The Lynne and William Frankel Center
for Computer Science
Department of Computer Science
Ben Gurion University of the Negev
Tel: 08-6428032 Fax: 08-6428021
fradmin@cs.bgu.ac.il

Distinguished Lecturer Series

Prof. Jeffrey Ullman

Stanford W. Ascherman Professor of Computer Science (Emeritus)

Cluster-Based Computation of Relational Joins

The prevalence of large racks of interconnected processor nodes forces us to take another look at how to exploit parallelism when taking the join of large relations. Sometimes, there is a gain in total cost to be had by distributing pieces of each relation to several different nodes and computing the join of several large relations at once. The optimization problem is to pick the degree of replication of each relation, under the constraint that the total number of compute-nodes is fixed. We set up this problem as a nonlinear optimization and show that there is always a solution (which must be approximated by rounding to the nearest integers). For some of the most common types of join -- star joins and chain joins -- we give closed-form solutions to the optimization problem. Finally, we point out that the join algorithm we propose can be implemented using features already present in Hadoop, the open-source implementation of map-reduce.

Jeffrey D. Ullman is currently the Stanford W. Ascherman Professor of Computer Science (Emeritus) at Stanford University, as well as CEO of the Gradiance Corporation.
He received a Bachelor of Science degree in Engineering Mathematics from Columbia University and his Ph.D. in Electrical Engineering from Princeton University. Previously he worked at Bell Labs for several years and then as a professor at Princeton University. Since 1979 he has been a professor at Stanford University. In 1995 he was inducted as a Fellow of the Association for Computing Machinery and in 2000 he was awarded the Knuth Prize.

His research interests include database theory, data integration, data mining, and education using the information infrastructure. His textbooks on compilers, data structures, theory of computation, and databases are regarded as standards in their fields. He is one of the founders of the field of database theory, and was the doctoral advisor of an entire generation of students who later became leading database theorists in their own right. He was the Ph.D. advisor of Sergey Brin who was one of the co-founders of Google. Professor Ullman also served on Google's technical advisory board.

12:00-13:00 on Tuesday, 16 June, 2009—Saal Auditorium, Alon Bldg (37/202)
12:00-13:00 יום ג' 16 ביוני 2009—באודיטוריום סאל בבנין אלון (202/37)