

Armazenamento e Indexação

Banco de Dados: Teoria e Prática

André Santanchè e Patrícia Cavoto
Instituto de Computação - UNICAMP
Setembro de 2015

Recomendações de Leitura

- (Silberschatz, 2006, cap. 11)
- (Ramakrishnan, 2003, cap. 8)
- (Elmasri, 2011, cap. 11 e 12)

Onde Armazenamos Dados?

Onde Armazenamos Dados?

- Memória RAM
- Disco
 - HD
 - CD / DVD
- Fita magnética
- Solid State Drive (SDD)
 - usa circuitos integrados como a memória - sem partes mecânicas
 - retém os dados sem a necessidade de energia
 - Interface equivalente a de um disco

Exercício 1

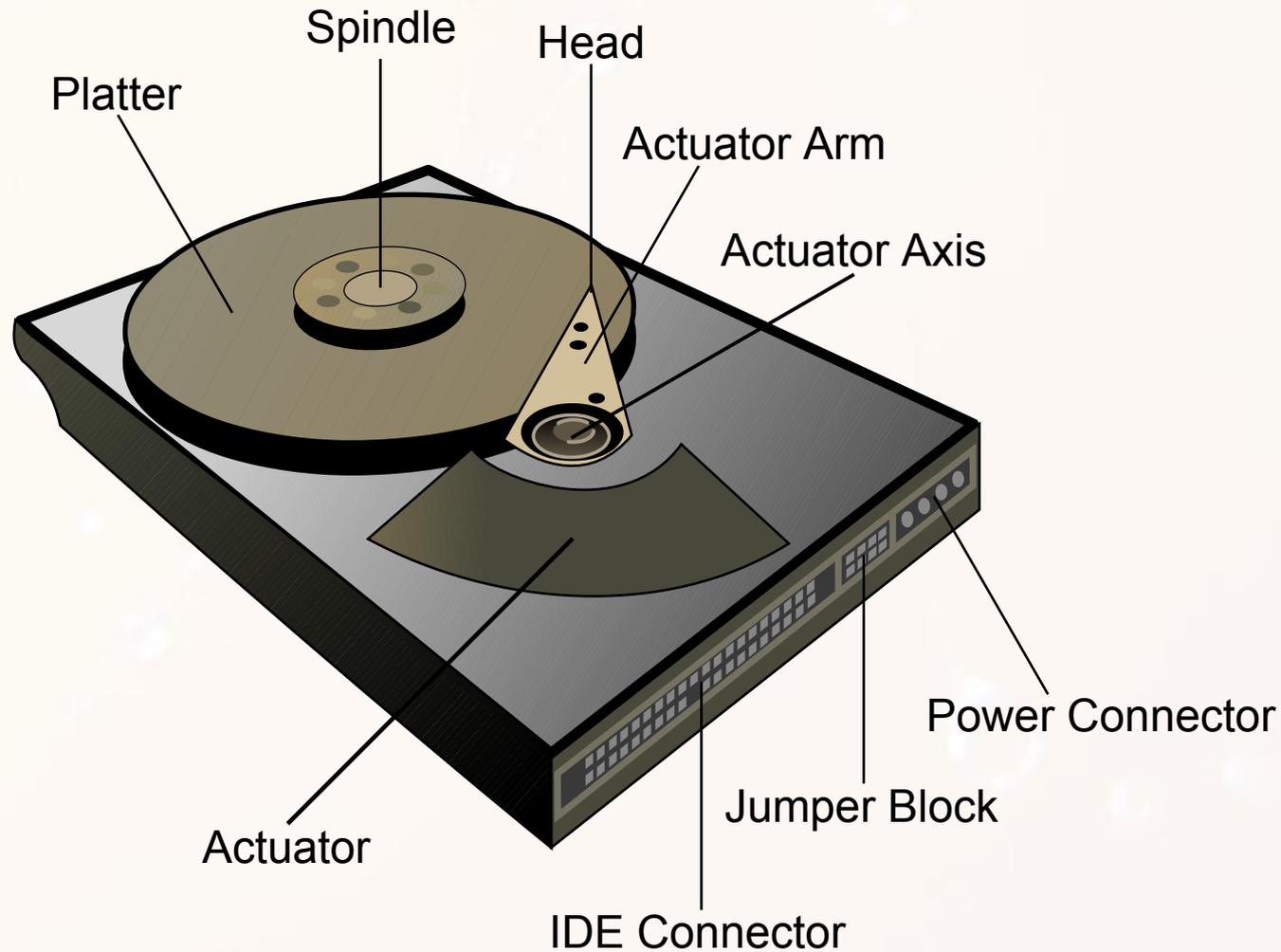
- Para cada item abaixo, liste suas vantagens e desvantagens como opção de tecnologia para armazenamento de dados num SGBD. Dê exemplos de dados que se adequariam à tecnologia.
 - a) Memória RAM
 - b) Disco Magnético
 - c) Fita Magnética

Hierarquia de Armazenamento

- Armazenamento Primário
 - Operado diretamente pela CPU
 - Exemplos: memória RAM, cache
- Armazenamento Secundário
 - Usualmente mais barato e mais lento
 - Não operado diretamente pela CPU
 - Exigem intermediação de armazenamento primário
 - Exemplos: disco, fita magnética

(Elmasri, 2011)

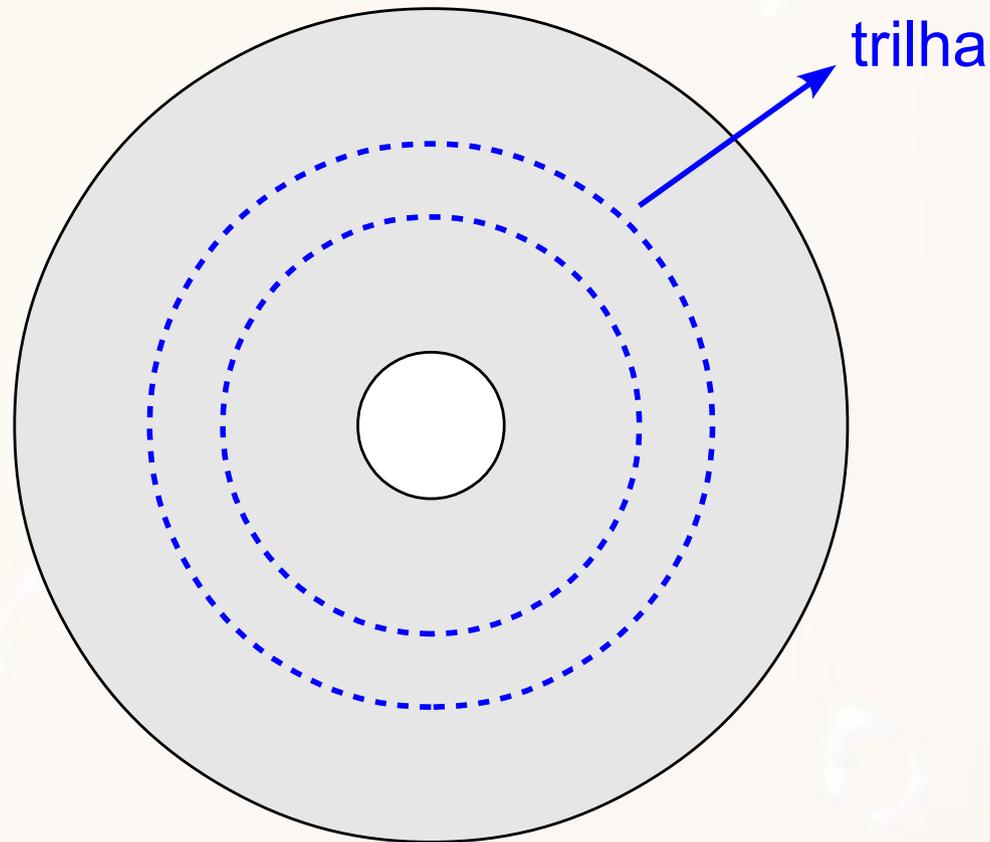
Estrutura do Disco



By Surachit [http://en.wikipedia.org/w/index.php?title=File:Hard_drive-en.svg]

