

MATA54 - Estruturas de Dados e Algoritmos II

Hashing Linear

Flávio Assis

Versão gerada a partir de slides do Prof. George Lima

IC - Instituto de Computação

Salvador, agosto de 2021

E se em tabelas hashing de tamanho m fixo $\alpha \rightarrow 1$?

- ▶ Cadeias de sondagem tendem a ficar maiores \rightarrow desempenho cai
- ▶ Necessidade de construir nova tabela hashing com maior valor de m

E se o valor de m for muito maior que n ?

- ▶ Uso desnecessário de espaço
- ▶ Conveniente diminuir o tamanho de m

Abordagem:

Alterar dinamicamente a tabela hashing para se adaptar ao número de chaves: o espaço de espalhamento cresce ou diminui em função do número de chaves

Dois Enfoques de Hashing Dinâmico

- ▶ **Hashing extensível** [Fagin, R; Nievergelt, J.; Pippenger, N; Strong, H. R. "Extendible Hashing-A Fast Access Method for Dynamic Files". ACM TODS, 4(3):315-344, 1979.]
- ▶ **Hashing linear** [Litwin, W. "Linear hashing: A new tool for file and table addressing". 6th Conference on Very Large Databases. pp 212-223, 1980.]

Hashing Linear

Multiplas funções hashing e alocação linear de espaço

$h_\ell(k)$, $\ell = 0, 1, 2, \dots$, mapeam chaves a listas de páginas

$$h_\ell(k) = k \bmod 2^\ell m$$

em que m é um número inicial de páginas

Nível	Número de Páginas
0	m
1	$2m$
2	$4m$
3	$8m$
...	...

Características

- ▶ Não precisa de armazenar índices ou diretórios
- ▶ Resolução de colisões \Rightarrow listas encadeadas de páginas
- ▶ Número de listas é função do fator de carga
- ▶ Tamanho da lista L_i é função das colisões em L_i
- ▶ Ao criar nova lista para nível $\ell + 1$, registros de uma lista no nível ℓ são redistribuídos

Informação necessária

- ▶ Nível ℓ
- ▶ Limites toleráveis para fator de carga: $\alpha^{\min}, \alpha^{\max}$
- ▶ Número inicial de listas m
- ▶ Marcador N para a próxima lista a sofrer redistribuição
 - ▶ Se $h_\ell(k) < N$, usar $h_{\ell+1}(k)$; caso contrário, usar $h_\ell(k)$

Ilustração (Inserção)

Exemplo

Inserir chaves 8, 11, 10, 15, 17, 25, 44 e 12.

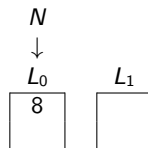
Considerar tamanho de página igual a 2, $\alpha^{\max} = 0.8$ e $m = 2$.

$$\ell = 0$$

$$N = 0$$

$$h_0(8) = 8 \bmod 2^0 \cdot 2 = 0$$

Inserção da chave **8**



$$\alpha = \frac{1}{4} = 0.25 \quad \alpha \leq \alpha^{\max} \quad \text{Ok!}$$

Ilustração (Inserção)

Exemplo

Inserir chaves 8, 11, 10, 15, 17, 25, 44 e 12.

Considerar tamanho de página igual a 2, $\alpha^{\max} = 0.8$ e $m = 2$.

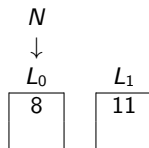
$$\ell = 0$$

$$N = 0$$

$$h_0(8) = 8 \bmod 2^0 \cdot 2 = 0$$

$$h_0(11) = 11 \bmod 2^0 \cdot 2 = 1$$

Inserção da chave **11**



$$\alpha = \frac{2}{4} = 0.50 \quad \alpha \leq \alpha^{\max} \quad \text{Ok!}$$

Ilustração (Inserção)

Exemplo

Inserir chaves 8, 11, 10, 15, 17, 25, 44 e 12.

