Aula 7: Caso Médio da Busca Linear

- Distribuição das entradas
- Caso médio da busca linear não ordenada
- Caso médio da busca linear ordenada

Definição de Complexidade do Caso Médio (recordação)

- A: algoritmo
- $E = \{ E_1, \ldots, E_m \}$: conjunto de todas as entradas de A
- $t(E_i)$: número de passos efetuados por A quando a entrada é E_i
- $p(E_i)$: probabilidade de ocorrência da entrada E_i

$$\boxed{\text{C.M.} = \text{Complexidade do caso médio} = \sum_{i=1}^{m} p(E_i)t(E_i)}$$

Caso Médio da Busca Linear Não Ordenada (pelo Método 1)

- <u>Observação Fundamental</u>: entradas distintas que tenham a chave procurada na mesma posição podem ser consideradas como idênticas.

  - Exemplo: Sejam *E* e *E'* as seguintes entradas:

E

6	-1	4	7	3	9

E'

3	-2	8	7	1	10

  - Observe que para a Busca Linear, *E* e *E'* são entradas idênticas (indistingüíveis) quando a chave procurada é x = 7.

Caso Médio da Busca Linear Não Ordenada (pelo Método 1)

Conclusão: A Busca Linear não ordenada numa lista com n elementos só reconhece $n+1$ entradas distintas, a saber:

- entradas em que a chave procurada está na posição 1
- entradas em que a chave procurada está na posição 2

⋮

- entradas em que a chave procurada está na posição n
- entradas em que a chave não se encontra na lista

Sejam:

- E_i, $1 \leq i \leq n$, uma entrada em que a chave procurada ocupa a i-ésima posição da lista
- E_0 entrada que corresponde à busca sem sucesso

O conjunto de entradas é, portanto:

$$\boxed{E = \{ E_0, E_1, \ldots , E_n \}}$$

Busca Linear Não Ordenada: Número de Passos de Cada Entrada

- Observe que $\begin{cases} t(E_k) = k, 1 \leq k \leq n \\ t(E_0) = n \end{cases}$

- Probabilidades das Entradas:
  - Seja q $(0 \leq q \leq 1)$ a probabilidade de sucesso da busca. Supondo que as entradas $E_1, \ldots, E_n$ tenham a mesma probabilidade, temos:

$$\begin{cases} p(E_k) = \dfrac{q}{n}, 1 \leq k \leq n \\ p(E_0) = 1 - q \end{cases}$$

Cálculo Final da Complexidade Média da Busca Linear Não Ordenada

$$\text{C.M.} = \sum_{k=0}^{n} p(E_k)t(E_k) = p(E_0)t(E_0) + p(E_1)t(E_1) + \ldots + p(E_n)t(E_n) =$$

$$= (1-q)\cdot n + \frac{q}{n}\cdot 1 + \frac{q}{n}\cdot 2 + \ldots + \frac{q}{n}\cdot n =$$

$$= (1-q)\cdot n + \frac{q}{n}(1+2+\ldots+n) =$$

$$= n - q\cdot n + \frac{q}{n}\cdot\frac{n(n+1)}{2} = \frac{2n-2qn+qn+q}{2} = \boxed{\frac{2n-qn+q}{2}}$$

Casos Particulares

- q = 1 (a chave sempre se encontra na lista)

$$\text{C.M.} = \frac{n+1}{2} \approx \frac{n}{2}$$

- q = 0 (a chave nunca se encontra na lista)

$$\text{C.M.} = n$$

Caso Médio da Busca Linear Ordenada

- Representemos uma entrada genérica E com n elementos da seguinte forma:

E

y_1	y_2	y_3	...	$y_{n\text{-}1}$	y_n

- Observe que a chave procurada x pode satisfazer $2n + 1$ casos possíveis, a saber :

$$
\begin{array}{l}
x < y_1 \\
x = y_1 \\
y_1 < x < y_2 \\
x = y_2 \\
y_2 < x < y_3 \\
x = y_3 \\
\quad\vdots \\
x = y_{n-1} \\
y_{n-1} < x < y_n \\
x = y_n \\
x > y_n
\end{array}
$$

Busca Linear Ordenada: Descrição do Conjunto de Entradas

- As $2n + 1$ entradas distintas podem ser descritas como:

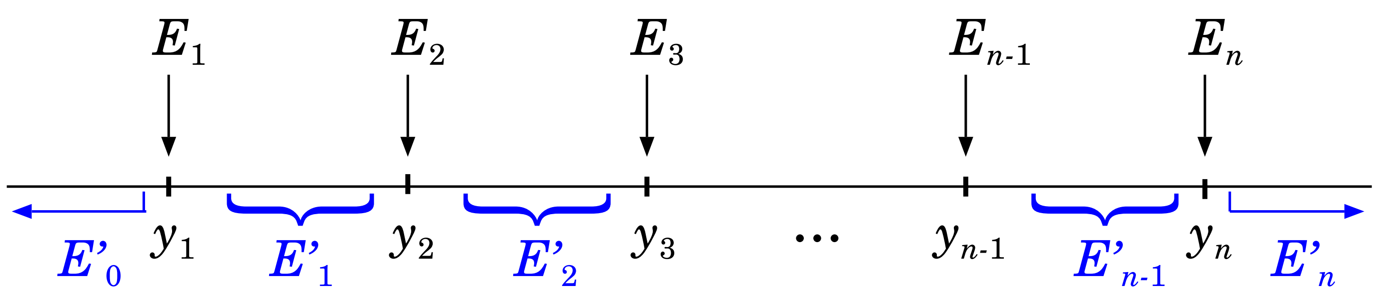
  E_k = entrada em que $x = y_k$, $1 \leq k \leq n$

