Combinatorial Scientific Computing and Petascale Simulations (CSCAPES)

SciDAC Applied Math Institute

Alex Pothen and Assefaw Gebremedhin, Purdue University

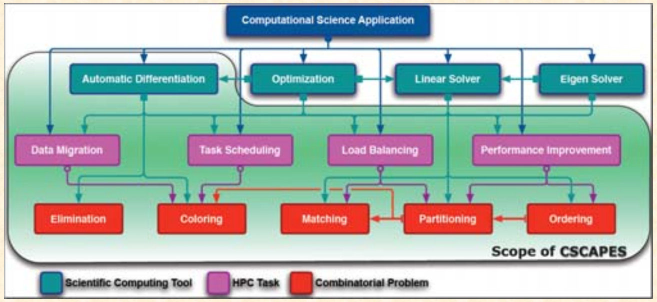
Erik Boman, Cedric Chevalier, Karen Devine and Bruce Hendrickson, Sandia National Laboratories

Paul Hovland, Boyana Norris, Ilya Safro and Jean Utke, Argonne National Laboratory

Umit Catalyurek and Doruk Bozdag, Ohio State University

Florin Dobrian, Conviva Corporation

Michelle Strout, Colorado State University

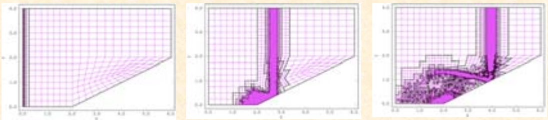


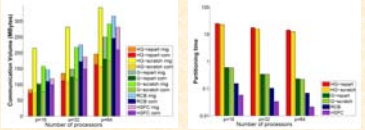
Combinatorial Scientific Computing (CSC)

CSC is concerned with the development and use of discrete algorithms in computational science and engineering. Graph and hypergraph algorithms are among the fundamental tools of CSC. They play a crucial enabling role in applications that require parallelization, optimization, solution of differential equations, mesh generation, etc. The CSCAPES Institute is founded on the recognition of this critical role. The diagram above shows the key research areas in CSCAPES and their relationship to a typical computational science application; an arrow from A to B indicates that A in some sense uses B.

Load Balancing

In parallel computation, data and tasks need to be mapped to processors such that workload is balanced and communication cost is low. In dynamic or adaptive applications, these goals must be met as computation and communication change over time. Zoltan is a software toolkit for parallelization and load balancing.
A hypergraph-based dynamic repartitioning method with superior performance compared to earlier methods has recently been added to Zoltan. This work earned the Best Algorithms Paper Award at IPDPS07.
The capabilities of Zoltan are being extended in several other ways to support petascale applications.

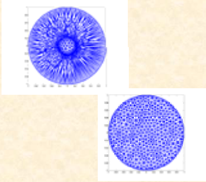




Performance comparison of hypergraph repartitioning (HG-repart) and several graph- and coordinate-based methods in an adaptive mesh refinement problem from Sandia's ALEGRA physics code. The top row shows the final element meshes at time-steps 0, 54, and 108.
The bottom row shows the reduced total communication volume obtained with hypergraph repartitioning, as well as repartitioning times.

Performance Improvement

Modern microprocessors are highly sensitive to spatial and temporal locality of data. In a mesh smoothing application, for example, reordering the vertices and elements of a mesh can lead to a significant improvement in single processor performance. CSCAPES is developing several data and iteration reordering algorithms based on hypergraph models.



CSCAPES Mission and Recent Activities

▪ Research and Development

- Provide load balancing and parallelization toolkits for petascale computation.
- Develop advanced automatic differentiation capabilities for complex applications.
- Advance the state-of-the-art in large-scale graph and sparse matrix computations.
- Develop data and iteration reordering algorithms for improving single processor performance.

▪ Training and Outreach

- In collaboration with SciDAC CETs and SAPs, support scientific applications with underlying graph and hypergraph algorithms. Current collaborators include ITAPS, TOPS, COMPASS and NREL.
- Organize workshops, tutorials and short courses in CSC. Recent activities include CSCAPES Workshop (Santa Fe, June 2008), minisymposium on CSC at SIAM Conference on Discrete Math (June 2008), and CSCAPES tutorial at the SciDAC 2007 Conference.
- Train researchers in CSC skills. Currently, 2 post-docs, 5 graduate and 1 undergraduate students are being trained.

