Mathematisches Forschungsinstitut Oberwolfach

Report No. 19/2003

Komplexitätstheorie

April 27th – May 3rd, 2003

Complexity theory is concerned with the study of the intrinsic difficulty of computational tasks. It is a central field of theoretical computer science. The 15th Oberwolfach Conference on Complexity Theory was organized by Joachim von zur Gathen (Paderborn), Oded Goldreich (Rehovot), and Claus Peter Schnorr (Frankfurt).

The meeting consisted of five general sessions, and in addition special sessions on the following topics:

- Algebraic Complexity
- Cryptography
- Lattices
- Pseudorandomness
- Proof Complexity
- Extractors/Derandomization
- List-Decoding

The organizers and participants thank the *'Mathematisches Forschungsinstitut Oberwolfach'* for making this conference possible.

The abstracts below are listed in alphabetical order.

Abstracts

Primes in P

WORK BY: AGRAWAL, KAYAL AND SAXENA
PRESENTED BY: ADI AKAVIA

The famous problem "Is PRIMES in P?" was solved in August 2002 by Agrawal, Kayal and Saxena, who devised a polynomial time algorithm for determining whether an integer is prime or composite. In this talk, their algorithm is presented and analyzed.

Proving Hard-Core Predicates Using List Decoding

ADI AKAVIA
(joint work with Shafi Goldwasser and Muli Safra)

We introduce a unifying framework for proving that predicate P is hard-core for a one-way function f, and apply it to a broad family of functions and predicates, reproving old results in an entirely different way as well as showing new hard-core predicates for well known one-way function candidates.

Our framework extends the list-decoding method of Goldreich and Levin for showing hard-core predicates. Namely, a predicate will correspond to some error correcting code, predicting a predicate will correspond to access to a corrupted codeword, and the task of inverting one-way functions will correspond to the task of list decoding a corrupted codeword.

A characteristic of the error correcting codes which emerge and are addressed by our framework, is that codewords can be approximated by a small number of heavy coefficients in their Fourier representation. Moreover, as long as corrupted words are close enough to legal codewords, they will share a heavy Fourier coefficient. We list decode such codes, by devising a learning algorithm applied to corrupted codewords for learning heavy Fourier coefficients.

For codes defined over $\{0,1\}^n$ domain, a learning algorithm by Kushilevitz and Mansour already exists. For codes defined over Z_N, which are the codes which emerge for predicates based on number theoretic one-way functions such as the RSA and Exponentiation modulo primes, we develop a new learning algorithm. This latter algorithm may be of independent interest outside the realm of hard-core predicates.

Proving Integrality Gaps Without Knowing the Linear Program

SANJEEV ARORA

During the past decade we have had much success in proving (using probabilistically checkable proofs or PCPs) that computing approximate solutions to NP-hard optimization problems such as CLIQUE, COLORING, SET-COVER etc. is no easier than computing optimal solutions.

After the above notable successes, this effort is now stuck for many other problems, such as METRIC TSP, VERTEX COVER, GRAPH EXPANSION, etc.

In a recent paper with Béla Bollobás and László Lovász we argue that NP-hardness of approximation may be too ambitious a goal in these cases, since NP-hardness implies a lowerbound —assuming $P \neq NP$— on *all* polynomial time algorithms. A less ambitious goal might be to prove a lowerbound on restricted families of algorithms. Linear and semidefinite programs constitute a natural family, since they are used to design most approximation algorithms in practice. A lowerbound result for a large subfamily of linear programs may then be viewed as a lowerbound for a restricted computational model, analogous say to lowerbounds for monotone circuits

The above paper showed that three fairly general families of linear relaxations for vertex cover cannot be used to design a 2-approximation for Vertex Cover. Our methods seem relevant to other problems as well.

This talk surveys this work, as well as other open problems in the field. The most interesting families of relaxations involve those obtained by the so-called *lift and project* methods of Lovász-Schrijver and Sherali-Adams.

Proving lowerbounds for such linear relaxations involve elements of combinatorics (i.e., strong forms of classical Erdős theorems), proof complexity, and the theory of convex sets.

References

[1] S. Arora, B. Bollobás, and L. Lovász. Proving integrality gaps without knowing the linear program. *Proc. IEEE FOCS 2002*.

[2] S. Arora and C. Lund. Hardness of approximations. In [3].

[3] D. Hochbaum, ed. *Approximation Algorithms for NP-hard problems.* PWS Publishing, Boston, 1996.

[4] L. Lovász and A. Schrijver. Cones of matrices and setfunctions, and 0-1 optimization. *SIAM Journal on Optimization*, 1:166–190, 1990.

[5] H. D. Sherali and W. P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zeroone programming problems. *SIAM J. Optimization*, 3:411–430, 1990.

How to Go Beyond the Black-Box Simulation Barrier

Boaz Barak

The simulation paradigm is central to cryptography. A simulator is an algorithm that tries to simulate the interaction of a (possibly cheating) party with an honest party, without knowing the private input of this honest party. Almost all known simulators use the strategy of the possibly cheating party as a black-box. We present the first constructions of non-black-box simulators. Using these new non-black-box techniques we obtain several results that were previously proven to be impossible to obtain using black-box simulators.

Specifically, assuming the existence of collision resistent hash functions, we construct a new zero-knowledge argument system for NP that satisfies the following properties:

1. This system has a constant number of rounds with negligible soundness error.

2. It remains zero knowledge even when composed concurrently n times, where n is the security parameter.

Simultaneously obtaining 1 and 2 has been proven to be impossible to achieve using black-box simulators.

3. It is an Arthur-Merlin (public coins) protocol.

Simultaneously obtaining 1 and 3 has also been proven to be impossible to achieve with a black-box simulator.

4. It has a simulator that runs in strict polynomial time, rather than in expected polynomial time.

This is the first protocol to simultaneously obtain 1 and 4. Following this work it was shown that obtaining 1 and 4 is also impossible to achieve with a black-box simulator.

Derandomization in Cryptography

BOAZ BARAK

(joint work with Shien Jin Ong and Salil Vadhan)

We give two applications of Nisan–Wigderson-type ("non-cryptographic") pseudorandom generators in cryptography. Specifically, assuming the existence of an appropriate NW-type generator, we construct:

1. A one-message witness-indistinguishable proof system for every language in NP, based on any trapdoor permutation. This proof system does not assume a shared random string or any setup assumption, so it is actually an "NP proof system."
2. A noninteractive bit commitment scheme based on any one-way function.

