Algorithms, Fall '23, Activity 7: Select

In the randomized Select algorithm the pivot is chosen randomly:

SELECT-RAND($A[1 \ldots n], k$)

1. Let $x = A[i]$, where i is a random number from $\{1, \ldots, n\}$
2. Rearrange A so that elements smaller than x are listed first, followed by elements equal to x, and then followed by elements larger than x
3. Let j_1 and j_2 be the first and the last position of x in the rearranged A
4. If $k < j_1$: return SELECT-RAND($A[1 \ldots j_1 - 1], k$)
5. If $k \geq j_1$ and $k \leq j_2$: return x
6. If $k > j_2$: return SELECT-RAND($A[j_2 + 1 \ldots n], k - j_2$)

We will intuitively analyze this algorithm's running time. For simplicity, assume all elements in A are different. For an element $y \in A$, we will use the term *rank* of y to indicate y's index had A been sorted.

(a) Suppose $A = [15, 2, 10, 3, 9, 7, 1, 13, 4, 6, 16, 11, 8, 5, 12, 14]$ and $k = 4$. What is Select(A, k) (you do not have to trace the algorithm)?

(b) Trace step 1 of the randomized Select's pseudo code: Using a random number generator for i, obtain x. Repeat this four times (i.e., generate four different x's, using the same A), and for each x state its rank. Also state whether the rank is between $n/4$ and $3n/4$.

x				
rank				
rank between $n/4$ and $3n/4$				

(c) For any input A, what is the probability that x chosen in step 1 has rank between $n/4$ and $3n/4$?

(d) If an event happens with probability 1/2, it, on average, happens every other time. Suppose we are lucky and it happens every time. Under this assumption, trace the first three recursive calls of SELECT-RAND for A and k from part (a).

In particular, the r-th row of the table below corresponds to the r-th recursive call of SELECT-RAND. State the values of A, k, and n for the current recursive call. Choose x such that its rank is between $\lceil n/3 \rceil$ and $\lfloor 3n/4 \rfloor$. If you wish, you are welcome to add more rows and trace it all the way to the base case.

A	k	n	x	Rank of x	Rearranged A

(e) Give the recurrence for the SELECT-RAND under this lucky assumption (when x always happens to have rank between $n/4$ and $3n/4$):

(f) Evaluate this recurrence, using any technique:

Note: We are lucky, on average, every other time. Therefore, even when not always lucky, the overall running time will be, in expectation, about twice as long as the bound you found in part (f). Therefore, it will be the same big-Oh bound.