Spring 2009 Theory Seminar

Spring 2009 Theory Seminar

Monday 11:00-12:00, in room SC 3405

Theory seminar gnome: Sungjin Im (im3 at uiuc dot edu)

To receive weekly reminders and announcements, please sign up for the theory-seminar mailing list.

As usual, every theory student is expected to speak at least once, but non-theory students are also welcome to volunteer. Email Sungjin for a speaking slot. Student speakers are encouraged to register for CS 591JE (CRN 35949).

Schedule

- February 2 - Hemanta Maji talked about basic concepts in Information Theory and its application.

  Fun with Information Theory.
  Hemanta Maji

  This talk will mainly focus on developing basic concepts in Information Theory and a few rudimentary applications. The aim of the talk is to develop fundamental concepts, definitions and notations; and apply them to some well known problems in Data Compression and Kolmogorov Complexity.
  Data Compression: We will characterize optimal codes and bounds on their length. We will also see the optimality proof of Huffman codes.
  Kolmogorov Complexity: We will define Kolmogorov complexity and deal with concepts like "Complexity of integers" and "Algorithmically Random and incompressible sequences".
  To summarize, we will learn to converse in the language of Information Theory. The talk will be based on chapters 1, 2, 5 and 7 from the book "Elements of Information Theory", T. M. Cover and J. A. Thomas.

- February 9 - Dan Schreiber presented the following paper.

  Dan Schreiber

  Given $m$ copies of the same problem, does it take $m$ times the amount of resources to solve these $m$ problems? This is the direct sum problem, a fundamental question that has been studied in many computational models. We study this question in the simultaneous message (SM) model of communication introduced by Yao. Like many other things, we have two parties Alice and Bob, with inputs $x,y$, they wish to compute $f(x,y)$. In the simultaneous message model they both send one message to a third party, the Referee, who then computes the function value based on the messages received from Alice and Bob. The protocol can be either deterministic, or randomized with a probability $\varepsilon$ of error.
  The equality problem for $n$-bit strings is well known to have SM complexity $\Theta(\sqrt{n})$. We prove that solving $m$ copies of the problem has complexity $\Omega(m\sqrt{n})$, which matches obvious upper bounds; the best lower bound provable using previously known techniques is $\Omega(\sqrt{mn})$.
  We introduce a new notion of informational complexity which is related to SM complexity and has nice direct sum properties. This notion is used as a tool to prove the above results; it appears to be quite powerful and may be of independent interest.

- February 16 - Sariel Har-Peled talked about extracting randomness.

  Extracting randomness.
  We will survey classical results about how to extract randomness from a biased coin. Consider a coin that comes up head with probability $p$, we will show that given $n$ flips of this coin,
  (a) one can extract $(1-\delta)n H(p)$ unbiased coin flips from it,
  (b) one can not extract more $n^*H(p)$ from such a sequence.
  Here $H(p)$ is the entropy for a biased coin with probability $p$ for head.

We present a new method for proving strong lower bounds in communication complexity. This method is based on the notion of the conditional information complexity, of a function which is the minimum amount of information about the inputs that has to be revealed by a communication protocol for the function. While conditional information complexity is a lower bound on communication complexity, we show that it also admits a direct sum theorem. Direct sum decomposition reduces our task to that of proving conditional information complexity lower bounds for simple problems (such as the AND of two bits). For the latter, we develop novel techniques based on Hellinger distance and its generalizations. Our paradigm leads to two main results: (1) An improved lower bound for the multi-party set-disjointness problem in the general communication complexity model, and a nearly optimal lower bound in the one-way communication model. As a consequence, we show that for any real \( k > 2 \), approximating the \( k \)-th moment in the data stream model requires essentially \( (n^{(1-2/k)}) \) space; this resolves a conjecture of Alon et al. (J. Comput. System Sci. 58(1) (1999) 137). (2) A lower bound for the \( L_p \) approximation problem in the general communication model: this solves an open problem of Saks and Sun (in: Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC), 2002, pp. 360-369). As a consequence, we show that for \( p > 2 \), approximating the \( L_p \) norm to within a factor of \( n^p \) in the data stream model with constant number of passes requires \( (n^{(1-4/2-p)}) \) space.

- March 2 - Alina and Mike talked about their research problems and we discussed about the problem session.
- March 9 - Chandra Chekuri talked about graph sparsification.

