

Homework 7

Released 12/2/2016

Due 8:00pm 12/14/2016 in Moodle

Note: *LaTeX* template courtesy of UC Berkeley EECS dept.

Instructions. You may work in groups, but you must individually write your solutions yourself. List your collaborators on your submission.

If you are asked to design an algorithm as part of a homework problem, please provide: (a) the pseudocode for the algorithm, (b) an explanation of the intuition for the algorithm, (c) a proof of correctness, (d) the running time of your algorithm, and (e) justification for your running time analysis.

Submission instructions. This assignment is due by 8:00pm on 12/14/2016 in Moodle. **Note the unusual submission deadline.** Please submit a pdf file. You may submit a scanned handwritten document, but a typed submission is preferred.

1. **Max-Cut (20 points).** In this problem we will prove that the MAX-CUT problem is NP-Complete. This is surprising since as we saw in lecture, MIN-CUT can be solved in polynomial time using Network Flows.

MAX-CUT is the following problem: Given an undirected graph $G = (V, E)$ with nonnegative edge capacities $w_{u,v}$ for $(u, v) \in E$ and a number c , decide if there exists a cut in G with capacity at least c . Recall that a cut is a set of vertices $S \subset V$ and the capacity of the cut is $\sum_{(u,v), u \in S, v \notin S} w_{u,v}$.

We will also need to describe a variant of SAT that will be a useful intermediate problem for the reduction.

NAE- k -SAT is the following problem: Given a boolean formula where each clause has exactly k terms, decide if there is an assignment ϕ such that each clause has at least one true term *and* one false term?

- (a) Prove that 3-SAT \leq_P NAE-3-SAT (Hint: you may want to go through NAE-4-SAT).
- (b) Prove that NAE-3-SAT \leq_P MAX-CUT.

2. **3-Coloring (K&T Ch13.Ex1) (10 points).** Consider the optimization version of 3-COLORING: Given a graph $G = (V, E)$, color each node with one of three colors to maximize the number of edges whose incident vertices have different colors. We say that an edge (u, v) is satisfied if the colors assigned to u and v are different.

Suppose that c^* is the maximum number of satisfied edges. Give a polynomial time algorithm that produces a 3-coloring that satisfies $\frac{2}{3}c^*$ edges. If your algorithm is randomized, the expected number of satisfied edges should be at least $\frac{2}{3}c^*$.

3. **Randomized Approximate Median (K&T Ch13.Ex15) (20 points).** Given a set $S = \{a_1, \dots, a_n\}$ of n distinct numbers, we say that x is an ϵ -approximate median if at least $(1/2 - \epsilon)n$ elements of S are less than x and $(1/2 + \epsilon)n$ elements of S are greater than x .

Here is an algorithm for approximate median: Select a subset $S' \subset S$ uniformly at random, compute the median of S' and return that as an approximate median of S .

Show that there is an absolute constant c , independent of n , such that if your sample S' consists of c elements ($|S'| = c$) then with probability at least 0.99 the algorithm returns a (0.05)-approximate median.

4. **Representative Sets (K&T Ch11.Ex2) (20 points)** We are given a set of n points P that are related by a (symmetric) distance function $d(\cdot, \cdot)$ where $d(p, q) \geq 0$ is the distance between $p, q \in P$.

We also have a parameter Δ that is used as a cutoff for *similarity*: two points p, q are deemed similar if $d(p, q) \leq \Delta$.

A subset $S \subset P$ is *representative* if for every $p \in P$, there is an $s \in S$ with $d(p, s) \leq \Delta$. We'd like to find a representative set S that is as small as possible.

- (a) Design a polynomial time algorithm that approximates the minimum size representative set to within a $O(\log n)$ factor. This means that if the minimum size representative set has size s^* , your algorithm should return a representative set of size $O(s^* \log n)$.
 - (b) This problem looks familiar to the Center Selection (or k -Center) problem from lecture. Why doesn't the k -center approximation algorithm solve this problem?
5. **Hitting Sets (K&T Ch8.Ex5 and K&T Ch11.Ex4) (30 points).** Given a set $A = \{a_1, \dots, a_n\}$ and a collection $B_1, \dots, B_m \subset A$ we say that $H \subset A$ is a *hitting set* for the collection if H contains at least one element from each B_i , that is $|H \cap B_i| > 0$.

The HITTING-SET problem is the following: Given a set $A = \{a_1, \dots, a_n\}$, subsets $B_1, \dots, B_m \subset A$, and a number k , is there a hitting set $H \subset A$ of size at most k .

- (a) **(10 points)** Prove that HITTING-SET is NP-Complete.
- (b) **(20 points)** In the optimization version of HITTING-SET, the goal is to find a hitting set $H \subset A$ with minimum size $|H|$. Let $b = \max_i |B_i|$ denote the maximum size of any of the target sets B_i . Design a polynomial-time approximation algorithm for the hitting set optimization problem whose total size is at most b times the minimum possible.