Computational Complexity (10w5028)

Paul Beame (University of Washington),
Stephen Cook (University of Toronto),
Russell Impagliazzo (University of California, San Diego),
Valentine Kabanets (Simon Fraser University),
Avi Wigderson (Institute for Advanced Study, Princeton)

August 1–6, 2010

1 Overview of the Field

Computational Complexity Theory is the field that studies the inherent costs of algorithms for solving mathematical problems. Its major goal is to identify the limits of what is efficiently computable in natural computational models. Computational complexity ranges from quantum computing to determining the minimum size of circuits that compute basic mathematical functions to the foundations of cryptography and security.

Computational complexity emerged from the combination of logic, combinatorics, information theory, and operations research. It coalesced around the central problem of "P versus NP" (one of the seven open problems of the Clay Institute). While this problem remains open, the field has grown both in scope and sophistication. Currently, some of the most active research areas in computational complexity are

- the study of hardness of approximation of various optimization problems (using probabilistically checkable proofs), and the connections to coding theory,
- the study of the role of randomness in efficient computation, and explicit constructions of "random-like" combinatorial objects,
- the study of the power of various proof systems of logic, and the connections with circuit complexity and search heuristics,
- the study of the power of quantum computation.

2 Recent Developments and Open Problems

The main focus of computational complexity is to understand *efficient* computation. One of the main open problems is the famous "P versus NP" question which asks if there is an efficient algorithm for solving such problems as Satisfiability (SAT): Given a propositional formula in n variables, decide if there is a setting of the variables to the truth values (True and False) so that the formula on the assignment evaluates to True. This problem has been the driving force behind many developments in complexity theory, and continues to be such. While the exact complexity of SAT remains unknown, researchers do come up with new results that shed more light on various aspects of this fundamental problem, and its counting version (given a propositional formula,

count the number of its satisfying assignments). Some of these new developments were also reported at the workshop (see below).

Understanding the complexity of approximation problems (e.g., given a conjunction of constraints find an assignment that satisfies approximately the largest number of the constraints) is also one of the main tasks of complexity theory, and has been pursued since 1990's using the machinery of Probabilistically Checkable Proofs (PCPs). While the known PCP results imply tight inapproximability results (under the hypothesis that $P \neq NP$) for a number of approximation problems, there are a few important exceptions for which the currently known techniques seem powerless. This led to the formulation of a conjecture (Unique Games Conjecture) whose truth would imply inapproximability results for quite a few new problems. Whether this conjecture is true or false has become one of the main open problems in modern complexity theory, with no apparent consensus by the experts of what outcome is more likely. A number of discussions on this conjecture were also held at the workshop.

One of the main open problems in communication complexity is the Direct Sum Conjecture which basically says that the amount of communication needed to solve k instances of a given problem should be k times the amount of communication needed to solve a single instance of the problem. The conjecture is still open. However, some new ideas have been recently introduced that may eventually lead to the resolution of the conjecture, and already have provided some nontrivial weaker statements. These ideas are based on compressing the communication protocol between two parties, and very information-theoretic in nature. This problem has also been discussed at the workshop.

Understanding efficient computation also involves understanding the role of randomness in computation. In particular, one of the basic questions is to construct pseudorandom generators (PRGs) that are efficient algorithms stretching a short truly random input string into a much longer string that "looks" random to a given class of observers. For the sufficiently general class of observers (say, when observers are themselves arbitrary efficient algorithms), no such construction is known (and it seems to be related to our lack of lower bounds for general models of computation). However, for sufficiently restricted classes of observers, some nontrivial PRG constructions are known. Some of the recent developments in the area have involved the class of polynomial threshold functions, showing that well-known constructions are actually pseudorandom for these functions, as well as giving new constructions of PRGs for these functions. Another important class of observers is the class of small-space algorithms, and is centered around the open problem of whether every small-space randomized algorithm can be made deterministic without increasing the space usage by more than a constant factor. While the general question remains open, some special cases have been recently solved. These PRG constructions have also been an important topic of discussion at the workshop.

3 Presentation Highlights

3.1 The Unique Games Conjecture

The famous result of Cook and Levin from 1970's introduced the important notion of NP-completeness. Many natural problems were later shown to be NP-complete, and hence unlikely to be efficiently solvable unless some well-studied hard problems (such as Satisfiability of propositional formulas, deciding if a graph has a large clique, and deciding if a graph is 3-colorable) are also efficiently solvable (in deterministic polynomial time).

