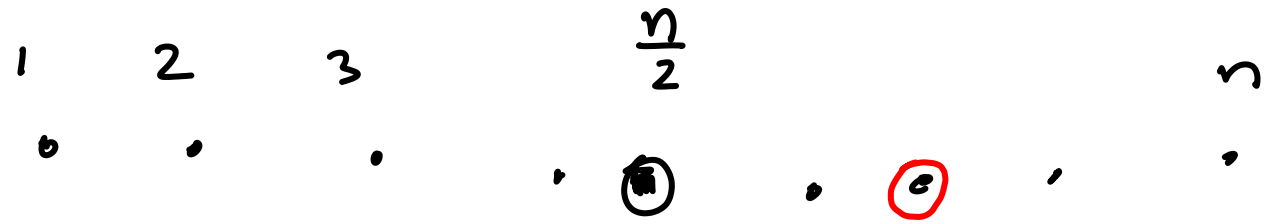
COL 702 Lecture 8 Aug 21



i is randomly chosen in $[1, \ldots n]$

Case 1 $i > \frac{n}{2}$ Recursively called for subproblem of size i with prob $\frac{1}{n}$ $i = \frac{n}{2}+1, \frac{n}{2}+2 \ldots n$

size of the recursively called subproblems lie in the range $\frac{n}{2}+j$ $1 \le j \le \frac{n}{2}-1$

Case 2 $i < \frac{n}{2}$ $i = 1, 2, \ldots \frac{n}{2}-1$

sizes of rec called subproblems

$\frac{n}{2}+1, \frac{n}{2}+2 \ldots \frac{n}{2}+\frac{n}{2}-1$

Let $T(n)$ be the expected running time of selecting median

$$T(n) = \frac{1}{n}\sum_{j=1}^{n/2} T\left(\frac{n}{2}+j\right) + \frac{1}{n}\sum_{j=1}^{n/2} T\left(\frac{n}{2}+j\right) + O(n)$$

splitting and partitioning

$T(n) = ?$

$$T(n) = \frac{2}{n} \sum_{j=1}^{\frac{n}{2}-1} T\left(\frac{n}{2}+j\right) + cn$$

Guess $T(n) : a \cdot n$ a is some constant and verify by induction

We must show that

$$a \cdot n \overset{?}{\geq} \frac{2}{n} \sum_{j=1}^{n/2-1} a\left(\frac{n}{2}+j\right) + cn$$

$$\frac{2}{n}\left[\frac{a}{4}\left(\frac{n}{2}\right)^2 + \frac{\frac{n}{2}\left(\frac{n}{2}+1\right)}{2}\right] + cn$$

$$a\left(\frac{n}{2} + \frac{n}{4}\right) + cn : \frac{3n}{4}a + cn$$

$\frac{3n}{4}a + cn \leq an \Rightarrow \frac{a}{4} \geq c \Rightarrow a \geq 4c$

For $k \neq \frac{n}{2}$ (median) must show $T(n,k) \leq T\left(n, \frac{n}{2}\right)$

Extend this analysis to Quicksort

Deterministic linear-time selection

Choosing an approximate median deterministically

Group all the n elements in groups of '5' (odd no. ≥ 5)

Sort these groups individually

sort 1	2	3	$\frac{n}{5}$
x_{11}	x_{21}		
x_{12}	x_{22}		
x_{13}	x_{23}		
x_{14}	x_{24}		
x_{15}	x_{25}		

$\alpha \cdot 5$ Total cost is $\frac{n}{5} \times \alpha 5 \sim \alpha n$

Consider the median of each group after sorting

$k = \frac{n}{10}$ $\longrightarrow$

$\tilde{x}_{11}$	$\tilde{x}_{12}$	$\tilde{x}_{13}$	$\tilde{x}_{14}$	$\tilde{x}_{15}$
$\tilde{x}_{12}$	$\tilde{x}_{22}$			$\tilde{x}_{25}$
$\vdots$				
$\tilde{x}_{k1}$	$\tilde{x}_{k2}$	$\tilde{x}_{k3}$	. .	$\tilde{x}_{k5}$

Choose the median among the medians
(of groups of 5)

Median of $\frac{n}{5}$ elements : say that is $\tilde{x}_{k,3}$

Claim $\tilde{x}_{k,3}$ has rank $\frac{n}{4} \leq \frac{3n}{4}$

HOW DO WE CHOOSE $\tilde{x}_{k,3}$?

$T(n)$: time to select median/any rank

$$T(n) = \alpha n + O(n) + T\left(\frac{n}{5}\right) + T\left(\frac{3n}{4}\right)$$

- αn: sorting groups of 5
- $O(n)$: rearranging for recursive calls
- $T\left(\frac{n}{5}\right)$: selecting approx median recursively
- $T\left(\frac{3n}{4}\right)$: recursive call to the right partition

$T(5) = O(1)$

$T(n) = ?$ Guess $T(n) = cn$ and verify