CSE 533: The PCP Theorem and Hardness of Approximation **(Autumn 2005)**

Lecture 20: Course summary and open problems

Dec. 7, 2005

Lecturer: Ryan O'Donnell *Scribe: Ioannis Giotis*

Topics we discuss in this lecture:

- Quick recap of tools, sister subjects
- Best known PCP systems
- Open problems in PCPs
- Open problems in hardness of approximation

1 Tools, theorems and related subjects

- Two tools that we used in the course that are of significant utility in many other areas of theoretical computer science are expanders and Fourier analysis.
- As for explicit theorems to take away from the class, the following three are the most useful:
  - **PCP theorem.** The classic PCP theorem is most useful for proving that optimization problems have no PTAS.
  - **E3LIN2.** The $1-\epsilon$ vs $1/2-\epsilon$ hardness of this problem is most useful for proving that problems have no approximation beyond a fixed constant factor; i.e., for statements like "such-and-such maximization problem has no $\frac{77}{78}$-approximation unless $\mathrm{P} = \mathrm{NP}$".
  - **Raz's Label Cover.** The hardness of this problem is often useful for showing strong or super-constant hardness-of-approximation results.
- Other areas on theoretical computer science that are related to PCPs and hardness of approximation:
  - Property testing.
  - Coding theory.
  - Approximation algorithms.
  - Fourier analysis.
  - Structural complexity.
  - Cryptography.
  - Pseudorandomness.

2 Best known PCP systems for NP

We summarize here some best known PCP systems with respect to various parameters.

- **Size:** Although in class we only cared about the witness being of polynomial size, we can take a closer look at PCP length in terms on n, the length of the classical proof for, say, 3SAT. The best known result with constantly many queries is due to Dinur[6] building on work of Ben-Sasson and Sudan [3]; the proof length can be $O(n \cdot \text{polylog}(n))$ bits. (Note: if we want PCPs for all of NTIME(n), we already suffer at least a $O(\log n)$ blowup in witness size in the best known version of Cook's theorem.)

- **Verifier time:** Normally in PCP systems, the verifier reads the entire input (statement), does polynomial time computation, and then queries only a sublinear number of bits of the proof. We may hope to have a verifier that also runs in sublinear time; however it then has no time to read the entire proof statement. However, if we are satisfied with Assignment Testers rather than PCPs, we can hope to run in sublinear time. The best known result along these lines is due to Ben-Sasson, Goldreich, Harsha, Sudan, and Vadhan [4] and gives a version of the previous result that also has polylogarithmic verifier running time.

- **Soundness with $O(1)$ queries to large alphabets:** Consider PCPs for NP with $O(\log n)$ randomness and constantly many queries over a large alphabet. In particular, when the alphabet size is m, we might hope to get soundness $\leq 1/m^{\Omega(1)}$. This seems harder and harder to achieve as m grows larger. The best result along these lines is due to Dinur, Fischer, Kindler, Raz, and Safra [7], who show such a result with m as large as $2^{\log^{1-\gamma} n}$. This holds for every $\gamma > 0$, although the smaller γ is, the more $O(1)$ queries must be made.

- **Soundness with 2 queries:** If we insist on using only 2 queries, the best result along the above lines is Raz's Parallel Repetition result [20], which works for every $m = O(1)$.

- **Soundness with q queries to bits:** Finally, consider the problem of achieving the lowest possible soundness while making q queries to a proof written in bits. Certainly we can't hope for soundness lower than 2^{-q}. The basic PCP theorem gives us $2^{-\Omega(q)}$. A recent algorithm for MAX-qCSP due to Hast [13] in fact shows we can't have soundness lower than $O(\frac{q}{\log q})2^{-q}$. The (essentially) best result known is from Samorodnitsky and Trevisan [21], which gives soundness $2^{2\sqrt{q}}2^{-q}$ (Engebretsen and Holmerin later got the 2 inside the square root). A recent work of the same authors [22] shows that, assuming the Unique Games Conjecture, this can be lowered to $(q+1)2^{-q}$ for infinitely many q.

3 Open problems in PCPs

Here is a brief list of important open problems in the theory of PCPs:

- Prove or disprove the Unique Games Conjecture. Currently there is no strong evidence for either possibility.

- Improve the soundness result in [**?**] to $m = poly(n)$. This is also known as the "sliding scale conjecture", stated by Bellare, Goldreich and Sudan [2] in 1995.