The NP-completeness theory is applicable to the case of *exact* algorithms, which are required to compute the exact optimal solutions to a given NP-problem. The subsequently developed theory of Probabilistically Checkable Proofs managed to extend NP-hardness results to the case of *approximation* algorithms, where an algorithm is only required to obtain an approximately optimal solution (to within a certain approximation factor). A large number of optimization problems were shown NP-hard to approximate (to within certain approximation factors). This theory, developed starting from 1990's, has been extremely successful in classifying a large number of approximation problems, often establishing tight approximation factors (where the problem is NP-hard to approximate with a better factor, and on the other hand, there is an efficient algorithm approximating the problem with slightly worse factor).

Despite this success, some important approximation problems still escape such tight classification. One approach to deal with this was formulated by Khot, and is known as the "Unique Games Conjecture". The conjecture involves systems of linear equations over finite fields, where each equation is over 2 variables.

Roughly, the conjecture states that it is NP-hard to distinguish the following two cases: (1) a system of such linear equations over a finite field (of appropriate size) where there is a vector that satisfies "many" equations, and (2) a system of such linear equations where no vector can satisfy more than "few" equations.

Lifting the restriction that each equation be over 2 variables to allow 3 variables per equation yields a well-known NP-hardness result due to Håstad. Khot conjectured that the same result is true even for 2 variables per equations, but this bold conjecture is still open.

At the workshop, there were a few discussions regarding the Unique Games Conjecture, attacking it from both sides: trying to prove it (see the talk by Moshkovitz), and trying to disprove it (see the talk by Steurer). There was also a talk describing interesting connections of the conjecture to the problem of expansion in graphs (see the talk by Raghavendra).

Below we list the abstracts of the relevant presentations.

P. Raghavendra *Approximating Graph Expansion: Connections, Algorithms and Reductions*

Approximating edge expansion, equivalently finding sparse cuts in graphs is a fundamental problem in combinatorial optimization that has received considerable attention in both theory and practice. Yet, the complexity of approximating edge expansion in graphs is poorly understood. Particularly, worse is the understanding of the approximability of the expansion of small sets in graphs. More formally, current algorithmic or hardness results do not settle the approximability of the following problem: Given a regular graph G and a very small constant c, find a set S of cn vertices in the graph such that minimum number of edges cross the set S. Recently it was shown that the complexity of this problem is closely tied to the Unique Games Conjecture. Furthermore, we show that the hardness of this problem is a natural assumption that generalizes the unique games conjecture, and yields hardness for problems like Balanced Separator and Minimum Linear arrangement.

D. Moshkovitz *Hardness of Approximately Solving Linear Equations Over Reals*

We consider the problem of approximately solving a system of homogeneous linear equations over reals, where each equation contains at most three variables. Since the all-zero assignment always satisfies all the equations exactly, we restrict the assignments to be "non-trivial". Here is an informal statement of our result: it is NP-hard to distinguish whether there is a non-trivial assignment that satisfies $1 - \delta$ fraction of the equations or every non-trivial assignment fails to satisfy a constant fraction of the equations with a "margin" of $\Omega(\sqrt{\delta})$. Unlike the well-studied case of linear equations over finite fields, for equations over reals, the best approximation algorithm known (SDP-based) is the same no matter whether the number of variables per equation is two or three.

Our result is motivated by the following potential approach to proving The Unique Games Conjecture:

1. Prove the NP-hardness of solving approximate linear equations over reals, for the case of three variables per equation (we prove this result).
2. Prove the NP-hardness of the problem for the case of two variables per equation, possibly via a reduction from the three variable case.
3. Prove the Unique Games Conjecture.

An interesting feature of our result is that it shows NP-hardness result that matches the performance of a non-trivial SDP-algorithm. Indeed, the Unique Games Conjecture predicts that an SDP-based algorithm is optimal for a huge class of problems (e.g. all CSPs by Raghavendra's result). (Joint work with Subhash Khot)

D. Steurer *Subexponential Algorithms for Unique Games and Related Problems*

We give a subexponential time approximation algorithm for the Unique Games problem: Given a Unique Games instance with optimal value $1 - \epsilon^6$ and alphabet size k, our algorithm finds in time $exp(k \cdot n^{\epsilon})$ a solution of value $1 - \epsilon$.

