Program

Wednesday November 22, 2017

Morning - Beit Hatefutzot, Zeevi Auditorium

Afternoon - Senate building, Yaglom Auditorium

8:30 - 9:00 **Gathering and Registration**

9:00 - 9:30 **Opening Remarks -** Adi Rosén, Sébastien Linden (French Embassy), Fadil Salih (Ministry of Science, Technology and Space)

9:30 - 10:10 **Magnus Halldórsson -** *Simple local algorithms for large independent sets*

"Simple" algorithms are valuable not only because they are more likely to be actually implemented, less resource consuming, and easier to understand. They also tend to embody features that allow them to be applied in different types of constrained models of computation, e.g., parallel, online, streaming, local, external memory, distributed. Even in yet-to-be-invented models of the future.

We ponder such pan-model algorithmics in the context of the max independent set problem in graphs, reporting on some very recent results (joint with various people) along with some historical nuggets.

10:10 - 10:40 **Talya Eden -** *Testing bounded arboricity*

Graphs with bounded arboricity includes, among others, bounded-degree graphs, all minor-closed graph classes (e.g. planar graphs, graphs with bounded treewidth) and randomly generated preferential attachment graphs. These family of graphs have been studied extensively in the past, in particular since for many problems they allow for much more efficient algorithms and/or better approximation ratios.

In this talk I will present a tolerant tester that distinguishes between graphs that are ϵ-close to having arboricity at most α and graphs that are $c \cdot \epsilon$ far from having arboricity at most 3α.

10:40 - 11:00 *Coffee break*

11:00 - 11:40 **Daniel Deutch -** *Explaining data-centric computation*

Data-intensive systems effectively rule our lives. Highly complex data analysis is employed in smart cities, in scientific experiments, to support medical decisions, for marketing, and in many other contexts. An artifact of this great progress is that decisions and actions, possibly with crucial effects, are often being made through complex computation, whose logic is usually not publicly available. This raises significant concerns of various flavors, such as: is a result based on private data? If so, was the data used only for legitimate purposes? If a result is of high importance, is it based on highly reliable data? Is the underlying computation correct?

These and similar concerns are all related to the provenance of data. Provenance intuitively explains a computation result by capturing the way in which different parts of the data are used, combined and manipulated by a query or an application. I will explain our research on provenance solutions for data analytics and the way it addresses questions such as the above.

11:40 - 12:20 **Dan Feldman -** *Secure Search on the Cloud via Coresets*

Secure Search is the problem of retrieving from a database table (or any unsorted array) the records matching specified attributes, as in SQL SELECT queries, but where the database and the query are encrypted. Secure search has been the leading example for practical applications of Fully Homomorphic Encryption (FHE) already in Gentry's seminal work in 2009; however, to the best of our knowledge all state-of-the-art secure search algorithms to date are realized by a polynomial of degree $\Omega(m)$ for m the number of records, which is typically too slow in practice even for moderate size m.

In this work we present the first algorithm for secure search that is realized by a polynomial of degree polynomial in $\log m$. We implemented our algorithm in an open source library based on HELib implementation for the BGV's FHE scheme, and ran experiments on Amazon's EC2 cloud. Our experiments show that we can retrieve the first match in a database of millions of entries in less than an hour; moreover, for a bounded number of matches (e.g., up to 40 matches) we can retrieve all matches in a rate of billions of entries per machine per minute.

We achieve our results via a novel paradigm that forges a link between cryptography and modern data reduction techniques known as coresets and sketches that turn very large databases into very small ones. The key idea is that even if we don't know how to efficiently answer a query securely on the cloud, it may suffice to compute only its corresponding coreset. Since the coreset is small, the client can quickly decode the desired answer after decrypting the coreset. As a central tool we design a novel sketch that returns the first strictly-positive entry in a (not necessarily sparse) array of non-negative reals; this sketch may be of independent interest.