- Give linear size PCPs for $3SAT$ or prove size lower bounds.

- (Less important.) What is the best possible soundness for a PCP for NP which is nonadaptive, has perfect completeness, and queries only 3 bits. I.e., for NP $\subseteq$ na-PCP$_{1,s}[3, O(\log n)]$, how small can s be? The best published result is by Håstad [14], $s = 3/4 + \epsilon$. An unpublished work of Khot and Saket claims $s < 3/4$. Zwick [24] gave an approximation algorithm for MAX-3CSP on satisfiable instances which satisfies a $5/8$ fraction of all clauses; he conjectures this is best possible, in which case $s = 5/8 + \epsilon$ is the best that could be achieved.

4 Open problems in hardness of approximation

There are still very many open problems in hardness of approximation; not every result is as sharp as the $7/8$-algorithm and $(7/8 + \epsilon)$-hardness of MAX-3SAT. Here we present a small selection of interesting such problems:

MAX-2LIN(2). The best unconditional hardness result known is $11/12 + \epsilon$, due to Håstad [14] using gadgets of Trevisan, Sorkin, Sudan and Williamson [23]. Assuming the Unique Games Conjecture, we get the same hardness as in MAX-CUT, $.878 + \epsilon$, where $.878$ is the Goemans-Williamson approximation factor. The GW algorithm also works for MAX-2LIN(2).

MAX-2SAT. The best unconditional hardness result is $21/22 + \epsilon$, due to the same authors as in MAX-2LIN(2). Assuming UGC, the hardness is $\beta + \epsilon$, where β is a certain trigonometric quantity equal to about $.943$ [18]. The best known algorithm, due to Lewin, Livnat and Zwick [19], achieves a factor of .9401. [18] conjecture that their factor β is optimal.

The question of $(1 - \epsilon)$-satisfiable instances for MAX-2SAT is also interesting. The best algorithm, due to Zwick [25], finds a $(1 - \Theta(\epsilon^{1/3}))$-satisfying assignment. On the other hand, the best known UGC-hardness result [18] shows that $(1 - \Theta(\epsilon^{1/2}))$ is hard. Unconditionally, only $(1 - \Theta(\epsilon))$ is known hard.

Min-Vertex-Cover. Given an undirected graph, this is the problem of finding as small a set of vertices as possible that touches every edge. There is a very easy greedy 2-approximation algorithm. Subject to UGC, this is best possible — Khot and Regev [16] give a $(2 - \epsilon)$ UGC-hardness result. The best unconditional hardness result known is $10\sqrt{5} - 21 + \epsilon \approx 1.36 + \epsilon$, in a very interesting paper of Dinur and Safra [9].

Coloring 3-colorable graphs. Given a graph, promised to be 3-colorable, this is the problem of coloring it with as few colors as possible. Coloring with 3 colors is of course NP-hard. Coloring with 4 colors is also known to be NP-hard, first proved using PCP technology by Khanna, Linial and Safra [15], and then later simplified to just a gadget (no PCPs) in [12]. The best known algorithm uses $n^{3/14}$ colors; this is due to Blum and Karger [5]. Assuming a Unique Games-like conjecture, Dinur, Mossel and Regev [8] have shown that using $O(1)$ colors in NP-hard.

Min-Feedback-Arc-Set. This is the problem of, given a directed graph, delete as few arcs as possible so that it becomes acyclic. This is known to have the same NP-hardness of approximation as Vertex-Cover (so 1.36-hardness is known). The best algorithm achieves a ratio of $O(\log n \log \log n)$ (Even, Naor, Schieber and Sudan [10]).

Sparsest-Cut. Equivalently, "Min-Edge-Expansion": Given an undirected graph, minimize $|E(S, \bar{S})|/|S|$ over all sets $S \subset V$, $|S| \leq |V|/2$. No hardness of approximation is known for this problem; i.e., it might have a PTAS. The best approximation algorithm achieves a factor of $O(\sqrt{\log n})$; this is from the notable paper of Arora, Rao and Vazirani [1].

Min-Bisection. This is the same as MIN-CUT, except you are required to produce a cut that partitions the graph into exactly equal parts ($n/2$ vertices on both sides). The best approximation algorithm has factor $O(\log^{3/2} n)$, due to Feige and Krauthgamer [11]. The best hardness result is an extremely weak one, due to Khot [17]: There is no $(1+\epsilon)$-approximation algorithm unless NP is contained in DTIME($2^{n^{\epsilon'}}$).

