

12-B4

## 12. 排序

选取

LinearSelect

邓俊辉

[deng@tsinghua.edu.cn](mailto:deng@tsinghua.edu.cn)

## linearSelect()

Let  $Q$  be a small constant

0. if (  $n = |A| < Q$  ) return trivialSelect(  $A, k$  )

1. else divide  $A$  evenly into  $n/Q$  subsequences (each of size  $Q$ )

2. Sort each subsequence and determine  $n/Q$  medians //e.g. by insertionsort

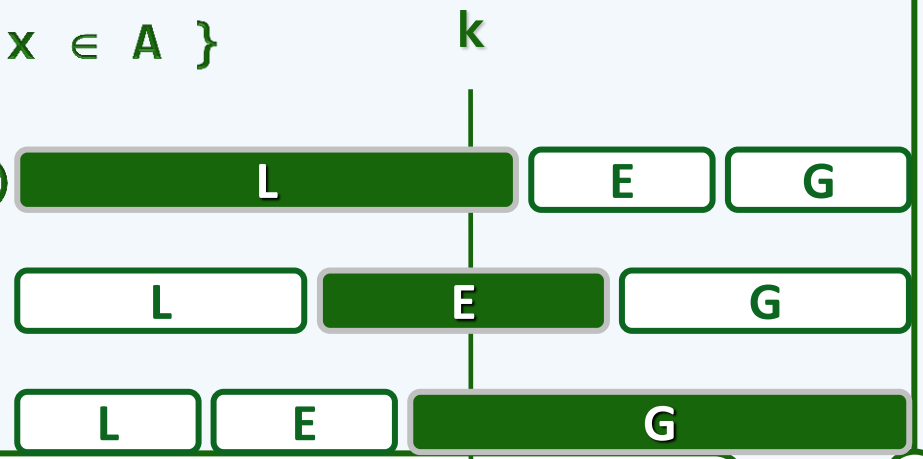
3. Call linearSelect to find  $M$ , median of the medians //by recursion

4. Let  $L / E / G = \{ x < / = / > M \mid x \in A \}$

5. if (  $k \leq |L|$  ) return linearSelect(  $L, k$  )

if (  $k \leq |L| + |E|$  ) return  $M$

return linearSelect(  $G, k - |L| - |E|$  )



## 复杂度

- ❖ 将linearSelect()算法的运行时间记作 $T(n)$
- ❖ 第0步： $O(1)$  =  $O(Q \log Q)$  //递归基：序列长度 $|A| \leq Q$
- ❖ 第1步： $O(n)$  //子序列划分
- ❖ 第2步： $O(n)$  =  $Q^2 \times n/Q$  //子序列各自排序，并找到中位数
- ❖ 第3步： $T(n/Q)$  //从 $n/Q$ 个中位数中，递归地找到全局中位数
- ❖ 第4步： $O(n)$  //划分子集L/E/G，并分别计数 —— 一趟扫描足矣
- ❖ 第5步： $T(3n/4)$  //为什么...

## 复杂度

❖ 在某种意义上，如上所确定的  $M$  必然 **不偏不倚**

至少各有  $n/4$  个元素，**不小**于/**不大**于  $M$

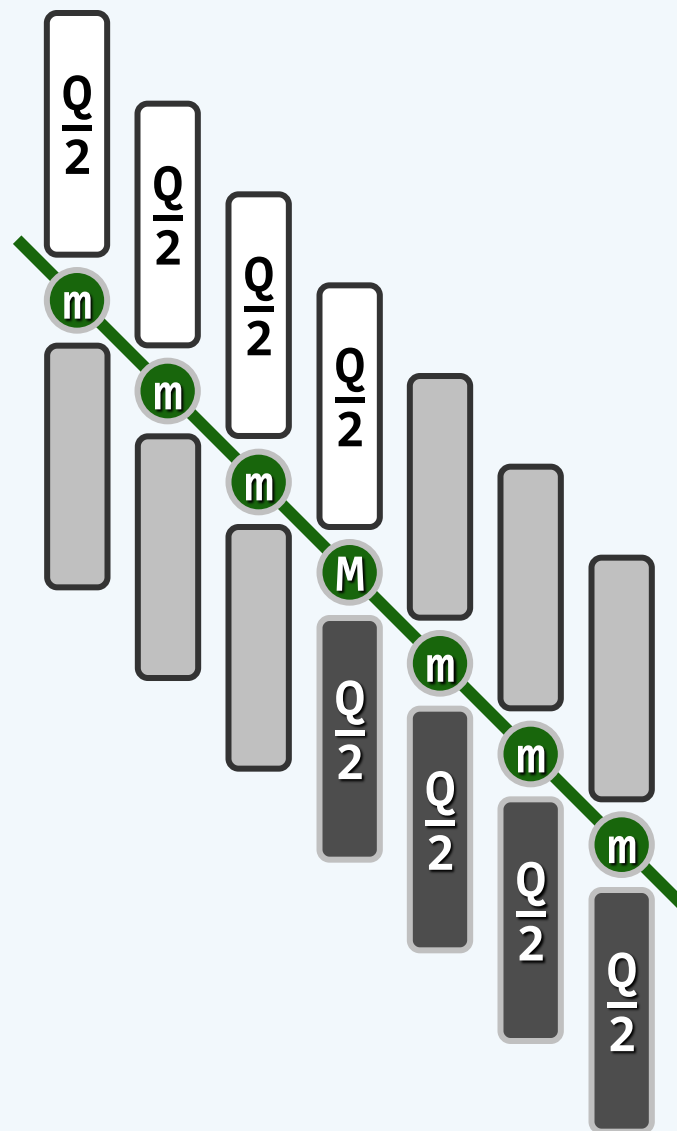
❖  $n/Q$  个中位数中，**至少半数** 不小于  $M$

而它们又各自不大于至少  $Q/2$  个元素

$$\frac{n/Q}{2} * \frac{Q}{2} = \frac{n}{4}$$

❖  $\min(|L|, |G|) + |E| \geq n/4$

$\max(|L|, |G|) \leq 3n/4$



## 复杂度

$$\diamond T(n) = \boxed{o(n)} + T(\boxed{n/Q}) + T(\boxed{3n/4})$$

❖ 为使之解作线性函数，只需保证

$$\boxed{n/Q} + \boxed{3n/4} < n$$

或等价地

$$\boxed{1/Q} + \boxed{3/4} < 1$$

❖ 比如，若取  $\boxed{Q = 5}$ ，则存在常数  $c$ ，使得

$$T(n) = cn + T(\boxed{n/5}) + T(\boxed{3n/4})$$

$$T(n) = o(\boxed{20c n}) = o(\boxed{n})$$

