For more than fifty years, the Department of Computer Science at the University of Illinois has quietly led a revolution that has redefined the meaning of computing. Our students and faculty have designed and built the world’s fastest computers, created the user interfaces that popularized the World Wide Web and helped make distributed collaboration possible, invented the compilation techniques for automatic programming parallelization, co-founded the field of computer arithmetic, discovered new techniques for numerical solution of stiff ordinary differential equations, and carried out seminal work in parallelizing compilers.

The department offers an array of powerful computers and software for instruction, as well as research laboratories maintained by individual faculty. All systems are connected to a high-speed network with multiple wireless networks also available. The department also provides computer clusters, printers, file services, and other technology services for all of its users. In early 2004, the department relocated to the Thomas M. Siebel Center for Computer Science, which serves as a living laboratory for exploring and evaluating 21st century computing environments.

Current research areas include the following: algorithms and theory; artificial intelligence in the areas of machine learning, vision, and robotics; computer architecture and compilers; databases and information systems; graphics and visualization; human–computer interfaces and social computing; networking; operating systems and security; parallel processing; programming languages, formal systems, and software engineering; real-time and embedded systems; bioinformatics; and scientific computing.

The department is also home to the Center for Simulation of Advanced Rockets (CSAR) and the Center for Advanced Research in Information Security (CARIS). It has many collaborative ties with units throughout campus, including the National Center for Supercomputing Applications (NCSA), Beckman Institute for Advanced Science and Technology, The Illinois Trust Institute (ITI), the Institute for Genomic Biology (IGB), and the Coordinated Science Laboratory (CSL).

M. Snir was department head during the reporting period.

Faculty and Their Interests

Vikram Adve
Compilers, software reliability, performance analysis, computer architecture

Sarita Adve
Computer architecture, low-power design, adaptive systems, real-time and network processing, performance evaluation methods, parallel computing

Gul A. Agha
Developing new abstractions for building open distributed systems and reasoning about their behavior, parallelism, coordination, real-time behavior

Eyal Amir
Automated reasoning and machine learning

Brian P. Bailey
User interface tools that better support early design tasks, systems and environments that help users maintain information awareness, tools for multimedia authoring and design, interfaces that foster social interaction, human–computer interaction

Geneva G. Belford, Emeritus
Databases and information systems, distributed systems

Stephen Bond
Numerical analysis and scientific computing, with applications in statistical mechanics, and biochemistry; understanding methods that bridge the temporal and spatial scales in multiscale biomolecular modeling using techniques from geometric integration and adaptive finite element methods

Marco Caccamo
Real-time operating systems, real-time scheduling and resource management, wireless sensor networks, quality of service control in next-generation digital infrastructures
Roy H. Campbell
Security, distributed operating systems, ubiquitous computing

Kevin C.-C. Chang
Databases, Internet information access, and digital libraries, with focuses on information integration of heterogeneous sources, Internet query processing, web databases, and ranked top-k query processing

Chandra Chekuri
Design and analysis of algorithms, combinatorial optimization, approximation algorithms, scheduling, graphs, networks

Gerald DeJong
Artificial intelligence

Eric deSturler
Iterative methods, eigenvalue problems, large-scale optimization

AnHai Doan
Databases, data integration and sharing, data mining, information discovery on the Web, efficient use and maintenance of meta-data, schema matching, machine learning

Jeffrey G. Erickson
Algorithms, data structures, and lower bounds; computational and discrete geometry

Margaret Fleck
Automated reasoning and machine learning

David Forsyth
Artificial intelligence, computer vision, machine learning

Michael Garland
Computer graphics, geometric modeling, human–computer interaction, visualization

Maria J. Garzaran
Compilers, software reliability, parallel computer architecture, thread-level speculation

William D. Gropp
High performance scientific computing, with particular emphasis on parallel computing

Elsa Gunter
Formal systems

Carl A. Gunter
Security, networks, software engineering, programming languages

Indranil Gupta
Distributed systems, distributed protocols, probabilistic protocols, design methodologies, sensor networks

Jiawei Han
Database systems, data mining, data warehousing, stream data mining, web mining, spatiotemporal data mining, biodata mining

Mehdi T. Harandi
Artificial intelligence, information systems, HCI, software engineering

Sariel Har-Peled
Algorithms, data structures, computational geometry, clustering, learning, computer graphics

Luddy Harrison
System architecture

John C. Hart
Computer graphics, computational topology

Michael T. Heath
Scientific computing, parallel computing

Anil N. Hirani
Numerical analysis, discrete exterior calculus, differential geometry, computational mechanics, computational astrodynamics

Julia Hockenmaier
Computational linguistics, computational biology, automated reasoning, machine learning, natural language processing

Jennifer C. J. Hou
Design, analysis, and implementation of self-adjusting protocols for wireless networks, enabling technology for large-scale network simulation and emulation, wireless-enabled cyber physical space for healthcare, design, fundamental property analysis, evaluation of data-centric wireless sensor networks

Sheldon H. Jacobson
Operations research, discrete optimization, heuristic design and analysis, applied probability, health care, aviation security, homeland security
Ralph E. Johnson
Object-oriented design, design patterns, frameworks, software architectures

Laxmikant V. Kale
Numerical, parallel, and scientific computing, operating systems

Samuel N. Kamin
Programming languages, software components, functional programming applied to scientific computation, denotational semantics, program specification and verification, domain-specific languages

Karrie Karahalios
Human–computer interfaces

Samuel T. King
Security, operating systems, experimental software systems, virtual machines

Negar Kiyavash
Digital rights management, computer, communication and multimedia security, cryptography

Robin Kravets
Mobile computing and communication, location management, power management, transport protocols, ad hoc networks, personal area networks

Steven M. LaValle
Robotics, motion planning, computational geometry, artificial intelligence, computational biology, computer vision, computer graphics, control theory

Haiyun Luo
Networking and distributed systems

Darko Marinov
Software engineering, programming languages, software testing

Jose Meseguer
Formal executable specification and verification; software composition, reflection, and metaprogramming; object-oriented specification and software architecture; concurrent, distributed, and mobile computing; logical frameworks and formal interoperability; logical and semantic foundations

Klara Nahrstedt
Quality-of-Service (QoS) management, integration of guaranteed and best effort services for audio/video/DATA traffic, QoS-aware resource management, QoS routing, multimedia security, soft real-time scheduling, middleware support for distributed multimedia applications

Luke Olson
Numerical partial differential equations (PDEs), numerical linear algebra, high-performance computing

David A. Padua
Computer architecture and systems, parallel computing, compilers

Madhusudan Parthasarathy
Software engineering, formal methods

Lenny Pitt
Artificial intelligence, theoretical computing

Jean A. Ponce
Computer vision, robotics, computer graphics

Manoj M. Prabhakaran
Cryptography, other topics in theoretical computer science

Grigore Rosu
Software and software related aspects; design, semantics, and implementation of programming and specification languages; automated software engineering and formal methods, especially “push-button” techniques for certification, monitoring, synthesis, and modularization; automated reasoning about computer systems, applications of logics, theorem proving; algorithms, (co)algebra, category theory

Dan Roth
Artificial intelligence, theoretical computing

Lui Sha
Distributed real-time computing systems, dynamic real-time architecture, Quality-of-Service (QoS) driven resource management, security and fault tolerance in networked embedded systems

Saurabh Sinha
Gene regulation, comparative genomics, sequence analysis, evolution
Marc Snir
Large-scale parallel and distributed systems, parallel computer architecture, grid computing, parallel programming

Josep Torrellas
Parallel and sequential computer architecture, processor-memory integration, thread-level speculation, low power design, reliability

Mehesh Viswanathan
Analysis and validation of software systems, including design of efficient algorithms, characterization of computational limitations, development of formal models for system specification, and implementation of software tools for program analysis

Marianne S. Winslett
Databases, security, parallel computation

Yizhou Yu
Data-driven graphical methods, computer animation, mesh editing, image and video processing

Chengxiang Zhai
Text processing and management, statistical natural language processing, machine learning, bioinformatics

Yuanyuan Zhou
Operating systems, file and storage systems, architecture, distributed systems, parallel systems, system support for database

Craig Zilles
Computer architecture, dynamic optimization, compiler construction, simulation methodologies, computer science education

Artificial Intelligence: Machine Learning, Vision, and Robotics

Autonomous Car Project
Q. Zhang (Ag. & Biol. Engr.)
University of Illinois Department of Computer Science; University of Illinois College of Engineering; University of Illinois Department of Electrical & Computer Engineering

This project will equip a stock car with sensors and actuators and will integrate them through software and special-purpose hardware. The result of the project would be an autonomous car that could drive in urban situations and could perform tasks such as driving from one place to another in town, avoiding obstacles on the road, adhering to the proper lanes of traffic, stopping in intersections as needed, turning and proceeding in intersections while not risking approaching cars, passing slow-moving vehicles, and parking. The resulting car would compete in the DARPA Urban Challenge in November 2007. The resulting technology could revolutionize the way we live, with potential for fewer traffic accidents, cheaper services (e.g. transportation of goods, street cleaning) and others.

Goal Achievement in Partially Known, Partially Observable Domains
E. Amir,* A. Chang
Defense Advanced Research Projects Agency (DARPA), HR0011-05-1-0040

We present a decision-making algorithm for agents that act in partially observable domains that they do not know fully. Making intelligent choices in such domains is very difficult because the effects of actions may not be known a priori (partially known domain), and features may not always be visible (partially observable domain). Nonetheless, we show that an efficient solution is achievable in STRIPS domains by using traditional planning methods. This solution interleaves planning and execution carefully. Computing each plan takes time that is linear in the planning time for the fully observable, fully known domain. The number of actions that it executes is bounded by a polynomial in the length of the optimal plan in the fully observable, fully known domain. Our theoretical results and preliminary experiments demonstrate the effectiveness of the algorithm.

* Denotes principal investigator.
Knowledge-Based Learning Approach to Cognitive Computing
E. Amir,* J. Dejong
Defense Advanced Research Projects Agency, HR0011-05-1-0040; Information Processing Technology Office

The brittleness of first-order symbolic logic has lead to the ascension of statistics as the dominant paradigm in AI. But the statistical paradigm has been unable to support the componential knowledge and the rich combinatorial inference of symbolic logic. Successful cognitive computing will need to combine the power of logic with the robustness that underlies statistics. The research in this project explores approaches to knowledge representation and automated reasoning that employ the inferential power of first-order logic but are supported by a new semantics that embraces the real-world robustness of statistical inference. The approach that this project develops follows a new form of explanation-based learning, and acquires both EBL concepts and partitioned axiom sets. If successful, we believe the research will lead to new algorithms that efficiently appreciate and adapt to their deployment contexts. They will learn quickly to supply a new user with what is most needed to optimize their joint problem-solving behavior.

Learning Partially Observable Action Models: Efficient Algorithms
E. Amir,* A. Chang, D. Shahaf
Defense Advanced Research Projects Agency (DARPA), HR0011-05-1-0040; DAF Air Force Research Laboratory Award, FA8750-04-2-0222; University of Illinois

We present tractable, exact algorithms for the effects of learning actions and preconditions in partially observable domains. Our algorithms maintain a propositional logical representation of the set of possible action models after each observation and action execution. The algorithms perform exact learning of preconditions and effects in any deterministic action domain. This includes STRIPS actions and actions with conditional effects. In contrast, previous algorithms rely on approximations to achieve tractability, and do not supply approximation guarantees. Our algorithms take time and space that are polynomial in the number of domain features, and can maintain a representation that stays compact indefinitely. Our experimental results show that we can learn efficiently and practically in domains that contain over thousands of features (more than 21000 states).

MPE and Partial Inversion in Lifted Probabilistic Variable Elimination
E. Amir,* D. Roth, R. Braz
Cycorp; Advanced Research and Development Activity (ARDA) Advanced Question Answering for Intelligence (AQUAINT) Program; National Science Foundation, ITRIIS-0085980; Defense Advanced Research Projects Agency (DARPA), HR0011-05-1-0040

It is often convenient to represent probabilistic models in a first-order fashion, using logical atoms such as partners (X; Y ) as random variables parameterized by logical variables. (de Salvo Braz, Amir, & Roth 2005), following (Poole 2003), give a lifted variable elimination algorithm (FOVE) for computing marginal probabilities from first-order probabilistic models (belief assessment, or BA). FOVE is lifted because it works directly at the first-order level, eliminating all the instantiations of a set of atoms in a single step, in some cases independently of the number of these instantiations. Previous work could treat only restricted potential functions. There, the instantiations of atoms cannot constrain each other: predicates can appear at most once, or logical variables must not interact across atoms. In this paper, we present two contributions. The first one is a significantly more general lifted variable elimination algorithm, FOVE-P, which covers many cases where atoms share logical variables. The second contribution is to use FOVE-P for solving the most probable explanation (MPE) problem, which consists of calculating the most probable assignment of the random variables in a model. The transition from BA to MPE is straightforward in propositional models, but the lifted first-order case is harder. We introduce the notion of lifted assignments, a distribution of values to a set of random variables rather than to each individual one. Lifted assignments are cheaper to compute while being as useful as regular assignments over that group. Both contributions advance the theoretical understanding of lifted probabilistic inference.

Scaling Up First-Order Logical Reasoning with Graphical Structure
E. Amir,* D. Ramachandran, I. Gammer
eyal@uiuc.edu
National Science Foundation, IIS 05-46663

We present polynomial-time algorithms that translate first-order logic (FOL) theories to smaller propositional encodings than achievable before in polynomial time. For example, we can sometimes reduce the number of propositions to $O(|P| + |C|)$, or $O(|P|k \cdot \log |P|)$, for $|P|$ predicates of arity $k$ and $|C|$ constant symbols. The guarantee depends on availability of some graphical

* Denotes principal investigator.
structure in the FOL representation. Our algorithms accept all FOL theories, and preserve soundness and completeness (sometimes requiring the domain closure assumption). Our experiments show significant speedup in inference with a SAT solver on real-world problems. Our results address a common approach that translates inference and decision problems that originate in FOL into propositional logic, later applying efficient SAT solvers. Standard translation techniques result in very large propositional encodings ($O(|P||C|k)$ for predicates of arity ($k$) that are often infeasible to solve. Our approach scales up inference for many objects, and has potential applications in planning, probabilistic reasoning, and formal verification.

**Domain Knowledge, Explanation-Based Control, and Reinforcement Learning**
G. DeJong,* A. Laud, Q. Sun, V. Moskovich
*Office of Naval Research, N00014-01-1-0063*

Prior research has shown that complex skill-like decision policies can be acquired by combining an inferential symbolic reasoning with a numeric component. The approach is called Explanation-Based Control. Reinforcement Learning (RL) is a popular machine learning approach that also automatically acquires skill-like policies. However, it is difficult or impossible to exploit a domain expert’s knowledge. This means that RL cannot learn complex policies in an example-efficient manner. This research explores combining the two approaches to achieve greater example efficiency on the practical side and a clearer conceptual foundation theoretically.

**Reasoning about Partially Observed Actions**
M. Nance,* E. Amir, A. Vogel
*DAF Air Force Research Laboratory, FA8750-04-2-0222*

Partially observed actions are observations of action executions in which we are uncertain about the identity of objects, agents, or locations involved in the actions (e.g., we know that action move [$?o$, $?x$, $?y$] occurred, but do not know $?o$, $?y$). Observed-action reasoning is the problem of reasoning about the world state after a sequence of partial observations of actions and states. In this paper we formalize observed-action reasoning, prove intractability results for current techniques, and find tractable algorithms for STRIPS and other actions. Our new algorithms update a representation of all possible world states (the belief state) in logic using new logical constants for unknown objects. A straightforward application of this idea is incorrect, and we identify and add two key amendments. We also present successful experimental results for our algorithm in Blocks-world domains of varying sizes and in Kriegspiel (partially observable chess). These results are promising for relating sensors with symbols, partial-knowledge games, multi-agent decision making, and AI planning.

**ITR: An Integrated Approach to 3-D Photography Using Shape, Texture, and Motion Cues**
J. Ponce*
*National Science Foundation, IIS-03-12438; Beckman Institute for Advanced Science and Technology*

The ability to acquire realistic visual models of complex 3-D scenes from collections of images, known as 3-D photography, is becoming useful for a wide array of IT applications, including film, game, and Web-content production, TV advertising, electronic commerce, teleconferencing, and architectural walkthroughs. This project addresses some of the scientific challenges to this technology. This would involve advances in three research areas: construction of topological mesh models of complex surfaces from weakly-calibrated photographs; automated matching and registration of photographs of textured surfaces taken from different viewpoints; and development of projective and Euclidean structure-from-motion techniques capable of handling large numbers of images in an efficient and uniform manner.

**Toward Category-Level Object Recognition**
J. Ponce,* Y. LeCun (New York Univ.)
*National Science Foundation, IIS-0535152/0535166*

In cooperation with New York University, Courant Institute

This project addresses the problem of category-level object recognition in images. Our aim is to develop effective methodologies for three purposes: representing object classes (a person, a car, a vegetable); learning the corresponding object models from cluttered sample images in a semi-supervised manner (i.e., the images themselves are labeled by the names of the objects they contain, but individual pixels and/or image features are not); and efficiently and robustly recognizing instances of these models in novel images despite clutter, occlusion, viewpoint and illumination changes, and individual variations within each class.

**Toward True 3-D Object Recognition**
J. Ponce*
*National Science Foundation, IIS-03-08087*

This project addresses four fundamental instances of the 3-D object recognition problem: modeling rigid 3-D objects from a small set of unregistered pictures and

* Denotes principal investigator.
recognizing them in cluttered photographs taken from unconstrained viewpoints; representing and recognizing nonuniform texture patterns under nonrigid transformations; modeling and recognizing articulated objects in image sequences, with applications to the identification of shots that depict the same scene (shot matching) in video clips; and learning and recognizing part-based descriptions of object classes in photographs and video clips.

**Large-Scale Temporal Associative Memory**
S. R. Ray,* S. Swarup
*University of Illinois*

The basic problem addressed is to research methods for learning to store and recognize a large body of temporal sequential percepts (presented in any order) utilizing neuroscience compatible techniques. Further, the system must be capable of stable learning and forgetting of randomly presented samples. In spite of a decade or more of research in artificial neural networks, temporal sequence associative memory is still very poorly understood. Researchers are studying, analyzing, and simulating a number of different models.

**Model of Superior Colliculus with a Spatiotemporal Neuron Model**
S. R. Ray,* C. Seguin, T. Anastasio
*Defense Advanced Research Projects Agency*

The superior colliculus is a region of the central nervous system that fuses input data from several modes (visual, auditory, somatosensory, and so forth) and computer orienting output for directing attention, such as head rotation. This project uses a neuron model with trainable temporal response. Spatiotemporal maps are learned by self-organizing algorithms that simulate known functions of the superior colliculus. The result is expected to have engineering applications.

**Neural Network Methods for Temporal Sequence Processing**
S. R. Ray,* H. Shah
*University of Illinois*

The storage of long temporal vector sequences and their recognition and recall is a subject of considerable interest both in engineering application and in neuroscience. A new method for storing long temporal sequences that can be content-addressed and retrieved from the starting point is under study. The method applies multilayer "chunking" of invertible vectors in feature map architectures. This multilayer method has the potential to store a quantity of distinct sequences in the same elements and to demonstrate the property of replay of long experiential memories discovered by Wilder Penfield in human subjects some decades ago.

**Self-Aiming Camera**
S. R. Ray,* T. Anastasio, P. Patton
*University of Illinois*

The superior colliculus in vertebrates appears to be a primary agent in deciding the direction of saccades and head-turning actions. Multisensory information, especially visual, audible, and somatosensory, feeds the two-layer neural network that comprises the superior colliculus. This research team has developed and continues to develop models of the SC that learn to adapt to variations in the head dimensions and visual/auditory properties. To demonstrate the overall system, including learning, researchers are constructing a camera and microphone system that will supply input to the model SC and respond to its directives. The camera/microphone system is fully operative on a physically fixed frame.

**Context-Sensitive Natural Language Inferences**
D. Roth,* A. Carlson
*National Science Foundation, IIS-9801638; IBM Corp*

The future of intelligent human–machine interaction is in the ability to perform context-sensitive inferences. These are knowledge intensive tasks that are difficult to make without a significant learning component. This research studies a learning approach that targets knowledge intensive language understanding related tasks and directly addresses the issue of scalability. It is tailored to large-scale processes in terms of data and computation. The approach developed can be applied to support a variety of inferences of the sort required in intelligent human–machine interactions, as demonstrated in this project by development of a system that exhibits a wide coverage and accurate context-sensitive spelling correction.

