03 Median & Akra Bazzi Theorem

Tuesday, January 14, 2025 13:51

- 1st half: D&C for finding median
- 2nd half: Akra Bazzi theorem (Master Theorem, more advanced)

Input: $A[0 \dots n-1]$, $1 \le k \le n$

Return the k-th smallest entry of A

Example: - sort the array // $O(n \log n)$

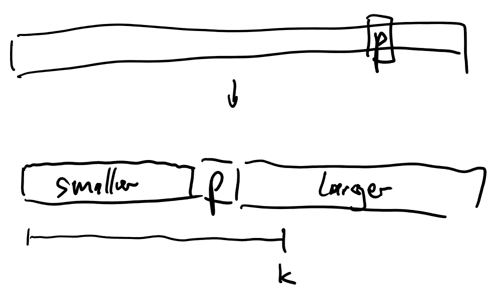
- return $A[k]$

Today: $O(n)$

```
kSmallest (A[0.. n-1], k)
  if n == 1
    if k == 1
      return A[0]
    else
      return failure // input didn't make sense

  find some "good" p = A[i]
  smaller = []
  larger = []
  for i = 0 ... n-1
    if A[i] ≤ p
      add A[i] to smaller
    if A[i] > p
      add A[i] to larger

  if #smaller < k
    return kSmallest (larger, k - #smaller )
  if #smaller ≥ k
    return kSmallest (smaller, k)
```



time:
in worst case, p could be the smallest/largest entry. then we recurse on array of length n-1

$\Rightarrow$ n recursions

$\Rightarrow O(n^2)$ total

We call p good if

$\frac{3}{10} n \le \#\text{smaller} \le \frac{7}{10} n \Rightarrow \frac{3}{10} n \le \#\text{larger} \le \frac{7}{10} n$

at least $\frac{3}{10} n$ entries smaller

$$\Rightarrow T(n) = T\left(\tfrac{7}{10} n\right) + O(n)$$

$$= O(n)$$

[——— |p| ———

at least $\frac{3}{10}n$ entries smaller

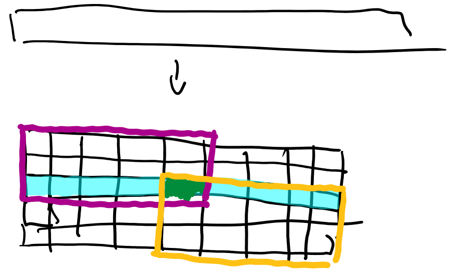
$= O(n)$

<u>if we can find a good p</u>

How to find a good p?

- split A into groups of size 5 // $O(n)$
- sort each group // $O(n)$
  - we now have $\frac{n}{5}$ medians (one for each group)

    $M[0 \ldots \frac{n}{5}-1]$



is at most

is at least

- find the median of M

  $\left(\text{kSmallest}\left(M, \frac{n}{5} / 2 \right) \right)$

$\underline{3 \cdot \frac{1}{2}\frac{n}{5} \leq (\#\text{ of } A[i] \text{ with } A[i] \leq p)}$

$\underline{\frac{3}{5} \cdot \frac{1}{2} \cdot n \leq \#\text{ of } A[i] \text{ with } A[i] \geq p}$

$\underbrace{3 \cdot \frac{1}{2} \cdot \frac{n}{5}}_{= \frac{3}{10}n} \leq \#\text{ smaller} \leq \underline{\left(1 - \frac{3}{5} \cdot \frac{1}{2}\right)}n$

$= \frac{7}{10}n$

$T(n)$ = time of kSmallest on arrays of length n

$$= \underbrace{O(n) + T\left(\tfrac{n}{5}\right)}_{\text{finding } p} + \underbrace{O(n)}_{\substack{\text{finding arrays} \\ \text{- smaller} \\ \text{- larger}}} + \underbrace{T\left(\tfrac{7}{10}n\right)}_{\substack{\text{recursion on} \\ \text{smaller \underline{or} larger}}}$$

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

after break $= O(n)$

$$T(n) = T\left(\frac{n}{5}\right) + T\left(\frac{3}{10}n\right) + \Omega(n)$$

$= \Omega(n)$

2:50

$$T(n) = T\left(\frac{1}{5} \cdot n\right) + T\left(\frac{7}{10} \cdot n\right) + O(n)$$

"Guess and prove"

Guess $T(n) = O(n)$

Guess $T(n) = O(n)$

then prove by induction over n that

there exists c such that $T(n) \le c \cdot n \quad \forall n$

How can we "guess" correctly?

Example: $T(n) = T(\frac{1}{2} \cdot n) + T(\frac{2}{3} n) + O(n)$

Assume $T(n) = c \cdot n^x$ for some x

$$c \cdot n^x = c(\tfrac{1}{2} \cdot n)^x + c(\tfrac{2}{3} \cdot n)^x + O(n)$$

$$= c n^x \cdot \left((\tfrac{1}{2})^x + (\tfrac{2}{3})^x \right) + O(n) \qquad // :(n^x \cdot c)$$

$$1 = (\tfrac{1}{2})^x + (\tfrac{2}{3})^x + O(n^{1-x})$$

use $x > 1 \rightarrow \approx (\frac{1}{2})^x + (\frac{2}{3})^x$

if $x < 1$ ($1-x > 0$)
then $1 \approx O(n^{1-x})$ cannot work

$1 = (\frac{1}{2})^x + (\frac{2}{3})^x$ if $x \approx 1.29\ldots$

if $x > 1$

Next: Prove $T(n) = O(n^{1.3})$ ie there is c such that $T(n) \le c \cdot n^{1.3}$

General $T(n) = \left(\sum_{i=1}^{k} a_i \cdot T(\frac{n}{b_i}) \right) + O(n^c)$

$// \; k=1 \quad T(n) = a \cdot T(\frac{n}{b}) + O(n^c)$

Let x be such that $1 = \sum_{i=1}^{k} a_i \cdot \frac{1}{b_i^x}$

then $T(n) = \begin{cases} O(n^c) & \text{if } c > x \\ O(n^c \log n) & \text{if } c = x \\ O(n^x) & \text{if } c < x \end{cases}$

$//$ if $k=1$
$x = \log_b(a)$

Akra Bazzi Theorem

Proof by induction, one proof for each case

Example case 3) $T(n) = O(n^x)$ if $c < x$

$T(n) \le \bar{c} \cdot n^x$

Example case 3) $T(n) = O(n^x)$ if $c < x$

$$T(n) \le \bar{c} \cdot n^x$$

$$T(n) = \sum_{i=1}^{k} a_i \, T\left(\frac{n}{b_i}\right) + n^c$$

$$\le \sum_{i=1}^{k} a_i \cdot \bar{c} \cdot \left(\frac{n}{b_i}\right)^x + n^c$$

$$= n^x \cdot \bar{c} \cdot \sum_{i=1}^{k} a_i \frac{1}{b_i^x} + n^c$$

$$= n^x \cdot \bar{c} \cdot 1 + n^c$$

$$= n^x \cdot \bar{c} \cdot \underbrace{\left(1 + \frac{n^c}{n^x}\right)}$$

This is > 1, so we did not manage to prove T(n) <= n^x \bar{c}
There is a corrected proof in the Piazza post-lecture recap.