k-ésimo menor elemento

CLRS Sec 9.3

k-ésimo menor

Problema: Encontrar o k-ésimo menor elemento de $A[1\,.\,.\,n]$.

Suponha $A[1\,.\,.\,n]$ sem elementos repetidos.

Exemplo: 33 é o 4o. menor elemento de:

1									10	
22	99	32	88	34	**33**	11	97	55	66	A

1			4						10	
11	22	32	**33**	34	55	66	88	97	99	ordenado

k-ésimo menor

Recebe $A[p\,.\,.\,r)$ e k tal que $1 \leq k \leq r-p$
e devolve valor do k-ésimo menor elemento de $A[p\,.\,.\,r)$.

SELECT-ORD (A, p, r, k)
1 ORDENE (A, p, r)
2 **devolva** $A[p - 1 + k]$

SELECT-ORD pode ser implementado de modo a consumir tempo $\mathrm{O}(n \lg n)$ onde $n := r - p$.

k-ésimo menor

Recebe $A[p\,..\,r)$ e k tal que $1 \leq k \leq r-p$
e devolve valor do k-ésimo menor elemento de $A[p\,..\,r)$.

SELECT-ORD (A, p, r, k)
1 ORDENE (A, p, r)
2 **devolva** $A[p-1+k]$

SELECT-ORD pode ser implementado de modo a consumir tempo $\mathrm{O}(n \lg n)$ onde $n := r - p$.

Dá para fazer melhor?

Relembremos o Particione

Rearranja $A[p\,..\,r)$ de modo que $p \leq q < r$ e $A[p\,..\,q) \leq A[q] < A(q\,..\,r)$

$\textsc{Particione}\ (A, p, r)$

```
x ← A[r−1]            ▷ x é o "pivô"
i ← p
para j ← p até r − 2 faça
    se A[j] ≤ x
        então A[i] ↔ A[j]
              i ← i + 1
A[i] ↔ A[r − 1]       ▷ troca com o "pivô"
devolva i
```

Invariantes:

(i0) $A[p\,..\,i) \leq x$ (i1) $x < A[i\,..\,j)$ (i2) $A[r-1] = x$

Relembremos o Particione

Rearranja $A[p \mathinner{..} r)$ de modo que $p \le q < r$ e $A[p \mathinner{..} q) \le A[q] < A(q \mathinner{..} r)$

```
PARTICIONE (A, p, r)
    x ← A[r−1]            ▷ x é o "pivô"
    i ← p
    para j ← p até r − 2 faça
        se A[j] ≤ x
            então A[i] ↔ A[j]
                  i ← i + 1
    A[i] ↔ A[r − 1]       ▷ troca com o "pivô"
    devolva i
```

Consumo de tempo: $\Theta(n)$ onde $n := r - p$.

Algoritmo SELECT determinístico

Recebe $A[p\,.\,.\,r)$ e k tal que $1 \leq k \leq r{-}p$
e devolve valor do k-ésimo menor elemento de $A[p\,.\,.\,r)$.

```
SELECT (A, p, r, k)     ▷ devolve k-ésimo menor de A[p..r)
    se r − p = 1 então devolva A[p]
    q ← PARTICIONE (A, p, r)
    num≤pivô ← q − p + 1     ▷ nº de elementos em A[p..q]
    se k = num≤pivô
        então devolva A[q]
    se k < num≤pivô
        então devolva SELECT (A, p, q, k)
        senão devolva SELECT (A, q+1, r, k−num≤pivô)
```

Consumo de tempo no pior caso é $\Theta(n^2)$, onde $n := r - p$.

Seleção em tempo linear

Como fazer algo melhor?

Seleção em tempo linear

Como fazer algo melhor?

Vamos usar de novo divisão e conquista.

Veremos o algoritmo BFPRT,
de Blum, Floyd, Pratt, Rivest e Tarjan.

Seleção em tempo linear

Como fazer algo melhor?

Vamos usar de novo divisão e conquista.

Veremos o algoritmo ~~BFPRT~~ MoM,
de Blum, Floyd, Pratt, Rivest e Tarjan,
também conhecido como "Median of Medians"

Seleção em tempo linear

Como fazer algo melhor?

Vamos usar de novo divisão e conquista.

Veremos o algoritmo ~~BFPRT~~ MoM,
de Blum, Floyd, Pratt, Rivest e Tarjan,
também conhecido como "Median of Medians"

Se o pivô do PARTICIONE for a mediana do vetor,
qual seria o consumo de tempo do SELECT?

Seleção em tempo linear

Como fazer algo melhor?

Vamos usar de novo divisão e conquista.

Veremos o algoritmo ~~BFPRT~~ MoM,
de Blum, Floyd, Pratt, Rivest e Tarjan,
também conhecido como "Median of Medians"

Se o pivô do PARTICIONE for A mediana do vetor,
qual seria o consumo de tempo do SELECT?

A Mediana

Para os propósitos desta aula, vamos redefinir mediana:

A mediana de $A[1 \,..\, n]$ é o $\lfloor \frac{n+1}{2} \rfloor$-ésimo menor elemento de $A[1 \,..\, n]$

A Mediana

Para os propósitos desta aula, vamos redefinir mediana:

A mediana de $A[1\,..\,n]$ é o $\lfloor\frac{n+1}{2}\rfloor$-ésimo menor elemento de $A[1\,..\,n]$

```
MEDIANA-ORD (A, p, r)    ▷ mediana de A[p . . r)
    n ← r − p
    devolva SELECT-ORD (A, p, r, ⌊(n+1)/2⌋)
```

A Mediana

Para os propósitos desta aula, vamos redefinir mediana:

A mediana de $A[1\,..\,n]$ é o $\lfloor \frac{n+1}{2} \rfloor$-ésimo menor elemento de $A[1\,..\,n]$

MEDIANA-ORD (A, p, r) $\triangleright$ mediana de $A[p\,..\,r]$
1 $n \leftarrow r - p$
2 **devolva** SELECT-ORD $(A, p, r, \lfloor \frac{n+1}{2} \rfloor)$

MEDIANA-ORD pode ser implementado de modo a consumir tempo $\mathrm{O}(n \lg n)$, onde $n := r - p$.

