CS:5350 Homework 9, Spring 2016

Due in class on Tue, Apr 26

Collaboration: You are welcome to form groups of size 2 and work on your homeworks in groups. Of course, you are not required to work in groups. Every group should make one submission and names of both group members should appear on the submission and both students in a group will receive the same score. Other than the TA and the professor, you can only discuss homework problems with your group partner. Collaboration can be positive because talking to someone else about these problems can help to clarify your ideas and you will also (hopefully) get to hear about different ways of thinking about the problem. On the other hand, collaboration can be negative if one member of the group rides on work being done by the other member – please avoid this situation. If your solutions are (even partly) based on material other than what has been posted on the course website, you should clearly acknowledge your outside sources.
Late submissions: No late submissions are permitted. You will receive no points for your submission if your submission is not turned in at the beginning of class on the due date.
Evaluation: Your submissions will be evaluated on correctness *and* clarity. Correctness is of course crucial, but how clearly you communicate your ideas is also quite important.

1. You are given an array $A[1..n]$ of numbers. Suppose that the elements in A are all distinct from each other. Your task is to partition A into two subarrays, let us call these B and C, such that: (i) $|B|, |C| \leq 2n/3$ and (ii) every element in B is less than every element in C. Here is a simple randomized algorithm for this problem:

   Step 1: Pick an index i from $\{1, 2, \ldots, n\}$ uniformly at random.

   Step 2: Compare each element $A[j]$ ($j \neq i$) to $A[i]$ and append $A[j]$ to B if $A[j] < A[i]$; otherwise append $A[j]$ to C. Finally, append $A[i]$ to B.

   Now answer the following questions.

   (a) What is the maximum probability that this algorithm will produce an incorrect answer?

   (b) Use the idea of *probability amplification* to design a simple algorithm that uses the above-described algorithm as a subroutine and has a maximum probability of error bounded above by $1/100$. What is the running time of this algorithm?

   (c) Resolve part (b), but now we want to reduce the the maximum probability of error to $1/n$.

2. Suppose that at every iteration of Karger's min-cut algorithm, instead of choosing a random *edge* for contraction, we choose two vertices u and v uniformly at random from all unordered vertex pairs, and contract the u-v pair. (Note that u and v need not be connected by an edge.) Show that there are graph examples on which the probability that this modified algorithm finds a min-cut is exponentially small. More precisely, show that there is a family $\mathcal{F}$ of graphs such that for every n large enough, $\mathcal{F}$ contains a graph G_n with n vertices such that the modified Karger's algorithm finds a min-cut on G_n with probability at most $1/a^{n/b}$ for constants $a > 1$ and $b \geq 1$.

3. Suppose we have a sequence of items passing by one at a time. At all times, we want to maintain a sample of one item with the property that it is *uniformly* distributed over all items that have been seen. Moreover, we want to accomplish this without knowing the total number of items in advance or storing all of the items we see.

   Consider the following algorithm, which stores just one item in memory at all times and uses a counter. When the first item appears, it is stored in the memory. When the kth item appears, it replaces the item in memory with probability $1/k$. Explain why this algorithm solves the problem.
   Note: This is a simple example of a *data streaming algorithm* called *reservoir sampling.*

4. Problem 6 from Chapter 9 in Prof. Erickson's notes (Page 11).
5. Problem 8 from Chapter 9 in Prof. Erickson's notes (Pages 11-12).
6. Problem 11(a), (b), and (c) from Chapter 9 in Prof. Erickson's notes (Pages 12-13).
7. Problem 13 from Chapter 9 in Prof. Erickson's notes (Pages 13-14).
