# Lecture 9 Randomized Algorithm II

Tuesday, September 26, 2017 2:48 PM

- Randomized Quick Sort

$$\mathbb{E}[X_n] = \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{E}[X_{i-1}] + \mathbb{E}[X_{n-i}] + A\cdot n\right)$$

- Induction Hypothesis: $\mathbb{E}[X_n] \le C\cdot n\cdot \log_2 n$

  $C$ is a parameter that we determine later.

- Base Case: easy
- Induction Step: Assume for all $t < n$, we have $\mathbb{E}[X_t] \le C\cdot t\cdot \log_2 t$

for $\mathbb{E}[X_n]$ use the recursion

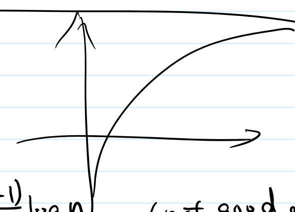
$$\mathbb{E}[X_n] = \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{E}[X_{i-1}] + \mathbb{E}[X_{n-i}]\right) + A\cdot n$$

$$= \frac{1}{n}\left(\mathbb{E}[X_0] + \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_{n-1}] + \mathbb{E}[X_{n-1}] + \mathbb{E}[X_{n-2}] + \cdots + \mathbb{E}[X_0]\right) + An$$

$$= \frac{2}{n}\sum_{i=1}^{n}\mathbb{E}[X_{i-1}] + An$$

$$\le \frac{2}{n}\sum_{i=1}^{n} C\cdot(i-1)\log_2(i-1) + An \qquad \text{(apply IH)}$$

think want $\le C\cdot n\log_2 n$



idea 1: $\log_2(i-1) \le \log_2 n$

$$\sum_{i=1}^{n} C\cdot(i-1)\log_2(i-1) \le \sum_{i=1}^{n} C\cdot(i-1)\cdot\log_2 n = C\frac{n(n-1)}{2}\log_2 n$$

(not good enough)

idea 2 $\quad i \le \frac{n}{2} \quad \log_2(i-1) \le \log_2\frac{n}{2}$

$n \ge i > \frac{n}{2} \quad \log_2(i-1) \le \log_2 n$

$$= \frac{2}{n}\left(\sum_{i=1}^{\frac{n}{2}} C\cdot(i-1)\log_2(i-1) + \sum_{i=\frac{n}{2}+1}^{n} C(i-1)\log_2(i-1)\right) + A\cdot n$$

$$\le \frac{2}{n}\left(\sum_{i=1}^{\frac{n}{2}} C(i-1)\underbrace{\log_2\frac{n}{2}}_{\log_2 n - 1} + \sum_{i=\frac{n}{2}+1}^{n} C(i-1)\log_2 n\right) + A\cdot n$$

$$= \frac{2}{n}\left(\sum_{i=1}^{n} C(i-1)\log_2 n - \sum_{i=1}^{\frac{n}{2}} C(i-1)\right) + A\cdot n$$

$$= \frac{2}{n}\cdot C\frac{n(n-1)}{2}\log_2 n - \frac{2}{n}\cdot C\cdot\frac{\frac{n}{2}\left(\frac{n}{2}-1\right)}{2} + A\cdot n$$

$$\approx C\cdot n\log_2 n - \frac{C\cdot n}{4} + An$$

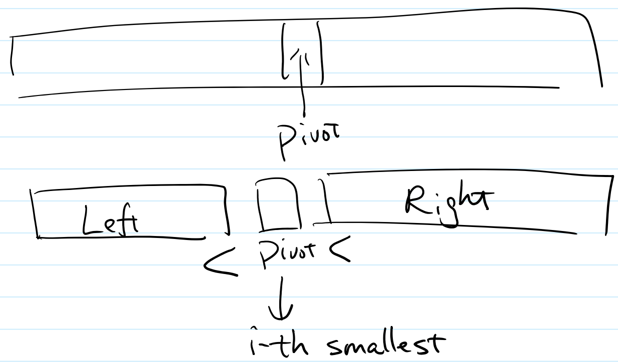
want this $\le C\cdot n\log_2 n$ need $C \ge 4\cdot A$

