Randomized Quicksort

Given an array A of n distinct elements:

> RANDOMIZED-QUICKSORT (A):
> - Select a pivot $p \in A$ uniformly at random.
> - Find $L = \{a \in A : a < p\}$, $R = \{a \in A : a > p\}$
> - RANDOMIZED-QUICKSORT (L) and RANDOMIZED-QUICKSORT(R)

Theorem. For a set A of n distinct elements, the expected number of comparisons is $O(n \log n)$.

Proof.

1. For each pair of indices i and j (with $i < j$), define an indicator random variable X_{ij} such that:

$$X_{ij} = \begin{cases} 1 & \text{if elements } a_i \text{ and } a_j \text{ are compared} \\ 0 & \text{otherwise} \end{cases}$$

Then the total number of comparisons is:

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

Recall: We seek $E[X]$.

2. Let's analyze the probability that elements a_i and a_j are compared, i.e. $\Pr[X_{ij} = 1]$.

Exercise: Each pair of elements is compared at most once. Why?

Key Observation:

Once an element between a_i and a_j is chosen as a pivot, then a_i and a_j cannot be compared in any subsequent step. Why?

That is, a_i and a_j are only compared if the first pivot chosen from $A_{ij} = \{a_i, a_{i+1}, \ldots, a_{j-1}, a_j\}$ is either a_i or a_j.

Hence:

$$\Pr[X_{ij} = 1]$$
"a_i and a_j are compared"

$$= \Pr[\text{"}a_i \text{ or } a_j \text{ is first pivot chosen from } A_{ij}\text{"}]$$

$$= \frac{2}{|A_{ij}|}$$ (pivots are selected uniformly at random)

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

3 Finally, taking the expectation:

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

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

$$= 2 \sum_{i=1}^{n-1} \left(\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n-i+1} \right)$$

$$\leq 2 \sum_{i=1}^{n-1} \left(\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \right)$$

$$\leq 2 \sum_{i=1}^{n-1} \ln n$$ (harmonic sum)

$$< 2n \ln n$$

as desired.

□

Exercise: Bound the total running time of RANDOMIZED-QUICKSORT given the above conclusion. What other work is done?