Selection

Monday, February 3, 2025 9:47 AM

Last Time:
- Randomized Quicksort
- Quick select

HW 3 → due Friday

Today:
- Select

Select

Input: List of items $A[1 \ldots n]$
Target rank r

Output: The r^{th} smallest item

---

1 5 8 6 2 9 3

$r = 6$

Quickselect $(A[1 \ldots n], r)$

1. pick a pivot a_p
2. partition around the pivot $a_p \to \hat{p}$
3. If $r = \hat{p}$, then a_p is the r^{th} smallest
   If $r < \hat{p}$, then quickselect $(A[1 \ldots (\hat{p}-1)], r)$
   If $r > \hat{p}$, then quickselect $(A[(\hat{p}+1) \ldots n], r - \hat{p})$

---

1 5 8 6 2 9 3

$r = 6$

1 2 3 5 8 6 9

1 2 3 5 8 6 9

$\hat{p} = 3$

select the rank 3 thing from here

Efficiency

$$T(n) = \theta(n) + \max\{T(n-\hat{p}), T(\hat{p})\}$$

partition

worst-case: $T(n) = \theta(n) + T(n-1) \implies \theta(n^2)$

In practice: Randomly choose a pivot.

Randomized quickselect $\rightarrow$ $\boxed{\theta(n) \text{ expected runtime}}$

If $\hat{p} = n/2 \longrightarrow T(n) = \theta(n) + T(n/2)$

$\Downarrow$

$\boxed{\theta(n)}$

n

$n/2$

$n/4$

$n/8$

$\log_2 1 \rightarrow 0$

n^0 vs n^1

$$T(n) = \theta(n) + 2\,T(n/2)$$

---

mom_select(A[1...n], r)

if $n < 1000$: do brute-force.

otherwise:

$\theta(n) \longrightarrow$ ① Divide A into $\lceil n/5 \rceil$ subarrays of 5 items each $M_1, M_2 \ldots M_{n/5}$.

$\theta(n) \longrightarrow$ ② Find the median of each of the subarrays, and make a list $M[1 \ldots n/5]$

$\theta(n) \longrightarrow$ ② Find the median of each of the subarrays, and make a list $\underline{M[1 \ldots n/5]}$

$T(\frac{n}{5}) \longrightarrow$ ③ $a_p \leftarrow$ mom_select $(\underline{M[1 \ldots n/5]}, \; r = \frac{n}{10})$.

$\theta(n) \longrightarrow$ ④ $\hat{p} \leftarrow$ partition (A, a_p)

⑤ Run mom_select on $\underline{A[1 \ldots (\hat{p}-1)], r}$

or

$\underline{A[\hat{p}+1, \ldots n], r - \hat{p}}$

We don't need $\hat{p} = n/2 \ldots$

let $\hat{p} = \frac{n}{\beta}$ for some $\beta > 1$ $\Rightarrow \frac{n}{4}$

$\max\{T(n-\hat{p}), T(\hat{p})\}$

$\downarrow$

$$n - \frac{n}{\beta} = \frac{(\beta-1)n}{\beta} = \frac{n}{\left(\frac{\beta}{\beta-1}\right)} = \frac{n}{c}, \quad c > 1$$

$$T(n) = \theta(n) + T\left(\frac{n}{c}\right) \qquad c > 1$$

$\log_c 1 = 0$ $\quad n^0$ v.s. n^1

$T(n)$ is $\theta(n)$ by case 3.

$A \Rightarrow$ | a_1 a_2 a_3 a_4 a_5 | b_1 b_2 b_3 b_4 b_5 | c_1 c_2 c_3 c_4 c_5 | ...

| a_1 a_2 a_3 | b_1 b_2 b_3 | ......

$M =$

| . . . | b_1, b_2, b_3 | . . .

$$M = \{ a_2, b_4, c_3, \dots \} \rightarrow \frac{n}{5}$$

$$a_p \rightarrow b_4$$

What is the rank of a_p?

a_p is the median of M

$\rightarrow$ a_p is larger than $\frac{n}{10}$ items

each of these is bigger than 2 other things...

a_p is rank at least $\frac{3n}{10}$

a_p is rank at most $\frac{7n}{10}$

$$\frac{3n}{10} < \hat{p} < \frac{7n}{10}$$

Efficiency

$$T(n) = \Theta(n) + T(n/5) + \max\{T(n-\hat{p}), T(\hat{p})\}$$

$$\frac{3n}{10} < \hat{p} < \frac{7n}{10}$$

$$T(n) = \Theta(n) + T(n/5) + T(7n/10)$$

$$\boxed{T(n) = \theta(n) + T(n/5) + T\left(\frac{7n}{10}\right)}$$

$f(n)$ — n — n

$f(n/5)$ $f\left(\frac{7n}{10}\right)$ $\Rightarrow$ $\frac{n}{5}$ $\frac{7n}{10}$ $\Rightarrow \left(\frac{9}{10}\right)n$

$f\left(\frac{n}{5}\cdot\frac{1}{5}\right)$ $f\left(\frac{n}{5}\cdot\frac{7}{10}\right)$ . . $\frac{n}{5}\cdot\frac{1}{5}$ $\frac{n}{5}\cdot\frac{7}{10}$ $\frac{7n}{10}\cdot\frac{1}{5}$ $\frac{7n}{10}\cdot\frac{7}{10}$ $\Rightarrow \left(\frac{9}{10}\right)^2 n$

$\vdots$

$\left(\frac{9}{10}\right)^3 n$

$$\boxed{T(n) \text{ is } \theta(n)}$$

$5 \Rightarrow T(n) = \theta(n) + T(n/5) + T\left(\frac{7n}{10}\right)$

$3 \Rightarrow T(n) = \theta(n) + T(n/3) + T\left(\frac{2n}{3}\right)$

n — n

$n/3$ $2n/3$ $\Rightarrow$ n $\Rightarrow$ $\boxed{\theta(n \log n)}$

$n/3 \cdot \frac{1}{3}$ $n/3 \cdot \frac{2}{3}$ $\frac{2n}{3}\cdot\frac{1}{3}$ $\frac{2n}{3}\cdot\frac{2}{3}$ — n

$\vdots$
