Recall Randomized Select 
\((A, p, q, i)\)

if \(p \geq q\) return \(A[p]\)

\(r := \text{Random Partition}(A[p, q])\)
if \(i = r - p + 1\) return \(A[r]\)
if \(i < r - p + 1\)
return Randomized Select 
\((A, p, r - 1, i)\)

// \(i > r - p + 1\)
return Randomized Select 
\((A, r + 1, q, i - (r - p + 1))\)

Let \(E(n)\) be the expected time of Randomized Select
on \(n\) items (worst case)

\[
E(n) \preceq \Theta(n) + \frac{1}{n} \sum_{k=0}^{n-1} \max(E(k), E(n-k-1)) \leq \Theta(n) + \frac{2}{n} \sum_{k=\lceil \frac{n}{2} \rceil}^{n-1} E(k)
\]

\(\therefore E(n) = O(n)\) (Substitution method)
Deterministic selection is nlg n time sort then random.

Randomized select in expected time $\Theta(n)$

Deterministic select in linear time $\Theta(n)$

Lemma: Suppose $\alpha, \beta > 0$ (real numbers) and $\alpha + \beta < 1$.

Suppose $T(n) = T(\alpha n) + T(\beta n) + \Theta(n)$

Then $T(n) = \Theta(n)$.

[If $\alpha + \beta = 1$ then $T(n) = \Theta(n \log n)$]

Proof: Substitution method.

For upper bound, say that $T(n) \leq T(\alpha n) + T(\beta n) + cn$

for all large enough $n$ ($c > 0$ is a constant).

Prove that there is a constant $d > 0$ such that $T(n) \leq dn$ for all large enough $n$.

Inductive case. Assume $n$ is large enough, and that $T(m) \leq dm$ for the relevant $m < n$.

\[
T(n) \leq T(\alpha n) + T(\beta n) + cn
\leq d\alpha n + d\beta n + cn
= d(\alpha + \beta)n + cn
= [d(\alpha + \beta) + c]n
\]
\[
\left[ \frac{d(x+\beta) + c}{d} \right]^n \\
\leq \frac{d}{n} \\
\text{provided} \\
d(x+\beta) + c \leq d \\
equiv: \\
d \geq \frac{c}{1 - (x+\beta)} > 0 \\
\text{constant} > 0
\]

For base case, choose \( d \) big enough to handle enough of the small values of \( n \) to get the induction rolling.

\[
\text{DSelect}(A, p, q, i) \\
\text{// returns } i^{th} \text{ smallest element of } A[p..q] \\
\text{// Precondition: } 1 \leq i \leq q-p+1 \\
\text{// Strategy:} \\
\quad - \text{find a pivot close to the median (guaranteed)} \\
\quad - \text{partition} \\
\]

\[
\begin{array}{cccccc}
B_1 & B_2 & B_3 & \cdots & \cdots & B_{n/5} \\
5 & 5 & 5 & 5 & 5 & \leq 5 \\
\end{array}
\]

\( \text{let } x_i \text{ be the median of } B_i \)

\( \text{\( O(n) \) time} \)
Let $T(n)$ be time for DSelect on $n$ items

Compute the median of $M$

$X_1, X_2, \ldots, X_n$

Call it $m$

(Recursive call to

$DSelect(M), \frac{n}{5}, \frac{n}{10}$)

Time so far:

$\mathcal{O}(n) + T\left(\frac{n}{5}\right)$

$\uparrow$

medians of blocks into $M$

$m = \text{median of } M$

Use $m$ as the pivot
to partition $A$

Claim that there are

$\geq \frac{3n}{10}$ elements of $A$

less than $m$

and at least $\frac{3n}{10}$ elements

$\geq m$.

In $M$ there are

$m$ med of $M - \text{size } \frac{n}{5}$

So $M$ has $\frac{n}{10}$ elements

$\leq m$

$\frac{n}{10}$ many $x_i \leq m$
For each $x_i$ there are $2^n$ elements of $B_i$, besides $x_i$ that are
$\leq x_i$, so are $\leq m$
So at least $\frac{3n}{10}$ elements in $A \leq m$
Similarly, at least
$\frac{3n}{10}$ elements in $A$
that are $\geq m$

Partition: $\geq \frac{3n}{10} - 1 \geq \frac{3n}{10} - 1$

Now recurse into the appropriate sublist.

The sublist has
$\leq \frac{7n}{10}$ elements

(time = $T(\frac{7n}{10})$)

Recap: Total time is

$T(n) = T(n) + T(\frac{n}{5})$

partition

&

block medians

$\alpha = \frac{1}{5}$

$\beta = \frac{7}{10}$

$\alpha + \beta = \frac{9}{10} < 1$
Site-3 blocks:

\[ T(n) = \Theta(n) + T\left(\frac{n}{3}\right) \]

\[ 2 \cdot \frac{n}{6} \quad \text{m} \quad + \quad T\left(\frac{2n}{3}\right) \]

\[ \frac{n}{3} \]

\[ T(n) = \Theta(n) + T\left(\frac{n}{7}\right) \]

\[ \frac{4n}{14} \quad \text{m} \quad + \quad T\left(\frac{5n}{7}\right) \]

\[ \frac{2n}{14} \]

Not really practical.

(Randomized Select used more in practice)