Mathematisches Forschungsinstitut Oberwolfach

Report No. 52/2009

DOI: 10.4171/OWR/2009/52

Complexity Theory

Organised by
Peter Bürgisser (Paderborn)
Joachim von zur Gathen (Bonn)
Oded Goldreich (Rehovot)
Madhu Sudan (Cambridge, MA)

November 15th – November 21st, 2009

Abstract. Computational Complexity Theory is the mathematical study of the intrinsic power and limitations of computational resources like time, space, or randomness. The current workshop focused on recent developments in various sub-areas including arithmetic complexity, Boolean complexity, communication complexity, cryptography, probabilistic proof systems, pseudorandomness, and quantum computation. Many of the developements are related to diverse mathematical fields such as algebraic geometry, combinatorial number theory, probability theory, quantum mechanics, representation theory, and the theory of error-correcting codes.

Mathematics Subject Classification (2000): 68-06, 68Q01, 68Q10, 68Q15, 68Q17, 68Q25, 94B05, 94B35.

Introduction by the Organisers

The workshop *Complexity Theory* was organized by Peter Bürgisser (Universität Paderborn), Joachim von zur Gathen (B-IT, Bonn), Oded Goldreich (Weizmann Institute), and Madhu Sudan (MIT). The workshop was held on November 15th–21st 2009, and attended by approximately 50 participants spanning a wide range of interests within the field of Computational Complexity. The plenary program, attended by all participants, featured few long lectures as well as short (5-minute) reports by almost all participants. In addition, intensive interaction took place in smaller groups.

The Oberwolfach Meeting on Complexity Theory is marked by a long tradition and a continuous transformation. Originally starting with a focus on algebraic and

Boolean complexity, the meeting has continuously evolved to cover a wide variety of areas, most of which were not even in existence at the time of the first meeting (in 1972). While inviting many of the most prominent researchers in the field, the organizers try to identify and invite a fair number of promising young researchers.

Computational complexity (a.k.a. complexity theory) is a central field of computer science with a remarkable list of celebrated achievements as well as a vibrant research activity. The field is concerned with the study of the *intrinsic complexity* of computational tasks, and this study tends to *aim at generality*: it focuses on natural computational resources, and considers the effect of limiting these resources on the class of problems that can be solved. Computational complexity is related to and has substantial interaction with other areas of mathematics such as algebra, analysis, combinatorics, geometry, number theory, optimization, probability theory, and quantum computation.

The workshop focused on several sub-areas of complexity theory and its nature may be best illustrated by a brief survey of some of the meeting's highlights.

Efficient Simulation of Quantum Mechanics. The power of the standard model of quantum computation (QC), demonstrated by Shor's celebrated quantum algorithm for integer factorization, presses the fundamental question of whether this standard model is feasibly realizable. In the meeting, Scott Aaronson presented a different evidence to the difficulty of (classically) simulating a quantum mechanical (QM) system. His fundamental result exhibits a specific distribution that arises in QM and is easily generated by a QC, and he provides strong evidence that no (classical) probabilistic polynomial-time algorithm can generate it. Namely, such an algorithm would imply that any counting problem can be efficiently solved using an oracle to $\mathcal{NP}$ (i.e., $\mathcal{P}^{\#\mathcal{P}}$ would equal $\mathcal{BPP}^{\mathcal{NP}}$). This holds even if the classical algorithm only approximates the said distribution of the QM system. Furthermore, the QM system is a very simple and special one; it consists of a system of identical, non-interacting bosonic particles. This contrasts with the efficient simulation of a system of fermions shown by Valiant.

Kakeya Sets and Extractors. In 1999 Wolff posed the finite field analogue to the Kakeya problem, conjecturing that for every $K \subseteq \mathbb{F}_q^n$ that contains a line in every direction it holds that $K = \Omega(|\mathbb{F}_q^n|)$, where the constant hidden in the Omega-notation may depend on n. This analogue was observed to be related to the design of randomness extractors, hence the complexity theoretic interest in it. In the meeting, Zeev Dvir surveyed a recent series of works that settle this conjecture and obtained almost the optimal constant in the Omega-notation. Furthermore, he showed that the proof techniques are indeed applicable to the analysis of constructions of randomness extractors, yielding improvements in some of the parameters of such constructions. Recall that a (k, ϵ)-randomness extractor is a function $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ such that for every random variable X of min-entropy at least k, when s is selected uniformly in $\{0,1\}^t$ it holds that $E(X, s)$ is ϵ-close to the uniform distribution over $\{0,1\}^m$.

Locally Decodable Codes of Sub-exponential Length. An error-correcting code is called locally decodable if, given a corrupted codeword, any bit in the original message can be correctly reconstructed (with high probability) based on a constant number of probes. Locally Decodable Codes (LDC) are closely related to (multi-server) Private Information Retrieval (PIR) schemes, which are of interest to cryptography. In the meeting, Klim Efremenko presented a 3-query LDC of sub-exponential length, thus improving on a breakthrough result of Yekhanin (which was presented in the 2007 meeting). Furthermore, in contrast to Yekhanin's construction, the current construction scheme has the pleasing feature of benefiting from more queries (i.e., it yields shorter lengths when more queries are allowed). The analogue of the main result for PIR yields a three-server scheme for n-bit long databases with communication $\exp(\widetilde{O}(\sqrt{\log n}))$, improving over Yekhanin's bound of $n^{1/30000000}$.