**Inference with Classifiers**
D. Roth,* V. Punyakanok
*National Science Foundation, Information Technology Research, IIS-00-85836*

In many situations it is necessary to make decisions that depend on the outcomes of several classifiers in a way that provides a coherent inference that satisfies some constraint. These constraints might arise from the sequential nature of the data or other domain specific constraints. Researchers are studying two general approaches for this problem and are evaluating those in the context of an important inference problem in natural language—identifying phrase structure. The first approach

* Denotes principal investigator.
studied is a Markovian approach that extends standard HMMs to allow the use of a rich observation structure and of general classifiers to model state-observation dependencies. The second is an extension of constraint satisfaction formalisms.

Intermediate Knowledge Representations that Facilitate Learning
D. Roth,* S. Agarwal, C. Cumby, W. T. Yih, D. Zimak

_National Science Foundation, Information Technology Research, IIS 00-85836; NSF IIS 00-85980; Office of Naval Research, Multidisciplinary Research Program of the University Research Initiative Award_

Learning becomes easy once the correct input representation has been chosen. A representation that produces linearly separable point sets is an example. Several projects are aimed at automatically generating intermediate representations to aid supervised learning algorithms, developing methods that allow the use of relational representations and of learning relational definitions, and developing a flexible knowledge representation language that can be used along with feature-efficient learning algorithms. Applications of this general knowledge representation paradigm are studied in the context of learning in the natural language domain and visual recognition.

Learning Coherent Concepts
D. Roth,* A. Garg, V. Punyakanok

_National Science Foundation, CAREER Award, IIS-9984168_

This research seeks to develop an integrated view (theoretical understanding, algorithms development, and experimental evaluation) for learning coherent concepts. These are learning scenarios that are common in cognitive learning, where multiple learners co-exist and may learn different functions on the same input, but there are mutual compatibility constraints on outcomes. This effort will consist of developing a learning theory for these situations and of studying algorithmic ways to exploit them in natural language inferences. The theoretical study concentrates on developing semantics for the coherency conditions and study of it from a learning theory point of view. The goal is to understand ways that learning can become easier and more robust in these situations. The algorithmic study concentrates on developing ways to exploit coherency and makes use of several important problems in natural language processing as a testbed for investigating chaining of coherent classifiers and inferences that rely on the outcomes of several classifiers.

Learning from Data and Additional Knowledge Sources
D. Roth*

_National Science Foundation, IIS-9801638_

The majority of the work in learning assumes that the learner interacts with the world via examples. This research focuses on situations in which the learner can interact with many information sources. Some of these may supply the learner with examples, while others may supply other types of information. In text-understanding related tasks, for example, in addition to examples (text) the learner may make use of dictionaries, a thesaurus, experts of various sorts, various general and domain-specific taxonomies, and other sources. In other cases, the additional information can be viewed as cross-modality information. This research is a study of how learning algorithms can make use of this additional information in various learning models and for various applications.

Learning Theory
D. Roth*

_National Science Foundation, IIS-9801638; National Science Foundation CAREER Award, IIS-9984168_

This is a study of several questions in learning theory that pertain to the problem of learning a number of concepts from common data. The study encompasses interactions and constraints imposed on outcomes. The main direction is the development of a theoretical framework for learning coherent concepts. This is a study of learning situations in which learning does not occur in isolation. Rather, the input is observed and processed by multiple learners. The goal is to understand what ways learning becomes easier and more robust in these situations.

Learning to Perform Knowledge-Intensive Inferences
D. Roth*

_National Science Foundation, IIS-9801638_

The goals of this research are to study an integrated theory of learning, knowledge representation, and reasoning and to evaluate it on large-scale, knowledge-intensive inferences in the natural language domain. Recent studies within the learning-to-reason framework have shown that there is much to gain from studying these issues within a unified framework. This research investigates some of the fundamental issues within this framework, concentrating on a probabilistic setting.
Two main lines of research are being pursued. One focus is on developing a coherent learning theory account of the major statistical approaches to learning in natural language. This is an attempt to develop better learning methods and an understanding of the role of learning in natural language inferences. A second focus is on the study of knowledge representations and learning techniques for various language-understanding related tasks. The emphasis is on learning techniques that tolerate data of high dimensionality and on incorporating additional knowledge. Projects include context-sensitive spelling correction, prepositional phrase attachment, part of speech tagging, shallow parsing, and applications to information extraction.

Robust Methods for Natural Language-Based Human and Machine Interaction
D. Roth,* X. Li, V. Punyakanok, W. T. Yih
Office of Naval Research, Multidisciplinary Research Program of the University Research Initiative Award

This research is directed toward the development of robust, freestyle, and adaptive natural language-based human and machine interaction. These include developing robust methods for identifying phrases in sentences. This is a fundamental technology that underlies the ability to extract key phrases and perform shallow parsing of sentences. The research also targets methods for information extraction, a robust identification of functional phrases in sentences. This is part of an effort to construct a more abstract representation of sentences that will be used to respond to queries and/or access a knowledge base. Another focus is on question–answering. This approach integrates the two other methods with relevant learning, knowledge representation, and inference methods in the natural language domain to a preliminary open domain question–answering system.

The Role of Experience in Natural Language
D. Roth,* K. Bock, J. Cole, G. Dell, C. Fisher, S. Garnsey, A. Goldberg, S. Levinson
National Science Foundation, KDI SBR 98-73450

An integrated multiparadigm approach to the study of learning mechanism in language production and comprehension is studied. The language processing system is constantly changing. It adapts quickly to recent experience while continuing to reflect the accumulated experience of a lifetime of speaking, listening, reading, and writing. This project integrates research efforts in psycholinguistics, linguistics theory, and computational models of learning in an attempt to address the mechanisms that enable the language processor to adapt to experience. In addition, the research will suggest learning mechanisms for language processing technology, particularly for rapid adaptation to changing linguistic environments.

The SNoW Learning Architecture
D. Roth,* A. Carlson
National Science Foundation, IIS-9801638; NSF KDI SBR 98-73450

A learning architecture and algorithms that are tailored for learning on large-scale, knowledge-intensive problems are being developed. The SNoW learning architecture is a sparse network of linear units over a common predefined or incrementally learned feature space. It is tailored for learning in domains in which the potential number of features taking part in decisions is very large, but may be unknown a priori. Preliminary versions of SNoW have already been used successfully on a variety of large-scale learning tasks in NLP and in the visual processing domain, including face detection and object recognition.

Computer Architecture and Compilers

ALP: Efficient Support for All Levels of Parallelism for Complex Media Applications
S. Adve,* A. Li, R. Sasanka; Y.-K. Chen, E. Debes (Intel)
AMD Corp.; Intel Corp.; National Science Foundation,
CCR-02-09198, EIA-02-24453

Real-time execution of contemporary multimedia applications needs processing power and energy efficiency that surpasses the capabilities of current superscalar processors. Many recent projects have proposed architectures for media applications focused on exploiting their large amounts of data-level parallelism (DLP). Our studies have shown that in advanced applications, the DLP is often interspersed with control, and architectures focusing solely on DLP are unlikely to suffice. This project is developing a general-purpose architecture for multimedia applications, called ALP, which seamlessly supports multiple forms of parallelism, including instruction-level parallelism, thread-level parallelism, and various forms of DLP such as sub-word SIMD instructions, vectors, and streams. ALP combines conventional and evolutionary techniques to provide energy-efficient, high-performance for an increasingly important class of
applications, without the need for radically different programming paradigms.

**Compiler Support for Performance Modeling of Parallel and Distributed Programs**

V. Adve*

*Defense Advanced Research Projects Agency, N66001-97-C-8533*

Researchers are developing compiler techniques that enable fast, accurate, and automatic performance modeling of highly scalable applications. One focus of this effort is a compiler-generated program representation that allows one to automate a wide range of analytical, simulation, and hybrid models of parallel programs. A second focus is to use additional compiler analysis, together with this representation, to enable efficient simulation of highly scalable applications. One such compiler technique achieved 10-2000x reduction in memory usage and 2-10x reduction in simulation time for the simulation of large message-passing programs. This work is part of a broader collaboration with five other universities.

**Dynamic Energy and Temperature Management for Multimedia Applications**

S. Adve,* X. Li, J. Srinivasan

*AMD Corp.; Motorola Inc.; National Science Foundation, CCR-00-96126, EIA-02-24453*

Power consumption is a first-class design constraint in current systems. The problem manifests itself in terms of energy (which affects battery life) and temperature (which affects cooling costs and system timing). A promising approach to reducing energy consumption and temperature on current hardware is to make the hardware adaptive. The hardware dynamically responds to changes in the application profile to ensure that only the resources required are used and only the performance desired is delivered, thereby reducing energy consumption and temperature. This project has developed control algorithms for adaptive processors for dynamic energy and temperature management for the increasingly important real-time multimedia application domain.


S. Adve,* A. F. Harris, C. J. Hughes, D. L. Jones, R. H. Kravets, K. Nahrstedt, D. G. Sachs, V. Varadhan, W. Yuan

*AMD Corp.; National Science Foundation, CCR-02-05638, EIA-02-24453*

Mobile devices that primarily process multimedia data are expected to become a dominant computing platform for a variety of application domains. Their design must consider demanding, dynamic, and multiple resource constraints, with energy as a first-class resource. However, the ability of multimedia applications to trade off output quality for system resources offers an opportunity for the design of systems where each system layer can adapt in response to resource or application changes. Reaping the full benefits of a system with multiple adaptive layers, however, requires a careful coordination of those adaptations. This project is developing a cross-layer adaptive system and framework to reduce energy consumption while preserving application quality within available computation and bandwidth resources. The final system prototype will integrate adaptations in the hardware, network, operating system, and application layers, opening up sources of energy savings not possible before.

**RAMP: Lifetime Reliability Aware Microprocessors**

S. Adve,* X. Li, J. Srinivasan; P. Bose, J. Rivers (IBM)

*AMD Corp.; IBM Corp.; National Science Foundation, EIA 02-24453*

Ensuring long processor lifetimes by limiting failures due to hard errors is a critical requirement for all microprocessor manufacturers. This project observes that CMOS technology scaling and increased on-chip temperatures will make lifetime reliability targets difficult to meet, that current reliability qualification methodologies (based on worst-case operating conditions) are overly conservative and unnecessarily constrain performance, and that lifetime reliability awareness at the architectural design stage can mitigate this problem by optimizing processors for the expected operating conditions, dynamically tracking workload-specific lifetime reliability estimates, and providing back-up techniques to handle unexpected conditions. Based on these observations, this project is developing lifetime reliability models and architectural techniques that will enable meeting the required lifetime reliability targets of future scaled processors.

**SAFECode: A Compiler System for Enforcing Memory Safety in C Programs**

V. Adve,* D. Dhurjati, S. Kowshik

*National Science Foundation, EIA-00-93426, CCR-02-09202, and CNS-04-06351; SRC MARCO/DARPA consortium*

Programming errors that lead to illegal memory accesses are a dominant source of vulnerabilities that lead to security violations in modern computer systems. Such errors can include buffer overruns, dangling pointer errors, and others. The SAFECode compiler automatically transforms

* Denotes principal investigator.
programs to prevent all kinds of memory errors in production code, through a combination of static and run-time checking. For safety techniques used in production code, the run-time checking introduces very low overheads. These low overheads are made possible by a novel approach for making hard-to-find errors harmless, without actually eliminating those errors. The compiler also ensures that a specific approach to program analysis can be made sound, despite the presence of hard-to-detect errors, enabling static checking and verification tools to give guarantees of soundness for the properties they claim to prove. SAFECODE also provides additional techniques for debugging difficult-to-find errors like dangling pointers, through more expensive but complete techniques.

Secure Virtual Architecture (SVA): Compiler Techniques for Operating System Security and Reliability

V. Adve,* A. Lenharth,* J. Criswell, D. Dhurjati
National Science Foundation, EIA-00-93426, CNS-04-06351, CCF-04-29561; DARPA/MARCO Gigascale Systems Research Center (GSRC)

Despite major developments in technologies for system security and reliability, today's operating systems and user processes remain vulnerable to an extensive range of attacks. At the root cause of such vulnerabilities is an important class of bugs: implementation errors, e.g., memory access errors or missing security checks. These are not addressed even by powerful, security-conscious design approaches like security kernels or virtual machine monitors. Moreover, design weaknesses in legacy kernels can magnify the potential risk when a successful exploit occurs, e.g., allowing device drivers or other kernel extensions to operate with full privilege within the kernel address space. This project is developing a compiler-based virtual machine (Safe LLVA) that is capable of hosting a standard C/C++-based operating system and all its applications, and exploring how this organization can improve overall system security and provide new security capabilities. Such a virtual machine can apply a range of compiler-based safety checking techniques on legacy kernel code. The immediate goal of this work is to apply compiler-based techniques for memory safety from the SAFECODE project to kernel code to prevent attacks via memory access errors. More broadly, the combined system will also address techniques for reducing kernel implementation errors, improving application security by enforcing least-privilege, and protecting critical application data from the kernel even if it is compromised.

Virtual Instruction Set Computers (VISC)

V. Adve,* R. Bocchino, J. Criswell, D. Dhurjati, T. Lattner,
National Science Foundation, EIA-00-93426, CCF 04-29561; DARPA/MARCO Gigascale Systems Research Center (GSRC)

Virtual Instruction Set Computer (VISC) architectures define two separate instruction sets (one to serve as the representation of stored programs, and another to control the hardware), and use a hardware-specific translation layer to optimize and execute code on the hardware. This approach lends great flexibility to hardware and compiler design, enabling close cooperation between the translator and the microarchitecture on each individual chip (since each distinct processor design has its own specific translator). This is because the hardware instruction set can expose microarchitecture details to the translator and, conversely, can rely on the translator to produce code with predefined assumptions. This project is defining a novel virtual instruction set for VISC architectures called Low Level Virtual Architecture (LLVA) and a translation strategy that permits both offline and online translation or a combination thereof. The project is also exploring compiler and architecture techniques to exploit the VISC framework with the rich virtual instruction set for improved processor performance and reliability.

The LLVM Compiler System

National Science Foundation, EIA-00-93426, CCR-02-09202, CNS-04-06351, CCF 04-29561; DARPA/MARCO Gigascale Systems Research Center (GSRC)

Modern applications are increasing in size, change their behavior significantly during execution, support dynamic extensions and upgrades, and often have components written in multiple different languages. While some applications have small hot spots, others spread their execution time evenly throughout the application. In order to maximize the efficiency of all of these programs, program optimization must be performed throughout the lifetime of a program, including compile-time, link-time, install-time, run-time, and "idle-time" profile-guided optimization between runs. The LLVM compiler project provides a framework to make "lifelong" program analysis and transformation available for arbitrary software, and in a manner that is transparent to programmers. The LLVM system retains key program information in a language-independent format that can be used for lifelong * Denotes principal investigator.
compilation, and exploits this representation to provide a unique combination of capabilities. LLVM is available freely in open-source form and has attracted a large community of active users and contributors in industry, academia, and the open-source community.

SoftCheck: Compiler and Run-Time Technology for Efficient Fault Detection and Correction Low nm-Scale Multicore Chips
M. Garzaran,* J. Yu
IBM PERCS W0133930

In collaboration with researchers from Texas A&M University

We are developing SoftCheck, a flexible and efficient Runtime and Compiler System (RCS) technology to cost-effectively detect and recover from hardware faults in upcoming multicore chips. SoftCheck will implement RCS algorithms that avoid full instruction replication within or across threads—this is key to acceptance in the energy- and cost-conscious commodity market. It will also provide knobs to select the desired performance vs. error-coverage tradeoff. SoftCheck is composed of a wide range of novel, efficient techniques. The fault-detection techniques include cost-effective techniques for a thread to exhaustively check itself, partially check itself, partially check many other threads in the core, and (other crosscutting, often multiprocessor-related, issues. The fault-correction techniques include disable clusters in a core, disable complete cores, and potentially, dynamically recompile to use other hardware.

Logic Synthesizers for VLSI Chip Design
S. Muroga,* V. Jayasena
University of Illinois

Automated design of logic networks is indispensable for designing a microprocessor with millions of transistors on a single chip. As the line width becomes less than 18 microns, signal propagation increasingly takes more time on connections than logic gates. So, logic networks are designed and the layout is made on a chip, and then logic networks are redesigned with consideration of connection length. The cycle of design and layout must be repeated until a satisfactory outcome is achieved, unlike the former approach of dealing with the layout once and only after logic networks are designed. Researchers are developing new algorithms for this new approach.

Analysis of Irregular Memory Accesses
D. Padua,* Y. Lin
Defense Advanced Research Projects Agency, DABT63-95-C-0097

This project aims to develop compiler techniques to extract precise information about irregular memory accesses. Traditional loop optimization methods require array subscripts to be closed-form expressions of loop indices. However, in sparse/irregular programs, array subscripts usually do not have this form, leaving many codes unoptimized as a result. Researchers have developed a demand-driven interprocedural array property analysis method for indirectly accessed arrays and a single-indexed array analysis method for irregular single-indexed arrays. The team has successfully improved the precision of data dependence and privatization tests, which has resulted in significant performance improvements for the collection of sparse codes.

Automatic Generation of Efficient Sorting Routines
D. Padua,* M. Garzaran, X. Li, B. Garber
National Science Foundation, CSR-AES 0509432

This project is part of a larger project on library generation technology. While the performance of the dense linear algebra and signal processing algorithms implemented by the best known library generators of today does not depend on the values of the input data, the best sorting algorithm is a function of characteristics of the input data such as its standard deviation and the degree to which it is partially sorted. Therefore, a sorting library generator must not only tune each algorithm to the target machine, but also compute a function that, at runtime, selects the best algorithm as a function of the input data. The generators implemented in this project compute such a function using machine learning strategies. They produce sequential and parallel sorting routines that are, on average, significantly faster than any other widely available sorting routine.

Automatic Generation of Signal Processing Algorithms
D. Padua,* M. Garzaran, J. Brodman, G. Ren
National Science Foundation, ITR-NGS 0325687; Defense Advanced Research Projects Agency

In collaboration with researchers from Carnegie-Mellon and Drexel

A generator of signal processing libraries, called SPIRAL, has been developed. The code produced by the generator is as efficient as hand-written codes. SPIRAL accepts descriptions of algorithms and their variants in the form of formulas. This makes the system extensible, a feature seldom found in library generators. A back-end compiler

* Denotes principal investigator.
implemented at Illinois applies machine-independent source-to-source optimizations and extends the search to the space of coding options such as the degree of loop unrolling and scheduling strategies. Transformations that take advantage of SSE and Altivec vector devices have recently been incorporated into the compiler in collaboration with Peng Wu of IBM Research.

**Compiler and Run-Time Environment for Numerical Java**

D. Padua,* L. V. Kale,* J. Decker, Z. Sura, D. Wong, P. Wu

*National Science Foundation, DMS 98-78945*

In this project, compiler and run-time techniques will be developed to achieve a performance comparable to that of Fortran programs for numerical Java applications. The system consists of a Java restructurer, a JVM that supports vector operations efficiently, an optimized native numerical library, and a source-to-source Fortran 95 to Java translator. The restructurer automatically identifies macro-operations (vector operations and BLAS routines) in the source program and replaces them with calls to library routines. The source-to-source translator provides a tool to convert legacy Fortran codes to Java and to produce a convincing testing benchmark suite for the system.

**Compilers for Processors in Memory**

D. Padua,* J. Torrellas,* J. Lee

*IBM Partnership Award*

Processors in memory (PIM) technology integrate processor logic and DRAM in the same chip. The goal of this research is to explore compiler techniques for FlexRAM, a computer system that consists of a commodity host microprocessor and several PIM chips. The research team is exploring compiler techniques to identify macro-operations (such as vector operations) from the source code. The strategy is to partition the source code into several functions based on the profitability of running the functions on the PIM chips or the host processor. This research encompasses performance prediction, parallelism extraction, cache locality enhancement, synchronization overhead reduction, and load balancing.

**Compile-Time Performance Prediction**

D. Padua,* D. A. Reed,* C. Cascaval

*Defense Advanced Research Projects Agency, N66001-97-C-8532*

As part of the Delphi project, University of Illinois researchers have developed compiler techniques to predict the performance of scientific codes on different architectures. The compiler is both the primary tool for estimating the performance and the beneficiary of the results obtained from predicting the program behavior at compile time. The research team developed a simple compile-time performance prediction model that, when augmented with profiling data from very light instrumentation, can be accurate within an average of 20% of the measured performance for codes using both dense and sparse computational methods on uniprocessor machines. A next step is to extend this model to multiprocessors.

