Selection Problem

➪ **Order statistic**

0	1	2	3	4	5	6	7	8
4	1	10	9	7	12	8	2	15

Do we need to sort the whole list to determine the position of just one element?

➪ **What is the median?**

➪ **Similar questions**

✎ What's the first? Second? Last?

✎ Generally, what's the kth value?

Given k, what is *ordered* $a[k]$?

👁 ✎ Convention

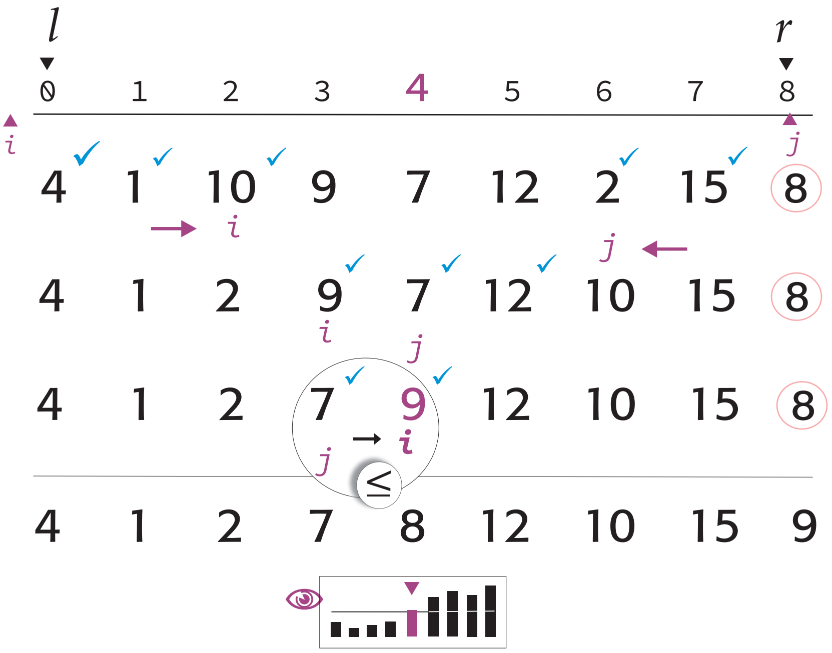
Selection Problem
Solution Idea

➪ [List] partition

What if the median happens to be picked as pivot in the **partition** procedure of a famous *quicksort*?

Ideally

```
$p \leftarrow a[r]$, $i \leftarrow l-1$, $j \leftarrow r$
loop
    while $a[{+}{+}i] < p$ do
    while $a[{-}{-}j] > p$ do
    if $i \geq j$ then break
    swap $a[i]$, $a[j]$
swap $a[i]$, $a[r]$
```



Quiz
At most, how many comparisons are needed for scans to cross? Why?

Selection Problem
How Much Better?

Brute-force assumes a presort based on a *quicksort* with the same partition scheme.

Exercise
Count key comparisons in previous slide by hand to verify the number 10 (failed comparisons also count).

Quiz
How many key-comps are needed if selection or bubble sort were used?

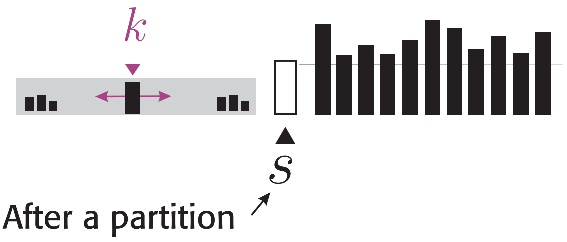
Number of key comparisons

- ✎ In this case: 10 instead of 28
- ✎ In general?

Selection Problem

Variable Decrease Algorithm

🗝 Element whose order is earlier than s must have a smaller value.



After a partition

Basic procedure

🕔 *Student setup iteration*

Exercise
Design non recursive and recursive algorithms using Slide 1 convention (*hint*: a tail recursion is easy).

✎ Partition

Split exactly at desired order (if its value happens to be picked as pivot).

