Review

`DSelect` $\rightarrow$ Deterministic Select

Goals of DSelect:

- Try and get the median of the medians

Runtimes

	DSelect	RSelect
Best		
Worst	$O(n)$	
Average	$O(n)$	$O(n)$

How many recursive calls does DSelect make? **2**

One for computing the median of C, and one for the partitioning step.

First Step

$$\leq \frac{n}{5}120 = 24n$$
$$\Rightarrow O(n)$$

Generalized

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

Rough Recurrence

Let $T(n) =$ max runtime of DSelect on an input array of length n.

$$\exists c \geq \text{ such that:}$$
$$T(1) = 1$$

$$T(n) \leq \underbrace{cn}_{\text{Sorting + Partition}} + \underbrace{T\left(\frac{n}{5}\right)}_{\text{Recurse to get good pivot}} + \underbrace{T(?)}_{\text{Recurse in Line 6/7}}$$

Goal: for every input array of length n, the runtime of DSelect is at most cn.

Note: Not as good as RSelect in practice. Reasons:

1. Constant hidden in the Big-O
2. Extra memory to store the C-array

Why do we use groups of 5?

Using an even number would make it harder to find the median of the medians.

Using 5, we know that the median of the medians is bigger than 3 out of 5 (60%) of th eelements in $\approx 50\%$ of the groups.

Bigger than 30% of all elements.

Sorted

Number of groups for n elements: $\lceil \frac{n}{5} \rceil = 6$
$\frac{1}{2}\lceil \frac{n}{5} \rceil + 1$

At least half of the groups have a median that is bigger than the median of the medians.

2/5

$$\text{At least } 3\left(\lceil \frac{1}{2} \lceil \frac{n}{5} \rceil \rceil - 2\right) \text{ elements are } > m$$