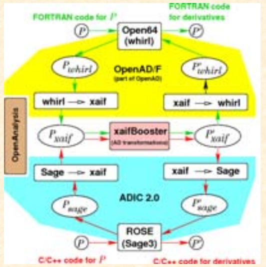
Selected Recent Publications

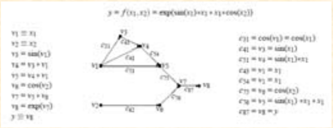
- Bozdag, Gebremedhin, Manne, Boman and Catalyurek. *A framework for scalable greedy coloring on distributed-memory parallel computers*. Journal of Parallel and Distributed Computing, 68(4):515-535, 2008.
- Wolf, Boman and Hendrickson. *Optimizing sparse matrix-vector multiplication by corner partitioning*. PARA 2008.
- Bucker, Bischof, Hovland, Naumann and Utke (editors). *Automatic Differentiation: applications, theory and implementations*. AD 2008, LNCSE 64.
- Gebremedhin, Pothen and Walther. *Exploiting sparsity in Jacobian computation via coloring and Automatic Differentiation: a case study in a Simulated Moving Bed process*. AD 2008.
- Catalyurek, Boman, Devine, Bozdag, Heaphy and Riesen. *Hypergraph-based dynamic load balancing for adaptive scientific computations*. IPDPS 2007. Won the Best Algorithms Paper Award.
- Gebremedhin, Tarafdar, Manne and Pothen. *New acyclic and star coloring algorithms with application to computing Hessians*. SIAM Journal on Scientific Computing 29(3):1042-1072, 2007.
- Hovland, Norris, Strout and Utke. *Term graphs for computing derivatives in imperative languages*. Electronic Notes on Theoretical CS, 2007.
- Pothen, Gebremedhin, Dobrian, Boman, Devine, Hendrickson, Hovland, Norris, Utke, Catalyurek and Strout. *Combinatorial algorithms for petascale science*. SciDAC Review, 5:26-35, 2007.

Further Information

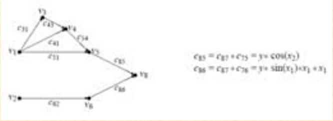
www.cscapes.org

OpenAD System Architecture





Decomposition of a function evaluation and its representation as a computational graph.
Graph resulting after vertex v_7 is eliminated (bottom).



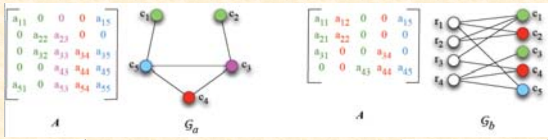
Automatic Differentiation (AD)

AD is a technology for computing analytic derivatives of functions specified by computer programs. Efficient derivative accumulation can be posed as a transformation (vertex or edge elimination) problem on a computational graph. Sparsity exploitation in large-scale derivative matrix computation gives rise to a variety of graph coloring problems. Checkpointing in irregular computation is another source of combinatorial problems in AD. CSCAPES is developing efficient algorithms for these problems. It is also extending the capabilities of the Argonne-housed AD tools ADIC and ADIFOR within the OpenAD framework, and plans to provide users with integrated load balancing and coloring capabilities.

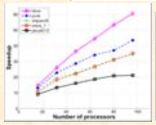
Graph Coloring

Graph coloring problems are powerful abstractions for minimizing the execution time, memory, and storage space needed in computing sparse derivative matrices using automatic differentiation. Graph coloring is also useful in discovering concurrency in parallel computation. CSCAPES researchers are developing novel sequential as well as parallel algorithms for a variety of coloring problems. Implementations of the parallel algorithms are being made available via Zoltan.

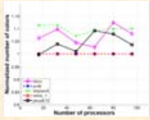
Structurally orthogonal column partition and its representation as a distance-2 coloring in the adjacency graph of a Hessian (left) or bipartite graph of a Jacobian.



Performance of a recently developed parallel distance-2 coloring algorithm on select applications.



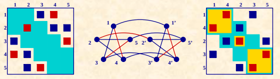
graph name	$\lvert V\rvert$	$\lvert E\rvert$	Degree max	colors (seq)
pkustk13	94,893	3,260,967	299	303
shipsec5	179,860	4,966,618	125	140
pwtk	217,918	5,708,253	179	180
inline_1	503,712	18,156,315	842	843
ldoor	952,203	22,785,136	76	112



Max degree is a lower bound on the optimal number of colors needed for distance-2 coloring. The parallel algorithm (far right) gave solutions very close to the sequential version, which in turn gave near optimal solutions (see last two columns of the table).

Matching

Various kinds of matchings in graphs are needed in several scientific or high-performance computing contexts. Examples include sparse linear systems (numerical preprocessing and block triangular decomposition) and graph partitioning (coarsening phase of multilevel methods). Traditional matching algorithms compute optimal solutions in superlinear time, but current trends are toward algorithms that find approximate solutions faster. CSCAPES is developing parallel matching algorithms based on approximation techniques.



Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the U. S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. The work at Argonne is supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, U.S. Department of Energy under Contract W-31-109-Eng-38.
The CSCAPES Institute is supported by the U.S. Department of Energy's Office of Science through grant DE-FC-0206-ER-25774, as part of its SciDAC program.