**Compiling for PC Clusters**

D. Padua,* J. Hoeflinger, Y. M. Kim, J. Zhu

*National Science Foundation, ACI 96-19019 COOP*

In this project, researchers are retargeting the Polaris parallelizing compiler to a network of Sun workstations and a Windows NT-based cluster using the TreadMarks virtual shared memory system to provide a shared view of memory. Researchers ported TreadMarks to the NT cluster and made it compatible with FORTRAN77 codes. The basic functionality of the compiler for TreadMarks has been built. Some programs from the Spec95 and Perfect Benchmarks have been tested and valid results obtained. Next steps are to measure the performance of a group of benchmark codes and design optimizations to improve their performance.

**Hierarchically Tiled Arrays**

D. Padua,* B. Fraguela, M. Garzaran, G. Bikshandi, J. Guo, D. Hoeflinger

*National Science Foundation, NGS 0103610, CSR-AES 0509432*

In collaboration with IBM Research members George Almasi, Calin Cascaval, and Christoph Von Praun

Tiling is perhaps the most widely used programming strategy for data distribution in parallel algorithms and locality enhancement of sequential and parallel algorithms. The objective of this project is to contribute to programmers' productivity by making available building blocks that represent and manipulate tiles in a natural manner. The core of the proposed approach is a class of objects, hierarchically tiled arrays (HTAs), which are arrays that may be recursively tiled. Operations on HTAs represent, in compact form, data movements and operations on all or some of their components. When HTA tiles are distributed across a parallel machine, HTA operations represent communication and parallelism in an intuitive, easy-to-read manner. MATLAB, C++, and X10 HTA implementations have been developed. Using these libraries, researchers have programmed a large collection

* Denotes principal investigator.
of programs, including the NAS parallel benchmarks. The resulting codes demonstrate an immense readability advantage of HTA programs over other forms of parallel programs.

**Signal Processing Algorithm Compiling Engine (SPACE)**

D. Padua,* A. Huang, J. Xiong

*Defense Advanced Research Projects Agency, DABT63-98-1-0004*

Fast signal processing algorithms have been obtained by properly factorizing the transformation matrix into several matrices. The Tensor Product Language (TPL) was designed to describe linear transformations. Using predefined parameterized matrices and their operations, TPL can represent several well-known signal processing algorithms. The goal of this project (which is part of a larger project involving researchers from Carnegie-Mellon, Drexel, and the University of Southern California) is to build a compiler to generate efficient programs from TPL formulas. Issues to be studied include techniques to predict the effect of loop unrolling/rerolling on performance and optimization techniques to very long basic blocks.

**A MATLAB Just-in-Time Compiler (MaJIC)**

D. Padua,* G. Almasi, L. DeRose

*National Science Foundation, ACI98-70687*

In cooperation with researchers at Cornell University, a University of Illinois team is developing MaJIC, a just-in-time compiler for MATLAB. The main goal of the project is to maintain a MATLAB-like working environment while achieving significant speedups. The analysis techniques under investigation are type analysis, dependence analysis (for generating SMP code), code signature generation, and array bounds check analysis (to eliminate superfluous array bounds checks). The compilation and execution phases in MaJIC overlap and can be executed in parallel on multiprocessor machines. Researchers also are considering stepwise refinement of compiled code to improve performance.

**Los Alamos Computer Science Institute: Compilation, Systems, and Performance Evaluation of Large Scale Parallel Machines**

D. A. Reed, * C. Mendes, K. Pattabiraman, Y. Zhang

*U.S. Department of Energy, W-7405-ENG-36*

The objective of this Los Alamos Computer Science Institute project is to develop compiler and run-time technology that will help application developers achieve a high fraction of peak performance on large-scale parallel computing systems. As one of the LACSI partners, University of Illinois at Urbana-Champaign researchers seek to develop a new generation of software tools that integrate compilation research with real-time performance measurement, adaptive control systems, and intelligent data visualization and control. The goal of this work is to create an intelligent performance toolkit that allows software developers to create and runtime systems to negotiate performance contracts among software components and hardware/software systems.

**Architectural and Application Support to Crack the Protein Folding Problem**


*National Science Foundation, EIA 00-81307; IBM Corp.*

Solving the protein folding problem would have an enormous pay-off: providing a true understanding of diseases and the discovery of more effective drugs. Luckily, advances in processor-memory integration allow us to build computers that, thanks to being more integrated, faster and cheaper, may solve the problem. In this project, researchers in computer hardware, software, and computational biology team up with IBM researchers to design improved algorithms for protein folding and to design a Processor-In-Memory architecture that can speed up the solution of the problem. The work involves using the IBM Blue Gene prototype and will help design its next generation system.

**Architectural Support for Speculative Thread-Level Parallelization**

J. Torrellas,* M. Cintra, J. Martinez, M. Prvulovic, M. Garzaran

*National Science Foundation, EIA 99-75018, CCR 99-70488, EIA 00-81307, EIA 00-72102; IBM Corp.; Intel Corp.*

Speculative thread-level parallelization is a technique to execute in parallel code that the compiler cannot fully analyze. In this technique, a program is dynamically divided into tasks and assigned to different threads. The threads execute in parallel, optimistically assuming that sequential semantics will not be violated. As the threads run, the data that they access are tracked. If a dependence violation is detected, the offending threads are stopped. Then a repair action re-executes the offending tasks, possibly after recovering some old, safe state. In this project, researchers examine the support for this technique in both chip multiprocessors and scalable multiprocessors.

* Denotes principal investigator.
Automatically and Manually Programming a Server
J. Torrellas,* D. Padua, F. Fraguela, J. Lee, Y. Solihin
National Science Foundation, EIA 99-75018, CCR 99-70488, EIA 00-81307, EIA 00-72102; IBM Corp.; Intel Corp.

This project addresses the problem of how to program a server with an intelligent memory system, such as FlexRAM. Such a machine includes two classes of processors: one or several off-the-shelf powerful processors and a myriad of simple processors in the memory system. Researchers examine two approaches to program this machine. The first approach is to implement compiler algorithms so that the compiler can automatically map the code onto the architecture. The second approach is to design language constructs and a programming methodology that make it possible and easy for a programmer to directly write code for the machine.

Dynamically Managing Energy in Chips for Energy and Efficiency
J. Torrellas,* M. Huang, J. Renau, S. Yoo
National Science Foundation, EIA 99-75018, CCR 99-70488, EIA 00-81307, EIA 00-72102; IBM Corp.; Intel Corp.

While technology is delivering increasingly sophisticated and powerful chip designs, it is also imposing alarmingly high energy requirements on the chips. Many techniques have been proposed to manage the energy consumed, including voltage scaling and various forms of storage reconfiguration. In this project, the different techniques are compared. A framework that applies techniques dynamically, in a fine-grained manner and according to a given policy is built. Goals are to maximize energy savings without extending application execution time beyond a given tolerable limit and to guarantee that the temperature remains below a given limit while minimizing any resulting slowdown.

FlexRAM: An Intelligent Memory Architecture System
J. Torrellas,* D. Padua, D. A. Reed, M. Huang, J. Lee, J. Renau, Y. Solihin, S. Yoo
National Science Foundation, EIA 99-75018, CCR 99-70488, EIA 00-81307, EIA 00-72102; IBM Corp.; Intel Corp.

Major advances in Merged Logic DRAM (MLD) technology coupled with the popularization of memory-intensive applications provide fertile ground for architectures based on Processors-in-Memory (PIM). In the FlexRAM project, researchers use PIM chips as the memory of an intelligent server. If such a server runs an application without recompiling it, the intelligent memory appears as plain memory. However, if the application is recompiled, the memory becomes a high-performance accelerator. In this project, the focus is on issues related to architecture design, energy dissipation, programming, compilation, run-time system, and applications for the FlexRAM intelligent memory system.

ITR: Automatic On-the-fly Detection, Characterization, Recovery, and Correction of Software Bugs in Production Runs
J. Torrellas,* J. Han, S. Midkiff, Y. Zhou
National Science Foundation, ITR-03-25603

Software bugs continue to be frequent in computer systems, accounting for as much as 40% of computer system failures. Many bugs that occur in production runs are hard to reproduce when the program is recompiled with heavy instrumentation and executed in a debugging run. Consequently, it is necessary to provide low-overhead debugging support that can be used in production runs. This work proposes to develop a novel, comprehensive debugging system for automatic, on-the-fly debugging of production runs. The system supports all the steps of debugging, namely bug detection, characterization, recovery, and correction. The system integrates innovations in computer hardware, operating system, data mining, and compiler support.

M3T: Morphable Multithreaded Memory Tiles
Defense Advanced Research Projects Agency, F30602-01-C-0078

The M3T system is a novel malleable computing system composed of polymorphous hardware and polymorphous software that can adapt to changing mission demands. The system is built by tiling M3T processor chips. Each processor chip is composed of many general-purpose RISC cores interleaved with memory blocks. Cores and memories are reconfigured at run-time, allowing the chip to morph into a superscalar, VLIW, systolic array, MIMD, SIMD, or fault-tolerant engine, or even a combination of these. The system has a compiler that reconfigures the platform (hardware and run-time system) for which it is generating code.

* Denotes principal investigator.
PERCS: A Highly Productive Computer System
J. Torrellas,* R. Johnson, D. Padua, M. Snir
Defense Advanced Research Projects Agency; IBM Corporation, INT IBM NBC PERCS

In cooperation with IBM, we are designing a highly adaptable multiprocessor that will yield up to 8 Petaflop/second of sustained performance and a 10-time reduction in time-to-solution in the year 2010-2012 timeframe. The vision centers on the theme of adapting the system layers to the application requirements. Adaptability enhances the technical efficiency of the system as well as its ease of use. Among the issues that we are studying are novel memory subsystem design, architectural support for debuggability and fault tolerance, architectural support for adaptability, and programming languages and compilers to enhance the ease of use.

Using Architectural Support for Speculation to Provide Fault Tolerance
J. Torrellas,* M. Garzaran, M. Prvulovic, Z. Zhang
National Science Foundation, EIA 99-75018, CCR 99-70488, EIA 00-81307, EIA 00-72102; IBM Corp.; Intel Corp.

Microprocessors occasionally suffer transient faults. This research is aimed at finding ways to detect transient faults and recover from them. While past work has addressed this problem in different ways, speculative parallelization opens up a potential new approach to detect and recover from transient faults. Under speculative parallelization, speculative threads must keep their memory state buffered until they are proved to be correct. In addition, they must be able to roll back to a safe state after a violation is detected. These same architectural supports can be used to provide fault detection and fault recovery. This project extends existing architectural support for speculation to provide inexpensive support for fault tolerance.

Macroscopic Data Structure Analyses and Transformations
V. Adve,* Q. Yi, K. Kennedy (Rice Univ.)
U.S. Department of Energy ASCI Academic Strategic Alliances Program, B347884

Managing performance on deep memory hierarchies is widely considered to be a critical open problem for high-performance systems. This research team is exploring a novel class of compiler transformations that provides improved locality at multiple levels of memory hierarchy simultaneously. The transformations exploit the property that recursive algorithms have identical reuse patterns at each level of recursion, providing a hierarchy of working sets. Researchers are developing compiler algorithms to transform existing loop-based codes into efficient recursive form automatically. Such a transformation has wide applicability, including automatic blocking for multiple levels of cache hierarchy and improving communication locality in shared memory codes.

Enhancing Single-Thread Performance through Hardware-Supported Software Speculation
C. Zilles,* N. Neelakantam
National Science Foundation, CCR 03-11340, EIA-02-24453; Intel Corp.

Hardware-based speculative execution has proven to be an effective technique for improving single-thread program execution, but continuing to scale pure hardware solutions is proving to be too inefficient for future power constrained systems. In this work, we are exploring architectures where hardware provides the support for speculative execution and recovery, but the decision of where to speculate and how to speculate are controlled by the software.

Databases and Information Systems

AIM: Supporting Efficient Top-k Query Processing
K. C.-C. Chang,* S. Hwang
University of Illinois

Supporting database queries that return ranked results is critical for decision support applications as well as multimedia and information retrieval. Such ranked queries order results by combining the scores of fuzzy predicates that are evaluated by different sources, which can be a local database, a multimedia subsystem, or a Web source. This project develops efficient algorithms for processing ranked queries.

MetaQuerier: Exploring and Integrating the Deep Web
K. C.-C. Chang,* B. He, C. Li, Z. Zhang
National Science Foundation, CAREER Award IIS-0133199, ITR Award IIS-0313260; National Center for Supercomputing Applications Faculty Fellows Award, 2003

This research seeks to enable effective access to structured information sources on the Internet. Over the past few years, a significant and increasing amount of information has become hidden on the "deep" Web, behind the query interfaces of searchable databases. A metquery system would help users in finding and querying online databases effectively and uniformly. First, to make the deep Web systematically accessible, the MetaExplorer will discover sources to build a "database of online databases" to help users find sources useful for their purposes. Second, to
make the deep Web uniformly usable, the MetaIntegrator will help users interact with online databases to ask queries.

**WISDM: Web Indexing and Search for Dynamic Mining**
K. C.-C. Chang,* T. Cheng, K. Pham, B. Zhao, H. Zhong

*Microsoft Accelerating Search in Academic Research Grant; IBM Faculty Fellow Awards; Google Research Grant*

This research seeks to enable “data-aware” search over the Web, to deepen the reach of search techniques into the vast amount of data on the Internet. The immense scale and wide spread has rendered the Web as an ultimate information repository—as not only the source where we find but also the destinations where we publish our information. These dual forces have enriched the Web with all kinds of data, much beyond the conventional page view of the Web as a corpus of HTML pages, or “documents.” Consequently, the Web is now a rich collection of data-rich pages, on the “surface Web” of static URLs (e.g., personal or company homepages) as well as the “deep Web” of database-backed contents (e.g., flights from aa.com). The richness of data, while a promising opportunity, has challenged us for effectively finding data we need. We propose to build novel search systems, to facilitate the user’s quest of data on the Web.

**iMAP: Discovering Complex Matches between the Schemas of Databases**
A. Doan,* R. Dhamankar, Y. Lee, A. Halevy, P. Domingos

*University of Illinois*

Establishing semantic matches between disparate data sources is fundamental to numerous data sharing efforts. Manually creating matches is extremely tedious and error-prone. Hence, many recent works have focused on automating the matching process. To date, however, virtually all of these works deal only with one-to-one (1-1) matches, such as address = location. They do not consider the important class of more complex matches, such as address = concat(city,state) and room_price = room_rate * (1 + tax_rate). In this project we describe a solution that applies learning and search techniques from the AI field to semiautomatically discover both 1-1 and more complex matches with high accuracy.

**MOBS: A Mass Collaboration Approach to Integrate Data Across Disparate Sources**
A. Doan,* R. McCann, A. Kramnik, W. Shen, V. Varadarajan, O. Sobulo

*University of Illinois*

Today, data integration systems are largely built by hand, a labor-intensive and error-prone process. A conceptually new solution to this problem is to consider a data integration system as having a finite set of parameters whose values must be set. To build the system the administrators construct and deploy a system "shell,” then ask the users to help the system "automatically converge" to the correct parameter values. Therefore, the burden of system development is lifted from administrators and spread "thinly” over many users. We apply this approach to the problem of schema matching in the context of data integration. Experiments are conducted with both real and synthetic users to show the promise of the approach.

**DataScope: Viewing Database Contents in Multiresolution at Your Finger Tips**
J. Han*

*National Science Foundation, IIS-06-42771*

The goal of this research project is to investigate issues in the design and development of a multiresolution data viewing system, called DataScope, which interacts with database systems based on visual comprehension and exploration of database contents in a flexible, user-friendly, multidimensional, and multiresolution way. The approach consists of investigation of the design principles and implementation considerations, including structures of dimensions, ordering of attribute values, layout of data values, flexible selection of layouts, attributes, and constraints, rich information display, and automated visualization interface generation and customization. The research integrates the technologies in database systems, data warehouse, data visualization, human–computer interaction, and data mining. It emphasizes efficient implementation of the functions for visual exploration of database contents and scale-up of such functions for real-time usage. The research results are to be published in research forums on database systems, data mining, and human–computer interaction, and will be integrated into the graduate education at the University of Illinois at Urbana-Champaign. This project is expected to have broad applications in many fields that need to explore huge amounts of data, including business, industry, government agencies, scientific research, and education.

* Denotes principal investigator.
Endowing Biological Databases with Analytical Power: Indexing, Querying, and Mining of Complex Biological Structures
J. Han,* J. X. Zhou
National Science Foundation, DBI-05-15813

A new-generation biological database system should be analysis-based and should support efficient and scalable indexing and accessing of complex biological structures, extracted from multiple, inter-related biological graphs and networks, and also facilitate cross-relational and cross-database analysis since biological data are usually generated at different institutes and/or under different environments. We propose to perform in-depth research and development of new, powerful, and scalable indexing, query processing, and data mining methods for construction of scalable, efficient, and analysis-based, heterogeneous biological database systems.

MAIDS: Mining Alarming Incidents from Data Streams
J. Han,* M. Welge (National Center for Supercomputing Applications)
Office of Naval Research

There are many applications that require handling data in the form of data streams. Data streams are dynamic, high in volume, potentially infinite, and require multidimensional analysis. These characteristics pose a challenge for data analysis of data streams. This project is researching and developing the MAIDS system, which mines alarming incidents from data streams, with the following major analysis functions: multiresolution modeling using a tilted time window framework, multidimensional analysis using a stream ”data cube” model, online stream classification, online frequent pattern mining, online clustering of data streams, and stream mining visualization. Moreover, we are exploring intrusion detection and other kinds of applications for this technology.

Mining Dynamics of Data Streams in Multidimensional Space
J. Han*
National Science Foundation, IIS-03-8215

Stream data processing and mining represent an important, emerging class of data-intensive applications, characterized by dynamic data flows, only adaptive to single-scan algorithms, but often demanding fast responses. Most stream data resides at the primitive abstraction level, but most interesting patterns may be discovered at certain high levels of abstraction in multidimensional space. This project will develop effective, efficient, and scalable methods for mining the dynamics of data streams in multidimensional space, including the discovery of changes, trends, and evolutions of characteristics, clusters, classification models, and frequent patterns in data streams. Several important applications will be explored, including network intrusion detection, telecommunication and Web data flow analysis, and financial data flow analysis.

Mining Hidden and Evolving Linkages across Multirelational Data
J. Han*
IBM Faculty Award

Industry firms and government agencies have built up huge databases or other information repositories, which contain a large number of semantically interrelated relations, with voluminous structured or semistructured data and heterogeneous definitions. These data repositories can be constantly changing and evolving, bringing in an additional dimension of complexity and heterogeneity to the system. Since data across multiple relations and/or platforms contain rich semantic information, it is interesting and challenging to automatically, effectively, and efficiently perform linkage analysis and structure discovery to find hidden and evolving linkages across multiple relations, as well as in various kinds of data repositories across multiple platforms. Mining hidden and evolving linkages across multirelational and/or multiplatform data is one of the most demanding and challenging tasks in data mining, information retrieval, homeland security, digital government, bioinformatics, Web search and mining, and many other applications.

Mining Sequential and Structured Patterns: Scalability, Flexibility, Extensibility, and Applicability
J. Han*
National Science Foundation, IIS-02-9199

Sequential and structured pattern mining finds frequent subsequences or substructures in large sequence, structured, or semistructured data sets. These are important data mining tasks with broad applications, including analysis of customer purchase sequences, analysis of Web page structures and Web traversal patterns, understanding of disease treatments, scientific experiments, transportation and production processes, discovery of unusual events for intrusion, discovery of DNA sequences and molecule structures, and so on. The project is aimed at developing effective, efficient, and scalable methods for mining such patterns and exploring their important applications.

* Denotes principal investigator.
Mining Unusual Patterns in Data Streams
J. Han*
IBM Faculty Award (2002-2003)

Traditional data mining systems and methods assume that data resides on disks or in main memory and that a data mining process can scan the data sets multiple times to uncover interesting patterns. However, dynamic environments often quickly generate a tremendous amount of stream data so that the volume of data is too huge to be stored on disks or is too expensive to be fetched and scanned multiple times from disk. Moreover, applications may require real-time mining of unusual patterns in data streams, including finding unusual network or telecommunication traffic, pattern mining in video surveillance, and detecting suspicious online transactions or unexpected terrorist activities. We propose to develop efficient and scalable stream data mining methods for discovery of unusual patterns in data streams.

