get customized information about university transportation and Metro, remember where they parked, or get directions to an office, a nearby restroom, or vending machines in a building. The network could help chart routes for employees or students who use wheelchairs or have other specific needs, Agrawala says.

Agrawala and his team are working with the university police on public safety applications for the network as well. By pressing a button on one's handheld device, a user could allow a police dispatcher to see and hear what is going on during an emergency and accurately relay his or her location. "We want to work on applications that improve the quality of life for the university community," Agrawala says, whether that means helping students in emergencies or with more everyday tasks like keeping in touch with friends, finding a shuttle bus, or ordering pizza.

A laboratory environment

Any device with WiMAX ability should be able to connect to the MAXWell Lab's WiMAX network. However, each type of device is slightly different, Agrawala notes, so his group is working on customizing connections for different wire-

less devices. As the network runs, his team will monitor and learn from the experience. He adds, "Being a lab environment, we can afford to distort our pictures and create specific kinds of test scenarios."

"In addition to applications we develop, any member of the WiMAX forum or other companies can bring its applications to campus and test them here," Agrawala says. A long list of companies, including Intel, Cisco, XOHM (the arm of Sprint that has established a WiMAX network over Baltimore), Nokia, Motorola, and several smaller companies have contacted Agrawala, as have government labs including the Naval Research Lab, the Laboratory of Telecommunication Sciences, and the National Institute of Standards and Technology. State agencies including those in charge of state highways and the port of Baltimore have also expressed interest in testing applications on the network. So have on-campus offices, including the Office of Information Technology, the Department of Public Safety, and Facilities Management, as well as University Police.

Ron Resnick, the president of the WiMAX Forum told Reuters, "As WiMAX technology is deployed, innovative applications will evolve … Establishing the MAXWell Lab promotes such development and provides access to the tremendous expertise and success in wireless communications and the development of advanced computer applications at the University of Maryland."

In the future, when WiMAX networks and interactive applications are widespread in communities, Agrawala envisions all sorts of conveniences. In a shopping center, a network could guide a shopper to the clothing he wants; in a hospital, a patient's wristband could communicate with computers to help schedule time with medical personnel; in an amusement park, visitors could define the types of rides they like and be guided to them efficiently; one might never need to search for one's car again. "If in the future, this kind of connectivity is available in every place, the kinds of uses you can make are just phenomenal," Agrawala says. He sees the campus as "a reasonable-sized community where we can demonstrate the feasibility of these idesas."

Among the many applications WiMAX networks may have in the future are systems that can guide firefighters and other first responders in emergencies by giving them floor plans and directing them to the specific locations where people need help. Image from the National Institute of Standards and Technology, by Tim McEvoy.

STOP: SOMA Terror Organization Portal for Understanding and Predicting Violent Behavior

Joining forces with social scientists, V.S. Subrahmanian's group has developed a number of analytical tools intended for use by defense agencies. One of the newest is STOP, the SOMA Terror Organization Portal, which allows users to query automatically learned rules about the behavior of potential terrorist groups and to forecast those groups' behavior.

Subrahmanian's group developed the computational side of SOMA, which stands for Stochastic Opponent Modeling Agents, as a language that deduces rules about how people behave in different cultural contexts. SOMA offers predictions about how groups will behave not over fine time scales but over about a year or so, Subrahmanian explains. To develop SOMA, Subrahmanian has collaborated with Jonathan Wilkenfeld, a professor of government and politics, in a study shepherded by Subrahmanian's student Amy Sliva. Sliva is working toward a Ph.D. in computer science and a master's degree in international security and economic policy at the university's school of public policy.

"We essentially provide input into SOMA," Wilkenfeld says about his group's contribution. Wilkenfeld is director of the University of Maryland's Center for International Development and Conflict Management, or CIDCM, which for more than 20 years has maintained a database called MAR, for minorities at risk, about groups around the world that face discrimination. MAR's research director, Amy Pate, is also heavily involved in this joint project. For almost 300 groups, the MAR database tracks variables such as groups' organizational structures and grievances. The data in MAR has primarily been collected by human coders who research and find their own sources of information. Human coding and analysis are very time-consuming. By joining the political science project with computer science, the process of sorting through data is becoming more automated and efficient.