We also obtain subexponential algorithms with similar approximation guarantees for Small-Set Expansion and Multi Cut. For Max Cut, Sparsest Cut and Vertex Cover, our techniques lead to subexponential algorithms with improved approximation guarantees on subclasses of instances. Khot's Unique Games Conjecture (UGC) states that it is NP-hard to achieve approximation guarantees such as ours for Unique Games. While our result stops short of refuting the UGC, it does suggest that Unique Games is significantly easier than NP-hard problems such as Max 3-SAT, Label Cover and more, that are believed not to have subexponential algorithms achieving a non-trivial approximation ratio.

The main component in our algorithms is a new kind of graph decomposition that may have other applications: We show that by changing an ϵ fraction of its edges, any regular graph on n vertices can be broken into disjoint parts such that the stochastic adjacency matrix of each part has at most n^{ϵ} eigenvalues larger than $1 - \epsilon^6$. (Joint work with Sanjeev Arora and Boaz Barak.)

3.2 Complexity of Counting

In the pioneering work of from the late 1970s, Valiant studied the complexity of counting problems (such as #SAT: Given a propositional formula, compute the number of its satisfying assignments), and proved the computing the Permanent of a 0-1 matrix is a complete problem for this class of counting problems. It is widely believed that there is no efficient algorithm to solve such counting problems. However, in certain special cases (for restricted counting problems), it was observed that efficient algorithms are possible, and are essentially some determinant computations for appropriate matrices. Given such fairly non-intuitive efficient algorithms, one may ask what other counting problems can be efficiently solved via some kind of reduction to a problem solvable by these known algorithms. Recently, Valiant proposed a way to formalize such reductions, and defined the notion of "Holographic algorithms". Using this approach, he was able to solve efficiently some new counting problems which were not known efficiently solvable before. Naturally it is an important question to understand the limitations of this "holographic method", and this has been tackled in a number of recent papers by Valiant and other researchers.

At the workshop, Valiant gave a talk on the current status of holographic algorithms.

L. Valiant *Holographic Algorithms*

First we briefly review some recent dichotomy results, including some that strictly generalize constraint satisfaction problems, that showcase the power of the holographic method. We go on to define the notion of diversity for families of finite functions, and express the limitations of a class of holographic algorithms in terms of limitations on diversity. In particular, we show, by a new but very classical looking combination of counting and algebraic methods, that the class of elementary holographic algorithms, which has yielded novel polynomial time algorithms for such problems as special cases of Boolean Satisfiability, is insufficient for expressing general Boolean Satisfiability. We suggest that the question of how far this lower bound argument can be extended is of some general interest.

We go on to explore the power of nonelementary polynomial time holographic algorithms by describing such algorithms for certain parity problems for which no polynomial time algorithms were previously known. These algorithms compute the parity of the following quantities for degree three planar undirected graphs: the number of 3-colorings up to permutation of colors, the number of connected vertex covers, and the number of induced forests or feedback vertex sets. These holographic algorithms, besides being nonelementary, use bases of more than two components and thereby potentially evade the Cai-Lu Collapse Theorem.

3.3 Complexity of SAT

The problem SAT (satisfiability of a propositional formula) is one of the most famous NP-complete problems, and has been a popular problem to study for a long time. At the workshop, there were a couple of new results regarding the complexity of SAT for formulas (by Santhanam), as well as the impossibility of certain "compression" of SAT (by van Melkebeek).

D. van Melkebeek *Satisfiability Allows No Nontrivial Sparsification Unless The Polynomial-Time Hierarchy Collapses*

Consider the following two-player communication process to decide a language L: The first player holds the entire input x but is polynomially bounded; the second player is computationally unbounded but does not know any part of x; their goal is to cooperatively decide whether x belongs to L at small cost, where the cost measure is the number of bits of communication from the first player to the second player.

For any integer $d \geq 3$ and positive real ϵ we show that if satisfiability for n-variable d-CNF formulas has a protocol of cost $O(n^{d-\epsilon})$ then coNP is in NP/poly, which implies that the polynomial-time hierarchy collapses to the third level. The result even holds for conondeterministic protocols, and is tight as there exists a trivial deterministic protocol for $\epsilon = 0$. Under the hypothesis that coNP is not in NP/poly, our result implies tight lower bounds for parameters of interest in several areas, including sparsification, probabilistically checkable proofs, instance compression, and kernelization in parameterized complexity.

