# Median Finding: Recursion

Wednesday, 23 August 2023 12:03 PM

**I/P:** Array $S = (x_1, \ldots, x_n)$ of $n$ <u>distinct</u> integers
**O/P:** $\lfloor n/2 \rfloor^{th}$ - smallest number (median)

Easy: Sort $S$, return $\lfloor n/2 \rfloor$-th smallest entry
Takes time $O(n \log n)$
Will give $O(n)$ time algo for this, using recursion

(Recursion is top-down. Induction is bottom-up)

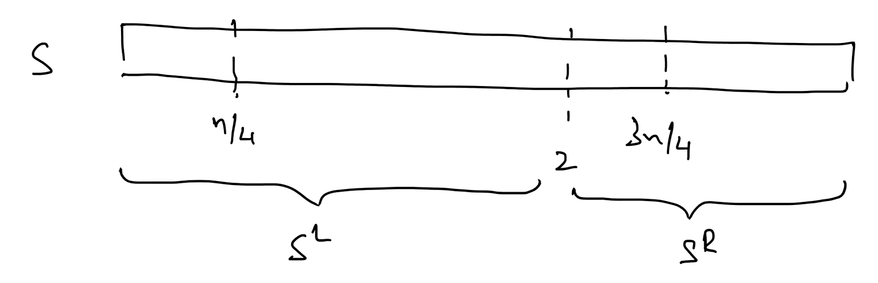
**2 ideas:** (i) Stronger algo: given $S$, s.t. $|S| = n$, & $k \leq n$, try & find $k^{th}$ smallest no., not just median

(ii) Approximate-Split: given $S$, s.t. $|S| = n$, returns $z \in S$ s.t. $z$ is "approx-median": $z$ is $k^{th}$ smallest elt. for $\frac{n}{4} \leq k \leq \frac{3n}{4}$

in linear time.

Algo.: **Rank-Find $(S, k)$:**

let $n \leftarrow |S|$
If $(n \leq 10)$
  sort $S$
  return $k^{th}$ smallest elt.
$z \leftarrow$ Approx-Split $(S)$ ← $O(n)$
$S^L \leftarrow \{x \in S : x \leq z\}$, $S^R \leftarrow \{x \in S : x > z\}$ $O(n)$




$n^L \leftarrow |S^L|$, $n^R \leftarrow |S^R|$
If $(k \leq n^L)$
  $\mu \leftarrow$ Rank-Find $(S^L, k)$
else $\mu \leftarrow$ Rank-Find $(S^R, k - S^L)$ $T(3n/4)$
Return $\mu$

Time taken for Rank-Find:

$T_{RF}(n) = O(n) + T_{AS}(n) + T_{RF}(3n/4)$ (since $|S^L|, |S^R| \leq 3n/4$)
$= O(n)$ (if Approx Split takes linear time)