Considerar tamanho de página igual a 2, $\alpha^{\max} = 0.8$ e $m = 2$.

$$\ell = 0$$

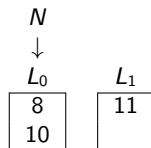
$$N = 0$$

$$h_0(8) = 8 \bmod 2^0 \cdot 2 = 0$$

$$h_0(11) = 11 \bmod 2^0 \cdot 2 = 1$$

$$h_0(10) = 10 \bmod 2^0 \cdot 2 = 0$$

Inserção da chave **10**



$$\alpha = \frac{3}{4} = 0.75 \quad \alpha \leq \alpha^{\max} \quad \text{Ok!}$$

Ilustração (Inserção)

Exemplo

Inserir chaves 8, 11, 10, 15, 17, 25, 44 e 12.

Considerar tamanho de página igual a 2, $\alpha^{\max} = 0.8$ e $m = 2$.

Inserção da chave **15**

$$\ell = 0$$

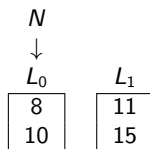
$$N = 0$$

$$h_0(8) = 8 \bmod 2^0 \cdot 2 = 0$$

$$h_0(11) = 11 \bmod 2^0 \cdot 2 = 1$$

$$h_0(10) = 10 \bmod 2^0 \cdot 2 = 0$$

$$h_0(15) = 15 \bmod 2^0 \cdot 2 = 1$$



▶ $\alpha = \frac{4}{4} = 1.00$ $\alpha > \alpha^{\max}$

Quebra de página!

Ilustração (Inserção)

Exemplo

Inserir chaves 8, 11, 10, 15, 17, 25, 44 e 12.

Considerar tamanho de página igual a 2, $\alpha^{\max} = 0.8$ e $m = 2$.

$$\ell = 0$$

$$N = 1$$

$$h_0(8) = 8 \bmod 2^0 \cdot 2 = 0$$

$$h_0(11) = 11 \bmod 2^0 \cdot 2 = 1$$

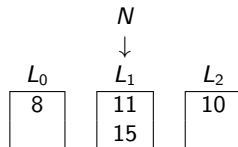
$$h_0(10) = 10 \bmod 2^0 \cdot 2 = 0$$

$$h_0(15) = 15 \bmod 2^0 \cdot 2 = 1$$

$$h_1(8) = 8 \bmod 2^1 \cdot 2 = 0$$

$$h_1(10) = 10 \bmod 2^1 \cdot 2 = 2$$

Inserção da chave **15**



- ▶ Distribuição das chaves das páginas apontadas por N com base em h_{l+1}
- ▶ Avanço de N
- ▶ $\alpha = \frac{4}{6} = 0.67$ **Ok!**

Ilustração (Inserção)

Exemplo

Inserir chaves 8, 11, 10, 15, 17, 25, 44 e 12.

Considerar tamanho de página igual a 2, $\alpha^{\max} = 0.8$ e $m = 2$.

$$\ell = 0$$

$$N = 1$$

$$h_0(8) = 8 \bmod 2^0 \cdot 2 = 0$$

$$h_0(11) = 11 \bmod 2^0 \cdot 2 = 1$$

$$h_0(10) = 10 \bmod 2^0 \cdot 2 = 0$$

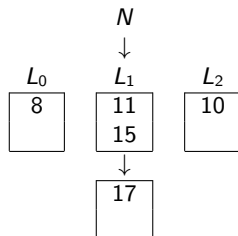
$$h_0(15) = 15 \bmod 2^0 \cdot 2 = 1$$

$$h_0(17) = 17 \bmod 2^0 \cdot 2 = 1$$

$$h_1(8) = 8 \bmod 2^1 \cdot 2 = 0$$

$$h_1(10) = 10 \bmod 2^1 \cdot 2 = 2$$

Inserção da chave **17**



- ▶ $h_0(17) \geq N$: **Ok!**
- ▶ Encadeamento de páginas
- ▶ $\alpha = \frac{5}{8} = 0.63$ $\alpha \leq \alpha^{\max}$: **Ok!**

Ilustração (Inserção)

Exemplo

Inserir chaves 8, 11, 10, 15, 17, 25, 44 e 12.