By reduction similar results holds for other NP-complete problems. For the vertex-cover problem on n-vertex d-regular hypergraphs the above statement holds for any integer $d \geq 2$. The case $d = 2$ implies that no nontrivial parameterized vertex deletion problem on standard graphs can have kernels consisting of $O(k^{2-\epsilon})$ edges unless coNP is in NP/poly. Kernels consisting of $O(k^2)$ edges are known for several problems in the class, including vertex cover, bounded-degree deletion, and feedback vertex set.

Our approach refines the framework developed in recent papers showing that certain parameterized languages do not have protocols of cost bounded by any polynomial in the parameter unless coNP is in NP/poly. We study parameterized problems that do have protocols of polynomial cost, and show that no polynomial cost of lower degree than the current best is achievable unless coNP is in NP/poly. In order to obtain our tight bounds we exploit a result from additive combinatorics, namely the existence of high-density subsets of the integers without nontrivial arithmetic progressions of length three. (Joint work with Holger Dell.)

R. Santhanam *New and Improved Upper Bounds for Formula Satisfiability and TQBF*

I will describe what appears to be a new technique for bounding the running time of algorithms for satisfiability, based on proving concentration versions of results about random restrictions. I will show how this gives a running time upper bound of $2^{n-\Omega(n)}$ for a simple and natural algorithm for formula satisfiability on formulae of linear length, and explain how the technique also gives a strong average-case lower bound for Parity against linear-size formulae. I will pose some questions relevant to extending this line of research to satisfiability and lower bounds for polynomial-size constant- depth circuits. If time permits, I will also mention a memoization-based technique that beats brute force search for QBF satisfiability on formulae with a bounded number of variable occurrences.

3.4 Small-space computation

One of the basic open questions in computational complexity is whether every problem in P (solvable in polynomial time) can be solved also in small (logarithmic) space. In the complexity language, the question is whether $P = LOGSPACE$. Cook has recently launched a project trying to separate these two classes, by considering a very special problem in P (tree-evaluation problem), and trying to show that it cannot be solved by logspace-bounded algorithms. The final result seems still distant, but there has been some partial progress, which was reported by Wehr (a student of Cook).

D. Wehr *A lower bound for a restricted model of log-space computation*

I'll show how to solve a problem posed in [Gal,Koucky,McKenzie "Incremental branching programs" 2006] regarding a restricted model of small-space computation, tailored for solving the P-complete GEN problem. They define two variants of incremental branching programs, the syntactic variant defined by a restriction on the graph-theoretic paths in the program, and the more-general semantic variant in which the same restriction is enforced only on the consistent paths (those that are followed by at least one input). They used a lower bound for monotone circuits to show that exponential size is required for the syntactic variant, but left open the problem of superpolynomial lower bounds for the semantic variant. I'll give the main part of the proof of an exponential lower bound for the semantic variant; it is a generalization of a lower bound argument for a similar restricted model of computation tailored for solving the Tree Evaluation Problem, which appeared in [Braverman, Cook, McKenzie, Santhanam, Wehr "Fractional pebbling and thrifty branching programs" 09].

3.5 Oblivious computation

An oblivious algorithm is an algorithm whose "behavior" (say in terms of memory access) is independent of a given input (and so by observing the memory locations queried by the algorithm, one has no information about the input of the algorithm). There has been recently a renewed interest in efficient constructions of oblivious algorithms. At the workshop, Beame reported on the lower bound for making a given algorithm into an oblivious one.

P. Beame *Making RAMs Oblivious Requires Superlogarithmic Overhead*

We prove a time-space tradeoff lower bound of $T = \Omega(n \log(n/S) \log\log(n/S))$ for randomized oblivious branching programs to compute 1GAP, also known as the pointer jumping problem, a problem for which there is a simple deterministic time n and space $O(\log n)$ RAM (random access machine) algorithm. In a recent STOC paper, Ajtai derived simulations of general RAMs by randomized oblivious RAMs with only

a polylogarithmic factor increase in time and space. Our lower bound implies that a superlogarithmic factor increase is indeed necessary in any such simulation. (Joint work with Widad Machmouchi.)

3.6 Quantum Computation

What is the power of quantum computation? Is there some problem that can be efficiently solved on a quantum computer (even with today's technology) but cannot be efficiently solved by classical computers? Aaronson addressed this question in his talk at the workshop.