The specific NW-type generator we need is a hitting set generator fooling *nondeterministic circuits*. It is known how to construct such a generator if $ETIME = DTIME(2^{O(n)})$ has a function of nondeterministic circuit complexity $2^{\Omega(n)}$ (Miltersen and Vinodchandran, FOCS '99).

Our witness-indistinguishable proofs are obtained by using the NW-type generator to derandomize the ZAPs of Dwork and Naor (FOCS '00). To our knowledge, this is the first construction of an NP proof system achieving a secrecy property.

Our commitment scheme is obtained by derandomizing the interactive commitment scheme of Naor (J. Cryptology, 1991). Previous constructions of noninteractive commitment schemes were only known under incomparable assumptions.

Formula Caching Proof Systems

PAUL BEAME

(joint work with Russell Impagliazzo, Toniann Pitassi, and Nathan Segerlind)

A fruitful connection between algorithm design and proof complexity is the formalization of the DPLL approach to satisfiability testing in terms of tree-like resolution proofs. We consider extensions of the DPLL approach that add some version of memorization, remembering formulas the algorithm has previously shown unsatisfiable. Various versions of such formula caching algorithms have been suggested for satisfiability and stochastic satisfiability. We formalize this method, and characterize the strength of various versions in terms of proof systems. These proof systems seem to be both new and simple, and have a rich structure. We compare their strength to several studied proof systems: tree-like resolution, regular resolution, general resolution, and Res(k). We give both simulations and separations.

Algebras of Minimal Rank over Arbitrary Field

MARKUS BLÄSER

Let $R(A)$ denote the rank (also called bilinear complexity) of a finite dimensional associative algebra A. A fundamental lower bound for $R(A)$ is the so-called Alder–Strassen bound $R(A) \geq 2\dim A - t$, where t is the number of maximal twosided ideals of A. The class of algebras for which the Alder–Strassen bound is sharp, the so-called algebras of minimal rank, has received a wide attention in algebraic complexity theory. As the main contribution of this work, we characterize all algebras of minimal rank over arbitrary fields. This finally solves an open problem in algebraic complexity theory. An algebra is of minimal rank, if and only if

$$A \cong C_1 \times \cdots \times C_s \times \underbrace{k^{2\times 2} \times \cdots \times k^{2\times 2}}_{u \text{ times}} \times B$$

where $C_1, \ldots, C_s$ are local algebras of minimal rank with $\dim(C_\sigma/\operatorname{rad} C_\sigma) \geq 2$ (as characterized by Büchi and Clausen), that is, $C_\sigma \cong k[X]/(p_\sigma(X)^{d_\sigma})$ for some irreducible polynomial p_σ with $\deg p_\sigma \geq 2$, $d_\sigma \geq 1$, and $\#k \geq 2\dim C_\sigma - 2$, and B fulfils $B/\operatorname{rad} B = k^q$ and is minimal rank, that is, there exist $w_1, \ldots, w_m \in \operatorname{rad} B$ with $w_i w_j = 0$ for $i \neq j$ such that

$$\operatorname{rad} B = \mathsf{L}_B + Bw_1B + \cdots + Bw_mB = \mathsf{R}_B + Bw_1B + \cdots + Bw_mB$$

and $\#k \geq 2N(B) - 2$. Any of the integers s, u, or m may be zero and the factor B is optional. Above, L_B and R_B denote the left and right annihilator of $\operatorname{rad} B$ and $N(B) = \max\{s \mid (\operatorname{rad} B)^s \neq \{0\}\}$.

Optimal Lower Bound for Polynomial Multiplication

Peter Buergisser

(joint work with Martin Lotz)

We prove lower bounds of order $n \log n$ for both the problem to multiply polynomials of degree n, and to divide polynomials with remainder, in the model of bounded coefficient arithmetic circuits over the complex numbers. These lower bounds are optimal up to order of magnitude. The proof uses a recent idea of R. Raz [Proc. 34th STOC 2002] proposed for matrix multiplication. It reduces the linear problem to multiply a random circulant matrix with a vector to the bilinear problem of cyclic convolution. We treat the arising linear problem by extending J. Morgenstern's bound [J. ACM 20, pp. 305-306, 1973] in a unitarily invariant way. This establishes a new lower bound on the bounded coefficient complexity of linear forms in terms of the singular values of the corresponding matrix. In addition, we extend these lower bounds for linear and bilinear maps to a model of circuits that allows a restricted number of unbounded scalar multiplications.

Counting Classes and Computational Complexity of Algebraic and Topological Invariants

Peter Buergisser

(joint work with Felipe Cucker)

We define counting classes $\#P_R$ and $\#P_C$ in the BSS-setting of computations over the real and complex numbers, respectively. The problems of counting the number of solutions of semialgebraic and complex algebraic sets, respectively, are natural complete problems in these classes. It turns out that these classes capture the complexity of computing some basic invariants in algebraic topology (over R) and algebraic geometry (over C). In fact, the computation of the Euler characteristic of (certain) real algebraic varieties is $FP_R^{\#P_R}$-complete, while the computation of the geometric degree of complex algebraic sets is $FP_C^{\#P_C}$-complete. We also define new counting complexity classes in the discrete (classical) model via taking Boolean parts of the classes above, and show that the discrete versions of the Euler characteristic and geometric degree problem are complete in these classes.

Arthur and Merlin Take a Walk

ANNE CONDON
(joint work with Michael Saks and Joseph Wong)

Arthur-Merlin games model situations in which Merlin must devise a strategy to win against Arthur, who merely flips a coin to determine his moves. In this talk, we consider the following very simple example of such a game. Merlin is placed at the left end of a finite set of points on a line, and will perform a random walk on the points. A finite set U of probabilities is given. Before starting, Merlin can choose, for each point r, a probability $p(r)$ from the set U. The mapping p from the set of points on the line to U is called Merlin's strategy; if there are k points, then Merlin has $|U|^k$ possible strategies. Once the strategy is fixed, the walk proceeds as follows. When Merlin is on point r, Arthur flips a coin which produces heads with probability $p(r)$. Then Merlin moves left if the outcome is heads and moves right otherwise.

Let's suppose that Merlin's goal is to fall off the left end of the line before falling off the right end. Typically, one is interested in knowing what strategy Merlin would use to maximize his probability of success. If, for example, $U = 1/3, 3/4$, the answer is easily found: obviously Merlin would choose to move left from every point with probability 3/4, and a gambler's ruin analysis would reveal the probability of success.

The question we ask is different. We consider the set $S_U(n)$ of Merlin's success probabilities, taken over all possible strategies on a line with n points. We ask: does $S_U(n)$ tend to a limit (suitably defined for sets) as n goes to infinity?