Estrutura do Disco

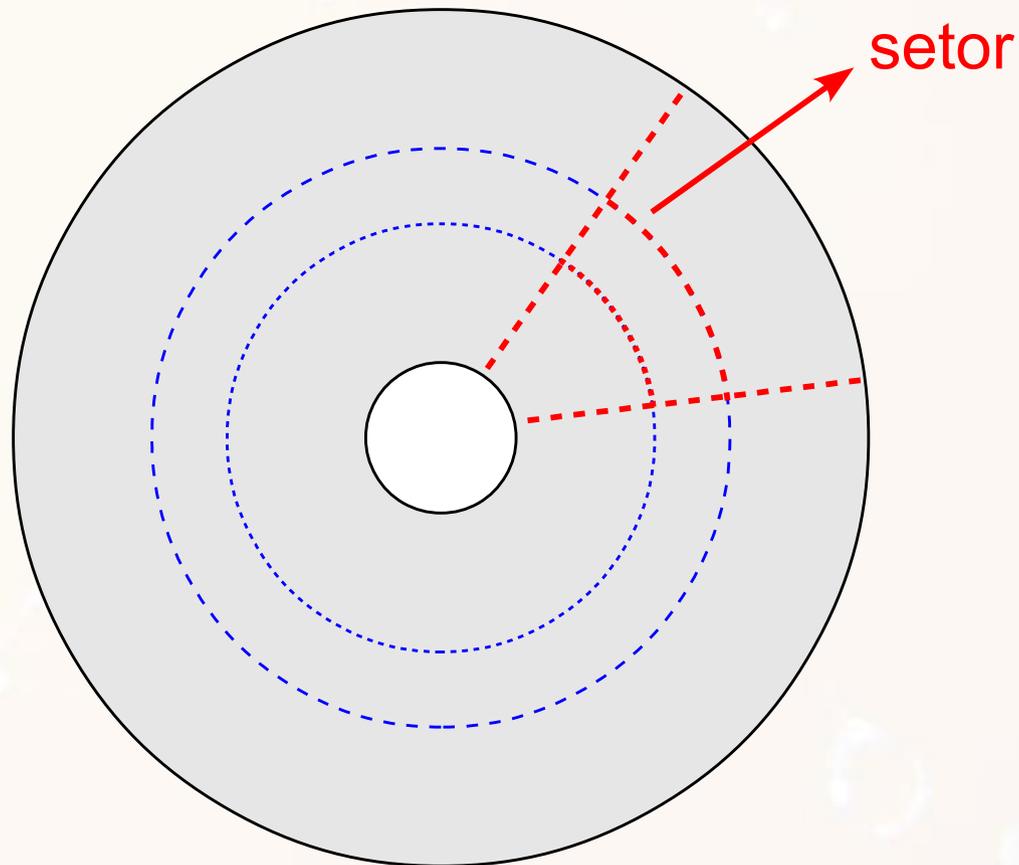
Trilha



- Círculos magnéticos sobre a superfície
- Local onde são armazenados os dados

Estrutura do Disco

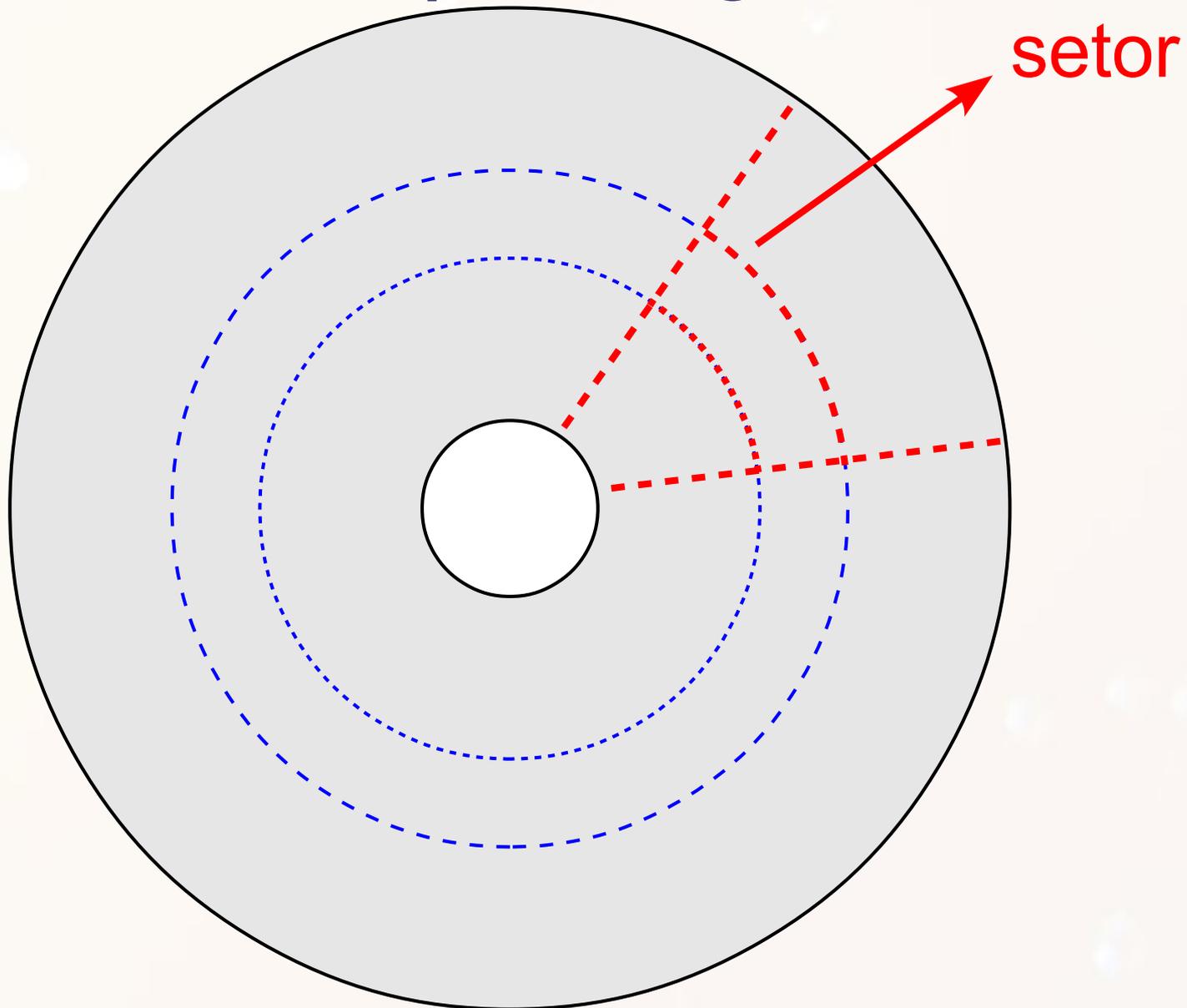
Setor



- Unidades de divisão da trilha
- Menor unidade de leitura/gravação

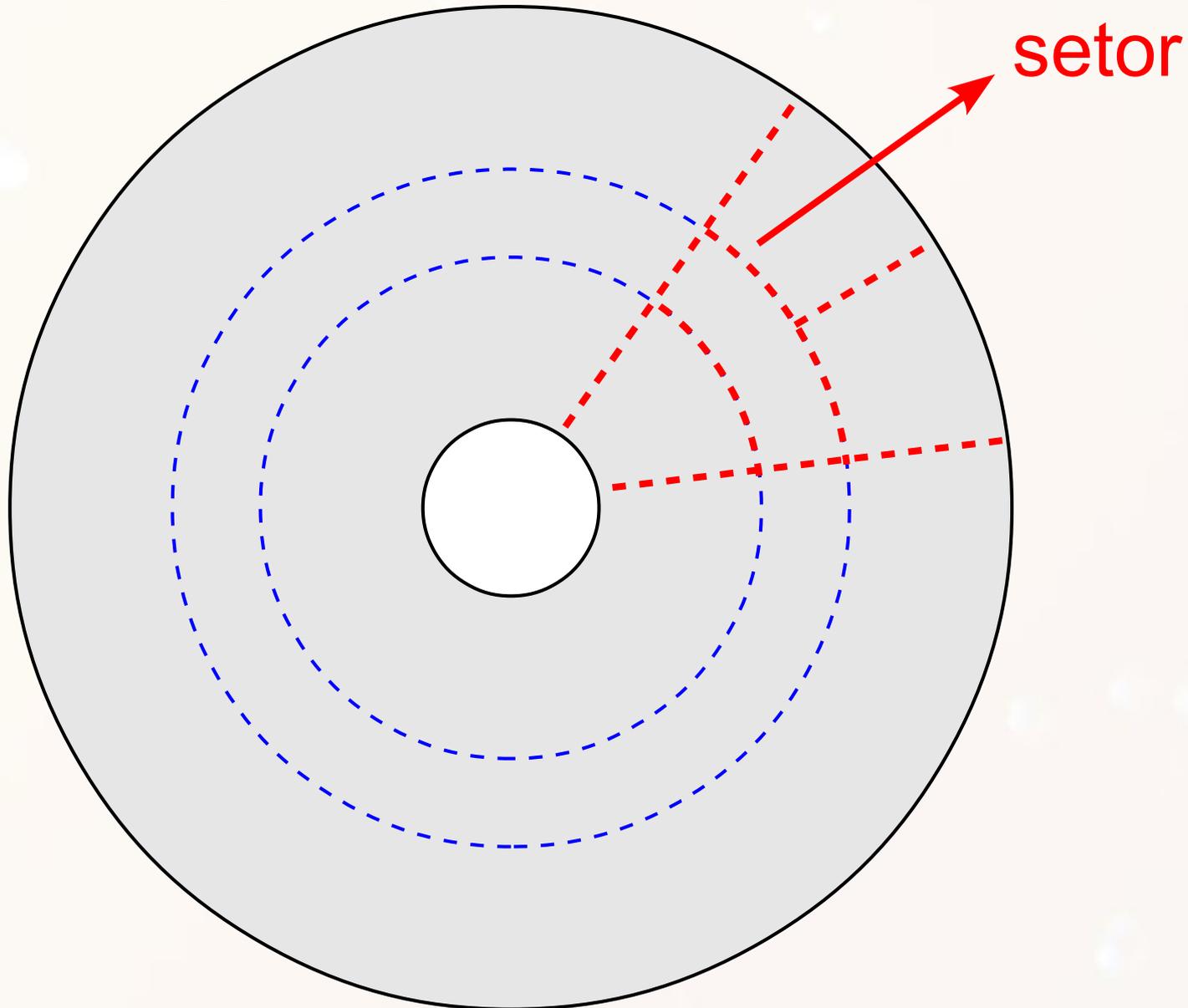
Setor

Divisão por Ângulo Fixo



Setor

Divisão por Densidade Constante



Bloco de Disco ou Página

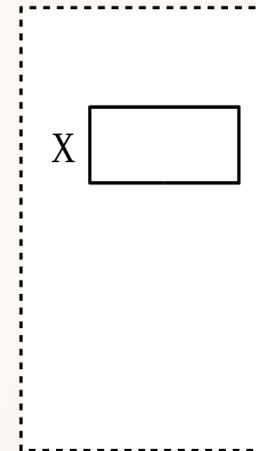
- Organização feita pelo SO sobre o disco
- Unidade de trabalho para o SO

Armazenamento Secundário intermediado pelo Primário

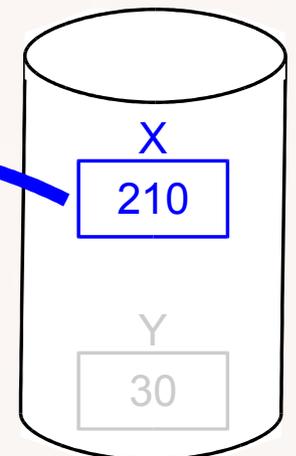
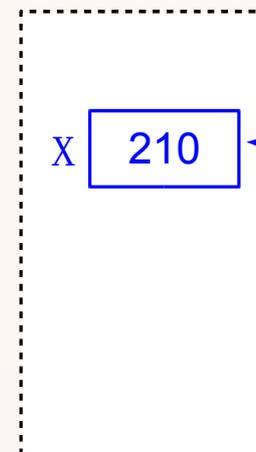
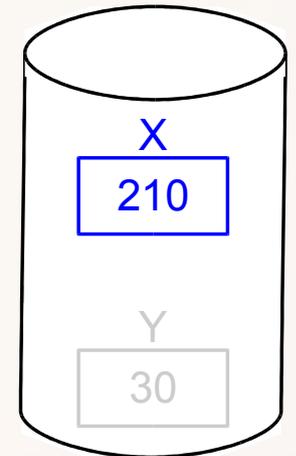
Operação de Leitura Como Abstraímos

- ler(X)

Memória Principal



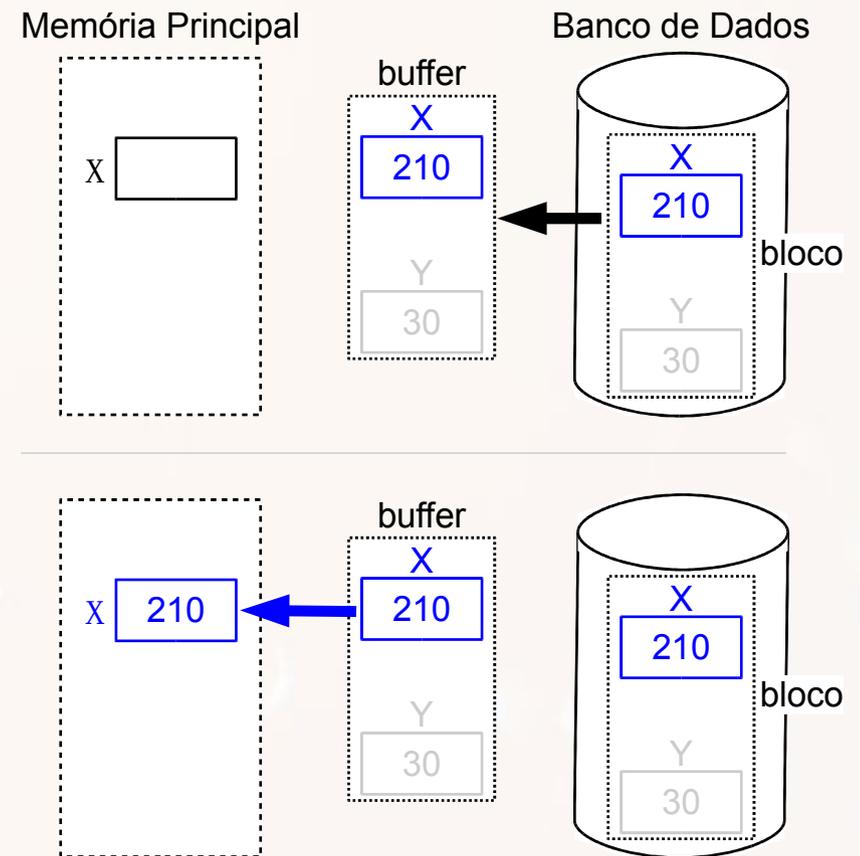
Banco de Dados



Operação de Leitura Como Acontece

■ ler(X)

