



Hashing

Sandy Porto

Introdução

Tratamento de
Colisões

Rehashing

Parte II: Hashing

Estrutura de Dados II

Sandy Porto
sandyporto@gmail.com

Universidade Federal de Sergipe

10 de Janeiro de 2017



Conteúdo

Hashing

Sandy Porto

Introdução

Tratamento de
Colisões

Rehashing

1 Introdução

2 Tratamento de Colisões

3 Rehashing



Introdução

Relembrando

Hashing

Sandy Porto

Introdução

Tratamento de
Colisões

Rehashing

- Na *organização direta* há uma *relação* entre o valor da chave C e o endereço E do registro.
- **Hashing** é um dos métodos de organização e acesso a registros em arquivos.

$$f(\text{chave}) = \text{endereço}$$

- A chamada *função hash* gera um provável endereço do registro a partir do valor da chave.
- O provável endereço é chamado *endereço base*.
- A função Hash mais utilizada: $f(ch) = ch \bmod M$, onde M deve ser um número primo.



Introdução

Exemplo

Hashing

Sandy Porto

Introdução

Tratamento de
Colisões

Rehashing

$$f(ch) = ch \bmod 13$$

| Chave | Atributo 1 | Atributo 2 | Atributo 3 | Atributo 4 |
|-------|------------|------------|------------|------------|
| 852 | .. | .. | .. | .. |
| 456 | .. | .. | .. | .. |
| 789 | .. | .. | .. | .. |
| 145 | .. | .. | .. | .. |
| 236 | .. | .. | .. | .. |
| 053 | .. | .. | .. | .. |
| 548 | .. | .. | .. | .. |
| 365 | .. | .. | .. | .. |
| 447 | .. | .. | .. | .. |
| 951 | .. | .. | .. | .. |



Introdução

Exemplo

Hashing

Sandy Porto

Introdução

Tratamento de
Colisões

Rehashing

$$f(ch) = ch \bmod 13$$

| Chave | Endereço |
|-------|----------|
| 852 | 07 |
| 456 | 01 |
| 789 | 09 |
| 145 | 02 |
| 236 | 02 |
| 053 | 01 |
| 548 | 02 |
| 365 | 01 |
| 447 | 05 |
| 951 | 02 |

O endereço de algumas chaves é o mesmo, e agora?



Introdução

Colisão

Hashing

Sandy Porto

Introdução

Tratamento de
Colisões

Rehashing

- **Colisão** é o nome dado quando duas ou mais chaves distintas geram o mesmo valor de endereço base.
 - *Hashing perfeito*: uma função que não gere colisões. É a melhor solução?
- Tratamento de Colisões:
 - As funções hash mais utilizadas na literatura não são perfeitas.
 - Então, depois de aplicada a função hash é necessário fazer o tratamento de colisões.
- Métodos utilizados:
 - Listas Encadeadas
 - Cada endereço base da tabela hash é a cabeça de uma lista encadeada
 - Endereçamento Aberto
 - A chave colidida será colocada em algum espaço em branco na tabela hash de acordo com uma função de deslocamento.