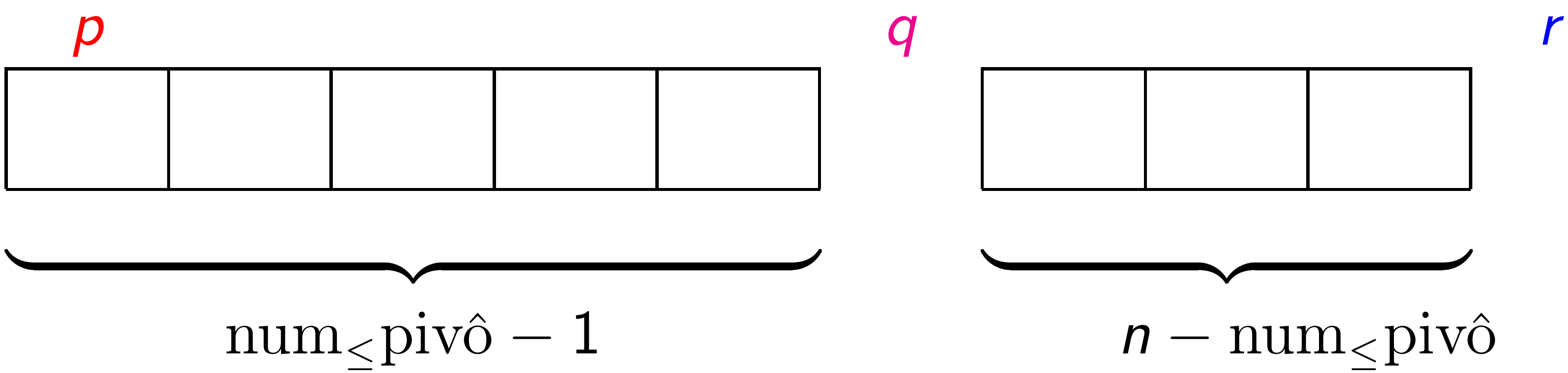
Select-MoM

Recebe $A[p\,.\,.\,r)$ e k tal que $1 \leq k \leq r-p$ e devolve o k-ésimo menor elemento de $A[p\,.\,.\,r)$.

$\text{SELECT-MOM}(A, p, r, k)$

```
se r − p = 1
    então devolva A[p]
q ← PARTICIONE-MOM(A, p, r)
num≤pivô ← q − p + 1        ▷ nº de elementos em A[p . . q]
se k = num≤pivô
    então devolva A[q]
se k < num≤pivô
    então devolva SELECT-MOM(A, p, q, k)
    senão devolva SELECT-MOM(A, q+1, r, k−num≤pivô)
```

Particione-MoM



Rearranja $A[p \mathinner{.\,.} r)$ e devolve um índice q, com $q \in [p \mathinner{.\,.} r)$,
tal que $A[p \mathinner{.\,.} q) \leq A[q] < A[q{+}1 \mathinner{.\,.} r)$ **e**

$$\max\{\mathrm{num}_{\leq}\mathrm{piv\hat{o}} - 1, n - \mathrm{num}_{\leq}\mathrm{piv\hat{o}}\} \leq \frac{7n}{10} + 3,$$

onde $n = r - p$ e $\mathrm{num}_{\leq}\mathrm{piv\hat{o}} = q - p + 1$.

Suponha que $P(n) :=$ consumo de tempo máximo do algoritmo PARTICIONE-MOM quando $n = r - p$

Consumo de tempo

$T(n) :=$ consumo de tempo máximo do algoritmo SELECT-MoM quando $n = r - p$ e $i = \mathrm{num}_{\leq}\mathrm{piv\hat{o}}$

linha	consumo de todas as execuções da linha
1–2	$= \Theta(1)$
3	$= P(n)$
4–7	$= \Theta(1)$
8	$\leq T(i-1)$
9	$\leq T(n-i)$

$$
\begin{aligned}
T(n) &\leq \Theta(1) + P(n) + \max_{1 \leq i \leq n} \{T(i-1), T(n-i)\} \\
&\leq \Theta(1) + P(n) + T(\lceil \tfrac{7n}{10} \rceil + 3)
\end{aligned}
$$

Particione-MoM

$n := r - p$

Suponha que n é um múltiplo de 5.

PARTICIONE-MOM (A, p, r)
1 $\quad n \leftarrow r - p$ $\quad \triangleright$ Usamos vetor auxiliar $B[0\,..\,\lceil n/5 \rceil)$
2 $\quad n_B \leftarrow \lceil n/5 \rceil$
3 $\quad$ **para** j **em** $[0\,..\,n_B)$ **faça**
4 $\qquad B[j] \leftarrow$ MEDIANA-ORD $(A, p + 5j, p + 5j + 5)$

Particione-MoM

$n := r - p$

Suponha que n é um múltiplo de 5.

PARTICIONE-MOM (A, p, r)

$n \leftarrow r - p$ $\quad \triangleright$ Usamos vetor auxiliar $B[0\,..\,\lceil n/5 \rceil)$

$n_B \leftarrow \lceil n/5 \rceil$

para j **em** $[0\,..\,n_B)$ **faça**

  $B[j] \leftarrow$ MEDIANA-ORD $(A, p + 5j, p + 5j + 5)$

Particione-MoM

$n := r - p$

Suponha que n é um múltiplo de 5.