- encontra bloco X no disco
- copia bloco para buffer da memória principal (se ainda não estiver lá)
- copia o item X do buffer para a variável X da memória principal

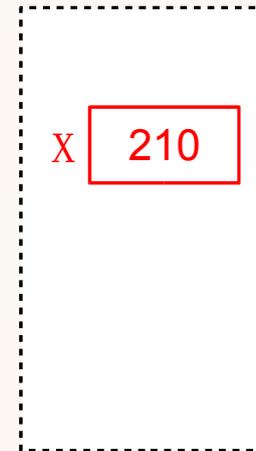


(Elmasri, 2010)

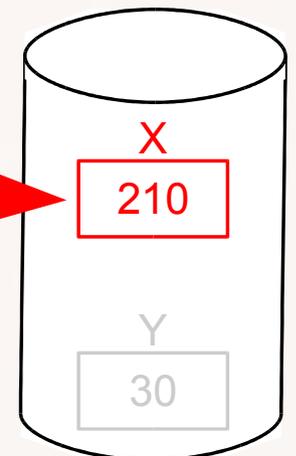
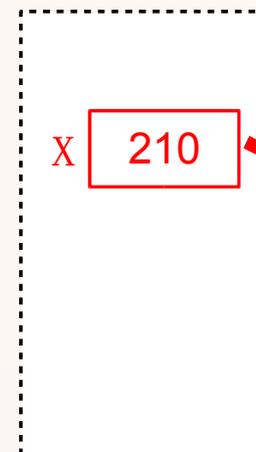
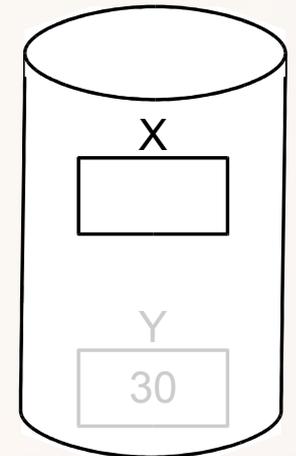
Operação de Gravação Como Abstraímos

- gravar(X)

Memória Principal



Banco de Dados

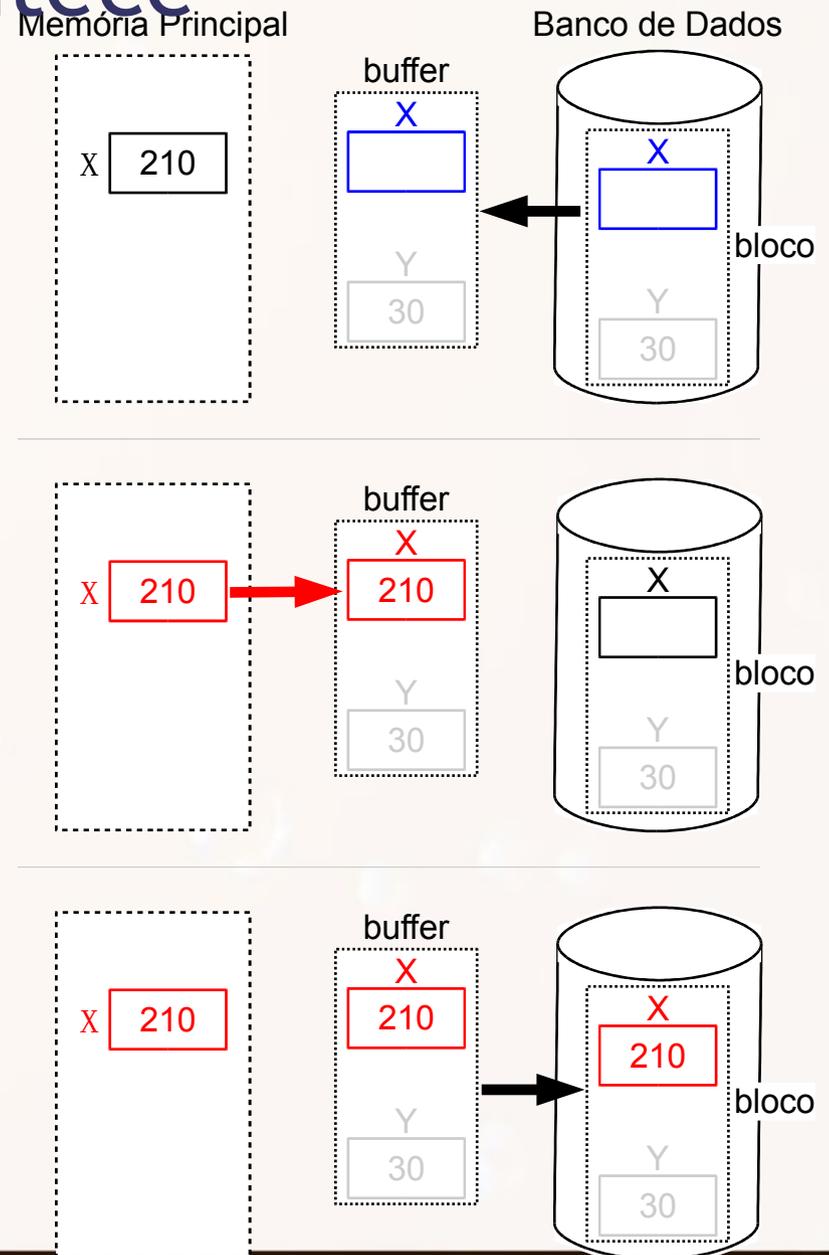


Operação de Gravação Como Acontece

■ gravar(X)

- encontra bloco X no disco
- copia bloco para buffer da memória principal (se ainda não estiver lá)
- copia variável X da memória principal para o buffer
- atualiza o buffer no disco

(Elmasri, 2010)



Arquivos e Registros

- Um arquivo de registros - abstração para SGBD
(Ramakrishnan, 2003)
- **Arquivo:** abstração criada pelo SO para os blocos de disco
- **Registro:** abstração de subdivisão do arquivo criada pela aplicação ou SGBD para o arquivo

Organização de Arquivos

- Heap
 - sem ordenação
 - gravação em qualquer posição
- Sequencial
 - gravação em ordem sequencial
- Hash
 - uso de função de hash

(Silberschatz, 2006)

Hashing

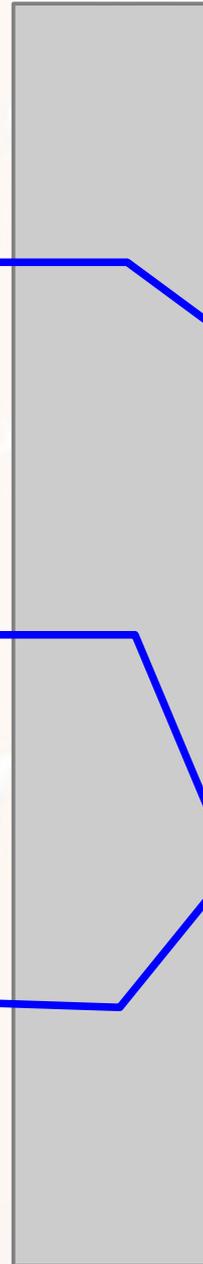
função
hash

buckets

1234 Quincas

9876 Doriana

6543 Asdrúbal



1

...

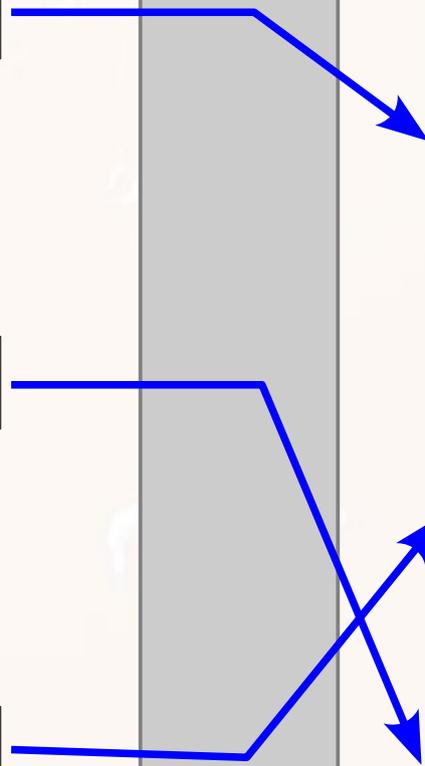
12 1234/Quincas

...

65 6543/Asdrúbal

...

98 9876/Doriana



função hash

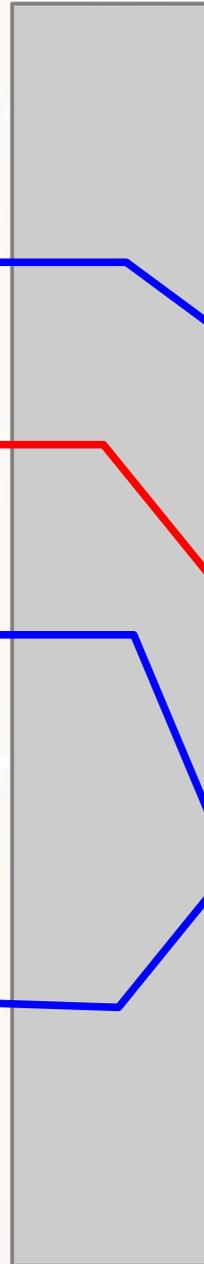
buckets

1234 Quincas

6500 Lucinda

9876 Dorianana

6543 Asdrúbal



1

...

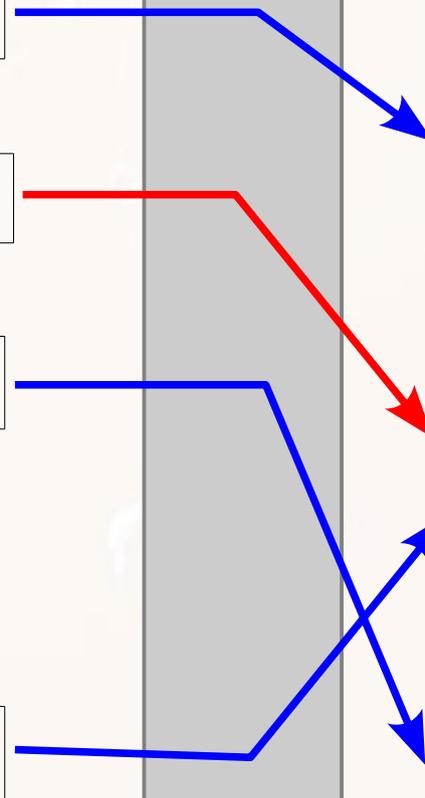
12 1234/Quincas

...