Considerar tamanho de página igual a 2, $\alpha^{\max} = 0.8$ e $m = 2$.

$$\ell = 0$$

$$N = 1$$

$$h_0(8) = 8 \bmod 2^0 \cdot 2 = 0$$

$$h_0(11) = 11 \bmod 2^0 \cdot 2 = 1$$

$$h_0(10) = 10 \bmod 2^0 \cdot 2 = 0$$

$$h_0(15) = 15 \bmod 2^0 \cdot 2 = 1$$

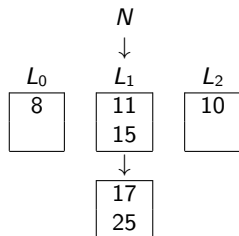
$$h_0(17) = 17 \bmod 2^0 \cdot 2 = 1$$

$$h_0(25) = 25 \bmod 2^0 \cdot 2 = 1$$

$$h_1(8) = 8 \bmod 2^1 \cdot 2 = 0$$

$$h_1(10) = 10 \bmod 2^1 \cdot 2 = 2$$

Inserção da chave **25**



▶ $h_0(25) \geq N$: **Ok!**

▶ $\alpha = \frac{6}{8} = 0.75$ $\alpha \leq \alpha^{\max}$: **Ok!**

Ilustração (Inserção)

Exemplo

Inserir chaves 8, 11, 10, 15, 17, 25, 44 e 12.

Considerar tamanho de página igual a 2, $\alpha^{\max} = 0.8$ e $m = 2$.

$$\ell = 0$$

$$N = 1$$

$$h_0(8) = 8 \bmod 2^0 \cdot 2 = 0$$

$$h_0(11) = 11 \bmod 2^0 \cdot 2 = 1$$

$$h_0(10) = 10 \bmod 2^0 \cdot 2 = 0$$

$$h_0(15) = 15 \bmod 2^0 \cdot 2 = 1$$

$$h_0(17) = 17 \bmod 2^0 \cdot 2 = 1$$

$$h_0(25) = 25 \bmod 2^0 \cdot 2 = 1$$

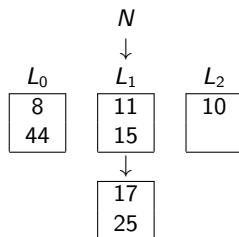
$$h_0(44) = 44 \bmod 2^0 \cdot 2 = 0$$

$$h_1(8) = 8 \bmod 2^1 \cdot 2 = 0$$

$$h_1(10) = 10 \bmod 2^1 \cdot 2 = 2$$

$$h_1(44) = 44 \bmod 2^1 \cdot 2 = 0$$

Inserção da chave **44**



▶ $h_0(44) < N$: calcula-se h_1 (próximo nível)

▶ $\alpha = \frac{7}{8} = 0.88$ $\alpha > \alpha^{\max}$ **Quebra!**

Ilustração (Inserção)

Exemplo

Inserir chaves 8, 11, 10, 15, 17, 25, 44 e 12.

Considerar tamanho de página igual a 2, $\alpha^{\max} = 0.8$ e $m = 2$.

$$\ell = 1$$

$$N = 0$$

$$h_1(8) = 8 \bmod 2^1 \cdot 2 = 0$$

$$h_1(10) = 10 \bmod 2^1 \cdot 2 = 2$$

$$h_1(44) = 44 \bmod 2^1 \cdot 2 = 0$$

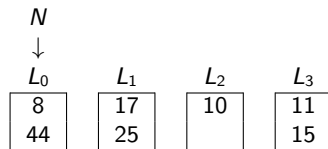
$$h_1(11) = 11 \bmod 2^1 \cdot 2 = 3$$

$$h_1(15) = 15 \bmod 2^1 \cdot 2 = 3$$

$$h_1(17) = 17 \bmod 2^1 \cdot 2 = 1$$

$$h_1(25) = 25 \bmod 2^1 \cdot 2 = 1$$

Inserção da chave **44**



► Incrementa-se o nível

► $\alpha = \frac{7}{8} = 0.88$ $\alpha > \alpha^{\max}$ **Nova quebra!**

Ilustração (Inserção)

Exemplo

Inserir chaves 8, 11, 10, 15, 17, 25, 44 e 12.