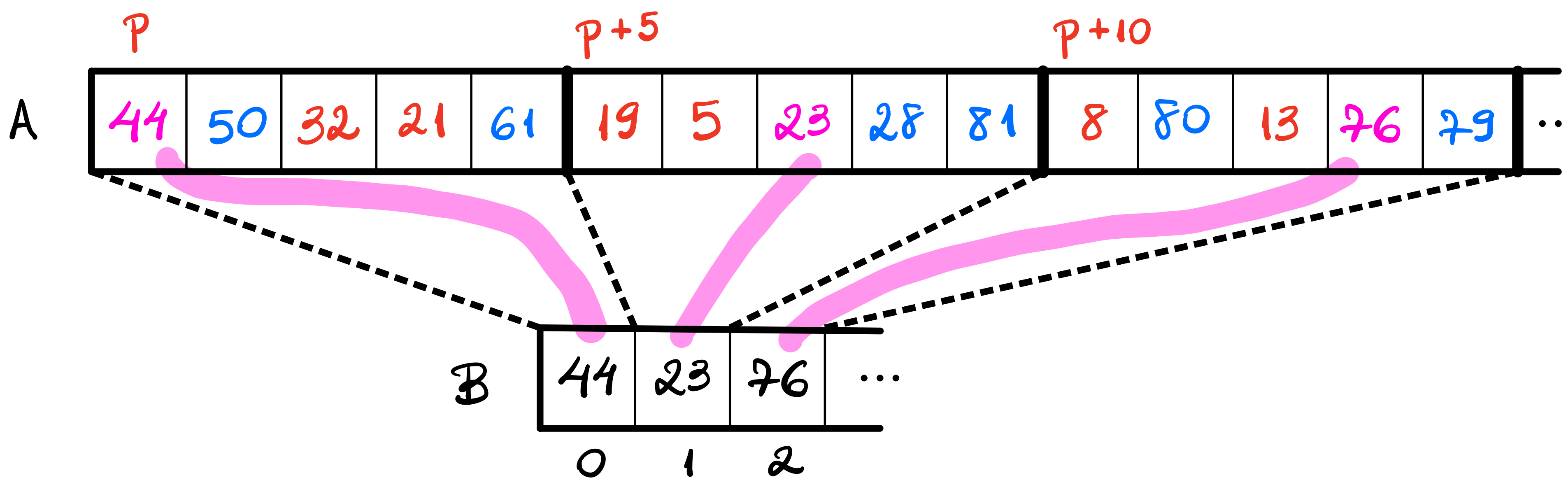
PARTICIONE-MOM (A, p, r)

$n \leftarrow r - p$ $\triangleright$ Usamos vetor auxiliar $B[0\,..\,\lceil n/5 \rceil)$

$n_B \leftarrow \lceil n/5 \rceil$

para j **em** $[0\,..\,n_B)$ **faça**

$\quad B[j] \leftarrow$ MEDIANA-ORD $(A, p + 5j, p + 5j + 5)$



Particione-MoM

$n := r - p$

Suponha que n é um múltiplo de 5. Se não, troque $p + 5j + 5$ na linha 4 por $\min\{p + 5j + 5, r\}$.

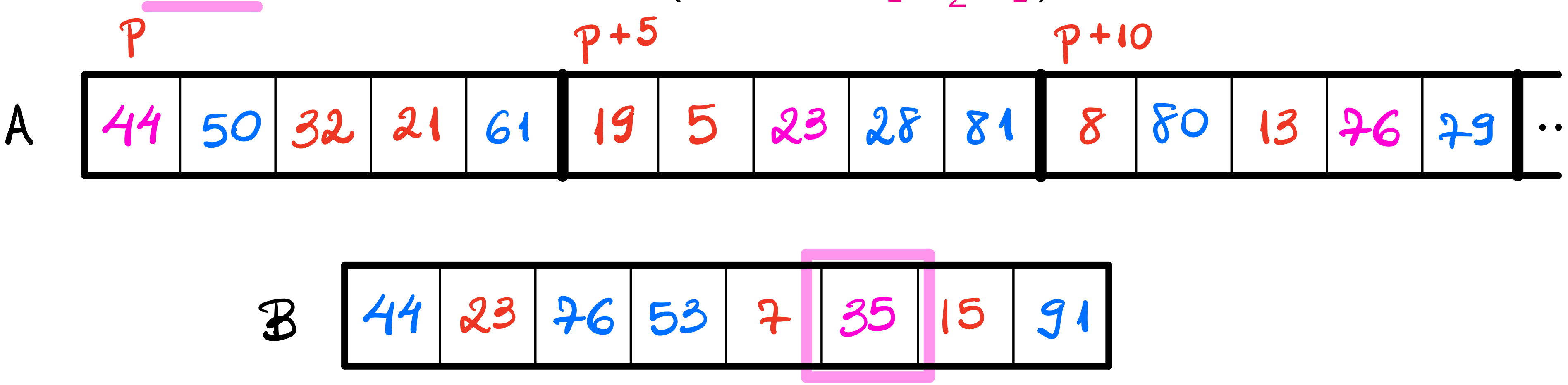
```
PARTICIONE-MoM (A, p, r)
1   n ← r − p        ▷ Usamos vetor auxiliar B[0 . . ⌈n/5⌉)
2   n_B ← ⌈n/5⌉
3   para j em [0 . . n_B) faça
4       B[j] ← MEDIANA-ORD (A, p + 5j, p + 5j + 5)
```

Particione-MoM

$n := r - p$

Suponha que n é um múltiplo de 5. Se não, troque $p + 5j + 5$ na linha 4 por $\min\{p + 5j + 5, r\}$.

PARTICIONE-MOM (A, p, r)
1 $\quad n \leftarrow r - p \qquad \triangleright$ Usamos vetor auxiliar $B[0 \,..\, \lceil n/5 \rceil)$
2 $\quad n_B \leftarrow \lceil n/5 \rceil$
3 $\quad$ **para** j **em** $[0 \,..\, n_B)$ **faça**
4 $\qquad B[j] \leftarrow$ MEDIANA-ORD $(A, p + 5j, p + 5j + 5)$
5 $\quad \text{MoM} \leftarrow$ SELECT-MOM $(B, 0, n_B, \lfloor \frac{n_B+1}{2} \rfloor) \quad \triangleright$ median of medians



Particione-MoM

$n := r - p$

Suponha que n é um múltiplo de 5. Se não, troque $p + 5j + 5$ na linha 4 por $\min\{p + 5j + 5, r\}$.

