Seminar Series Supported by Jeffrey and Holly Ullman

BGU Computer Science Theory Day
22 May, 2008

10:30   Coffee and tagging

11:00   Property Testing or Approximate Decision Making
Dana Ron, Tel-Aviv University

Abstract: **In broad terms, property testing is the study of the following class of problems: Given the ability to perform (local) queries concerning a particular object (e.g., a function, or a graph), the task is to determine whether the object has a predetermined (global) property (e.g., linearity or 3-colorability), or is far from having the property (i.e., should be modified significantly so as to obtain the property).**

Thus property testing is a relaxation of exact decision making, and can be viewed as approximate decision making. As such, property testing algorithms are expected to be much more efficient than the corresponding exact decision procedures. In particular, we would like the algorithm to inspect only a small (randomly selected) part of the whole object, and to run in time sublinear (or even independent) of the size of the object, where a small probability of failure is allowed.

Property testing of functions was first explicitly defined by Rubinfeld and Sudan in the context of program testing, where their focus was on testing algebraic properties of functions. The study of properties of combinatorial objects, and in particular graphs, was initiated by Goldreich, Goldwasser and Ron. In recent years, property testing has been studied in many additional contexts.

In this talk I will not attempt to provide a comprehensive survey, but rather I shall describe several examples that will give the flavor of research on property testing, and also mention some recently studied extensions.

11:45   Local Testing of Codes
Irit Dinur, Weizmann Institute

12:30   Lunch

13:30   Auditing Algorithms for Data Privacy
Nina Mishra, Microsoft Research Silicon Valley Search Labs
and University of Virginia

Abstract: **Private data sets are ubiquitous and the need to extract valuable information from this data is real. We consider the online query auditing problem: given a sequence of queries that have already been posed about a private data set, their corresponding answers and given a new query, deny the answer if privacy can be breached or give the true answer otherwise. We uncover the fundamental problem that query denials leak information—a problem that was overlooked in previous work.**

Because of this oversight, some of the previously suggested auditors can be used by an attacker to compromise the privacy of a large fraction of the individuals in the data. To overcome this problem, we introduce a new model called simulatable auditing where query denials provably do not leak information. We present a general framework for designing auditing algorithms and show how a variety of problems can be solved in this framework including auditing sum queries, auditing max queries and the release of private contingency tables.

Based on joint work with Krishnaram Kenthapadi, Kobbi Nissim and Shubha Nabar

14:15   Coffee Break

14:30   A Fully Dynamic Distributed Algorithm for Maintaining Sparse Spanners
Michael Elkin, Ben-Gurion University

Abstract: **Motivated by real world networks and use of algorithms Currently, there are no known explicit algorithms for the great majority of problems in the dynamic distributed message-passing model. Instead, most state-of-the-art dynamic distributed algorithms are constructed by composing a static algorithm for the problem at hand with a simulation technique that converts static algorithms into dynamic ones. We argue that this powerful methodology does not provide satisfactory solutions for many important dynamic distributed problems, and consequently, it is desirable to develop algorithms for these problems from scratch.**

We develop a fully dynamic distributed algorithm for maintaining sparse spanners. Our algorithm improves drastically the quiescence time of the state-of-the-art algorithm for the problem. Moreover, we show that the quiescence time of our algorithm is optimal up to a small constant factor. In addition, our algorithm improves significantly upon the state-of-the-art algorithm in all efficiency parameters; specifically, it has smaller quiescence message and space complexities, and smaller local processing time.

The talk is based on a paper published this summer in PODC 2007.

15:15   Effective Sampling from Large Data Sets
Haim Kaplan, Tel-Aviv University

Abstract: **Sampling is an essential tool for management of massive data sets. It preserves, for future queries, the essence of data which may be transient and too large to store or access. We survey some recent sampling techniques which are particularly suitable for estimating weights of subsets with small variance.**

Organizers:

Amos Beimel, Ben-Gurion University
Kobbi Nissim, Ben-Gurion University