65 6543/Asdrúbal

...

98 9876/Dorianana



função hash

buckets

1234 Quincas

6500 Lucinda

9876 Doriana

6543 Asdrúbal

1

...

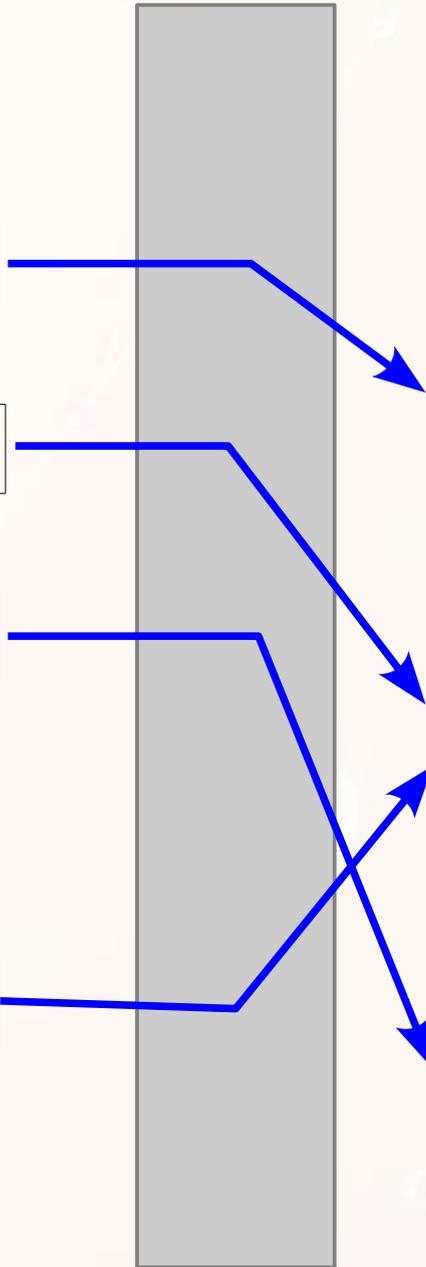
12 1234/Quincas

...

65 6543/Asdrúbal 6500/Lucinda

...

98 9876/Doriana



função hash

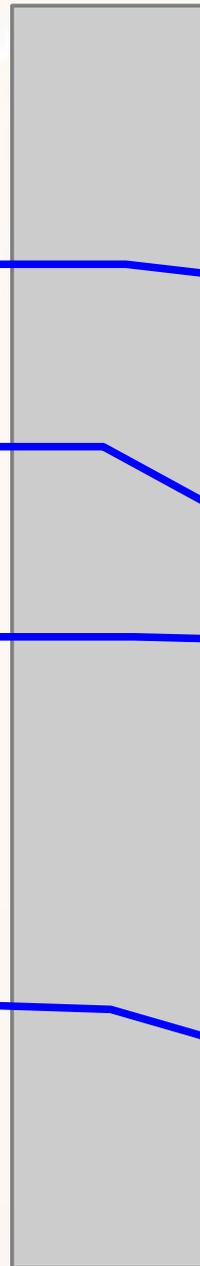
buckets

1234 Quincas

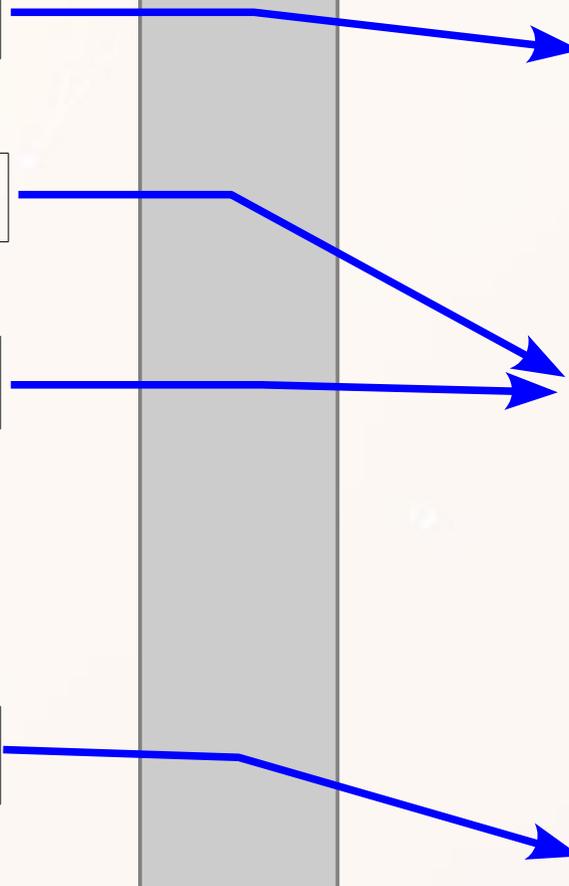
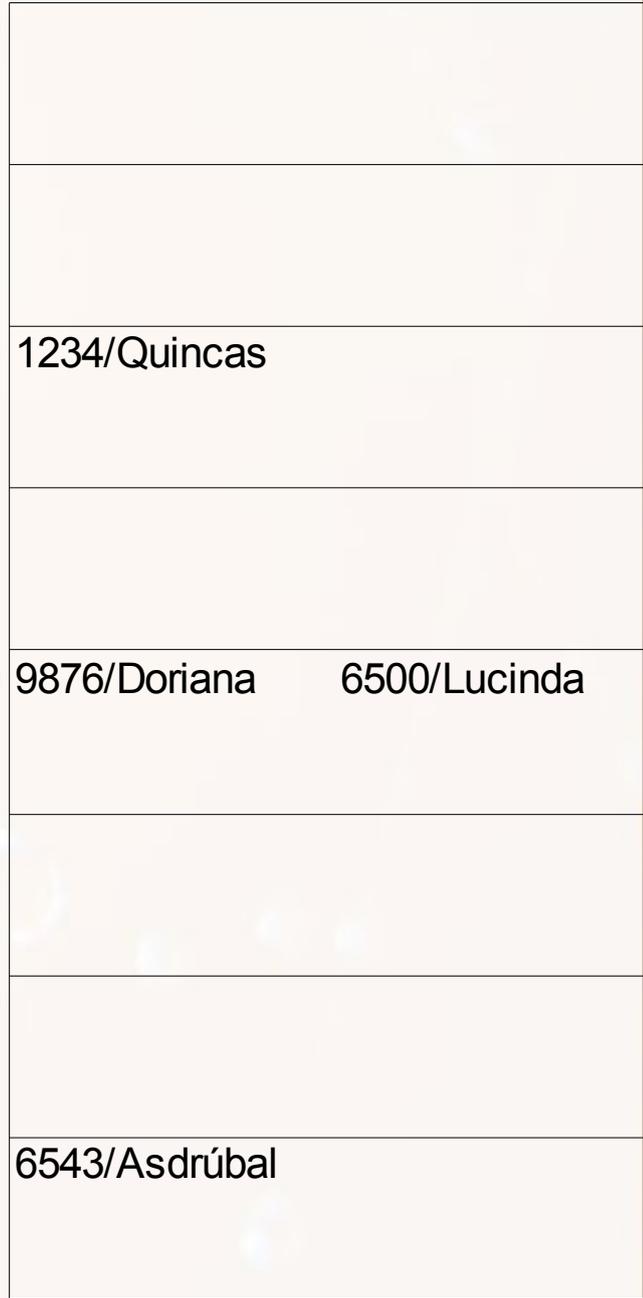
6500 Lucinda

9876 Doriana

6543 Asdrúbal



000
001
010
011
100
101
110
111



Índice

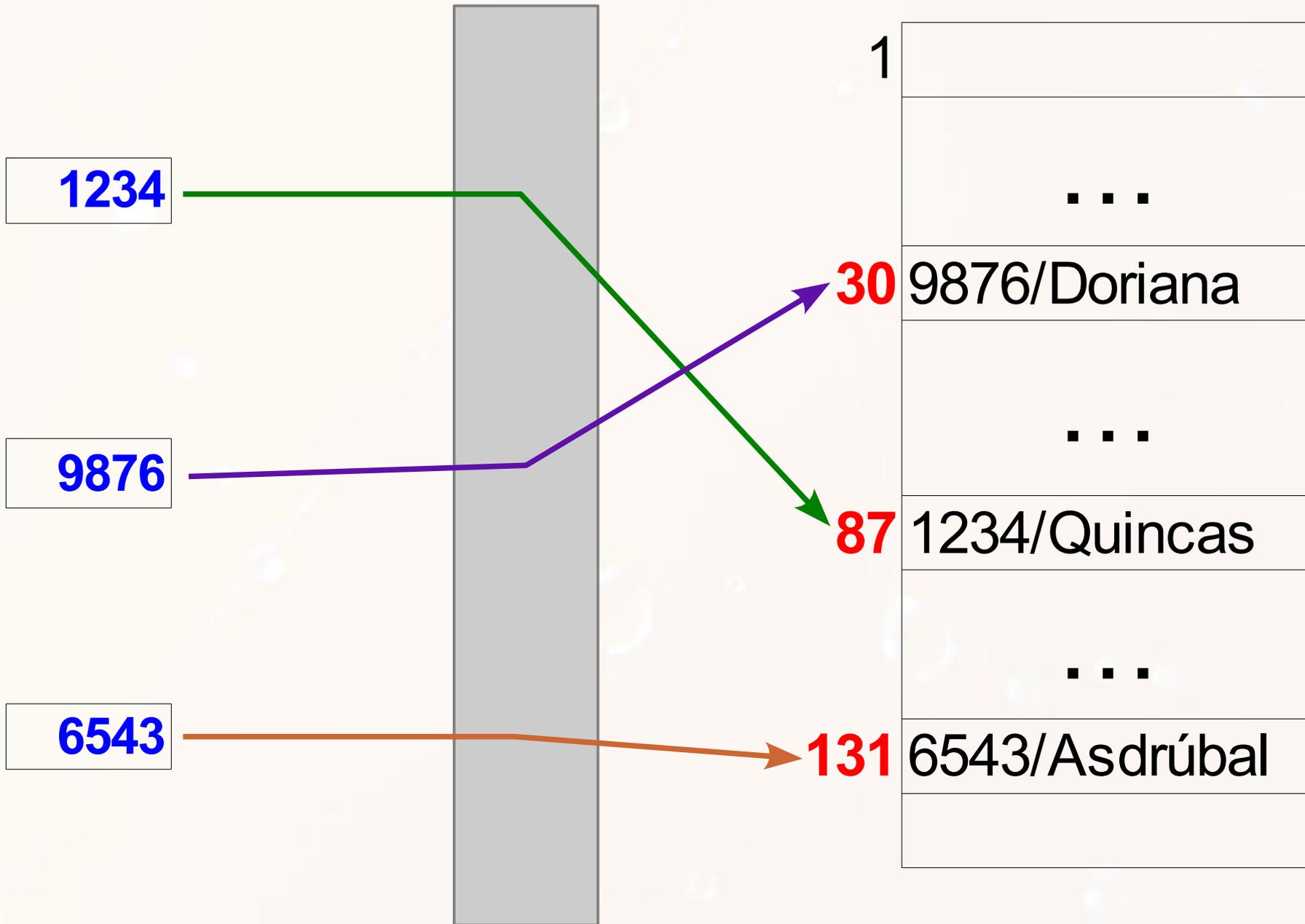
Índice

- Estrutura de dados
- Organiza registros
- Otimiza certas operações de recuperação

