



Institute for Scientific Computing Research



Seminar Series Abstracts

(in reverse chronological order)

ISCR Seminar Series

The ISCR hosts on-site seminars by visiting experts at the request of Laboratory researchers. Approximately one seminar is held per week throughout the year.

Name, Affiliation, Seminar Date	Page
Hans de Sterck, University of Colorado, Boulder, August 29, 2003.....	36
Michael McCracken, UC San Diego, August 14, 2003.....	37
Sam Madden, UC Berkeley, August 11, 2003.....	38
Karen Karavanic, Portland State University, August 8, 2003.....	39
Said Elghobashi, UC Irvine, August 7, 2003.....	40
Christoph Reisinger, University of Heidelberg, July 25, 2003.....	41
Chen Ding, University of Rochester, July 18, 2003.....	42
Randolph Bank, UC San Diego, July 17, 2003.....	43
Radek Grzeszczuk, Intel Corporation, July 11, 2003.....	44
Charles Hansen, University of Utah, June 19, 2003.....	45
Andrew de Laix, Wolfram Research, Inc., June 11, 2003.....	46
Alok Choudhary, Northwestern University, May 15, 2003.....	47
Christian Früh, UC Berkeley, May 13, 2003.....	48
Avideh Zakhori, UC Berkeley, May 13, 2003.....	48
John Trangenstein, Duke University, April 30, 2003.....	49
Bruce Rankin, Canadian Information Processing Society, April 14, 2003.....	50
Scott Banachowski, UC Santa Cruz, April 7, 2003.....	51
Katharine Gurski, National Institute of Standards and Technology, April 4, 2003.....	52
Patrick Langley, Institute for the Study of Learning and Expertise, April 4, 2003.....	53
Lenny Oliker, Lawrence Berkeley National Laboratory, March 26, 2003.....	54
Joel Saltz, Ohio State University, March 17, 2003.....	55
Frank Behrendt, Berlin University of Technology, March 5, 2003.....	56

Phil Mucci, University of Tennessee at Knoxville, February 28, 2003.	57
Michael Walker, General Atomics, February 19, 2003.	58
Pedro Diniz, University of Southern California, February 18, 2003.	59
Ivan Ovcharenko, Lawrence Berkeley National Laboratory, February 18, 2003.	60
Eamonn Keogh, UC Riverside, February 14, 2003.	61
Douglas Arnold, University of Minnesota, February 7, 2003.	62
Douglas Arnold, University of Minnesota, February 6, 2003.	63
Stan Rifkin, Master Systems, February 3, 2003.	64
Alberto Paoluzzi, Third University of Rome, Italy, January 13, 2003.	65
Ion Muslea, University of Southern California, January 10, 2003.	66
Alex Pothén, Old Dominion University, January 6, 2003.	67
Reagan Moore, San Diego Supercomputer Center, December 13, 2002.....	68
Rudolf Eigenmann, Purdue University, December 11, 2002.....	69
John Perry, Foothill Deanza College, December 6, 2002.	70
Tushar Mohan, Lawrence Berkeley National Laboratory, December 3, 2002.	71
Gregory Andrews, University of Arizona, November 8, 2002.	72
Dana Knoll, Los Alamos National Laboratory, October 28, 2002.....	73
Dana Knoll, Los Alamos National Laboratory, October 29, 2002.....	74
David Lopez, UC San Diego, October 18, 2002.....	75
Sutanu Sarkar, UC San Diego, October 18, 2002.	75
Sergey Kirshner, UC Irvine, October 8, 2002.	76
Charles Hansen, University of Utah, October 3, 2002.	77

August 29, 2003

Numerical Conservation Properties of Least- Squares Finite Element Methods

Hans de Sterck

University of Colorado at Boulder

Email: desterck@newton.colorado.edu

Abstract

This talk presents results from our ongoing study of least-squares finite element methods (LSFEMs) for nonlinear hyperbolic conservation laws. The conservation law is reformulated by introducing the flux vector, or the associated flux potential, explicitly as additional dependent variables. The standard least-squares finite element procedure is then applied to the reformulated equations using $H(\text{div})$ -conforming finite element spaces and a Gauss–Newton nonlinear solution technique with grid continuation. Numerical results are presented for the one-dimensional Burgers equation on adaptively refined space–time domains, indicating that the $H(\text{div})$ -conforming finite element methods converge to the entropy-weak solution of the conservation law. The $H(\text{div})$ -conforming LSFEMs do not satisfy a discrete exact conservation property in the sense of Lax and Wendroff. However, weak conservation theorems that are analogous to the Lax–Wendroff theorem for conservative finite difference methods are proved for the $H(\text{div})$ -conforming LSFEMs. These results illustrate that discrete exact conservation in the sense of Lax and Wendroff is not a necessary condition for numerical conservation but can be replaced by minimization in a suitable continuous norm. This is joint research with Luke Olson, Tom Manteuffel, and Steve McCormick.

Speaker's web page: <http://amath.colorado.edu/faculty/desterck>

Institution web page: <http://www.colorado.edu/>

A Framework for Practical Applications Performance Modeling

Michael McCracken

University of California, San Diego

Email: mccrack@cs.ucsd.edu

Abstract

This talk describes technical advances that have enabled a framework for performance modeling to become practical for analyzing the performance of HPC applications; also, studies of some applications are exhibited to indicate the kinds of insights the framework can provide. In previous work, we presented a framework for performance modeling and prediction that is faster than cycle-accurate simulation and more informative than simple benchmarking and that was shown useful for performance investigations via proof-of-principle studies applied to short-running kernels. Here, we provide an update on the subsequent investigations we have carried out to advance the framework. The methods are now extensible to longer-running applications. The dominant cost in the framework has always been the time required to gather “application signatures” via instrumented application traces. We describe how we reduced tracing time while preserving the accuracy of resulting performance models. The framework is used to model the performance of three applications, POP (Parallel Ocean Program), NLOM (Navy Layered Ocean Model), and Cobalt60 run with exemplary data inputs at various processor counts on several HPC platforms (with Power3, Power4, Alpha, and Cray X1 processors). We show how the models can be used to develop sensitivity profiles of these applications to quantify how their performance would increase from improvements in the underlying architecture. This research was done in collaboration with Allan Snaveley.

Speaker's web page: <http://www.cs.ucsd.edu/~mmccrack/>

Research web page: <http://www.sdsc.edu/PMaC/>

Institution web page: <http://www.ucsd.edu/>

August 11, 2003

Query Processing in Sensor Networks

Sam Madden

University of California, Berkeley

Email: madden@cs.berkeley.edu

Abstract

Many of the emerging applications for sensor networks are focused on data collection and monitoring in remote environments. Unfortunately, existing tools for building such applications require users of these networks, who often are not trained computer scientists, to write low-level, embedded C code. Such deployments frequently become mired in the difficulties of coding power-management, routing, and storage features in these volatile distributed environments.