Considerar tamanho de página igual a 2, $\alpha^{\max} = 0.8$ e $m = 2$.

$$\ell = 1$$

$$N = 1$$

$$h_1(8) = 8 \bmod 2^1 \cdot 2 = 0$$

$$h_1(10) = 10 \bmod 2^1 \cdot 2 = 2$$

$$h_1(44) = 44 \bmod 2^1 \cdot 2 = 0$$

$$h_1(11) = 11 \bmod 2^1 \cdot 2 = 3$$

$$h_1(15) = 15 \bmod 2^1 \cdot 2 = 3$$

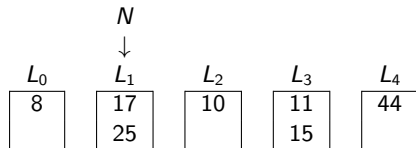
$$h_1(17) = 17 \bmod 2^1 \cdot 2 = 1$$

$$h_1(25) = 25 \bmod 2^1 \cdot 2 = 1$$

$$h_2(8) = 8 \bmod 2^2 \cdot 2 = 0$$

$$h_2(44) = 44 \bmod 2^2 \cdot 2 = 4$$

Inserção da chave **44**



► $\alpha = \frac{7}{10} = 0.70$ $\alpha \leq \alpha^{\max}$ **Ok!**

Ilustração (Inserção)

Exemplo

Inserir chaves 8, 11, 10, 15, 17, 25, 44 e 12.

Considerar tamanho de página igual a 2, $\alpha^{\max} = 0.8$ e $m = 2$.

$$\ell = 1$$

$$N = 1$$

$$h_1(8) = 8 \bmod 2^1 \cdot 2 = 0$$

$$h_1(10) = 10 \bmod 2^1 \cdot 2 = 2$$

$$h_1(44) = 44 \bmod 2^1 \cdot 2 = 0$$

$$h_1(11) = 11 \bmod 2^1 \cdot 2 = 3$$

$$h_1(15) = 15 \bmod 2^1 \cdot 2 = 3$$

$$h_1(17) = 17 \bmod 2^1 \cdot 2 = 1$$

$$h_1(25) = 25 \bmod 2^1 \cdot 2 = 1$$

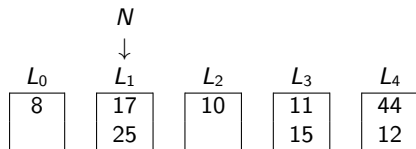
$$h_1(12) = 12 \bmod 2^1 \cdot 2 = 0$$

$$h_2(8) = 8 \bmod 2^2 \cdot 2 = 0$$

$$h_2(44) = 44 \bmod 2^2 \cdot 2 = 4$$

$$h_2(12) = 12 \bmod 2^2 \cdot 2 = 4$$

Inserção da chave **12**



► $h_1(12) < N$: calcula-se $h_2(12)$ (próximo nível)

► $\alpha = \frac{8}{10} = 0.80$ $\alpha \leq \alpha^{\max}$ **Ok!**

Realocação de Registros - Inserção

Quando há quebra de páginas, os registros na página foram distribuídos entre apenas duas páginas: a **página atual** e a **nova página criada**.

Realocação de Registros - Inserção

Quando há quebra de páginas, os registros na página foram distribuídos entre apenas duas páginas: a **página atual** e a **nova página criada**.

Isto é sempre verdade? Por quê?

Realocação de Registros - Inserção

Quando há quebra de páginas, os registros na página foram distribuídos entre apenas duas páginas: a **página atual** e a **nova página criada**.

Isto é sempre verdade? Por quê?

Quando há a quebra de página de índice i , para quais páginas os registros que estão nesta página podem ser realocados?

Procedimento de Inserção

Inserir registro com chave k

Considere que k não existe em nenhuma das listas

1. $i \leftarrow h_\ell(k)$
2. Se $i < N$, então registros em L_i foram redistribuídos. Neste caso, considerar $i \leftarrow h_{\ell+1}(k)$
3. Inserir k na lista L_i
4. Se $\alpha > \alpha^{max}$, então limite do fator de carga foi violado
 - 4.1 Criar lista $L_{N+2^\ell m}$
 - 4.2 Redistribuir todos os registros com chaves k^* em L_N entre L_N e $L_{N+2^\ell m}$. Se $h_{\ell+1}(k^*) \neq N$, k^* pertence à lista $L_{N+2^\ell m}$. Caso contrário, k^* permanece a L_N
 - 4.3 $N \leftarrow N + 1$
 - 4.4 Se $N \geq 2^\ell m$, então todas as listas pertencem ao nível $\ell + 1$
 - 4.4.1 $N \leftarrow 0$
 - 4.4.2 $\ell \leftarrow \ell + 1$