✎ If split index is k, stop

$s > k$ ✎ If split after desired order discard right sub-list (fig) else discard left

Selection Problem
Example: Median

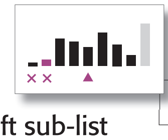
```
$p \leftarrow a[r]$, $i \leftarrow l - 1$, $j \leftarrow r$
loop
    while $a[{+}{+}i] < p$ do
    while $a[{-}{-}j] > p$ do
    if $i \geq j$ then break
    swap $a[i]$, $a[j]$
swap $a[i]$, $a[r]$
```

Stop when $s = 4$ ($= k$)

	0	1	2	3	4	5	6	7	8
i									j
	4	1	10	9	7	12	8	2	15
								j	i
$s > k$	4	1	10	9	7	12	8	2	15
$s = 8$									
	i							j	
	4	1	10	9	7	12	8	2	
	i	j							
	1	4	10	9	7	12	8	2	
	j	i							
$s < k$	1	2	10	9	7	12	8	4	
$s = 1$									

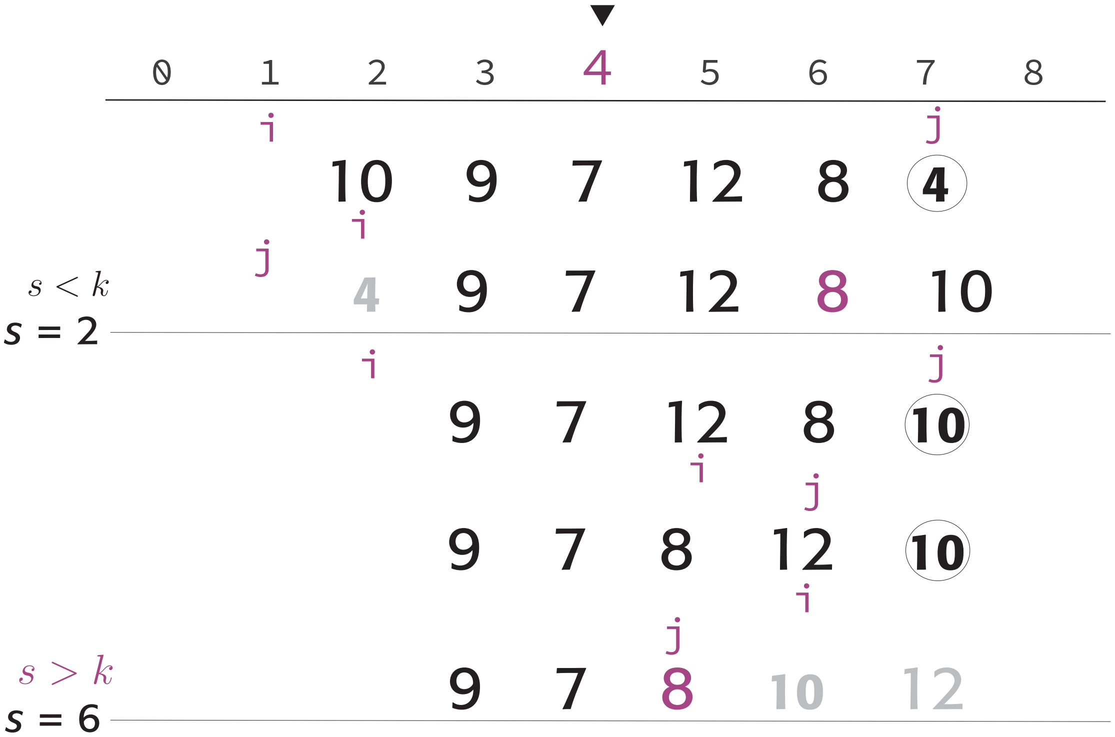
Discard right sub-list

Element whose order is 4 must be smaller than $a[s{=}8]$, therefore is in left sublist.



Discard left sub-list

Selection Problem
📖 Example

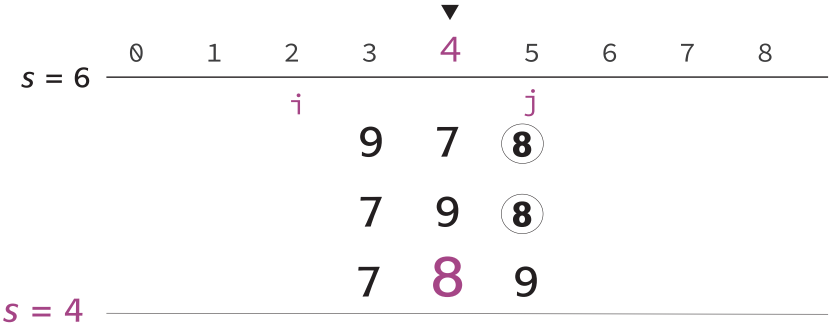


Discard left sub-list

Discard right sub-list

Variable Decrease-Conquer Example

Given $k = 4$, what is *ordered* $a[4]$? 8



Thinking Exercise
Note in this case how many times a partition had to be performed (iterated). How many times are needed to completely sort the list via a quicksort?

