



KTH Computer Science
and Communication

Homework VIII, Theoreticians Toolkit 2009/2010

Due on Tuesday March 30 at 15.15 (send an email to osven@kth.se). Solutions to many homework problems, including problem one on this set, is available on the internet, either in exactly the same formulation or with some minor perturbation. It is *not acceptable* to copy such solutions. It is hard to make strict rules on what information from the internet you may use and hence whenever in doubt contact Ola Svensson. You are, however, allowed to discuss problems in groups with up to three students, but solutions should be handed in individually.

- 1 (25 p) Johan proved that there is a particularly simple PCP verifier for SAT. For every $\epsilon > 0$ it uses the $O(\log n)$ random bits to compute three positions in the proof, say i, j and k , and a bit b , and accepts iff

$$y(i) + y(j) + y(k) = b \pmod{2}.$$

Here $y(i)$ is the i 'th bit in the proof y . The verifier has completeness $1 - \epsilon$ and soundness $1/2 + \epsilon$. In other words,

- If ϕ is a satisfiable SAT instance then there is a proof y that makes the verifier accept with probability $1 - \epsilon$.
- If ϕ is a not satisfiable SAT instance then for any proof the verifier accepts with probability at most $1/2 + \epsilon$.

Use the above described PCP verifier to show the vertex cover problem¹ is NP-hard to approximate within a factor of c , for some c that you wish to maximize.

- 2 (30p) Prove that the following two statements are equivalent.

(A) $NP = PCP[\log n, 1]$

(B) There is a constant $\epsilon > 0$ so that it is NP-hard to distinguish whether a 3-SAT formula is satisfiable or any assignment satisfies at most a fraction $1 - \epsilon$ of the clauses.

¹Vertex Cover is the problem where we are given an undirected graph $G(V, E)$ and we wish to find a subset $V' \subseteq V$ of minimum cardinality such that each edge is covered. An edge $\{u, v\} \in E$ is covered if $u \in V'$ or $v \in V'$.

3 (30p) In class we used the technique called “Dual-fitting” to show that the greedy algorithm for Set Cover gives a H_n approximation. One advantage of the LP approach is that it is fairly easy to generalize. One such generalization of the Set Cover problem is

- **Multiset Multicover:** Each element, e , needs to be covered a specified integer number r_e of times. Further, we are given collection of multisets, rather than sets, of U . A multiset S contains a specified number $M(S, e)$ of copies of each element e in S . We make the reasonable assumption that the instance satisfies $\forall S, e : M(S, e) \leq r_e$. The objective again is to cover all elements up to their coverage requirements at minimum cost. We will assume that the cost of picking a set S k times is k times the cost of picking the set once.

Your task is to first improve the analysis for greedy Set Cover and then extend it to Multiset Multicover.

3a (10p) Show that the dual-fitting-based analysis for the greedy Set Cover algorithms actually establishes an approximation guarantee of H_k , where k is the size of the largest set in the given instance.

3b (20p) Give an algorithm for Multiset Multicover that finds a solution of cost at most H_m times the optimal cost, where m is the size of the largest multiset in the given instance (the size of a multiset counts elements with multiplicity).