(Ramakrishnan, 2003)

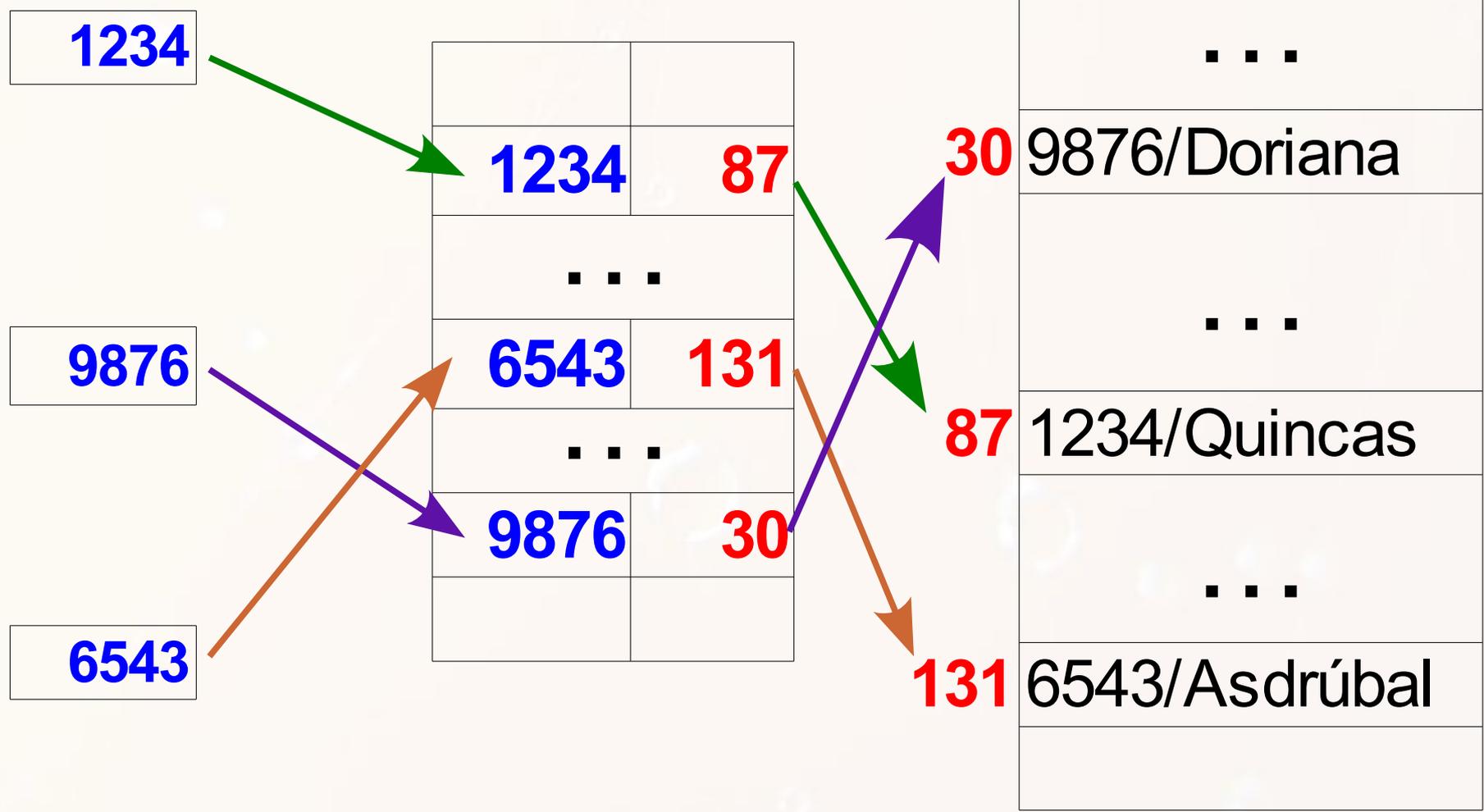
índice

registros



índice

registros



Exercício 2

- Nas aulas anteriores, discutimos sobre redundância de informação e seus potenciais problemas. Índices são estruturas que introduzem redundância no banco de dados. Descreva o impacto da introdução deste tipo de redundância em termos de:
 - a) Consistência dos dados
 - b) Velocidade de leitura
 - c) Velocidade de gravação

Arquivos e Indexação

- Entrada de índice (*data entry*) → registros armazenados em um índice
 - entrada de índice: k^*
 - chave: k
 - Alternativas para k^*
 - (1) registro completo de chave k
 - (2) $(k, \text{rid}) \rightarrow \text{rid} = \text{id do registro de chave } k$
 - (3) $(k, \text{rid-list}) \rightarrow \text{rid-list} = \text{lista de registros de chave } k$
- (Ramakrishnan, 2003)

Exercício 3

Como este índice seria criado em SQL?

```
CREATE CONSTRAINT ON (p:PATO) ASSERT p.id IS  
UNIQUE
```

Índice Único

- Índice cujas chaves não tem duplicatas

Índices de Agrupamento (Clustering)

- chave de busca (índice) => ordem dos registros (arquivo)
- modelo (1): k^* é o registro de dados

Índices Primários e Secundários

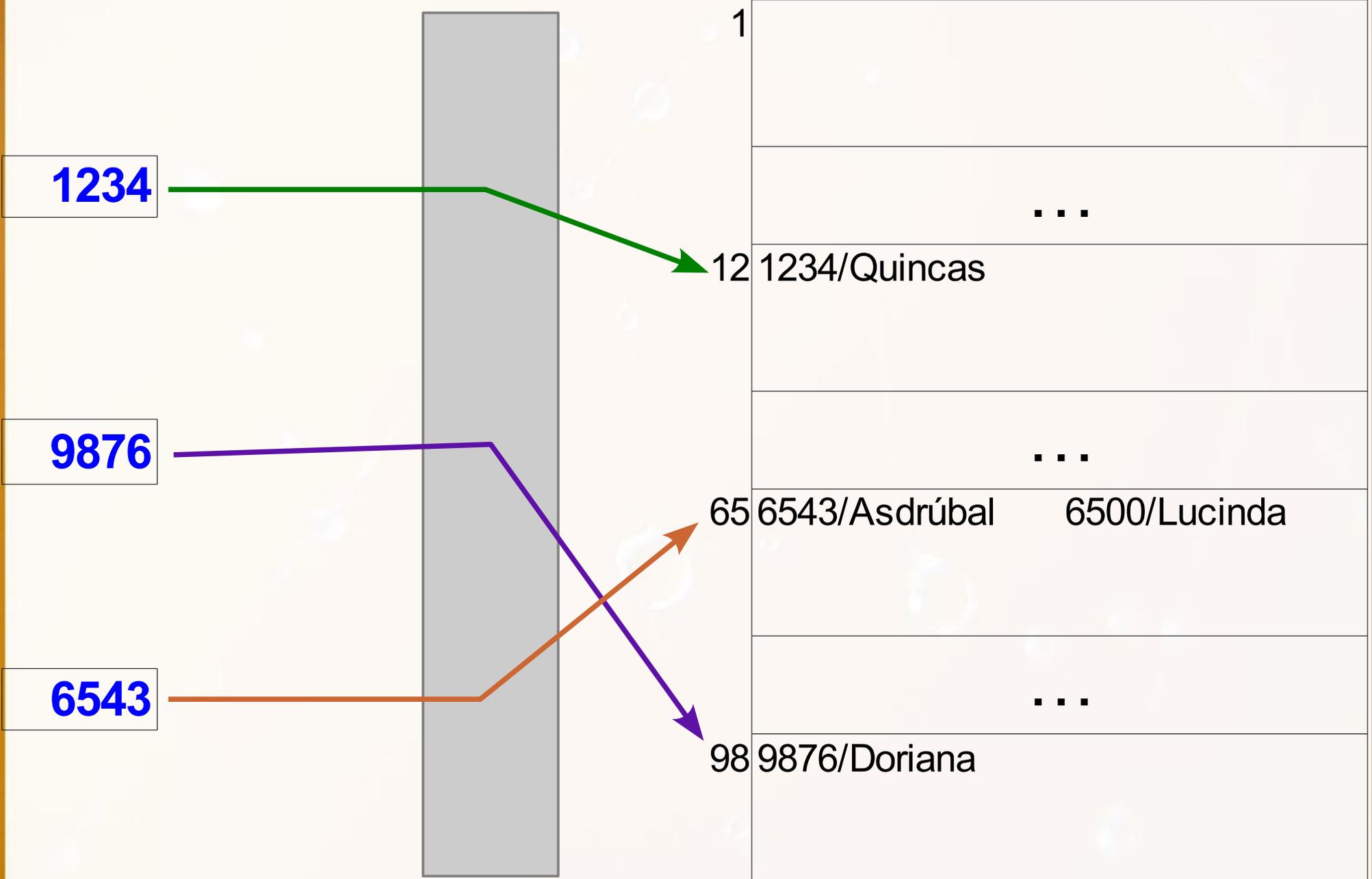
- Índice primário
 - índice de agrupamento com chave primária
 - índice único
- Índice secundário
 - índice de não agrupamento
 - índice não necessariamente único

Índices Densos e Esparsos

- Denso
 - um registro associado a cada entrada do índice
- Esparso
 - mais de um registro associado a cada entrada do índice

índice

buckets registros



Exercício 4

- Em uma relação com 5 atributos, qual o número máximo possível de índices de agrupamento? Justifique.

