

CS/ECE 374 A (Spring 2020)
Homework 10 (due Apr 23 Thursday at 10am)

Instructions: As in previous homeworks.

Problem 10.1: We are given n intervals $[a_1, b_1], \dots, [a_n, b_n]$, which are pairwise disjoint (i.e., non-overlapping). The intervals are not given in sorted order. We are also given a number k (which could be less than or greater than n). Your objective is to pick k numbers p_1, \dots, p_k , such that each p_i lies in $[a_1, b_1] \cup \dots \cup [a_n, b_n]$, so as to maximize the separation

$$\Delta := \min_{i,j: i \neq j} |p_i - p_j|.$$

(In these times, social distancing is important!)

For example: for the intervals $[1, 4], [5, 8], [11, 12], [13, 17], [19, 23]$ and $k = 4$, one solution is to pick the numbers $1, 8, 15, 23$, which has separation $\Delta = 7$ (this appears to be optimal). Note that some intervals may contain none of the numbers you pick; on the other hand, some intervals could contain multiple numbers.

- (a) (7.5 points) First consider the decision problem: decide whether there is a solution with separation at least r , for a given value r (and if so, output one such solution). Give an efficient greedy algorithm to solve this decision problem. Analyze its running time as a function of n and k .
(*Important Note:* for greedy algorithm, you must give a full proof of correctness.)
- (b) (2.5 points) Using (a), give an algorithm (the fastest you can achieve) to solve the original problem. You may assume that all a_i and b_i are integers in $\{1, \dots, U\}$, and you may analyze the running time in terms of n , k , and U .

Problem 10.2: We consider one formulation of a clustering problem where we want to divide a dataset into two clusters. Specifically, given a set P of n points, we want to partition P into two subsets P_1, P_2 to minimize the following cost function:

$$c(P_1, P_2) = \max\{\text{diam}(P_1), \text{diam}(P_2)\},$$

where (in the corrected version)

$\text{diam}(P_i)$ is the smallest value r such that for every $p, q \in P_i$, there is a path from p to q where every edge uv in the path has $d(u, v) \leq r$.

Give an algorithm to solve this problem in $O(n^2)$ time. Prove correctness of your algorithm.
(*Hint:* consider the largest MST edge...)

Added bonus: 2.0 bonus points if you also have a correct algorithm with $O(n^2)$ or $O(n^2 \log n)$ running time for the original version of the problem with $\text{diam}(P_i) = \max_{p,q \in P_i} d(p, q)$ instead.

Problem 10.3: For each of the following languages, either prove that it is undecidable, or describe an algorithm that decides this language. (Note that you cannot use Rice Theorem in solving this problem.)

- (a) $\{\langle M, N \rangle \mid L(M) \cap L(N) \text{ is context-free, where } M \text{ is a Turing machine and } N \text{ is an NFA}\}$
- (b) $\{\langle M, R \rangle \mid L(M) \subseteq L(R), \text{ where } M \text{ is a DFA and } R \text{ is a regular expression}\}$
- (c) $\{\langle M \rangle \mid \text{for every } i \geq 1, M \text{ halts on exactly two strings of length } i, \text{ where } M \text{ is a Turing machine}\}$