Joint work with Adi Akavia and Hayim Shaul

12:30 - 14:00 *Lunch* (Please see our list for a selection of restaurants on campus.)

14:00 - 14:40 **Boaz Patt-Shamir -** *The space complexity of packet routing on trees*

We study the size of buffer storage required by lossless packet routing algorithms. We assume that packet injection into the network is bounded such that in any window of t time units, at most $t\rho + \sigma$ packets are injected, for some given parameters $\sigma \geq 0$ and $0 \leq \rho \leq 1$. (This is the adversary of the Adversarial Queuing Theory.)

We give an overview of a sequence of results about the buffer space required to avoid overflows in the special case that the network is a tree, and all packets are destined for the root. In particular, we give (1) an extremely simple centralized algorithm that requires $O(\sigma+\rho)$ buffers (which is optimal), (2) a simple, but unintuitive local algorithm that requires $O(\sigma + \log n)$ buffers (which is the best possible for local algorithms), and (3) an algorithm with locality $O(\log n)$ that requires the optimal $O(\sigma+\rho)$ buffer space. (An algorithm is said to have locality d if an action at a node depends only on the current state of nodes at distance at most d.) We also prove a few lower bounds on the storage requirement in case of other topologies.

Joint work with Avery Miller (U. Manitoba) and Will Rosenbaum (TAU).

14:40 - 15:20 **Allan Borodin -** *Online bipartite matching revisited*

The seminal Karp, Vazirani and Vazirani (KVV) results on biparite matching provide a definitive analysis of the competitive ratio for deterministic and randomized online algorithms for bipartite matching. Namely, that no deterministic (respectively, randomized) online algorithm can provide a better ratio than 1/2 (resp, $1 - 1/e$). These results are with respect to the traditional one pass online model. We revisit this problem in the context of more permissive models that extend the online model. We compare results for bipartite matching with similar results for max-sat.

15:20 - 16:00 **Amos Korman -** *Sequential Computation without Feedback*

We introduce the dependent doors problem as an abstraction for situations in which one must perform a sequence of possibly dependent decisions, without receiving feedback information on the effectiveness of previously made actions. Informally, the problem considers a set of d doors that are initially closed, and the aim is to open all of them as fast as possible. To open a door, the algorithm knocks on it and it might open or not according to some probability distribution. This distribution may depend on which other doors are currently open, as well as on which other doors were open during each of the previous knocks on that door. The algorithm aims to minimize the expected time until all doors open. Crucially, it must act at any time without knowing whether or which other doors have already opened. In this work, we focus on scenarios where dependencies between doors are both positively correlated and acyclic.

The fundamental distribution of a door describes the probability it opens in the best of conditions (with respect to other doors being open or closed). We show that if in two configurations of d doors corresponding doors share the same fundamental distribution, then these configurations have the same optimal running time up to a universal constant, no matter what are the dependencies between doors and what are the distributions. We also identify algorithms that are optimal up to a universal constant factor. For the case in which all doors share the same fundamental distribution we additionally provide a simpler algorithm, and a formula to calculate its running time. We furthermore analyse the price of lacking feedback for several configurations governed by standard fundamental distributions. In particular, we show that the price is logarithmic in d for memoryless doors, but can potentially grow to be linear in d for other distributions.

We then turn our attention to investigate precise bounds. Even for the case of two doors, identifying the optimal sequence is an intriguing combinatorial question. Here, we study the case of two cascading memoryless doors. That is, the first door opens on each knock independently with probability p_1. The second door can only open if the first door is open, in which case it will open on each knock independently with probability p_2. We solve this problem almost completely by identifying algorithms that are optimal up to an additive term of 1.

This talk is based on a joint work with Yoav Rodeh, that appeared in ICALP 2017.