PARTICIONE-MoM (A, p, r)

1 $n \leftarrow r - p$ $\triangleright$ Usamos vetor auxiliar $B[0 \,..\, \lceil n/5 \rceil)$
2 $n_B \leftarrow \lceil n/5 \rceil$
3 **para** j **em** $[0 \,..\, n_B)$ **faça**
4 $\quad B[j] \leftarrow$ MEDIANA-ORD $(A, p + 5j, p + 5j + 5)$
5 $\mathrm{MoM} \leftarrow$ SELECT-MoM $(B, 0, n_B, \lfloor \frac{n_B+1}{2} \rfloor)$ $\triangleright$ median of medians
6 Encontre ℓ em $[p, r)$ tal que $A[\ell] = \mathrm{MoM}$
7 $A[\ell] \leftrightarrow A[r-1]$
8 **devolva** PARTICIONE (A, p, r)

Consumo de tempo do Particione-MoM

$P(n) :=$ consumo de tempo máximo do algoritmo PARTICIONE-MoM quando $n = r - p$

linha	consumo de todas as execuções da linha
1–2	$= \Theta(1)$
3–4	$= \lceil n/5 \rceil\, \Theta(1) \ = \ \Theta(n)$
5	$\leq T(\lceil n/5 \rceil)$
6	$= \Theta(n)$
7	$= \Theta(1)$
8	$= \Theta(n)$

$$P(n) \quad \leq \Theta(n) + T(\lceil n/5 \rceil)$$

Consumo de tempo do Select-MoM

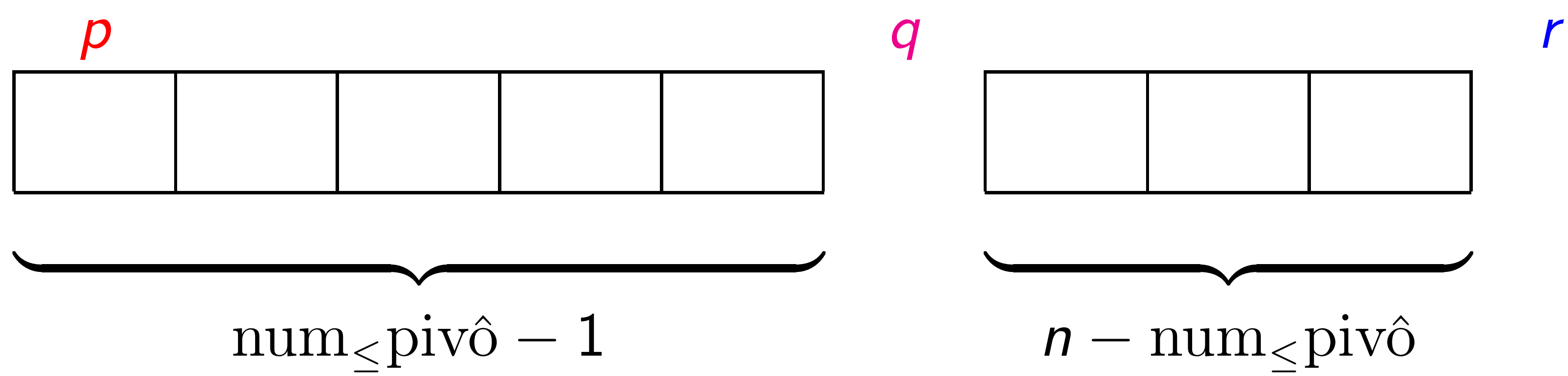
$T(n) :=$ consumo de tempo máximo do algoritmo SELECT-MoM quando $n = r - p$

Temos que

$$
\begin{aligned}
T(1) &= \Theta(1) \\
T(n) &\leq \Theta(1) + P(n) + T\left(\left\lceil \frac{7n}{10} \right\rceil + 3\right) \\
&\leq \Theta(1) + \Theta(n) + T\left(\left\lceil \frac{n}{5} \right\rceil\right) + T\left(\left\lceil \frac{7n}{10} \right\rceil + 3\right) \\
&= \Theta(n) + T\left(\left\lceil \frac{2n}{10} \right\rceil\right) + T\left(\left\lceil \frac{7n}{10} \right\rceil + 3\right)
\end{aligned}
$$

para todo $n \geq 2$.

Particione-MoM



Rearranja $A[p \mathinner{..} r)$ e devolve um índice q, com $q \in [p \mathinner{..} r)$, tal que $A[p \mathinner{..} q) \leq A[q] < A[q{+}1 \mathinner{..} r)$ **e**

$$\max\{\text{num}_{\leq}\text{pivô} - 1, n - \text{num}_{\leq}\text{pivô}\} \leq \frac{7n}{10} + 3,$$

onde $n = r - p$ e $\text{num}_{\leq}\text{pivô} = q - p + 1$.

Particione-MoM

1							$\left\lceil \frac{n}{5} \right\rceil$
3	10	5	2	21	42	8	
6	14	19	12	32	51	13	
7	15	23	35	44	53	76	91
95	20	28	38	50	87	79	93
99	24	81	64	61	89	80	

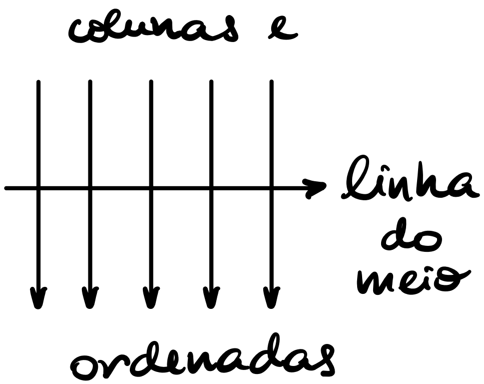
A

44	50	32	21	61	19	5	23	28	81	8	80	13	76	79	...

B

44	23	76	53	7	35	15	91

Particione-MoM



1							$\lceil \frac{n}{5} \rceil$
3	10	5	2	21	42	8	
6	14	19	12	32	51	13	
7	15	23	35	44	53	76	91
95	20	28	38	50	87	79	93
99	24	81	64	61	89	80	

A	44	50	32	21	61	19	5	23	28	81	8	80	13	76	79	...

