DECIPHERING HUMAN SCRIPT

School of Engineering and Applied Sciences
University at Buffalo *The State University of New York*
www.eng.buffalo.edu

Mark H. Karwan
Professor and Dean

Computing and IT **is our fourth publication in a six-issue *Igniting Ideas* series highlighting our school's research focus areas. Each area contributes to our strategic objective of performing high-quality research while preparing future researchers for industrial, academic, and government positions.**

Represented in this issue are hardware and software advances that improve the world around us; e.g., our research in machine recognition of characters, biometrics, communications, and cyber security improve our postal systems, security systems, cellular phones, and computing systems. As has been our practice, we have also included noteworthy educational and applications projects related to this issue's theme.

On a personal faculty note, I am pleased to announce that so far this year we have heard that three of our junior faculty—Paul DesJardin and Venkat Krovi of MAE and Hung Quang Ngo of CSE—have been awarded 2004 NSF CAREER Awards. These promising academics follow in a rich tradition of our young faculty, who have won 15 awards since the series was initiated in 1997.

I trust that you will enjoy reading about the accomplishments of our faculty and staff in *Computing and IT.*

Igniting Ideas Series available at: www.eng.buffalo.edu/IgnitingIdeas

Produced by the University at Buffalo Office of Creative Services, University Communications. February 2004. 03-EAS-004.

Talk It Up

Senior-level students in the Department of Computer Science and Engineering who take Software Engineering and other capstone design courses have invented the UB Talker, a product designed to be a working artificial speaking device while providing socially relevant and real-world experiences.

The device uses a touch-screen, menu-driven laptop computer to select words and phrases to help speech-impaired persons communicate. Displays can be preprogrammed or entered in an on-screen keyboard during conversation. The device uses context awareness, phrase completion, and an evolving intelligence that prepares responses in anticipation of choice. Statistics are derived using frequency of use, most recently used, time of day, day of week, and time of year to present most-likely phrases to users. Over time, it adapts to a user's patterns and habits, and becomes conversational and increasingly easier to use.

Speech-impaired persons at two local sites are using the device. One stroke patient who is unable to speak but has some motor skills used the device to make his first phone call in more than 20 years. Colleague UB groups are using the UB Talker as a test and research platform to study natural language synthesis and advanced adaptive technology design.

Leading the effort are Kris Schindler, Michael Buckley, and Helene Kershner of the Department of Computer Science and Engineering.

School of Engineering and Applied Sciences
412 Bonner Hall
Buffalo, NY 14260

Nonprofit org.
U.S. Postage
Paid
Buffalo, N.Y.
Permit No. 311

SOMETHING TO write home ABOUT

Looking Back on Two Decades of Scholarship and Breakthroughs at UB's CEDAR

Twenty years ago, Sargur N. Srihari, SUNY Distinguished Professor and CEDAR executive director, had an idea that handwriting could be recognized by machine. Since then, this idea has blossomed into world-renowned cutting-edge research on pattern recognition, machine learning, and information retrieval technology that is conducted at UB's Center of Excellence for Document Analysis and Recognition.

The center's work has generated such advances that today CEDAR is a leading developer of automated mail whose accomplishments include:

- Software that reads handwritten mail at blur speed—20 pieces per second with an accuracy rate of over 98 percent
- Annual savings to the United States Postal Service (USPS) estimated at $150 million per year

Further evidence of success is sustained funding from USPS over the past two decades and designation of CEDAR as a Center of Excellence by USPS.

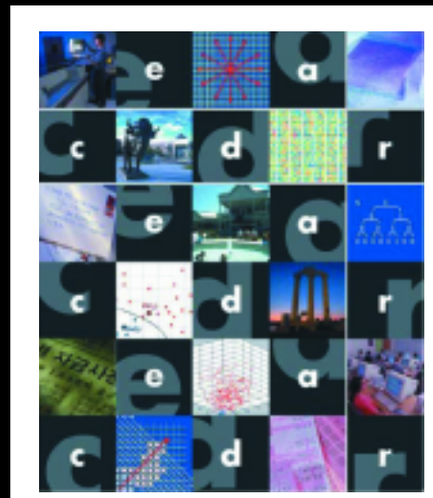
Srihari is joined by CEDAR associate director and professor of computer science and engineering Venugopal Govindaraju and CEDAR's research staff, as well as associate professor of computer science and engineering Rohini K. Srihari.

