Análise de Algoritmos

Slides de Paulo Feofiloff

[com erros do coelho e agora também da cris]

Análise probabilística

CLRS 5.1, C.2, C.3

Máximo

Problema: Encontrar o elemento máximo de um vetor $A[1\,..\,n]$ de números inteiros positivos distintos.

MAX (A, n)
1 $\quad max \leftarrow 0$
2 $\quad$ **para** $i \leftarrow 1$ **até** n **faça**
3 $\qquad$ **se** $A[i] > max$
4 $\qquad\quad$ **então** $max \leftarrow A[i]$
5 $\quad$ **devolva** max

Quantas vezes a linha 4 é executada?

Máximo

Problema: Encontrar o elemento máximo de um vetor $A[1\,..\,n]$ de números inteiros positivos distintos.

```
MAX (A, n)
1    max ← 0
2    para i ← 1 até n faça
3        se A[i] > max
4            então max ← A[i]
5    devolva max
```

Quantas vezes a linha 4 é executada?
Melhor caso, pior caso, caso médio?

Máximo

Problema: Encontrar o elemento máximo de um vetor $A[1\,..\,n]$ de números inteiros positivos distintos.

MAX (A, n)
1 $\quad max \leftarrow 0$
2 $\quad$ **para** $i \leftarrow 1$ **até** n **faça**
3 $\qquad$ **se** $A[i] > max$
4 $\qquad\quad$ **então** $max \leftarrow A[i]$
5 $\quad$ **devolva** max

Quantas vezes a linha 4 é executada?
Melhor caso, pior caso, caso médio?

Suponha que $A[1\,..\,n]$ é permutação aleatória uniforme de $1, \ldots, n$

Cada permutação tem probabilidade $1/n!$.

Um pouco de probabilidade

$(S, \Pr)$ espaço de probabilidade
$S =$ conjunto finito (eventos elementares)
$\Pr\{\}$ = (distribuição de probabilidades) função de S em $[0,1]$ tal que

- p1. $\Pr\{s\} \geq 0$;
- p2. $\Pr\{S\} = 1$; e
- p3. $R, T \subseteq S, R \cap T = \emptyset \Rightarrow \Pr\{R \cup T\} = \Pr\{R\} + \Pr\{T\}$.

$\Pr\{U\}$ é abreviação de $\sum_{u \in U} \Pr\{u\}$.

Um pouco de probabilidade

$(S, \Pr)$ espaço de probabilidade
$S =$ conjunto finito (eventos elementares)
$\Pr\{\} =$ (distribuição de probabilidades) função de S em $[0, 1]$ tal que

p1. $\Pr\{s\} \geq 0$;

p2. $\Pr\{S\} = 1$; e

p3. $R, T \subseteq S, R \cap T = \emptyset \Rightarrow \Pr\{R \cup T\} = \Pr\{R\} + \Pr\{T\}$.

$$\Pr\{U\} \text{ é abreviação de } \sum_{u \in U} \Pr\{u\}.$$

No problema do máximo:

- S é o conjunto das permutações dos números em $A[1 \mathinner{.\,.} n]$;
- na distribuição uniforme, para cada $s \in S$, $\Pr\{s\} = 1/n!$.

Mais um pouco de probabilidade

Um evento é um subconjunto de S.

Mais um pouco de probabilidade

Um evento é um subconjunto de S.

No problema do máximo, eventos são subconjuntos de permutações de $A[1\,..\,n]$.

Exemplo.

$$U := \{\text{permutações de } A[1\,..\,n] \text{ em que } A[n] \text{ é máximo}\}$$

é um evento de S.

Mais um pouco de probabilidade

Um evento é um subconjunto de S.

No problema do máximo, eventos são subconjuntos de permutações de $A[1\,..\,n]$.

Exemplo.

$$U := \{\text{permutações de } A[1\,..\,n] \text{ em que } A[n] \text{ é máximo}\}$$

é um evento de S.

Se $\Pr\{\}$ é distribuição uniforme, então

$$\Pr\{U\} = ???.$$

Mais um pouco de probabilidade

Um evento é um subconjunto de S.

No problema do máximo, eventos são subconjuntos de permutações de $A[1\,..\,n]$.

Exemplo.

$$U := \{\text{permutações de } A[1\,..\,n] \text{ em que } A[n] \text{ é máximo}\}$$

é um evento de S.

Se $\Pr\{\}$ é distribuição uniforme, então

$$\Pr\{U\} = 1/n.$$

Mais um pouco de probabilidade

Uma variável aleatória é uma função númerica definida sobre os eventos elementares.

Mais um pouco de probabilidade

Uma variável aleatória é uma função númerica definida sobre os eventos elementares.

Exemplo de variável aleatória

$X(A) :=$ número de execuções da linha 4 em MAX(A, n)

Mais um pouco de probabilidade

Uma variável aleatória é uma função númerica definida sobre os eventos elementares.