B	44	23	76	53	7	35	15	91

Particione-MoM

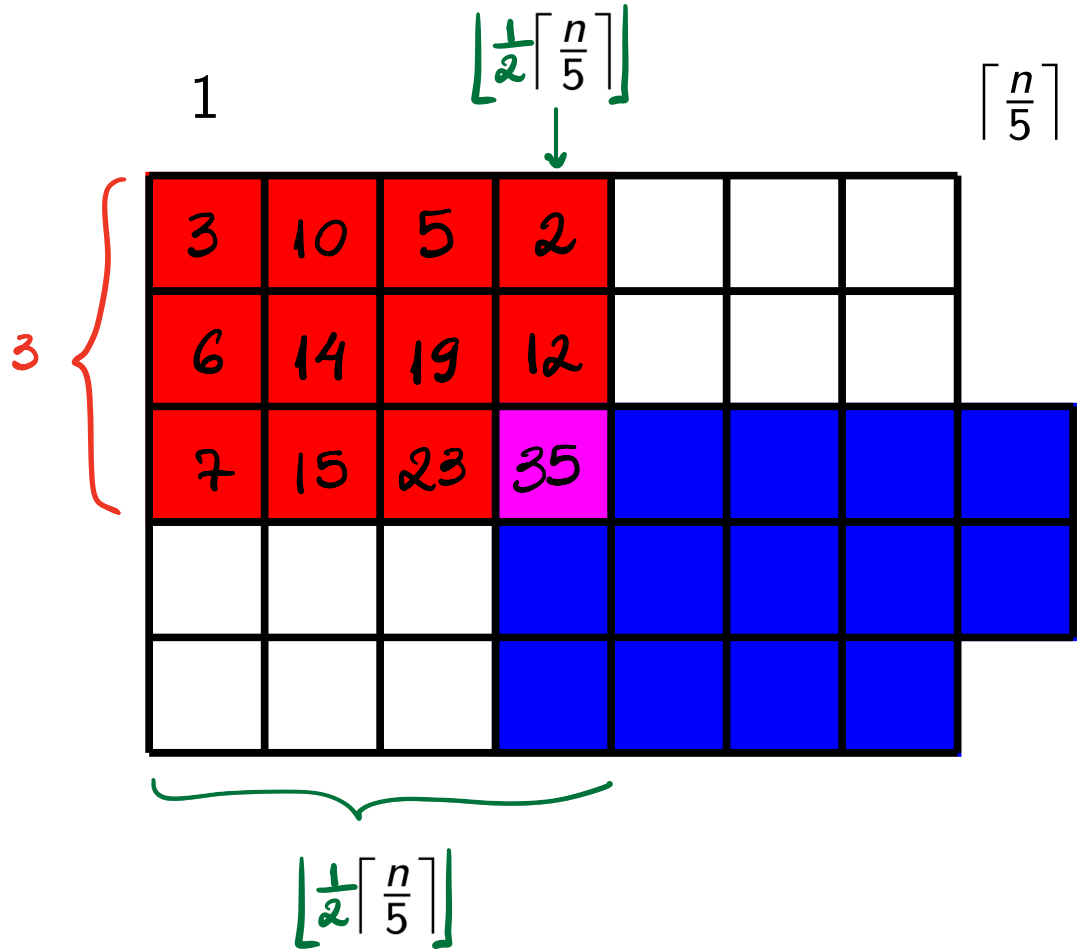
1							$\left\lceil \frac{n}{5} \right\rceil$
3	10	5	2	21	42	8	
6	14	19	12	32	51	13	
7	15	23	35	44	53	76	91
95	20	28	38	50	87	79	93
99	24	81	64	61	89	80	

"deleta" região azul (incl. pivô)

se $k < \mathrm{num}_{\leq}\mathrm{piv\hat{o}}$
    então devolva SELECT-MoM (A, p, q, k)
    senão devolva SELECT-MoM $(A, q{+}1, r, k{-}\mathrm{num}_{\leq}\mathrm{piv\hat{o}})$

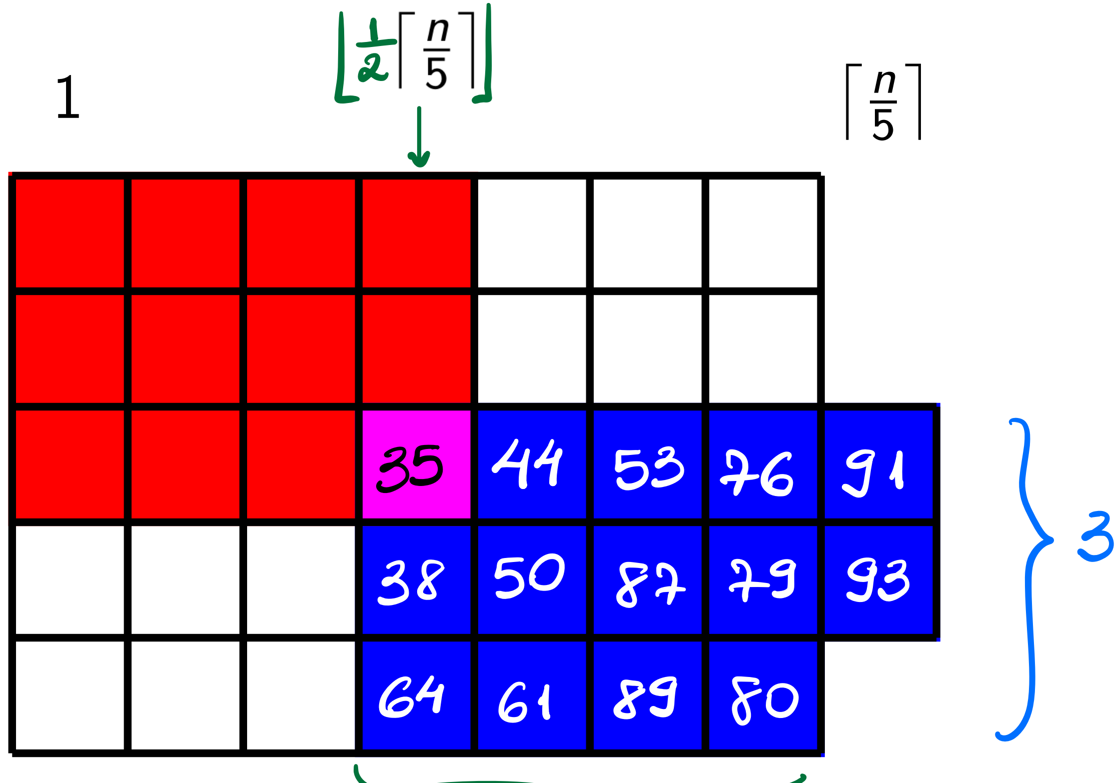
"deleta" região vermelha (incl. pivô)

Particione-MoM



$$\text{vermelhos} = 3 \cdot \left\lfloor \frac{1}{2} \left\lceil \frac{n}{5} \right\rceil \right\rfloor \geq 3 \cdot \left(\left\lceil \frac{1}{2} \left\lceil \frac{n}{5} \right\rceil \right\rceil - 1 \right) = 3 \left\lceil \frac{1}{2} \left\lceil \frac{n}{5} \right\rceil \right\rceil - 3$$