While the analysis of scanned paper documents, particularly handwritten ones, continues to be CEDAR's core research topic, there are now several research projects involving document types that are "born digital." Applications explored also include document analysis and recognition, forensic document examination, textual information retrieval, biometrics, medical informatics, and biomedical information retrieval.

Selected CEDAR milestones include:

- 1995 - Name and Address Block Reader (NABR) installed at most U.S. Internal Revenue Service processing centers
- 1998 - First Handwritten Address Interpretation (HWAI) technology deployed by the USPS
- 1999 - Handwriting systems incorporated by Lockheed Martin in systems for Australia Post and the United Kingdom Royal Mail
- 2002 - Handwriting Individuality Analysis work cited, leading to the acceptance of handwriting evidence in U.S. federal courts
- 2002 - Image Evaluation System deployed by USPS to evaluate recognition systems for letter mail and subsequently for flats and parcels
- 2003 - CEDAR-FOX, a system for questioned document examination and handwriting processing, is made available to the law enforcement community

CEDAR has been a conduit for scholarly activity, as well as individual accomplishment for scientists and students alike. Research scientists at CEDAR hold 10 registered patents and have published more than 500 technical papers. The organization has supported more than 500 students, resulting in the award of several hundred master's degrees and more than 40 doctoral degrees.



Leading the Pack: More and More Wired at UB

UB, already listed among *Yahoo*'s "Top Wired" campuses in the U.S., continually seeks innovative ways to evolve its computing and IT infrastructure.

Over the past few years, the school has been involved in university initiatives to establish UB as one of the largest academic supercomputing facilities in the nation. Several projects, including the installation of Dell, IBM, and Sun machines in the Center for Computational Research, have helped UB accomplish this goal. UB distinguished professor of computer science and engineering and CCR director Russ Miller leads the group.

At the school level, a major state initiative funded the creation of the New York State Center for Engineering Design and Industrial Innovation. This center, a top-end visualization facility led by Christina Bloebaum, professor for competitive product and process design in the mechanical and aerospace engineering department, is heavily involved in industrial and scientific research.

Another recent project constructed a remote sensing satellite ground station on campus to collect environmental field data from around the world. A cooperative project involving faculty from engineering, geography, and geology, this resource will be a valued campus-wide tool for research and instruction. Charles Brunskill, director of Science and Engineering Node Services, joined project director Mohammed Sultan, professor of geology; Joseph Atkinson, professor of civil, structural, and environmental engineering; and Chris Renschler, assistant professor of geography.

The school's latest student-oriented venture is a system that allows students to create electronic portfolios of their academic activities throughout their tenure at the university and show prospective employers or graduate schools their selected work on a CD or Web site. Brunskill and SENS staffer Molly Ives Brower are implementing this system with Sandy Peters, associate director of UB academic services, and instructional designer Roberta Sullivan.

Revealing the Real You

The UB School of Engineering and Applied Sciences has just created the Center for Unified Biometrics and Sensors (CUBS) to foster the development of radically different biometric technologies geared toward developing applications to improve comfort, convenience, and security for personal and commercial use. The center establishes a unified biometric framework to facilitate the development of next-generation biometric systems. This framework will allow evaluation of market needs and constraints in choosing the biometric technologies best suited for a particular application. Researchers will develop new biometric technologies from proof-of-concept to product-readiness, including usability studies and educational outreach to evaluate and mitigate ethical and legal concerns.

Since the terrorist attacks of 9/11, physical biometrics—such as an individual's height, weight, shape of the iris in the eye, vein structure, and hand geometry—have become increasingly important for security applications because they cannot be easily faked. This technology will also reach to industrial and forensic applications.

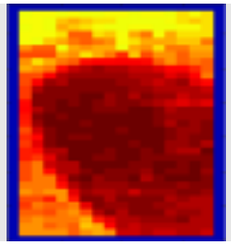
CUBS leverages researchers from UB's SEAS, College of Arts and Sciences, School of Informatics, and School of Medicine and Biomedical Sciences to form a multidisciplinary team to address some of our country's most critical needs. Venu Govindaraju, professor of computer science and engineering, directs the new center.

