Question 0 concerns hashing, presented by Professor Martel.

The remaining problems concern material on randomized algorithms. Questions 1 and 2 can be done with only the material from Section 13.2, discussed in class today. The remaining questions rely on material in Sections 13.4 and 13.5, which we will discuss on Tuesday Nov. 6. You can wait until Tuesday to start on those problems, or before Tuesday, read ahead, or watch the videos on my website that cover that material.

0. Suppose our universe \( U \) has five elements, \( a, b, c, d, e \), and we are hashing to a table of size 3. Give a collection of three hash functions which is universal (so for any pair of elements at most one hash function causes them to collide), but also for the set of elements \( \{a, b, c\} \) every function in the collection causes at least one collision, so there is no perfect hash function in the collection for \( \{a, b, c\} \). Describe the hash functions by a table, e.g. \( h_1(a) = 0, \ h_1(b) = 0, \ h_1(c) = 1, \ldots \) NOT by a normal functional description such as \( h(x) = x \mod 3 \).

1. In class on Thursday, we discussed a randomized algorithm to find a global min-cut. That algorithm successively picks at random an edge \( (u, v) \) in the current graph \( G' \), and contracts that edge, i.e., merges nodes \( u \) and \( v \). Suppose instead, that the algorithm successively picks at random pair of nodes, \( u \) and \( v \) in \( G' \), whether or not there is an actual edge \( (u, v) \) in \( G' \), and then merges nodes \( u \) and \( v \). This is done until there are only two super-nodes left, which defines a bipartition of the nodes, and hence defines a cut in \( G \). We are again concerned with the probability that the cut returned by this algorithm is a global min-cut.

   a) Without doing a formal analysis, do you think this algorithm will be more likely or less likely, or equally likely to find a global min-cut, compared to the algorithm we discuss in class. Give a hand-waving explanation.

   b) Do a formal analysis of the modified algorithm to get the best probabilistic bound you can. That is, prove that the algorithm will return a global min-cut with probability \( f(n) \) for a function \( f \) of \( n \) which is as large as you can obtain. For the algorithm discussed in class, \( f(n) = \left( \frac{1}{n^2} \right) \).

2. Solve the randomized version of problem 1 in chapter 13 of the book.

3. Do problem 7 in chapter 13 of the book. (needs material in 13.4)

4. On page 731 of the book the second paragraph has the statement
\[ X = X_0 + X_1 + X_2 + ..., \text{ where } X_j \text{ is the expected number of steps spent by the algorithm in phase } j. \]

I think there is an error in this statement. If you understand the analysis, you should be able to spot it. What is it, and why is it an error, how is it fixed? (needs material in 13.5)

5a. Suppose in the analysis of Randomized Select (S,k) we defined a central splitter to be one that is larger than at least one sixth of the elements and smaller than at least one sixth of the elements. Then what would the analysis show about the expected number of comparisons done by the algorithm? That is, before we showed it was at most 8n. What is it now?

5b. Given your answer to part 5a, how would you answer the question asked in class: “Why 1/4?” (needs material in 13.5)