Buscar registro com chave k

1. $i \leftarrow h_\ell(k)$
2. Se $i < N$, então registros em L_i foram redistribuídos. Neste caso, considerar $i \leftarrow h_{\ell+1}(k)$
3. Pesquisar k na lista L_i

Ilustração (Remoção)

Exemplo

Considere o arquivo após a inserção das chaves 8, 11, 10, 15 e 22, com tamanho de página igual a 2, $\alpha^{\max} = 0.85$, $\alpha^{\min} = 0.40$ e $m = 2$.

$$\ell = 0$$

$$N = 1$$

$$h_0(8) = 8 \bmod 2^0 \cdot 2 = 0$$

$$h_0(11) = 11 \bmod 2^0 \cdot 2 = 1$$

$$h_0(10) = 10 \bmod 2^0 \cdot 2 = 0$$

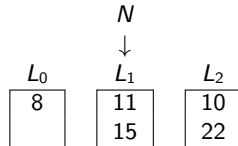
$$h_0(15) = 15 \bmod 2^0 \cdot 2 = 1$$

$$h_0(22) = 22 \bmod 2^0 \cdot 2 = 0$$

$$h_1(8) = 8 \bmod 2^1 \cdot 2 = 0$$

$$h_1(10) = 10 \bmod 2^1 \cdot 2 = 2$$

$$h_1(22) = 22 \bmod 2^1 \cdot 2 = 2$$



► $\alpha = \frac{5}{6} = 0.83$

Ilustração (Remoção)

Exemplo

Considere o arquivo após a inserção das chaves 8, 11, 10, 15 e 22, com tamanho de página igual a 2, $\alpha^{\max} = 0.85$, $\alpha^{\min} = 0.40$ e $m = 2$.

$$\ell = 0$$

$$N = 1$$

$$h_0(8) = 8 \bmod 2^0 \cdot 2 = 0$$

$$h_0(11) = 11 \bmod 2^0 \cdot 2 = 1$$

$$h_0(10) = 10 \bmod 2^0 \cdot 2 = 0$$

$$h_0(15) = 15 \bmod 2^0 \cdot 2 = 1$$

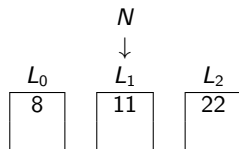
$$h_0(22) = 22 \bmod 2^0 \cdot 2 = 0$$

$$h_1(8) = 8 \bmod 2^1 \cdot 2 = 0$$

$$h_1(10) = 10 \bmod 2^1 \cdot 2 = 2$$

$$h_1(22) = 22 \bmod 2^1 \cdot 2 = 2$$

Remoção das chaves **10** e **15**:



► $\alpha = \frac{3}{6} = 0.50$ $\alpha \geq \alpha^{\min}$ **Ok!**

Ilustração (Remoção)

Exemplo

Considere o arquivo após a inserção das chaves 8, 11, 10, 15 e 22, com tamanho de página igual a 2, $\alpha^{\max} = 0.85$, $\alpha^{\min} = 0.40$ e $m = 2$.

$$\ell = 0$$

$$N = 1$$

$$h_0(8) = 8 \bmod 2^0 \cdot 2 = 0$$

$$h_0(11) = 11 \bmod 2^0 \cdot 2 = 1$$

$$h_0(10) = 10 \bmod 2^0 \cdot 2 = 0$$

$$h_0(15) = 15 \bmod 2^0 \cdot 2 = 1$$

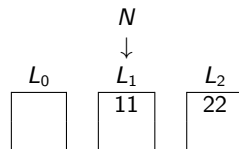
$$h_0(22) = 22 \bmod 2^0 \cdot 2 = 0$$

$$h_1(8) = 8 \bmod 2^1 \cdot 2 = 0$$

$$h_1(10) = 10 \bmod 2^1 \cdot 2 = 2$$

$$h_1(22) = 22 \bmod 2^1 \cdot 2 = 2$$

Remoção da chave **8**:



$$\blacktriangleright \alpha = \frac{2}{6} = 0.33 \quad \alpha < \alpha^{\min}$$

Eliminar página!