In this talk, I will discuss how many of these difficulties can be overcome by providing users with a simple declarative interface where short, SQL-like queries are pushed into the network. Such queries concisely express a user's data needs, freeing him or her from the details of implementation and execution. In addition to dramatically simplifying the task of sensor-network programming, this approach enables the system to transparently optimize in-network query execution to minimize overall power consumption in ways that even sophisticated programmers may miss.

I will summarize the query processing features of TinyDB, a query processor for sensor networks we have developed at Berkeley, focusing on a framework for executing and optimizing aggregation queries. I will discuss current deployments that are under way at Berkeley, along with new features that are being incorporated to accommodate these deployments. I will include a brief demonstration of the system.

Speaker's web page: <http://www.cs.berkeley.edu/~madden/berkeley.htm>

Institution web page: <http://www.berkeley.edu/>

August 8, 2003

Automated Comparative Profiling of Parallel Cluster Applications

Karen Karavanic

Portland State University

Email: karavan@cs.pdx.edu

Abstract

The PPerfDB research project is developing scalable methods for diagnosing the performance of large-scale parallel applications using data from more than one program execution. PPerfXchange, a tool component under development, allows geographically dispersed data stores to be included as sources of data for PPerfDB. PPerfXchange models each site's performance data as XML documents, based upon a global schema, allowing client applications to retrieve performance data from remote sites using XQuery.

To gain insights for our continued work developing PPerfDB, we recently completed a comparative performance study of two MPI implementations on a dual-processor Linux cluster. I will present the results from this study and discuss our conclusions. I will also present our ongoing research into the issues related to storing and sharing performance data as part of an automated performance diagnosis environment.

Speaker's web page: <http://www.cs.pdx.edu/~karavan/research.html>

Institution web page: <http://www.pdx.edu/>

August 7, 2003

DNS of Particle-Laden Turbulent Flows

Said Elghobashi

University of California, Irvine

Email: selghoba@uci.edu

Abstract

Turbulent flows laden with particles are encountered in nature and numerous engineering applications. When the volume fraction of the dispersed phase is large enough, the turbulence properties (e.g., kinetic energy and dissipation spectra, strain rates) of the carrier flow undergo modifications, which, in turn, affect the dispersion characteristics of the particles (or bubbles). The objective of this seminar is to discuss the results of direct numerical simulations of two flows—particle-laden isotropic turbulence and particle-laden turbulent homogeneous shear flows—with emphasis on the physical mechanisms responsible for turbulence modifications.

Speaker's web page: <http://kolmog.eng.uci.edu/>

Institution web page: <http://www.uci.edu>

Sparse-Grid and Multigrid Methods for Multidimensional Problems

Christoph Reisinger

University of Heidelberg

Email: christoph.reisinger@iwr.uni-heidelberg.de

Abstract

Recently, there has been an increasing focus on problems where the dimensionality exceeds that which can be accommodated by standard tensor-product grid representations. Prominent examples include quantum mechanics (Schroedinger equation) and option pricing (Black–Scholes equation).

To deal with the “curse of dimensionality” encountered on, for example, Cartesian grids, we represent the solution on so-called sparse grids. In the combination technique, the solution is then extrapolated from small anisotropic grids with log-linear complexity. We can show that the convergence deteriorates only by logarithmic factors as well.

For a scalable method in a parallel setting, this demands from the linear (multigrid) solver not only optimal (linear) complexity that is robust with respect to the underlying equations, but also the anisotropy of the grids. Numerical results for block relaxation methods are shown for parabolic equations and free-boundary-value problems.

Institution web page: http://www.uni-heidelberg.de/index_e.html

July 18, 2003

Compiler Management of Global and Dynamic Data Reuse

Chen Ding

University of Rochester

Email: cding@cs.rochester.edu

Abstract

Reusing data in cache is critical to achieving high performance on modern machines. Many programs reuse a large amount of data in changing patterns across the whole program. Examples include an astronomical simulation of a changing galaxy in time steps, a compiler pulling together a set of functions in multiple passes, and a game-playing program calculating a sequence of moves through many searches. The question our work addresses is how well one can analyze and organize data reuses across large, dynamic program phases.

This presentation will describe several studies we conducted in the past seven years. First, I will describe a two-step strategy for improving global and dynamic cache reuse. The first step fuses computation on the same data to enable caching of repeated accesses. The second step groups data used by the same computation to effect contiguous access to memory. I will show how much we can automate this strategy across the whole program, over the entire data, and throughout program execution. Next, I will present recent results in predicting reuse distance pattern and demonstrate its use with a 3D web applet. For a wide range of benchmark programs, the applet predicts the cache miss rate and its changes for all cache sizes and all program inputs. Throughout the talk, I will also review classical studies that have laid the foundation for this and other work in memory hierarchy management.

Speaker's web page: <http://www.cs.rochester.edu/~cding/>

Research web page: <http://www.cs.rochester.edu/~cding/Research>

Institution web page: <http://www.rochester.edu/>

July 17, 2003

A Domain Decomposition Solver for a Parallel Adaptive Meshing Paradigm

Randolph Bank

University of California, San Diego

Email: reb@sdna1.ucsd.edu

Abstract

We describe a domain decomposition algorithm for use in the parallel adaptive meshing paradigm of Bank and Holst. Our algorithm has low communication, makes extensive use of existing sequential solvers, and exploits in several important ways data generated as part of the adaptive meshing paradigm. Numerical examples illustrate the effectiveness of the procedure. This work was done in collaboration with Shaoying Lu.

Speaker's web page: <http://www.scicomp.ucsd.edu/~reb/>

Research web page: <http://www.scicomp.ucsd.edu/~reb/>

Institution web page: <http://www.scicomp.ucsd.edu/>

July 11, 2003

Nonlinear Optimization Framework for Image-Based Modeling on Programmable Graphics Hardware

Radek Grzeszczuk

Intel Corporation

Email: Radek.Grzeszczuk@intel.com

Abstract

Graphics hardware is undergoing a change from fixed-function pipelines to more programmable organizations that resemble general purpose stream processors. In this paper, we show that certain general algorithms, not normally associated with computer graphics, can be mapped to such designs. Specifically, we cast nonlinear optimization as a data-streaming process that is well matched to modern graphics processors. Our framework is particularly well suited for solving image-based modeling problems since it can be used to represent a large and diverse class of these problems using a common formulation. We successfully apply this approach to two distinct image-based modeling problems: light field mapping approximation and fitting the Lafortune model to spatial bidirectional reflectance distribution functions. Comparing the performance of the graphics hardware implementation to a CPU implementation, we show a more than 5-fold improvement.

Speaker's web page: http://www.intel.com/research/people/bios/grzeszczuk_r.htm

Research web page: <http://www.intel.com/research/>

Institution web page: <http://www.intel.com>

PC-Based Visualization: Hardware and Software Solutions

Charles Hansen

University of Utah

Email: hansen@cs.utah.edu

Abstract

Clusters of PCs provide the opportunity for high-performance visualization methods of large-scale simulation data at lower cost than currently available. This talk will focus on three projects at Utah that leverage a PC cluster or PC graphics programmability applicable to clusters.

