Approx. Algo. HW2

Justin Yip (justin)

September 19, 2008

1 Answer for 3.3

1.1 Set cover problem

Input: a collection of subsets of a universe $U = \{1, .., n\}$, $\mathcal{S} = \{S_1, \ldots, S_k\}$ and a cost function $c : \mathcal{S} \rightarrow \mathcal{Q}^+$.

Output: a subcollection of $\mathcal{S}$ that covers all elements of U

Obj: minimize the cost of the subcollection.

1.2 Construction

We can reduce it to a Directed Steiner Tree (DST) problem.

We construct an instance of the DST problem with a graph $G = (V, E)$ and a required set $R \subseteq V$. V is partitioned into three layers. The first layer V_1 contains only the root node r. The second layer V_2 contains k nodes $v_1, \ldots, v_k$, a node v_i corresponds to the set S_i in the collection $\mathcal{S}$. The third layer V_3 contains n nodes, $v_{-1}, \ldots, v_{-n}$, the node v_{-i} corresponds to the element i in U.

The required set R is the union of the first and third layer.

For every node v_i in the second layer, there is an directed edge from the root r to it, the weight of the edge is $c(S_i)$. For every element j in the set S_i, there is a zero-weight edge from v_i to v_{-j}.

1.3 Correctness

Here we prove an solution in the set cover problem corresponds to an solution in the DST problem.

$\Rightarrow$ An solution in the set cover problem is a sub-collection C that covers all elements in the universe. Here we show the output $T = (V', E')$ of the DST problem with input $G = (V, E), R$ has the same cost. For each set S_i in C, the edge (r, v_i) is included in T , correspondingly, all edges from v_i to the third layers are also included, hence r is covered. Since the union of all sets in C is the universe U, by the construction of the graph G, all nodes in the third layer are also covered. Additionally, as there is no cycle in G, T has no cycle as wel. Hence, T is a solution.

$\Leftarrow$ An output $T = (V', E')$ of the DST problem with input $G = (V, E), R$ is a tree that contains all required nodes R. If a second layer node v_i is in T, the set S_i belongs to the output C of the set cover problem. And by the construction, it is trivial that that the union of the sub-collection C covers all elements in the universe.

1.4 Approximation factor preservation

Here we show that the optimal value of the set cover problem is equals to the optimal value of the DST problem.

In the set cover problem, each set S_i incur a cost of $c(S_i)$, it corresponds to the cost of the edge (r, v_i) in the DST problem. Since edges between the second and third layer has zero cost, the cost of the tree equals to the cost of the sub-collection.

Since $O(\log n)$ factor approximation is the best one can hope for, from the above approximation factor preserved reduction, we show that it is also unlikely to achieve an approximation algorithm with guarantee better than $O(\log n)$.