Jon Wilkenfeld and V.S. Subrahmanian

Amy Sliva

Automated reasoning with data

When Subrahmanian, Wilkenfeld, Sliva, and Pate started the project in 2005, well into the post-9/11 world, with the United States at war in Iraq and Afghanistan, Subrahmanian thought his AI and database background could help provide insight on collecting data and reasoning with it. One of his motivations was to explore and expand the range of questions people were addressing. "I want to make sure that we are not missing out on important hypotheses," Subrahmanian says.

At the same time, Wilkenfeld and Pate focused on collecting increasingly fine-grained information about organizations such as Hamas, Islamic Jihad, and PKK, the Kurdish group that aims to establish an independent Kurdish state. "We were trying to understand the differences between those organizations that operated through political means versus those that became violent or became terrorist," Wilkenfeld says. Data collection and analysis had become very time-consuming, he says. "V.S. and I began to talk about how methodologies in computer

science could a) help make us more efficient and b) help us build better models," Wilkenfeld says.

As Subrahmanian and Wilkenfeld were discussing how they could work together, Subrahmanian recruited Amy Sliva, an entering computer science graduate student who was finishing college at Georgetown where she had minored in political science. She started leading the day-to-day operation of the project the summer before her official start as a graduate student at the University of Maryland.

The SOMA Terror Organization Portal helps analysts sort out what factors are associated with an increased likelihood that a group will resort to terrorism. It turns out that those factors include the extent to which a group is associated with religion and the extent of the support a group receives from its diaspora. By looking at multiple such factors, SOMA can develop more reliable predictions about whether a group will engage in terrorism and perhaps help intelligence officers concentrate on the right people. While the predictions don't come with any guarantees, Wilkenfeld says, "If you were a government and interested in which organizations to monitor, then you would do well to pay attention to these findings."

The information fed into SOMA at this point remains manually collected and curated, but SOMA automatically extracts predictions. Besides predicting what groups might become violent, SOMA can forecast groups' behavior over specific periods of time. For example, the SOMA approach revealed that the frequency of Hezbollah attacks against Israel has historically been low during years when the group has been involved in elections in Lebanon. However, Sliva and Subrahmanian emphasize that SOMA isn't an academic exercise intended to develop theories about what motivates and drives various groups; instead, it's meant to predict the behavior of specific groups in the real world. "If I'm a guy on the street, I don't want to know what terrorist groups will do in general but what one will do next year," Subrahmanian says. For example, SOMA has formulated almost 15,000 very specific rules on Hezbollah behavior when it comes to transnational attacks. "The rules themselves can be complex, and there is complex interplay among the rules," Subrahmanian says. "You can't make a prediction based on a single rule, but you can make a prediction based on the totality of the rules." A user can provide data for a list of variables to extract predictions about a group's behavior in the coming year. In practical terms, SOMA can shed light on how outside incentives can help reduce groups' transition into violence.

Work on STOP and SOMA has been funded by the Departments of Defense and Homeland Security as well as by the National Science Foundation. Individuals from a number of organizations have accounts on the SOMA Terror Organization Portal, including the Army, Air Force, Marines, Department of State, the federal contractor SAIC, and others.

The cultural reasoning architecture in SOMA has been applied to problems such as how outside incentives could reduce the cultivation of opium and other drugs in Afghanistan. This photograph shows agents of the U.S. Drug Enforcement Agency burning hashish in Afghanistan. Courtesy of DEA.