Índice Mapa de Bits

Índice de Hash

Hashing Extensível

função hash

buckets

1234 Quincas

6500 Lucinda

9876 Doriana

6543 Asdrúbal

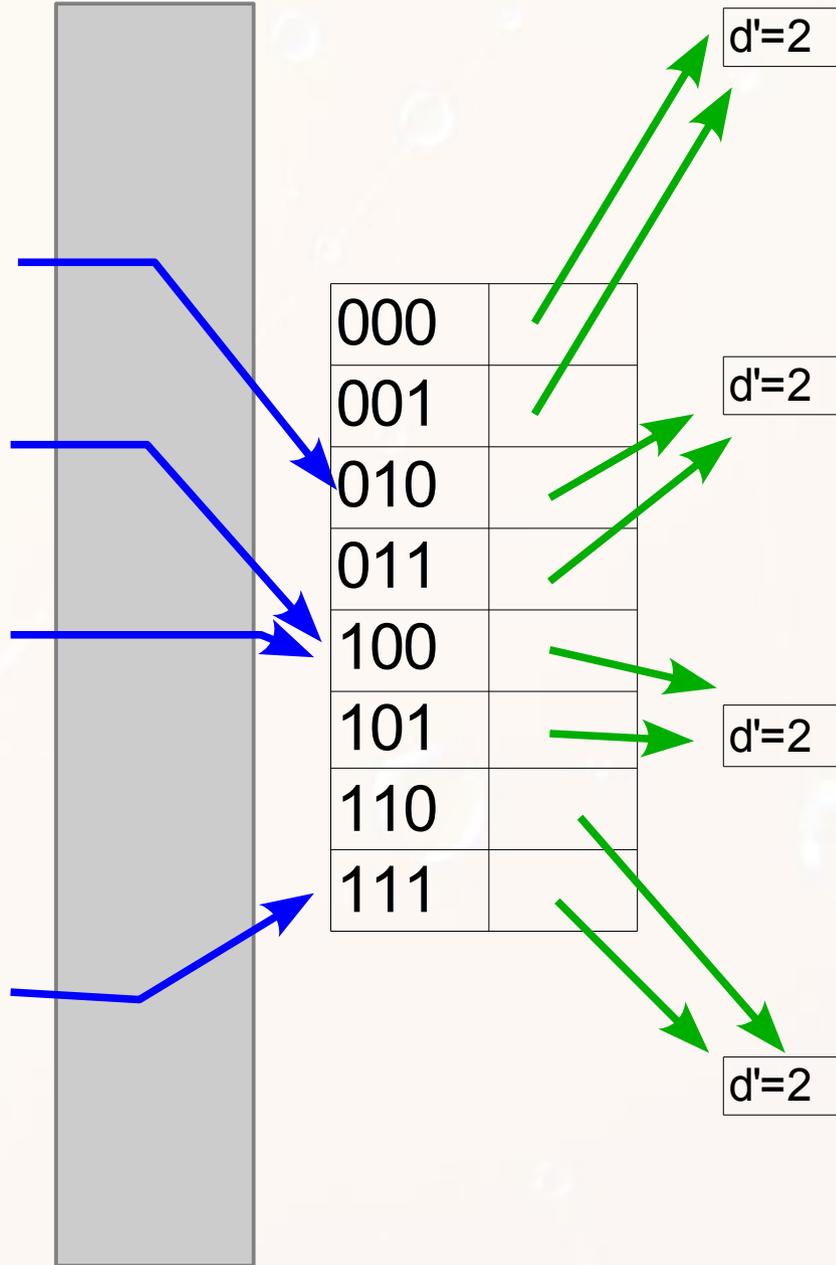
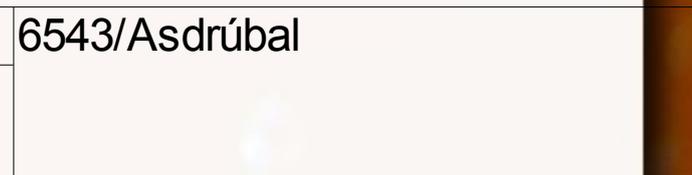
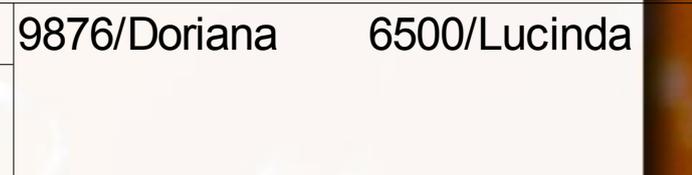
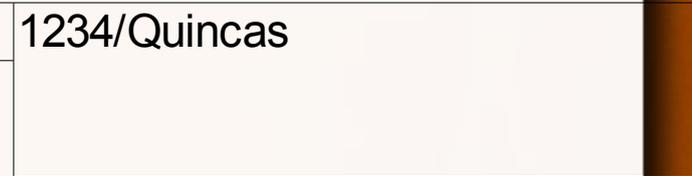
000	
001	
010	
011	
100	
101	
110	
111	

d'=2

d'=2

d'=2

d'=2



função hash

buckets

1234 Quincas

6500 Lucinda

9876 Doriana

8500 Dóris

6543 Asdrúbal

000	
001	
010	
011	
100	
101	
110	
111	

d'=2

d'=2

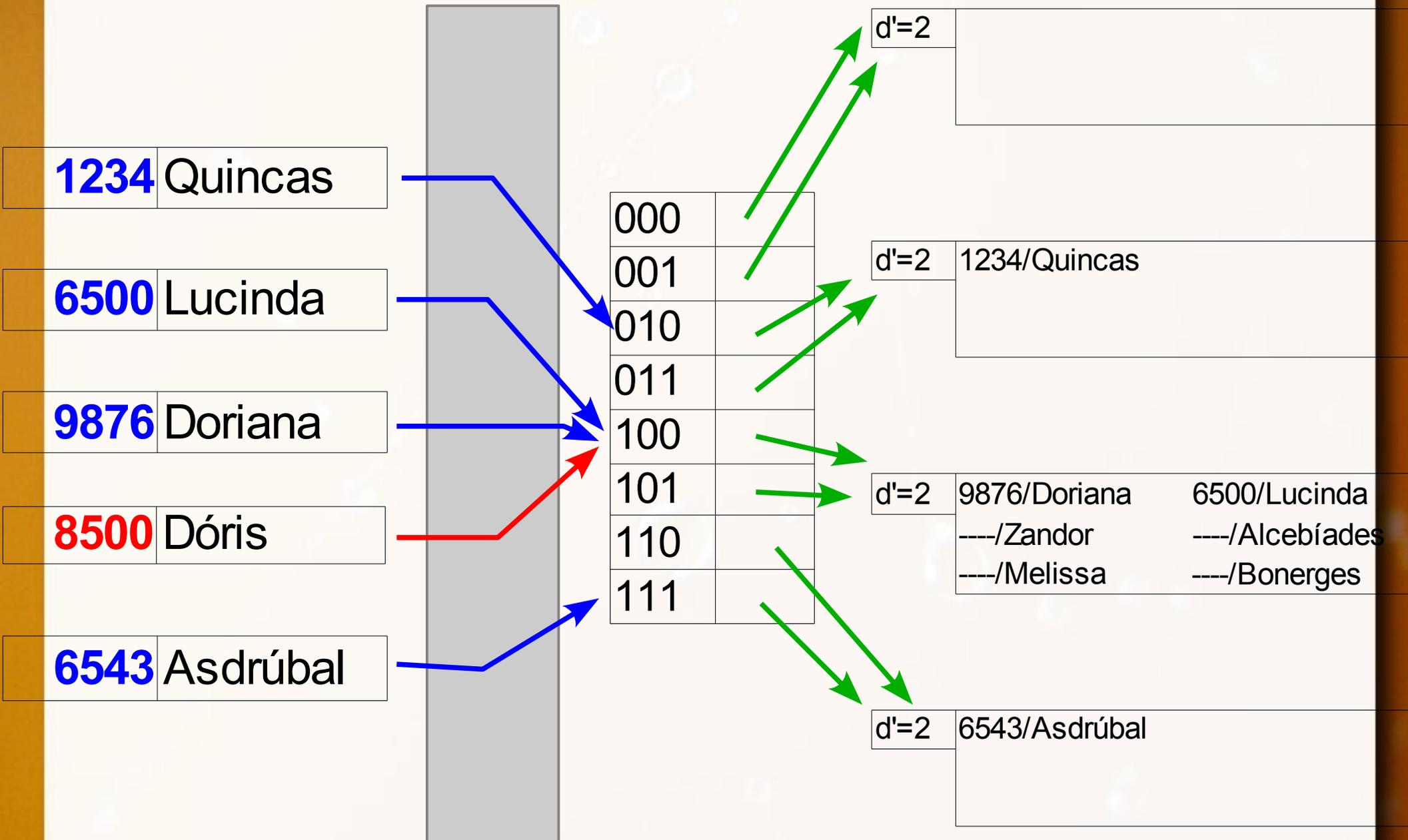
d'=2

d'=2

1234/Quincas

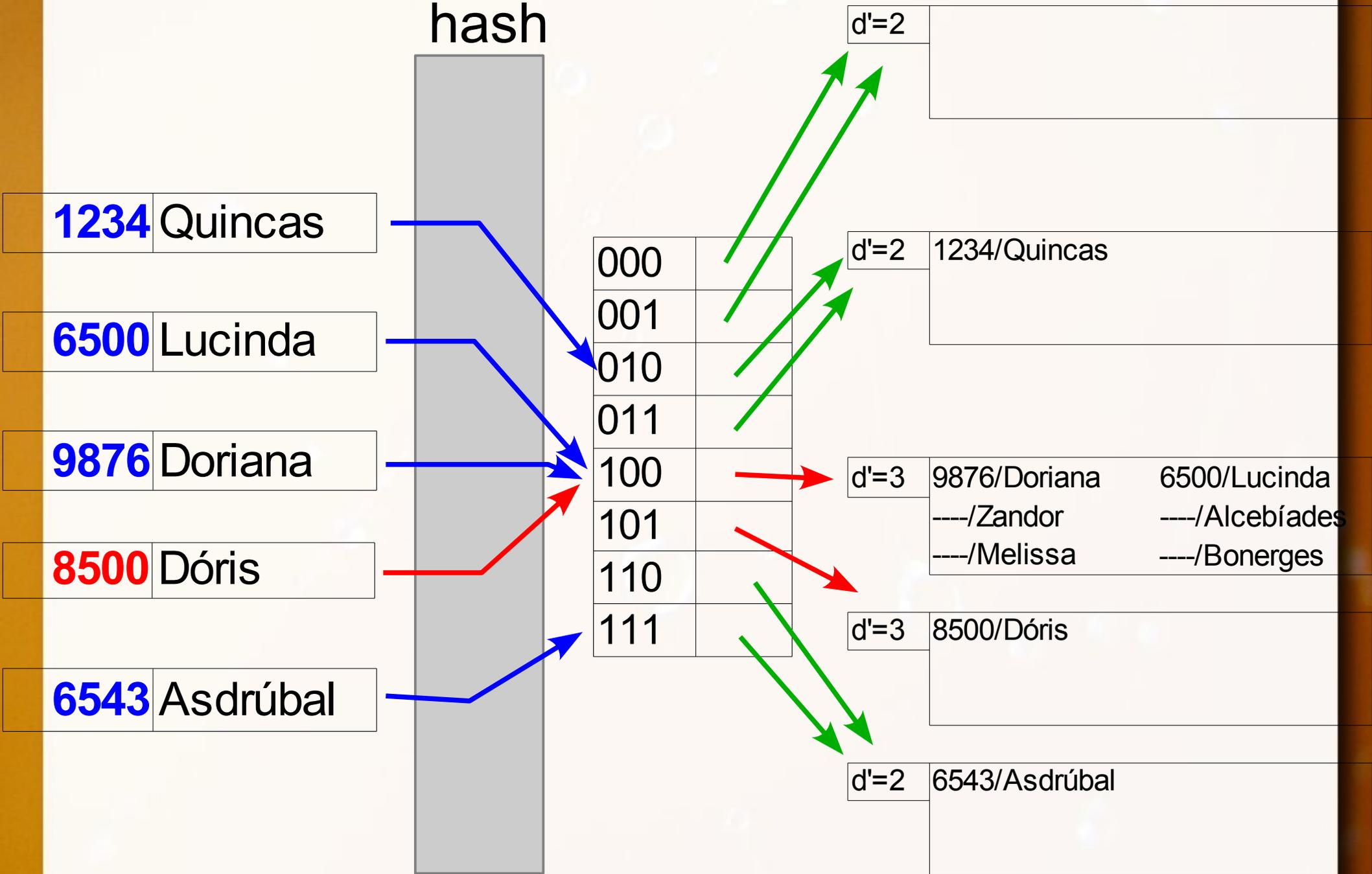
9876/Doriana 6500/Lucinda
----/Zandor ----/Alcebíades
----/Melissa ----/Bonerges