In another talk, Umans considered a related question: Is quantum computation more powerful than what is captured by a classical complexity class PH (polynomial-time hierarchy)? No answer is known at the moment. Moreover, there is not even any evidence of the advantage of quantum computation over PH in the form of some relativized (oracle) construction. Umans discussed the problem of constructing such an oracle, and relates it some other interesting questions in classical complexity.

S. Aaronson *The Computational Complexity of Linear Optics*

We propose a linear-optics experiment that might be feasible with current technology, and argue that, if the experiment succeeded, it would provide evidence that at least some nontrivial quantum computation is possible in nature. The experiment involves generating reliable single-photon states, sending the photons through a random linear-optical network, and then reliably measuring the photon number in each mode. The resources that we consider are not known or believed to be universal for quantum computation; nevertheless, we show that they would allow the solution of certain sampling and relational problems that appear to be intractable for classical computers.

Our first result says that, if there exists a polynomial-time classical algorithm that samples from the same probability distribution as our optical experiment, then $P^{\#P} = BPP^{NP}$, and hence the polynomial hierarchy collapses to the third level. Unfortunately, this assumes an extremely reliable experiment. While that could in principle be arranged using quantum error correction, the question arises of whether a noisy experiment would already have interesting complexity consequences. To address this question, we formulate a so-called "Permanent-of-Gaussians Conjecture" (PGC), which says that it is $\#P$-hard to approximate the permanent of a matrix A of independent $N(0,1)$ Gaussian entries, with high probability over A; as well as a "Permanent Anti-Concentration Conjecture" (PACC), which says that $|Per(A)| >= sqrt(n!)/poly(n)$ with high probability over A. We then show that, assuming both the PGC and the PACC, polynomial-time classical simulation even of noisy linear-optics experiments would imply a collapse of the polynomial hierarchy. (Joint work with Alex Arkhipov)

C. Umans *Pseudorandom generators and the BQP vs. PH problem*

It is a longstanding open problem to devise an oracle relative to which BQP does not lie in the Polynomial-Time Hierarchy (PH). We advance a natural conjecture about the capacity of the Nisan- Wigderson pseudo-random generator [NW94] to fool AC^0, with MAJORITY as its hard function. Our conjecture is essentially that the loss due to the hybrid argument (which is a component of the standard proof from [NW94]) can be avoided in this setting. This is a question that has been asked previously in the pseudorandomness literature [BSW03]. We then show that our conjecture implies the existence of an oracle relative to which BQP is not in the PH. This entails giving an explicit construction of unitary matrices, realizable by small quantum circuits, whose row-supports are nearly-disjoint. Our framework generalizes the setting of [Aar09], and remains a viable approach to resolving the BQP vs. PH problem after the recent proof [Aar10] that the Generalized Linial-Nisan Conjecture of [Aar09] is false. (Joint work with Bill Fefferman)

3.7 Error-correcting codes

Error-correcting codes have become a central tool and an object of study in computational complexity. At the workshop, Dvir reported on an interesting connection between the classical complexity problem (on matrix rigidity) and certain (locally self-correctable) codes. In another talk, Guruswami showed a beautiful result establishing the tight list-decodability property of random linear codes, which shows that random linear codes achieve essentially the same list-decodability parameters (up to constant factors) as random non-linear codes.

Z. Dvir *On matrix rigidity and locally self-correctable codes*

We describe a new approach for the problem of finding rigid matrices, as posed by Valiant [Val77], by connecting it to the, seemingly unrelated, problem of proving lower bounds for locally selfcorrectable

codes. This approach, if successful, could lead to a non-natural property (in the sense of Razborov and Rudich [RR97]) implying super-linear lower bounds for linear functions in the model of logarithmic-depth arithmetic circuits.

Our results are based on a lemma saying that, if the generating matrix of a locally decodable code is not rigid, then it defines a locally self-correctable code with rate close to one. Thus, showing that such codes cannot exist will prove that the generating matrix of any locally decodable code (and in particular Reed Muller codes) is rigid.

V. Guruswami *List decodability of random linear codes*

For every fixed finite field F_q, $0 < p < 1 - 1/q$, and $\epsilon > 0$, we prove that with high probability a random subspace C of F_q^n of dimension $(1-h_q(p)-\epsilon)n$ has the property that every Hamming ball of radius pn has at most $O(1/\epsilon)$ elements of C. (Here $h_q(x)$ is the q-ary entropy function.) This answers a basic open question concerning the list-decodability of linear codes, showing that a list size of $O(1/\epsilon)$ suffices to have rate within ϵ of the information-theoretic limit $1 - h_q(p)$. This matches up to constant factors the list-size achieved by general (non-linear) random codes, and gives an exponential improvement over the best previously known list-size bound of $q^{O(1/\epsilon)}$.