MotionEye: Querying and Mining Large Datasets of Moving Objects (NSF/SEIII)
J. Han,* O. Wolfson
National Science Foundation, IIS-05-13678

With huge amount of moving object data accumulated or streaming in, we propose to design and implement a MotionEye system for querying and hypothesis validation in moving-object databases and data streams and for data mining in moving-object databases and data streams respectively. The project will enable the development of more advanced information systems in homeland security, law enforcement, traffic control, and other domains that deal with moving objects.

On-Line Mining of Strange Moving Objects for Security Protection
J. Han*
Boeing, ITI Boeing RPS-14

We propose to investigate the issues related to the design and development of innovative methods for analyzing and mining of spatiotemporal information to find typical patterns and uncover suspicious motion in large datasets of moving objects. We will work together with Boeing researchers and application developers, study the real moving object data adaptable to Boeing applications, especially on security protection, develop scalable and effective data mining solutions to such data sets, and build a Motion-Guard system that can detect anomaly effectively on the fly.

SGER: DataScope: Viewing Database Contents in Multiresolution at Your Finger Tips
J. Han*
National Science Foundation, IIS-06-42771

The goal of this research project is to investigate issues in the design and development of a multiresolution data viewing system, called DataScope, which interacts with database systems based on visual comprehension and exploration of database contents in a flexible, user-friendly, multidimensional, and multiresolution way. The approach consists of investigation of the design principles and implementation considerations, including structures of dimensions; ordering of attribute values; layout of data values; flexible selection of layouts; attributes, and constraints; rich information display; and automated visualization interface generation and customization. The research integrates the technologies in database systems, data warehouse, data visualization, human-computer interaction, and data mining. It emphasizes an efficient implementation of the functions for visual exploration of database contents and scale-up of such functions for real-time usage. The research results are to be published in research forums on database systems, data mining, and human computer interaction, and data mining. It emphasizes an efficient implementation of the functions for visual exploration of database contents and scale-up of such functions for real-time usage. The progress of the project and the research results will be disseminated via the project website (http://www.cs.uiuc.edu/~hanj/projs/datascope.htm).

Structure Discovery and Database Integration by Data Mining
J. Han*
IBM Faculty Award (2003-2004)

Our world is filled with huge amounts of heterogeneous and interconnected data, and it is challenging to automatically, effectively, and efficiently perform knowledge discovery, information exchange, and database integration on various kinds of data repositories across multiple platforms. This project investigates the issues for the discovery and integration of database structures, linkages, and contents, and develops efficient mechanisms. We will especially focus on the automatic discovery of knowledge across multiple relations and the building-up of semantic linkages across multiple databases and information repositories to facilitate information analysis.

* Denotes principal investigator.
Data Management, Storage, and Query Systems for Scientific Applications
M. Winslett,* S. Mitra, A. Termehchy
U.S. Department of Energy, Center for Simulation of Advanced Rockets; National Science Foundation, ITR; U.S. Department of Energy, Scientific Discovery through Advanced Computing

Large-scale scientific applications pose new challenges for data management, including a unique environment and requirements for long-term data storage, unusual requirements for large-scale movement of data across the Internet, and a need for very lightweight and data-format-independent facilities for buffering, caching, parallel I/O, indexing, and querying. The Panda and Godiva projects are addressing all of these needs in close collaboration with scientists and engineers from University of Illinois at Urbana-Champaign and beyond.

User-Centered Adaptive Information Retrieval
C. X. Zhai,* X. Shen, B. Tan
National Science Foundation IIS CAREER Award, IIS-0347933

Existing information retrieval systems are inherently nonoptimal because the retrieval decision is primarily based on the current query and the document collection without considering information about the user and search context. For example, a traveler and a programmer may use the same word "java" to search for different information, but the current search systems would return the same results. This project seeks to break this limitation of the existing retrieval methods and formally develop a new retrieval paradigm called user-centered adaptive information retrieval, in which user information and search context are both exploited to improve retrieval performance. We will develop a client-side personalized search agent to optimize search accuracy by utilizing the previous queries, viewed documents, and other information about the user.

Educational Research

Building Communities: Recruiting and Retention of Underrepresented Groups in Computer Science
S. Kamin,* T. Wentling, C. Heeren
National Science Foundation, CNS-0420505

This project endeavors to increase the number of women and minority students entering and the percentage of these students remaining in the undergraduate programs of Computer Science Department. The University of Illinois leads five other central Illinois partners—computer science departments of Parkland Community College, Eastern Illinois University, Illinois State University, Heartland Community College, and Bradley University—in developing programs for women and minority students. These include outreach to high schools in our region and around the state, computer-oriented competitions for high school and college women, teacher training workshops, special orientation programs, and other activities.

Educational Uses of the Tablet PC
S. Kamin,* C. Peiper, B. Capitanu, C. Heeren
Microsoft, Hewlett-Packard

The Tablet PC holds great promise as an educational tool for several reasons. Its light weight allows it to be carried easily to class and its flat format allows it to be used in classes without blocking the student’s view. Because it is shaped like a paper pad, it is easier to read from than a laptop, so can effectively replace textbooks. Its writable surface allows for taking notes more easily and more quietly than a computer with a keyboard. We are developing a Tablet PC-based system, called eFuzion, for classroom use, either by a professor alone or in a tablet-equipped classroom. eFuzion allows for virtually unlimited interaction among students and professors. It is scriptable, so that instructors can customize their interactions to suit their teaching style and course content and educational researchers can test a variety of usage modalities without excessive development costs.

Graphics and Visualization

Automatic Global Multiscale Surface Parameterization
M. Garland*
University of Illinois

Many algorithms that operate on 2-manifold surfaces are most easily formulated over a planar parametric domain. This requires a parametric mapping of the plane onto the 2-manifold in question. Free-form polygonal meshes are not necessarily equipped with such a parameterization, and constructing a suitable one is a challenging problem. We are focused on developing efficient multiscale (multigrid) methods over irregular triangle meshes embedded in a 3-D Euclidean space. We also address the problem of constructing a global conformal structure over the surface, represented by a holomorphic 1-form. This global structure provides a global parameterization of the surface mesh, and provides a very convenient means of dealing with manifolds of arbitrary genus.

* Denotes principal investigator.
Everyday Shape Modeling  
M. Garland*  
* Denotes principal investigator.

National Science Foundation, ITR CCR-0086084

There are many tools available for modeling 3-D shapes. However, they are almost universally targeted toward experts and often exhibit steep learning curves. This project aims to produce tools that allow nonexperts to easily work with digital representations of 3-D shapes, just as the typical user can currently work easily with digital images. We emphasize simple interface modalities, such as sketching, that come naturally to most users. We are developing the technologies necessary for users to model shape by example. That is, a user first demonstrates a desired editing operation on a part of the model. The modeling system then automatically extrapolates the modification to the shape as a whole.

Processing Massive Polygon Meshes  
M. Garland*  
* Denotes principal investigator.

National Science Foundation, CCR-0098170

The past decade has seen the production of ever-larger polygonal datasets that users expect to manipulate and visualize. Extremely accurate laser scanners can produce precise models of real-world objects. Similarly, large-scale scientific simulations, such as computational fluid dynamics, produce huge amounts of data. Complex human artifacts, such as buildings and airplanes, are frequently modeled on computers in full detail. In all these cases, multigigabyte datasets containing billions of individual polygons are used. Our aim is to develop the means to store, process, and transmit extremely complex 3-D data efficiently. Our current work is focused on automatic data reduction, hierarchical data storage, and efficient streaming data representations.

Real-Time Texture Synthesis  
M. Garland*  
* Denotes principal investigator.

National Science Foundation, ITR CCR-0086084

Virtually all 3-D environments make extensive use of texture images to achieve a convincing level of realism. Large surfaces must frequently be covered with textures simulating a diverse range of materials such as wood, sand, brick, water, grass, paper, and patinated copper. From the design perspective, it is convenient to be able to generate large amounts of surface texture from a single example image. In this project, we are working on the development of near real-time methods that can generate arbitrary amounts of texture on-demand from sample images. We address the problems of generating texture images and of directly texturing 3-D surfaces from the given example.

Simplification and Multiresolution Modeling of Polygonal Surfaces  
M. Garland*  
* Denotes principal investigator.

University of Illinois

Extremely precise polygonal surface models are now widely available, in large part due to advances in scanning technology. Models of this complexity, perhaps several million polygons, frequently contain far more detail than required for the target application. This project is focused on the design and implementation of efficient algorithms for the automatic simplification of highly detailed polygonal surface models into faithful approximations containing fewer polygons. Using such simplification algorithms, researchers can also construct new multiresolution representations for surfaces that allow the application to treat an object at many different levels of detail.

Fluid Simulation and Control  
Y. Yu*  
* Denotes principal investigator.

University of Illinois at Urbana-Champaign

Many interesting natural phenomena involve fluid flows. Our goal in this work is to propose methods that produce physically plausible motion of fluids, which at the same time, assumes recognizable static or dynamic objects. Such techniques have many applications in advertising and filmmaking. For instance, "The Lord of the Rings" included scenes of water horses emerging from a flooding river. If we take smoke as an example, its density always tends to drift from a nonuniform distribution to a uniform one. Solving the proposed problem requires the maximum level of control of this process while maintaining a believable fluid appearance. We have developed an effective solution to this problem by looking for a careful compromise between controllability and physics equations. The basic idea is based on tracking the surface of target objects using velocity constraints and the Level Set Method. We have also developed a companion technique for controlling rapidly moving liquids, a simulation technique for flows constrained to a curved mesh surface, and a controllable motion synthesis technique for lightweight objects immersed in a gaseous medium.

Mesh Processing and Editing  
Y. Yu*  
* Denotes principal investigator.

University of Illinois at Urbana-Champaign

Mesh editing is a very important research topic in graphics. It is a routine practice for people in the game and movie industries to create new mesh objects by reusing existing ones, especially when the new objects do not exist so that
3-D scanning is infeasible. During the course of our research on fluid simulation, we unexpectedly discovered that the math behind fluids could also be applied to mesh editing. This discovery gave rise to a new technology that is based on gradient field manipulation, which implicitly modifies mesh vertices. This new technique is capable of producing desirable results with significantly less user interaction. This is because the gradient is a differential property that can be modified locally. Subsequent reconstruction from the modified gradient can give rise to a global effect, which would otherwise require much more user interaction. Based on this idea, we have developed algorithms for multiple types of mesh editing, including mesh deformation, merging, and smoothing. Recently, we have also developed a fast multigrid algorithm to accelerate the editing of large-scale meshes.

**Texture Analysis and Synthesis**

Y. Yu*

*University of Illinois at Urbana-Champaign*

Texture modeling has drawn much attention from researchers since the dawn of computer graphics because they contribute significantly to the photorealism of synthetic objects and images. My group is investigating analysis and synthesis tools for both 2-D and 3-D textures. We have introduced a new technique called feature-based texture synthesis. It is based on the observation that the human visual system is most sensitive to edges, corners, and other high-level features in images. 3-D textures refer to spatially varying geometric details that are relatively small but still visible. The presence of such small-scale details gives rise to a rich set of visual effects, including mutual shadowing, interreflection, occlusion and foreshortening, in addition to varying surface normal orientations. We have developed a new method for recovering the geometric details of 3-D textures by exploiting shadows. The rendered synthetic images of the recovered 3-D textures are comparable to the original photographs. We have also developed a technique for synthesizing bi-directional texture functions (BTFs), which form an image-based representation of 3-D textures.

**Visual Data Analysis and Representation**

Y. Yu*

*University of Illinois at Urbana-Champaign*

With advances in imaging technologies as well as physically based solid and fluid simulation technologies, new visual data of multiple dimensions have been produced at an unprecedented rate and scale. These new technologies bring new challenges to existing visual data modeling and processing techniques. One of the fundamental and challenging problems is how to efficiently represent, analyze, and visualize such a vast and ever-growing amount of visual data. Visual data exhibit two important intertwined characteristics. First, they are comprise of signals at many different scales or frequencies. Second, these signals have spatially inhomogeneous magnitudes. Multilinear models based on tensor approximation have caught researchers' attention recently. They are capable of generating a more compact representation of multidimensional data than traditional dimensionality reduction methods, such as principal component analysis (PCA). My group is investigating novel and effective tensor approximation techniques to exploit the aforementioned characteristics of visual data. The potential applications of our new technology include medical and scientific data visualization, data-driven rendering, and synthesis by examples.

**Human–Computer Interfaces**

**Developing Computational Tools that Amplify Human Creativity**

B. P. Bailey*

*University of Illinois*

Since its inception, the field of human-computer interaction has concentrated on developing techniques that enhance user productivity. Our work is pioneering a new direction of research in the field to understand how to enhance creativity. Creativity is a valued commodity throughout many work domains where producing creative outcomes is both expected and desired. Our research is investigating the design, implementation, and evaluation of interface representations, interaction techniques, and supporting tools that foster creativity. Our current work is grounded in the design domains, as these domains are replete with creative activity. Examples of recent projects include interactive sketching tools that support rapid exploration of behavioral ideas, interfaces that enable and encourage working with multiple design ideas in parallel, knowledge management systems that support design reuse, and interfaces that enhance example finding behavior.

**Principles and Systems for Managing Notifications**

B. P. Bailey*

*University of Illinois*

While engaged in memory-intensive tasks on the computer (e.g. programming, writing, and designing), users often receive notifications from peripheral applications. While these may provide valued information, such notifications typically interrupt the user’s ongoing task at inappropriate
moments. Our research involves conducting controlled experiments to better understand costs of interruption caused by notifications; examining cognitive theories from psychology and other disciplines to probe strategies for identifying less disruptive moments; building systems that automatically detect these moments during user tasks and that defer pending notifications until those moments; and deploying our systems in authentic environments to understand how they affect user behavior. We are also expanding our research to focus on reducing the number and severity of interruptions caused by mobile devices in everyday social situations.

**User Interface for Multiple Display Environments**

B. P. Bailey*

*University of Illinois*

Since the 1980s, the design of window managers has assumed a single person, single display configuration. However, recent advances in collaborative computing technology now make it possible to network multiple computers to form a unified workspace. Such workspaces are called multiple display environments. These types of computing environments introduce new challenges for the design of window managers because multiple users must now be able to manage application windows across multiple displays—breaking the traditional assumptions. To address these challenges, our research investigates the design, implementation, and evaluation of new interfaces and interaction techniques for collaborative multiple display environments. For example, we have developed a new interaction metaphor called a world-in-miniature that enables users to quickly and easily manage windows among any networked display. Our ongoing work is extending and enhancing this basic metaphor with new accommodations for privacy control and mechanisms for coordinating access to shared windows.

**Adaptive/Reflective Middleware System**

R. Campbell,* K. Nahrstedt,* R. Kravets,* L. Sha,* J. Tanner, P. DeRose

*Defense Advanced Research Projects Agency Grant, INT NBCH 1030017*

Lockheed-Martin; BBN Technologies; Johns Hopkins University; Scientific Research Corporation; Telecordia

The ARMS project investigates Multi-Layer Resource Management for next generation Navy battleships. This collaborative effort is developing a system that encompasses many major aspects of computing resource management. Our team's focus is the management of the human element of mission tasks, identifying and tasking the best user for each task, and locating the best hardware to perform that task within a distributed system platform. Intelligent algorithms dynamically adjust to current conditions when selecting users and also learn from a user’s previous performance. We are integrating current research in pervasive computing, allowing users to perform their tasks more efficiently by allowing them mobility that existing systems lack.

**Networking**

**Monitor and Control: Toward Dependable COTS-Based Real-Time Embedded Systems**

M. Caccamo,* G. Rosu*

*National Science Foundation, CNS-0720512*

The research tasks of this project aim at building a sound knowledge base and systematic design framework for dependable and predictable integration of next generation COTS-based embedded systems. In modern computer architectures, peripherals autonomously can initiate data transfers and contend for bus transactions. Unfortunately, temporal isolation guaranteed at the CPU level can be globally violated whenever the CPU tries to access the shared bus, which is currently locked by another master device. To guarantee the dependable and predictable behavior of next generation embedded systems, the following research tasks are pursued. First, the research introduces the novel idea of a "hardware server," implemented on a customized smart bridge. In the architecture being developed, a smart bridge separates a group of peripherals from the rest of the system, shielding the system itself from undesirable behaviors of peripherals. The smart bridge is implemented using an FPGA-based full system-on-chip (SoC). Second, relevant bus transactions from untrusted components are monitored since it can be very difficult or impossible to be shown that commercial off-the-shelf (COTS) devices satisfy system assumptions. The rationale is to optimistically assume that all assumptions hold and then to monitor the runtime behavior of COTS components against their assumed specifications. If violations are detected, then an appropriate recovery measure is taken. Monitoring is decentralized: events are filtered and communicated to FPGA-based monitors by corresponding smart bridges. In addition, specifications are synthesized into low-level monitors via translations to intermediate timed automata.

* Denotes principal investigator.
Ad hoc Wireless Communication Between Vehicles
R. H. Campbell,* S. Myagmar
Motorola, Inc.

We propose an ad hoc routing protocol with location service for vehicle-to-vehicle communication. As an example of feasibility, we developed and tested application prototypes of voice chat, location filtering, and roadside information service for moving vehicles. Our protocol takes into account the motion of vehicles on a highway. It broadcasts location updates only when the velocity or direction of a vehicle "space reservation" to avoid transmission collisions.

Efficient Resource Management for Controlled-Mobility Wireless Networks
E. Frazzoli,* L. Sha, P. R. Kumar, M. Caccamo, N. A. Neogi
National Science Foundation, CCF-0325716

A controlled-mobility wireless network (CMWN) is defined as a network of embedded devices endowed with computation, communication, and motion capabilities. The purpose of this project is the development of a new conceptual framework for the design, development, and operation of efficient and reliable networks with such characteristics.

IT-Based Collaboration Framework for Preparing against, Responding to, and Recovering from Disasters Involving Critical Physical Infrastructures
I. Gupta,* F. Pena-Mora,* G. Robinson, N. Contractor, A. Hollingshead
National Science Foundation Grant, CMS-0427089

In a new and novel study, scientists are looking to nature—specifically, to ants, bees, and viruses—for ways to improve human collaboration during disaster relief efforts. Supported by a five-year, $2.37 million grant from the National Science Foundation, and supplemented by the University of Illinois at Urbana-Champaign, the research team is attempting to draw inspiration from the collaboration patterns that honeybees and ants use, and the spread patterns that viruses typically take. The ideas the researchers develop will augment current collaboration among first responders to extreme events involving critical physical infrastructures. Epidemiological algorithms can be used in large groups of participants to spread, collect, and search for information. The resulting software systems can scale to networks with hundreds or thousands of first responders and withstand unresponsive participants and poor communication channels.

Systematic Design of Distributed Protocols—from Methodologies and Toolkits to Systems
I. Gupta*
National Science Foundation CAREER Award, CNS-0448246

Gupta and his research group will study, invent, and discover systematic "design methodologies" to enable creation of a variety of distributed systems that are scalable and robust under failures and attacks. Newly developed systems will include cooperative web caching, adaptive Grid computing paradigms, peer-to-peer systems, persistent distributed file systems, and disaster response and recovery networks. Protocols for these systems will be created by using novel methodologies such as ones that translate differential equations into equivalent distributed protocols, or those that compose existing distributed protocols in a systematic manner. This project will enable systematic translation of many natural phenomena from non-Computer Science fields, such as biology, into practicable and robust distributed system designs. The project integrates research and educational activities.

Transparent Data Sharing in Large-Scale, Dynamic Distributed Systems
I. Gupta,* G. Antoniu (INRIA), A.-M. Kermarrec
University of Illinois, Centre National de la Recherche Scientifique, l’Institut National de Recherche en Informatique et en Automatique Collaboration Grant

Distributed environments such as the Grid are large-scale and dynamic, for instance, they involve a large number of participating workstations that are joining and leaving at high rates. This collaborative project seeks to develop novel solutions for transparent and decentralized access to data in dynamic, large-scale, and distributed environments. The project will develop algorithms and strategies for replication, group membership, data consistency, resource discovery, and network size estimation. Many of these ideas are well-understood in fault-tolerant distributed systems research, but have not been explored for Grid environments.

Data-Centric Sensor Networks
J. C. Hou,* L. Sha, P. R. Kumar, N. Li, H. Zhang
National Science Foundation, Special Projects in Networking, ANI-0221357

In this research project, we first lay an integrated framework in which a comprehensive solution can be designed that comprises a set of component solutions at

* Denotes principal investigator.
each layer to achieve the targeted goals of data-centric sensor networks. Then, we consider under this unified framework, research issues along the following thrusts of research: hierarchical cluster formation and routing; topology control and power management; Quality-of-Service provisioning within/between clusters; MAC design for timely dissemination of delay-sensitive data; and empirical study with the use of Motes.