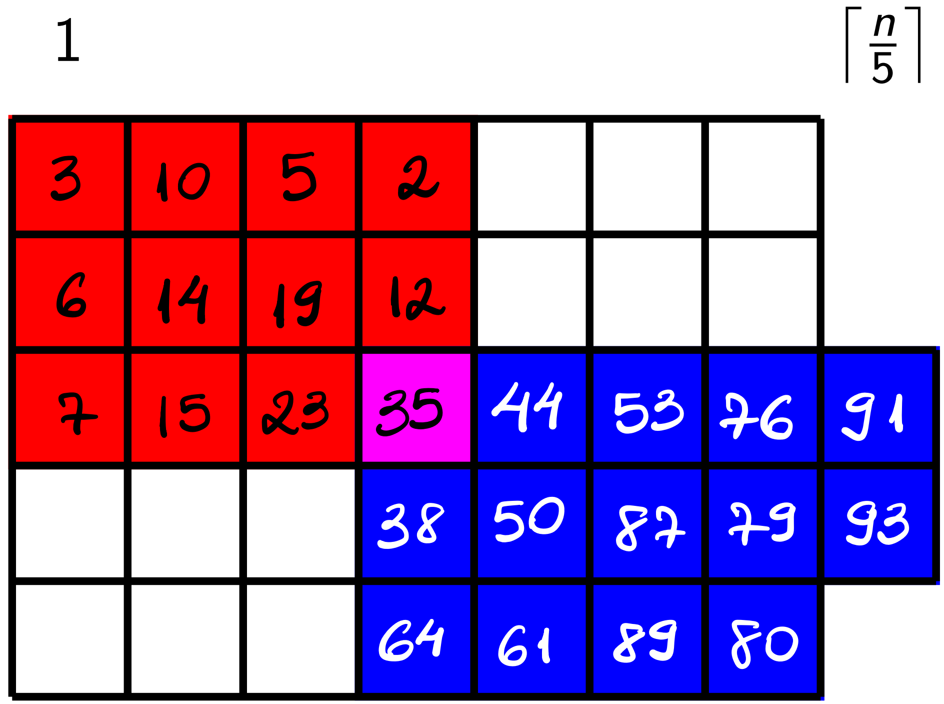
Particione-MoM



$$= \left\lceil \frac{n}{5} \right\rceil - \left\lfloor \frac{1}{2} \left\lceil \frac{n}{5} \right\rceil \right\rfloor = \left\lceil \frac{1}{2} \left\lceil \frac{n}{5} \right\rceil \right\rceil$$

$$\text{azuis} \geq 3 \cdot \left\lceil \frac{1}{2} \left\lceil \frac{n}{5} \right\rceil \right\rceil \geq 3 \cdot \left\lceil \frac{1}{2} \left\lceil \frac{n}{5} \right\rceil \right\rceil - 3$$

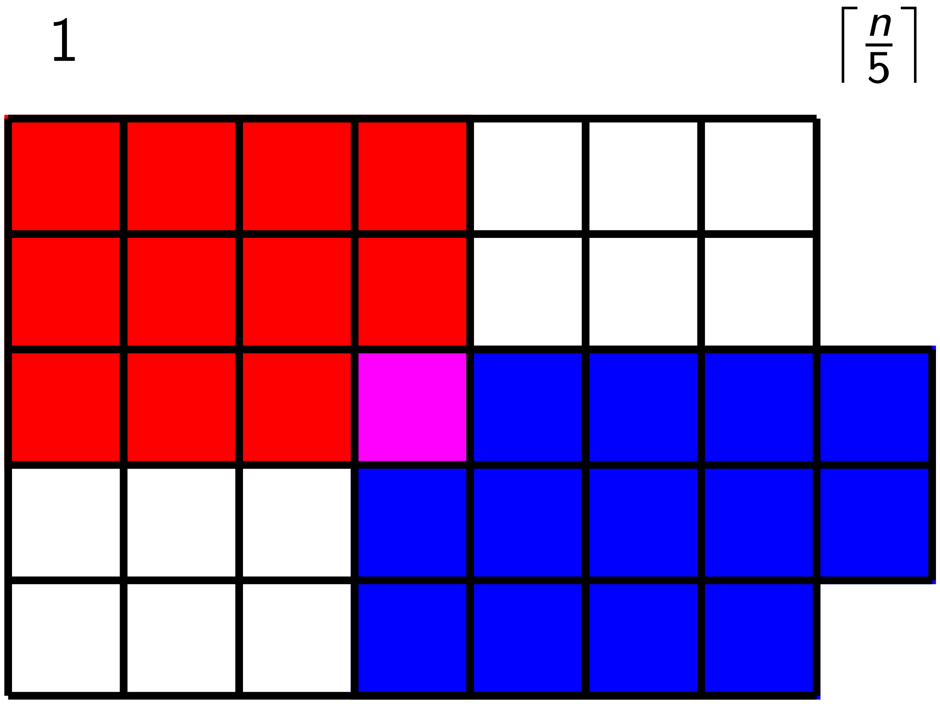
Particione-MoM



$$\text{vermelhos} \geq 3 \left\lceil \frac{1}{2} \left\lceil \frac{n}{5} \right\rceil \right\rceil - 3$$

$$\text{azuis} \geq 3 \cdot \left\lceil \frac{1}{2} \left\lceil \frac{n}{5} \right\rceil \right\rceil - 3$$