6543/Asdrúbal



função hash

buckets



1234 Quincas

6500 Lucinda

9876 Doriana

8500 Dóris

6543 Asdrúbal

000

001

010

011

100

101

110

111

d'=2

d'=2

d'=3

d'=3

d'=2

1234/Quincas

9876/Doriana 6500/Lucinda
---/Zandor ---/Alcebíades
---/Melissa ---/Bonerges

8500/Dóris

6543/Asdrúbal

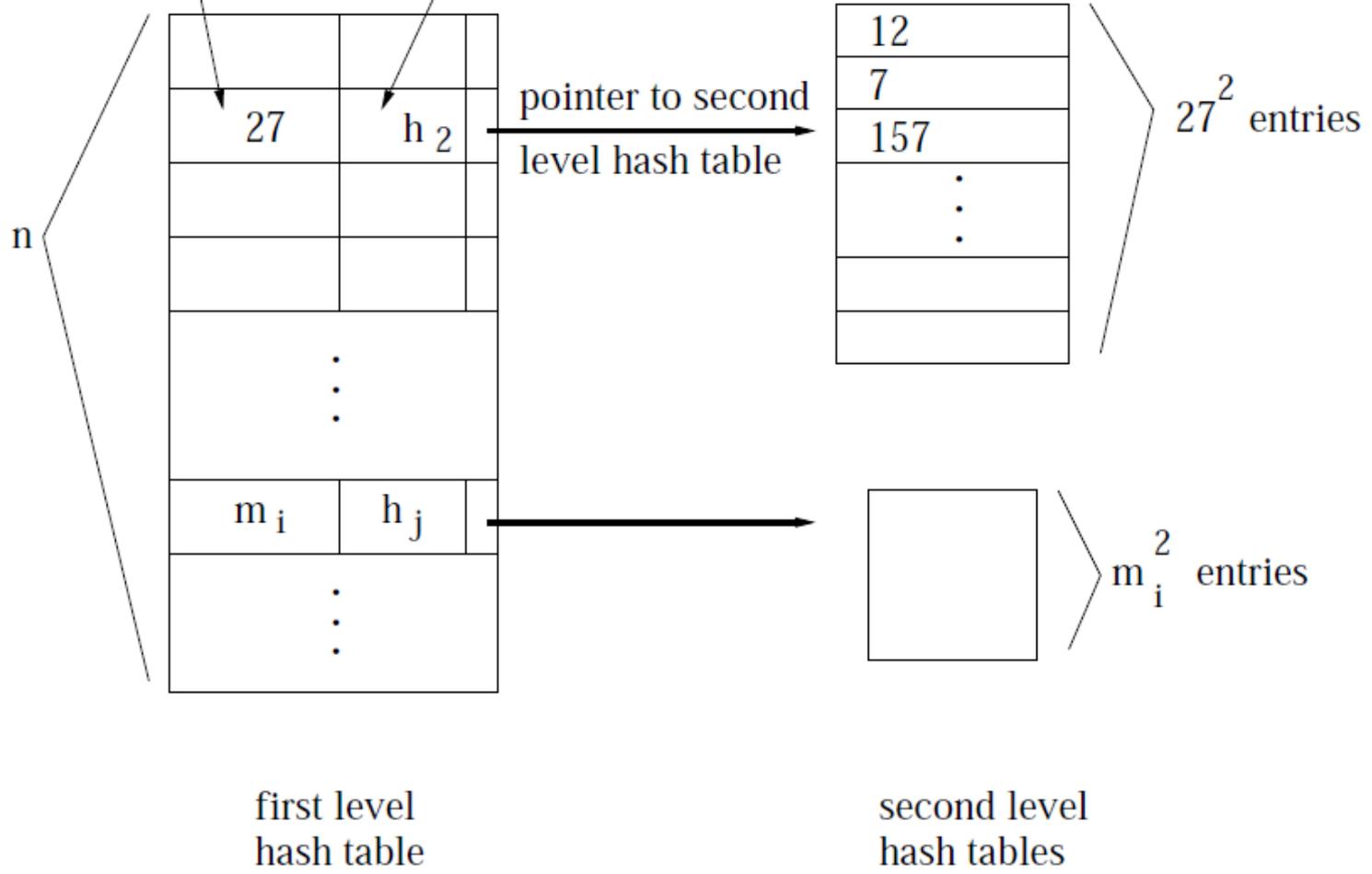
Índices Multiníveis

Hashing Dinâmico

Perfect Hashing

how many elements hash to this slot

a hash function from family H

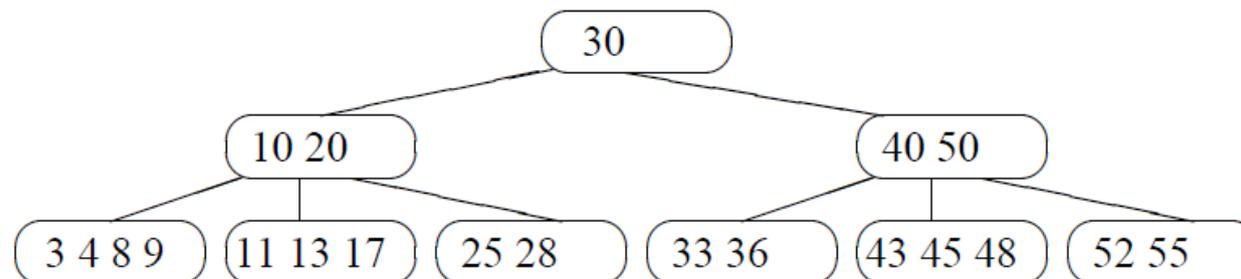


(Demaine, 2003)

Árvores B

Árvores B

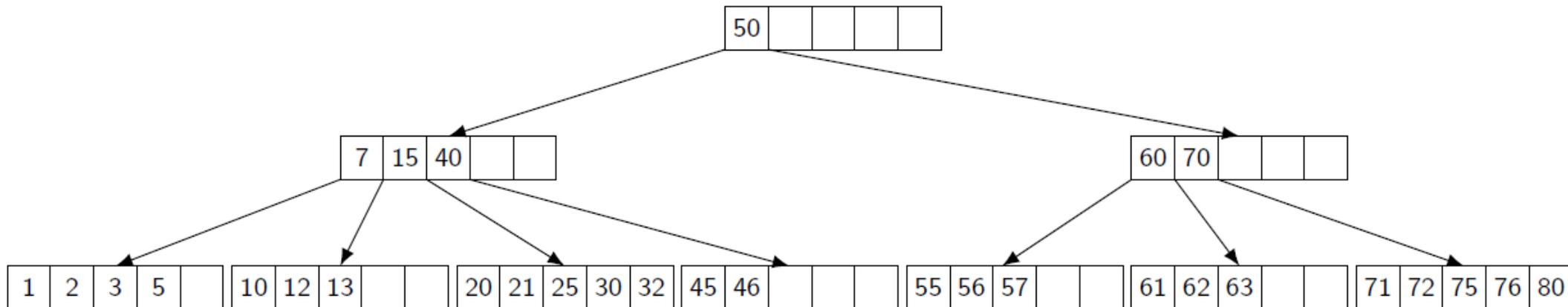
- Árvores n -árias: mais de um registro por nodo.
- Em uma árvore B de ordem m :
 - página raiz: 1 e $2m$ registros.
 - demais páginas: no mínimo m registros e $m + 1$ descendentes e no máximo $2m$ registros e $2m + 1$ descendentes.
 - páginas folhas: aparecem todas no mesmo nível.
- Registros em ordem crescente da esquerda para a direita.
- Extensão natural da árvore binária de pesquisa.
- Árvore B de ordem $m = 2$ com três níveis:



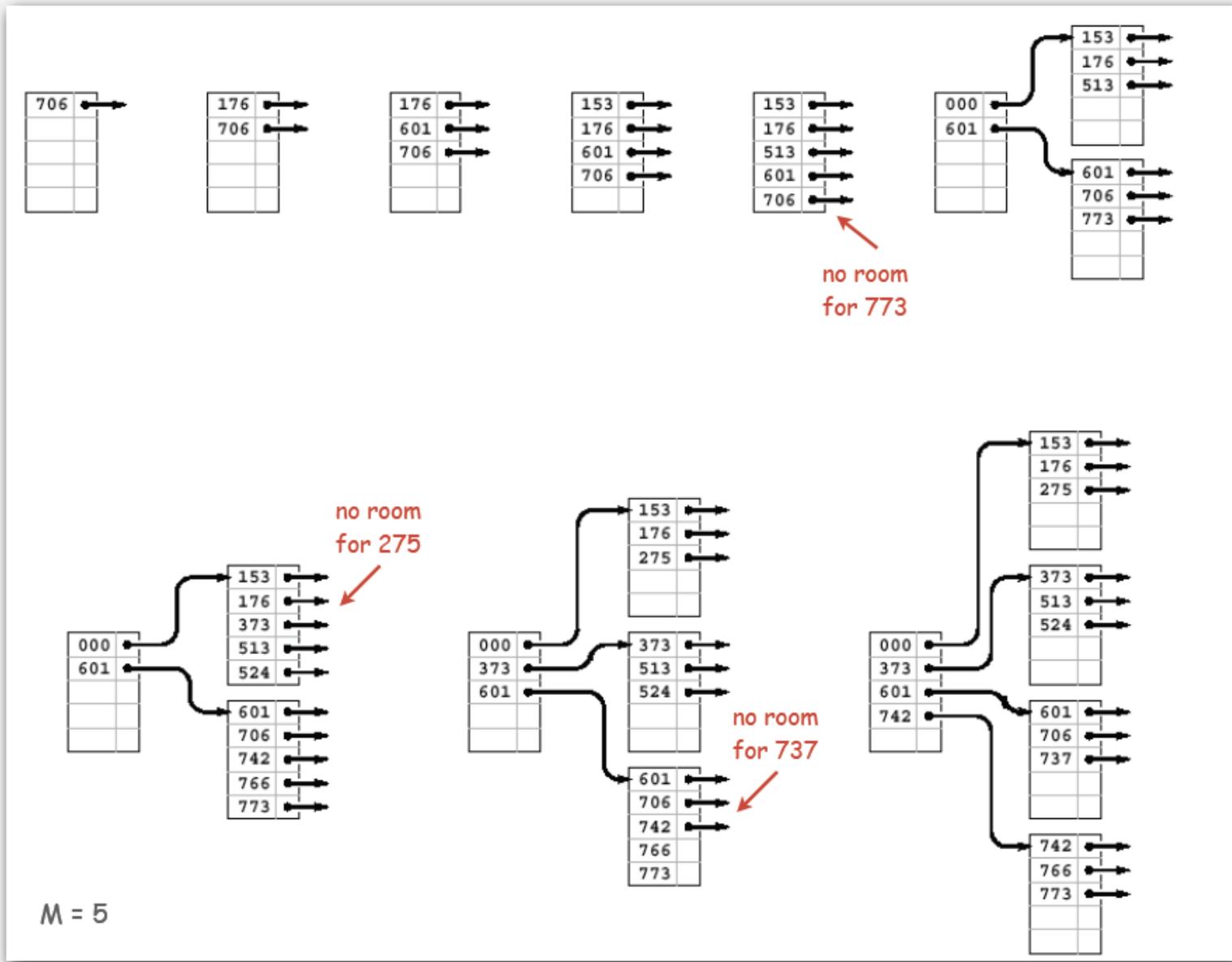
(Almeida, 2010)

Exemplo de árvore B de ordem 5

Neste caso, cada nó tem no mínimo dois e no máximo cinco registros de informação.