**Graph Sparsification.**  
Chandra Chekuri

We discuss several results which give a sparse (approximate) representation of the cuts of an undirected graph. The goal is to highlight the statement of a recent remarkable result of Batson, Spielman and Srivastava on the existence of linear sized sparsifiers with very strong properties. We hope to give some details of the not so recent result of Benczur and Karger that shows the following. For any undirected graph \( G \) there is a (weighted) graph \( H \) on the same vertex set with \( O(n \log n/\epsilon^2) \) edges such that \( H \) approximates all the cuts in \( G \) to within a \((1+\epsilon)\) factor. A recent application to finding perfect matchings in \( d \)-regular graphs by Goel and Khanna will also be mentioned.

- March 16 - Sungjin Im talked about scheduling policies in the broadcast setting for several objective function, which is a joint work with Chandra Chekuri and Ben Moseley.

**Longest Wait First for Broadcast Scheduling.**  
Sungjin Im

We consider online algorithms for broadcast scheduling. In the pull-based broadcast model there are \( n \) unit-sized pages of information at a server and requests arrive online for pages. When the server transmits a page \( p \), all outstanding requests for that page are satisfied. There is a lower bound of \( (n) \) on the competitiveness of online algorithms to minimize average flow-time; therefore we consider resource augmentation analysis in which the online algorithm is given extra speedover the adversary. The longest-wait-first (LWF) algorithm is a natural algorithm that has been shown to have good empirical performance. Edmonds and Pruhs showed that LWF is \( 6 \) speed \( O(1) \)-competitive using a very complex analysis; they also showed that LWF is not \( O(1) \)-competitive with less than 1.618-speed. In this paper we make several contributions to the analysis of LWF and broadcast scheduling: (1) An intuitive and easy to understand analysis of LWF that shows that it is \( O(1/3) \) competitive for average flow-time with \( 4 \)-speed. (2)LWF is \( O(1/3) \) competitive for average flow-time with \( 3.4 \)-speed. We use our insights to prove that a natural extension of LWF is \( O(1) \) speed \( O(1) \)-competitive for more general objective functions such as aver-age delay-factor and \( L_k \) norms of delay-factor (for fixed \( k \)). These metrics generalize average flow-time and \( L_k \) norms of flowtime respectively and ours are the first non-trivial results for these objective functions.

- March 23 - Cora Borradaile will describe her work on approximation algorithms for planar graphs.

**A guide to polynomial-time approximation schemes for connectivity problems in planar graphs.**  
Glencora Borradaile, University of Waterloo

We present a framework for designing polynomial-time approximation schemes for network design problems such as Steiner tree and 2-edge connectivity in planar graphs. For a fixed \( \epsilon \), a polynomial-time approximation scheme finds, in polynomial time, a solution whose value is within \( 1+\epsilon \) of the optimal solution. In this talk I will overview the framework and discuss how to use it to solve a variety of connectivity problems.

- March 30 - Kyle Fox talked about a constructive proof of the General Lovasz Local Lemma.

**A constructive proof of the general Lovasz Local Lemma by Robin A. Moser and Gábor Tardos**  
Kyle Fox

Lovász's Local Lemma is often used to prove the existence of certain structures by means of the probabilistic method. Unfortunately, the local lemma is non-constructive and generally does not provide an efficient way to find the structures it guarantees to exist. We present a new result of Moser and Tardos that gives a simple randomized algorithm for constructively finding these structures. This new result applies to nearly all known applications of the general Local Lemma.

- April 6 - Jeff Erickson.
### Planar maximum flows and parametric shortest paths

Jeff Erickson

I’ll describe a reformulation of the O(n log n)-time planar maximum flow algorithm of Borradaile and Klein *SODA 2006, to appear in JACM* as a certain parametric shortest path problem in the dual graph. In this formulation, the algorithm maintains a shortest-path tree as a kinetic data structure, similar to the algorithm of Chambers and Cabello *SODA 2008*. This dual formulation leads to a simpler and tighter analysis of the algorithm. If there’s time, I’ll also discuss some potential generalizations, with pretty pictures but no theorems.

### On Computing Quadratic-Size Kernels for the Feedback Vertex Set Problem

April 13 - Md. Abul Hassan Samee.

**Fixed parameter tractable algorithms are often found to be quite effective in practice to deal with NP-Hard optimization problems. Computation of ‘small’ kernels is one of the possible first steps towards the design of efficient fixed parameter tractable algorithms. In this talk, I will present a recent result by Thomasse *SODA 2009* on computing quadratic-size kernels for the Feedback Vertex Set problem.**

### A Zero-One Law for Deterministic 2-Party Secure Computation

April 20 - Mike Rosulek.