The main technical ingredient in our proof is a strong upper bound on the probability that m random vectors chosen from a Hamming ball centered at the origin have too many (more than $O(m)$) vectors from their linear span also belong to the ball. (Joint work with Johan Hastad (KTH) and Swastik Kopparty (MIT).)

3.8 Computational Learning

A basic task in computational learning is to "learn" an object (say, a halfspace in a high-dimensional space) by having access to possibly noisy data (say, an oracle which answers if a given point is in the halfspace or not). Some of the main approaches to learning involve the algebraic techniques based on low-degree polynomial representations of the objects one needs to learn. While these techniques were successful for some classes of objects (e.g., halfspaces), they don't seem to help with others (e.g., intersections of halfspaces).

There were two talks at the workshop that addressed this issue. Sherstov explained his negative result (saying that low-degree techniques won't help for learning an intersection of two halfspaces). Klivans showed some new results on approximately representing intersections of "regular" halfspaces (a special case of halfspaces), which in particular imply a new learning algorithm for intersections of such regular halfspaces.

A. Sherstov *Symmetrization Without Symmetries*

We prove that the intersection of two halfspaces on the n-cube cannot be sign-represented by a polynomial of degree less than $\Theta(n)$, which matches the trivial upper bound and solves an open problem due to Klivans (2002). This result shows that intersections of halfspaces are not amenable to learning by perceptron-based techniques, which have been successful in other cases (halfspaces, DNF formulas, read-once formulas). A mostly complete proof will be presented with emphasis on a key technical component, a method for symmetrizing a Boolean function f without any symmetries by averaging f over suitable sections over the n-cube.

A. Klivans *An Invariance Principle for Polytopes*

Let X be randomly chosen from $\{-1, 1\}^n$, and let Y be randomly chosen from a standard n-variate Gaussian. For any polytope P formed by the intersection of k halfspaces, we prove that $|Pr[X \in P] - Pr[Y \in P]| \leq polylog(k) \cdot \Delta$, where Δ is a parameter that is small for polytopes formed by the intersection of "regular" halfspaces (i.e., halfspaces with low influence). The novelty of our invariance principle is the polylogarithmic dependence on k. Previously, only bounds that were at least linear in k were known.

We give two important applications of our main result:

1. A bound of $polylog(k) \cdot \epsilon^{O(1)}$ on the noise sensitivity of intersections of k regular halfspaces (previous work gave bounds linear in k). This gives the first quasipolynomial-time algorithm for learning intersections of regular halfspaces.

2. The first pseudorandom generators (with polylogarithmic seed length) for regular polytopes. This gives an algorithm for approximately counting the number of solutions to a broad class of integer programs (including dense covering programs and contingency tables).

(This is joint work with Prahladh Harsha and Raghu Meka)

3.9 Communication Complexity

The communication complexity is concerned with the amount of communication (say, between two parties) needed for jointly solving a certain computational task (say, computing some function of two inputs, where one input is known to the first party, and the other input to the second party). There are various models of communication one can define, and a large number of lower and upper bound results are known. Still, a number of important questions remain open.

At the workshop, there were several discussions of different aspects of communication complexity. Regev showed a tight lower bound for a well-known problem of distinguishing two strings that are either close or far in the Hamming distance. Braverman discussed a new approach to an old problem in communication complexity (direct sum conjecture) via compressing a communication protocol so that the new compressed protocol uses the amount of communication that is closer to the information-theoretic lower bounds. Finally, Rao addressed an issue of error-correction in interactive communication.

O. Regev *Tight Bound for the Gap Hamming Distance Problem*

We consider the Gap Hamming Distance problem in communication complexity. Here, Alice receives an n-bit string x, and Bob receives an n-bit string y. They are promised that the Hamming distance between x and y is either at least $n/2 + \sqrt{n}$ or at most $n/2 - \sqrt{n}$, and their goal is to decide which is the case. The naive protocol requires n bits of communication and it was an open question whether this is optimal. This was shown in several special cases, e.g., when the communication is deterministic [Woodruff'07] or when the number of rounds of communication is limited [Indyk-Woodruff'03, Jayram-Kumar-Sivakumar'07, Brody-Chakrabarti'09, Brody-Chakrabarti-R-Vidick-deWolf'09]. Here we settle this question by showing a tight lower bound of $\Omega(n)$ on the randomized communication complexity of the problem. The bound is based on a new geometric statement regarding correlations in Gaussian space, related to a result of C. Borell from 1985, which is proven using properties of projections of sets in Gaussian space. (Partly based on a joint paper with Amit Chakrabarti.)