Exciting projects on the research agenda include:

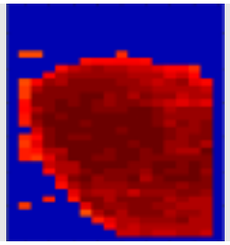
- Pathobiometric systems that can automatically track illnesses in livestock, such as mad cow disease, by analyzing aerial images of large cattle herds, along with other farm data
- A state-of-the-art automated fingerprint identification system
- Biometrics systems modeled after the methods that animals use to understand their environments

With funding from the National Science Foundation, the New York State Office of Science, Technology, and Academic Research (NYSTAR), and several companies, CUBS will bridge academia and industry in designing, developing, and prototyping biometric devices for commercialization.

Original Image



Energy Map



Threshhold Map

A Fresh Approach to Genomic Data Mining

A truly multidisciplinary team of researchers has been formed at UB to develop novel genomic data-mining algorithms. Led by principal investigator Aidong Zhang, professor of computer science and engineering, and co-PIs Maurizio Trevisan, dean of the UB School of Public Health and Health Professions, and William Jusko, professor of pharmaceutics, the project is laying the groundwork for the establishment of a National Institutes of Health biomedical computing center. Other investigators include Jan Chomicki, Jian Pei, and Xin He, computer science and engineering; Richard Almon, biology; Murali Ramanathan, pharmaceutics; Yulan Liang, School of Public Health and Health Professions; and Steven Pruitt and Ping Liang, Roswell Park Cancer Institute.

The team will provide expertise and infrastructure to merge research activities of computational and biomedical scientists. Specifically, they will develop specialized data-mining algorithms to analyze patterns and relationships in gene-expression data. They will also investigate novel approaches to conduct pattern measurements, cluster analysis and validation, time series analysis, and integrated dimensional analysis, as well as study the relationship between clustering and gene regulation.

The development of techniques to effectively analyze genomic datasets is a critical step in work on the medical applications of bioinformatics. The synergistic merging of computer science expertise and biomedical expertise is expected to advance biomedical research to the next level. Working side by side, with a clear understanding of the research approaches each uses, computational and biomedical scientists will be in a position to make the critical contributions necessary to reap the benefits of new treatments for common diseases with unmet medical needs.

Top 20 in Funding

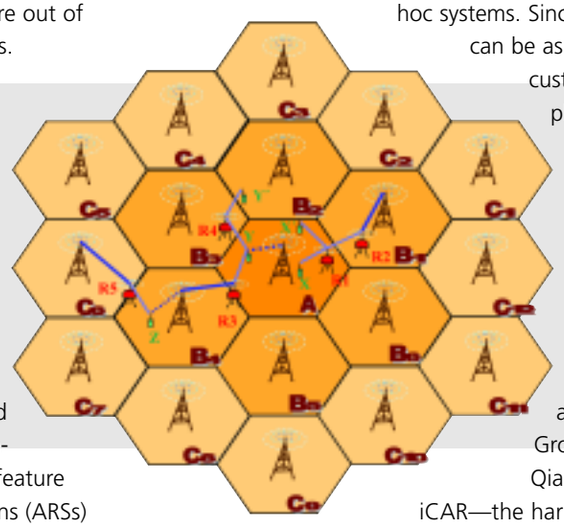
The UB Department of Computer Science and Engineering, with 24 faculty, ranks 20th in R&D expenditures (approx. $7.5 million yearly) among 94 Ph.D.-granting computer departments in U.S. public universities and colleges, using the most recent available NSF data. UB ranks 16th on a per-faculty basis.

Hope for Tangled Lifelines

A cell phone can be a lifeline in an emergency—unless calls cannot be completed when lines get jammed at peak times or if users are out of range of repeater stations in coverage areas.