Spatial Databases and a Better Way to Calculate Nearest Neighbors in a Road Network

Hanan Samet started working on computational methods for representing spatial information more than 30 years ago, long before it was considered a fruitful area of research and long before interactive on-line maps and GPS devices in cars. Samet was intrigued by the question of how to sort two-dimensional information, and he's pursued that theme ever since. Most recently, he and his students have developed an efficient method for finding nearest neighbors where the objects lie on a spatial network such as a road network. A paper in which Samet and his students described this work was given a best paper award at the 2008 Association for Computing Machinery (ACM) Special Interest Group on Management of Data (SIGMOD) conference.

Precomputing distances

The conventional method for finding, say, the closest pizza restaurant to one's house is to sort the restaurants by the air distance from the house rather than to rank the restaurants by the distance along the shortest actual routes taken to reach them. Programs that then calculate the shortest routes between two locations typically use the Dijkstra method, which was conceived in 1959. This method visits almost every node or road intersection that is reachable via a shorter distance from one's house when trying to find the closest pizza restaurant in terms of distance traveled.

Samet's group developed an entirely different approach for calculating nearest neighbors. "What we did is precompute the shortest paths between all nodes or intersections, which includes the location of the source and the set of objects from which the nearest neighbors are drawn," Samet says. Given a starting location (one's house) and the locations of the set of objects (i.e., the restaurants) from which the nearest neighbors are to be drawn, a program chooses the next node to be visited on the shortest routes to them, which turns out to be a very small set. For example, leaving one's house, one may have a choice of approaching one of four surrounding intersections. In Samet's approach, all the destinations in the city and indeed the world would already be linked to one of the four intersections.

This may sound like a storage-intensive approach to finding the nearest neighbors, since the size of the set of possible destinations, or N, is huge and must be stored for each possible location (which is also a destination). However, it's far less taxing than previous methods because Samet and his group represent the polygonal regions that bound the sets of intersections reachable on the shortest paths from each intersection with quadtrees, which, by appealing to their dimension-reducing property, reduce the storage requirements for the shortest paths for all pairs of destinations, of which there are N^2, from the order of N^3 to N^1.5. This behavior was observed to hold by looking at real road networks. Samet also proved it with a simple 8-line proof. Precomputing all of the shortest paths in order to solve one query certainly isn't an effective approach, but precomputing to set up solutions for many future queries is very efficient. The beauty of the method, Samet notes, is that it is both practical and logically compelling. "My motto is get the right answer, not any answer," he says.

A graphical query language

One of Samet's key insights early on was that to retrieve geographic data it's best to use a query language that is graphical. He and his students developed graphical methods to interact with spatial data as well as appropriate query processing algorithms that only retrieve as much data as is necessary. That enabled them to create "spatial spreadsheets" and "spatial browsers" to make large amounts of geographic information accessible and usable. The systems can quickly process and retrieve large amounts of spatial data and associated non-spatial data according to users' requirements, stopping the computation once just enough data has been obtained rather than unnecessarily executing some queries to completion when only a few answers are needed.

Hanan Samet

Also early on, Samet decided to use quadtrees to represent spatial information. Using quadtrees, two-dimensional information can be sorted in grids of boxes of varying sizes, with a judicious use of additional indexes to organize the contents of each box of the grid. Note that the boxes in a quadtree are denser where objects are more densely packed on a map.

"Quadtree methods are simple and easy to implement," says Samet. Quadtree methods are now widely used by geographic information systems, or GIS, developers at

Google, the United Parcel Service, and by game developers.

"When he started 30 years ago, disk drives were very small, and the concept of putting a map of the entire United States in a car was inconceivable," says Richard Snodgrass, a professor of computer science at the University of Arizona. "Partly because of storage becoming so cheap and partly because of the data structures he developed over the years, we have GPS systems and other devices."

Michael Jones, the chief technologist for Google Earth, says, "Quadtrees are a big part of our database processing. They're a very efficient way of organizing information. Google Earth is the world's most widely used graphics application, and it takes advantage of Dr. Samet's quadtrees."

