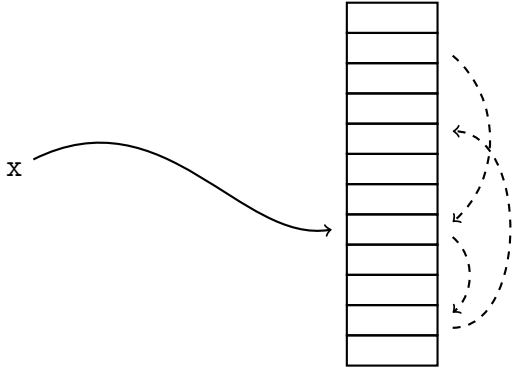


Análise do endereçamento aberto

Na aula passada, nós vimos que a busca sem sucesso na tabela de dispersão com 90% de ocupação executa tempo médio $O(1)$



Porque a busca fica ping-pongando na tabela até bater em uma posição vazia.

E isso acontece após 10 passos, em média — (*i.e.*, $O(1)$).

Legal.

Mas a busca com sucesso acaba antes.

Porque ela encontra o seu elemento antes de bater em uma posição vazia.

Mas qual é o tempo médio da busca com sucesso?

Bom, nós começamos observando que a inserção também precisa encontrar uma posição vazia para colocar o novo elemento.

Daí que, no momento em que a tabela está 90% ocupada, a inserção inspeciona 10 posições da tabela, em média — (*i.e.*, $O(1)$)

E esse também é o tempo da busca por esse elemento.

Porque a busca sempre realiza o mesmo caminho que foi percorrido no momento da inserção.

Mas, elementos que entram na tabela mais cedo encontram a sua posição mais rápido — (*porque a tabela tinha mais posições vazias*)

Logo, o seu tempo de busca vai ser menor.

E isso nos dá a ideia de que o tempo médio da busca com sucesso é menor do que o tempo médio da busca com sucesso.

Sim, mas como é que a gente calcula isso?

Bom, nós vamos calcular isso imaginando o seguinte experimento

- primeiro a gente insere m elementos em uma tabela de tamanho n
- daí, a gente faz a busca por cada um desses elementos
- e soma os tempos de busca de todos os elementos e divide por m , para obter o tempo médio da busca com sucesso

Mas o tempo da busca é igual o tempo da inserção.

Logo, a gente pode somar os tempos de inserção.

Certo.

Imagine que já foram inseridos k elementos na tabela, e nós estamos a ponto de inserir mais um.

Nesse momento, a probabilidade de encontrar uma posição vazia é $\frac{n-k}{n}$.

Logo, nesse momento, a inserção vai inspecionar

$$\frac{1}{(n-k)/n} = \frac{n}{n-k} \quad \text{posições (em média)}$$

Legal.

Agora, basta somar isso para k variando de 0 até $m-1$

$$\begin{aligned} \sum_{k=0}^{m-1} \frac{n}{n-k} &= n \cdot \sum_{k=0}^{m-1} \frac{1}{n-k} \\ &= n \cdot \underbrace{\left(\frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \dots + \frac{1}{n-m+1} \right)}_{(*)} \end{aligned}$$

Daí, para estimar a soma entre parênteses, a gente faz a seguinte esperteza

$$\begin{aligned} (*) &= \left(\frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \dots + \frac{1}{2} + 1 \right) \\ &\quad - \left(\frac{1}{m-n} + \frac{1}{m-n-1} + \dots + \frac{1}{2} + 1 \right) \end{aligned}$$

E daí, para estimar essas somas, a gente faz mais uma esperteza

$$\begin{aligned} &1 \underbrace{+}_{\simeq 1} \frac{1}{2} \underbrace{+ \frac{1}{3}}_{\simeq 1} + \underbrace{\frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}}_{\simeq 1} + \underbrace{\frac{1}{8} + \dots + \frac{1}{15}}_{\simeq 1} + \dots + \frac{1}{n} \\ &= O(\log n) \end{aligned}$$

Quer dizer, a soma dos tempos de inserção é (aproximadamente)

$$n \cdot \left(\log n - \log(n-m) \right)$$

Quando $m = 9n/10$ (i.e., 90% de ocupação), isso dá

$$\begin{aligned} n \cdot \left(\log n - \log(n/10) \right) &= n \cdot \left(\log n - \log n + \log 10 \right) \\ &\simeq 3,32 \cdot n \end{aligned}$$

E dividindo isso por n , nós descobrimos que

→ *A busca com sucesso inspeciona 3,32 posições em média*
— (*na tabela com 90% de ocupação*)

O pior caso

Certo.

Mas isso é apenas o tempo médio.

E isso não significa que todas as buscas vão levar esse tempo.

Quer dizer, a busca por alguns elementos vai ser mais rápida.

E a busca por outros elementos vai levar mais tempo.

Daí, faz sentido perguntar

- *Qual é o pior caso?*

(. . .)