To the rescue comes CSE professor Chunming Qiao, whose pioneering work, called iCAR (**i**ntegrated **C**ellular and **A**d hoc **R**elaying system), focuses on how to make high-speed wireless connectivity available virtually anywhere, anytime.

iCAR was introduced to address the specific problem of having limited data-rate and scarce bandwidth resources in today's cell phone systems, coupled with the problem of having limited coverage and mobility support using wireless ad hoc technology, such as WLAN (802.11). A unique feature of iCAR is the use of ad hoc relaying stations (ARSs) that relay voice or data calls from a location where the downlink quality (e.g., signal strength or bit rate) is low, because either the cellular coverage is poor or the area is currently congested with many other calls, to a nearby location with a much better downlink quality. Results obtained by Qiao and his team showed that iCAR can effectively improve the performance of a conventional cellular system. The concept of iCAR can also be applied to extending the coverage and improving the mobility support and scalability of wireless ad hoc systems. Since each ARS can be as simple as a custom-made cell phone, it is portable and can be quickly deployed and effectively used as a part of a rescue operation in a disaster area (e.g., Ground Zero).



Qiao's work on iCAR—the harbinger of the push toward convergence between cellular and wireless LAN technologies—has been featured in such magazines as *Business Week* and *Wireless Europe,* as well as on international Web sites, including Canadian Broadcasting Corporation (CBC) and New Scientists (UK). The work is funded in part by the National Science Foundation's ITR program and by the Nokia Research Center.

Guarding Against Cyberattacks

The UB Center of Excellence in Information Systems Assurance Research and Education (CEISARE)—designated a Center of Excellence by the National Security Agency—conducts research to protect computing networks from cyberattacks, and trains computer scientists and engineers in these techniques.

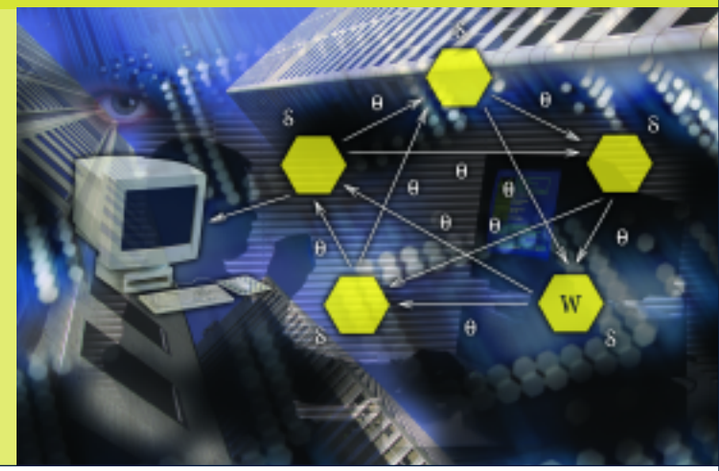
CEISARE, under the direction of associate professor of computer science and engineering Shambhu Upadhyaya, is a leader in academic homeland security efforts. The center's activities include:

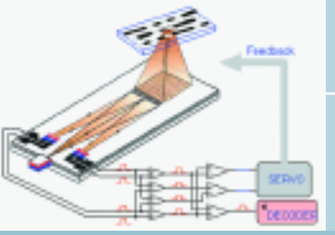
- Promoting and conducting multidisciplinary research in Information Assurance (IA) and cyberwarfare
- Collaborating with companies throughout the nation engaged in security research
- Educating and training the next generation of IA specialists

A blend of engineers, scientists, mathematicians, and management and legal specialists, the group addresses technical issues in cybersecurity, as well as legal implications.

The center has received generous financial support from the National Security Agency, the Defense Advanced Research Projects Agency, the Air Force Research Labs, the Information Institute at Rome, New York, and Telcordia Technologies of Morristown, New Jersey. Recent funding includes two research grants from NSA/ARDA: Data Fusion and Intrusion Detection, and Protecting Documents from Insider Threats; a third recent grant from the Department of Defense provides for two new and two returning student scholarships, funding that comes with the center's NSA designation.

Much of CEISARE's research is conducted in the affiliated SPIDER—**S**ecurity, **PrI**vacy, **DE**pendability **R**esearch Lab, also directed by Upadhyaya.





Award-winning Applets

Professor of electrical engineering Chu Wie's semiconductor applet service was recognized with a 2003 MERLOT (Multimedia Education Resource Learning and Online Teaching) Classics Award. His site provides more than 60 simulations, animations, and tutorials applets on semiconductors and device physics, as well as semiconductor device manufacture and operation. Topics range from crystal structure and electronic energy bands through device fabrications to circuit design and simulation. The applets provide an interactive environment that allows students to explore effects of changing parameters on the operation of devices.