Real-time ray-tracing for isosurfacing has proven to be the most interactive method for large-scale scientific data. However, the limitation of requiring a tightly coupled DSM computer with a large number of processors has led to a bottleneck in the usability of such methods. We have developed a mechanism for allowing PC clusters to perform the same tasks as a tightly coupled Origin system. In particular, we are able to render the Lawrence Livermore National Laboratory Richmeyer–Meshkov instability dataset at multiple frames per second. Illumination in volume rendering has been limited because of the standardization on the ancient Phong shading model. We have developed a translucency model for volumetric rendering. Recently, we have extended this work to include an evaluation of a Gaussian approximation to the rendering equation that we call the Gaussian transfer function method. This is similar to preintegrated volume rendering for multidimensional transfer functions. The ability to interactively segment volumes for feature extraction and region-of-interest analysis is possible with the programmability of the modern graphics cards (GPU). We have developed a method for interactive level-set computation for segmentation and coupled this with real-time volume rendering. The contribution is the applicability of sparse methods to a coupled GPU and CPU implementation.

Speaker's web page: <http://www.cs.utah.edu/~hansen/>

Institution web page: <http://www.utah.edu/>

June 11, 2003

The Mathematica Platform

Andrew de Laix

Wolfram Research, Inc.

Email: delaix@wolfram.com

Abstract

Mathematica is a comprehensive platform for software development that combines a large library of algorithms, an advanced programming language, and an interactive notebook interface into a single package. This talk will offer a conceptual overview of Mathematica, focusing on its key components: the algorithm knowledge base, the symbolic programming language, and the document-centered interface. The algorithm knowledge base contains over 4000 functions forming a wide-ranging toolbox for building software. The symbolic programming language offers high-level constructs for manipulating not only numbers and arrays, but also complex tree structures, allowing one to write compact and flexible code. Mathematica's notebook interface enables one to combine text, mathematical typesetting, interactive elements, and code into a single document. Software, documentation, graphics, and interactive examples can be combined into a notebook for a unified presentation. Together, these elements constitute a flexible platform for developing a wide range of software solutions.

Institution web page: <http://www.wolfram.com/>

May 15, 2003

Scientific Data Management: A Case for Using High-level Information for I/O Optimizations

Alok Choudhary

Northwestern University

Email: choudhar@ece.northwestern.edu

Abstract

Management, storage, and efficient access to and analysis of huge amounts of data that are likely to be generated and/or used in various phases of large-scale scientific experiments and simulations are challenging tasks. Current data management and analysis techniques do not measure up to the challenges posed by such requirements in term of performance, scalability, ease of use, and interfaces.

In this talk, we will describe the research and development at Northwestern University to address the above problems. We will present optimizations techniques that use high-level access pattern information to enable optimizations in parallel file accesses. In particular, we describe techniques that can reduce or eliminate locking by explicitly managing I/O concurrency, which is an important step in achieving true scalability. We present performance results by incorporating these techniques in MPI-IO layer. In the second part, we will present the architecture and implementation of a metadata management system that allows the user to store, analyze, and use access patterns, relationships among data sets, data analysis, and I/O optimizations for scientific applications. We will describe the use of automatic I/O optimization techniques that can be incorporated into applications in a seamless fashion. Finally, we present recent results on data mining techniques that we are developing for online data analysis for scientific applications.

Speaker's web page: <http://www.ece.northwestern.edu/~choudhar>

Institution web page: <http://www.northwestern.edu/>

May 13, 2003

Automated Reconstruction of 3D City Models By Merging Ground- Based and Airborne Views

Avideh Zakhor

&

Christian Früh

University of California, Berkeley

Email: avz@eecs.berkeley.edu,
frueh@eecs.berkeley.edu

Abstract

Three-dimensional (3D) models of urban environments are useful in a variety of applications such as urban planning, computer games, special effects for the movie industry, training and simulation for urban terrorism scenarios, and virtual heritage conservation. Currently, acquisition of 3D city models is difficult and time consuming; hence, commercially available models typically take months to create, usually require significant manual intervention, and lack the level of detail necessary for many applications.

We present a fast approach to completely automated reconstruction of textured 3D city models with both high details at ground level and complete coverage for bird's-eye view. While driving under normal traffic conditions on public roads, a vehicle equipped with two 2D laser scanners and a digital camera can acquire close-range laser scans and images of the facades at the ground level. Furthermore, a far-range digital surface map (DSM) containing complementary roof and terrain shape is created from airborne laser scans, then triangulated, and finally texture mapped with aerial imagery. While an initial estimate for the acquisition vehicle's motion is derived from scan-to-scan matching, the final pose with respect to the DSM is determined by using Monte Carlo localization, hence globally registering the facade data with the DSM. We have developed a framework of data processing algorithms that copes with imperfections inherent in city laser scans such as foreground occlusions and reflections from glass surfaces and is capable of reconstructing both geometry and texture of the facades. The resulting facade models and the airborne mesh from the DSM are merged to a single consistent model by removing redundant parts and filling gaps. The developed algorithms are evaluated on a data set acquired in downtown Berkeley, California, and we present the resulting textured 3D city model.

Avideh Zakhor web page: <http://www-video.eecs.berkeley.edu/~avz>

Christian Früh web page: <http://www-video.eecs.berkeley.edu/~frueh/>

Research web page: <http://www-video.eecs.berkeley.edu>

Institution web page: <http://www.berkeley.edu/>

Adaptive Mesh Refinement for Miscible Displacement

John Trangenstein

Duke University

Email: johnt@math.duke.edu

Abstract

Multicomponent flow in porous media with random heterogeneity often involves concentration fronts that are difficult to resolve with finite difference methods. In order to concentrate the computational effort in modeling such problems, we will describe a dynamically adaptive mesh refinement algorithm.

The flow equations are of two types: convection-dominated mass conservation equations and an elliptic pressure equation representing a combination of Darcy's law and incompressibility. The mass conservation equations are integrated by formally second-order accurate donor cell upwind finite difference schemes. One such scheme, a modification of the corner-transport upwind scheme given an incompressible velocity field, is being developed with help from Jeffrey Hittinger. The pressure equation is discretized by a hybrid mixed finite element method. The permeability field is generated by a Markov random field.

Application of these numerical methods to hierarchical adaptive grids requires the application of several discrete numerical conservation laws for communication between scales. However, iterative solution of the linear system for the pressure field on the adaptive grid is significantly more complicated.