DAWN: Dynamic, Ad-Hoc Wireless Networks
J. C. Hou*
Army Research Office, ARO-W911NF-05-1-0246

In this project, we will strive toward two research thrusts. Along the theory thrust, we will, on the one hand, study the performance limits of wireless ad-hoc networks with respect to connectivity, lifetime, critical power, and their implication on protocol design, and on the other hand, devise localized and distributed algorithms for power control (a.k.a. topology control) whose performance approaches derived theoretical bounds. In particular, we will consider topology control under the physical model where the SINR determines the link quality (and hence the neighbor relation). Along the simulation thrust, we will develop a number of optimization techniques to reduce the overheads for event processing in wireless network simulation and engage in incorporating model checking in J-Sim so as to enable network performance evaluation and verification in the same framework. We will also study, through a tight, synergistic loop composed of measurements, modeling/analysis, and designed experiments for validation, how multipath fading, interference, and signal attenuation due to terrain and obstacles affect the channel/link behavior in IEEE 802.11-based wireless networks, and devise/incorporate channel models into simulators to provide high-fidelity simulation environments to the R&D community.

Design, Theoretical Validation, and Empirical Evaluation of a Class of Multicast Congestion Control Schemes
J. C. Hou,* Y. Gao, Y. Ge
National Science Foundation, ANI-0073725 (subcontract to Ohio State University)

In this research project, we explore the use of robust feedback control theory to design and implement a rate-based congestion control framework for multicasts, with the following design objectives: scalability, capability to adjust source sending rates to achieve TCP-friendliness, and (weighted) fairness in an analytically provable manner, capability to handle independent losses of the same packet, capability to deal with dynamic traffic/membership changes, and minimal router support.

Detection, Identification, and Tracking in Sensor-Cyber Multidomain Networks: Perimeter Surveillance and Plume Detection Applications
J. C. Hou*
Oak Ridge National Laboratories, UT-Battelle 4000052947

Under the national SensorNet initiative, Oak Ridge National Lab (ORNL), in conjunction with its university collaborators, has carried out initial deployments of a detection, identification, and tracking sensor-cyber network (DITSCN) in the Washington D.C. and Memphis Port areas. We have also identified three grand systems challenges that remain to substantially enhance the performance and robustness of these networks. We will carry out five synergistic research tasks: network formation by static sensor selection, placement, and coverage; mobile sensor coverage; protocols for QoS-aware sensor-cyber communication; sensor tasking protocol with temporal/spatial uncertainty management; and algorithms for plume modeling, detection, and tracking under uncertainty.

High-Fidelity, Integrated Simulation and On-line Management for Networked Communications Systems
J. C. Hou,* J. Meseguer
Boeing Company, 2006-06086-00-00

Several key issues must be addressed, especially in the case of large-scale networked communications systems, in order to realize the notion of simulation/emulation-assisted network management and control. In this project (and its proposed research and development efforts), we aim to address these key technical challenges, and to develop a scientific and engineering foundation for realistic representation of communications and network components and their low-level interaction. We will use J-Sim coupled with a metalinguage Maude for system verification and validation, as our base environment for carrying out the R&D efforts. Both software environments have been designed and developed in the Department of Computer Science at University of Illinois at Urbana Champaign.

Multiple Time-Scale Traffic for Next-Generation Internets
J. C. Hou,* G. He, H. Lim
National Science Foundation, ITR ANI-00-82861

In this research project, we address the following two-pronged problem: how to exploit the multiple time-scale

* Denotes principal investigator.
property of network protocols to facilitate effective coordination and integration of disjoint network controls for end-to-end Quality-of-Service (QoS), and how to exploit the multiple time-scale nature of Internet workload to achieve workload-sensitive traffic controls. The first problem is comprised of two key issues: sufficiency or separation conditions under two network controls and the effective coupling of protocols when time-scale separation is not available. The second program consists of three key issues: short-lived connection management using lightweight optimistic control, long-lived connection management using connection duration prediction and multilayered feedback control, and QoS amplification through workload-sensitive, end-to-end, and per-hop control.

Next-Generation Network Protocols and Services for Real-Time Systems
J. C. Hou,* C. Hu, A. Sobeit
Multidisciplinary Research Program of the University Research Initiative, AFOSR F49620-00-1-0330 (subcontract to University of California, Santa Barbara)

In this research project, we categorize network protocols into traffic specification and Quality-of-Service (QoS) translation/negotiation, admission control, unicast/multicast routing, resource reservation and adaptation, congestion and flow control, local traffic control and run-time packet scheduling, fault detection and fast restoration, security, and traffic regulation and QoS monitoring. We address the problems of provisioning QoS and fault tolerance in each category.

On Providing Quality-of-Service Control for Core-Based Multicast Routing
J. C. Hou,* J.-K. Lee
National Science Foundation, Advanced Networking Infrastructure and Research, ANI-9804993

We investigate the Quality-of-Service (QoS) extension to core-based multicast routing protocols such as Core-Based Tree, Simple Multicast, and Protocol Independent Multicast Sparse Mode. Specifically, we propose a set of enhancements in the member join/leave and state update/refresh procedures to facilitate the deployment of additive (e.g., end-to-end delay bound), multiplicative (e.g., packet loss ratio along a path), and concave (e.g., minimum bandwidth available) QoS. We also implement a proof-of-concept prototype in the FreeBSD UNIX operating system and conduct an empirical study on a lab testbed.

Toward Building a Performance-Predictable Wireless Mesh Network
J. C. Hou*
National Science Foundation, NSF CNS-06-26584

In this project, we aim to take a bottom-up approach, and tackle issues with the better definition and characterization of wireless links and their implications for higher-layer protocol design and optimization. We would like to understand how, and to what extent, wireless links are affected by PHY/MAC attributes and other environmental factors; characterize the behavior of wireless links in such a way that they become amenable to rigorous analysis and reasoning; and identify control knobs in the MACPHY layers with which the network capacity can be optimized. To make such a study with synergistic measurement, characterization, and design components, we will work with Champaign-Urbana Wireless Mesh Network (CUWiN), substantially extend/strengthen its operational software infrastructure, and lay a virtual device driver on top of firmware. The virtual device driver will include several functional modules that enable dynamic tuning of control knobs to improve network capacity, and export PHY/MAC attributes through well-defined APIs to facilitate cross-layer design and optimization. Specifically, we will conduct research through a synergistic measurement-design-experimentation loop along the following closely-related research and development thrusts: channel behavior understanding and modeling; better characterization of wireless links and its representation in virtual coordinates; capacity improvement through power control, carrier sense threshold tuning, and spatial diversity; and modular, virtual device driver that encompasses research components in CUWiN.

A Component-Based Software Environment for Simulating and Synthesizing Network Protocols in Large-Scale Networks
J. C. Hou,* R. Campbell, L. Kung, H. Kim
University of Illinois

In this project, we propose to design, implement, and evaluate a component-based software environment for a wide variety of emerging network architectures and applications. The environment expedites execution and simulates, emulates, and synthesizes network protocols and services in a systematic manner. We follow three research thrusts: We extend JavaSim to include base classes and packages for grid networking technologies. We investigate issues of parallelizing real-time process driven simulation engines and explore the use of fluid models, network calculus models, and rescaling techniques to
expedite simulation. In a related project, we build a software-programmable router platform, called CROSS, that is dynamically extensible, configurable, and able to predictably process network flows that require QoS-aware access to multiple resources. We will leverage JavaSim components as building blocks for CROSS/Linux router services, and realize differentiated multicast and secure video proxy systems as CROSS services.

An Integrated Simulation Environment for Robust, Multiple Time-Scale Traffic Control in Next-Generation Internets
J. C. Hou,* D. Chi, H. Kim, L. Kung
Network Modeling and Simulation, Defense Advanced Research Projects Agency, Information Processing Technology Office (subcontract to University of California, Berkeley); Cisco University Research Program

In this research project, we develop a component-based, compositional Java network simulation environment, called JavaSim. JavaSim is built upon an innovative autonomous component architecture that closely mimics the IC design architecture and a generalized packet switched network model that defines the generic structure of a node (either an end host or a router) and the generic network components. We also study (in the JavaSim simulation environment) the issues of fluid model-based simulation and large-scale parallel network simulation. JavaSim also provides a script interface to allow integration with different script languages such as Perl, Tcl, or Python.

An Open, Dependable, and Evolvable Software Infrastructure for Assisted-Living
J. C. Hou,* C. Gunter, L. Sha, M. Caccamo, K. Karahalios
National Science Foundation, CNS-05-09268

This project aims to design, implement, and evaluate infrastructure technologies to support assisted living, emphasizing: software infrastructure that allows disparate technologies, software components, and wireless devices of different protocol families to work together in a low cost, dependable, and secure fashion with predictable properties; and an interface that adapts this software infrastructure for technology-naive users and incorporates social transluency to provide user-controlled privacy.

To achieve the goal, the project seeks to develop the following: dependable, robust, plug-and-play infrastructure that coordinates sensing, localization, communication, and event/data management; quality-of-service annotation and analysis technologies that make assumptions explicit and machine checkable for COTS component interfaces; role-based access control and workflow modeling techniques for a rigorous security and privacy framework; QoS manager that handles the surges of workload and adjusts QoS settings of difference services; interference mitigation techniques that allow various wireless devices and protocol stacks to co-exist harmoniously and share the unlicensed frequency band, and a human computer interaction (HCI) component that addresses how directed and ambient information, reminders, and alerts should be presented to users and to what extent users can control the degree of information disclosure.

A Cross-Layer Approach to Energy Conservation in Mobile Ad Hoc Networks
R. Kravets*
National Science Foundation, ANIR 0347468

Current approaches to energy conservation in ad hoc networks narrowly focus on optimizations at only one layer, and some only focus on a specific mechanism in that layer. By ignoring the impact within a layer and on other layers, many approaches become ineffective and may even increase energy consumption. This research integrates optimizations across layers in the innovative Pulsar framework, resulting in network-wide energy-efficient communication. Pulsar captures the relationships between layers and provides energy-efficient communication using intelligent MAC and routing protocols and distributed algorithms to manage high-level network goals.

Application/System Quality-of-Service (QoS) Interface Capabilities
K. Nahrstedt,* W. J. Jeon, B. Kalter, J. H. Seo
National Aeronautics and Space Administration, NAG 2-1250

Researchers are investigating application-system Quality-of-Service (QoS) interface capabilities for visual tracking distributed applications. The interface between the application and the underlying QoS-aware resource management system must provide several important functionalities: application QoS application programming interface, translation between the application QoS into the system QoS parameters, integrated reservation coordination policies and protocols to avoid and prevent deadlock situations, adaptation policies and their application enforcement, and others. These functions will reside in the end-system management entity called the QoS Broker, which represents the application/system interface for provision of end-to-end QoS guarantees.

* Denotes principal investigator.
H-Media: The Holistic-Multistream Environment for Distributed Immersive Applications
K. Nahrstedt,* R. Campbell,* M. Caccamo
National Science Foundation, CNS-0720702

Three-dimensional tele-immersive collaborative environments are becoming a reality. The emerging tele-immersive (TI) technology empowers and enables collaborative interactions and a plethora of new applications among geographically distributed sites. TI technology allows creation of a cyber TI room, where geographically separated users can jointly perform physical activities such as dance or exercise. This project is working to take this vision further and allow users to participate in simultaneous TI sessions and to cyber-walk between TI rooms. To achieve the vision for the TI rooms, the underlying cyber-physical infrastructure must consider both streams of 3-D data as a first class object in its design and in its deployment, and holistic end-to-end management of the multistream environments for each TI room. Hence, the project is developing a holistic multistream environment for distributed applications (H-MEDIA). Researchers will investigate the following: system architectures with correlated multistreaming; real-time virtualization of resources for resource isolation between individual TI rooms and switching (cyber-walk) between rooms; end-to-end configurable, robust, and fault-tolerant virtual networks for different rooms; and adaptive configuration and system management that will yield customizable, stable, adaptable, available, and robust individual TI rooms. H-MEDIA research will have impact on communities in computer science and also on medical, social science, and other domains. The H-MEDIA project will also result in educational benefits such as involving graduate students research in very novel TI technologies, inclusion of undergraduate students, and impact on education in other disciplines such as new teaching of choreography in TI environments, as well as many others.

Hybrid Adaptive Algorithms for End System Middleware
K. Nahrstedt,* B. Kalter, B. Li

Current distributed multimedia applications demand Quality-of-Service (QoS) from the supporting system. However, within the QoS demands, lower level transport facilities may not constantly provide guaranteed QoS without perturbation. In this scenario, researchers are investigating hybrid adaptive algorithms in the middleware level of end systems to perform QoS adaptation on a critical QoS metric. The research concentrates on analysis of QoS adaptation in dependence of system resource availability changes by applying theories from digital control systems.

QoS Routing
K. Nahrstedt,* J. Qian, L. K. Shan

The task of Quality-of-Service (QoS) routing is to find a path in the network that satisfies constraints on such metrics as bandwidth, delay jitter, and cost. This study focuses on QoS routing algorithms and their design within routers. The problem of finding a path with constraints on two or more additive metrics (delay and delay jitter) is NP-complete. This research concentrates on heuristic algorithms and study of the family of distributed and hierarchical routing algorithms to solve the multiconstrained routing problem. The QoS routing solutions are applied to point-to-point as well as multicasting scenarios.

QoS-Aware Resource Management
K. Nahrstedt,* K. Kim, A. K. Viswanathan, J. Wang
Partnerships for an Advanced Computational Infrastructure

Operating systems and communication systems need new algorithms, services, and protocols to support processing of audio/visual streams according to Quality-of-Service (QoS) specification. This project concentrates on the CPU brokerage service with advanced reservation, admission, scheduler, and adaptation control for soft real-time and non-real-time tasks. At the communication level, researchers provide IntServ bandwidth brokers in the edge networks and DiffServ brokers within the backbone routers to provide end-to-end guarantees.

Operating Systems and Security

Choices: A Reliable and Secure Operating System for Mobile Devices
R. H. Campbell,* J. Hou,* Z. Anwar
National Science Foundation, CNS 03-05537

Distributed denial of service, man-in-the-middle attacks, message tampering, eavesdropping, and replaying threaten to cripple the Internet infrastructure. They are especially harmful to killer applications for the Internet such as voice over IP (VoIP) and voice over wireless. There is a need to develop innovative strategies to detect, mitigate, and counter these threats. Unfortunately various key components are required to realistically model a large VoIP infrastructure and study its vulnerable spots. J-Sim is a
composable and extensible network simulation and emulation environment. We extend J-Sim to include representative security mechanisms/policies for VoIP such as IPsec, firewalls, Media Gateways, Soft Switches, key distribution and authentication mechanisms, and popular VoIP protocol stacks such as RTP, SIP, and H.248. We also provide various attacker models and IDS mechanisms to allow vendors to plug in their VoIP components and test them for vulnerabilities in a controlled and simulated environment before actual deployment. In addition we are exploring the use of virtualization techniques and reference monitors to choose secure paths for VoIP information flows.

**Composing Security in Large-Scale Cyber-Infrastructures**

R. H. Campbell,* S. R. Katasani  
National Science Foundation

Present day cyber infrastructures like the power grid are very complex assortments of various devices with different security requirements and differing ability to provide security. Not all the devices in the system can provide the necessary security according to the enterprise policies, but the presence of these devices is indispensable. In such a scenario, it is really difficult for the system managers or the system administrators to monitor and maintain these devices. We are developing a methodology that will allow a system administrator or a manager automatically to analyze the system and reason about the security and decide the necessary security measures to install. In order to achieve this goal, we developed a representation for modeling the cyber assets of an organization based on workflows and common information models. In our modeling formalism an organization can be defined as a combination of subjects, objects, services, tasks, and communication protocols.

**ITR: Active Information Spaces Based on Ubiquitous Computing**

National Science Foundation, CCF-00-86094

The project researches a new form of operating system to manage a model of computing called an Active Space. This model integrates physical spaces that contain ubiquitous computers into a computational environment that supports human activity and applications. The physical space, augmented with communicating computer devices, becomes a distributed computing system. Gaia, an operating system for Active Spaces, will accommodate diversity by exploiting standards for interoperation and cooperation. System services track, authenticate, and support mobile users with reconfigurable graphics, multimedia, and Active Space applications. A unifying object bus, component model, and adaptive stream model extends plug and play to distributed mobile computers within physical spaces like cities, buildings, and rooms. Active Spaces have the potential for creating multibillion dollar industries. Automated surgery, collaboration, and engaged learning are a few of the compelling examples.

**Mobile Sensor-Network Authentication**

R. Campbell,* V. Welch,* C. Andrews, P. Naldurg, H. Khurana  
Office of Navy Research

Authentication is a critical security requirement for sensor network nodes and provides a high quality of assurance in a hostile deployment scenario, when it is important for a data-gathering source to confidently verify the origin of sensor data. We focus on higher-end sensors that have significantly more processing power and memory than first-generation sensors. We question some of the existing weak cryptographic protocols and investigate the feasibility of using limited public-key encryption to address the sensor origin authentication problem. The challenge of key distribution is addressed by imprinting sensors with public key certificates, as well as the corresponding private keys, before being deployed. A line-of-sight transmission may be used to update the key.

**Security and Configuration of Software Defined Radios**

R. H. Campbell,* S. Myagmar  
Various Donors

Reconfigurability of software defined radios (SDR) supports integration and co-existence of multiple radio access technologies on a general-purpose radio equipment. An SDR terminal is able to switch its operating mode by configuring its radio parameters and component composition to suit the appropriate radio access technology, user preferences, and local conditions. The main challenges are how to provide a methodology to dynamically and securely configure software components originating from several, different vendors, and how to remotely attest the validity of the radio configuration to external parties such as a network operator or service provider.

* Denotes principal investigator.
Pediatric Vaccine Formulary Optimization and Analysis
S. H. Jacobson,* S. N. Hall, R. A. Proano
National Science Foundation, DMI-0457176

The objective of this project is to formulate and analyze operations research models and algorithms for addressing childhood immunization vaccine formulary design issues. Healthcare decision-makers are being faced with an overwhelming number of pediatric vaccine choices, with no basis for comparing and evaluating their different vaccine selection options. This project formulates discrete optimization problems, models, and dynamic programming algorithms that capture the key features of the recommended childhood immunization schedule, as well as the restrictions and requirements imposed by the ACIP and the Food and Drug Administration (FDA). It also uses these models and algorithms to assess the economic and societal impact of new vaccines and immunization policy requirements, as well as providing a decision-making tool for healthcare administrators based on the problems and algorithms obtained, including the vaccine completion problem, the limited budget problem, and the balancing problem.

A Heuristic Design Information Sharing Framework for Hard Discrete Optimization Problems
S. H. Jacobson,* L. A. McLay, S. N. Hall, H. Kaul, G. Kao
Air Force Office of Scientific Research, FA9550-04-1-0110

The objective of this project is to study and develop simultaneous generalized hill climbing (SGHC) algorithms as an algorithmic framework for information sharing in discrete optimization problems. This framework is used to design hybrid heuristics that allow different heuristics to share information when addressing the same problem, and to design hybrid neighborhood functions that allow different neighborhood functions to share information when using the same heuristic applied to the same problem. This project introduces the SGHC algorithm framework for information sharing across sets of related discrete optimization problems, provides guidelines on how to use this framework to design hybrid heuristics that effectively trade off convergence versus finite-time performance, provides guidelines on how to use this framework to design hybrid neighborhood functions for heuristics that enhance their performance, and computationally evaluates the application of SGHC algorithms to a variety of test bed, large-scale, real-world discrete optimization problems. The primary application for this research is a search and rescue problem, where several possible search and rescue strategies must be considered to determine the optimal strategy.

A Study of Aviation Access Control Security Systems
S. H. Jacobson,* L. McLay, A. Nikolaev, A. J. Lee
National Science Foundation, DMI-0114499

International terrorism inflicted on the nation's aviation system poses a significant threat to the economic and political infrastructure of the United States. Aviation security technologies in airports throughout the United States provide an important line of defense against such threats. It is a challenge to determine how to optimally determine which security technologies to purchase as well as where to deploy such technologies and how to use them most effectively. The objective of this research project is to develop operations research models and algorithms to address these questions. The results of this project will be used to develop strategies to improve the security of the entire national airspace system through a systematic process of cost-effectively allocating aviation security resources.