The spatial spreadsheets that Samet and his group developed can save the results of different queries and layer them on top of each other. The results for one query can be used as the input for subsequent queries. His system gives users flexibility in asking questions and deciding how to ask them.

Spatial browsers developed by Samet's team can be used to access information in spatial spreadsheets. As one example of how a spatial browser can help test hypotheses and make spatial correlations, Samet and his students used a spatial browser to reanalyze epidemiologist John Snow's famous 1854 map showing that cholera outbreaks in a region of London clustered around a single water pump. Comparing the locations of homes of people who died of cholera with the locations of water pumps clearly confirmed that deaths grouped around the suspect pump.

Vinton Cerf, a Turing award winner known as a founding father of the internet and currently employed as the Chief Internet Evangelist at Google, has tracked Samet's work since serving as his Ph.D. thesis advisor. Samet "is a tireless and comprehensive researcher with energy to spare," Cerf says. "He's persistent and determined. He pioneered the field of spatial information management in the form of quadtrees and their derivatives and is the dean of its school of thinking. His most recent book is a magnum opus on the topic, and the application software he has developed maintains the highest fidelity of representation of spatial relationships."

"The theme of my work is sorting in space—or location, location, location," says Samet. The engine behind mapping tools is a sorting mechanism, Samet explains, and he continues to devote himself to building and refining that engine.

A map by John Snow, one of the founding fathers of epidemiology, famously showed that an 1854 cholera outbreak in London clustered around a single water pump. Samet and his students used a spatial browser to reanalyze the data and confirm the findings computationally.

News In Brief News In Brief News In Brief News in Brief

• Ashok Agrawala's Video-911 project received massive press coverage - both in the print and TV media. It was covered by the Diamondback, News 8, the Washington Post, the Washington Times, Tweakers, and Telecompaper.

• Dr. Agrawala was also featured in a Washington Times article on his pioneering MyeVue work.

• The HCIL Open House and Ben Bederson's work was featured in the Washington Post.

• Rama Chellappa is a Fellow of the Optical Society of America. Dr Chellapa was recognized for "pioneering and sustained contributions to image and video based pattern recognition and computer vision".

• Rita Colwell's work on using satellite based methods to map and predict global cholera outbreaks was featured by BBC News.

• Dr. Colwell also received an honorary degree from the University of Oslo. In addition, she gave an invited special lecture at the Royal Society of Medicine in London.

• Larry Davis and grad student Abhinav Gupta's work on describing video content through speech (with applications to baseball) was featured on Discovery and in IEEE Spectrum.

• Dr. Davis and V. S. Subrahmanian were interviewed on the Kojo Nnamdi show. The topic was the future of artificial intelligence.

• Bonnie Dorr, and her students Matthew Snover and Nitin Madnani (in collaboration with Rich Schwartz at BBN Technologies) participated in the first ever NIST Metric MATR workshop to evaluate and compare automatic machine translation evaluation metrics. Their submission, TERp (Translation Edit Rate plus), was evaluated by its ability to automatically predict the quality of a translation. TERp was one of the top performing metrics at the workshop, and had the highest Pearson correlation coefficient, with human judgments in 9 of the 45 test conditions---more than any other metric. In addition, in 33 of the 45 test conditions, TERp was statistically indistinguishable from the top metric---again more than any other. Overall, TERp was consistently one of the best performing metrics in the workshop.

• Lise Getoor and student Mustafa Bilgic received the best student paper award at the 14th ACM SIGKDD International Conference on Knowledge Discovery + Data Mining (KDD)for the paper "Effective Label Acquistion for Collective Classification."

• Mike Hicks and his former student Iulian Neamtiu were quoted in an MIT Technology Review article on a company developing "live" (on-the-fly) updating for Linux.

• David Jacobs' work on Digital Field Guides to enable researchers to identify plants from photos of leaves has been featured in both the NY Times and CNN.

• Dr. Jacobs' work on recognizing trees from iPhone pictures was also described in a CNN Science-Tech blog post.