In the talk we describe the answer to this question, motivate why we are interested in this problem, and introduce several related open problems.

Relativized Propositional Calculus

STEPHEN COOK

Motivation: Complexity lower bounds and independence results are easier in a relativized setting. It seems reasonable to define a relativized setting for the propositional calculus in order to prove lower bounds.

We extend the syntax of ordinary propositional calculus to the language **PC(R)** by adding a single relation symbol R intended to stand for a unary relation on strings $\{0,1\}^*$. We add the formation rule

$$\text{if } A_1, ..., A_n \text{ are formulas, } n \geq 0, \text{ then } R(A_1, ..., A_n) \text{ is a formula}$$

An interpretation τ assigns 0 or 1 (false or true) to each atom, and assigns the set $R^\tau \subseteq \{0,1\}^*$ to R. Then

$$R(A_1, ..., A_n)^\tau = 1 \iff A_1^\tau ... A_n^\tau \in R^\tau$$

This syntax and semantics is essentially the same as that defined by Ben-David and Gringauze in [1].

We say that A is *valid* iff $A^\tau = 1$ for all structures τ, and A is *satisfiable* iff $A^\tau = 1$ for some structure τ.

Theorem 1: The satisfiability problem for **PC(R)**formulas is in **NP** (and hence **NP**-complete).

The proof system **PK(R)** is Gentzen's sequent system **PK**, except formulas are allowed to be **PC(R)**formulas, and in addition to the axiom scheme $A \rightarrow A$, and the axioms $\rightarrow 1$ and $0 \rightarrow$, we add the axiom scheme

AX : $\quad \neg A \vee B, A \vee \neg B, R(\vec{C}, A, \vec{D}) \rightarrow R(\vec{C}, B, \vec{D})$

which asserts that if A and B are equivalent, then one can be substituted for the other as an argument of R.

Theorem 2: $\mathbf{PK(R)}$ is sound and complete. Further every valid sequent S has a $\mathbf{PK(R)}$ proof π with $O(2^{|S|})$ sequents, where each sequent in π has length $O(|S|)$, where $|S|$ is the total number of symbols in S.

Remark: In counting the number of sequents in a proof, we do not count weakenings and exchanges.

Now we extend the language of $\mathbf{PC(R)}$ to $\mathbf{QPC(R)}$ by allowing quantifiers $\forall x$ and $\exists x$, for an atom x. The semantics are obtained in the obvious way by letting x range over $\{0,1\}$.

Theorem 3: The satisfiability problem for $\mathbf{QPC(R)}$ is complete for $\mathbf{NEXP}$. The same is true for the satisfiability problem restricted to $\Pi_1^q(R)$ formulas.

The proof system $\mathbf{G(R)}$ is the system G of quantified propositional calculus described in section 4.6 of Krajicek's book [2], extended so that formulas are allowed to be $\mathbf{QPC(R)}$ formulas, and we allow the axiom scheme $\mathbf{AX}$ above. In other words, $\mathbf{G(R)}$ is obtained from $\mathbf{PK(R)}$ by extending the definition of formula, and allowing the four quantifier rules of $\mathbf{LK}$ ([2], page 58).

Theorem 4: $\mathbf{G(R)}$ is sound and complete.

Example: Krajicek ([2], p 223) shows that the weak pigeonhole principle $PHP(R)_a^{a2}$ is not provable in the relativized theory $S2_2(R)$. By adapting the proof, we can obtain a stronger, nonuniform version of this result, asserting that the $\mathbf{QPC(R)}$ formulas $\langle PHP(R)_n^{n2}\rangle$ do not have polynomial size proofs in the tree-like system $G_2^*(R)$.

References

1. Shai Ben-David and Anna Gringauze, On the Existence of Optimal Propositional Proof Systems and Oracle-Relativized Propositional Logic. Manuscript, pp 1-12.

2. Jan Krajicek, *Bounded Arithmetic, Propositional Logic, and Complexity Theory.* Cambridge, 1995.

Vertex Cover Inapproximability

Irit Dinur

(joint work with Muli Safra)

allowing us to reach the hardness factor of 1.36, but not 2-epsilon.

Abstract: In this talk we discuss the hardness of approximating the minimum vertex cover, to within factor 1.36. This result builds upon the PCP theorem and specifically the BGS-Hastad paradigm of composing a PCP "outer verifier" with an "inner verifier" that is the long-code. We introduce a generalization of the long-code called the biased long code and present new techniques for analyzing it, relying on analysis of influence of variables on Boolean functions and on Erdos-Ko-Rado theorems on intersecting families of finite sets. We discuss a special 'uniqueness' property required from any PCP outer verifier for successful composition with the biased-long-code in our setting. We present a system which partially possesses this property,

In this talk we discuss the hardness of approximating the

On the Complexity of Real Solving

Marc Giusti

(joint work with Bernd Bank, Joos Heintz and Luis Miguel Pardo)

Let $F_1, \dots, F_p$ be n–variate polynomials with rational coefficients and with degree at most d. Suppose that they are represented by a division–free arithmetic circuit over the rationals of size L and non–scalar depth ℓ. Furthermore, assume that they form a regular sequence of the polynomial ring, that they generate a radical ideal, and that they define a non–empty and smooth real algebraic variety S.

Then there exists an arithmetic network $\mathcal{N}$ with "=" and "<" decision gates over the rationals, which finds a (suitably encoded) representative point for each connected component of S. The size and non–scalar depth of $\mathcal{N}$ are bounded by $\binom{n}{p} L^2 (nd\delta)^{O(1)}$ and $O(n(\ell + \log nd)\log\delta)$, respectively, where $\delta \leq d^n\, p^{n-p}$ is the (suitably defined) degree of the real interpretation of the polynomial equation system $F_1 = \cdots = F_p = 0$.

In order to prove this result we introduce new notions of generalized polar varieties of equidimensional closed algebraic subvarieties of the real and complex affine and projective spaces.

Improved Certification of Random Unsatisfiable $2k$-SAT Instances by Approximation Techniques

Andreas Goerdt

(joint work with Andre Lanka, Frank Schädlich)

Random 4-SAT instances with a linear number of clauses are unsatisfiable with high probability. Experiments show that formulas with approximately $10n$ clauses are unsatisfiable, n being the number of variables. However, given such an instance F we only know that F *should* be unsatisfiable. We are interested in an *efficient* algorithm certifying this. That is, we look for an algorithm which either gives the output "unsatisfiable" in which case the input formula is really unsatisfiable, or it gives the output "fail" in which case the input formula either can be satisfiable or not. Moreover, to be of any use the algorithm should give the output "unsatisfiable" with high probability with respect to the probability space considered.