Particione-MoM



$$\max\{\mathrm{num}_{\leq}\mathrm{piv\hat{o}} - 1, n - \mathrm{num}_{\leq}\mathrm{piv\hat{o}}\} \leq n - \left(3 \left\lceil \frac{1}{2} \left\lceil \frac{n}{5} \right\rceil \right\rceil - 3\right)$$

$$\leq n - \left(\frac{3n}{10} - 3\right) = \frac{7n}{10} + 3$$

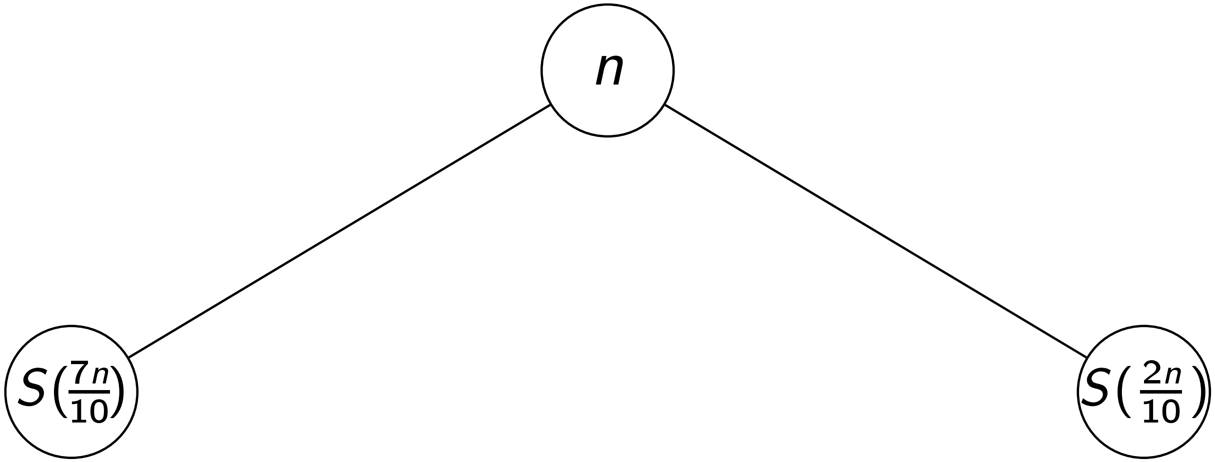
Um chute para a classe O?

Árvore da recorrência:

$$S(n)$$

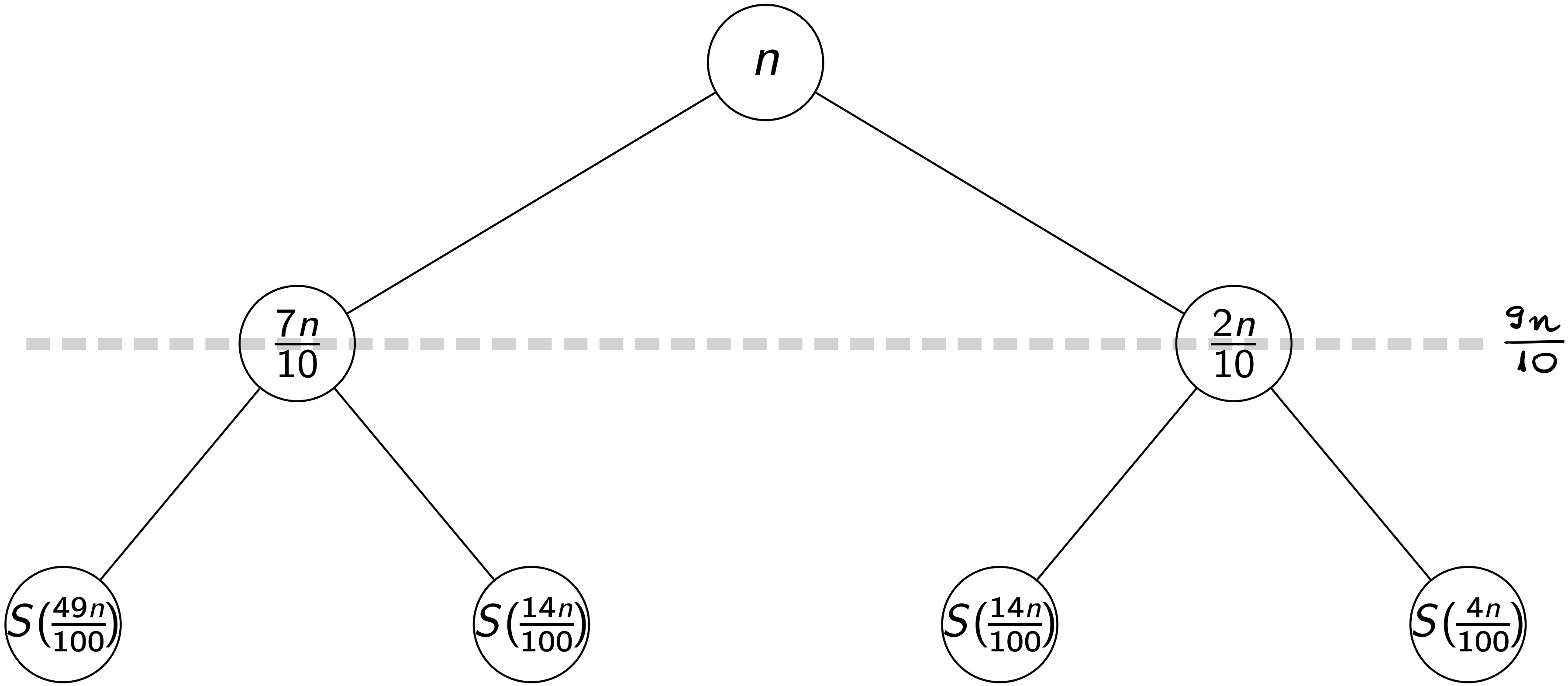
Um chute para a classe O?

Árvore da recorrência:



