

Estruturas de Dados II

- Aula 8
 - Tabelas de dispersão (hash tables)

Prof. DSc Carlos de Salles

www.deinf.ufma.br/~csalles

csalles@deinf.ufma.br



Tabelas de Dispersão

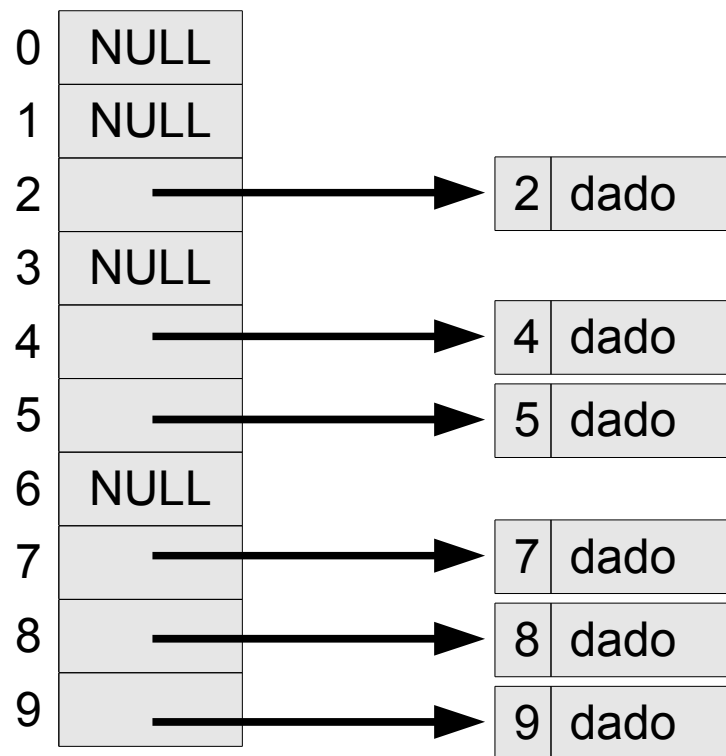
- É como um vetor cujas chaves **não** são apenas inteiros e os valores **não** estão necessariamente consecutivos na memória
- Suporta as primitivas/operações de dicionário de dados:
 - Insert
 - Search
 - Delete

Definições Prévias

- Assuma que numa tabela de dispersão, você tem:
 - Um universo U de $|U|$ chaves, $U=\{k_1, k_2, k_3, \dots, k_u\}$
 - Um conjunto K de $|K|$ chaves sendo utilizadas, $K=\{k_{i1}, k_{i2}, k_{i3}, \dots, k_{ik}\}$, onde i pode variar de 1 a $|U|$
 - Para cada chave do universo U , pode haver um único respectivo dado satélite (há um valor para cada chave)
 - Importante: há tabelas com multivalores, em que uma chave do universo U pode estar relacionada a mais de um valor. Exemplo: multiset da STL de C++

Tabelas de Acesso Direto

- Há uma **tabela de acesso direto** com $|U|$ entradas, uma para cada chave U do universo



Tabelas de Acesso Direto

- Primitivas

- Insert

```
function insert(T, x) T[ key[x] ] = x end
```

- Search

```
function search(T, k) return T[ k ] end
```

- Delete

```
function delete(T, x) T[ key[x] ]=nil end
```

Tabelas de Acesso Direto

- Problemas:
 - Se o universo U for grande demais, uma tabela de tamanho $|U|$ pode ser impraticável ou até impossível
 - Suponha as chaves sendo os nomes de pessoas
 - Suponha as chaves sendo os códigos de alunos
 - E se $|K|$ for pequeno e $|U|$ for grande? A tabela de acesso direto vai ter muitas entradas com valor NULL, ocupando memória desnecessariamente