**Mike Rosulek**

In two-party secure computation, two mutually distrusting parties engage in a protocol to achieve some computational task (e.g., computing a function of joint inputs, generating shared randomness, playing poker). In the Universal Composition (UC) security framework, the protocol is deemed secure if it achieves the “same effect” as if the two parties had directly interacted with a completely trusted third party who performs the task on their behalf. In this framework, cryptographic tasks are completely specified by the code of such an interactive trusted third party.

In the UC framework, it was known that almost all interesting tasks do not have secure protocols. We call the tasks that do have secure protocols the “trivial” tasks. Many other tasks were known to be complete; that is, if we allow protocols to access a trusted party for that task, then there exist secure protocols for every other task. In this work, we show a surprising “zero-one” result about the complexity of two-party computation tasks. Namely, every deterministic task is either trivial or complete – there are no intermediate levels of “cryptographic complexity”.

One aspect of the result that I will highlight is how to take an arbitrary, possibly interactive task and use it to obtain a secure protocol for another arbitrary, possibly interactive task. Previous works classifying 2-party secure computation were almost entirely restricted to non-interactive tasks. Our approach is to model interactive tasks as finite state machines and develop automata-theoretic tools to reason about their cryptographic potential. I will also discuss how the zero-one result breaks down when considering even larger classes of tasks.

A joint work with Hemanta Maji and Manoj Prabhakaran, currently in submission

### UFP in Paths and Trees and Column-Restricted Packing Integer Programs

April 27 - Alina Ene.

**Alina Ene**

In UFP we are given an edge-capacitated graph $G=(V,E)$ and $k$ request pairs $R_1, \ldots, R_k$ where each $R_i$ consists of a source-destination pair $(s_i, t_i)$, a demand $d_i$ and a weight $w_i$. The goal is to find a maximum weight subset of requests that can be routed unsplittably in $G$. Inspired by the recent work of Bansal et al. on UFP on a path, we consider UFP on paths as well as trees. I will present a simple $O(\log n)$ approximation for UFP on trees when all weights are identical; this yields an $O(\log^* n)$ approximation for the weighted case. These are the first non-trivial approximations for UFP on trees. I will also touch upon an LP relaxation for UFP on paths that has an integrality gap of $O(\log^* n)$; previously there was no relaxation with $O(n)$ gap.

A joint work with Chandra Chekuri and Nitish Korula

### May 4 - Amir Nayyeri gave a practice talk for STOC 09.

**May 4 - Amir Nayyeri**
Homology Flows, Cohomology Cuts
Amir Nayyeri

We describe the first algorithms to compute Maximum flows in surface-embedded graphs in near-linear time. Given an undirected graph embedded on an orientable surface of genus $g$, with two specified vertices $s$ and $t$, our algorithm computes a maximum $(s,t)$-flow in in $O(g^7 n \log^2 n \log^2 C)$ time for integer capacities that sum to $C$, or in $(g \log n)O(g) n$ time for real capacities. Except for the special case of planar graphs, for which an $O(n \log n)$-time algorithm has been known for 20 years, the best previous time bounds for maximum flows in surface-embedded graphs follow from algorithms for general sparse graphs. Our key insight is to optimize the relative homology class of the flow, rather than directly optimizing the flow itself. A dual formulation of our algorithm computes the minimum-cost cycle or circulation in a given (real or integer) homology class.

A joint work with Erin W. Chambers and Jeff Erickson, and it will appear in STOC ’09.

Suggested papers/results

We plan to cover Information theory and its applications to proving lowerbounds for a few weeks and to shift to important theoretical results achieved last year (or recently.) Please add some papers here!

Information Theory and its Applications

- "Informational Complexity" for communication complexity lowerbounds

- Sampling lower bounds
  - Z. Bar-Yossef, Sampling Lower Bounds via Information Theory, STOC 2003 (slides)

- Cryptography related topics
  - L. Csirmaz, On the impossibility of graph secret sharing, Cryptology 2009

- Etc
  - R. Cilibrasi and P. Vitanyi, The Google Similarity Distance, IEEE Trans. Knowledge and Data Engineering, 2007. (Also see this)

Important Theoretical Achievements

Previous Semesters

- Fall 2008 Theory Seminar
- Spring 2008 Theory Seminar
- Fall 2007 Theory Seminar
- Spring 2007 Theory Seminar
- Fall 2006 Theory Seminar