Exemplo de variável aleatória

$X(A) :=$ número de execuções da linha 4 em MAX(A, n)

"$X = k$" é uma abreviação de $\{s \in S : X(s) = k\}$

Esperança $\mathrm{E}[X]$ de uma variável aleatória X

$$\mathrm{E}[X] = \sum_{k \in X(S)} k \cdot \Pr\{X = k\} = \sum_{s \in S} X(s) \cdot \Pr\{s\}$$

Mais um pouco de probabilidade

Uma variável aleatória é uma função númerica definida sobre os eventos elementares.

Exemplo de variável aleatória

$X(A) :=$ número de execuções da linha 4 em MAX(A, n)

"$X = k$" é uma abreviação de $\{s \in S : X(s) = k\}$

Esperança $\mathrm{E}[X]$ de uma variável aleatória X

$$\mathrm{E}[X] = \sum_{k \in X(S)} k \cdot \Pr\{X = k\} = \sum_{s \in S} X(s) \cdot \Pr\{s\}$$

Linearidade da esperança: $\mathrm{E}[\alpha X + Y] = \alpha \mathrm{E}[X] + \mathrm{E}[Y]$

De volta ao máximo

Problema: Encontrar o elemento máximo de um vetor $A[1\,..\,n]$ de números inteiros distintos.

MAX (A, n)
1 $\quad max \leftarrow 0$
2 $\quad$ **para** $i \leftarrow 1$ **até** n **faça**
3 $\qquad$ **se** $A[i] > max$
4 $\qquad\quad$ **então** $max \leftarrow A[i]$

---

5 $\quad$ **devolva** max

Quantas vezes a linha 4 é executada no caso médio?

Suponha que $A[1\,..\,n]$ é permutação aleatória uniforme de $1, \ldots, n$

Cada permutação tem probabilidade $1/n!$.

Exemplos

$A[1\,..\,2]$	linha 4
1,2	2
2,1	1
$\mathrm{E}[X]$	3/2

$A[1\,..\,3]$	linha 4
1,2,3	3
1,3,2	2
2,1,3	2
2,3,1	2
3,1,2	1
3,2,1	1
$\mathrm{E}[X]$	11/6

Mais um exemplo

$A[1\,..\,4]$	linha 4	$A[1\,..\,4]$	linha 14
1,2,3,4	4	3,1,2,4	2
1,2,4,3	3	3,1,4,2	2
1,3,2,4	3	3,2,1,4	2
1,3,4,2	3	3,2,4,1	2
1,4,2,3	2	3,4,1,2	2
1,4,3,2	2	3,4,2,1	2
2,1,3,4	3	4,1,2,3	1
2,1,4,3	2	4,1,3,2	1
2,3,1,4	3	4,2,1,3	1
2,3,4,1	3	4,2,3,1	1
2,4,1,3	2	4,3,1,2	1
2,4,3,1	2	4,3,2,1	1

$\mathrm{E}[X]$ 50/24

Variáveis aleatórias

$$X \quad = \quad \text{número total de execuções da linha 4}$$

Variáveis aleatórias

X = número total de execuções da linha 4

$$X_i = \begin{cases} 1 \text{ se "}max \leftarrow A[i]\text{" é executado} \\ 0 \text{ caso contrário} \end{cases}$$

X = número total de execuções da linha 4

$= X_1 + \cdots + X_n$

Variáveis aleatórias

X = número total de execuções da linha 4

$$X_i = \begin{cases} 1 \text{ se ``}max \leftarrow A[i]\text{'' é executado} \\ 0 \text{ caso contrário} \end{cases}$$

X = número total de execuções da linha 4

$= X_1 + \cdots + X_n$

Esperanças:

