CS-E400204: Approximation Algorithms	**Spring (Fifth Period) 2022**

Assignment 2 (April 28): From Steiner Trees to Set Cover

Due Date: May 5 — *Due Time: 23:59 pm*

Exercise 1 - Greedy for Multiway Cut

A natural greedy algorithm for calculating a multiway cut for given terminals $s_1, \ldots, s_k$ is the following: Based on G, calculate cheapest s_i–s_j cuts for all pairs s_i, s_j that are still connected and remove the cheapest of these cuts. Repeat until all pairs s_i, s_j are separated.

Show that this algorithm has approximation ratio $2 - 2/k$.

[1 point]

Exercise 2 - Factor-f approximation for unweighted Set Cover

Let $(U, \mathcal{S})$ be an unweighted Set Cover instance in which each element from U is contained in at most f sets from $\mathcal{S}$. We refer to f as *frequency.* (For frequency $f = 2$, we obtain Vertex Cover.)

Consider the following iterative algorithm:

Algorithm 1 SimpleGreedySetCover(U, $\mathcal{S}$)

$\mathcal{S}' \leftarrow \emptyset$
while $U \neq \emptyset$ **do**
  mark arbitrary element $u \in U$
  $\mathcal{R} \leftarrow \{S \in \mathcal{S} \mid u \in S\}$
  $\mathcal{S}' \leftarrow \mathcal{S}' \cup \mathcal{R}$
  $U \leftarrow U \setminus \bigcup \mathcal{R}$ *// remove all elements from U that are covered by $\mathcal{R}$*
end while
return $\mathcal{S}'$

Show that this algorithm yields a factor-f approximation, which generalizes the factor-2 approximation for Vertex Cover from the lecture.

[1 point]

Exercise 3 - **Maximum k-Coverage** *(Bonus)*

In this exercise, we consider the Maximum k-Coverage problem, which is a variant of the Set Cover problem introduced in the lecture. Let U be a set, $\mathcal{S}$ be a set of subsets of U, and k be a natural number. The problem Maximum k-Coverage asks for a subset $\mathcal{S}' \subseteq \mathcal{S}$ with $|S'| = k$, such that the number $|\bigcup \mathcal{S}'|$ of elements covered by $\mathcal{S}'$ is maximized.

For this problem, we consider a greedy algorithm, which iteratively performs the following step: In each iteration, select a set from $\mathcal{S}$ that contains the largest number of uncovered elements. After k iterations, the algorithm returns the set of sets selected this way.

In the following, we analyze the quality of this algorithm. Let ALG_i be the number of elements that are covered by the sets that have been selected in the steps up to and including step $i \in \{1, \ldots, k\}$. We set $\mathrm{ALG}_0 = 0$. Accordingly, the total number of covered elements (i.e. after k iterations) is ALG_k.

a) Show that $\mathrm{ALG}_1 \geq \frac{1}{k} \cdot \mathrm{OPT}$, i.e. that the greedy algorithm covers at least $\frac{1}{k} \cdot \mathrm{OPT}$ elements with the set selected in the first step.

b) Show by induction that for each $i \in \{0, \ldots, k\}$, the following holds: $\mathrm{OPT} - \mathrm{ALG}_i \leq \left(1 - \frac{1}{k}\right)^i \cdot \mathrm{OPT}$.

c) Show that the greedy algorithm has approximation ratio $\left(1 - \frac{1}{e}\right) \approx 0.63$.

[2 bonus points]