Security and Privacy in Open Systems
M. Winslett,* A. Lee, L. Olson, M. Rosulek
Defense Advanced Research Projects Agency, National Science Foundation Information Technology Research (medium), NSF-Information Technology Research (large)

The need for rapid response to opportunities and threats has become the impetus for a move to computing systems where resources are shared across organizational boundaries: open systems. Example open systems range from supply chain management to the semantic web, grid, and peer-to-peer computing. The TrustBuilder project is addressing the unique security needs of open systems, with a focus on the development of trust negotiation as an approach to authorization and authentication in open systems.

Parallel Processing

Data Parallel Programming
L. V. Kale*
University of Illinois

Highly regular, array-oriented computations constitute a significant majority of computation-intensive scientific and engineering applications. Data parallel languages, such as High Performance Fortran, simplify and support the parallelizations of such applications. This project explores novel techniques for making such languages more efficient and flexible. One of the techniques developed

* Denotes principal investigator.
causes different parallel subcomputations to overlap, increasing the tolerance of the performance to communication latencies. A rich set of intrinsics is being developed. The system provides flexible intermodule connectivity, so a high degree of reuse of parallel software becomes possible. Also, data parallel components can be integrated with nondata parallel components at will.

**Debugging and Performance Feedback for Parallel Programs**

L. V. Kale*

*University of Illinois*

Debugging parallel programs and improving their performance is a daunting task. Researchers are developing debugging and performance feedback tools that can maintain and exploit refined application-specific data better than is possible with traditional techniques. Such tools can be significantly beneficial in developing and improving parallel programs. This work is carried out using the Charm/Charm++ object parallel programming system. It is being extended to other languages in a multilingual framework. By exploiting "specificity" at language constructs, and by recording specific events, it becomes possible to provide visual feedback on performance and to suggest improvement paths via expert analysis.

**Dynamic Load-balancing Strategies**

L. V. Kale*

*University of Illinois*

A small grain of parallelism is essential and natural in executing irregular symbolic computations. The efficiency of a parallel processing system depends on how uniformly these granules of action are distributed to processors. Researchers have developed a dynamic load-balancing scheme that speedily distributes newly created work to the "needy" processors. It employs a corrective redistribution component and saturation control (not moving pieces of work around while everyone has sufficient work). Also, new strategies appropriate for the current generation of multicomputers are being designed. Some of these strategies support prioritized load balancing and control memory requirements simultaneously.

**Highly Parallel Discrete Event Simulation**

L. V. Kale,* T. Wilmarth

*University of Illinois*

Many complex systems are characterized by the presence of asynchronous events. A flexible manufacturing system, a digital circuit, and vehicular traffic in cities are examples of such systems. Neither analytical solutions nor time-marching simulations are appropriate for modeling such systems; they are modeled by discrete event simulations. The efficiency of many industrial processes is dependent on their speedy modeling. Researchers are engaged in developing a parallel computing-based approach for this challenging problem. The objective is to develop easy-to-use modeling tools and enable their simulation on highly parallel supercomputers and in workstation clusters.

**Java Extensions for Parallel Computing**

L. V. Kale,* M. Bhandarkar

*University of Illinois*

Java has emerged as a dominant object-oriented language that could replace or augment C++ for building large-scale parallel applications. Several studies show that Java boosts programmer productivity because of its well-conceived design, garbage collection, and standard libraries. For this project, researchers implemented a prototype parallel extension to Java that provides dynamic creation of remote objects with load balancing and object groups. The language constructs are based on those of Charm++. The prototype is implemented using the Converse interoperability framework, which makes it possible to integrate, in a single application, parallel libraries written in Java with modules in other parallel languages.

**Parallel and Distributed Object-Oriented Programming**

L. V. Kale,* M. Bhandarkar, R. Brunner

*University of Illinois*

This project extends earlier research on Charm, a portable parallel programming language based on message-driven objects. Message-driven execution makes it possible to adaptively overlap computation and communication, even across multiple modules. Charm supports dynamic creation and load balancing of parallel objects and specific information-sharing abstractions. Its branched chare and chare array constructs facilitate interfacing parallel modules and implementation of distributed data structures. Charm is highly suitable for irregular parallel computations. Several applications and libraries are written using Charm. Current research includes support for heterogeneity and client-server environments.

**Parallel Execution of Speculative Computations**

L. V. Kale*

*University of Illinois*

A large class of interesting computational problems has the following property: if one attempts to solve them in parallel, one often ends up solving subproblems that are not needed or that are not solved in a sequential execution.

* Denotes principal investigator.
This leads to anomalous behavior. The speedups may vary from sublinear to superlinear from run to run and may increase or decrease with the addition of processors. This research is aimed at obtaining consistent and monotonically increasing speedups for such computations. Such problems arise in state-space search, branch-and-bound, game-tree search, planning, and theorem proving. Each of them requires a different set of techniques. Successful in obtaining satisfactory results for state-space search, the research team is working on other problems.

Run-Time Framework for Multilingual Interoperability among Parallel Languages

L. V. Kale,* M. Bhandarkar, J. Yelon, R. Brunner
University of Illinois

To tackle the difficult problem of parallel software development, many parallel languages are being developed, each with its own unique features and advantages. To benefit from this multitude of languages, one should be able to compose modules written in different languages into a single application program. This is difficult because of different scheduling models assumed by each language run time. This research is aimed at developing Converse, a framework for facilitating such interoperability. Converse also simplifies development of the run-time systems for new languages. Converse includes components supporting flexible threads, message passing, and processor scheduling. Several languages are being implemented using Converse.

CADRE: A National Facility for High-Performance I/O Characterization and Optimization

D. A. Reed,* G. Wang, D. Wells, Y. Zhang
National Science Foundation, EDA 99-75248

To catalyze research on I/O system design, analysis, and optimization for scalable, parallel, and distributed systems, this work will create a national facility to disseminate I/O characterization as well as optimization tools and data for quantitative study of I/O systems. The work will extend, document, and distribute a multilevel I/O characterization toolkit with logical and physical I/O instrumentation, statistical and visual data analysis tools, and documentation; instrumented I/O libraries, including MPI-I/O and HDF; instrumented, I/O-intensive applications; documented I/O traces in a portable data metaformat; and I/O trace analyses based on statistical analyses, hidden Markov models, and time series analysis that can be used for constructing configurable I/O benchmarks.

Human–Computer Interactions and Interfaces

Collaboration Support for the Development and Evolution of Complex Systems

D. A. Reed,* S. Kaplan,* H. George, Y. Huang, C. Lu, B. Schaeffer, E. Shaffer, S. Whitmore, S. Wu
Defense Advanced Research Projects Agency, F30602-96-2-0264

Complex systems are developed and evolved by groups, often separated in time and space and using disparate tools and techniques. No circumscribed tool suite can encompass all future needs. Instead, software support and collaboration environments must themselves evolve and grow to meet the needs of collaborative groups. As a solution to this problem, this research is focused on developing an extensible collaboration framework that supports the full range of activities undertaken by current team members, integrates the wide variety of tools and artifacts they use to realize their goals, and supports the evolution of the development team and its processes over time.

Intelligent Information Spaces: A Testbed to Explore and Evaluate Intelligent Devices and Augmented Realities

D. A. Reed,* R. H. Campbell, R. Kravets, M. D. Mickunas, K. Nahrstedt, L. Sha
National Science Foundation, EIA 99-72884

To support information environments where ubiquitous, intelligent devices unobtrusively share data, preferences, and contexts about users and their movement among environments, this project is developing interoperable component architectures for device coordination, seamless object communication for user quality of service, and adaptive user context and modality management. The goal is to define a software architecture capable of enabling a mobile, responsive, and contextual information environment where a broad collection of high-end data display and visualization systems, low-power mobile devices, and "smart" devices with widely varying capabilities are seamlessly integrated using dynamically tailored software components.

Intelligent, Adaptive Performance Analysis (Intelligent High-Performance Computing on Toys)

D. A. Reed,* P. DeRose, C. Lu, C. Steffen, D. Wells
National Science Foundation, ACI 02-19597

This research explores the utility and performance of low-cost game system clusters for scientific computing and high resolution visualization. Assessment will be conducted using a suite of adaptive performance analysis tools that support both offline and online performance.

* Denotes principal investigator.
optimization and their application to a suite of scientific and visualization codes. Research will focus on three areas: offline, multilevel performance instrumentation of applications and system software; online, adaptive selection of multiversion code execution; and experimental assessment using large-scale scientific applications and visualization software. Performance analysis software developed via this project will be distributed widely, both in binary and source code forms.

Language-Directed Performance Prediction and Analysis
D. A. Reed,* D. Padua,* R. Aydt, K. Mahesh, Y. Zhang
Defense Advanced Research Projects Agency, N66001-97-C-8532

A multi-institution collaborative project

Achieving a large fraction of peak performance on parallel and metacomputing systems has proven difficult. To address this challenge, researchers are developing an integrated performance modeling, measurement, analysis, and prediction environment that will allow application and system developers to explore the performance implications of software and hardware design choices for extant systems, hypothetical systems, and combinations of the two. This work is based on compiler-supported program annotation, symbolic performance scalability models, integrated performance instrumentation, and comparative analysis of multiple system configurations.

Scalable Performance Analysis Tools
D. A. Reed,* R. A. Aydt, D. Israel, J. Karim, J. Mainzer, B. Schaeffer, J. Wendling
National Computational Science Alliance

A multi-institution collaborative project

As part of the national Partnerships for an Advanced Computational Infrastructure (PACI), this research team is focusing on three interrelated areas of performance analysis for scalable parallel systems: augmentation of current tools for performance analysis of parallel codes; analysis of the interactions among application I/O patterns, file systems, and I/O hardware configurations; and design and testing of an adaptive resource management infrastructure that uses real-time performance data to interactively and automatically choose and configure runtime resource management policies. The results of this work are being shared throughout the national PACI alliance.

Parallel Programming Patterns
M. Snir*
U.S. Department of Energy, DE-FG02-03ER25560

We study and classify programming patterns that are prevalent in high-end scientific codes and that describe key activities of programmers that develop high-end scientific codes. We next examine how well these programming patterns are supported by several new parallel programming environments. This work is part of the activities of the Center for Programming models for Scalable Parallel Computing.

Programming Languages, Formal Systems, and Software Engineering

Actor Coordination Abstractions, Semantics, and Implementation
G. Agha,* C. Varela
University of Illinois

This research focuses on the complexity of expressing interaction and coordination in Web-based computing. Researchers are working on providing high-level mechanisms to manage the complexity of scaling up computations over the Web, piggy-backing on the availability of Java byte-code for portability. The project defines several actor-based abstractions (casts, directors, messengers) to effectively harness the power of the World Wide Web as a global computing infrastructure. Groups of actors, or casts, represent an abstraction unit for naming, synchronization, migration, composition, and load balancing. Each cast contains a director, and intercast communication is performed via special actors named messengers.

Agent Generation and Control
G. Agha,* N. Jamali, P. Thati
U.S. Air Force Office of Scientific Research, F49620-97-1-0382

Agents provide a natural abstraction for using geographically distributed computational and memory resources. Agents are autonomous mobile actors that may be invoked to satisfy specific goals that may require traveling across physical and economic boundaries. Agents and agent ensembles can exhibit resource consumptive or otherwise unsafe behavior, raising security and resource management concerns. Agents must, therefore, be limited by the resources they consume in pursuing a goal. The project is developing concepts necessary to provide linguistic and system support for defining multiagent

* Denotes principal investigator.
architectures. A related goal is to extend the mathematical theory of actors to allow reasoning about multiagent systems.

Customizable Coordination Services for Large-Scale Network Embedded Systems
G. Agha,* P. Chang, P. Thati, R. Ziaei
Defense Advanced Research Projects Agency,
F49620-97-1-0382

The focus of this research is on developing application independent services to coordinate large scale network embedded systems. The coordination services will use customization and composition to enable dynamic adaptation in uncertain environments. The approach is to define algorithms that are based on stochastic models of system behavior, which enable the research team to represent the incompleteness in information about the current global system state as well as the unpredictability of the environment. The operational model uses probabilistic transitions rather than simple nondeterministic interleavings of actions, and it explicitly accounts for duration of transitions. The goal is to develop algorithms that provide for coordination in real-time and that guarantee the desired properties with sufficiently high probabilities. Examples include algorithms for approximate consensus (such as approximate synchrony), recovery, and hierarchical coordination. The algorithms will be implemented to provide a code basis for application independent coordination services. The implementation strategy is to build a repository of basic coordination services using reflective middleware. The goal will be to derive more complex algorithms based on simpler core resource management services.

Parametric Models for Large-Scale Agent Systems
G. Agha,* N. Jamali, P. Thati, R. Ziaei
Defense Advanced Research Projects Agency,
F30602-00-2-0586

A goal of this research is to develop mathematical models to support the analysis and modeling of complex, large-scale agent systems. Instead of simple nondeterminism, the new theory will represent behavior stochastically. Moreover, instead of the current approach of using input-output behavior of individual agents, it will allow the behavior to be parametric in terms of variables that represent aggregated behavior of large numbers of agents. The operational model uses probabilistic transitions over an abstract representation of the current state of the system. The use of statistical techniques on this model for aggregating behaviors opens up the possibility of studying conditions under which either a stable equilibrium or chaotic behavior may occur. Another goal is to develop a radically different logical framework for expressing properties of large-scale agent systems. The framework is inspired by Quantum Logics, which allow the expression of testable properties. This is in contrast to the usual algebraic approach that assumes every sentence (whether testable or not) can be assigned a truth value. Specifically, this research will enable macroscopic properties to be expressed without implying assertions about how they arise.

Software Architectures for Distributed Systems
G. Agha,* M. Astley
University of Illinois; National Science Foundation, CCR 9619522

The term middleware describes a set of services for integrating components of a distributed application, such as coordination and communication mechanisms. Recently, middleware services have been developed that support fault-tolerance, security, and other high-level policies. Such services have a fixed semantics, their implementation being influenced by the semantics of the application and the nature of the execution environment. The goal of this research is to provide a modular framework for developing middleware services. The project is formulating theoretical, linguistic, and run-time support for developing the needs of a particular application. Particular attention is paid to placement and mobility issues and vertical integration requirements.

Specifying and Deriving Mobile Systems
G. Agha,* P. Thati, R. Ziaei
U.S. Army, JHU 8812-48151

This research is focused on studying formal methods for specifying and verifying distributed software systems. The objective is to use automated deduction tools to reason about certain properties of mobile agents in open distributed systems. More specifically, security issues in authentication protocols and agent design are being studied. The project is formalizing an appropriate semantic framework that captures the fundamental properties of mobile computing and simplifies the task of reasoning. A specification language and logic will be developed based on the semantic framework. Finally, automated reasoning environments will be explored to find a suitable platform to implement the reasoning system.
Studies in Distributed Database Systems
G. G. Belford*
University of Illinois

Distributed databases can provide enhanced data availability to support applications such as office automation and computer-aided design and manufacturing. But first, further work is needed on problems such as maintaining the consistency of replicated data in an efficient, reliable manner. This research addresses these and related problems by the development of provably correct, reliable protocols and by the analysis and simulation of competing protocols and algorithms.

Run-Time Code Generation for the Masses
S. Kamin,* B. Aktemur, P. Morton, M. Katelman
National Science Foundation, CCR 0306221

The goal of this project is to explore an easy-to-use approach to developing run-time code generators for Java. The technique we employ is "compositional compilation," which provides power similar to macros (that is, "generative components"), while allowing for run-time code generation, which we believe is essential to making code generation practical. Our previous research resulted in the Jumbo system, which consists of a Java compiler and an API to support run-time code generation. We are studying ways of generating code more efficiently and allowing for safe use of software components. We are also applying the method of compositional compilation to other languages.

The Memory Model for Java
W. W. Pugh* (Univ. of Maryland), S. Adve
University of Maryland; University of Illinois

A multi-institution collaborative project

The memory consistency model for a multithreaded programming language determines the ease of programming and possible hardware and compiler optimizations. The Java programming language is perhaps the first commercially successful language to incorporate threads as first class objects. Unfortunately, the memory model of Java is incompletely and incorrectly specified. An expert group of seven researchers/companies has been established to fix the Java memory model. It has resulted in a new memory model specification for Java that is the first memory model specification to incorporate known hardware and compiler optimizations while preserving the security and safety features of the Java language. This specification will impact all programmers of multithreaded Java programs, Java compiler writers, and hardware designers.

Scalable Formal Methods for Multidimensional Components
G. Rosu,* J. Meseguer*
National Science Foundation, CCR 0234524

This research investigates scalable formal methods and their combinations as a way to reduce the gap between formal methods and software practice. These methods include domain-specific certification and runtime verification and monitoring. Various prototypes are developed: safety policy certifier for units of measurement; coordinate frame safety certifier; optimality state estimation certifier; and runtime verification and monitoring prototypes. These are experimentally evaluated using the NASA-HDCP testbed. This research is expected to lead to advances in software technology and to benefit advanced education. Novel combinations of software synthesis, certification, and monitoring lead to new, powerful, dependable software development methodologies that ensure safe execution with little or no overhead.

Real-Time and Embedded Systems
SoD: A Feedback-Based Architecture for Highly Reliable Embedded Software
T. Abdelzaher,* L. Sha, M. Caccamo, D. Marinov
National Science Foundation, CNS-0613665

The focus of this project is on developing a theory and architectural framework for incorporating feedback control as a main principle of software engineering. This proposal develops scientific foundations, tools, and architectural design principles to improve the reliability and reduce the development cost of mission-critical software. Two common approaches to ensure reliable system behavior are: scientific foundations, tools, and architectural principles for ensuring elimination of errors; and foundations, tools, and architectural principles for ensuring tolerance to such errors. While the first solution is sufficient, building completely error-free large-scale systems has been an elusive exercise as evidenced by experiences with today's engineering artifacts (e.g., recalls on vehicles, cascading power blackouts, and so forth). Consequently, a new approach to robust software design, development, and certification is needed that addresses the increasing cost and safety implications of the current practices. Feedback control has proven to be a very successful tool for ensuring correct behavior of complex, poorly modeled systems in the presence of uncertainty. This project applies principles of feedback control and stability envelopes to the engineering of software systems.

* Denotes principal investigator.
SoD-TEAM: A Feedback-Based Architecture for Highly Reliable Embedded Software
T. Abdelzaher,* L. Sha,* M. Caccamo,* D. Marinov*
National Science Foundation, CNS-0613665

The focus of this project is on developing a theory and architectural framework for incorporating feedback control as a main principle of software engineering. This proposal develops scientific foundations, tools, and architectural design principles to improve the reliability and reduce the development cost of mission-critical software. Two common approaches to ensure reliable system behavior are: scientific foundations, tools, and architectural principles for ensuring elimination of errors; and foundations, tools, and architectural principles for ensuring tolerance to such errors. While the first solution is sufficient, building completely error-free large-scale systems has been an elusive exercise as evidenced by experiences with today's engineering artifacts (e.g., recalls on vehicles, cascading power blackouts, and so forth). Consequently, a new approach to robust software design, development, and certification is needed that addresses the increasing cost and safety implications of the current practices. Feedback control has proven to be a very successful tool for ensuring correct behavior of complex, poorly modeled systems in the presence of uncertainty. This project applies principles of feedback control and stability envelopes to the engineering of software systems.

M. Caccamo*
National Science Foundation, CNS-0237884

This CAREER project is a combined research and educational project in the area of resource management for highly dynamic real-time systems with physical constraints. The research project investigates three avenues. The first is collaborative scheduling: Distributed rate adaptation and collaborative resource reclaiming techniques are explored to mitigate the effects of highly dynamic workloads in a distributed system with tightly coupled real-time tasks running on different nodes. The second avenue is prioritized medium access control (MAC) with rate-adaptive messages: This research considers how to prioritize medium access to guarantee bounded delay, and how to implement distributed rate adaptation. The third avenue is template design: Off-line template design will be explored for managing multiple physical constraints, such as energy and thermal bounds in real-time CPU and network resource scheduling.

ITR: Collaborative Research: Efficient Resource Management for Controlled-Mobility Wireless Networks
E. Frazzoli,* L. Sha, P. R. Kumar, M. Caccamo, N. A. Neogi
National Science Foundation, CCF-0325716

A controlled-mobility wireless network (CMWN) is defined as a network of embedded devices endowed with computation, communication, and motion capabilities. The purpose of this project is the development of a new conceptual framework for the design, development, and operation of efficient and reliable networks with such characteristics. The overall objectives of the basic research proposed in this project are: the development of a conceptually sound, consistent, and complete framework for the analysis of the interactions between competing computation, communication, and motion control requirements, arising in the design of CMWNs; the design, analysis, and performance characterization of distributed algorithms and communication protocols for provably efficient and adaptive CMWNs; the specification, implementation, and verification of software, middleware, and networking services for the deployment of representative CMWNs.