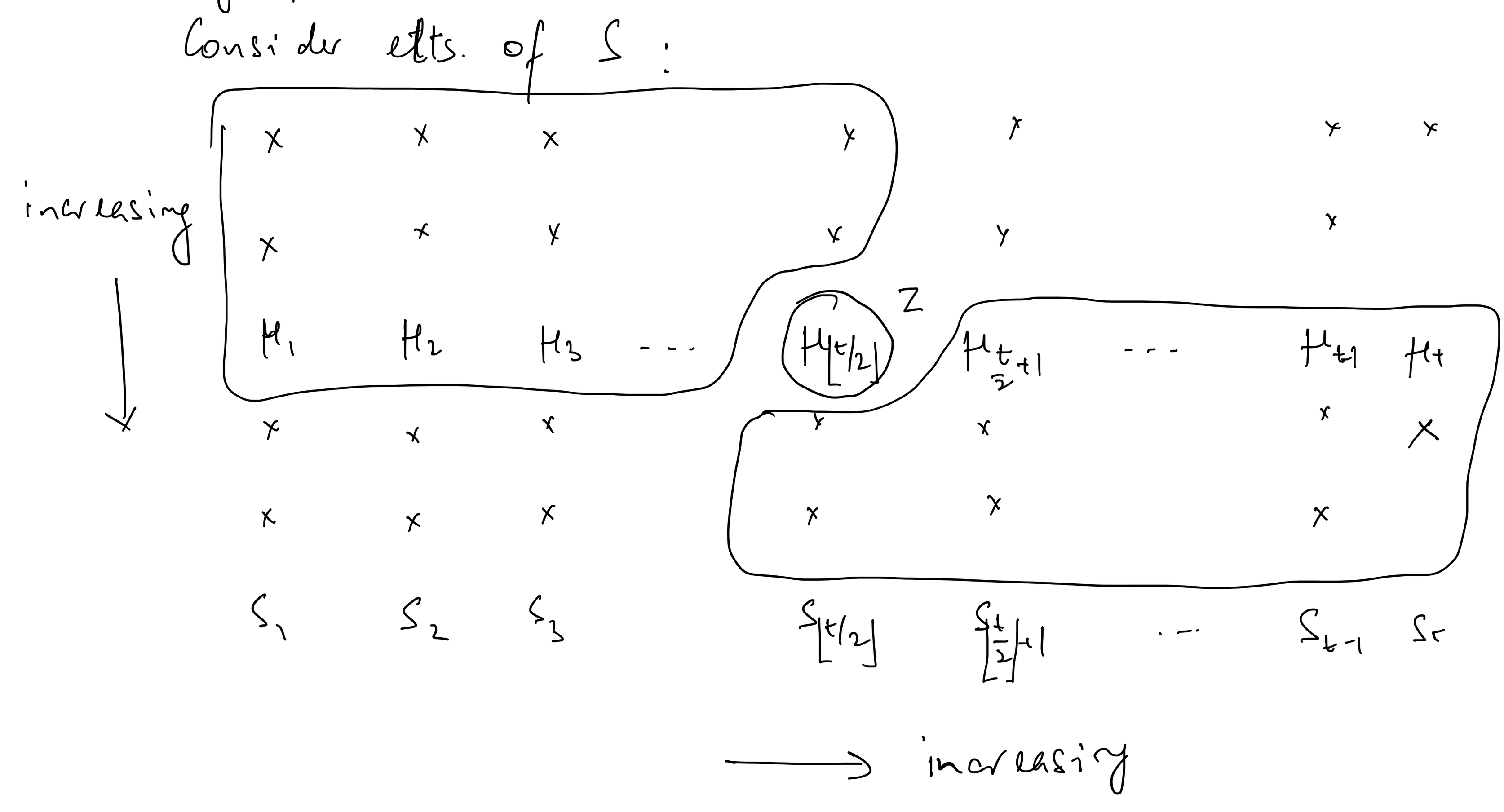
So, it remains to implement Approx-Split to run in linear time.

I/P: Array $S = (x_1, \ldots, x_n)$ of $n$ distinct integers
O/P: $z \in S$ s.t. $z$ is $k^{th}$ smallest elt., where
$\frac{n}{4} \leq k \leq \frac{3n}{4}$

Step 0: If $n \leq 60$, sort $S$, return median $\mu$
Step 1: Partition $S$ into $\lceil n/5 \rceil$ arrays, of 5 elts. each, except the last which has $\leq 5$ elts.
Let $t = \lceil \frac{n}{5} \rceil$, & $S_1, S_2, \ldots, S_t$ are the diff.

sets (note: not sorted!) ← $O(n)$
Step 2: In each $S_i$, find median $\mu_i$ ← $O(n)$
Step 3: Using earlier algo (Rank-Find $(S, k)$) find median ($\lfloor \frac{t}{2} \rfloor^{th}$ smallest elt.) of $\mu_1, \ldots, \mu_t$ (say $z$)
Return $z$ $T_{RF}(n/5)$

**Claim:** # elts. in $S$ smaller than $z$ lies b/w $\frac{n}{4}$ & $\frac{3n}{4}$
Proof by picture:
Consider elts. of $S$:




# elts. with value $\leq z$ is at least

$$3 \times \left\lfloor \frac{t}{2} \right\rfloor \geq 3 \times \left(\frac{t}{2} - 1\right)$$
$$\geq 3 \times \left(\frac{1}{2}\frac{n}{5} - 1\right) \quad \left(\text{since } t = \left\lceil \frac{n}{5} \right\rceil\right)$$
$$= \frac{3n}{10} - 3 \geq \frac{n}{4} \quad \text{for } n \geq 60$$

# elts. with value $\geq z$ is at least $\frac{n}{4}$ for $n \geq 60$

Hence, $z$ has rank $k$, where $\frac{n}{4} \leq k \leq \frac{3n}{4}$ ∎

Time Complexity:
$T_{AS}(n) = O(n) + T_{RF}(n/5)$
Overall:
$$T_{RF}(n) = O(n) + T_{AS}(n) + T_{RF}(3n/4)$$
$$= O(n) + T_{RF}(n/5) + T_{RF}(3n/4)$$
$$= O(n) \quad \text{(prove!)}$$

PROBLEM 1: Modify algorithm to run even if numbers are not distinct.

PROBLEM 2: Prove $T(n) = O(n) + T(n/5) + T(3n/4)$ gives $T(n) = O(n)$