CS 570: Analysis of Algorithms – H6

Submitted by: Indronil Bhattacharjee

Exercise 7.2-6

Without loss of generality, assume that the entries of the input array are distinct. Since only the relative sizes of the entries matter, we may assume that A contains a random permutation of the numbers 1 through n.

Now, $0 < \alpha \leq 1/2$. Let k denote the number of entries of A which are less than A[n]. PARTITION produces a split more balanced than $1 - \alpha$ to α if and only if $\alpha n \leq k \leq (1 - \alpha)n$.

This happens with probability $= \frac{(1-\alpha)n - \alpha n + 1}{n}$

$$= 1 - 2\alpha + \frac{1}{n}$$

$$\approx 1 - 2\alpha$$

Therefore, the probability of having a better partition, where the split is more balanced, is approximately $1-2\alpha$.

Exercise 7.4-4

We'll use the lower bound for the expected running time

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1} \\
&= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k} \\
&\geq \sum_{i=1}^{n-1} 2 \ln (\mathrm{n} - \mathrm{i} + 1) \\
&= 2 \ln \sum_{i=1}^{n-1} \mathrm{n} - \mathrm{i} + 1 \\
&= 2 \ln(n!) \\
&= \frac{2}{\lg e} \lg(n!) \quad \geq \mathrm{c} n \lg n \\
&= \Omega(n \lg n)
\end{aligned}$$

The first step defines the expected value E[X] as the sum of probabilities of all possible outcomes weighted by the number of comparisons for each outcome. The nested summation represents all possible pairs of indices (i, j) where i < j. Then simplified using the fact that j - i + 1 = k (the number of elements between i and j) and replacing j with k + i - 1.

The inequality sign ≥ is introduced, which comes from the fact that each term in the inner summation is at least 2/k. This inequality holds because 2/k is minimized when k=1 and the natural logarithm function is monotonically increasing.

Then we use the property of logarithms where the sum of logarithms is equal to the logarithm of the product. Finally, Converting from natural logarithm to base-2 logarithm. Finally c is a constant which is $\frac{2}{\lg e}$ and nlgn is a lower bound for lg(n!).

This derivation shows that the expected value of X is at least proportional to nlgn, which means that the average case time complexity of whatever algorithm X represents is at least Ω(nlgn). Therefore RANDOMIZED-QUICKSORT's expected running time is Ω(n lg n).