$\mathrm{E}[X_i]$ = probabilidade de que $A[i]$ seja máximo em $A[1\,..\,i]$

$= 1/i$

Esperança

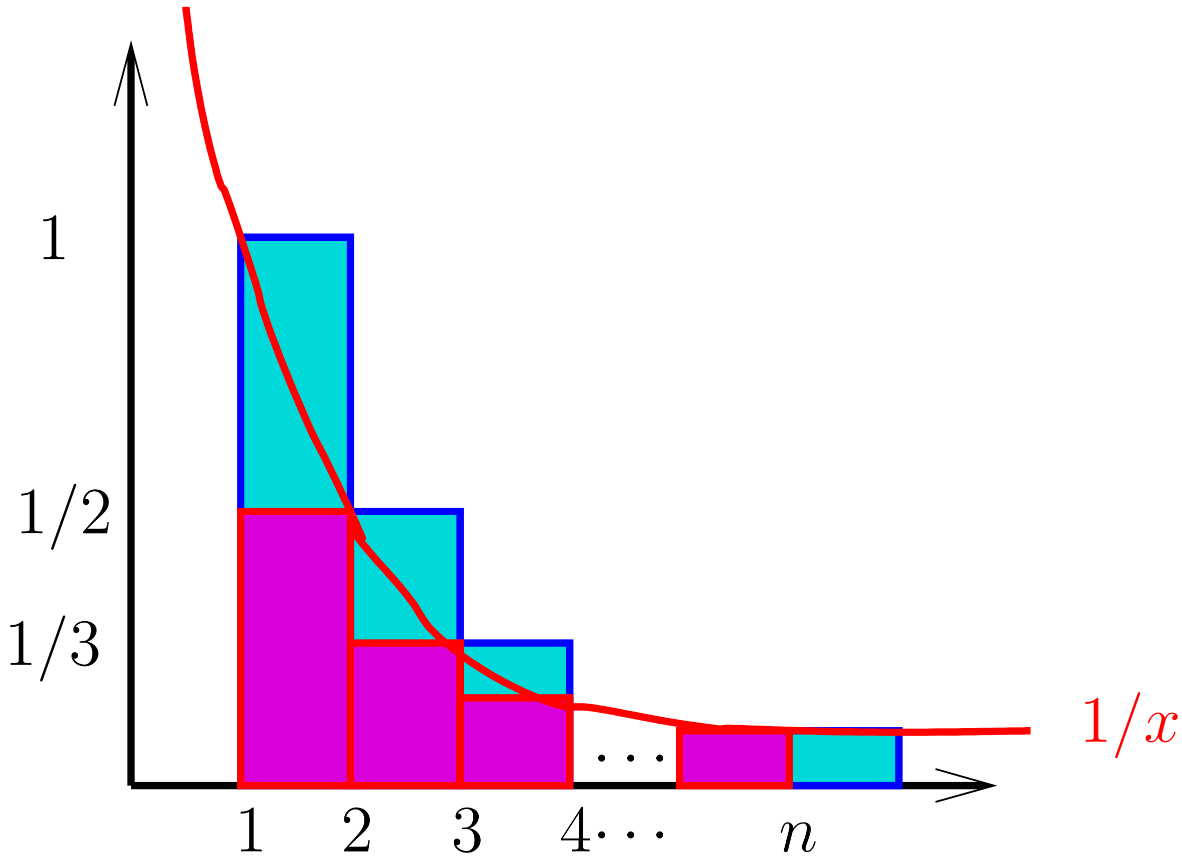
$$
\begin{aligned}
\mathrm{E}[X] &= \mathrm{E}[X_1 + \cdots + X_n] \\
&= \mathrm{E}[X_1] + \cdots + \mathrm{E}[X_n] \\
&= 1/1 + \cdots + 1/n \\
&< 1 + \ln n \\
&= \Theta(\lg n)
\end{aligned}
$$

$$2.92 < \frac{1}{1} + \cdots + \frac{1}{10} < 2.93 < 3.30 < 1 + \ln 10$$

$$5.18 < \frac{1}{1} + \cdots + \frac{1}{100} < 5.19 < 6.60 < 1 + \ln 100$$

$$9.78 < \frac{1}{1} + \cdots + \frac{1}{10000} < 9.79 < 10.21 < 1 + \ln 10000$$

Série harmônica



$$\ln n = \int_1^n \frac{dx}{x} < H_n = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n}$$

$$< 1 + \int_1^n \frac{dx}{x} = 1 + \ln n.$$

Experimentos

Para cada valor de $n = 252, 512, 1024, \ldots$ foram geradas 10, 100 ou 200 amostras de seqüencias de inteiros através do trecho de código

```
for (i = 0; i < n; i++){
  v[i]=(int)((double)INT_MAX*rand()/(RAND_MAX+1
}
```

onde `rand()` é a função geradora de números (pseudo-)aleatórios da biblioteca do `C`.

A coluna $\mathrm{E}[\hat{X}]$ nas tabelas a seguir mostra o número médio de vezes que a linha 4 do algoritmo MAX foi executada para cada valor de n e cada amostra de seqüências.

Experimentos (10)

n	$E[\hat{X}]$	$1 + \ln n$
256	7.20	6.55
512	6.90	7.24
1024	7.30	7.93
2048	7.10	8.62
4096	10.20	9.32
8192	9.00	10.01
16384	10.80	10.70
32768	11.00	11.40
65536	12.50	12.09
131072	12.60	12.78
262144	13.20	13.48
524288	13.20	14.17
1048576	12.80	14.86
2097152	13.90	15.56
4194304	14.90	16.25
8388608	17.90	16.94

Experimentos (100)

n	$E[\hat{X}]$	$1 + \ln n$
256	5.92	6.55
512	6.98	7.24
1024	7.55	7.93
2048	8.39	8.62
4096	8.97	9.32
8192	9.26	10.01
16384	10.44	10.70
32768	11.32	11.40
65536	11.66	12.09
131072	12.38	12.78
262144	13.17	13.48
524288	13.56	14.17
1048576	14.54	14.86
2097152	15.10	15.56
4194304	15.61	16.25
8388608	16.56	16.94

Experimentos (200)

n	$E[\hat{X}]$	$1 + \ln n$
256	6.12	6.55
512	6.86	7.24
1024	7.38	7.93
2048	7.96	8.62
4096	8.87	9.32
8192	9.41	10.01
16384	10.28	10.70
32768	10.92	11.40
65536	11.31	12.09
131072	12.37	12.78
262144	12.92	13.48
524288	13.98	14.17
1048576	14.19	14.86
2097152	15.62	15.56
4194304	15.74	16.25
8388608	17.06	16.94