It is known that random 4-SAT instances with $\mathrm{poly}(\log n) \cdot n^2$ clauses can be efficiently certified as unsatisfiable. Here we show that $C \cdot n$ clauses, where C is a sufficiently large constant are sufficient.

Our algorithm has two stages. In the first stage it certifies the property that an assignment can only satisfy the input formula F, if it makes an odd number of literals true in all clauses up to an asymptotically negligible rest. This implies that the graph obtained from F by making each all-positive clause an edge between its two *pairs* of variables has a cut comprising almost all edges. However, when F is random the graph is a random graph it has only a cut with about one half of all edges. The MAX-CUT approximation algorthim of Goemans/Williamson certifies that we have no cut consisting of all edges and thus that F is unsatisfiable. Instead of the MAX-CUT algorithm one might also use a MIN-BISECTION approximation algorithm of Krauthgamer and Feige.

The technique described can be extended to $2k$-SAT instances giving a bound of $C \cdot n^k$ clauses. For the odd case for example $k = 3$ we still have the bound of $\text{poly}(\log n) \cdot n^{3/2}$ clauses.

Locally Testable Codes and PCPs of Almost-Linear Length

ODED GOLDREICH
(joint work with Madhu Sudan)

Locally testable codes are error-correcting codes that admit very efficient codeword tests. Specifically, using a constant number of (random) queries, non-codewords are rejected with probability proportional to their distance from the code.

Locally testable codes are believed to be the combinatorial core of PCPs. However, the relation is less immediate than commonly believed. Nevertheless, we show that certain PCP systems can be modified to yield locally testable codes. On the other hand, we adapt techniques we develop for the construction of the latter to yield new PCPs. Our main results are locally testable codes and PCPs of almost-linear length. Specifically, we present:

o Locally testable (linear) codes in which k information bits are encoded by a codeword of length approximately $k \cdot \exp(\sqrt{\log k})$. This improves over previous results that either yield codewords of exponential length or obtained almost quadratic length codewords for sufficiently large non-binary alphabet.

o PCP systems of almost-linear length for SAT. The length of the proof is approximately $n \cdot \exp(\sqrt{\log n})$ and verification in performed by a constant number (i.e., 19) of queries, as opposed to previous results that used proof length $n^{1+O(1/q)}$ for verification by q queries.

The novel techniques in use include a random projection of certain codewords and PCP-oracles, an adaptation of PCP constructions to obtain "linear PCP-oracles" for proving conjunctions of linear conditions, and a direct construction of locally testable (linear) codes of sub-exponential length.

On the (In)security of the Fiat-Shamir Paradigm

SHAFI GOLDWASSER
(joint work with Yael Tauman)

In 1986, Fiat and Shamir introduced a general method for transforming secure 3-round public-coin identification schemes into digital signature schemes, which are efficient and hopefully secure against chosen message attacks. This is significant, because all known constructions which guarantee such security are substantially more inefficient and complicated in design.

In 1996, Pointcheval and Stern proved that signature schemes obtained by the Fiat-Shamir transformation are secure in the so-called "Random Oracle Model." The question is: does the proof of security of the Fiat-Shamir transformation in the Random Oracle Model, imply that the transformation yields secure signature schemes in the "real world"?

In this paper, we answer this question negatively. We show that if one way functions exist, then there exist secure 3-round public-coin identification schemes for which the Fiat-Shamir methodology produces **insecure** digital signature schemes for **any** implementation of the Random Oracle Model by a function ensemble. Obviously, if one functions do not exist, then all digital signature schemes are insecure.

Among other things, we make new usage of Barak's technique for taking advantage of non black-box access to a program, this time in the context of digital signatures.

Homomorphic Public-Key Cryptosystems and Encrypting Boolean Circuits

Dimitri Grigoriev

(joint work with Ilia Ponomarenko)

Homomorphic cryptosystems are designed for the first time over any finite group. Applying Barrington's construction we produce for any boolean circuit of the logarithmic depth its encrypted simulation of a polynomial size over an appropriate finitely generated group.

List Decoding with Side Information

Venkatesan Guruswami

Consider the problem of communicating a binary string over a channel that can corrupt an arbitrarily chosen fraction p of symbols. It is well-known that one can use error-correcting codes, specifically those with relative distance at least $2p$, to cope with such channels and ensure correct communication. However, this distance requirement imposes a limit $p < 1/4$ on the error threshold that can be handled. For larger values of p ($1/4 < p < 1/2$), a notion called "list decoding" becomes necessary and appropriate. Under list decoding, the decoding algorithm is allowed to output a small list of codewords that are close to the noisy received word. This, however, is not entirely satisfactory when the receiver needs to unambiguously determine the original transmitted message. We consider one possible scenario that would permit disambiguating between the elements of the list, namely where the sender of the message provides some hopefully small amount of side information about the transmitted message on a separate auxiliary channel that is noise-free (or has a noise threshold less than $1/4$). This setting becomes meaningful and useful when the amount of side information that needs to be communicated is much smaller than the length of the message.

We study what kind of side information is necessary and sufficient in the above context. The short, conceptual answer is that the side information must be randomized and the message recovery is with a small failure probability.

Specifically, we prove that deterministic schemes which guarantee correct recovery of the message provide no savings and essentially the entire message has to be sent as side information. However, there exist randomized schemes which only need side information of length logarithmic in the message length. In fact, in the limit of repeated communication of several messages, the amortized amount of side information needed per message can be a constant independent of the message length or the failure probability. Concretely, we can correct up to a fraction $(1/2-\gamma)$ of errors for binary codes using only $2\log(1/\gamma)+O(1)$ amortized bits of side information per message, and this is in fact the best possible (up to additive constant terms).

Derandomizing Polynomial Identity Tests Means Proving Circuit Lower Bounds

Valentine Kabanets

(joint work with Russell Impagliazzo)

We show that derandomizing Polynomial Identity Testing is, essentially, equivalent to proving circuit lower bounds for $NEXP$. More precisely, we prove that if one can test in polynomial time (or, even, nondeterministic subexponential time, infinitely often) whether a given arithmetic circuit over integers computes an identically zero polynomial, then either (i) $NEXP \not\subseteq P/poly$ or (ii) Permanent is not computable by polynomial-size arithmetic circuits. We also prove a (partial) converse: If Permanent requires superpolynomial-size arithmetic circuits, then one can test in subexponential time whether a given arithmetic formula computes an identically zero polynomial.