16:00- 16:30 *Coffee break*

16:30 - 17:10 **Avi Cohen -** *Formation Games in Social Networks*

The Preferential Attachment model is one of the most commonly used models for describing the evolution of real life networks in general, and social networks in particular, and there have been many empirical observations that justify the use of Preferential Attachment in the social networks framework. Yet, the reason for the suitability of Preferential Attachment is still unclear. In this talk we address this question by showing that the Preferential Attachment rule naturally emerges in the context of evolutionary network formation, as the unique Nash equilibrium of a simple social network game. In this game, each node aims at maximizing its degree in the future, representing its social capital in the “society” formed by the nodes and their connections. The proof exploits new connections between Preferential Attachment, random walks, and Young’s Lattice.

Joint work with Chen Avin, Pierre Fraigniaud, Zvi Lotker and David Peleg.

17:10 - 17:50 **Claire Mathieu -** *Online k-compaction*

Given, at each time $t = 1, 2, \ldots n$, a new file of length $l(t)$ and a read rate $r(t)$, an online k-compaction algorithm must maintain a collection of at most k files, choosing (at each time t, without knowing future inputs) some of the files to merge into one, thereby incurring a merge cost equal to the total length of the merged files and a read cost equal to the read rate $r(t)$ times the number of files present at time t. The goal is to minimize the total cost over time. k-compaction algorithms are a key component of log-structured merge trees, the file-based data structure underlying NoSQL databases such as Accumulo, Bigtable, Cassandra, HBase, and others.

We initiate the theoretical study of k-compaction algorithms. We formalize the problem, consider worst-case, average-case and competitive analysis (per-instance optimality), and propose new algorithms that are optimal according to these metrics.

This is joint work with Carl Staelin, Neal E. Young, and Arman Yousefi.

20:00 - **Workshop Dinner -** *at Maganda, 26, Rabbi Meir St.*

Thursday November 23, 2017

All day - Porter Building, Hall 101

9:00 - 9:40 **Benny Applebaum -** *Exponentially-Hard gap-CSP and local PRG via Local Hardcore Functions*

The gap-ETH assumption (Dinur 2016; Manurangsi and Raghavendra 2016) asserts that it is exponentially-hard to distinguish between a satisfiable 3-CNF formula and a 3-CNF formula which is at most 0.99-satisfiable. We show that this assumption follows from the exponential hardness of finding a satisfying assignment for smooth 3-CNFs. Here smoothness means that the number of satisfying assignments is not much smaller than the number of Òalmost-satisfyingÓ assignments. We further show that the latter (Òsmooth-ETHÓ) assumption follows from the exponential hardness of solving constraint satisfaction problems over well-studied distributions, and, more generally, from the existence of any exponentially-hard locally-computable one-way function. This confirms a conjecture of Dinur (ECCC 2016). We also prove an analogous result in the cryptographic setting. Namely, we show that the existence of exponentially-hard locally-computable pseudorandom generator with linear stretch (el-PRG) follows from the existence of an exponentially-hard locally-computable Òalmost regularÓ one-way functions. None of the above assumptions (gap-ETH and el-PRG) was previously known to follow from the hardness of a search problem. Our results are based on a new construction of general (GL-type) hardcore functions that, for any exponentially-hard one-way function, output linearly many hardcore bits, can be locally computed, and consume only a linear amount of random bits. We also show that such hardcore functions have several other useful applications in cryptography and complexity theory.

9:40 - 10:20 **Amnon Ta-Shma -** *Almost Optimal eps bias*

The question of finding an epsilon-biased set with close to optimal support size, or, equivalently, finding an explicit binary code with distance $1/2 - \epsilon$ and rate close to the Gilbert-Varshamov bound, attracted a lot of attention in recent decades. In this work we solve the problem almost optimally and show an explicit epsilon-biased set over k bits with support size $O(k/epsilon^{2+o(1)})$. This improves upon all previous explicit constructions which were in the order of k^2/ϵ^2, k/ϵ^3 or $(k/\epsilon^2)^{5/4}$. The result is close to the Gilbert-Varshamov bound which is $O(k/\epsilon^2)$ and the lower bound which is $Omega(k/\epsilon^2 log(1/\epsilon))$.