M. Braverman *Compression, information and direct sum for communication complexity*

We will present a tight three-way connection between three types of results related to the randomized two-party communication complexity of a problem:

1. Direct sum theorems, relating the communication complexity of computing many copies of a function to the complexity of computing one copy;

2. The information complexity of a problem, which is the smallest amount of information (as opposed to communication) the parties need to exchange to solve the problem; and

3. Compression theorems, which show how to convert two party communication protocols closer to the information-theoretically optimal bounds.

We will then use these connections along with new compression schemes to derive new results in communication complexity. Based on two joint works, one with [Boaz Barak, Xi Chen, and Anup Rao], and the second one with [Anup Rao].

A. Rao *Recovering from Maximal Errors in Interactive Communication*

We show that it is possible to encode any communication protocol between two parties so that the protocol succeeds even if a $(1/4 - \epsilon)$ fraction of all symbols transmitted by the parties are corrupted adversarially, at a cost of increasing the communication in the protocol by a constant factor (the constant depends on epsilon). No encoding can tolerate a 1/4 fraction of errors in the interactive setting, if the communication is to remain bounded in terms of the original communication of the protocol. This improves on an earlier result of Schulman, who showed how to recover when the fraction of errors is at most 1/240. (Joint work with M. Braverman)

3.10 Pseudorandom generators

Constructing pseudorandom generators is one of the basic tasks in computational complexity, and is an open problem for many models of computation. However, some progress has been made for certain restricted models.

Some recent such progress has been reported on by Zuckerman (for threshold functions), Lovett (for constant-depth modular circuits), Yehudayoff (for regular branching programs of constant width), and Pudlak (for group products). Also, Viola discussed the complexity of generating distributions of the form $h(x)$ for a random x, where h is some function from m to n bits, as well as some applications to succinct data structures and pseudorandom generators.

D. Zuckerman *Pseudorandom Generators for Polynomial Threshold Functions*

We study the natural question of constructing pseudorandom generators (PRGs) for low-degree polynomial threshold functions (PTFs). We give a PRG with seed-length $\log n/\epsilon^{O(d)}$ fooling degree d PTFs with error at most ϵ. Previously, no nontrivial constructions were known even for quadratic threshold functions and constant error ϵ. For the class of degree 1 threshold functions or halfspaces, we construct PRGs with much better dependence on the error parameter ϵ and obtain the following results.

1. A PRG with seed length $O(\log n \log(1/\epsilon))$ for $\epsilon > 1/poly(n)$.
2. A PRG with seed length $O(\log n)$ for $\epsilon > 1/poly(\log n)$. Previously, only PRGs with seed length $O(\log n \log^2(1/\epsilon)/\epsilon^2)$ were known for halfspaces. We also obtain PRGs with similar seed lengths for fooling halfspaces over the n-dimensional unit sphere.

The main theme of our constructions and analysis is the use of invariance principles to construct pseudorandom generators. We also introduce the notion of monotone read-once branching programs, which is key to improving the dependence on the error rate ϵ for halfspaces. These techniques may be of independent interest. (Joint work with R. Meka)

S. Lovett *Pseudorandom generators for $CC^0[p]$ and the Fourier spectrum of low-degree polynomials over finite fields*

In this paper we give the first construction of a pseudorandom generator with seed length $O(\log n)$, for $CC^0[p]$, the class of constant-depth circuits with unbounded fan-in MODp gates, for some prime p. More accurately, the seed length of our generator is $O(\log n)$ for any constant error $\epsilon > 0$. In fact, we obtain our generator by fooling distributions generated by low degree polynomials, over F_p, when evaluated on the Boolean cube. This result significantly extends previous constructions that either required a long seed [LVW93] or that could only fool the distribution generated by linear functions over F_p, when evaluated on the Boolean cube [LRTV09, MZ09]. Enroute of constructing our PRG, we prove two structural results for low degree polynomials over finite fields that can be of independent interest:

1. Let f be an n-variate degree d polynomial over F_p. Then, for every $\epsilon > 0$ there exists a subset S of variables of size depending only on d and ϵ, such that the total weight of the Fourier coefficients that do not involve any variable from S is at most ϵ.
2. Let f be an n-variate degree d polynomial over F_p. If the distribution of f when applied to uniform zero-one bits is ϵ-far (in statistical distance) from its distribution when applied to biased bits, then for every $\delta > 0$, f can be approximated over zero-one bits, up to error δ, by a function of a small number (depending only on ϵ, δ and d) of lower degree polynomials.