Since Polynomial Identity Testing is a $coRP$ problem, we obtain the following corollary: If $RP = P$ (or, even, $coRP \subseteq \cap_{\epsilon>0} NTIME(2^{n^\epsilon})$, infinitely often), then **NEXP** is not computable by polynomial-size arithmetic circuits. Thus, establishing that $RP = coRP$ or $BPP = P$ would require proving superpolynomial lower bounds for Boolean or arithmetic circuits. We also show that any derandomization of RNC would yield new circuit lower bounds for a language in $NEXP$.

Our techniques allow us to prove an unconditional circuit lower bound for a language in $NEXP^{RP}$: we prove that either (i) Permanent is not computable by polynomial-size arithmetic circuits, or (ii) $NEXP^{RP} \not\subseteq P/poly$.

On the Complexity of the Determinant

Erich Kaltofen
(joint work with Gilles Villard)

We present new algorithms for computing the determinant of an $n \times n$ matrix A with integer entries of maximal bit length $\log \|A\|$. With fastest known matrix multiplication algorithms we achieve asymptotic running time of $(n^{2.697263} \log \|A\|)^{1+o(1)}$ bit operations, where the $o(1)$ term in the exponent represents polylogarithmic factors in n and $\log \|A\|$. A version that employs the classical matrix multiplication algorithm and classical greatest common divisor algorithms on matrix polynomials has bit complexity $(n^{3+1/3} \log \|A\|)^{1+o(1)}$. Our algorithms are randomized of the Las Vegas kind and return either "failure" or the correct determinant, the latter with probability no less than 1/2.

Our baby steps/giant steps technique in the context of the Wiedemann determinant algorithm was employed by Kaltofen in 1992 for asymptotically reducing the number of additions, subtractions and multiplications required for computing the determinant of a matrix over an arbitrary commutative ring, that is, without divisions. Again by blocking we can improve the complexity to $O(n^{2.697263})$ ring operations. The division-free algorithm is deterministic.

This work was reported in Oberwolfach in January 2001 at the Finite Field Conference. At that time, due to incomplete analysis our exponents were higher. In the meantime, Arne Storjohann has proposed an algorithm of bit complexity $(n^{2.3755} \log \|A\|)^{1+o(1)}$. His algorithm seems not lead to an improvement of the division-free complexity of the determinant, as it is based on Hensel lifting. For polynomial entries, Jeannerod and Villard achieve the corresponding complexity $(n^{2.3755} \deg A)^{1+o(1)}$ field operations by diagonalization. Their algorithm requires an input matrix in generic position, and is thus superseded by Storjohann's method. However, their technique also applies to the matrix inverse with $(n^3 \deg A)^{1+o(1)}$.

Our methods yield speedups for other problems on matrices with integer or polynomial entries or with entries from an abstract commutative ring, such as the adjoint matrix, the characteristic polynomial (via Hensel lifting), and the Smith normal form.

Algebraic Attacks against Combiners with Memory

Matthias Krause

(joint work with Frederik Armknecht)

Recently, algebraic attacks turned out to be surprisingly useful to attack secret key cryptosystems, e.g. AES, Serpent and LILI-128. This paper extends the use of algebraic attacks to combiners with memory. A (k, l)-combiner consists of k parallel linear feedback shift registers (LFSRs), and the nonlinear filtering is done via a finite automaton with k input bits and l memory bits. It is shown that for (k, l)-combiners, nontrivial canceling relations of degree at most $\lceil k(l+1)/2 \rceil$ exist. This makes algebraic attacks possible. Also, a general method is presented to check for such relations with an even lower degree. This allows to show the invulnerability of certain (k, l)-combiners against this kind of algebraic attacks. On the other hand, this can also be used as a tool to find improved algebraic attacks.

Inspired by this method, the E_0 keystream generator from the Bluetooth standard is analyzed. As it turns out, a secret key can be recovered by solving a system of linear equations with $2^{23.07}$ unknowns. To our knowledge, this is the best published attack on the E_0 keystream generator yet.

A preliminary version of this paper will appear in the CRYPTO 2003 proceedings.

Almost Perfect Lattices

Daniele Micciancio

Two important quantities associated to any point lattice are the packing radius and the covering radius. The packing (resp. covering) radius is the largest r such that spheres of radius r centered at lattice points do not intersect (resp. cover the entire space.) In analogy with perfect codes, we say that a lattice is almost perfect if the covering radius is not much bigger than the packing radius radius.. (In a perfect code the two radii are exactly the same, but it is easy to see that for lattices equality can never hold.) It is known that lattices with constant covering-packing ratio exist, but the proof does not lead to efficient constructions. The lattice of all n-dimensional points with integer coordinates has covering radius $\sqrt{n}$ times larger than the packing radius. In this talk we present efficient constructions of infinite families of lattices with covering-packing ratio asymptotically smaller than $\sqrt{n}$. In particular, we present 1) an efficiently constructible family of lattices with ratio $\sqrt{n/\log n}$ such that the closest vector problem in the lattices can be solved in polynomial time, and 2) an explicit (and efficiently constructible) family of lattices with ratio bounded by $n^{3/8}$. The bound on the covering ratio in the latter family is proved giving a (polynomial time) algorithm that for any input point, finds a lattice point within distance $n^{3/8}$. However, our algorithm does not always return the closest lattice point. Finding a polynomial time closest vector point algorithm for this lattice would give improved lattice based cryptographic hash functions with average-case/worst-case connection.

The Complexity of Fighting Spam

Moni Naor

(joint work with Cynthia Dwork and Andrew Goldberg)

Consider the following simple technique for combatting spam:

If I don't know you, and you want your e-mail to appear in my inbox, then you must attach to your message an easily verified "proof of computational effort", just for me and just for this message.

If the proof of effort requires, say, 10 seconds to compute on a typical machine, then the economics of sending spam are radically altered, as a single machine can send only 8,000 messages per day.

This talk describes recent work on the choice of functional challenges that can be used to yield easily verifiable proofs of computational effort, where most of the work is in retrieving information from memory.

Combinatorial Principles Based on Games

Pavel Pudlak

We consider two players A and B playing k copies of the same finite game G simultaneously. We shall say that the players play G on k boards. A starts by playing the first move on all boards, then B plays second moves on all boards etc.

The following is, trivially, a valid principle:

There do not exist two strategies f and g such that (1) f is a strategy for A to win on at least one board; (2) g is a strategy for B to win on at least one board.

