

# Organização e Recuperação da Informação

## Tabelas hash

Jander Moreira  
jander@dc.ufscar.br

Universidade Federal de São Carlos  
Departamento de Computação

2014

- ① Visão geral
- ② Funções hash
- ③ Atividades

# Visão geral

### Tabelas hash:

- Base no cálculo da posição da tabela em que o registro deve ficar
- Evitação de comparações sucessivas entre os valores das chaves (como na pesquisa sequencial e binária)

## Operação da tabela:

- Cálculo do endereço a partir do valor da chave

$$e = h(c)$$

- Função de cálculo de endereço: **função hash**
- Validade do endereço  $e$ : dentro dos limites da tabela
- Função  $h$ :
  - Hash
  - Espalhamento
  - Aleatorização
  - ...

Operação da tabela:

- Situação ideal:  $h(c_1) \neq h(c_2)$  para  $c_1 \neq c_2$   
(função hash perfeita)
- Situação real: **colisões**

# Funções hash

Funções hash perfeitas:

- $h(c_1) \neq h(c_2), c_1 \neq c_2$
- Difícil representação por uma fórmula concisa

Outras funções para hash:

- Aceitação de colisões



### Opções para formulação de funções hash:

- Divisão modular:  
módulo da divisão pelo número de posições da tabela
- Dobramento:  
quebra da chave em partes e combinação dessas partes
- Meio do quadrado: dígitos do meio do valor da chave ao quadrado
- Extração: uso de parte da chave
- Mudança de base: representação da chave em outra base numérica

Escolha a opção que você julga mais importante quando se projeta uma função hash.

- (A) Redução da velocidade de cálculo do endereço
- (B) Vínculo adequado da função como o tratamento de colisões escolhido
- (C) Redução do número de colisões
- (D) Ignorar as características das chaves da coleção que será usada, focando em um tratamento de colisões

Algumas considerações sobre o projeto de funções hash:

- Conhecimento das características das chaves que serão usadas
- Manutenção do tempo de cálculo da função sem impacto significativo
- Redução, tanto quanto possível, do número de colisões
- Ponderação sobre o tamanho máximo da tabela e sua relação com a função hash

# Atividades

Dinâmica da atividade:

- (i) Formar grupos de trabalho
- (ii) Nomear um representante do grupo
- (iii) Realizar a tarefa solicitada
- (iv) Apresentar os resultados à classe, em caso do grupo ser escolhido para apresentação
- (v) Discussão geral com a classe sobre as exposições feitas

## Atividade

Projete uma função hash para RAs de alunos, considerando a pesquisa por qualquer RA de 1 até 70.000. Como não se dispõe da fórmula de cálculo do dígito de controle, deve-se considerar apenas a parte sequencial do número de registro.

Considere que a aplicação que usará a tabela espera que cerca de 2000 alunos sejam inseridos na tabela, de forma que seu tamanho deve ser apropriado.

Levante as seguintes informações (estimativas):

- Quantas colisões ocorrem para alguns experimentos com dados aleatórios
- Avaliar se a modificação do tamanho da tabela altera significativamente o número de colisões
- Estimar se, no tempo proporcionado, a função elaborada pelo grupo foi satisfatória

## Atividade

Para os grupos escolhidos:

- A função projetada
- Os resultados obtidos
- As considerações do grupo sobre os resultados obtidos