  E'_0 = entrada em que $x < y_1$

  E'_k = entrada em que $y_k < x < y_{k+1}$, $1 \leq k \leq n\text{-}1$

  E'_n = entrada em que $x > y_n$

- Observe o Diagrama:



O conjunto de entradas é, portanto:

$$\boxed{E = \{ E'_0, E_1, E'_1, E_2, E'_2, E_3, E'_3, \ldots , E_{n\text{-}1}, E'_{n\text{-}1}, E_n, E'_n \}}$$

Busca Linear Ordenada: Número de Passos de Cada Entrada

- Observe que $\begin{cases} t(E_k) = k, \quad 1 \leq k \leq n \\ t(E'_k) = k + 1, \; 0 \leq k \leq n \end{cases}$

- Exemplo: Suponha que deseja-se procurar a chave $x = 7$ numa lista com $n = 6$ elementos. Então:

  - Uma entrada que representa E_3 é

-1	6	7	9	11	12

$t(E_3) = 3$

(A busca faz 3 comparações neste caso)

  - Uma entrada que representa E'_3 é

0	4	6	8	11	15

$t(E'_3) = 4$

(A busca faz 4 comparações neste caso)

Busca Linear Ordenada: Probabilidades das Entradas

- Seja q $(0 \leq q \leq 1)$ a probabilidade de sucesso da busca. Supondo que as entradas $E_1, \ldots, E_n$ sejam equiprováveis, temos:

$$p(E_k) = \frac{q}{n}, (1 \leq k \leq n)$$

- Além disso, supondo que as entradas $E'_0, \ldots, E'_n$ sejam equiprováveis, temos:

$$p(E'_k) = \frac{1-q}{n+1}, (0 \leq k \leq n)$$

Cálculo Final da Complexidade Média da Busca Linear Ordenada

- $\text{C.M.} = \sum_{k=1}^{n} p(E_k)t(E_k) + \sum_{k=0}^{n} p(E'_k)t(E'_k) = \sum_{k=1}^{n} \frac{q}{n} \cdot k + \sum_{k=0}^{n} \frac{1-q}{n+1} \cdot (k+1)$

  - Efetuando os cálculos, vem: $\boxed{\text{C.M.} = \frac{n-q+2}{2}}$

- Observação: O valor da C.M. para a Busca Linear Ordenada é tal que

$$\boxed{\frac{n+1}{2} \leq C.M. \leq \frac{n+2}{2}}$$

ou seja, $\text{C.M.} \approx \frac{n}{2}$ para qualquer valor de q

- Verifique a validade dos cálculos acima para a C.M.

Tempo: 5 minutos

Solução

$$\text{C.M.} = \sum_{k=1}^{n} \frac{q}{n} \cdot k + \sum_{k=0}^{n} \frac{1-q}{n+1} \cdot (k+1) =$$

$$= \frac{q}{n}[1+2+\cdots+n] + \frac{1-q}{n+1}[1+2+\ldots+(n+1)] =$$

$$= \frac{q}{n} \cdot \frac{n(n+1)}{2} + \frac{1-q}{n+1} \cdot \frac{(n+1)(n+2)}{2} =$$

$$= \frac{q(n+1)}{2} + \frac{(1-q)(n+2)}{2} =$$

$$= \frac{qn+q+n+2-qn-2q}{2} =$$

$$= \boxed{\frac{n-q+2}{2}}$$

Exercício

Determinar a expressão da complexidade média de uma busca não ordenada de n chaves (n par), em que as probabilidades de busca das chaves de ordem ímpar são iguais entre si, sendo esse valor igual ao dobro da probabilidade de qualquer chave par. Supor, ainda, que a probabilidade de a chave se encontrar na lista é igual a q .

Tempo: 20 minutos

Solução

De acordo com o enunciado do problema, temos:

$$p(E_1) = p(E_3) = \ldots = p(E_{n-1}) = p \qquad \text{(pois } n \text{ é par)}$$

$$p = 2 \cdot p(E_2) = 2 \cdot p(E_4) = \ldots = 2 \cdot p(E_n)$$

Logo, $p(E_2) = p(E_4) = \ldots = p(E_n) = \dfrac{p}{2}$

<u>Calculando o valor de p</u> :

$$\sum_{i=1}^{n} p(E_i) = q \Rightarrow p(E_1) + p(E_2) + \ldots + p(E_{n-1}) + p(E_n) = q$$

$$\Rightarrow \frac{n}{2} \cdot p + \frac{n}{2} \cdot \frac{p}{2} = q \Rightarrow \boxed{p = \frac{4q}{3n}}$$

Solução

Calculando o valor de C.M. :

$$\text{C.M.} = p(E_0)t(E_0) + p(E_1)t(E_1) + p(E_2)t(E_2) + \ldots + p(E_{n-1})t(E_{n-1}) + p(E_n)t(E_n)$$

$$\text{C.M.} = (1-q)n + \frac{4q}{3n} \cdot 1 + \frac{2q}{3n} \cdot 2 + \ldots + \frac{4q}{3n} \cdot (n-1) + \frac{2q}{3n} \cdot n$$

Complete os cálculos finais e encontre:

$$\text{C.M.} = n + \frac{q}{3} - \frac{qn}{2}$$

Exercício Final

- Determinar a expressão da complexidade média de uma busca não ordenada, supondo que a probabilidade da busca de qualquer chave, exceto a última, é igual a metade da probabilidade da chave seguinte na lista. Supor também que a probabilidade de a chave se encontrar na lista é igual a q .