Ilustração (Remoção)

Exemplo

Considere o arquivo após a inserção das chaves 8, 11, 10, 15 e 22, com tamanho de página igual a 2, $\alpha^{\max} = 0.85$, $\alpha^{\min} = 0.40$ e $m = 2$.

$$\ell = 0$$

$$N = 0$$

$$h_0(8) = 8 \bmod 2^0 \cdot 2 = 0$$

$$h_0(11) = 11 \bmod 2^0 \cdot 2 = 1$$

$$h_0(10) = 10 \bmod 2^0 \cdot 2 = 0$$

$$h_0(15) = 15 \bmod 2^0 \cdot 2 = 1$$

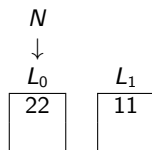
$$h_0(22) = 22 \bmod 2^0 \cdot 2 = 0$$

$$h_1(8) = 8 \bmod 2^1 \cdot 2 = 0$$

$$h_1(10) = 10 \bmod 2^1 \cdot 2 = 2$$

$$h_1(22) = 22 \bmod 2^1 \cdot 2 = 2$$

Remoção da chave **8**:



- ▶ Realocam-se as chaves que estão em L_2 (página de maior índice)
- ▶ Atualiza-se N
- ▶ $\alpha = \frac{2}{4} = 0.50 \geq \alpha^{\min}$ **Ok!**

Realocação de Registros - Remoção

Quando há eliminação de um página, os registros foram realocados **em uma mesma página**.

Realocação de Registros - Remoção

Quando há eliminação de um página, os registros foram realocados **em uma mesma página**.

Isto é sempre verdade? Por quê?

Realocação de Registros - Remoção

Quando há eliminação de um página, os registros foram realocados **em uma mesma página**.

Isto é sempre verdade? Por quê?

Quando há a eliminação de uma página de índice i , para qual página os registros que estão nesta página serão realocados?

Procedimento de Remoção

Remover registro com chave k

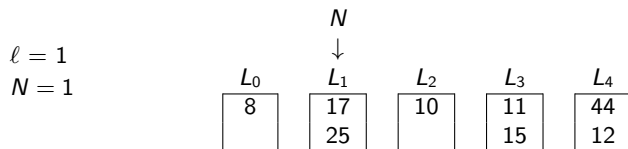
Considere que k existe em alguma das listas

1. $i \leftarrow h_\ell(k)$
2. Se $i < N$, então registros em L_i foram redistribuídos. Neste caso, considerar $i \leftarrow h_{\ell+1}(k)$
3. Remover k da lista L_i
4. Se $\alpha < \alpha^{min}$, então limite do fator de carga foi violado
 - 4.1 $N \leftarrow N - 1$
 - 4.2 Se $N < 0$, então todas as listas pertencem ao nível ℓ . É necessário retornar ao nível anterior
 - 4.2.1 $\ell \leftarrow \ell - 1$
 - 4.2.2 $N \leftarrow 2^\ell - 1$
 - 4.3 Transferir registros da lista $L_{N+2^{\ell_m}}$ para a lista L_N usando h_ℓ
 - 4.4 Remover lista $L_{N+2^{\ell_m}}$

Exercício

A partir do estado abaixo, mostre o estado final do arquivo após a inserção das chaves 22, 37, 30 e 21.

Tamanho de página igual a 2, $\alpha^{\max} = 0.80$ e $m = 2$.



Exercício

Resposta:

	N							
	\downarrow							
	L_0	L_1	L_2	L_3	L_4	L_5	L_6	L_7
$\ell = 2$	$\boxed{8}$	$\boxed{17}$	$\boxed{10}$	$\boxed{11}$	$\boxed{44}$	$\boxed{37}$	$\boxed{22}$	$\boxed{15}$
$N = 0$	$\boxed{}$	$\boxed{25}$	$\boxed{}$	$\boxed{}$	$\boxed{12}$	$\boxed{21}$	$\boxed{30}$	$\boxed{}$

$$\alpha = \frac{12}{16} = 0.75$$