• Jonathan Katz was one of 12 professors selected to be a member of the DARPA Computer Science Study Panel (CS2P) for 2009. This is a multi-year program, consisting of a funded educational experience to familiarize the participants with DoD practices, challenges and risks, and up to three years of funded research to explore and develop technologies that have the potential to transition innovative and revolutionary computer science and technology advances to the government.

• UMIACS research scientist Ugur Kuter, former UMIACS postdoc Guillaume Infantes, and ISR postdoc Florent Teichteil-Königsbuch's program RFF won the Fully Observable Probabilistic track of the 2008 International Planning Competition.

•Xu Liu, a CS and UMIACS PhD Candidate working with Dave Doermann placed second in the ACM Student Research Competition associated with ASSETS2008 in Halifax. His work on a mobile currency reader for the visually impaired is being beta tested nationwide by the National Federation for the blind, and has qualified for the Grand Finals in April (see

News In Brief News In Brief News In Brief News in Brief

article starting on back page of this publication).

• Doug Oard's work on processing legal documents in connection with the TREC benchmarks has garnered some good press in the ABA (American Bar Association) journal.

• Mihai Pop was highlighted and interviewed on the NSF's Cloud Computing press release.

• The paper "Using Static Analysis to Find Bugs" by Nathaniel Ayewah, Bill Pugh, David Hovemeyer, David Morgenthaler and John Penix was selected by IEEE Software's editorial and advisory boards as one of their 25th-Anniversary Top Picks.

• Louiqa Raschid was named one of this year's 27 ACM Distinguished Scientists.

• Phil Resnilk was interviewed on Federal News Radio discussing cloud computing and its relevance to R&D in language technology.

• Steven Salzberg was interviewed about the swine flu in Mexico on WTOP.

• Dr. Salzberg's work on the cow genome was highlighted in Science News.

• Dr. Salzberg was also quoted in an article in Science Magazine about the FBI investigation of the anthrax attacks on the US Capitol.

• Hanan Samet has been awarded the 2009 University Consortium for Geographic Information Science (UCGIS) Research Award for his research contribution to GIS and for his landmark book *Multidimensional and Metric Data Structures*

• A news article published in Nature has coverage of some of Ben Shneiderman's work (and that of his colleagues) on disaster response systems.

• V.S. Subrahmanian was named a fellow by the American Association for the Advancement of Science (AAAS) in recognition of his contributions in computer science and multidisciplinary computing, for techniques to implement multiple data sources, software programs and automatically build group behavioral models and forecast group behaviors.

