

Homework 1

1) Coding:

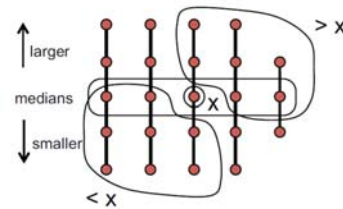
Write programs to implement the following $\text{SELECT}(S, i)$ algorithm. Any programming language is allowed. However, Python/C++/Java is recommended.

```
 $\text{SELECT}(S, i)$   
1  Pick  $x \in S$   $\triangleright$  cleverly  
2  Compute  $k = \text{rank}(x)$   
3   $B = \{y \in S \mid y < x\}$   
4   $C = \{y \in S \mid y > x\}$   
5  if  $k = i$   
6      return  $x$   
7  else if  $k > i$   
8      return  $\text{Select}(B, i)$   
9  else if  $k < i$   
10     return  $\text{Select}(C, i - k)$ 
```

Picking x Cleverly

Need to pick x so $\text{rank}(x)$ is not extreme.

- Arrange S into columns of size 5 ($\lceil \frac{n}{5} \rceil$ cols)
- Sort each column (bigger elements on top) (linear time)
- Find “median of medians” as x



2) Testing:

Generate an n -sized sequence S of random integers ($n \geq 1000$).

a) Will the $\text{SELECT}(S, i)$ algorithm find out the i^{th} smallest element correctly?

You can verify this by outputting x (step 6), e.g., given $i = 250/500/750$, respectively. (HINT: you can check the correctness of x by sorting the S and returning the i^{th} element.)

b) Do we have to arrange S into columns of size 5? Should we get the same result if we arrange S into columns of other size, e.g., 3 or 7?

Please compare the results and the running time when you use the $\text{SELECT}(S, i)$ algorithm to look for the median (i.e., $i = n/2$) under different column-size settings (e.g., 3, 5 and 7).

Please send your homework report (code + result) to the teacher's email box: 22799204@qq.com before February 27.

Please name your report file in the form of "StudentID+Alg+HW1.pdf".