CS 577 # Assignment 5 – Divide and Conquer Spring 2022

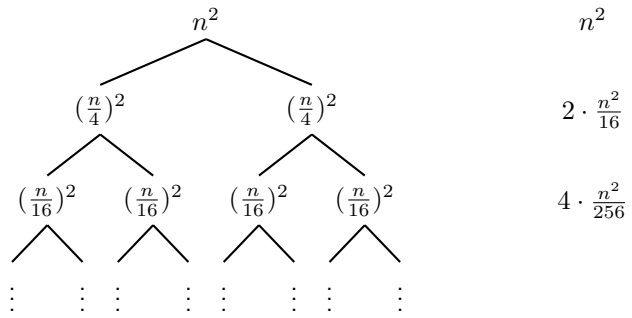
Answer the questions in the boxes provided on the question sheets. If you run out of room for an answer, add a page to the end of the document.

Name: _____ Wisc id: _____

Divide and Conquer

1. *Erickson, Jeff. Algorithms (p.49, q. 6).* Use recursion trees to solve each of the following recurrences.
   (a) $C(n) = 2C(n/4) + n^2$

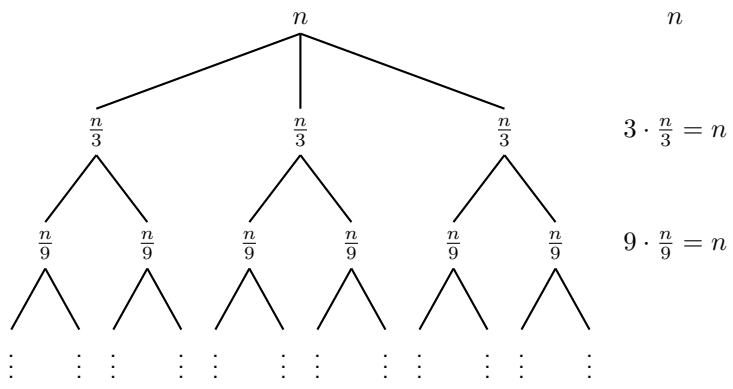
Solution:



There are $O(log_4 n)$ levels. Work at level i is $\frac{2^i}{16^i} n^2 = \frac{2^i}{2^{4i}} n^2 = \frac{1}{2^{3i}} n^2 = \frac{1}{8^i} n^2$.
Total work is $n^2(\sum_{i=0}^{O(log_4 n)} \frac{1}{8^i})$.
By sum of *infinite* geometric series, $\frac{1}{8} + \frac{1}{64} + \ldots = \frac{1}{1-\frac{1}{8}} = \frac{8}{7}$, so $n^2(\sum_{i=0}^{O(log_4 n)} \frac{1}{8^i}) \leq n^2(1 + \frac{8}{7})$.
Runtime is $O(n^2)$

   (b) $E(n) = 3E(n/3) + n$

Solution:



There are $O(log_3 n)$ levels. The work at each level is $O(n)$, so runtime is $O(n \log_3 n) \equiv O(n \log n)$

2. *Kleinberg, Jon. Algorithm Design (p. 246, q. 1).* You are interested in analyzing some hard-to-obtain data from two separate databases. Each database contains n numerical values—so there are $2n$ values total—and you may assume that no two values are the same. You'd like to determine the median of this set of $2n$ values, which we will define here to be the nth smallest value.

   However, the only way you can access these values is through queries to the databases. In a single query, you can specify a value k to one of the two databases, and the chosen database will return the kth smallest value that it contains. Since queries are expensive, you would like to compute the median using as few queries as possible.

   Give an algorithm that finds the median value using at most $O(\log n)$ queries.

   Solution:

   Label the databases A and B. Keep track of lower bounds in A_{lower} and B_{lower} (initialized to 1) and upper bounds in A_{upper} and B_{upper} (initialized to n). Let a and b be the results of querying A and B, respectively.

   Base Case: If $A_{upper} \leq A_{lower}$, the median is $min(a, b)$.

   Recursive Case:

   Query A on $k_1 = \frac{A_{upper}+A_{lower}}{2}$ and B on $k_2 = \frac{B_{upper}+B_{lower}}{2}$ (assume integer division).

   If $a > b$, set $A_{upper} = k_1$ and $B_{lower} = k_2 + 1$ (and vice versa for $a < b$).

3. *Kleinberg, Jon. Algorithm Design (p. 246, q. 2).* Recall the problem of finding the number of inversions. As in the text, we are given a sequence of n numbers $a_1, ..., a_n$, which we assume are all distinct, and we define an inversion to be a pair $i < j$ such that $a_i > a_j$.

   We motivated the problem of counting inversions as a good measure of how different two orderings are. However, this measure is very sensitive. Let's call a pair a significant inversion if $i < j$ and $a_i > 2a_j$. Give an $O(n \log n)$ algorithm to count the number of significant inversions between two orderings.

   Solution:

   Use the inversion counting algorithm stated in *Algorithm Design*, except in the merge-and-count method, change "If b_j is the smaller element" to "if $b_j < 2a_i$".

4. *Kleinberg, Jon. Algorithm Design (p. 246, q. 3).* You're consulting for a bank that's concerned about fraud detection. They have a collection of n bank cards that they've confiscated, suspecting them of being used in fraud.

   It's difficult to read the account number off a bank card directly, but the bank has an "equivalence tester" that takes two bank cards and determines whether they correspond to the same account.

   Their question is the following: among the collection of n cards, is there a set of more than $\frac{n}{2}$ of them that all correspond to the same account? Assume that the only feasible operations you can do with the cards are to pick two of them and plug them in to the equivalence tester. Show how to decide the answer to their question with only $O(n \log n)$ invocations of the equivalence tester.

   Solution:

   Return values: card that represents majority element in a group and number of cards in group equivalent to that element, or no majority element.

   Divide: Split the cards into two halves and recursively find the majority element of each half.

   Merge: Since the majority element in a subproblem must also be the majority element in at least one of its halves, consider the majority element (if any) returned by each half.

   - If they're the same, return this as the majority element.
   - Otherwise, test each against all elements in the subproblem and return the element that is found to match the majority.
     - Return "none" if neither candidate matches a majority of elements, or if there were no candidates to begin with.

   Base case: Only one card to compare. No need to invoke equivalence tester, but this card is the majority element of its group (size 1).

5. Implement the optimal algorithm for inversion counting in either C, C++, C#, Java, or Python. Be efficient and implement it in $O(n \log n)$ time, where n is the number of elements in the ranking.

   The input will start with an positive integer, giving the number of instances that follow. For each instance, there will be a positive integer, giving the number of elements in the ranking. A sample input is the following:

```
2
5
5 4 3 2 1
4
1 5 9 8
```

   The sample input has two instances. The first instance has 5 elements and the second has 4. For each instance, your program should output the number of inversions on a separate line. Each output line should be terminated by a newline. The correct output to the sample input would be:

```
10
1
```