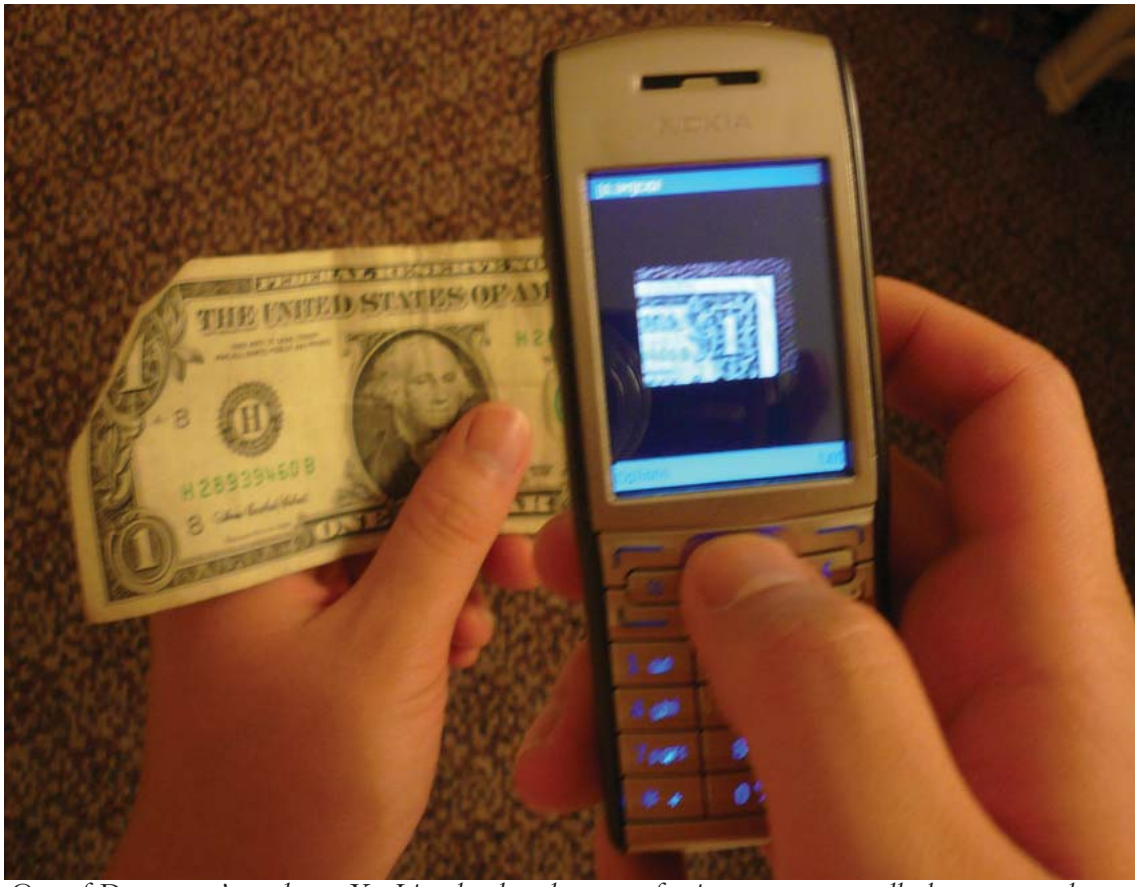
• The paper "Using Histograms to Better Answer Queries to Probabilistic Logic Programs" by Matthias Broecheler, Gerardo Simari and V.S. Subrahmanian has been named the recipient of the Best Student Paper Award at the 2009 International Conference on Logic Programming.

• Amitabh Varshney has been elected as the Chair of the IEEE Visualization and Graphics Technical Committee (VGTC) for 2008 - 2011 term. VGTC provides technical leadership and organization for technical activities in the areas of visualization, computer graphics, virtual and augmented reality, and interaction.

• The University of Maryland hosted a workshop on parallel computing systems on May 29, 2009, titled "Theory and Many-Cores (T&MC): What Does Theory Have to Say About Many-Core Computing?" The workshop was organized by Uzi Vishkin (ECE/UMIACS). The objective of the workshop was to explore opportunities for theoretical computer science research and education in the emerging era of many-core computing, and develop understanding of the role that theory should play in it.

• Min Wu was elected to serve as Vice President for Finance of the IEEE Signal Processing Society (SPS), for a three year term of 2010-2012.

Visual Aspects of Document Analysis



One of Doermann's students, Xu Liu, developed a way of using cameras on cell phones to read currency, an application that is being deployed by the National Federation for the Blind.

Before the Laboratory for Language and Media Processing, or LAMP, was set up in 1996, the community of researchers who analyzed document images had limited interaction with the community of researchers who analyzed text documents, explains David Doermann, co-director of LAMP and an associate research scientist. Established with Department of Defense funding, LAMP brought together researchers at the University of Maryland who analyze the image content of scanned documents with those who analyze their linguistic content.

Document analysis in the real world

Doermann himself has a background in image processing. His mission at LAMP is to analyze large, heterogeneous collections of documents, whether business letters, newspaper articles, or handwritten notes. "A lot of

continued on page 2.

UNIVERSITY OF MARYLAND
Institute for Advanced Computer Studies
2119 A.V. Williams Bldg.
University of Maryland
College Park, MD 20742-3311

Nonprofit Org.
U.S. Postage
PAID
ICM