CS 466/666 Fall 2018 due Tuesday November 6, 5 PM

A. Lubiw

ASSIGNMENT 5

DO NOT COPY. ACKNOWLEDGE YOUR SOURCES.

1. [5 marks] Consider the following randomized algorithm to find the minimum of n distinct integers. Take a random order $x_1, \ldots, x_n$ of the input numbers. Set $m = \infty$. For $i = 1 \ldots n$, if $x_i < m$ then update m to x_i.

   Use backwards analysis to argue that the expected number of times you update m is $O(\log n)$. (You may use without proof the fact that the Harmonic numbers grow as $O(\log n)$.)

2. [10 marks] This problem is about a weighted version of the Set Cover problem. The input consists of: a set E of n elements, where each $e \in E$ has a non-negative weight $w(e)$; a collection $S_1, \ldots, S_k$ where $S_i \subseteq E$; and a number t. The problem is to pick t of the subsets S_i to maximize the sum of the weights of the elements covered.

   The goal of this question is to show that the obvious greedy algorithm has approximation factor $1 - \frac{1}{e}$ where e is Euler's constant. The greedy algorithm first picks a set that covers the maximum weight of elements, then deletes those elements and repeats until t sets have been chosen. In other words, at each step the algorithm picks a set that has the maximum weight of uncovered elements. Note that ties are broken arbitrarily.

   Let W^* be the weight of the elements covered by the optimum solution. Let W_i, for $i = 1, \ldots t$, be the weight of the elements covered by the first i sets of the greedy algorithm. Thus W_t is the final weight of the greedy solution.

   (a) [2 marks] (Warm up questions to gain intuition.) Find an example with $t = 2$ and unit weights where $W_2 = \frac{3}{4}W^*$. Explain why you cannot have an example with $t = 2$ and unit weights where $W_2 \leq \frac{1}{2}W^*$.

   (b) [4 marks] Prove that $W_1 \geq W^*/t$. More generally, prove that $W_i - W_{i-1} \geq (W^* - W_{i-1})/t$.

   (c) [4 marks] Prove by induction that $W_i \geq (1 - (1 - \frac{1}{t})^i)W^*$.

   (d) [0 marks] From part (c), the weight of the greedy solution, W_t, is at least $(1-(1-\frac{1}{t})^t)W^*$. Since $\lim_{t\to\infty}(1-(1-\frac{1}{t})^k t) = 1-\frac{1}{e}$ and $1-(1-\frac{1}{t})^t$ is decreasing, thus $1-(1-\frac{1}{t})^t \geq 1-\frac{1}{e}$, so the approximation ratio of the greedy algorithm is $1 - \frac{1}{e}$.

   In fact $1 - \frac{1}{e}$ is the best approximation factor possible for this problem (unless P=NP).