We consider the following computational problem. Given circuits for the strategies f and g and a circuit for the game G (which determines who wins in the final position), decide which of the two (1) or (2) above fails (if both fail, we do not care).

1. We observe that if the number of games k and the number of rounds are constant, then it can be decided in polynomial time. This follows easily from the provability of the principle in $S_1 2$ and Buss's witnessing theorem for this theory, but to extract an explicit algorithm from this argument seems difficult.

2. Therefore we construct such a polynomial algorithm directly. Here is the idea. Take the k first moves given by the strategy f. Apply g to all the k^k k-tuples obtained from these k elements. Thus we obtain k^k k-tuples for the second moves. Take again all possible k tuples from these elements and apply f to it etc. Thus we eventually obtain exponentially many final positions, but the exponential function depends only on the number of games and rounds, hence it is still finite. Restrict G to this finite set; let G' denote this restriction. Now suppose, w.l.o.g., that A has a winning strategy f' in G'. Then by playing f' on each of the k boards we will beat g, because no elements outside of the domain of f' will be produced in this way.

3. We note that there are other combinations of games for which the first argument works, but for which we are not able to apply the second one.

4. Due to time constraints the following part was not presented in the talk. It is possible to further reduce the above principle to a principle saying that an explicit hypergraph cannot be two-colored. The advantage is that we can avoid mentioning circuits; instead the hypergraph is parametrized by three numerical parameters (which correspond to the number of games, the number of rounds, and the number of possible moves in a round).

Pseudo-Random Generators for All Hardnesses

Chris Umans

Given a function $f : \{0,1\}^{\log n} \to \{0,1\}$ with circuit complexity s, we construct a pseudo-random generator that stretches a random seed of length $O(\log n)$ into a string of $m = s^{\Omega(1)}$

pseudo-random bits that fool circuits of size m. The construction works for any hardness s, giving an optimal hardness vs. randomness tradeoff with a direct and self-contained proof. A key element in our construction is an augmentation of the standard low-degree extension encoding that exploits the field structure of the underlying space in a new way.

Deterministic Polynomial Identity Testing in Non Commutative Models

Ran Raz

(joint work with Amir Shpilka)

We give a deterministic polynomial time algorithm for polynomial identity testing in the following two cases:

1) Non Commutative Arithmetic Formulas: The algorithm gets as an input an arithmetic formula in the non-commuting variables $x_1, ..., x_n$ and determines whether or not the output of the formula is identically 0 (as a formal expression).

2) Pure Arithmetic Circuits: The algorithm gets as an input a pure arithmetic circuit (as defined by Nisan and Wigderson) in the variables $x_1, ..., x_n$ and determines whether or not the output of the circuit is identically 0 (as a formal expression).

We also give a deterministic polynomial time identity testing algorithm for non commutative algebraic branching programs as defined by Nisan. One application is a deterministic polynomial time identity testing for multilinear arithmetic circuits of depth 3.

Finally, we observe an exponential lower bound for the size of pure arithmetic circuits for the permanent and for the determinant. (Only lower bounds for the *depth* of pure circuits were previously known).

Vertex Cover Might be Hard to Approximate to within $2 - \epsilon$

Oded Regev

(joint work with Subhash Khot)

Based on a conjecture regarding the power of unique 2-prover-1-round games presented by Khot (STOC'02), we show that vertex cover is hard to approximate within any constant factor better than 2. We actually show a stronger result, namely, based on the same conjecture, vertex cover on k-uniform hypergraphs is hard to approximate within any constant factor better than k.

Private Computations in Networks: Topology versus Randomness

Rüdiger Reischuk

(joint work with Andreas Jakoby and Maciej Liśkiewicz)

In a distributed network, computing a function privately requires that no participant gains any additional knowledge other than the value of the function. We study this problem for incomplete networks and establish a tradeoff between connectivity properties of the network and the amount of randomness needed. First, a general lower bound on the number of random bits is shown. Next, for every $k \geq 2$ we design a quite efficient (with respect to randomness) protocol for symmetric functions that works in arbitrary k-connected networks. Finally, for directed cycles that compute threshold functions privately almost matching lower and upper bounds for the necessary amount of randmoness are proven.

Analysis of Boolean Functions and Various Applications

Muli Safra

Representing a Boolean function as a polynomial is only natural. It turns out that this representation, along with some related technology – for example the study of the Influence of variables on Boolean functions – gives insight to many aspects of such functions. This field was founded in a paper by Kahn, Kalai and Linial from '89, and has since shown applications in a wide array of fields, including Game Theory and Social Choice, Economics, Percolation, and Complexity theory.

The talk will survey the methodology and some of its applications, to Mechanism Design, Graph Properties and Complexity Theory. We would then consider some further applications, show the state of art in terms of known results in the field; and suggest open problems with their relevant applications.

Proof Systems and Chosen-Ciphertext Security

Amit Sahai

Zero-knowledge proofs, introduced by Goldwasser, Micali, and Rackoff, are fascinating constructs in which one party (the "prover") convinces another party (the "verifier") that some assertion is true, without revealing anything else to the verifier.

In this talk, we present a connection between the theory of zero-knowledge proofs and one of the classical notions in cryptography, encryption. In particular, we introduce the notion of simulation-sound zero knowledge, and show how the non-interactive form of this notion can be used to achieve a strong notion of security for encryption, namely adaptive chosen-ciphertext security, in a simple manner. We also present constructions of simulation-sound non-interactive zero-knowledge proofs for all NP languages, and discuss other applications of this notion.

LLL-type Lattice Reduction in $O(n^3 logn)$ Arithmetic Steps

Claus Schnorr

We modify the concept of LLL-reduction of lattice bases in the sense of Lenstra, Lenstra, Lovász (1982) towards a faster reduction algorithm. We introduce SLLL-bases and a corresponding algorithm of SLLL-reduction that organizes the reduction reduction in segments of size k. Local reduction of segments is done using local coordinates of dimension $2k$.

Our SLLL-bases approximate the successive minima of the lattice in the same way as LLL-bases. For integer lattices of dimension n given by a basis of length $2^{O(n)}$ SLLL-reduction runs in $O(n^{5+\varepsilon})$ bit operations for every $\varepsilon > 0$, compared to $O(n^{7+\varepsilon})$ for the original LLL and to $O(n^{6+\varepsilon})$ for the LLL-algorithms of Schnorr (1988) and Storjohann (1996). SLLL-reduction via iterated subsegments runs in $O(n^3 \log n)$ arithmetic steps and $O(n^{4.5+\varepsilon})$ bit operations.

