

Análise de Algoritmos

**Estes slides são adaptações de slides
do Prof. Paulo Feofiloff e do Prof. José Coelho de Pina.**

Quicksort e Select Aleatorizados

CLRS Secs 7.3, 7.4 e 9.2

Particione

Rearranja $A[p \dots d]$ de modo que $p \leq q \leq d$ e $A[p \dots q-1] \leq A[q] < A[q+1 \dots d]$

PARTICIONE (A, p, d)

```
1   $x \leftarrow A[d]$        $\triangleright x$  é o “pivô”
2   $i \leftarrow p-1$ 
3  para  $j \leftarrow p$  até  $d-1$  faça
4      se  $A[j] \leq x$ 
5          então  $i \leftarrow i + 1$ 
6               $A[i] \leftrightarrow A[j]$ 
7   $A[i+1] \leftrightarrow A[d]$ 
8  devolva  $i + 1$ 
```

Invariantes: no começo de cada iteração de 3–6,

(i0) $A[p \dots i] \leq x$ (i1) $A[i+1 \dots j-1] > x$ (i2) $A[d] = x$

Particione

Rearranja $A[p \dots d]$ de modo que $p \leq q \leq d$ e $A[p \dots q-1] \leq A[q] < A[q+1 \dots d]$

PARTICIONE (A, p, d)

```
1   $x \leftarrow A[d]$        $\triangleright x$  é o “pivô”
2   $i \leftarrow p-1$ 
3  para  $j \leftarrow p$  até  $d-1$  faça
4      se  $A[j] \leq x$ 
5          então  $i \leftarrow i + 1$ 
6               $A[i] \leftrightarrow A[j]$ 
7   $A[i+1] \leftrightarrow A[d]$ 
8  devolva  $i + 1$ 
```

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

Quicksort aleatorizado

PARTICIONE-ALEA(A, p, d)

1 $i \leftarrow \text{RANDOM}(p, d)$

2 $A[i] \leftrightarrow A[d]$

3 **devolva** PARTICIONE(A, p, d)

Quicksort aleatorizado

PARTICIONE-ALEA(A, p, d)

- 1 $i \leftarrow \text{RANDOM}(p, d)$
- 2 $A[i] \leftrightarrow A[d]$
- 3 **devolva** PARTICIONE(A, p, d)

QUICKSORT-ALE(A, p, d)

- 1 **se** $p < d$
- 2 **então** $q \leftarrow \text{PARTICIONE-ALE}(A, p, d)$
- 3 QUICKSORT-ALE($A, p, q - 1$)
- 4 QUICKSORT-ALE($A, q + 1, d$)

Quicksort aleatorizado

PARTICIONE-ALEA(A, p, d)

1 $i \leftarrow \text{RANDOM}(p, d)$

2 $A[i] \leftrightarrow A[d]$

3 **devolva** PARTICIONE(A, p, d)

QUICKSORT-ALE(A, p, d)

1 **se** $p < d$

2 **então** $q \leftarrow \text{PARTICIONE-ALEA}(A, p, d)$

3 QUICKSORT-ALE($A, p, q - 1$)

4 QUICKSORT-ALE($A, q + 1, d$)

Consumo esperado de tempo?

Quicksort aleatorizado

PARTICIONE-ALEA(A, p, d)

1 $i \leftarrow \text{RANDOM}(p, d)$

2 $A[i] \leftrightarrow A[d]$

3 **devolva** PARTICIONE(A, p, d)

QUICKSORT-ALE(A, p, d)

1 **se** $p < d$

2 **então** $q \leftarrow \text{PARTICIONE-ALEA}(A, p, d)$

3 QUICKSORT-ALE($A, p, q - 1$)

4 QUICKSORT-ALE($A, q + 1, d$)

Consumo esperado de tempo?

Basta contar o número esperado de comparações na linha 4 do PARTICIONE.