Defect-Tolerant System Integration and Evolution
L. Sha*
U.S. Office of Naval Research, Sha 2063

Large software systems are developed by integrating software components. Unfortunately, many complex software components often contain defects. On the other hand, the technology exists to develop modest-size software components with a high degree of confidence. Flight control software is an example. This research focuses on algorithms and architectures that can leverage simple high-assurance components to ensure the integrity of large distributed real-time systems in spite of faults in complex software components.

Dependable and Secured Embedded Systems
L. Sha,* V. Adve, M. Spong
National Science Foundation, CNS 0209202

Faults and attacks during upgrades can be classified into three categories: application level control logic faults or attacks; code, data, thread, or process access faults or attacks; and resource depletion faults or attacks. To protect against them, our work will focus on integrated compiler static analysis and runtime checks to enforce the resource usage limits and to protect code, data, thread, and processes; and advanced safety controllers that can protect against coordinated control logic faults or attacks.

* Denotes principal investigator.
Together with real-time scheduling technology, they form a foundation upon which applications can be upgraded without shutting down normal operation. Furthermore, the system stability can be maintained in spite of insider attacks masquerading as upgrades.

Quality of Surveillance and Control in Network Centric Warfare
L. Sha,* J. C. Hou,* M. Caccamo, W.-P. Chen, P. R. Kumar, R. Iyer, R. Zheng
Office of Naval Research, Multidisciplinary Research Program of University Research Initiative

In this project, we aim to develop a sound scientific foundation and technologies to allocate computing, sensing, and communication resources in a way that will enhance the quality of surveillance and control for the Department of Defense's vision of network centric cooperative engagement. We are working with the DoD community to develop model problems that embody the fundamental scientific and engineering challenges faced by DoD systems, including network of multifunction radars, distributed sensor network, and advanced avionics systems. We are working to solve these model problems, demonstrate the solutions, and transition the technologies to major DoD programs through technology transition partners.

Scientific Computing

Curriculum Development in Computational Materials Science and Nanoscale Science and Engineering
E. deSturler,* D. Ceperley, D. Johnson, R. Martin, T. Martinez, U. Ravaioli
National Science Foundation, EEC-0088101

This project encompasses development of course materials (lecture notes, demos, practical assignments, numerical experiments) on Krylov subspace methods for linear systems and eigenvalue problems, nonlinear problems, multigrid methods, and fast multipole methods for summer school (summer 2001 and summer 2004) and general curriculum in the College of Engineering.

Fast Linear Solvers for Large Sequences of Slowly Varying Linear Systems
E. deSturler,* P. Geubelle, R. Dodds, M. Parks, C. Siefert, M. Kilmer (Tufts University)
U.S. Department of Energy, B341494; University of Illinois Computational Science and Engineering Program (fellowship M. Parks)

This project aims at the efficient solution of large sequences of slowly changing linear systems arising in time-dependent and nonlinear finite element analysis, such as crack propagation and analysis of fatigue. Further applications involve nonlinear optimization in diffuse optical tomography and other optimization problems. More generally, this project also involves the analysis and development of very robust and efficient (parallel) linear solvers and preconditioners for very large, sparse problems. Such research is of paramount importance to address the huge systems and sequences of systems that result from the problems addressed in the ASCI project.

Iterative Methods and Matrix Function Approximation in Material Science Simulations
E. deSturler,* D. Johnson
National Science Foundation, Focused Research Group Grant, DMR-0325939

This research is focused on study of a variety of techniques related to iterative methods for linear systems for problems arising in material science. An important topic is formed by cheap methods to compute selected coefficients of the inverses of matrices.

Multigrid and Krylov Subspace Solvers for Parallel Simulations with Dynamic Adaptive Mesh Refinements
E. deSturler,* Z. Chen, S. Wang
NASA Computational Technologies Project Cooperative Agreement, NCC5-615

Dynamic adaptive mesh refinement is a key technology in solving problems that require a large variation of spatial resolution in different parts of the computational domain and rapid changes in the required resolution. Such dynamically changing meshes pose several hard problems for iterative solvers, both of Krylov subspace and multigrid type. We consider improvements to both types of solvers for simulations using dynamic adaptive mesh refinements. For Krylov subspace methods, one of the main problems is to devise preconditioners that remain efficient after the redistribution of grid blocks and effectively deal with long range interaction. For multigrid methods, two problems are strong anisotropy and jumps in the coefficients. The traditional methods to deal with these situations in

* Denotes principal investigator.
multigrid solvers cannot easily be used on adaptively refined meshes.

**Preconditioners for Generalized Saddle Point Problems**
E. deSturler,* A. Beaudoin, Robert Dodds, C. Siefert, L. Zhu, M. Parks  
*U.S. Department of Energy, B341494; University of Illinois Computational Science and Engineering Program (Fellowship M. Parks)

Saddle point problems and generalized saddle point problems pervade computational science and engineering. Among other areas, these problems arise in constrained optimization and in systems of partial differential equations with conservation laws. The linear systems arising from these problems lead to very poor converge for iterative solvers. As a consequence, research in effective preconditioners is of paramount importance. An interesting application of such preconditioners is in the parallel domain decomposition method FETI, which turns variational problems in solid and fluid mechanics into constrained optimization problems. The FETI method is one of the most popular parallel domain decomposition methods with very good numerical and parallel scalability.

**Center for Simulation of Advanced Rockets**  
M. T. Heath,* E. deSturler, X. Jiao, W. Cochran, C. Siefert  
*U.S. Department of Energy, B523819

The center carries out a wide variety of research in combustion, fluid dynamics, structural analysis, and computer science. Research in numerical computation focuses on solvers for large linear systems and on data transfer between disparate meshes. Work on linear solvers includes development of direct and iterative methods for symmetric and nonsymmetric linear systems and eigenvalue problems. Specific emphases include grid-based solvers, subspace reuse in Krylov iterative solvers, and preconditioners for indefinite systems. Also under development are algorithms for numerically accurate and physically conservative data transfer at component interfaces based on a common refinement of mismatched surface meshes.

**Adaptive Gravitational Representation for Fast Trajectory Propagation Near Small Bodies**  
A. N. Hirani,* A. Colombi, B. F. Villac  
*University of Illinois Department of Computer Science

In this project approximation schemes for gravitational force and potential near arbitrarily shaped small bodies such as asteroids and comets are being developed. One approximation that has been developed is polynomial interpolation with an adaptive spatial data structure near the asteroid, and spherical harmonics far from it. These data structures allow us to drive the approximation errors of the model to within user defined thresholds, while significantly reducing the run time of trajectory integrations about small bodies (about 100 times faster than what is used in current space mission design). Other approximations that are being developed include those that have smoothness and which are harmonic.

**Algebraic Topology and Exterior Calculus in Numerical Analysis**  
A. N. Hirani,* J. H. Chaudhry, A. Reichert, and K. B. Nakshatrala  
National Science Foundation, DMS-0645604

The main aims of this project are the numerical analysis of discrete exterior calculus and the development of algorithms and software for such discretizations. This is a class of numerical methods for solving partial differential equations and these methods attempt to preserve the geometric and algebraic structures of the physics being modeled. A main ingredient of this project is the use of algebraic topology and differential geometry to create and analyze such discretizations. This is being done in the context of physical applications.

**Well-Centered Meshing**  
A. N. Hirani,* D. Guoy, E. Ramos, E. Vanderzee, V. Zharnitsky  
University of Illinois, Computational Science and Engineering Program (fellowship E. Vanderzee)

A well-centered simplex is one that contains its circumcenter in its interior. Well-centered meshes have applications in scientific computing and other fields. This project involves development of theory and tools to create well-centered meshes starting from existing meshes using an optimization approach. In addition, triangulation of simple shapes and necessary and sufficient conditions for well-centeredness are also being developed.

**Multidimensional Computation of Electronic Properties of Semiconductor Microstructures of High-Performance Heterojunction Devices**  
National Science Foundation, EET-87-19100

The goal of this research is to design numerical programs for the simulation of multidimensional quantum wells. The numerical model consists of Poisson's equation coupled with Schrodinger's equation. Poisson's equation is solved

* Denotes principal investigator.
as an elliptic boundary value problem, while Schrodinger's equation defines an eigenvalue problem. The model is nonlinear and requires repeated solution of such eigenvalue problems. Researchers compare the performance of Lanczos procedures with subspace iteration and examine the use of inverse iteration. Consistency of the two problems is obtained in an outer iteration that will be accelerated with a conjugate gradient procedure.

**National Computational Infrastructure for Lattice Gauge Theory**
D. A. Reed,* C. Mendes, Y. Zhang  
*U.S. Department of Energy, DEFC02-01ER41205*

This initiative is developing the software and hardware infrastructure necessary to support large-scale numerical simulations of quantum chromodynamics (QCD). As part of this effort, this group is assessing the performance of software APIs, measuring application performance, and evaluating new software approaches for community QCD codes.

**Performance Evaluation Research Center**
D. A. Reed,* C. Mendes, Y. Zhang  
*U.S. Department of Energy, DAR 0517 ANTIC*

This research group is one of the participants developing performance analysis and optimization techniques for high-end computer systems and applications. By developing a software infrastructure for monitoring and collecting performance data, this research group is contributing to the development of benchmarks, models, analytical techniques, and tools for the evaluation, optimization, and utilization of memory hierarchy in high performance computer systems.

**NIH Resource for Macromolecular Modeling and Bioinformatics**
K. Schulten* (Physics), G. Budescu (Beckman Instit.), L. V. Kale, R. Skeel, J. DeSouza, D. Hardy, S. Kumar, M. Potnuru, W. Wang, G. Zhang, and others  
*National Institutes of Health, P41RR05969*

This project involves the development of computational technology for modeling of large macromolecular systems in realistic environments and its deployment in the biomedical community. The computer science part of the effort is directed toward the parallel molecular simulation program NAMD and the web-enabled collaborative environment BioCore. Research topics include parallelization techniques for scaling to thousands of processors and fast methods for electrostatic calculations including polarization effects.

**Numerical Methods for Molecular Dynamics**
R. Skeel,* E. Cyr  
*National Science Foundation, DMS-0204442*

The goal of this research is the analysis and development of propagators and especially integrators for molecular simulations. The emphasis is on the study of integrators for the kinetics of biomolecules with application to methods for conformational dynamics.

## Journal Articles

### Algorithms and Theory

Alber, D. M. and Olson, L. N.  
**Parallel coarse-grid selection.** *Numerical Linear Algebra with Applications, 14*:8, 611-643 (Oct. 2007).

Armstrong, D. E. and Jacobson, S. H.  

Chekuri, C. and Khanna, S.  
**Edge disjoint paths revisited.** *Association for Computing Machinery Transactions on Algorithms, 3*:4, 1-31 (Nov. 2007).

Chekuri, C. and Pal, M.  

Chekuri, C., Chakrabarti, A., Gupta, A., and Kumar, A.  
**Approximation algorithms for the unsplittable flow problem.** *Algorithmica, 47*:1, 53-78 (2007).

Chekuri, C., Chuzhoy, J., Lewin-Eytan, L., Naor, J., and Orda, A.  

Chekuri, C., Khanna, S., and Shepherd, F. B.  
**A note on multicommodity flows and treewidth.** *Algorithmica, 49*:3, 171-257 (Nov. 2007).

Chekuri, C., Mydlarz, M., and Shepherd, F. B.  

* Denotes principal investigator.
Chekuri, C., Oriolo, G., Scutella, M., and Shepherd, F. B. **Hardness of robust network design.** *Networks, 50:1*, 50-54 (Sep. 2007).

Cheong, O., Efrat, A., and Har-Peled, S. **Finding a guard that sees most and a shop that sells most.** *Discrete and Computational Geometry, 37:4*, 545-563 (May 2007).


**Artificial Intelligence: Machine Learning, Vision, and Robotics**


White, R., Crane, K., and Forsyth, D. A. Capturing and animating occluded cloth [art. no. 34]. *Association for Computing Machinery Transactions on Graphics, 26:3, 34* (Jul. 2007) (http://dx.doi.org/10.1145/1276377.1276420).


---

**Computer Architecture and Compilers**


---

**Databases and Information Systems**


Wang, J. Y., Han, J. W., and Li, C. Frequent closed sequence mining without candidate maintenance. *IEEE Transactions on Knowledge and Data Engineering, 19*: 8, 1042-1056 (Aug. 2007) (http://dx.doi.org/10.1109/TKDE.2007.1043).


### Distributed Systems


Educational Research


Graphics and Visualization


Human–Computer Interfaces


Networking


**Operating Systems and Security**


**Programming Languages, Formal Systems, and Software Engineering**


**Real-Time and Embedded Systems**


Scientific Computing


Books

Programming Languages, Formal Systems, and Software Engineering


Book Chapters

Artificial Intelligence: Machine Learning, Vision, and Robotics


Databases and Information Systems


Networking


Operating Systems and Security


Parallel Processing


Programming Languages, Formal Systems, and Software Engineering


Real-Time and Embedded Systems


Scientific Computing


Papers Presented at Conferences and Symposia

Algorithms and Theory


**Artificial Intelligence: Machine Learning, Vision, and Robotics**


Dawsey, W., Minsker, B., and Amir, E. **Real-time assessment of drinking water systems using a Bayesian network.** World Environmental and Water Resources Congress (Tampa, FL, May 2007).


Hill, D. J., Minsker, B., and Amir, E. **Real-time Bayesian anomaly detection for environmental sensor data.** 32nd Congress of the International Association of Hydraulic Engineering and Research (Venice, Italy, Jul. 2007).


Lim, S. H., Wang, L. L., and DeJong, G. Explanation-based feature construction. 20th International Joint Conference on Artificial Intelligence (Hyderabad, India, Jan. 2007).


Ramachandran, D. and Amir, E. Bayesian inverse reinforcement learning. 20th International Joint Conference on Artificial Intelligence (Hyderabad, India, Jan. 2007).

Richards, M. and Amir, E. Opponent modeling in Scrabble. 20th International Joint Conference on Artificial Intelligence (Hyderabad, India, Jan. 2007).

Shahaf, D. and Amir, E. Logical circuit filtering. 20th International Joint Conference on Artificial Intelligence (Hyderabad, India, Jan. 2007).


White, R., Crane, K., and Forsyth, D. Data-driven cloth animation. 34th International Conference and Exhibition on Computing Graphics and Interactive Techniques (San Diego, CA, Aug. 2007).

Computer Architecture and Compilers


Baugh, L. and Zilles, C. An analysis of I/O and syscalls in critical sections and their implications for transactional memory. 1st Association for Computing Machinery SIGPLAN Workshop on Languages, Compilers, and Hardware Support for Transactional Computing (Portland, OR, Aug. 2007).


**Databases and Information Systems**


Mei, Q., Shen, X., and Zhai, C. **Automatic labeling of multinomial topic models.** 13th Association for Computing Machinery SIGKDD International Conference on Knowledge Discovery and Data Mining (San Jose, CA, Aug. 2007). Proceedings of the 13th Association for Computing Machinery SIGKDD International Conference on Knowledge Discovery and Data Mining 490-499 (2007) (http://dx.doi.org/10.1145/1281192.1281246).


Tan, L., Yuan, D., and Zhou, Y. **iComment: Bugs or bad comments?** 21st Association for Computing Machinery Symposium on Operating Systems Principles (Stevenson, WA, Oct. 2007).


### Distributed Systems


### Educational Research

Winslett, M. *Managing scientific data: New challenges for database research.* 19th International Conference on Scientific and Statistical Database Management (Banff, AB, Jul. 2007).

### Graphics and Visualization


### Human–Computer Interfaces

Bergstrom, T. and Karahalios, K. **Conversa- tion clock: Visualizing audio patterns in co-located groups.** 40th Hawaii International Conference on Systems Science (Waikoloa, Big Island, HI, Jan. 2007).


Bergstrom, T. and Karahalios, K. **A user study for tabletop conversation interaction.** Computer/Human Interaction 2007 Conference (San Jose, CA, Apr. 2007).


Hailpern, J., Hinterbichler, E., Leppert, C., Cook, D. J., and Bailey, B. P. **TEAM STORM: Demonstrating an interaction model for working with multiple ideas during creative group work.** Association for Computing Machinery Conference on Creativity and Cognition (Washington, DC, Jun. 2007).


**Networking**


Crepaldi, R., Harris III, A. F., Kooper, R., Kravets, R., Maselli, G., Petrioli, C., and Zorzi, M. **Managing heterogeneous sensors and actuators in ubiquitous computing environments.** Association for Computing Machinery International MobiCom Workshop on Sensor Actor Networks (Montreal, QC, Sep. 2007).


Nahrstedt, K. *CA-AQM: Channel-aware active queue management for wireless networks*. IEEE International Conference on Communications (Glasgow, UK, Jul. 2007).


**Operating Systems and Security**


David, F. M., Carlyle, J. C., and Campbell, R. H. **Context switch overheads for Linux on ARM platforms.** Workshop on Experimental Computer Science (San Diego, CA, Jun. 2007).


Kapadia, A., Naldurg, P. and Campbell, R. H. **Distributed enforcement of unlikability policies: Looking beyond the Chinese wall.** IEEE Workshop on Policies for Distributed Systems and Networks (Bologna, Italy, Jul. 2007).


**Parallel Processing**


Programming Languages, Formal Systems, and Software Engineering

AlTurki, M. and Meseguer, J. Real-time rewriting semantics of Orc. 9th International Association for Computing Machinery SIGPLAN Conference on Principles and Practice of Declarative Programming (Wroclaw, Poland, Jul. 2007).


Chang, P. and Agha, G. Supporting reconfigurable object distribution for customized web applications. Association for Computing Machinery SIGAPP Symposium on Applied Computing (Seoul, South Korea, Mar. 2007).

Chang, P. and Agha, G. Towards context-aware Web applications. 7th International Federation for Information Processing International Conference on Distributed Applications and Interoperable Systems (Paphos, Cyprus, Jun. 2007).


Serbanuta, T.-F., Rosu, G., and Meseguer, J. **A rewriting logic approach to operational semantics.** 4th Workshop on Structural Operational Semantics (Wroclaw, Poland, Jul. 2007).


**Real-Time and Embedded Systems**


David, F. M., Carlyle, J. C., and Campbell, R. H. **Exploring recovery from operating system lockups.** USENIX Annual Technical Conference (Santa Clara, CA, Jun. 2007).

Kazman, R. and Agha, G. **Software technology track introduction.** 40th Hawaii International Conference on System Sciences (Waikoloa, Big Island, HI, Jan. 2007).


