

Handout 5: Homework 3

Instructor: Susan Hohenberger

TA: Ming Chuang

This assignment is due by the start of lecture on September 30, 2009. Please clearly indicate your collaborators.

1. (30 points) You are commissioned to design a new busline for the Baltimore-Washington Parkway, which connects Baltimore and Washington DC. There are n stops on the North-bound bus line. Commuters may begin their journey at any stop i and end at any other stop $i < j$. There are some obvious options:
 1. You can have one bus run from the southern-most point to the northern-most point as a traditional busline might run. The system would be cheap because it only requires n segments for the entire system. However, a person traveling from $i = 0$ to $j = n$ must travel through all n segments. This system will be slow.
 2. You can have a special bus run from every point to every other destination. No person will ever wait through any unnecessary segments no matter where they start and end. This system requires $\Theta(n^2)$ route segments and will be expensive.

Suggest a compromise solution. Specifically, your task is to:

- (a) Use a divide-and-conquer technique to design a bus system that uses $\Theta(n \lg n)$ route segments and which requires a person to wait through at most 1 extra segment when going from any i to any j (as long as $i \leq j$, i.e., we only consider North-bound routes for simplicity). In other words, a commuter can travel from any i to any j by using at most 2 route segments.
- (b) Informally argue that your solution is correct (can travel from any i to j in 2 segments).
- (c) Analyze your algorithm and argue that it indeed takes $\Theta(n \lg n)$ route segments.
2. (20 points) Use indicator random variables to solve the following problem. Suppose the professor asks the class to help grade their own homeworks. Each of n students hand their homework to the professor. Then, the professor distributes the homeworks in a random order with each student grading one of them. What is the expected number of students who grade their own homework?
3. (20 points) Boris is your own personal robot. He pours your cereal (although he sometimes spills it); he cleans your bathroom (although he sometimes leaves the water running); and most recently, he alphabetically sorted your gigantic collection of n DVDs. Given Boris' track record, you have doubts that the DVDs are in perfect sorted order. To check that they are sorted takes $\Omega(n)$ time.

You are in a hurry. So to test if your DVDs (in A) are sorted, you execute the following algorithm. Repeat k times: Choose i independently and uniformly at random such

that $1 < i < n$. Compare $A[i]$ with $A[i-1]$ and $A[i+1]$. If these DVDs are not sorted correctly, output false. If all k tests succeed, output true.

Is this a good approach? Clearly, if A is sorted, this algorithm will always output true. What if A is not sorted? Let's call an array *almost* sorted if it is possible to remove 10% (or less) of the items and have the resulting array be sorted. Suppose poor Boris did not even get A almost sorted. Will the above algorithm detect this fact with probability at least $3/4$ for some $k = o(n)$? Justify your answer. (Hint: Can you think of a sequence that is not almost sorted, but with only a small number of elements that will cause the algorithm to return false?)

4. (30 points) CLRS 9-2 [weighted median] subproblems (a), (b) and (c).