Conteúdo

Hashing

Sandy Porto

Introdução

Tratamento de
Colisões

Rehashing

1 Introdução

2 Tratamento de Colisões

3 Rehashing



Tratamento de Colisões

Listas Encadeadas

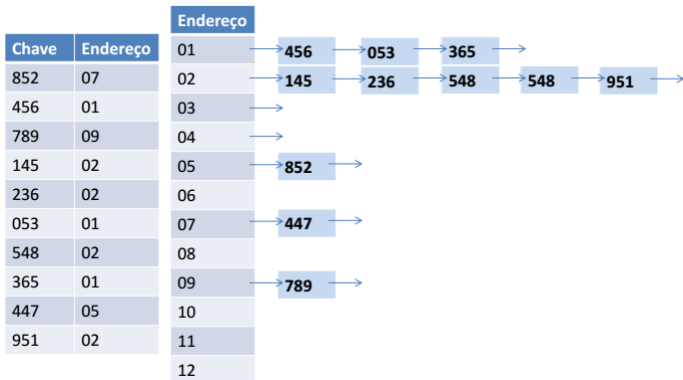
Hashing

Sandy Porto

Introdução

Tratamento de Colisões

Rehashing





Tratamento de Colisões

Endereçamento Aberto

Hashing

Sandy Porto

Introdução

Tratamento de
Colisões

Rehashing

$$f(ch) = ch \bmod 13, g(end) = end + 1$$

| Chave | Endereço |
|-------|----------|
| 852 | 07 |
| 456 | 01 |
| 789 | 09 |
| 145 | 02 |
| 236 | 02 |
| 053 | 01 |
| 548 | 02 |
| 365 | 01 |
| 447 | 05 |
| 951 | 02 |

| Endereço | Registro |
|----------|----------|
| 01 | |
| 02 | |
| 03 | |
| 04 | |
| 05 | |
| 06 | |
| 07 | |
| 08 | |
| 09 | |
| 10 | |
| 11 | |
| 12 | |



Tratamento de Colisões

Endereçamento Aberto

Hashing

Sandy Porto

$$f(ch) = ch \bmod 13, g(end) = end + 1$$

Introdução

Tratamento de
Colisões

Rehashing

| Chave | Endereço |
|-------|----------|
| 852 | 07 |
| 456 | 01 |
| 789 | 09 |
| 145 | 02 |
| 236 | 02 |
| 053 | 01 |
| 548 | 02 |
| 365 | 01 |
| 447 | 05 |
| 951 | 02 |

| Endereço | Registro |
|----------|----------|
| 01 | 456 |
| 02 | 145 |
| 03 | 236 |
| 04 | 053 |
| 05 | 548 |
| 06 | 365 |
| 07 | 852 |
| 08 | 447 |
| 09 | 789 |
| 10 | 951 |
| 11 | |
| 12 | |



Tratamento de Colisões

Endereçamento Aberto

Hashing

Sandy Porto

$$f(ch) = ch \bmod 13, \quad g(ch) = (ch \bmod 3) + 1$$

Introdução

Tratamento de
Colisões

Rehashing

| Chave | Endereço |
|-------|----------|
| 852 | 07 |
| 456 | 01 |
| 789 | 09 |
| 145 | 02 |
| 236 | 02 |
| 053 | 01 |
| 548 | 02 |
| 365 | 01 |
| 447 | 05 |
| 951 | 02 |

| Endereço | Registro |
|----------|----------|
| 01 | |
| 02 | |
| 03 | |
| 04 | |
| 05 | |
| 06 | |
| 07 | |
| 08 | |
| 09 | |
| 10 | |
| 11 | |
| 12 | |



Tratamento de Colisões

Endereçamento Aberto

Hashing

Sandy Porto

$$f(ch) = ch \bmod 13, g(ch) = (ch \bmod 3) + 1$$

Introdução

Tratamento de Colisões

Rehashing

| Chave | Endereço |
|-------|----------|
| 852 | 07 |
| 456 | 01 |
| 789 | 09 |
| 145 | 02 |
| 236 | 02 |
| 053 | 01 |
| 548 | 02 |
| 365 | 01 |
| 447 | 05 |
| 951 | 02 |

| Endereço | Registro |
|----------|----------|
| 01 | 456 |
| 02 | 145 |
| 03 | 053 |
| 04 | 365 |
| 05 | 236 |
| 06 | 447 |
| 07 | 852 |
| 08 | 548 |
| 09 | 789 |
| 10 | 951 |
| 11 | |
| 12 | |



Conteúdo

Hashing

Sandy Porto

Introdução

Tratamento de
Colisões

Rehashing

1 Introdução

2 Tratamento de Colisões

3 Rehashing



Rehashing

Hashing

Sandy Porto

Introdução

Tratamento de
Colisões

Rehashing

- Refazer a estrutura hash.
- Seja M o tamanho atual da estrutura e N o tamanho após o rehashing:
 - Se $N > M$, o objetivo é diminuir a frequência de colisões.
 - Se $N = M$, o objetivo é reorganizar a estrutura hash.
 - Se $N < M$, o objetivo é poupar espaço.