We will describe an iterative strategy that uses a conjugate gradient outer iteration, with multiplicative domain decomposition as its preconditioner. Multiplicative domain decomposition (which is similar to multigrid) is applied between levels of refinement. Within a level of refinement, we use a block Jacobi iteration equivalent to solve $2 \times 2 \times 2$ elliptic boundary value problems for the multigrid smoother. The 3D multigrid algorithm was completed and debugged with help from Robert Falgout and Panayot Vassilevski.

Numerical results for viscous fingering in 2D and 3D will be presented. The code has been parallelized, with significant help and code from Andrew Wissink and David Hysom. These results are being compared to the Todd-Longstaff mixing model to develop models of the error in its prediction of effective mixing. This work is being conducted in collaboration with Mike Christie of Heriot-Watt University.

Speaker's web page: <http://www2.math.duke.edu/faculty/johnt>

Institution web page: <http://www.duke.edu/>

April 14, 2003

What We Have Here Is a Failure to Communicate: Application Integration in Emergency Services

Bruce Rankin

City of Edmonton, Alberta, Canada

Email: Bruce.Rankin@edmonton.ca

Abstract

This presentation of the current state-of-the-art information systems for emergency first-response services gives insight into what works well in the real world and where there is still work to be done. The first surprising revelation is that no single information management system is used to respond to emergencies. Rather, a collection of systems is used by fire and emergency management services to get the needed information to the right people. A brief operations overview will describe the functionality required by these systems and provide a backdrop to a technical discussion outlining their capabilities. The technical discussion will focus on two examples used throughout the presentation: what happens during a 911 call and what happens during an extended emergency situation. These examples will highlight how information systems support logistical, administrative, operational, and strategic purposes. Integration issues will also be discussed, and I will conclude with goals for future computing environments.

Institution web page: <http://www.edmonton.ca/>

Supporting Soft Real-Time Processing in Best-Effort Systems

Scott Banachowski

University of California, Santa Cruz

Email: sbanacho@cse.ucsc.edu

Abstract

General-purpose computing platforms are commonly used to execute time-constrained applications—for example, running a multimedia player on a desktop system. These systems use best-effort scheduling algorithms that make no guarantees about the timing of resource allocation. The best-effort model is attractive for both developers and users because it doesn't require any interface for negotiating resource allocations. For the same reason, best-effort systems lack support for tasks with deadlines, limiting their usefulness to situations where the available resources far exceed the demands of applications.

This talk focuses on supporting soft real-time processing within the best-effort model. We present online techniques for inferring real-time behavior from tasks and improving responsiveness for those with periodic deadlines. The advantage of this approach is that a best-effort system can automatically provide soft real-time support without intervention from users or prior knowledge of tasks' timing constraints. We have created two CPU schedulers based on this technique that combine desirable aspects of both best-effort and soft real-time scheduling. The first scheduler provides accurate estimation of deadlines, and the second scheduler reduces scheduling latency while preserving the time-share notion of fairness.

Speaker's web page: <http://www.soe.ucsc.edu/~sbanacho/>

Institution web page: <http://www.ucsc.edu/>

April 4, 2003

An HLLC-Type Approximate Riemann Solver for Ideal Magnetohydro- dynamics

Katharine Gurski

National Institute of Standards and
Technology

Email: kgurski@nist.gov

Abstract

In the numerical simulation of magnetohydrodynamics (MHD), a balance is necessary between capturing the key features of the flow, limiting computational expense, and the robustness of the numerical method. This talk will present a method based on the HLLC approximate nonlinear Riemann solver for gas dynamics for the ideal MHD equations written in conservation form. The method is intended to be computationally inexpensive (compared to exact nonlinear solvers) and guaranteed in almost all cases to provide positive densities and pressures. The talk includes a brief introduction covering ideal MHD, Riemann problems, and the HLLC method for gas dynamics. Simulation results will concentrate on one-dimensional test cases.

Speaker's web page: <http://math.nist.gov/~KGurski>

Institution web page: <http://www.nist.gov>

Challenges in the Computational Discovery of Explanatory Scientific Models

Pat Langley

Stanford University & Institute for
the Study of Learning and Expertise

Email: langley@csli.stanford.edu

Abstract

The growing amount of scientific data has led to the increased use of computational discovery methods to understand and interpret them. However, most work has relied on knowledge-lean techniques such as clustering and classification learning, which produce descriptive rather than explanatory models, and it has used formalisms developed in AI or statistics, so that results seldom make contact with current theories or scientific notations. In this talk, I present a new approach to computational discovery that encodes explanatory scientific models as sets of quantitative processes, simulates these models' behavior over time, incorporates background knowledge to constrain model construction, and induces these models from time-series data in a robust manner. I illustrate this framework on data and models from earth science and microbiology, two domains in which explanatory process accounts occur frequently. In closing, I describe our progress toward an interactive software environment for the construction, evaluation, and revision of such explanatory scientific models. This talk describes joint work with Kevin Arrigo, Stephen Bay, Lonnie Chrisman, Dileep George, Andrew Pohorille, Javier Sanchez, Dan Shapiro, and Jeff Shrager.

Speaker's web page: <http://www.isle.org/~langley/>

Research web page: <http://csl.stanford.edu/~langley/>

Institution web page: <http://www-csli.stanford.edu/csl/>

March 26, 2003

Existing and Emerging Approaches for Addressing the Processor Memory Gap

Lenny Oliker

Lawrence Berkeley National
Laboratories

Email: loliker@lbl.gov

Abstract

The growing gap between sustained and peak performance for scientific applications has become a well-known problem in high-performance computing. This talk presents existing and emerging approaches for addressing this gap. First, we examine the NEC SX6 vector processor, which offers the potential to deliver a substantial increase in computing capabilities for a significant number of computational science codes. We compare the intranode performance between the SX6 vector processor and the cache-based IBM Power3/4 superscalar architectures across a number of key scientific computing areas. We evaluate low-level system characteristics using microbenchmarks and present performance results for numerical codes from scientific computing domains, including astrophysics, fusion energy, materials science, fluid dynamics, and molecular dynamics.

Next we examine two emerging microarchitectures designed for high-end media processing that have the potential to bridge the processor memory gap. The VIRAM architecture uses novel PIM technology to combine embedded DRAM with a vector coprocessor for exploiting its large bandwidth potential. The Imagine architecture, on the other hand, provides a stream-aware memory hierarchy to support the tremendous processing potential of the SIMD-controlled VLIW clusters. Our experiments isolate the set of application characteristics best suited for each architecture and show a promising direction toward interfacing leading-edge media processor technology with high-end scientific computations.