List size reduction sequence

$\boxed{9,\ 8,}\ 6,\ 5,\ 3$

Selection Problem
Quickselect

📖 ➪ **Lomuto partition**

Exercise
Compare partitions in terms of #comps, investigate resulting *quickselect* worst-case recurrences in *WolframAlpha*.

➪ Partition procedure

➪ Implementation style

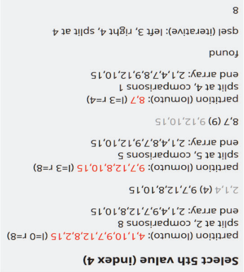
Quiz
When does the best and worst cases occur? What's the efficiency in each case?

➪ Compare efficiency

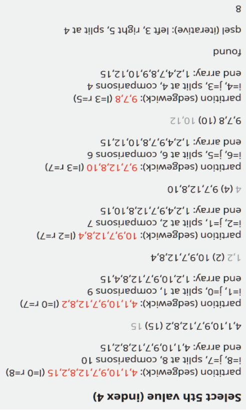
✎ Best and worst cases

✎ Average-case 👁 📖

Exercise
Compare *quickselect* and a brute-force median based on presort in terms of performance?



```
Select 5th value (index 4)

partition (lomuto): 4,1,10,9,7,12,8,2,15 (l=0 r=8)
split at 2, comparisons 8
end array: 2,1,4,9,7,12,8,10,15

2,1,4 (4) 9,7,12,8,10,15

partition (lomuto): 9,7,12,8,10,15 (l=3 r=8)
split at 5, comparisons 5
end array: 2,1,4,8,7,9,12,10,15

8,7 (9) 9,12,10,15

partition (lomuto): 8,7 (l=3 r=4)
split at 4, comparisons 1
end array: 2,1,4,7,8,9,12,10,15

found

qsel (iterative): left 3, right 4, split at 4

8
```



```
Select 5th value (index 4)

partition (sedgewick): 4,1,10,9,7,12,8,2,15 (l=0 r=8)
i=8, j=7, split at 8, comparisons 10
end array: 4,1,10,9,7,12,8,2,15

4,1,10,9,7,12,8,2 (15) 15

partition (sedgewick): 4,1,10,9,7,12,8,2 (l=0 r=7)
i=1, j=0, split at 1, comparisons 9
end array: 1,2,10,9,7,12,8,4,15

1,2 (2) 10,9,7,12,8,4

partition (sedgewick): 10,9,7,12,8,4 (l=2 r=7)
i=2, j=1, split at 2, comparisons 7
end array: 1,2,4,9,7,12,8,10,15

4 (4) 9,7,12,8,10

partition (sedgewick): 9,7,12,8,10 (l=3 r=7)
i=6, j=5, split at 6, comparisons 6
end array: 1,2,4,9,7,8,10,12,15

9,7,8 (10) 10,12

partition (sedgewick): 9,7,8 (l=3 r=5)
i=4, j=3, split at 4, comparisons 4
end array: 1,2,4,7,8,9,10,12,15

found

qsel (iterative): left 3, right 5, split at 4

8
```

Variable Decrease-Conquer
Class Exercise

```
$p \leftarrow a[r]$, $i \leftarrow l - 1$, $j \leftarrow r$
loop
    while $a[{+}{+}i] < p$ do
    while $a[{-}{-}j] > p$ do
    if $i \geq j$ then break
    swap $a[i]$, $a[j]$
swap $a[i]$, $a[r]$
```

0	1	2	3	4	5	6
19	12	8	12	15	11	10

Exercise
How many comparisons are needed to sort assuming a *quicksort*? Compare to *quickselect* if both use the same partition from previous example.

Quicksort performs 20 comps. vs. 13 for selection alg. (BTW, bubble sort would have cost 21 comps! why?, quicksekect with Lomuto 14).

✎ What's the third value? (*k*=2)

✎ What's the list size sequence?

✎ How many comparisons needed?