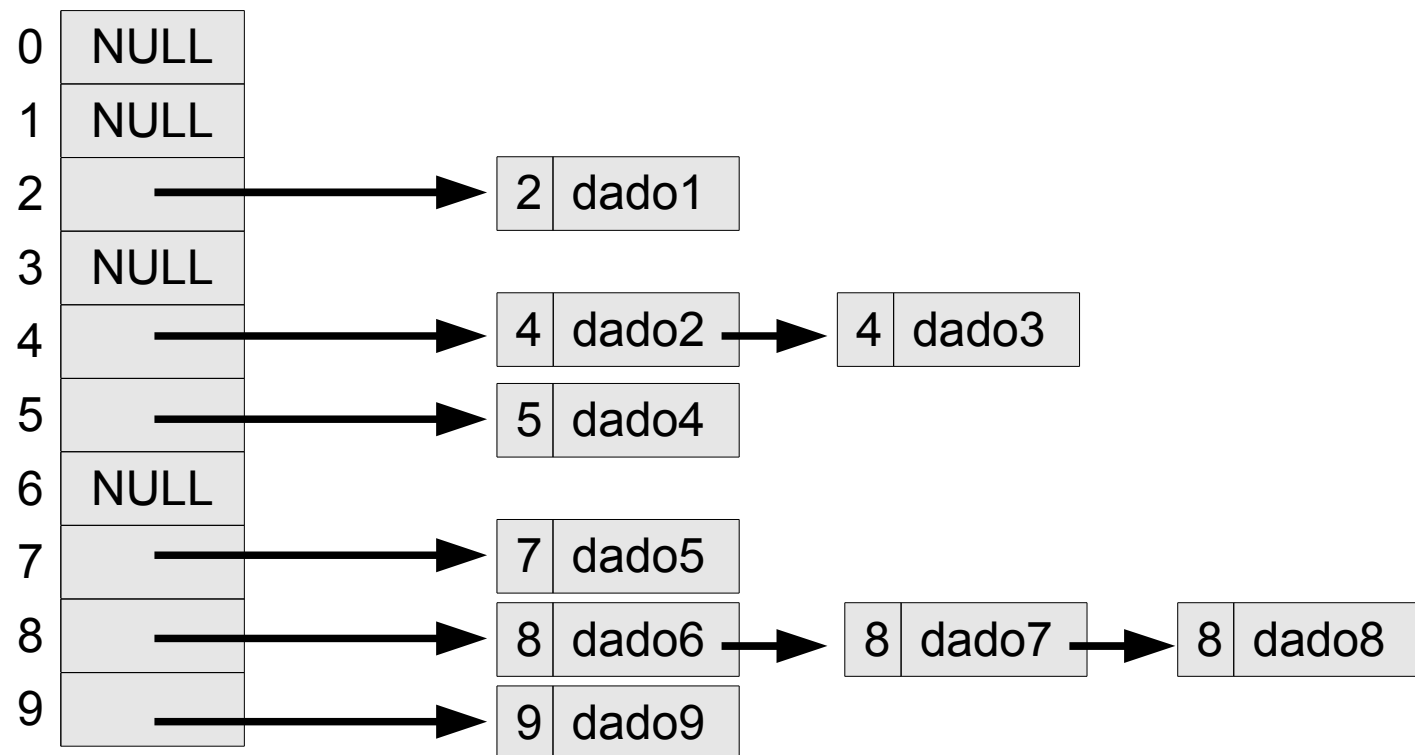
Técnica de Dispersão

- Uma **função de dispersão h** mapeia as chaves k , que variam de 0 a $|U|-1$ do universo U , em chaves 0 a $|T|-1$ da tabela T
- Colisões podem ocorrer!
 - Duas chaves k_i e k_j diferentes podem mapear em um mesmo valor de dispersão m . Isso caracteriza uma colisão

Resolvendo Colisões por Encadeamento

- Cada entrada da tabela T aponta para uma lista encadeada de valores
- Quando dois valores colidem, eles são inseridos na mesma lista encadeada

Tabelas de Dispersão com Listas Encadeadas



Tabelas de Dispersão com Listas Encadeadas

- Primitivas

- Insert

```
function insert(T, x)
    insira x no início da lista T[ h[key[x]] ]
end
```

- Search

```
function search(T, k)
    procure x na lista T[ h[key[x]] ]
end
```

- Delete

```
function delete(T, x)
    delete x da lista T[ h[key[x]] ]
end
```