Uniform Hardness Versus Randomness Tradeoffs for Arthur-Merlin

Ronen Shaltiel

(joint work with Danny Gutfreund and Amnon Ta-Shma)

Arthur-Merlin games are interactive protocols in which Merlin (who is all-powerful) convinces Arthur (who is probabilistic polynomial time) that an input x belongs to some language L. The class AM is the class of all languages L which have a convincing Arthur-Merlin protocol. It is a big open problem whether AM=NP, or in other words, whether Arthur can be "derandomized" and made deterministic. (Note that if Arthur is deterministic then the Arthur-Merlin protocol becomes an NP proof.) This problem is analogous to the BPP versus P question which asks whether any probabilistic poly-time algorithm can be "derandomized" and made deterministic.

Impagliazzo and Wigderson addressed the second question and showed that probabilistic algorithms are either very strong or very weak. We show an analogous result for Arthur-Merlin games: Either Arthur-Merlin protocols are very strong (and any language in $E = dtime(2^{O(n)})$ can be proved to a sub-exponential time Arthur) or Arthur-Merlin protocols are weak and "AM=NP on real life inputs". (More precisely, every language in AM has an NP procedure. This procedure does not necessarily answer correctly on every input. However, it is infeasible to come up with inputs on which the procedure fails.)

I'll start the talk by giving a survey on hardness versus randomness tradeoffs for both BPP and AM.

Simple Extractors for All Min-Entropies and a New Pseudo-Random Generator

Ronen Shaltiel

(joint work with Chris Umans)

We present a simple, self-contained extractor construction that produces good extractors for all min-entropies (min-entropy measures the amount of randomness contained in a weak random source). Our construction is algebraic and builds on a new polynomial-based approach introduced by Ta-Shma, Zuckerman, and Safra and avoids complicated recursions, iterations, and compositions that characterized much of the previous work.

Applying similar ideas to the problem of building pseudo-random generators, we obtain a new pseudo-random generator construction that is *not* based on the Nisan-Wigderson generator, and turns worst-case hardness *directly* into pseudo-randomness. The parameters of this generator match those of previous constructions and in particular are strong enough to obtain a new proof that $P = BPP$ if E requires exponential size circuits. Our construction also yields *optimal* hitting set generators closing the gap left by previous constructions.

The problem we solve can be thought of as (list)-decoding the Reed-Muller in a an unusual setup, and in particular we get a new list-decoding algorithm for the Reed-Muller code.

Lower Bounds for Matrix Product

Amir Shpilka

(joint work with Amir Shpilka)

We prove lower bounds on the number of product gates in bilinear and quadratic circuits that compute the product of two $n \times n$ matrices over finite fields. In particular we obtain the following results:

1. We show that the number of product gates in any *bilinear* (or *quadratic*) circuit that computes the product of two $n \times n$ matrices over $GF(2)$ is at least $3n2 - o(n2)$.
2. We show that the number of product gates in any *bilinear* circuit that computes the product of two $n \times n$ matrices over $GF(q)$ is at least $(2.5 + \frac{1.5}{q3-1})n2 - o(n2)$.

These results improve the former results of Bshouty from '89 and of Bläser from '99 who proved lower bounds of $2.5n2 - o(n2)$.

Locally Testable Cyclic Codes

Amir Shpilka

(joint work with László Babai, Amir Shpilka and Daniel Štefankovič)

? We consider binary linear codes, i. e., subspaces $C \leq GF(2^n)$. A family of codes C is *good* if $\dim(C) = \Omega(n)$ and the weight of each nonzero codeword is $\Omega(n)$.

The code C is *r-testable* if there exists a randomized algorithm which, given a word $x \in GF(2^n)$, adaptively selects r positions, checks the bits of x in the selected position, and makes a decision (accept or reject x) based on the bits found on the positions selected, such that

(i) if $x \in C$ then x is surely accepted;

(ii) if $dist(x, C) \geq \epsilon n$ then x is probably rejected. ("dist" refers to Hamming distance.)

A family of codes is *locally testable* if there exists a constant r such that all members of the family are r-testable. This concept arose from holographic proofs/PCPs.

A *cyclic code* is a linear code which is invariant under the cyclic shift of the coordinates. Cyclic codes play an important role in classical coding theory. It is a long-standing open problem whether there exist good cyclic codes. It is a more recent question whether there exist good, locally testable codes.

In this paper we address the intersection of these two questions.

Theorem. *There are no good, locally testable cyclic codes.*

In fact our result is stronger in that it replaces condition (ii) of local testability by the condition

(ii') if $dist(x, C) \geq \epsilon n$ then x has a positive chance of being rejected.

The proof involves methods from Galois theory, cyclotomy, and diophantine approximation.

On the Power of Quantum Proofs

Amir Shpilka

(joint work with Ran Raz and Amir Shpilka)

We study the power of quantum proofs, or more precisely, the power of Quantum Merlin-Arthur (QMA) protocols, in two well studied models of quantum computation: the black box model and the communication complexity model.

Our main results are obtained for the communication complexity model. For this model, we identify a complete promise problem for QMA protocols, the *Linear Subspaces Distance* problem. The problem is of geometrical nature: Each player gets a linear subspace of $\mathbb{R}^m$ and considers the sphere of unit vectors in that subspace. Their goal is to output 1 if the distance between the two spheres is very small (say, smaller than $0.1 \cdot \sqrt{2}$) and 0 if the distance is very large (say, larger than $0.9 \cdot \sqrt{2}$). We show that:

1. The QMA communication complexity of the problem is $O(\log m)$.
2. The (classical) MA communication complexity of the problem is $\Omega(m^\epsilon)$ (for some $\epsilon > 0$).
3. The (standard) quantum communication complexity of the problem is $\Omega(\sqrt{m})$.

In particular, this gives an exponential separation between QMA communication complexity and MA communication complexity.

For the black box model we give several observations. First, we observe that the block sensitivity method, as well as the polynomial method for proving lower bounds for the number of queries, can both be extended to QMA protocols. We use these methods to obtain lower bounds for the QMA black box complexity of functions. In particular, we obtain a tight lower bound of $\Omega(N)$ for the QMA black box complexity of a random function, and a tight lower bound of $\Omega(\sqrt{N})$ for the QMA black box query complexity of $NOR(X_1, ..., X_N)$. In particular, this shows that any attempt to give short quantum proofs for the class of languages $Co-NP$ will have to go beyond black box arguments.

