CRAIG SILVERSTEIN

Department of Computer Science
Stanford, CA 94305-9045
650 723-1787
http://www-cs-students.stanford.edu/~csilvers/

817 Santa Fe Ave
Stanford, CA 94305
650 494-3680
csilvers@cs.stanford.edu

Academic Interests

Information retrieval and data mining For two summers at Xerox PARC, I have worked with Jan Pedersen and Hinrich Schuetze on efficient clustering of large data sets using Scatter/Gather. As part of the advanced technology group at Verity, Inc., I modified the Verity search engine to support natural language queries. I am also active in the Data Mining Research Group at Stanford, having published papers on the use of correlation and causality in data mining.

Eperimental Analysis of Algorithms I have worked with Andrew Goldberg on extensive analysis of practical implementations for Dijkstra's shortest path algorithm. This work has led to the development of new algorithms, including one which improves on the best known time bounds for this problem. I am also analyzing perfect hashing algorithms.

Academic Experience

Stanford University, Stanford, CA
Candidate for degree of Ph.D. in Computer Science. Focus in information retrieval and data mining under the supervision of Rajeev Motwani.

Harvard University, Cambridge, MA
A.B. degree in Computer Science, with honors. June 1994. GPA: 14.33 (out of 15). GPA in Computer Science: 14.75 (out of 15).
Phi Beta Kappa (1993), recipient of Microsoft Technical Scholarship (1993), twice recipient of Derek Bok Award for Teaching Excellence (1992, 1993), captain of first-place team in Association for Computing Machinery Annual International Programming Contest (1993).

Teaching Experience

Teaching Assistant in Advanced Algorithms (1998),
Introduction to Computer Theory (1997),
and Network Optimization (1996) STANFORD UNIVERSITY
Taught an hour-long section each week, graded problem sets, helped compose and grade problem sets and exams.
Winter 1998, Autumn 1997, Winter 1996

Teaching Fellow in Introduction to Computer Theory (1993)
and Introduction to Programming (1992, 1991) HARVARD UNIVERSITY
Taught an hour-long section of 15-18 students a week. Graded problem sets, helped compose and grade exams, tutored on an individual basis.
Fall 1993, 1992, 1991

(over)

Work Experience

Consultant at Verity, Inc. SUNNYVALE, CA
As part of the Advanced Technology Group, developed a new technique for processing natural language queries and modified the Verity search process to incorporate this new technology.
Summer 1997 – present

Researcher at Xerox PARC PALO ALTO, CA
Worked with Jan Pedersen to develop almost-constant-time clustering algorithms for large data sets. I developed a technique to merge this fast clustering with boolean search (patent pending). I have also explored, with Hinrich Schuetze, the application of Latent Semantic Indexing to clustering.
Summer 1995 – January 1997

Software Design Engineer at Microsoft REDMOND, WA
In 1993, developed a segment of a communications interface for Microsoft's entry into the handheld computer market. In 1992, conceived, designed, and programmed one segment — a game — of a multimedia project involving works of Igor Stravinski.
Summer 1993, 1992

Publications

Beyond Market Baskets: Generalizing Association Rules to Correlations by S. Brin, R. Motwani, and C. Silverstein. To appear in *Data Mining and Knowledge Discovery*, 1998.

A Comparison of Projections for Efficient Document Clustering by H. Schütze and C. Silverstein. *Proceedings of the Twentieth Annual ACM SIGIR Conference on Research and Development in Information Retrieval.* July 1997, pp. 74–81.

Almost-Constant-Time Clustering of Arbitrary Corpus Subsets by J. Pedersen and C. Silverstein. *Proceedings of the Twentieth Annual ACM SIGIR Conference on Research and Development in Information Retrieval.* July 1997, pp. 60–66.

A Practical Perfect Hashing Algorithm by C. Silverstein. To appear in the *DIMACS Series in Discrete Mathematics and Theoretical Computer Science.*

Buckets, Heaps, Lists, and Monotone Priority Queues by B. V. Cherkassky, A. V. Goldberg, and C. Silverstein. To appear in *SIAM Journal on Computing.*

Implementations of Dijkstra's Algorithm Based on Multi-Level Buckets by A. V. Goldberg and C. Silverstein. Technical Report 95-187, NEC Research Institute, Inc., November 1995.

A Practical Evaluation of Kinetic Data Structures by J. Basch, L. J. Guibas and C. Silverstein. *Proceedings of the 13th ACM Symposium on Computational Geometry*, June 1997.

Constrained TSP and Low-Power Computing by M. Charikar, R. Motwani, P. Raghavan, and C. Silverstein. *Proceedings of WADS*, 1997.

Predicting Individual Book Use for Off-Site Storage Using Decision Trees by C. Silverstein and S. Shieber, *Library Quarterly*, July 1996, pp. 266–293.