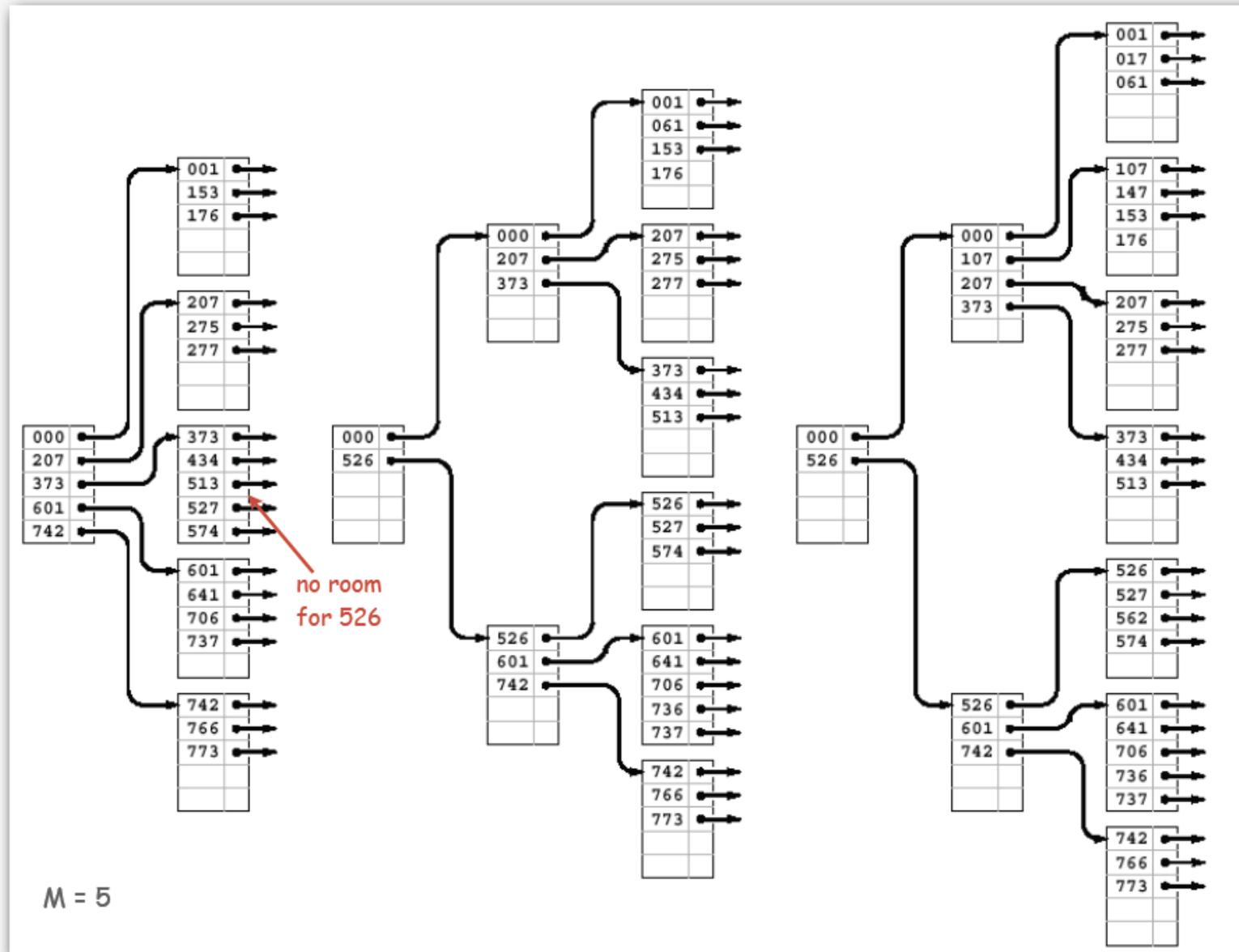


B-Tree Example



M = 5

B-Tree Example (cont)



$M = 5$

Números mínimos e máximos de registros

Árvore B de ordem 255:

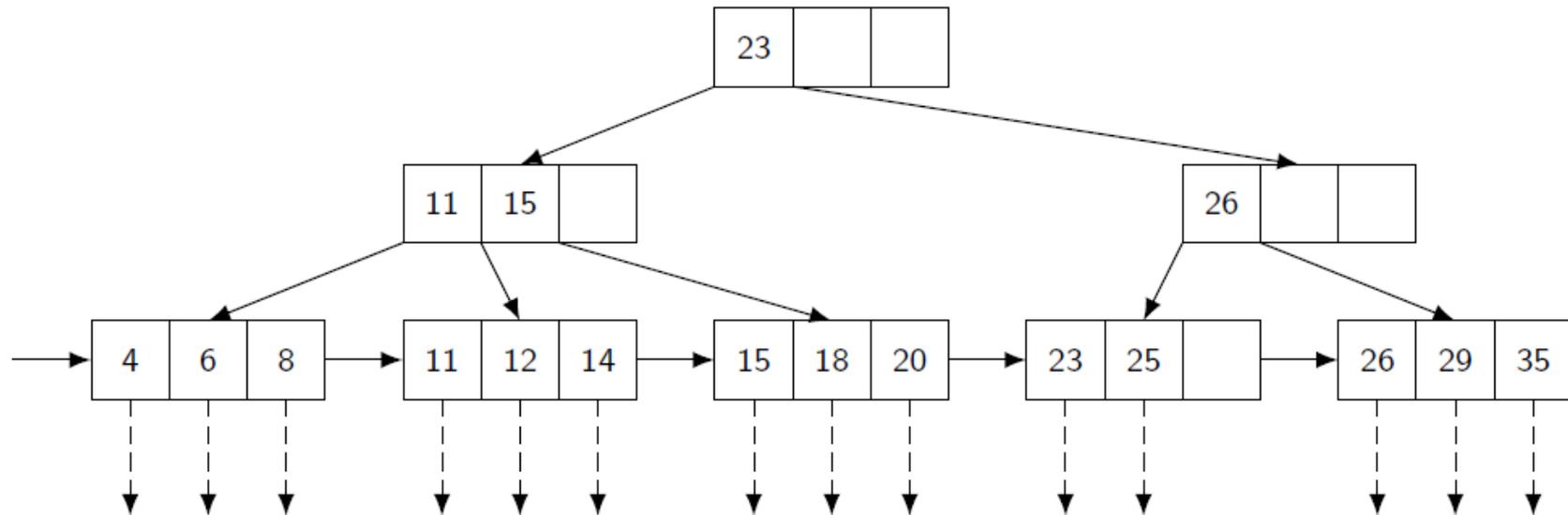
nível	mínimo		máximo	
	nós	registros	nós	registros
1	1	1	1	1×255
2	2	2×127	256^1	$256^1 \times 255$
3	2×128^1	$2 \times 128^1 \times 127$	256^2	$256^2 \times 255$
4	2×128^2	$2 \times 128^2 \times 127$	256^3	$256^3 \times 255$
5	2×128^3	$2 \times 128^3 \times 127$	256^4	$256^4 \times 255$
Total	4.227.331	536.870.911	4.311.810.305	1.099.511.627.775

Variantes de árvores B

- ▶ Árvores B*: o número de registros ocupados de um nó é no mínimo $\frac{2}{3}$ da sua capacidade.
- ▶ Árvores B⁺:
 - ▶ nós internos com chaves apenas para orientar o percurso
 - ▶ pares (*chave, valor*) apenas nas folhas
 - ▶ regra de descida:
 - ▶ subárvore esquerda: menor
 - ▶ subárvore direita: maior ou igual
 - ▶ apontadores em lugar de valores tornando mais eficiente a movimentação dos registros durante inserções e remoções
 - ▶ ligações facilitando percurso em ordem de chaves

Variantes de árvores B (cont.)

Exemplo de árvore B⁺ de ordem 3:



Setas tracejadas indicam apontadores para os valores da informação. A lista ligada das folhas permite percurso simples e eficiente em ordem de chaves.

Exercício 5

- Considere a relação Aluno(ra, curso, idade) que armazene estes dados para todos os alunos da Unicamp. Para cada uma das questões a seguir, defina qual o tipo de índice mais indicado.
 - a) `select * from Aluno where ra=5.`
 - b) `select * from Aluno where idade<70.`
 - c) `select * from Aluno where idade>27 and B<30.`
 - d) `select avg(idade) from Aluno.`
 - e) `select idade, count(*) from aluno
where curso="Computação"
group by idade`

Referências

- Almeida, Charles Ornelas , Guerra, Israel; Ziviani, Nivio (2010) **Projeto de Algoritmos** (transparências aula).
- Demaine, Erik. 6.897: Advanced Data Structures - Lecture 2 (notas de aula). Fevereiro, 2003.
- Elmasri, Ramez; Navathe, Shamkant B. (2005) **Sistemas de Bancos de Dados**. Addison-Wesley, 4^a edição em português.
- Elmasri, Ramez; Navathe, Shamkant B. (2011) **Sistemas de Bancos de Dados**. Addison-Wesley, 6^a edição em português.

Referências

- Ramakrishnan, Raghuram; Gehrke, Johannes (2003) **Database Management Systems**. McGraw-Hill, 3rd edition.
- Sedgwick, Robert; Wayne, Kevin (2008) Princeton University: **Algorithms**. Maio, 2008.
- Silberschatz, Abraham; Korth, Henry F.; Sudarshan, S. (2006) **Sistema de Banco de Dados**. Elsevier, Tradução da 5a edição.

Agradecimentos

- Luiz Celso Gomes Jr (professor desta disciplina em 2014) pela contribuição na disciplina e nos slides.
- Luana Loubet Borges pelos exercícios.

André Santanchè

`http://www.ic.unicamp.br/~santanche`

Licença

- Estes slides são concedidos sob uma Licença Creative Commons. Sob as seguintes condições: Atribuição, Uso Não-Comercial e Compartilhamento pela mesma Licença.
- Mais detalhes sobre a referida licença Creative Commons veja no link:
<http://creativecommons.org/licenses/by-nc-sa/3.0/>
- Fotografia da capa e fundo por
<http://www.flickr.com/photos/fdecomite/>
Ver licença específica em
<http://www.flickr.com/photos/fdecomite/1457493536/>