Constructing Low-Error 2-Query PCPs. Probabilistically Checkable Proofs (PCPs) are proofs that offer a trade-off between the number of locations inspected at random in the alleged proof and the statistical confidence in its validity. In the meeting, Irit Dinur presented a methodology for constructing (relatively short) PCPs in which verification is performed by two queries such that the error probability is inversely related to the length of the answers. The core of her new methodology is a new composition theorem that refers to "decodable PCPs" (a notion implicit in prior work). The resulting construction matches the parameters of the construction of Moshkovitz and Raz, but the current construction is significantly simpler. In contrast to prior results, her new constructions yield inapproximability results (for many natural optimization problems such as Max-Clique) in which approaching the "threshold of approximability" does not cause a deterioration in the complexity of the reduction.

Parallel Repetition of Interactive Protocols. It has been known for more than a decade that parallel repetition may fail to reduce the error in computationally-sound proof (a.k.a. argument) systems. In the meeting, Iftach Haitner presented a methodology for (slightly) modifying an interactive protocol such that parallel repetition does reduce the (observable) error in the *resulting* protocol. The modification amounts to having the verifier abort at random with probability 1/4 (i.e., after each round, the verifier aborts with probability $1/4r$, where r denotes the number of rounds). In case of abort, the verifier always accepts, which means that this modification increases the probability of error. The benefit of this modification is that the probability of cheating in the parallel execution is not sensitive to whether the verifier aborts in any typical individual copy, which establishes sufficient independence between the copies.

On the Best Possible Approximation of CSPs. Constraint Satisfaction Problems (CSPs) are specified by a finite set of finite predicates (e.g., 3-SAT is specified by the set of predicates on at most three variables that may be written as disjunctions of the corresponding literals). In the meeting, Prasad Raghavendra presented an approximation threshold result for any CSP, assuming the Unique

Game Conjecture (UGC). Specifically, for every CSP and every $\epsilon > 0$, there exists a polynomial-time algorithm that gets within a factor ϵ of the threshold beyond which approximation becomes UGC-hard. Furthermore, this seemingly optimal approximation threshold factor can be efficiently approximated.

New Notions of Computational Entropy. Omer Reingold and Salil Vadhan presented two complementary notions of "computational entropy" (a.k.a. pseudoentropy, akin to pseudorandomness which refers to distributions that are computationally indistinguishable from the uniform distribution on n-bit strings). The first notion, called next bit (or block) pseudoentropy measures the computational unpredictability of the next bit (given the previous bits). In contrast to the extreme case (of full unpredictability), in general next bit pseudoentropy does not yield the standard notion of pseudoentropy. Nevertheless, the new notion is instrumental for deriving an improved construction of pseudorandom generators based on any one-way function. The second notion, called inaccessible entropy, refers to the infeasibility of generating a next block that is as random as expected by an unbounded observer. For example, if some party sends a (statistically hiding) commitment to a random value, then when asked to reveal the value it can provide at most one possible value, whereas from the (unbounded) observer's point of view any value is possible. Indeed, the notion of inaccessible entropy is is related to statistically hiding commitment schemes, and is actually pivotal to their construction.

The Average-Case Complexity of k-Clique. For any constant k, constant-depth (unbounded fan-in) circuits of size $O(n^k)$ can distinguish n-vertex graphs having a clique of size k from graphs lacking such a clique. Ben Rossman's presentation addressed the average case complexity of this problem, where the input distribution corresponds to the standard random graph model with arbitrary edge density (which may be thought of as studying the problem at the threshold edge density, where the problem is not trivial). Interestingly, relatively tight lower and upper bounds, asserting that the size is $n^{(k/4)+\Theta(1)}$, can be obtained. The same holds when considering monotone circuits (of arbitrary depth).

Poly-Logarithmic Independence Fools AC^0 Circuits. Two decades ago, it was conjectured that poly-sized constant-depth circuits (of unbounded fan-in) cannot distinguish between any two poly-logarithmically independent distributions, and hence any such (poly-log independent) distribution is pseudorandom with respect to AC^0 circuits. Mark Braverman presented a proof of this conjecture. The proof combines two known approximation methods that yield different and incomparable approximations of AC^0 circuits by low degree polynomials.

Fast Polynomial Factorization and Modular Composition. Chris Umans presented an improved randomized algorithms for factoring univariate polynomials over a finite field. The source of the improvement is a new algorithm for modular composition of univariate polynomials that operates in nearly linear time. In the case of very big finite fields, the algorithm uses a sequence of reductions that first reduce to a multivariate problem, then lift to the integers, next reduces modulo

small primes, and finally applies a FFT. Most previous methods used only operations in the original field. As an interesting feature, the new method shows that (at the current state of knowledge) Boolean computations beat arithmetic ones for this algebraic problem.

Informal sessions. Besides the plenary formal program, intense interaction between the participants took place in smaller groups, as witnessed by the following list of afternoon or evening sessions.

- Structure problems and results on non-abelian groups (Wigderson)
- Polynomial identity testing (Shpilka)
- Compressing interactive communication (Rao)
- Security in steganography (Reischuk)
- Tutorial on group representations and matrix multiplication (Umans)
- Semantic communication (Sudan)
- On Smale's 17th problem (Cucker)
- Informal session of "going down hill" (Reingold)
- Semantic communication (Sudan)
- Open session on probabilistic proof systems, IP and PCP (Or Meir)
- Codes (Guruswami, Saraf, Kabanets, Kopparty, Impagliazzo)
- The Generalized Linial-Nisan Conjecture and BQP vs. PH (Aaronson)
- Sum of Squares (Koiran)
- Real polynomials for Boolean functions (Lovett, Beame)
- More on non-abelian groups (Wigderson)

The rest of this report. This report contains extended abstracts of the 15 long presentations as well as abstracts of 10 short communications.