The main technical tool we use is bias amplification with the s-wide replacement product. The sum of two independent samples from a biased set is ϵ^2 biased. Rozenman and Wigderson showed how to amplify the bias more economically by choosing two samples with an expander. Based on that they suggested a recursive construction that achieves sample size $O(k/\epsilon^4)$. We show that amplification with a long random walk over the s-wide replacement product reduces the bias almost optimally.

10:20 -11:00 **Pascal Koiran -** *An update on the fg+1 problem for Newton polygons*

Let f, g be bivariate polynomials with at most t monomials each. The product fgcan have up to t^2 monomials, but only $2t$ of them can appear as vertices of the Newton polygon of fg. If we add a constant (say, the constant 1) to this product, the determination of the maximum number of vertices of the corresponding Newton polygon becomes quite difficult. This is due to possible cancellations with the constant term of fg (which may expose some monomials from the interior of the Newton polygon of fg). This problem is motivated by a connection to an algebraic version of the P versus NP problem, as explained in my paper on the "tau-conjecture for Newton polygons" (joint work with Natacha Portier, Sébastien Tavenas and Stéphan Thomassé). In this talk I will present some connections with problems from combinatorial geometry, and some results obtained by William Aufort for his Master's degree. His internship report is available at: http://perso.ens-lyon.fr/pascal.koiran/Publis/aufort.pdf.

10:00 - 11:30 *Coffee break*

11:30 - 12:10 **Ami Paz -** *Quadratic and Near-Quadratic Lower Bounds for the CONGEST mode*

We present the first super-linear lower bounds for natural graph problems in the CONGEST model, answering a long-standing open question in the field of distributed graph algorithms. Specifically, we show that any exact computation of a minimum vertex cover or a maximum independent set requires a near-quadratic number of rounds in the CONGEST model, as well as any algorithm for computing the chromatic number of the graph. We further show that such strong lower bounds are not limited to NP-hard problems, by showing two simple graph problems in P which require a quadratic and near-quadratic number of rounds. Finally, we address the problem of computing an exact solution to weighted all-pairs-shortest-paths (APSP), which arguably may be considered as a candidate for having a super-linear lower bound. We show a simple linear lower bound for this problem, which implies a separation between the weighted and unweighted cases, since the latter is known to have a sub-linear complexity. We also formally prove that the standard Alice-Bob framework is incapable of providing a super-linear lower bound for exact weighted APSP, whose complexity remains an intriguing open question.

Based on a joint work with Keren Censor-Hillel and Seri Khoury.

12:10 - 12:50 **Guy Even -** *Faster and Simpler Distributed CONGEST-Algorithms for Testing and Correcting Graph Properties*

We consider the following problem introduced by [Censor-Hillel et al., DISC 2016]. Design a distributed algorithm (called an ϵ-tester) that tests whether the network over which the algorithm is running satisfies a given property (e.g., acyclic, bipartite) or is ϵ-far from satisfying the property. If the network satisfies the property, then all processors must accept. If the network is ϵ-far from satisfying the property, then (with probability at least 2/3) at least one processor must reject. Being ϵ-far from a property means that at least $\epsilon \cdot |E|$ edges need to be deleted or inserted to satisfy the property.

Suppose we have an ϵ-tester that runs in $O(Diameter)$ rounds. We show how to transform this tester to an ϵ-tester that runs in $O((\log |V|)/\epsilon))$ rounds. Since cycle-freeness and bipartiteness are easily tested in $O(Diameter)$ rounds, we obtain ϵ-testers for these properties with a logarithmic number of rounds.

Moreover, for cycle-freeness, we obtain a *corrector* of the graph that locally corrects the graph so that the corrected graph is acyclic. The corrector deletes at most $\epsilon \cdot |E| + distance(G, P)$ edges and requires $O((\log |V|)/\epsilon))$ rounds.

