 5/22/13

Randomized algorithms

Quicksort

A. To sort $a_1, a_2, \dots, a_n$: If $n > 1$,

1. Choose $p \in \{1, 2, \dots, n\}$ randomly and uniformly.
2. Let $L = \{a_i \mid a_i < a_p\}$,
   $E = \{a_i \mid a_i = a_p\}$,
   $G = \{a_i \mid a_i > a_p\}$.
3. Recursively sort and output L;
   Output E;
   Recursively sort and output G.

B. If unlucky $|L| = 0$ and $|E| = 1$, and time is $\Theta(n^2)$.

But what is the expected time? For simplicity, assume all a_i distinct.

Let r.v. R be the rank of a_p, i.e., $R = |L|$.

Let r.v. X_n be the number of comparisons to sort $a_1, a_2, \dots, a_n$.

if $n > 1$: $X_n = n - 1 + X_R + X_{n-1-R}$.

$$E[X_n] = n - 1 + E[X_R + X_{n-1-R}] \quad \text{(linearity)}$$

$$= n - 1 + \sum_{i=0}^{n-1} E[X_R + X_{n-1-R} \mid R = i] P(R = i) \quad \text{(law of total expectation)}$$

$$= n - 1 + \frac{1}{n} \sum_{i=0}^{n-1} E[X_i + X_{n-1-i}]$$

$$= n - 1 + \frac{1}{n}\left(\sum_{i=0}^{n-1} E[X_i] + \sum_{i=0}^{n-1} E[X_{n-1-i}]\right) \quad \text{(linearity)}$$

$$= n - 1 + \frac{2}{n} \sum_{i=0}^{n-1} E[X_i]$$

$$n E[X_n] = n(n-1) + 2 \sum_{i=0}^{n-1} E[X_i]$$

$$(n-1) E[X_{n-1}] = (n-1)(n-2) + 2 \sum_{i=0}^{n-2} E[X_i] \quad \text{(substitute } n \to n-1\text{)}$$

$$n E[X_n] - (n-1) E[X_{n-1}] = 2n - 2 + 2E[X_{n-1}]$$

$$n E[X_n] = 2n - 2 + (n+1) E[X_{n-1}]$$

$$\frac{E[X_n]}{n+1} = \frac{2n-2}{n(n+1)} + \frac{E[X_{n-1}]}{n} \leq \frac{2}{n} + \frac{E[X_{n-1}]}{n}$$

$$\leq \frac{2}{n} + \frac{2}{n-1} + \frac{E[X_{n-2}]}{n-1}$$

$$\leq \frac{2}{n} + \frac{2}{n-1} + \frac{2}{n-2} + \frac{E[X_{n-3}]}{n-2}$$

$$\leq \cdots$$

$$\leq \frac{E[X_1]}{2} + 2\sum_{i=2}^{n} \frac{1}{i} = 2\sum_{i=2}^{n} \frac{1}{i} \leq 2(H_n - 1) \leq 2\ln n$$

$$E[X_n] \leq 2(\ln n)(n+1) = 2n\ln n + 2\ln n \leq 1.4 n\log_2 n + O(\log n)$$