References

[1] S. Arora, S. Rao, and U. V. Vazirani. Expander flows, geometric embeddings and graph partitioning. In *STOC*, pages 222–231, 2004.

[2] M. Bellare, O. Goldreich, and M. Sudan. Free bits, PCPs, and nonapproximability — towards tight results. *SIAM J. Comput.*, 27(3):804–915, 1998.

[3] Ben-Sasson and Sudan. Simple PCPs with poly-log rate and query complexity. In *STOC: ACM Symposium on Theory of Computing (STOC)*, 2005.

[4] E. Ben-Sasson, O. Goldreich, P. Harsha, M. Sudan, and S. Vadhan. Short PCPs verifiable in polylogarithmic time. In *Proc. 20th IEEE Conference on Computational Complexity*, pages 120–134, San Jose, California, 12–15 June 2005.

[5] A. Blum and D. R. Karger. An $\tilde{O}(n^{3/14})$-coloring algorithm for 3-colorable graphs. *Inf. Process. Lett*, 61(1):49–53, 1997.

[6] I. Dinur. The pcp theorem by gap amplification. ECCC, TR05-046, 2005.

[7] I. Dinur, E. Fischer, G. Kindler, R. Raz, and S. Safra. PCP characterizations of NP: Towards a polynomially-small error-probability. In *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing (STOC'99)*, pages 29–40, New York, May 1999. Association for Computing Machinery.

[8] I. Dinur, E. Mossel, and O. Regev. Conditional hardness for approximate coloring, Apr. 13 2005.

[9] I. Dinur and S. Safra. On the hardness of approximating minimum vertex cover. *Annals of Mathematics*, 162(1), 2005.

[10] G. Even, J. Naor, B. Schieber, and M. Sudan. Approximating minimum feedback sets and multicuts in directed graphs. *Algorithmica*, 20(2):151–174, 1998.

[11] U. Feige and R. Krauthgamer. A polylogarithmic approximation of the minimum bisection. *SIAM J. Comput.*, 31(4):1090–1118, 2002.

[12] V. Guruswami and S. Khanna. On the hardness of 4-coloring a 3-colorable graph. *SIAM J. Discrete Math*, 18(1):30–40, 2004.

[13] G. Hast. Beating a random assignment. In *APPROX-RANDOM*, pages 134–145, 2005.

[14] J. Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.

[15] S. Khanna, N. Linial, and S. Safra. On the hardness of approximating the chromatic number. *Combinatorica*, 20(3):393–415, 2000.

[16] Khot and Regev. Vertex Cover might be hard to approximate to within $2-\epsilon$. In *Annual IEEE Conference on Computational Complexity*, volume 18, 2003.

[17] S. Khot. Ruling out PTAS for graph min-bisection, densest subgraph and bipartite clique. In *FOCS*, pages 136–145, 2004.

[18] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for max-cut and other 2-variable CSPs? In *FOCS*, pages 146–154, 2004.

[19] M. Lewin, D. Livnat, and U. Zwick. Improved rounding techniques for the MAX 2-SAT and MAX DI-CUT problems. In *IPCO*, pages 67–82, 2002.

[20] R. Raz. A parallel repetition theorem. *SIAM J. Comput*, 27(3):763–803, 1998.

[21] A. Samorodnitsky and L. Trevisan. A PCP characterization of NP with optimal amortized query complexity. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing, STOC'2000 (Portland, Oregon, May 21-23, 2000)*, pages 191–199, New York, 2000. ACM Press.

[22] A. Samorodnitsky and L. Trevisan. Gowers uniformity, influence of variables, and PCPs, Oct. 12 2005.

[23] L. Trevisan, G. B. Sorkin, M. Sudan, and D. P. Williamson. Gadgets, approximation, and linear programming. *SIAM J. Comput.*, 29(6):2074–2097, 2000.

[24] U. Zwick. Approximation algorithms for constraint satisfaction problems involving at most three variables per constraint. In *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA-98)*, pages 201–210, New York, Jan. 25–27 1998. ACM Press.

[25] U. Zwick. Finding almost-satisfying assignments. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC-98)*, pages 551–560, New York, May 23–26 1998. ACM Press.