Joint work with Reut Levi and Moti Medina.

12:50 - 14:00 *Lunch* (Please see our list for a selection of restaurants on campus.)

14:00 - 14:40 **Miklos Santha -** *On the Polynomial Parity Argument complexity of the Combinatorial Nullstellensatz*

The complexity class PPA consists of NP-search problems which are reducible to the parity principle in undirected graphs. It contains a wide variety of interesting problems from graph theory, combinatorics, algebra and number theory, but only a few of these are known to be complete in the class. Before this work, the known complete problems were all discretizations or combinatorial analogues of topological fixed point theorems.

In this talk we prove the PPA-completeness of two problems of radically different style. They are PPA-Circuit CNSS and PPA-Circuit Chevalley, related respectively to the Combinatorial Nullstellensatz and to the Chevalley-Warning Theorem over the two elements field. The input of these problems contain PPA-circuits which are arithmetic circuits with special symmetric properties that assure that the polynomials computed by them have always an even number of zeros. In the proof of the result we relate the multilinear degree of the polynomials to the parity of the maximal parse subcircuits that compute monomials with maximal multilinear degree, and we show that the maximal parse subcircuits of a PPA-circuit can be paired in polynomial time.

14:40 - 15:20 **Irit Dinur -** *Unique games conjecture - recent progress*

The unique games conjecture [Khot 2002] stands at the heart of the study of tight inapproximability results. The conjecture says that a certain gap constraint satisfaction problem is NP hard. A long line of research shows that the correctness of this conjecture paints a beautiful landscape in which one meta algorithm provides the best approximation for many different optimization problems. However, until recently, most of the progress on the unique games conjecture has been algorithmic, causing some experts to doubt its correctness.

I will describe recent works that point in the other direction, namely towards the correctness of the unique games conjecture.

Based on joint works with Subhash Khot, Guy Kindler, Dor Minzer, Muli Safra.

15:20 - 15:50 *Coffee break*

15:50 - 16:30 **Haim Kaplan -** *Voronoi diagrams on planar graphs, and computing the diameter in deterministic $\tilde{O}(n^{5/3})$ time*

We present a deterministic algorithm that computes the diameter of a directed planar graph with real arc lengths in $\tilde{O}(n^{5/3})$ time. This improves the recent breakthrough result of Cabello (SODA'17), both by improving the running time (from $\tilde{O}(n^{11/6})$), and by using a deterministic algorithm. It is in fact the first truly subquadratic *deterministic* algorithm for this problem. Our algorithm follows the general high-level approach of Cabello, but differs in the way it is implemented. Our main technical contribution is an explicit and efficient construction of additively weighted Voronoi diagrams on planar graphs (under the assumption that all Voronoi sites lie on a constant number of faces). The idea of using Voronoi diagrams is also borrowed from Cabello, except that he uses abstract Voronoi diagrams (as a black box), which makes the implementation more involved, more expensive, and randomized. Our explicit construction avoids these issues, and leads to the improved (and deterministic) running time.

Joint work with Pawel Gawrychowski, Shay Mozes, Micha Sharir, and Oren Weimann.

16:30 - 17:10 **Pierre Fraigniaud -** *Distributed Testing of Excluded Subgraphs*

Given a graph H, a graph G is H-free if it does not contain H as a subgraph. Distributed decision refers to the task in which, given a boolean predicate, a set of computing entities have to decide whether their inputs collectively satisfy the predicate. If yes, then all entities must accept, otherwise at least one entity must reject. This talk will survey the recent results related to deciding H-freeness obtained in the context of Distributed Property Testing in the CONGEST model. This latter model limits to $O(\log n)$ bits the amount of information to be exchanged at each round between neighbors, and Distributed Property Testing refers to the task where it is only required to distinguish inputs satisfying the predicate from inputs that are far from satisfying the predicate, with constant (arbitrarily large) probability.

17:15 - *Farewell drink*