Speaker's web page: <http://www.nersc.gov/~oliker/>

Institution web page: <http://www.lbl.gov/>

Middleware Support for Data Ensemble Analysis

Joel Saltz

Ohio State University

Email: saltz-1@medctr.osu.edu

Abstract

Dramatic decreases in the cost of storage, combined with equally dramatic improvements in network connectivity, will allow communities to collaboratively generate and analyze very large distributed datasets. We will describe application scenarios that motivate this work and provide a broad view of what advances in systems software are needed to make this vision a reality. In many application scenarios, datasets describe spatio-temporal regions. We will then describe techniques we have developed to support optimized distributed data storage, indexing, retrieval, and processing. Our approach is to develop systems software able to leverage knowledge of spatio-temporal descriptive metadata in a way that supports a broad range of application areas. We will describe techniques that target very large datasets distributed among storage systems located in multiple parallel machines and clusters. These techniques target spatio-temporal and relational queries directed against these large grid-based datasets. They also target multiple-query optimization techniques that involve identification of intermediate results shared between the queries, along with grid-based semantic caching and retrieval of these results.

Speaker's web page: <http://medicine.osu.edu/Informatics/saltz.html>

Research web page: <http://www.cs.umd.edu/projects/hpsl/chaos>

Institution web page: <http://www.osu.edu/>

March 5, 2003

From Elementary Reactions to Power Plants: In How Much Detail Should Combustion Be Modeled?

Frank Behrendt

Berlin University of Technology

Email: frank.behrendt@tu-berlin.de

Abstract

Modeling and simulation are an established third branch of science in addition to theory and experiment. While modeling and simulation have seen significant successes, principal limitations remain. This talk will address these limitations with respect to combustion processes.

Modeling and simulation of combustion aim at the development of more efficient, cleaner, and more economical combustion technologies. To achieve these developments through modeling, the underlying models have to undergo continuous improvement to increase their predictive power. But even with the multi-teraflops computers available today, only a limited range of scales can be resolved simultaneously. Consequently, to resolve microscopic and macroscopic scales at the same time, hierarchical models are required.

This talk will address hierarchical models for three cases—elementary chemical reaction mechanisms, droplet combustion, and turbulent flows. Principal approaches are discussed and their suitability evaluated. An integrated model for the burning chamber of a coal-fired power plant concludes the talk.

Speaker's web page: http://www.tu-berlin.de/fb6/ifu/rdh/RDH_eng/people.htm

Research web page: http://www.tu-berlin.de/fb6/ifu/rdh/RDH_eng/

Institution web page: <http://www.tu-berlin.de/>

PAPI and Dynamic Performance Analysis Technology

Phil Mucci

University of Tennessee, Knoxville

Email: mucci@cs.utk.edu

Abstract

In this talk, I will cover the status and direction of PAPI, Dynaprof, and related tool technologies. PAPI is an open, cross-platform interface to the performance analysis hardware found in most modern microprocessors. In its simplest form, this hardware consists of a small set of programmable event counters. In addition, it may contain some advanced features such as performance-monitor interrupts, instruction tracing, branch profiling, and address sampling. Monitoring these types of events has a variety of uses in application performance prediction, analysis, and subsequent optimization. PAPI provides two standardized interfaces, one for the application engineer looking for quick measurements and one for the tool developer, both providing maximum performance and flexibility. Dynaprof is a tool to allow the user to insert and remove performance instrumentation (such as PAPI) at runtime. Dynaprof uses either DPCL or DynInst to insert the instrumentation directly into the address space of the application. Dynaprof currently provides a PAPI probe for collecting hardware counter data and a wallclock probe for measuring elapsed time, both on a per-thread basis. Users may write their own probes, and that probe may use whatever output format is appropriate—for example, a real-time data feed to a visualization tool or a static data file dumped to disk at the end of the run. Future plans for Dynaprof will be discussed, including integration with the University of Wisconsin's Tool Daemon Protocol, Lawrence Livermore National Laboratory's Tool Gear System, the University of Oregon's TAU, Sandia National Laboratory's Vprof, and Rice University's PAPIprof.

Speaker's web page: <http://www.cs.utk.edu/~mucci>

Research web page: <http://icl.cs.utk.edu/projects/papi/>

Institution web page: <http://www.utk.edu>

February 19, 2003

A Data Mining Approach to Modeling of Fusion Plasmas

Michael Walker

General Atomics

Email: walker@fusion.gat.com

Abstract

Experimental fusion devices have a useful life of several years and generate enormous quantities of raw and processed data. For example, just the raw diagnostic data archived for the DIII-D fusion experiment now increases by over 400 gigabytes (compressed) per year. Traditional analysis approaches typically focus on a few plasma discharges at a time, sometimes only one, out of the thousands collected per year. An opportunity exists for development and application of software tools that can process large numbers of these discharges to look for patterns in the data, a process usually referred to as data mining.

A significant challenge to developers of scientific data mining technology is the need to provide methods that can derive knowledge from data in a form useful to the scientist within a particular discipline. In tokamak plasma physics, for example, knowledge representations range from experientially derived procedures for how to access various plasma regimes during experimental operations to detailed and sophisticated models of plasma behavior that combine multiple phenomena such as particle transport, effects of external magnetic fields, pressure, or radio-frequency heating. The most sophisticated representations often take the form of detailed simulations, while simpler empirical models, or scaling laws, are often used to provide intuitive understanding.

Empirical physics models are also often the first step in the development of theory derived from observed experimental behavior. This talk will discuss the process of deriving empirical physics models from experimental data when viewed as a computational problem. Overview descriptions of some applications of data mining tools to four representative fusion plasma physics problems will be given.

Institution web page: <http://www.generalatomics.com/>

DEFACTO: Combining Parallelizing Compiler Technology with Hardware Behavioral Synthesis

Pedro Diniz

University of Southern California

Email: pedro@isi.edu

Abstract

Field Programmable Gate Arrays (FPGAs) offer the promise of substantial performance improvements over conventional processors by allowing the implementation of application-specific data paths that exploit instruction-level parallelism or domain-specific numeric formats and operations. Unfortunately, FPGAs are still difficult to program, making them inaccessible to the average developer. The standard practice requires developers to express the application program in a hardware-oriented language such as Verilog or VHDL and to synthesize the hardware design using a variety of synthesis tools. Because of the complexity of the synthesis process, it is difficult to predict a priori the performance and space characteristics of the resulting design. For this reason, the developer usually engages in an iterative design process, examining the results and modifying the design to trade off performance for space.

In this presentation, we describe DEFACTO, a system that automatically maps computations written in high-level imperative programming languages such as C to multi-FPGA-based systems. DEFACTO combines parallelizing compiler technology with commercially available behavioral synthesis tools. We use synthesis estimation techniques to guide the application of high-level program transformations in the search of high-quality hardware designs, thereby avoiding the long compilation-synthesis design cycles. We illustrate the effectiveness of DEFACTO in exploring a wide space of implementation designs for a set of image processing computations. For these computations the system searches on average less than 0.3% of the design space while deriving an “optimal” implementation, leading up to four orders of magnitude reduction in design time. We also describe in detail the mapping of a digital image processing computation, the Sobel edge detection, for which DEFACTO yields a 60-fold reduction in design time with only a 59% increase in execution time as compared to a manual implementation of the same algorithm.