Scientific Computing


### Theses

#### Algorithms and Theory


#### Artificial Intelligence: Machine Learning, Vision, and Robotics


#### Computer Architecture and Compilers


Databases and Information Systems


Graphics and Visualization


Human–Computer Interfaces


**Networking**


**Operating Systems and Security**


**Parallel Processing**


Programming Languages, Formal Systems, and Software Engineering


Scientific Computing


Real-Time and Embedded Systems


Scientific Computing


Patents

Computer Architecture and Compilers


Awards and Honors

Vikram Adve
Best Paper Award, 15th Workshop on Parallel and Distributed Simulation, May 2001
C. W. Gear Outstanding Junior Faculty Award, University of Illinois Department of Computer Science, 2002
Associate Editor, ACM Transactions on Programming Languages and Systems, 2003-
Best Paper Award, Programming Language Design and Implementation, 2005

Sarita Adve
Research Initiation Award, National Science Foundation (NSF), 1994-1997
Faculty Early Career Development Program (CAREER) Award, NSF, 1995-1998
Associate Editor, ACM Transactions on Modeling and Computer Simulation, 1997-
Faculty Partnership Award, IBM Corporation, 1997-1999
Research Fellowship, Alfred P. Sloan Foundation, 1998-2000
Coauthor, one of forty-one papers chosen for an anthology of selected papers from the first twenty-five International Symposia on Computer Architecture (Sohi, editor), ACM Press, 1998
University Scholar, University of Illinois, 2004
Faculty Partnership Award, IBM, 2005-2006

**Gul A. Agha**
Young Investigator Award, U.S. Office of Naval Research, 1989
Incentive for Excellence Award, Digital Equipment Corporation Faculty Program, 1990
Fellow, University of Illinois Center for Advanced Study, 1992-1993
International Lecturer, Association for Computing Machinery (ACM), 1991-2000
Golden Core Member, Institute of Electrical and Electronics Engineers (IEEE) Computer Society, 1999
Meritorious Service Award, IEEE Computer Society, 1999
Editor-in-Chief, *ACM Computing Surveys*, 2000-
Fellow, IEEE Computer Society, 2002
Member, European Academy of Sciences, 2003
Arthur Schofstall Distinguished Lecturer, Renssleear Polytechnic Institute, 2003
ACM Recognition of Service Award, Editor in Chief of *ACM Computing Surveys*, 2007

**Eyal Amir**
Arthur L. Samuel Award for Best Thesis, Stanford University Computer Science Department, 2002

**Brian P. Bailey**
Research Contribution Award, University of Minnesota Department of Computer Science, 2001
Fellow, National Center for Supercomputing Applications, 2004
List of Teachers Ranked As Excellent by Their Students, University of Illinois, 2004, 2005, 2006
Engineering Council Award for Excellence in Student Advising, University of Illinois, 2005, 2007
Faculty Early Career Development Program (CAREER) Award, National Science Foundation, 2007

**Geneva G. Belford, Emeritus**
Distinguished Visitor, Institute of Electrical and Electronics Engineers Computer Society, 1982-1985
Halliburton Engineering Education Leadership Award, University of Illinois College of Engineering, 1986
Award for Outstanding Faculty Member, Dad's Association, University of Illinois, 1991
Graduate College Outstanding Mentor Award, 2005

**Stephen Bond**
Fellow, La Jolla Interfaces in Science, 2000
Accenture Engineering Council Award for Excellence in Advising, University of Illinois College of Engineering, 2005
Rose Award for Teaching Excellence, University of Illinois College of Engineering, 2006

**Marco Caccamo**
Faculty Early Career Development Program (CAREER) Award, National Science Foundation, 2003

**Roy H. Campbell**
Senior Visiting Research Fellowship at University of Newcastle upon Tyne, Science and Engineering Research Council of Great Britain, 1981-1983
Information Technology Committee, Illinois Terrorism Task Force, 2002-2003
Sohaib and Sara Abbasi Professor in Computer Science, University of Illinois, 2004-
Fellow, Institute of Electrical and Electronics Engineers (IEEE), 2005

**Kevin C.-C. Chang**
Best Paper selection, 26th International Conference on Very Large Data Bases, 2000
Faculty Early Career Development Program (CAREER) Award, National Science Foundation, 2002
Faculty Fellows Award, National Center for Supercomputing Applications, 2003
IBM Faculty Award, 2004
IBM Faculty Award, 2005

**Chandra Chekuri**
Arthur Samuel Thesis Award, Stanford University Computer Science Department, 1999
Presidents Award, Lucent Bell Labs, 2000
Associate Editor, *Journal of Computer System Sciences (JCSS)*, 2005-
Area Editor, *Encyclopedia of Algorithms*, Springer, 2006-

**Gerald DeJong**
Faculty Assistant Grant, Exxon Mobil Corporation, 1982
Arnold O. Beckman Research Award, University of Illinois Research Board, 1984
Faculty Recognition Grant, Alcoa Foundation, 1989
Fellow, American Association for Artificial Intelligence, 1992
International Scientist of the Year, International Biographical Centre, 2001
Accenture Award for Excellence in Advising, University of Illinois College of Engineering, 2007

**Eric deSturler**
Leslie Fox Prize, Institute of Mathematics and Its Applications, 1997
Invited Plenary Lecture, 15th Householder Symposium in Numerical Linear Algebra, 2002
Program Director, SIAM Activity Group on Supercomputing, 2003-2006
Co-chair, SIAM Conference on Computational Science and Engineering, 2005
(Invited) Plenary lecture, 16th Householder Symposium in Numerical Linear Algebra, 2005

**AnHai Doan**
William Chan Memorial Dissertation Award, University of Washington, 2003
Distinguished Doctoral Dissertation Award, Association for Computing Machinery, 2003
Faculty Early Career Development Program (CAREER) Award, National Science Foundation, 2004
Outstanding Alumnus Award, University of Wisconsin, 2004

**Jeffrey G. Erickson**
Mathematical Sciences Postdoctoral Research Fellowship, National Science Foundation (NSF), 1996-1998
Research Fellowship, Alfred P. Sloan Foundation, 1999-2001
Everitt Award for Teaching Excellence, University of Illinois College of Engineering, 2001
C. W. Gear Outstanding Junior Faculty Award, University of Illinois Department of Computer Science, 2001

Faculty Early Career Development Program (CAREER) Award, NSF, 2001-2006
Willett Faculty Scholar, University of Illinois College of Engineering, 2002-2005
Campus Award For Excellence in Undergraduate Teaching, University of Illinois at Urbana-Champaign, 2007

**David Forsyth**
National Science Foundation Research Initiation Award, 1992
National Science Foundation Young Investigator Award, 1992
Marr Prize, Best Paper, 1993 International Conference on Computer Vision, 1993
Fellowship, Okawa Foundation, 2003
Technical Achievement Award, Institute of Electrical and Electronics Engineers, 2006

**Maria J. Garzaran**
Best PhD Thesis Award, Universidad de Zaragoza in Spain, 2003

**Elsa Gunter**
Best Paper Award, Fourth International Conference on Requirements Engineering, 2000
European Association of Software Science and Technology Award for the Best Software Science Paper, European Joint Conferences on Theory and Practice of Software, 2001

**Carl A. Gunter**
Best Paper Award, Fourth International Conference on Requirements Engineering, 2000
Editor, *Journal of Computer Security*, November 2004

**Indranil Gupta**
Faculty Early Career Development Program (CAREER) Award, National Science Foundation, 2005
Incomplete List of Teachers Ranked as Excellent, University of Illinois Instructor and Course Evaluation System, Student Survey, 2003-2007

**Jiawei Han**
Associate Editor, *Data Mining and Knowledge Discover (DAMI)*, 1997-
Associate Editor, *Journal of Intelligent Information Systems (JIIIS)*, 1997-
Endowed University Professor, Simon Fraser University, 1999-2001
Outstanding Contribution Award, Institute of Electrical and Electronics Engineers (IEEE) International Conference on Data Mining, 2002
Board of Directors, Association for Computing Machinery (ACM) SIGKDD, 2001-
IBM Faculty Award, 2002-2004
Senior Member, IEEE, 2003-
Fellow, Association for Computing Machinery, 2004-
Technical Achievement Award, IEEE Computer Society, 2004
ACM SIGKDD Innovations Award, 2004
Editor-in-Chief, ACM Transactions on Knowledge Discovery from Data (ACM TKDD), 2005-
Best Student Paper Runner-Up Award, 2005 International ACM-SIGKDD Conference on Knowledge Discovery and Data Mining (KDD’05), 2005
Faculty Fellow, National Center for Supercomputing Applications, 2006-
Incomplete List of Teachers Ranked as Excellent, University of Illinois, 2002-2007

Mehdi T. Harandi
Best Paper Award, 3rd International Workshop on Computer-Aided Software Engineering, 1989
Senior Member, Institute of Electrical and Electronics Engineers, 1990
Best Paper Award, 5th Conference on Knowledge-based Software Assistant, 1990
Editor, Journal of Automated Software Engineering, 1992-
Recognition of Service Award, Association for Computing Machinery, 1997

John C. Hart
Research Initiation Award, National Science Foundation, 1993-1996
Associate Editor, ACM Transactions on Graphics, 2000-2002
Editor-in-Chief, ACM Transactions on Graphics, 2003-

Michael T. Heath
Eugene P. Wigner Postdoctoral Fellow, Oak Ridge National Laboratory, 1978-1980
Editor, SIAM News, 1988-
Editor, SIAM Journal on Scientific Computing, 1990-1995
Editor, International Journal of High-Performance Computing Applications, 1993-
Editor, SIAM Review, 1994-2002

Everitt Award for Teaching Excellence, University of Illinois College of Engineering, 1998
Fellow, Association for Computing Machinery (ACM), 2000
Recognition of Service Award, ACM, 2001
Member, European Academy of Sciences, 2002
Fulton Watson Copp Chair in Computer Science, 2003-
Engineering Council Award for Excellence in Advising, University of Illinois College of Engineering, 2003
Editor, SIAM Fundamentals of Algorithms, 2003-
Member, European Academy of Sciences, 2003

Anil N. Hirani
Faculty Early Career Development Program (CAREER) Award, Algebraic Topology and Exterior Calculus in Numerical Analysis, National Science Foundation, 2007

Jennifer C. J. Hou
Women in Science Initiative Award, University of Wisconsin, 1993-1995
Wisconsin/Hilldale Undergraduate/Faculty Research Fellowship, University of Wisconsin College of Engineering, 1995
Faculty Early Career Development Program (CAREER) Award, National Science Foundation, 1996-2001
Senior Member, Institution of Electrical and Electronics Engineers (IEEE), 1999-
Technical Program Co-Chair, IEEE Sixth Real-Time Technology and Application Symposium, 2000
Lumley Research Award, The Ohio State University, 2001
General Co-Chair, IEEE Seventh Real-Time Technology and Application Symposium, 2001
Cisco University Research Program Award, 2002-2003
Editor, Elsevier Computer Networks Journal, 2002-
Editor, IEEE Transactions on Wireless Communications, 2002-
Editor, Association for Computing Machinery, ACM Baltzer Wireless Networks WINET, 2002-
Editor, IEEE Transaction on Parallel and Distributed Systems, 2003-
Editor, IEEE Wireless Communications Magazine, 2003-
Editor, Wireless Ad Hoc and Sensor Networks: An International Journal, 2004-
Recognition of Service Award, Association for Computing Machinery, 2004
Technical Program Co-Chair, ACM/IEEE Third International Symposium and Information Processing in Sensor Networks, 2004
Editor, Foundations and Trends in Networking, 2005
Technical Program Co-Chair, First Wireless Internet Conference, 2005

**Sheldon H. Jacobson**
Research Initiation Award, National Science Foundation, 1994
Best Paper Award, Industrial Simulation Track, European Simulation Multiconference, 1997
First Place Application Award, Institute of Industrial Engineers (IIE) Operations Research Division, 1998
Willett Faculty Scholar Award, University of Illinois College of Engineering, 2002-2005
Guggenheim Fellowship, John Simon Guggenheim Memorial Foundation, 2003
Best Paper Award, IIE Transactions, 2003
Meritiorious Service Award, Operations Research, 2003
Award of Excellence, 13th Annual Communicator’s Award Competition, 2006
Incomplete List of Teachers Ranked as Excellent by their Students, University of Illinois, 2006

**Ralph E. Johnson**
Programming Language Achievement Award, Programming Language Design and Implementation, 2005

**Laxmikant V. Kale**
NCR Award of Excellence, 1989
C. W. Gear Outstanding Junior Faculty Award, University of Illinois Department of Computer Science, 1990
Gordon Bell Award, Supercomputing Conference 2002, 2002
Best Poster Award, Supercomputing Conference 2002, 2002
Editorial Board, SIAM Book Series: Computational Science and Engineering, 2004-

**Samuel N. Kamin**
Editorial Board, ACM Transactions on Programming Languages and Systems, 1993-1999
Stanley H. Pierce Faculty Award, University of Illinois College of Engineering, 2001
Charter Member, International Federation for Information Processing Working Group 2.11 (Program Generation), 2004
Campus Award for Excellence in Advising Undergraduate Students, University of Illinois, 2005

**Karrie Karahalios**
IBM Fellowship, 2000, 2001

**Robin Kravets**
Faculty Early Career Development Program (CAREER) Award, National Science Foundation (NSF), 2004

**Steven M. LaValle**
Finalist for Best Journal Paper of the Year Award, IEEE Transactions on Robotics and Automation, 1998
Alternative Winner: Presidential Early Career Award (PECASE), Directorate for Computer and Information Science, National Science Foundation, 2000
C. W. Gear Outstanding Junior Faculty Award, University of Illinois Department of Computer Science, 2003-2004
List of Teachers Ranked Excellent by Their Students, Daily Illini, 2006

**Darko Marinov**
Incomplete List of Teachers Ranked as Excellent, University of Illinois, 2006

**Jose Meseguer**
Member, International Federation for Information Processing (IFIP) Working Group 14.3 (Foundations of Systems Specification)
Member, IFIP Working Group WG1.6 (Term Rewriting)
Member, Gesellschaft für Informatik Working Group 0.1.7 (Specification and Semantics)
Postdoctoral Research Fellow, Spanish Ministry of Universities and Science, 1977-1979
Research Fellow, U.S.-Spain Joint Committee for Scientific Cooperation, 1979-1981
Associate Editor, Theory and Practice of Object Systems, 1994-2000
Editorial Board, University of Madrid Mathematics Journal, 1995-

**Klara Nahrstedt**
Weierstrass Prize, Weierstrass Institute of Mathematics, Berlin, 1985
Faculty Early Career Development Program (CAREER) Award, National Science Foundation (NSF), 1996
NASA Space Act Award, NSF, 1996
Editorial Board, IEEE Multimedia Newsletter, 1996
Xerox Award for Junior Faculty Research, University of Illinois College of Engineering, 1998
C. W. Gear Faculty Award, University of Illinois Department of Computer Science, 1999
Best Tutorial Paper Award, Institute of Electrical and Electronics Engineers (IEEE), for "An Overview of Quality of Service Routing for Next-Generation High-Speed Networks: Problems and Solutions," 1999
Campus Award for Innovation in Undergraduate Instruction Using Educational Technologies, University of Illinois, 2000
Associate Editor, *ACM Computer Communications Reviews*, 2000-
Editor-in-Chief, *ACM Multimedia Systems Journal*, 2001-
Ralph M. and Catherine V. Fisher Professorship, University of Illinois College of Engineering, 2002-
Service Award, Association for Computing Machinery (ACM), 2006, 2007
Chair, ACM SIG Multimedia, 2007

David A. Padua
Distinguished Visitor, Institute of Electrical and Electronics Engineers (IEEE) Computer Society, 1990-1993
Xerox Award for Faculty Research, University of Illinois College of Engineering, 1992
Editorial Board, *Journal of Parallel and Distributed Computing*, 1993-
Best Paper Award, International Conference on Supercomputing, 1995
Golden Core Member, IEEE Computer Society, 1997
Editorial Board, *International Journal of Parallel Programming*, 1998-
Editorial Board, *Electronic Journal of the Latin-American Center for Information Studies*, 1998-
Fellow, IEEE, 2000

Lenny Pitt
Research Initiation Award, National Science Foundation, 1988
C. W. Gear Outstanding Junior Faculty Award, University of Illinois Department of Computer Science, 1992

Everitt Award for Teaching Excellence, University of Illinois College of Engineering, 1999
University Distinguished Teacher/Scholar, University of Illinois, 2004

Jean A. Ponce
Xerox Award for Faculty Research, University of Illinois College of Engineering, 1993, 1998
Area Editor, *Computer Vision and Image Understanding*, 1994-
Beckman Associate, University of Illinois Center for Advanced Study, 1994-1995
Critical Research Initiative Planning Award, University of Illinois, 1996
Associate Editor, *Institute of Electrical and Electronics Engineers (IEEE) Transactions on Robotics Research*, 1998-2002
Editor-in-Chief, *International Journal of Computer Vision*, 2002-
Fellow, IEEE, 2003
Critical Research Initiative Planning Award, University of Illinois, 2005

Grigore Rosu
Irina Gorun-Bercovici Memorial Prize, Society of Mathematicians from Romania, 1997
Best Software Science Paper (ETAPS 2002), European Association for the Study of Sciences and Technology, 2002
Faculty Early Career Development Program (CAREER) Award, National Science Foundation, 2005
C. W. Gear Outstanding Junior Faculty Award, University of Illinois Department of Computer Science, 2004

Dan Roth
Derek Bok Excellence in Teaching Award, Harvard University, 1993
Best Paper, 16th International Joint Conference on Artificial Intelligence, 1999
IBM Faculty Equipment Award, 1999
Faculty Early Career Development Program (CAREER) Award, National Science Foundation, 2000
C. W. Gear Outstanding Junior Faculty Award, University of Illinois Department of Computer Science, 2000
Innovative Applications of AI Award, American Association of Artificial Intelligence, 2001
Xerox Award for Junior Faculty Research, University of Illinois College of Engineering, 2001
Honorable Mention: Campus Award for Excellence in Guiding Undergraduate Research, University of Illinois, 2002
Willett Faculty Scholar, University of Illinois Department of Computer Science, 2002-2005
Editorial Board, *Computational Intelligence*, 2003-
Action Editor, *Machine Learning Journal*, 2004-
Xerox Award for Faculty Research, Xerox Foundation, 2005

**Lui Sha**
Fellow, Institute of Electrical and Electronics Engineers
Associate Editor, *International Journal of Real-Time Systems*, 1992-
GE Scholar, University of Illinois Academy for Excellence in Engineering Education, 1999
Co-Chair, IEEE Real-Time and Embedded Technology and Applications Symposium, 2001
Associate Editor, *IEEE Transactions on Parallel and Distributed Systems*, 2001-2002

**Marc Snir**
Outstanding Innovation Award for work on Hierarchical Memory Model, IBM Corporation, 1989
Editorial Board, *Parallel Processing Letters*, 1992-
Member, IBM Academy of Technology, 1993
Outstanding Innovation Award for work on Parallel System Architecture and Software Structure, IBM Corporation, 1994
Fellow, Institute of Electrical and Electronics Engineers, 1996
Fellow, Association for Computing Machinery, 1999
Michael Faiman and Saburo Muroga Professor of Computer Science, University of Illinois, 2001-
Editorial Board, *Computing Surveys*, 2002-
Board of Directors, Computing Research Association, 2003-2006
Fellow, American Association for the Advancement of Science, 2006

**Josep Torrellas**
Research Initiation Award, National Science Foundation (NSF), 1993-1996
Young Investigator Award, NSF, 1994-1999
C. W. Gear Outstanding Junior Faculty Award, University of Illinois Department of Computer Science, 1997
IBM Partnership Award, 1997-1999

Xerox Award for Outstanding Faculty Research, University of Illinois College of Engineering, 1997, 2000
Vice-Chairman, Institute of Electrical and Electronics Engineers (IEEE) Technical Committee on Computer Architecture (TCCA), 2000
Editorial Board, *IEEE Computer Architecture Letters*, 2001-
Willett Faculty Scholar, University of Illinois Department of Computer Science, 2002-2005
Associate Editor, *ACM Transactions on Architecture and Code Optimization*, 2003-
Fellow, IEEE, 2004
Best Paper Award, 39th International Symposium on Microarchitecture, 2006

**Mehesh Viswanathan**
Faculty Early Career Development Program (CAREER) Award, National Science Foundation, 2005

**Marianne S. Winslett**
Presidential Young Investigator Award, National Science Foundation, 1989
Xerox Award for Junior Faculty Research, University of Illinois College of Engineering, 1990
Associate Editor (Funding News and On-Line Services), *ACM SIGMOD Record*, 1991-
University Scholar, University of Illinois, 1992
Associate Editor, *Journal of Intelligent Information Systems*, 1992-
Associate Editor, *ACM Transactions on Database Systems*, 1994-
Associate Editor, *IEEE Transactions on Knowledge and Data Management*, 1994-1998
Xerox Award for Senior Faculty Research, University of Illinois College of Engineering, 1997
Stanley H. Pierce Award, University of Illinois College of Engineering, 1999
Board Accredited Editor, *ACM SIGMOD Digital Review*, 1999-
Associate Editor (Distinguished Profiles in Databases), *ACM SIGMOD Record*, 2001-
Faculty Fellows Program, National Center for Supercomputing Applications, 2004-2005
Best Paper Award, Association for Computing Machinery, (ACM) Storage Security and Survivability Workshop, 2006
ACM Fellow, 2007

**Yizhou Yu**
Fellowship, Microsoft, 1998
Faculty Early Career Development Program (CAREER) Award, National Science Foundation, 2002
Best Paper Award, ACM SIGGRAPH/Eurographics Symposium on Computer Animation, 2005

Chengxiang Zhai
Presidential Early Career Award for Scientists and Engineers (PECASE), National Science Foundation, 2004

Yuanyuan Zhou
DuPont Fellowship, University of Virginia, 1992
IBM Shared University Research (SUR) Award, 2003
Faculty Early Career Development Program (CAREER) Award, National Science Foundation, 2004
IBM Faculty Award, 2004
W. Anita Borg Early Career Award, Computing Research Association, 2005
Alfred Sloan Fellowship, 2007

Craig Zilles
Fellowship, National Science Foundation, 1996
Fellowship, Intel Graduate, 2000
Wisconsin Distinguished Graduate Research Fellowship, 2001-2002
Faculty Early Career Development Program (CAREER) Award, National Science Foundation, 2004
Accenture Outstanding Undergraduate Advisor award, 2006
Accenture Outstanding Undergraduate Advisor award, 2007
Rose Undergraduate Teaching Award, 2007