Leading academics and practitioners have cited Wie's applets as being a great help in understanding semiconductor devices. The applets are being used in undergraduate instruction throughout the nation, including at the University of Notre Dame, the University of Texas at Austin, and Virginia Polytechnic Institute and State University. Wie's work has been noted in numerous publications, including *Science,* and may be seen on the Web at jas.eng.buffalo.edu.

Science in the Classroom

Teacher's Page

There's a new resource in the ongoing search for ways of presenting technology to pre-college students in an interesting and understandable fashion. Now high school teachers can direct their students to a Web site designed to help them learn scientific principles and topics that many students find elusive.

Called "Web-Based Technology Education," the project is a creation of David Shaw, professor of electrical engineering and head of UB's Center for Innovative Engineering Learning, in cooperation with colleagues from the College of Engineering and Applied Sciences at Stony Brook.

The goal of the project is to produce instructional modules to address the newly published national standards for technological literacy that specify content for the study of technology for high school students. School districts in parts of upstate New York and Long Island are participating in development and evaluation of the site's materials. At this time, modules exist for Music in the Digital World, Drying by Design, and Safety Light Systems (see www.eng.buffalo.edu/Shaw).

The project is financially supported by a grant from the National Science Foundation.

Acclaimed Algorithm

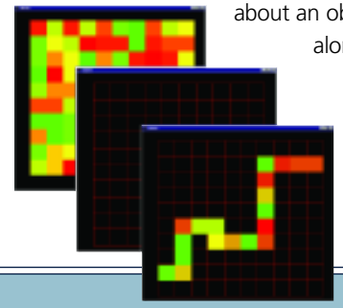
Research by UB distinguished professor of computer science and engineering Russ Miller, director of the UB Center for Computational Research, with Nobel Laureate Herbert Hauptman and Charles M. Weeks, both of the Hauptman-Woodward Medical Research Institute, resulted in a direct-methods procedure called "Shake-and-Bake" that is listed on the poster "The Top Ten Algorithms of the 20th Century" by *Computing in Science and Engineering.*

Miller was listed on *HPCWire*'s 2003 list of Top People and Organizations to Watch in high-performance computing.

Human Engineering Information

Human decision making under uncertainty includes work on computer and display technologies to aid decision makers (e.g., fatigued pilots interpreting computerized cockpit displays). Ann Bisantz, associate professor of industrial engineering, is investigating properties of graphical representations of uncertainty based on blurred or degraded icons, as well as visual, auditory, and tactile representations of spatially distributed uncertainty.

Her research reveals that people are able to successfully map iconic representations to underlying concepts of uncertainty, and that performance on dynamic decision-making tasks using such representations is similar to that of numeric representations. Further work explored participants' interpretation of icons by empirically generating fuzzy membership functions that mapped their interpretation of the icons' meaning to probabilities. These results, too, indicated that people generated membership functions with maximal values closely correlated to the intended numeric values; also, membership functions were reasonably similar across individuals. In combination, these are important findings because they suggest that such representations may be implementable and allow display designers to use a single icon to encode the uncertainty about an object's identity (e.g., hostile/friendly aircraft) along with other dimensions, such as location.



This work is supported by Bisantz's NSF CAREER grant and involves collaboration with the UB Center for Multisource Information Fusion and UB colleagues James Llinas and Thenkurussi Kesavadas.

Unlocking Knowledge

"We are drowning in information but starved for knowledge. This level of information is clearly impossible to be handled by present means. Uncontrolled and unorganized information is no longer a resource in an information society, instead it becomes the enemy."

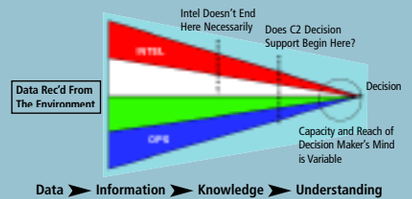
—John Naisbitt, *Megatrends* (1980)

This was true when it was written and is even more critical today. Addressing Naisbitt's "enemy," researchers at UB's Center for Multisource Information Fusion analyze monolithic data sets to identify patterns and seek meaning from massive information. Their focus is on multiple-source information processing environments, such as in multiple-sensor or multiple-instrumented systems. These environments occur frequently in defense applications for advanced surveillance and reconnaissance systems, as well as in robotics, civil infrastructure systems, medical monitoring systems, intrusion detection systems, intelligent transportation systems, and environmental monitoring applications.