This work has been sponsored by the Defense Advanced Research Project Agency (DARPA) under contract # F30602-98-2-0113.

Speaker's web page: <http://www.isi.edu/~pedro/>

Research web page: <http://www.isi.edu/asd/defacto>

Institution web page: <http://www.usc.edu>

February 18, 2003

Computational Approaches Towards Functional Annotation of the Human Genome

Ivan Ovcharenko

Lawrence Berkeley National Laboratory

Email: ovcharenko1@lbl.gov

Abstract

A first glance at the complete genomic sequence of humans revealed evidence for about 30,000 unique genes. The proteins encoded by these genes are responsible for constructing the human body from a single cell and regulating all the biochemical processes required for maintaining human health. While there is extensive knowledge about the function of many human genes from studies in model organisms and from hints from genetic disorders, the sequences required for the transcriptional regulation of these genes are poorly defined or as of yet unidentified. For example, why is BMP4, a gene responsible for teeth generation, shared by humans and birds, but the latter lack teeth? What sequences differentially regulate the expression pattern of BMP4 in birds and humans? How can we use computational tools to assist us in identifying regulatory DNA sequences encoded in the human genome?

Comparative sequence analysis has been proven to be an efficient and powerful approach to identify functional regions in the human genome. Genomes of different organisms have significantly diverged throughout the evolution of life on Earth, while orthologous genes and functional regulatory elements have been evolutionarily conserved. This talk will introduce computational strategies for aligning the human, mouse, and fugu genomes, and for predicting putative regulatory elements by combining transcription factor binding site prediction and the analysis of inter-species sequence conservation. I will also discuss how we can use the information derived from sequence alignments to understand the landscape of the human genome in terms of gene density, evolutionary conservation, and noncoding gene regulation.

Efficient Discovery of Previously Unknown Patterns and Relationships in Massive Time Series Databases

Eamonn Keogh

University of California at Riverside

Email: eamonn@cs.ucr.edu

Abstract

To date, the vast majority of research on time series data mining has focused on similarity search and, to a lesser extent, on clustering. We believe that these problems should now be regarded as essentially solved. From a knowledge-discovery viewpoint, several important unsolved problems in time-series data mining are more interesting, important, and challenging. In future work, we intend to address these problems. Our long-term goal is the creation of efficient algorithms to allow the extraction of knowledge in the form of patterns, anomalies, regularities, and rules from massive time-series datasets.

We will begin with a summary of the state of the art in time-series data mining, including a review of the multitude of representations proposed in the literature. We will then concretely define the three major problems we wish to address:

- enumerating frequently occurring patterns in time-series databases,
- the discovery of surprising or anomalous patterns in time series, and
- the discovery of causal/association rules in time series.

We will consider the limitations of the current work in these areas, enumerate the challenges faced, and sketch out a promising unpublished idea with implications for all these problems.

Speaker's web page: <http://www.cs.ucr.edu/~eamonn/>

Institution web page: <http://www.cs.ucr.edu>

February 7, 2003

Mixed Finite Elements for Elasticity

Douglas Arnold

University of Minnesota

Email: arnold@ima.umn.edu

Abstract

There have been many efforts, dating back four decades, to develop stable mixed finite elements for the stress-displacement formulation of the plane elasticity system. The key is to design a finite-element discretization for the space of square-integrable symmetric tensor fields with square-integrable divergence. This discretization enjoys a variety of properties. Although there are a number of well-known discretizations of $H(\text{div})$ vector fields with the analogous properties, such finite elements for symmetric tensor fields have proven very hard to design. We will present a new family of such elements, one for each polynomial degree quadratic and above. We will also analyze the obstructions to the construction of such elements, which account for the paucity of elements available. A star supporting role will be played by the de Rham sequence and related sequences of partial differential operators and their discrete analogues.

Speaker's web page: <http://www.ima.umn.edu/~arnold/>

Institution web page: <http://www.umn.edu/>

February 6, 2003

The Institute for Mathematics and Its Applications

Douglas Arnold

University of Minnesota

Email: arnold@ima.umn.edu

Abstract

The Institute for Mathematics and Its Applications is one of the world's foremost institutions devoted to increasing the impact of mathematics by fostering interdisciplinary work. In this informal presentation, the director of the Institute for Mathematics and Its Applications will give an overview of upcoming activities and opportunities at the IMA.

Speaker's web page: <http://www.ima.umn.edu/~arnold/>

Institution web page: <http://www.umn.edu/>

February 3, 2003

Software Quality Process: Is It a Square Peg in a Round Hole?

Stan Rifkin

Master Systems

Email: sr@Master-Systems.com

Abstract

We in the business of helping organizations change often label the members of our audience “bad characters” because they don’t agree with us. This presentation offers a different view: our clientele are not bad characters, and if we listened to them and acted on what is underneath what they say, the problem of implementation transforms from convincing, persuading, and cajoling to controlling wild horses that want to dash off with the new initiatives.

The key observation is that our software process improvement programs are not aligned with the strategy—stated or not—of our organizations, and those whom we label as resisters are actually performing a valuable service by pointing out the gaps between what drives the organization and what we are proposing.

There are only three strategies, and, oddly enough, many quality improvement agendas are not the same as what gets rewarded, so there is a natural tension between our proposals and what the organization does. This presentation offers concrete antidotes, surprisingly not by changing the rewards for quality but rather by changing the quality programs.

Speaker’s web page: <http://www.master-systems.com/Papers.ivnu>

Institution web page: <http://www.master-systems.com/>

Geometric Programming with a Functional Language

Alberto Paoluzzi

Third University of Rome, Italy

Email: paoluzzi@dia.uniroma3.it

Abstract

The aim of this talk is to introduce the new features of PLaSM, a design environment for geometric modeling and animation that supports rapid prototyping of geometric structures but does not deprive the user of direct control over the underlying geometric programming. PLaSM is a geometry-oriented extension of a subset of FL, developed by the Backus Functional Programming Group at IBM Almaden, that introduced an algebra over programs, in which a set of algebraic identities between functional expressions is established. PLaSM was recently extended with object-oriented features and with heuristic functions, allowing for some rule-based forward-chaining reasoning. The approach to the integration of the three paradigms, to our knowledge, is new. Joint work with CASC researchers is under way to design a new geometric kernel based on progressive combination and visualization of large-scale spatial indices. The PLaSM project (<http://www.plasm.net>) is open-source and multiplatform (windows, linux, macosx). A book on the subject is in press.

