Date Pg 0

CS-412 - Randomized Algorithms

[Source: CLRS] <u>The Hiring Problem</u>

<u>The Problem</u>: Suppose, we are hiring for a job and delegate the hiring task to a hiring agency.

The agency sends n candidates; one candidate on a single day.

Suppose, we hire the first candidate, say, x_1 with some quality score q_1. Let the cost of hiring a candidate is C_i (cost of interviewing).

Consider that the next day, we are sent the next candidate x_2 with a quality score q_2. If x_2 is a better candidate than x_1 (already hired), we hire x_2 ($\because q_2 > q_1$).

The cost of hiring x_2 will be $\underbrace{C_i}_{\text{interview}}$ plus the cost of firing x_1 and hiring x_2, i.e. (C_h).

Notice, for x_1, we only had to bear the cost C_i.

<u>Assumption</u>: C_i is <u>small</u> (i.e., the cost of interviewing a candidate)

i.e., $\boxed{C_i << C_h}$.

If we end up interviewing all n candidates, the total cost of interviewing will become $\boxed{n.C_i}$ or $\boxed{C_i.n}$

Now, suppose we hire the i^{th} candidate, if they're the best so far

If we hire only $m << n$ candidates, the total cost of hiring will be $\boxed{C_h.m}$

Hence, the total cost of hiring becomes $\boxed{O(\underbrace{C_i . n}_{\text{for sure}} + \underbrace{C_h . m}_{\text{depends upon the quality of Candidates}})}$

Worst-case: Hiring all n candidates, ie when $m = n$ ie when the candidates $x_1, x_2, ..., x_n$ come in the order s.t., $q_1 < q_2 < ... < q_n$ respectively.

So, the worst-case is $\boxed{O(C_h . n)}$

Best-case: When $q_1 > q_2 > ... > q_n$, the total cost will be $\boxed{O(C_i . n)}$ [The subsequent candidates aren't hired]

Average Case: If candidates arrive in an arbitrary order of quality,

Define: An indicator variable associated with some event 'A' s.t.

$$I_A = \begin{cases} 1, & \text{if 'A' occurs} \\ 0, & \text{otherwise.} \end{cases}$$

Lemma: Given a sample space S and an event A, let $X_A = I\{A\}$, then,

$$E[X_A] = Pr\{A\}$$

ie,

$$E[X_A] = E[I\{A\}]$$

$E[x] = \sum x \, Pr(x)$

$$= 1 . Pr(A) + 0 . Pr(A^c)$$

$$\therefore \boxed{E[X] = Pr(A)}$$

Refresher: Let X be a R.V. denoting the total no. of heads in n Bernoulli trials.

then,

$$E[X] = E\Big[\sum_{i=1}^{n} \underbrace{X_i}_{\text{Indicator R.V.}}\Big]$$

$$\Rightarrow E[X] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} Pr(X_i) \quad \{\text{from the Lemma}\}$$

$$= \sum_{i=1}^{n} (1/2)$$

$$\therefore \boxed{E[X] = n/2}$$

<u>Now, back to the Hiring Problem</u>

$$E[X] = \sum_{x} x \cdot Pr(X = x)$$

<u>Defn</u>: X_i to be an indicator R.V. s.t.

$$X_i = \begin{cases} 1, & \text{candidate } i \text{ is hired} \\ 0, & \text{otherwise} \end{cases}$$

Let,

$$X = X_1 + X_2 + \cdots + X_n$$

then,

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right]$$

$$= \sum_{i=1}^{n} E[X_i]$$

$$= \sum_{i=1}^{n} 1/i \quad \text{(from the Lemma)}$$

$$\therefore E[X] = \underbrace{\ln n + O(1)}_{H_n : \, n^{\text{th}} \text{ Harmonic number}}$$

ie, On average, when interviewing n candidates, we would hire approx. <u>$\ln(n)$</u> candidates

the avg. cost of hiring is $O(\underbrace{C_h \cdot \ln n}_{\text{Compared to } O(C_h n) \text{ worst-case!}})$