We also observe that for any boolean function $G(X_1, ..., X_N)$, if for both G and $\neg G$ there are QMA black box protocols that make at most T queries to the black box, then there is a classical deterministic black box protocol for G that makes $O(T^6)$ queries to the black box. In particular, this shows that in the black box model $QMA \cap Co-QMA = P$.

On the positive side, we observe that any (total or partial) boolean function $G(X_1, ..., X_N)$ has a QMA black box protocol with proofs of length N that makes only $O(\sqrt{N})$ queries to the black box.

Finally, we observe a very simple proof for the exponential separation (for promise problems) between QMA black box complexity and (classical) MA black box complexity (first obtained by Watrous).

On Worst-Case to Average-Case Reductions for NP Problems

LUCA TREVISAN
(joint work with Andrej Bogdanov)

We formalize a general notion of "worst-case to average-case" reduction that, in particular, contains the notion of random self-reduction as a special case.

We show that if an NP-complete problem had a non-adaptive worst-case to average-case reduction then NP would be contained in non-uniform AM, and the polynomial hierarchy would collapse to the third level. Feigenbaum and Fortnow show the same conclusion from the assumption that an NP-complete problem had a random self-reduction.

List Decoding Using the XOR Lemma

LUCA TREVISAN

We show that Yao's XOR Lemma, and its essentially equivalent rephrasing as a Direct Product Lemma, can be re-interpreted as a way of obtaining error-correcting codes with good list-decoding algorithms from error-correcting codes having weak unique-decoding algorithms. To get codes with good rate and efficient list decoding algorithms one needs a proof of the Direct Product Lemma that, respectively, is strongly derandomized, and uses very small advice.

We show how to reduce advice in Impagliazzo's proof of the Direct Product Lemma for pairwise independent inputs, which leads to error-correcting codes with $O(n^2)$ encoding length, $O(n^2)$ encoding time, and probabilistic $O(n)$ list-decoding time. (Note that the decoding time is sub-linear in the length of the encoding.)

Back to complexity theory, our advice-efficient proof of Impagliazzo's "hard-core set" results yields a (weak) uniform version of O'Donnell results on amplification of hardness in NP. We show that if there is a problem in NP that cannot be solved by BPP algorithms on more than a $1-1/(logn)^c$ fraction of inputs, then there is a problem in NP that cannot be solved by BPP algorithms on more than a $3/4+1/(logn)^c$ fraction of inputs, where $c>0$ is an absolute constant.

On Converting CNF to DNF

Ingo Wegener

(joint work with Peter Bro Miltersen and Jaikumar Radhakrishnan)

The best-known representations of boolean functions f are those as disjunction of terms (DNFs) and as conjunction of clauses (CNFs). It is convenient to define the DNF size of f as the minimal number of terms in a DNF representing f and the CNF size as the minimal number of clauses in a CNF representing f. This leads to the problem to estimate the largest gap between CNF size and DNF size. More precisely, what is the largest possible DNF size of a function f with polynomial CNF size? We show the answer to be $2^{n-\Theta(n/\log n)}$.

Expander Graphs - Where Combinatorics and Algebra Compete and Cooperate

Avi Wigderson

Expansion of graphs can be given equivalent definitions in combinatorial and algebraic terms. This is the most basic connection between combinatorics and algebra illuminated by expanders and the quest to construct them. The talk will survey how fertile this connection has been to both fields, focusing on recent results.

Extractors - Optimal to Constant Factors

Avi Wigderson

(joint work with Chi-Jen Lu, Omer Reingold and Salil Vadhan)

Randomness extractors are functions which extract almost uniform bits from sources of biased and correlated bits, using a short, truely random seed as a catalyst. Extractors play a fundamental role in the theory of pseudorandomness, and have a wide variety of applications. Thus coming up with explicit constructions of extractors has been the focus of a large body of work over the past decade.

This paper gives an explicit construction of extractors which work for sources on strings of length n that contain any min-entropy k. These extractors can extract any constant fraction of the min-entropy using a seed of length $O(\log n)$, and has an arbitrarily small constant error. This is the first construction that works for any min-entropy and is simultaneously optimal up to a constant factor in both the seed length and output length.

I will explain some of the old and new ideas which lead to the new construction. In particular we'll see new constructions of "mergers" from locally decodable error correcting codes, new constructions of "condensers" which have constant seed length, and how to compose these a non-constant number of steps via error reduction.

Derandomized "Low Degree" Tests via "Epsilon-Biased" Sets, with Applications to Short Locally Testable Codes and PCPs

Avi Wigderson

(joint work with Eli Ben-Sasson, Madhu Sudan and Salil Vadhan)

We present the first explicit construction of Probabilistically Checkable Proofs (PCPs) and Locally Testable Codes (LTCs) of fixed constant query complexity which have almost-linear size. Such objects were recently shown to exist (nonconstructively) by Goldreich and Sudan.

The key to these constructions is a nearly optimal randomness-efficient version of the Rubinfeld-Sudan low degree test. The original test uses a random line in the given vector space. The number of such lines is quadratic in the size of the space, which implied a similar blow up in previous constructions of LTCs. Goldreich and Sudan showed that *there exists* a nearly linear sized sample space of lines such that running the low-degree test on a random line from this collection is a good test. We give an explicit sample space with this property.

In a similar way we give a randomness-efficient version of the Blum-Rubinfeld-Sudan linearity test (which is used, for instance, in locally testing the Hadamard code).

Both derandomizations are obtained through epsilon-biased sets for vector spaces over finite fields. The sample space consists of the lines defined by the edges of the Cayley expander graph generated by the epsilon-biased set.

The analysis of the derandomized tests rely on alternative views of epsilon-biased sets — as generating sets of Cayley expander graphs for the low degree test, and as defining good linear error-correcting codes for the linearity test.

Deterministic Extractors for Bit-Fixing Sources and Exposure-Resilient Cryptography

David Zuckerman

(joint work with Jesse Kamp)

We give a linear-time deterministic algorithm which extracts $\Omega(n^{2\gamma})$ almost-random bits from sources where $n^{\frac{1}{2}+\gamma}$ of the n bits are uniformly random and the rest are fixed in advance. This improves on the previous constructions which required that at least $n/2$ of the bits be random. Our construction also gives explicit adaptive exposure-resilient functions and in turn adaptive all-or-nothing transforms. For sources where instead of bits the values are chosen from $[d]$, we show that for $d > 2$ we can extract a constant fraction of the randomness. We also give bounds on extracting randomness for sources where the fixed bits can depend on the random bits.

Edited by Adi Akavia