Speaker's web page: <http://www.dia.uniroma3.it/~paoluzzi/>

Research web page: <http://www.plasm.net/>

Institution web page: <http://www.dia.uniroma3.it/index-en.html>

January 10, 2003

Active Learning with Multiple Views

Ion Muslea

University of Southern California

Email: muslea@isi.edu

Abstract

L labeling training data for machine learning algorithms is tedious, time-consuming, and error-prone. Consequently, it is of utmost importance to minimize the amount of labeled data that is required to learn a target concept. In the work presented here, I focus on reducing the need for labeled data in multiview learning tasks. The key characteristic of multiview learning tasks is that the target concept can be independently learned within different views (i.e., disjoint sets of features that are sufficient to learn the concept of interest). For instance, robot navigation is a two-view learning task because a robot can learn to avoid obstacles based on either sonar or vision sensors. In my dissertation, I make three main contributions. First, I introduce Co-Testing, which is an active learning algorithm that exploits multiple views. Co-Testing is based on the idea of learning from mistakes. More precisely, it queries examples on which the views predict a different label: if two views disagree, one of them is guaranteed to make a mistake. In a variety of real-world domains, from information extraction to text classification and discourse-tree parsing, Co-Testing outperforms existing active learners. Second, I show that existing multiview learners can perform unreliably if the views are incompatible or correlated. To cope with this problem, I introduce a robust multiview learner, Co-EMT, which interleaves semisupervised and active multiview learning. My empirical results show that Co-EMT outperforms existing multiview learners on a wide variety of learning tasks. Third, I introduce a view validation algorithm that predicts whether or not two views are adequate for solving a new, unseen learning task. View validation uses information acquired while solving several exemplar learning tasks to train a classifier that discriminates between tasks for which the views are adequate and inadequate for multiview learning. My experiments on wrapper induction and text classification show that view validation requires a modest amount of training data to make high-accuracy predictions.

Speaker's web page: <http://www.isi.edu/~muslea/>

Research web page: <http://www.isi.edu/~muslea/papers.html>

Institution web page: <http://www.isi.edu>

Computational Problems in Proteomics: Statistics, Optimization, and Combinatorics

Alex Pothen

Old Dominion University

Email: pothen@cs.odu.edu

Abstract

Now that the genomes of many organisms have been sequenced, large-scale projects are under way to characterize the protein products of the genes (the proteome) and the multiprotein complexes that are responsible for the functions of the cell. High throughput, rapid, and automatable techniques are currently being developed to identify the tens of thousands of proteins (or more) involved. Among these are protein chips, novel mass spectrometric techniques, and nuclear magnetic resonance (NMR) methods.

Two computational problems in proteomics will be considered in this talk.

The first problem is to use the protein profiles of tissues, obtained via novel mass spectrometric techniques (SELDI and MALDI), to classify tissues into diseased or healthy specimens. Statistical classification and support vector machines are used to discover protein markers that characterize disease.

The second problem is to computationally represent multiprotein complexes and protein interaction networks using graphs and hypergraphs to enable algorithms for answering biological questions. We describe a new algorithm for identifying a “k-core” of a hypergraph (a subhypergraph in which every vertex belongs to at least “k” hyperedges of the subhypergraph) and use it to characterize core proteomes of yeast. The biological significance of core proteomes is that they are expected to have similar functions in related organisms.

Speaker's web page: www.cs.odu.edu/~pothen

Institution web page: www.cs.odu.edu/

December 13, 2002

Digital Libraries for Scientific Data Management

Reagan Moore

San Diego Supercomputer Center

Email: moore@sdsc.edu

Abstract

Across the data-grid, digital-library, and persistent-archive communities, a consensus is emerging on the software infrastructure needed for scientific data management. Each community has focused on a particular set of capabilities needed to manage data distributed across administration domains and multiple types of storage repositories. Data grids look at latency management and storage repository federation. Digital libraries look at discovery mechanisms for finding relevant material. Persistent archives examine the management of technology evolution. The common approach that is emerging is based upon:

- the use of logical name spaces to provide infrastructure-independent naming,
- storage repository abstractions to define the set of operations used to manipulate remote digital entities, and
- information repository abstractions to define the set of operations needed to manipulate collections stored in databases.

The common approach will be discussed, along with representative projects from each of these data-management communities.

Research web page: <http://www.sdsc.edu/~moore/rmoore.html>

Institution web page: <http://www.sdsc.edu/>

Tools and Benchmarks for Performance Evaluation of Applications in Science

Rudolf Eigenmann

Purdue University

Email: eigenman@ecn.purdue.edu

Abstract

I will give an overview of a number of projects that create programming environments and a performance evaluation infrastructure for computational applications. Two important parts of the environment are the Ursa Minor interactive performance tuning tool and the Polaris parallelizing compiler. Ursa Minor combines information gathered from sources such as compilers and runtime performance monitors and presents it in a uniform way to the user. It supports stepwise performance improvements, following a methodology of performance analysis and improvement for OpenMP programs. The Polaris compiler furnishes Ursa Minor with information about program characteristics and automatable parallelism. An important part of every performance evaluation infrastructure is the test program suite, i.e., benchmarks. We have participated in an effort to collect the largest possible public computational applications and to create benchmark suites for both research and industry. The most recent suite is just being released under the name SPEC HPC2002.

Speaker's web page: <http://ece.www.ecn.purdue.edu/~eigenman/Index.html>

Research web page: <http://ParaMount.www.ecn.purdue.edu/ParaMount/>

Institution web page: <http://ece.www.ecn.purdue.edu>

December 6, 2002

The Oxymoron of Computer "Science"

John Perry

Foothill Deanza College

Email: perry@voyager.deanza.fhda.edu

Abstract

The author gives a historical account of the field of computer science since the 1970s. The purpose of this historical review is to show how the field is, in fact, not a science at all but rather the practicing of pop psychology by charismatic Ph.D.'s, the imposition of empirically unproven paradigms and procedures on programmers, and a sociological phenomenon replete with elements of "hipness" and associated in-group and outcast behaviors. The author illustrates how the celebrityization of personalities, the religification of computers, and the lack of respect for true science have led to regressive software, regressive programming languages, regressive paradigmatic systems, and the death of common sense in a field allegedly allied with the "hard" sciences.

Institution web page: <http://www.deanza.edu/>

Detecting and Exploiting Spatial Regularity in Memory References

Tushar Mohan

University of Utah

Email: tushar@cs.utah.edu

Abstract

Rising processor speed, unaccompanied by corresponding reductions in memory access latency, has caused the performance of codes to be limited by memory accesses. Strided memory accesses, or streams, if known to exist in an application, can be targeted by a host of optimizations, such as stream prefetching, relocation, remapping, and vector loads. Undetected, they can be a significant source of memory stalls in loops. Existing stream-detection mechanisms either require special hardware, which may not gather enough stream statistics for subsequent analysis, or are confined to limited compile-time detection of array accesses in loops. Little formal treatment has been accorded to the subject; the concept of locality fails to capture the existence of streams in a program's memory accesses. In this thesis we define spatial regularity as a means to depict the presence of strided memory accesses. We develop measures to quantify spatial regularity and design and implement an online, parallel algorithm to detect streams, and hence regularity, in running applications. We identify critical program sections for regularity measurements by using PAPI—a performance measurement API—to access hardware performance counters portably. Dyninst's dynamic binary translation infrastructure is leveraged to perform selective and transitory instrumentation in the application. This development allows the user to limit the stream detection overhead at the cost of measurement accuracy. We use examples from real codes and popular benchmarks to illustrate how stream information can be used to effect profile-driven optimizations.