(Joint work with Partha Mukhopadhyay and Amir Shpilka.)

A. Yehudayoff *Pseudorandom generators for regular branching programs*

We give new pseudorandom generators for *regular* read-once branching programs of small width. A branching program is regular if the in-degree of every vertex in it is (0 or) 2. For every width d and length n, our pseudorandom generator uses a seed of length $O((\log(d) + \log\log(n) + \log(1/\epsilon))\log(n))$ to produce n bits that cannot be distinguished from a uniformly random string by any regular width d length n read-once branching program, except with probability ϵ. We also give a result for general read-once branching programs, in the case that there are no vertices that are reached with small probability. We show that if a (possibly non-regular) branching program of length n and width d has the property that every vertex in the program is traversed with probability at least p on a uniformly random input, then the error of the generator above is at most $2\epsilon/p^2$. (Joint work with Mark Braverman, Anup Rao, and Ran Raz)

P. Pudlak *Pseudorandom Generators for Group Products*

We will show that the pseudorandom generator introduced in [INW94] fools group products of a given finite group. The seed length is $O(\log n \log \frac{1}{\epsilon})$, where n the length of the word and ϵ is the precision. The result is equivalent to the statement that the pseudorandom generator fools read-once permutation branching programs of constant width. (Joint work with Michal Koucky and Prajakta Nimbhorkar.)

E. Viola *The complexity of distributions*

Complexity theory typically studies the complexity of computing a function $h(x) : \{0,1\}^m \to \{0,1\}^n$ of a given input x. We advocate the study of the complexity of generating the distribution $h(x)$ for uniform x, given random bits. We discuss recent work in this direction. This includes lower and upper bounds for various computational models (NC^0, decision trees, and AC^0) and the consequences of these bounds for succinct data structures and pseudorandom generators. (We expect the talk to be based on two papers "The complexity of distributions" and "Bounded-depth circuits cannot sample good codes," the latter co-authored with Shachar Lovett.)

3.11 Social choice

Economics and social choice theory are becoming the objects of study by computer scientists, who bring the computational perspective on the old issued studied by economists and social scientists. One example of such interaction between social choice theory and computer science was given at the workshop in a talk by Kindler.

G. Kindler *A Quantitative Proof of the Gibbard-Satterthwaite Theorem*

A social choice function f with n voters and q alternatives, takes as input a tuple of n full rankings of the alternatives, supposedly corresponding to the preferences of the voters, and outputs the winner alternative. We say that f is manipulable at a given voting profile if a voter who knows the rankings given by the others can change her own ranking in a way that does not reflect her true preferences, but which leads to a winner that is more favorable to her.

Gibbard and Satterthwaite proved that any social choice function which attains three or more values, and which is not a dictatorship, must be manipulable. We show a quantitative version of the theorem in the case where f is neutral, showing that f must be manipulable at a uniformly chosen voting profile with probability bounded below by (the inverse of) a polynomial in n and q. Our results also imply that manipulations cannot be completely hidden by making them computationally hard to find: a voter can randomly try different permutations and find a useful manipulation with non-negligible probability.

Out results extend those of Friedgut, Kalai and Nisan, which worked only for the case of 3 alternatives. The methods we use are quite different though, using a canonical-paths style geometric argument.

4 Outcome of the Meeting

The meeting has brought together some of the best researchers actively working in various areas of complexity. The richness of the field of complexity theory has been reflected in the wide range of topics discussed at the meeting: from classical problems on the complexity of SAT, to communication complexity, learning, quantum algorithms, error-correcting codes, pseudorandomness, and even social choice theory. While seemingly different, many of these areas share ideas, and contribute techniques useful in other areas.

The workshop provided a valuable venue for exchange of ideas between researchers working in different areas, and stimulating discussions. Some new results have already been obtained thanks to such discussions at the workshop, and more are likely to follow. Even more importantly, the meeting has been a source of enthusiasm and encouragement for many young researchers who will be shaping complexity theory in the near future.