A Randomized Quick Sort

Assume that elements of A are distinct.

Instead of choosing A[r] as the Pivot, we will select a randomly chosen element from the array A[P,.. r] as a Pivot.

Now, we expect the split of the input array to be reasonably well balanced on average.

```
RANDOMIZED-PARTITION(A, p, r)
    i = RANDOM(p, r)
    exchange A[r] with A[i]
    return PARTITION(A, p, r)
```

→ Chooses an index b/w P & r Uniformly random and independently.

```
RANDOMIZED-QUICKSORT(A, p, r)
    if p < r
        q = RANDOMIZED-PARTITION(A, p, r)
        RANDOMIZED-QUICKSORT(A, p, q - 1)
        RANDOMIZED-QUICKSORT(A, q + 1, r)
```

```
PARTITION(A, p, r)
x = A[r]
i = p - 1
for j = p to r - 1
    if A[j] ≤ x
        i = i + 1
        exchange A[i] with A[j]
exchange A[i + 1] with A[r]
return i + 1
```

Q: How many times PARTITION Subroutine is Called in the entire execution of the quick sort algorithm.

Sol: at most n calls

One call to Partition takes $O(1)$ time + amount of time that is Propotional to the number of iterations of the for loop (Lines 3-6)

Since each iteration of the for loop Performs a Comparison in line 4, Comparing the Pivot element to another element of the array A.

Therefore, if we can Count the total # of times line 4 is executed, we can bound the total time Spent in the for loop during the entire execution of QUICKSORT.

Let X be the total number of comparisons performed in all calls to PARTITION.

Then total time $O(n + X)$

atmost n calls to the PARTITION, each of which const amount of work.

Goal: To compute X.

Idea: Instead of finding # of comparisons in each call to PARTITION, we ~~can~~ find an overall bound on the total number of comparisons.

We rename the elements of the array ~~are~~ as

$z_1, z_2, \ldots z_n$

where z_i is the i^{th} smallest element of A

Ex: $A = [4, 6, 2, 3]$

then $z_1 = 2,\ z_2 = 3,\ z_3 = 4,\ z_4 = 6$

Let $Z_{ij} = \{z_i, z_{i+1}, \ldots z_j\}$ be the set of elements b/w z_i and z_j, inclusive.

Q1 Fix two elements z_i and z_j of i/p array. How many times can these two elements be compared during the execution of Quicksort?

Q2 When does the algorithm compare z_i and z_j

Let $X_{ij} = \begin{cases} 1 & \text{if } z_i \text{ is compared to } z_j \\ 0 & \text{o.w} \end{cases}$

Here too, we are considering whether the comparison takes place at any time during the execution of the algorithm, not just ~~one~~ during one call of PARTITION.

From Q1 above, each pair is compared at most once, we get

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}$$

$$E[X] = E\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right]$$

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[X_{ij}]$$

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \underbrace{Pr[z_i \text{ is compared to } z_j]}_{\text{Now, we will compute this quantity.}}$$

$$z_1 \ldots z_i \; z_{i+1} \ldots z_j \; z_{j+1} \ldots z_n$$

If a pivot x, is chosen such that $z_i < x < z_j$

then z_i and z_j cannot be compared at any subsequent time.

If, z_i is chosen as a pivot before any other element in Z_{ij}, then z_i will be compared with each element of Z_{ij}, except itself.

Similarly if z_j is choosen as a pivot before any other element in Z_{ij}, then z_j will be compared with each element of Z_{ij}, except itself.

$Pr[z_i \text{ is compared with } z_j]$

$= Pr[z_i \text{ or } z_j \text{ is first Pivot chosen from } Z_{ij}]$

$= Pr[z_i \text{ is first Pivot chosen from } Z_{ij}] +$

$Pr[z_j \text{ is first Pivot chosen from } Z_{ij}]$

$$= \frac{1}{j-i+1} + \frac{1}{j-i+1}$$

$$= \frac{2}{j-i+1}$$

$$\therefore \quad E[X] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1}$$

let $k = j-i$ then

$$E[X] = \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1}$$

$$< \sum_{i=1}^{n-1} \sum_{k=1}^{n} \frac{2}{k}$$

$$= \sum_{i=1}^{n-1} O(\log n)$$

$$= O(n \log n)$$

$\therefore$ The expected running time of Randomized quicksort is $O(n\log n)$ if the elements are distinct.