

Exercises 2

Handed Out: Apr. 8

Due: Apr. 15

1. Suppose you are given a collection A of n points on the real line. The problem is to find a **smallest collection S of unit-length intervals** that cover every point in A .

(Another way to think about this is the following. The points of A are the times at which trains arrive at a station. When a train arrives there must be someone manning the station. Due to union rules, each employee can work at most **one hour** at the station. The problem is to find a scheduling of employees that covers all the times in A and uses the smallest number of employees.)

- (a) Prove or disprove that the following algorithm correctly solves this problem.

Let I be the interval that covers the most number of points in A . Add I to the solution set S . Then recursively continue on the points in A not covered by I .

- (b) Prove or disprove that the following algorithm correctly solves this problem.

Let a_j be the smallest (leftmost) point in A . Add the interval $I = [a_j, a_{j+1}]$ to the solution set S . Then recursively continue on the points in A not covered by I .

2. You are given n files of length $\ell_1, \ell_2, \dots, \ell_n$ to be stored on a tape, where the i th file is accessed with **probability** p_i . Assume that each retrieval starts with the tape completely rewound, so the cost of each retrieval equals the total length of the preceding files in tape plus the length of the retrieved file. For example, if the files are put in the order $1, 2, \dots, n$, then the expected total cost of retrieval is

$$p_1\ell_1 + p_2(\ell_1 + \ell_2) + \dots + p_n(\ell_1 + \ell_2 + \dots + \ell_n)$$

Given the file lengths and their access probabilities, the problem is to determine the order in which the files should be stored on the tape to **minimize** the total expected retrieval cost.

For each of the following three algorithms, either give a proof that the algorithm correctly finds the optimal order, or give an example where it finds a sub-optimal order.

- (a) Order the files from shortest to longest on the tape. That is, if $\ell_i < \ell_j$, then i is stored before j .
- (b) Order the files from most-probable to least-probable. That is, if $p_i < p_j$, then i is stored after j .
- (c) Order the files from the smallest ratio of length to probability to the largest ratio. That is, if $\frac{\ell_i}{p_i} < \frac{\ell_j}{p_j}$, then i is stored before j .