Consumo esperado de tempo

Basta contar o número esperado de comparações na linha 4 do **PARTICIONE**.

```
PARTICIONE ( $A, p, d$ )
1   $x \leftarrow A[d]$             $\triangleright x$  é o “pivô”
2   $i \leftarrow p - 1$ 
3  para  $j \leftarrow p$  até  $d - 1$  faça
4      se  $A[j] \leq x$ 
5          então  $i \leftarrow i + 1$ 
6               $A[i] \leftrightarrow A[j]$ 
7   $A[i+1] \leftrightarrow A[d]$ 
8  devolva  $i + 1$ 
```

Consumo de tempo esperado

Suponha $A[p \dots r]$ permutação de $1 \dots n$.

X_{ab} = número de comparações entre a e b na
linha 4 de **PARTICIONE** do QUICKSORT-ALE;

Queremos calcular

$$\begin{aligned} X &= \text{total de comparações } "A[j] \leq x" \\ &= \sum_{a=1}^{n-1} \sum_{b=a+1}^n X_{ab} \end{aligned}$$

Consumo de tempo esperado

Supondo $a < b$,

$$X_{ab} = \begin{cases} 1 & \text{se primeiro pivô em } \{a, \dots, b\} \text{ é } a \text{ ou } b \\ 0 & \text{caso contrário} \end{cases}$$

Qual a probabilidade de X_{ab} valer 1?

Consumo de tempo esperado

Supondo $a < b$,

$$X_{ab} = \begin{cases} 1 & \text{se primeiro pivô em } \{a, \dots, b\} \text{ é } a \text{ ou } b \\ 0 & \text{caso contrário} \end{cases}$$

Qual a probabilidade de X_{ab} valer 1?

$$\Pr \{X_{ab}=1\} = \frac{1}{b-a+1} + \frac{1}{b-a+1} = \mathbb{E}[X_{ab}]$$

Consumo de tempo esperado

Supondo $a < b$,

$$X_{ab} = \begin{cases} 1 & \text{se primeiro pivô em } \{a, \dots, b\} \text{ é } a \text{ ou } b \\ 0 & \text{caso contrário} \end{cases}$$

Qual a probabilidade de X_{ab} valer 1?

$$\Pr \{X_{ab}=1\} = \frac{1}{b-a+1} + \frac{1}{b-a+1} = \mathbb{E}[X_{ab}]$$

$$X = \sum_{a=1}^{n-1} \sum_{b=a+1}^n X_{ab}$$

$$\mathbb{E}[X] = \text{????}$$

Consumo de tempo esperado

$$E[X] = \sum_{a=1}^{n-1} \sum_{b=a+1}^n E[X_{ab}]$$

$$= \sum_{a=1}^{n-1} \sum_{b=a+1}^n \Pr\{X_{ab}=1\}$$

$$= \sum_{a=1}^{n-1} \sum_{b=a+1}^n \frac{2}{b-a+1}$$

$$= \sum_{a=1}^{n-1} \sum_{k=1}^{n-a} \frac{2}{k+1}$$

$$< \sum_{a=1}^{n-1} 2 \left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} \right)$$

$$< 2n \left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} \right) < 2n (1 + \ln n)$$

CLRS (A.7), p.1060

Conclusões

O consumo de tempo esperado do algoritmo
QUICKSORT-ALE é $O(n \log n)$.

Do **exercício 7.4-4 do CLRS** temos que

O consumo de tempo esperado do algoritmo
QUICKSORT-ALE é $\Theta(n \log n)$.

Select aleatorizado

PARTICIONE-ALEA(A, p, d)

- 1 $i \leftarrow \text{RANDOM}(p, d)$
- 2 $A[i] \leftrightarrow A[d]$
- 3 **devolva** PARTICIONE(A, p, d)

SELECT-ALEA(A, p, d, k)

- 1 **se** $p < d$
- 2 **então** $q \leftarrow \text{PARTICIONE-ALEA}(A, p, d)$
- 3 **se** $k = q - p + 1$
- 4 **então devolva** $A[q]$
- 5 **se** $k < q - p + 1$
- 6 **então** SELECT-ALEA($A, p, q - 1, k$)
- 7 **senão**
- SELECT-ALEA($A, q + 1, d, k - (q - p + 1)$)

Select aleatorizado

PARTICIONE-ALEA(A, p, d)