Center accomplishments include:

- **Significantly improved ground-object tracking from combining multiple-sensor observations and a priori terrain and cultural-feature data**
- **More accurate and complete situational information on civil disaster conditions, derived from multiple field reports, to aid in casualty rescue and improved management of emergency-response equipment and vehicles**
- **Definition of a new and formally grounded methodology for the testing and evaluation of data fusion-capable aircraft**

This highly multidisciplinary technology is also applicable to a wide array of nonmilitary problems. For example, the center is conducting research with Penn State in condition-based maintenance, making maintenance scheduling more cost effective by integrating sensors into the equipment. An NIH project is on multispectral mammography.



Other grant sponsors include the U.S. Air Force Office of Scientific Research, the Office of Naval Research, the U.S. Army, and the National Security Agency. Top industrial partners are Lockheed-Martin and Boeing.

James Llinas, ATA professor of industrial engineering, is director of the center.

Going Grid

As the scientific community develops a national tera-scale infrastructure for high-performance computing, Bina Ramamurthy and Bharat Jayaraman, CSE research assistant professor and professor and chair, respectively, are participating in NSF's Course, Curriculum, and Laboratory Improvement Program to introduce grid computing concepts and to optimize existing computing and data resources.

Companies are eagerly embracing grid computing and promoting it under various names (e.g., utility computing, on-demand computing). Scientists and practitioners believe the grid developed for scientific computing will one day make computing freely available as yet another "utility," as accessible as the electrical power grid and the telephone grid.

The researchers are developing GridForce—Grid For Research Collaboration and Education, sponsored by NSF—a network of computational units cooperating to share compute cycles, data, and other resources using an open and standardized service-based framework.

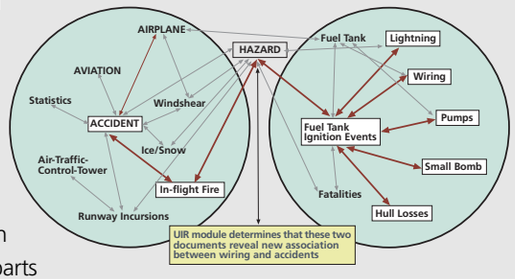
GridForce seeks to promote grid awareness and technical readiness in the workforce. Accomplishments so far include:

- **Introducing grid computing in the department's curriculum**
- **Building laboratory prototypes to support grid application development in the courses**
- **Conducting workshops for industry**

Future projects include grid-based program execution environments and a grid-based anomaly detection system.

Concept Chain Graphs: A Hybrid Framework for Text Mining

Extracting meaning from a Web site and linking it with information from related Web locations is the focus of work by Rohini K. Srihari, PI and CSE associate professor, and Miguel Ruiz, co-PI and assistant professor in the UB School of Informatics. They are analyzing documents to determine if the information content of the sum of the parts is greater than that of the individual parts. For example, could viewing different sites on related subjects yield a synergistic overview that could compromise security?



An overlooked consequence of multiple documents by multiple sources could be the unintended accessibility of critical security information, a special case of text mining referred to as unapparent information revelation.

While information analysts currently use search engines to sift through extensive document collections for such links, automated tools are far more efficient in exposing them, or at least generating plausible concept chains. In this scenario, a user purposefully surfs a Web site with the intention of exposing a particular link, or chain of links, between key concepts. The challenge is to decipher his ultimate information need based on his queries. For example, is he trying to gain information on how to sabotage aircraft?

The proposed solution combines the robustness and efficiency of information retrieval systems models with the higher granularity of information extraction models. Search engines represent document content by counting the frequency of specific words. Information extraction systems focus on entities, relationships between entities, and events. This project focuses on building concept chain graphs combining both systems in a probabilistic framework. Text mining is achieved by automatic detection and visualization of interesting paths and subgraphs within the concept chain graph.

The group has processed 15,000 FAA Web pages and constructed an initial concept chain graph. Next they will showcase a tool whereby the graph can be visualized, and connections between critical concepts can be viewed in terms of the documents and text snippets that connect them.

The National Science Foundation and the FAA currently fund their research.