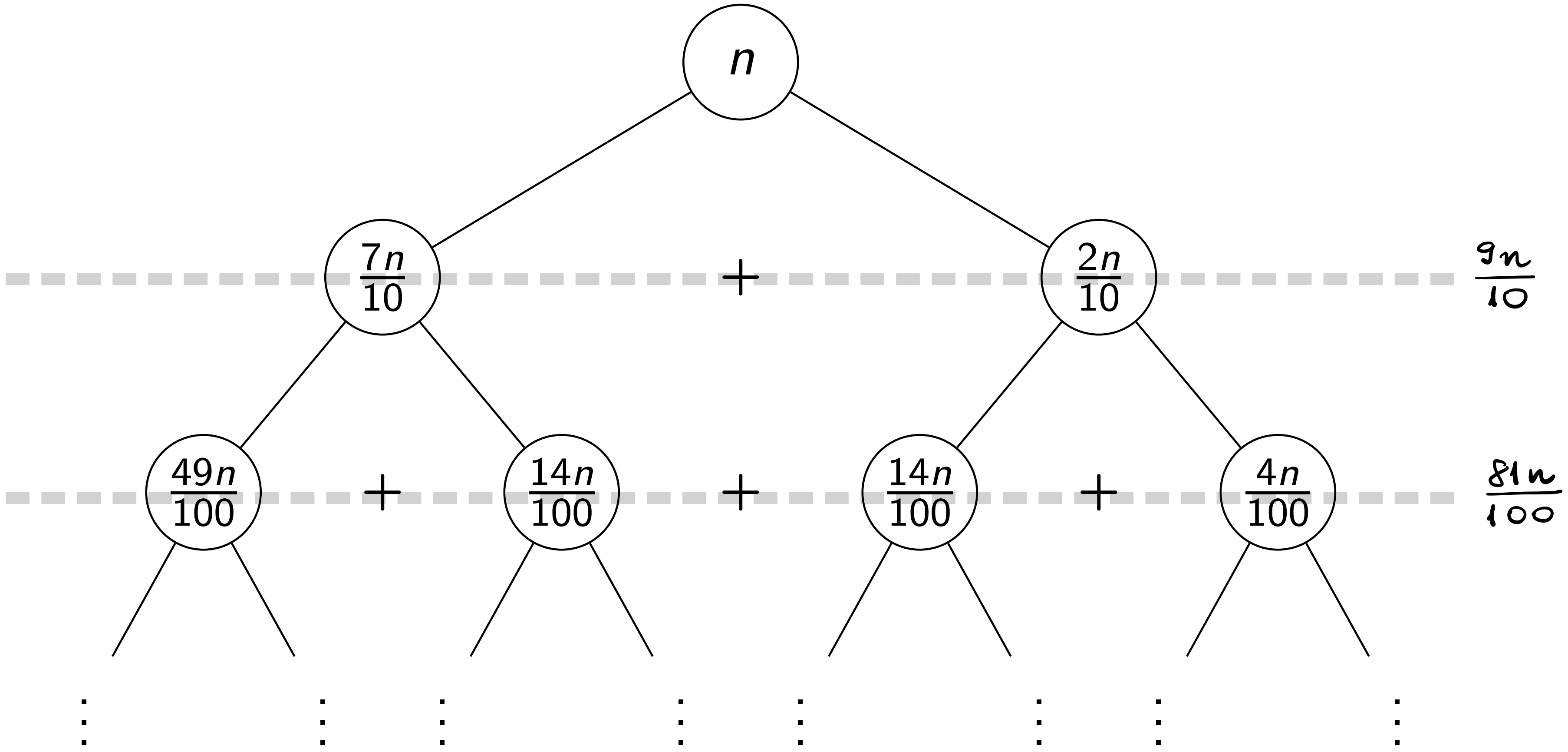
Um chute para a classe O?

Árvore da recorrência:



Um chute para a classe O?

Árvore da recorrência:



Contas

nível	0	1	2	...	$k-1$	k
soma	n	$\frac{9}{10}n$	$\frac{9^2}{10^2}n$	...	$\frac{9^{k-1}}{10^{k-1}}n$	$\frac{9^k}{10^k}n$

Contas

nível	0	1	2	$\ldots$	$k-1$	k
soma	n	$\frac{9}{10}n$	$\frac{9^2}{10^2}n$	$\ldots$	$\frac{9^{k-1}}{10^{k-1}}n$	$\frac{9^k}{10^k}n$

$$\frac{10^{k-1}}{9^{k-1}} < n \leq \frac{10^k}{9^k} \quad \Rightarrow \quad k = \lceil \log_{\frac{10}{9}} n \rceil$$

Contas

nível	0	1	2	$\ldots$	$k-1$	k
soma	n	$\frac{9}{10}n$	$\frac{9^2}{10^2}n$	$\ldots$	$\frac{9^{k-1}}{10^{k-1}}n$	$\frac{9^k}{10^k}n$

$$\frac{10^{k-1}}{9^{k-1}} < n \leq \frac{10^k}{9^k} \quad \Rightarrow \quad k = \lceil \log_{\frac{10}{9}} n \rceil$$

$$\begin{aligned}
S(n) &= n + \frac{9}{10}n + \cdots + \frac{9^{k-1}}{10^{k-1}}n + \frac{9^k}{10^k}n \\
&= \left(1 + \frac{9}{10} + \cdots + \frac{9^k}{10^k}\right)n \\
&= 10\left(1 - \frac{9^{k+1}}{10^{k+1}}\right)n \\
&< 10n
\end{aligned}$$

Consumo de tempo do Select-MoM

$T(n)$ pertence a mesma classe O que:

$$S(n) = 1 \quad \text{para todo } n < 100$$

$$S(n) = S\left(\left\lceil \frac{n}{5} \right\rceil\right) + S\left(\left\lceil \frac{7n}{10} \right\rceil + 3\right) + n \quad \text{para todo } n \geq 100$$

Vamos verificar que $S(n) \leq C\,n$ para todo $n \geq 1$, onde $C := 20$.

Prova: Se $n \in \{1, \ldots, 99\}$, então $S(n) = 1 < 20 \leq C\,n$.

Recorrência

Se $n \geq 100$, então

$$
\begin{aligned}
S(n) &\leq S\left(\left\lceil \frac{n}{5} \right\rceil\right) + S\left(\left\lceil \frac{7n}{10} \right\rceil + 3\right) + n \\
&\overset{\text{hi}}{\leq} C \cdot \left\lceil \frac{n}{5} \right\rceil + C \cdot \left(\left\lceil \frac{7n}{10} \right\rceil + 3\right) + n \\
&\leq C\left(\frac{n}{5} + 1\right) + C\left(\frac{7n}{10} + 4\right) + n \\
&= \tfrac{C}{5} n + C + \tfrac{7C}{10} n + 4C + n \\
&= \tfrac{2C}{10} n + \tfrac{7C}{10} n + n + 5C \\
&= \left(\tfrac{9C}{10} + 1\right) n + 5C \\
&\leq Cn \quad (\text{pois } n \geq 100 \text{ e } C = 20).
\end{aligned}
$$

Logo, $T(n)$ é $\mathrm{O}(n)$.

Conclusão

O consumo de tempo do SELECT-MoM é $O(n)$.