- 1 $i \leftarrow \text{RANDOM}(p, d)$
- 2 $A[i] \leftrightarrow A[d]$
- 3 **devolva** PARTICIONE(A, p, d)

SELECT-ALEA(A, p, d, k)

- 1 **se** $p < d$
- 2 **então** $q \leftarrow \text{PARTICIONE-ALEA}(A, p, d)$
- 3 **se** $k = q - p + 1$
- 4 **então devolva** $A[q]$
- 5 **se** $k < q - p + 1$
- 6 **então** SELECT-ALEA($A, p, q - 1, k$)
- 7 **senão**
- SELECT-ALEA($A, q + 1, d, k - (q - p + 1)$)

Consumo esperado de tempo?

Consumo de tempo esperado

Suponha $A[p \dots r]$ permutação de $1 \dots n$.

X_{ab} = número de comparações entre a e b na
linha 4 de **PARTICIONE** do **SELECT-ALEA**.

Observe que X_{ab} não é a mesma de antes.

De novo, queremos calcular

$$\begin{aligned} X &= \text{total de comparações } "A[j] \leq x" \\ &= \sum_{a=1}^{n-1} \sum_{b=a+1}^n X_{ab} \end{aligned}$$

Consumo de tempo esperado

Vamos supor que $k = n$.

Supondo $a < b$,

$$X_{ab} = \begin{cases} 1 & \text{se primeiro pivô em } \{a, \dots, n\} \text{ é } a \text{ ou } b \\ 0 & \text{caso contrário} \end{cases}$$

Qual a probabilidade de X_{ab} valer 1?

Consumo de tempo esperado

Vamos supor que $k = n$.

Supondo $a < b$,

$$X_{ab} = \begin{cases} 1 & \text{se primeiro pivô em } \{a, \dots, n\} \text{ é } a \text{ ou } b \\ 0 & \text{caso contrário} \end{cases}$$

Qual a probabilidade de X_{ab} valer 1?

$$\Pr \{X_{ab}=1\} = \frac{1}{n - a + 1} + \frac{1}{n - a + 1} = \mathbb{E}[X_{ab}]$$

Consumo de tempo esperado

Vamos supor que $k = n$.

Supondo $a < b$,

$$X_{ab} = \begin{cases} 1 & \text{se primeiro pivô em } \{a, \dots, n\} \text{ é } a \text{ ou } b \\ 0 & \text{caso contrário} \end{cases}$$

Qual a probabilidade de X_{ab} valer 1?

$$\Pr \{X_{ab}=1\} = \frac{1}{n - a + 1} + \frac{1}{n - a + 1} = \mathbb{E}[X_{ab}]$$

$$X = \sum_{a=1}^{n-1} \sum_{b=a+1}^n X_{ab}$$

$$\mathbb{E}[X] = \text{????}$$

Consumo de tempo esperado

$$\begin{aligned} \mathbb{E}[X] &= \sum_{a=1}^{n-1} \sum_{b=a+1}^n \mathbb{E}[X_{ab}] = \sum_{a=1}^{n-1} \sum_{b=a+1}^n \Pr\{X_{ab}=1\} \\ &= \sum_{a=1}^{n-1} \sum_{b=a+1}^n \frac{2}{n-a+1} \\ &= \sum_{a=1}^{n-1} \frac{2(n-a)}{n-a+1} \\ &< \sum_{a=1}^{n-1} 2 < 2n. \end{aligned}$$

Consumo de tempo esperado

$$\begin{aligned} E[X] &= \sum_{a=1}^{n-1} \sum_{b=a+1}^n E[X_{ab}] = \sum_{a=1}^{n-1} \sum_{b=a+1}^n \Pr\{X_{ab}=1\} \\ &= \sum_{a=1}^{n-1} \sum_{b=a+1}^n \frac{2}{n-a+1} \\ &= \sum_{a=1}^{n-1} \frac{2(n-a)}{n-a+1} \\ &< \sum_{a=1}^{n-1} 2 < 2n. \end{aligned}$$

Exercício: Refaça os cálculos para um k arbitrário.

Uma análise diferente

Suponha $A[p..r]$ permutação de $1..n$.