choose $C = 4 \cdot A$

$\leq C \cdot n \log_2 n.$ $\square$

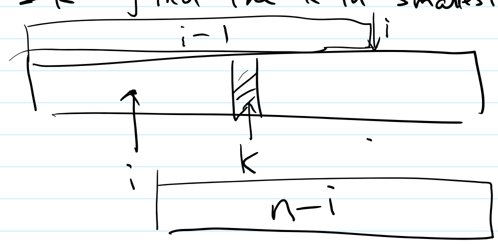
By induction we have $E[X_n] \leq 4 \cdot A n \log_2 n = O(n \log n)$

- Quick Selection
  - recall quick sort




if we want to find k-th smallest number

$$\begin{cases} i = k & \text{output the pivot number} \\ i < k & \text{find the } (k-i)\text{-th smallest number in right part} \\ i > k & \text{find the k-th smallest number in left part.} \end{cases}$$




- Rejection sampling

Claim: $\Pr[X = i \mid X \text{ is kept}] = q_i = \Pr[Y = i]$

Proof: $\Pr[X = i \mid X \text{ is kept}] = \dfrac{\Pr[X = i, X \text{ is kept}]}{\Pr[X \text{ is kept}]} = P_i \cdot \dfrac{q_i}{C P_i} = \dfrac{q_i}{C}$

$\Pr[X \text{ is kept}] = \sum_t \Pr[X = t, X \text{ is kept}] = \sum_t \frac{q_t}{C} = \frac{1}{C}$

$= q_i$ $\square$

- Coin Toss
  - every time we toss the biased coin twice $\begin{cases} HT \ TH & \text{succeed} \\ HH \ TT & \text{fail} \end{cases}$

$\Pr[\text{succeed}] = \Pr[HT] + \Pr[TH]$
$= p(1-p) + (1-p)p = 2p(1-p)$

Let $X$ be the number of tries that we need before we succeed

$$\Pr[X=1] = \underbrace{2p(1-p)}_{q}$$

$$\Pr[X=2] = \underbrace{(1-q)}_{\text{failed in first time}} q \longrightarrow \text{succeed in second time}$$

$$\Pr[X=i] = \underbrace{(1-q)^{i-1}}_{\text{failed in first } i-1 \text{ tries}} q$$

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \Pr[X=i]\, i$$

distribution : geomeric distribution

$$\mathbb{E}[X] = \frac{1}{q}$$

Computing $\mathbb{E}[X]$

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \Pr[X=i]\, i$$

$$= \sum_{i=1}^{\infty} \Pr[X \geq i]$$

See this

sum rows first = sum columns first

$$= \sum_{i=1}^{\infty} (1-q)^{i-1}$$

$$= \frac{1}{q}$$

$P_i = \Pr[X=i]$

| | | | | | row sum |
|---|---|---|---|---|---|
| $P_1$ | $P_2$ | $P_3$ | $P_4$ | | $\Pr[X \geq 1]$ |
| | $P_2$ | $P_3$ | $P_4$ | | $\Pr[X \geq 2]$ |
| | | $P_3$ | $P_4$ | $\cdots$ | $\Pr[X \geq 3]$ |
| | | | $P_4$ | | $\vdots$ |

Columnsum $P_1 \times 1$ $P_2 \times 2$ $P_3 \times 3$ $\cdots$

- Monte Carlo alg for area of a circle

Let $X_i = \begin{cases} 0 & i\text{-th point is not in circle} \\ 1 & i\text{-th point is in circle} \end{cases}$

$$\mathbb{E}[X_i] = \Pr[X_i = 1] = \frac{\text{area of circle}}{\text{area of square}} = P \quad \left(\frac{\pi}{4}\right)$$

Final output : Let $X = \sum_{i=1}^{n} X_i$ ($X$ = count in alg)

output $\frac{4 \cdot X}{n}$

$$\mathbb{E}[\text{output}] = \frac{4\,\mathbb{E}[X]}{n} = \frac{4 \sum_{i=1}^{n} \mathbb{E}[X_i]}{n} = 4 \cdot P = \text{area of circle}$$