Speaker's web page: <http://www.cs.utah.edu/~tushar/>

Institution web page: <http://www.utah.edu/>

November 8, 2002

Link-Time Optimization of Parallel Scientific Programs

Gregory Andrews

University of Arizona

Email: greg@CS.Arizona.EDU

Abstract

A binary rewriting system transforms a binary program into a different but functionally equivalent program. A link-time optimizer is a binary rewriting system that optimizes an object program to improve some aspect of its performance, such as execution time, code size, power consumption, or communication bandwidth. A significant benefit of a link-time optimizer is that it can perform whole program optimizations that cannot be done by compilers—e.g., those based on run-time values—or that usually are not done by compilers—e.g., because library source code is not available or is written in a different language.

The SOLAR project is developing binary rewriting techniques for software optimization at link time and run time. A particular interest is improving the performance of parallel scientific programs that use a communications library such as MPI. This talk will describe the link-time optimizer we have developed for the Pentium architecture and demonstrate the improvements it is able to make to already highly optimized application programs. The talk will also describe several useful tools that have been relatively easy to construct given our Pentium link-time optimizer.

Speaker's web page: <http://www.cs.arizona.edu/people/greg/>

Institution web page: <http://www.arizona.edu/>

On Balanced Approximations for Time Integration of Multiple Time Scale Systems

Dana Knoll

Los Alamos National Laboratory

Email: nol@lanl.gov

Abstract

The effect of various numerical approximations used to solve linear and nonlinear problems with multiple time scales is studied in the framework of modified equation analysis (MEA). MEA is used to study the effect of linearization and splitting in a simple nonlinear ordinary differential equation (ODE) and in a linear partial differential equation (PDE). Several forms of time differencings of the ODE and PDE are considered, and the resulting truncation terms are compared analytically and numerically. It is demonstrated quantitatively that both linearization and splitting can result in accuracy degradation when a computational time step larger than the competing (fast) time scales is employed. Many of the issues uncovered on the simple problems are shown to persist in more realistic applications.

Institution web page: <http://www.lanl.gov>

October 28, 2002

Physics-Based Preconditioning for Jacobian-Free Newton–Krylov Methods

Dana Knoll

Los Alamos National Laboratory

Email: nol@lanl.gov

Abstract

An important class of preconditioners for the Jacobian-free Newton–Krylov method is referred to as physics-based or PDE-based. The motivation behind this approach is that there exist numerous, legacy algorithms to solve nonlinear systems, both IVPs and BVPs. These algorithms typically were developed with some insight into the time scales or physical behavior of the problem. As a benefit of this insight, a reduced implicit system, or a sequence of segregated explicit or implicit systems, may be solved in place of the fully coupled system. Examples include the semi-implicit method for low-speed flow (stiff-wave problem), the SIMPLE algorithm for incompressible flow, Gummel's method for the semiconductor drift-diffusion equations, and numerous other structure-based operator splitting methods for reaction-diffusion systems. We will outline the concept of physics-based preconditioning and then focus on our own work on applications with stiff-wave phenomena.

Institution web page: <http://www.lanl.gov/>

Numerical Study of Burn Propagation in Inhomogeneous Mixtures

**David Lopez and
Sutanu Sarkar**

University of California, San Diego

Email: ssarkar@ucsd.edu

Abstract

In inertial confinement fusion applications, the reactant mix can be contaminated by inert material introduced by Rayleigh–Taylor hydrodynamic instabilities promoted at sites of imperfect surface finish of the capsule. We have performed multidimensional simulations to study burn-front propagation through a contaminated mixture. The flow field and chemistry are accurately resolved on an instantaneous space–time basis. Associated with the large density change across the burn front and consequent fluid acceleration, there is a natural instability, the so-called Darrieus–Landau instability discussed in premixed combustion literature. The coupling of contaminant and velocity fluctuations in the burn mix with the natural instability is found to significantly affect burn-front propagation. In the talk, we will discuss the simulation method and results illustrating the modified burn propagation in a contaminated mix.

Research web page: <http://www-mae.ucsd.edu/RESEARCH/SARKAR/sarkar.html>

Institution web page: <http://www.ucsd.edu/>

October 8, 2002

Learning to Classify Galaxy Shapes Using the EM Algorithm

Sergey Kirshner

University of California, Irvine

Email: skirshne@odessa.ics.uci.edu

Abstract

I will describe the application of probabilistic model-based learning to the problem of automatically identifying classes of galaxies, based on both morphological and pixel intensity characteristics. The EM algorithm can be used to learn how to spatially orient a set of galaxies so that they are geometrically aligned. This “orientation-model” is augmented with a mixture model on objects, and classes of galaxies can then be learned in an unsupervised manner using a two-level EM algorithm. The resulting models provide highly accurate classification of galaxies in cross-validation experiments.

Research web page: <http://www.ics.uci.edu/~skirshne/>

Institution web page: <http://www.uci.edu/>

Leveraging PC Graphics Cards for Advanced Visualization

Charles Hansen

University of Utah

Email: hansen@cs.utah.edu

Abstract

Recent advances in the programmability of PC graphics hardware have led to corresponding advances in direct volume rendering. In this talk, I will present an overview of two such advances: the use of multidimensional transfer functions and the incorporation of advanced shading techniques. Most direct volume renderings produced today employ one-dimensional transfer functions, which assign color and opacity to the volume based solely on the single scalar quantity that comprises the dataset. Though they have not received widespread attention, multidimensional transfer functions are a very effective way to extract specific material boundaries and convey subtle surface properties. However, identifying good transfer functions is difficult enough in one dimension, let alone two or three. I will outline an important class of three-dimensional transfer functions for scalar data (based on data value, gradient magnitude, and a second directional derivative), and describe a set of direct manipulation widgets that make setting such transfer functions intuitive and convenient. I will also describe how to use modern graphics hardware to interactively render with multidimensional transfer functions. The transfer functions, widgets, and hardware combine to form a powerful system for interactive volume exploration.

Direct volume rendering is a commonly used technique in visualization applications. Many of these applications require sophisticated shading models to capture subtle lighting effects and characteristics of volumetric data and materials. Many common objects and natural phenomena exhibit visual quality that cannot be captured using simple lighting models or cannot be solved at interactive rates using more sophisticated methods. I will describe a simple yet effective interactive shading model, which captures volumetric light attenuation effects to produce volumetric shadows and the subtle appearance of translucency. By coupling this shading with a technique for volume displacement or perturbation, one can achieve realistic interactive modeling of high-frequency detail for real and synthetic volumetric data.

Research web page: <http://www.cs.utah.edu/~hansen/>

Institution web page: <http://www.utah.edu/>