Seja

T_n = número de comparações realizadas em
uma chamada do **SELECT-ALEA** para $n := r - p + 1$.

Uma análise diferente

Suponha $A[p..r]$ permutação de $1..n$.

Seja

T_n = número de comparações realizadas em
uma chamada do **SELECT-ALEA** para $n := r - p + 1$.

Observe que $T_1 = 0$ e

$T_n \leq \max\{T_{q'-1}, T_{n-q'}\} + n - 1$ para $n \geq 2$,
onde $q' = q - p + 1$.

Uma análise diferente

Suponha $A[p..r]$ permutação de $1..n$.

Seja

T_n = número de comparações realizadas em
uma chamada do **SELECT-ALEA** para $n := r - p + 1$.

Observe que $T_1 = 0$ e

$T_n \leq \max\{T_{q'-1}, T_{n-q'}\} + n - 1$ para $n \geq 2$,
onde $q' = q - p + 1$.

Seja

$$X_h = \begin{cases} 1 & \text{se } \max\{k-1, n-k\} = h \\ 0 & \text{caso contrário.} \end{cases}$$

Temos que $T_n \leq \sum_{h=1}^{n-1} X_h T_h + n - 1$.

Uma análise diferente

Note que

$$E[X_h] = \Pr[X_h = 1] \leq \begin{cases} 0 & \text{se } h = 1, \dots, \lfloor n/2 \rfloor - 1 \\ \frac{1}{n} + \frac{1}{n} = \frac{2}{n} & \text{caso contrário.} \end{cases}$$

Uma análise diferente

Note que

$$E[X_h] = \Pr[X_h = 1] \leq \begin{cases} 0 & \text{se } h = 1, \dots, \lfloor n/2 \rfloor - 1 \\ \frac{1}{n} + \frac{1}{n} = \frac{2}{n} & \text{caso contrário.} \end{cases}$$

Para $n \geq 2$, como $T_n \leq \sum_{h=1}^{n-1} X_h T_h + n - 1$, temos que

$$\begin{aligned} E[T_n] &\leq \sum_{h=1}^{n-1} E[X_h T_h] + n - 1 \\ &= \sum_{h=1}^{n-1} E[X_h] E[T_h] + n - 1 \quad \text{pois } X_h \text{ e } T_h \text{ são indep.} \\ &\leq \frac{2}{n} \sum_{h=a}^{n-1} E[T_h] + n - 1 \quad \text{onde } a = \lfloor n/2 \rfloor \end{aligned}$$

Uma recorrência...

$E[T_n]$ é da mesma classe $O(\cdot)$
que a solução da seguinte recorrência:

$$\begin{aligned} S(1) &= 0, \quad \text{e} \\ S(n) &\leq \frac{2}{n} \sum_{h=1}^{n-1} S(h) + n - 1 \quad \text{para } n = 2, 3, \dots \end{aligned}$$

Uma recorrência...

$E[T_n]$ é da mesma classe $O(\cdot)$
que a solução da seguinte recorrência:

$$S(1) = 0, \quad \text{e}$$
$$S(n) \leq \frac{2}{n} \sum_{h=1}^{n-1} S(h) + n - 1 \quad \text{para } n = 2, 3, \dots$$

Vamos verificar que $S(n) \leq 4n$.

$S(n) \leq 4n$ **por indução em n**

$$S(1) = 0 \text{ e } S(n) \leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \text{ para } n = 2, 3, \dots$$

Se $n = 1$, temos que $S(1) = 0 < 4$.

$S(n) \leq 4n$ **por indução em n**

$$S(1) = 0 \text{ e } S(n) \leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \text{ para } n = 2, 3, \dots$$

Se $n \geq 2$, então

$$S(n) \leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \leq \frac{2}{n} \sum_{h=a}^{n-1} 4h + n - 1$$

$S(n) \leq 4n$ **por indução em n**

$$S(1) = 0 \text{ e } S(n) \leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \text{ para } n = 2, 3, \dots$$

Se $n \geq 2$, então

$$\begin{aligned} S(n) &\leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \leq \frac{2}{n} \sum_{h=a}^{n-1} 4h + n - 1 \\ &= \frac{8}{n} \left(\sum_{h=1}^{n-1} h - \sum_{h=1}^{a-1} h \right) + n - 1 = \frac{4}{n} (n(n-1) - a(a-1)) + n \end{aligned}$$

$S(n) \leq 4n$ **por indução em n**

$$S(1) = 0 \text{ e } S(n) \leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \text{ para } n = 2, 3, \dots$$

Se $n \geq 2$, então

$$\begin{aligned} S(n) &\leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \leq \frac{2}{n} \sum_{h=a}^{n-1} 4h + n - 1 \\ &= \frac{8}{n} \left(\sum_{h=1}^{n-1} h - \sum_{h=1}^{a-1} h \right) + n - 1 = \frac{4}{n} (n(n-1) - a(a-1)) + n \\ &= 4(n-1) - \frac{4a(a-1)}{n} + n - 1 \end{aligned}$$

$S(n) \leq 4n$ **por indução em n**

$S(1) = 0$ e $S(n) \leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1$ para $n = 2, 3, \dots$

Se $n \geq 2$, então

$$\begin{aligned} S(n) &\leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \leq \frac{2}{n} \sum_{h=a}^{n-1} 4h + n - 1 \\ &= \frac{8}{n} \left(\sum_{h=1}^{n-1} h - \sum_{h=1}^{a-1} h \right) + n - 1 = \frac{4}{n} (n(n-1) - a(a-1)) + n \\ &= 4(n-1) - \frac{4a(a-1)}{n} + n - 1 \\ &\leq 4(n-1) - \frac{4(n-1)(n-3)}{4n} + n - 1 \end{aligned}$$

$S(n) \leq 4n$ **por indução em n**

$$S(1) = 0 \text{ e } S(n) \leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \text{ para } n = 2, 3, \dots$$

Se $n \geq 2$, então

$$\begin{aligned} S(n) &\leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \leq \frac{2}{n} \sum_{h=a}^{n-1} 4h + n - 1 \\ &= \frac{8}{n} \left(\sum_{h=1}^{n-1} h - \sum_{h=1}^{a-1} h \right) + n - 1 = \frac{4}{n} (n(n-1) - a(a-1)) + n \\ &\leq 4(n-1) - \frac{4(n-1)(n-3)}{4n} + n - 1 \\ &= 4(n-1) + \frac{(n-1)n - (n-1)(n-3)}{n} \end{aligned}$$

$S(n) \leq 4n$ **por indução em n**

$$S(1) = 0 \text{ e } S(n) \leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \text{ para } n = 2, 3, \dots$$

Se $n \geq 2$, então

$$\begin{aligned} S(n) &\leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \leq \frac{2}{n} \sum_{h=a}^{n-1} 4h + n - 1 \\ &= \frac{8}{n} \left(\sum_{h=1}^{n-1} h - \sum_{h=1}^{a-1} h \right) + n - 1 = \frac{4}{n} (n(n-1) - a(a-1)) + n \\ &\leq 4(n-1) - \frac{4(n-1)(n-3)}{4n} + n - 1 \\ &= 4(n-1) + \frac{3(n-1)}{n} \end{aligned}$$

$S(n) \leq 4n$ **por indução em n**

$$S(1) = 0 \text{ e } S(n) \leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \text{ para } n = 2, 3, \dots$$

Se $n \geq 2$, então

$$\begin{aligned} S(n) &\leq \frac{2}{n} \sum_{h=a}^{n-1} S(h) + n - 1 \leq \frac{2}{n} \sum_{h=a}^{n-1} 4h + n - 1 \\ &= \frac{8}{n} \left(\sum_{h=1}^{n-1} h - \sum_{h=1}^{a-1} h \right) + n - 1 = \frac{4}{n} (n(n-1) - a(a-1)) + n \\ &\leq 4(n-1) - \frac{4(n-1)(n-3)}{4n} + n - 1 \\ &= 4(n-1) + \frac{3(n-1)}{n} \\ &< 4n - 4 + 3 < 4n. \end{aligned}$$

Conclusões

O consumo de tempo esperado do algoritmo
SELECT-ALEA é $O(n)$.