

# Árvores B

Ana Eliza Lopes Moura

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

1

## Situação Problema

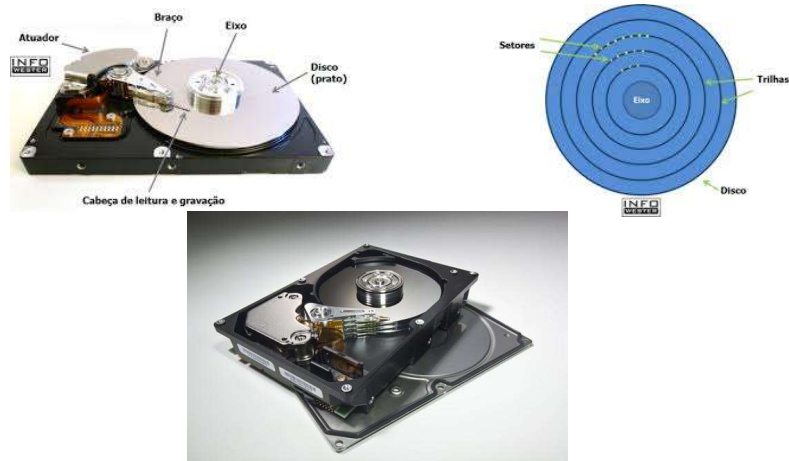
- Memória Principal
  - Volátil e limitada
- Aplicações
  - Grandes quantidades de informação
  - Armazenamento permanente de informações

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

2

## Situação Problema

### □ Memória Secundária: leitura por setor



3

## Situação Problema

### □ Estruturas de Dados para Armazenamento Secundário

- Árvores de Busca Binária
  - Adequado para memória principal
  - Ineficiente em memória secundária
    - Grande quantidade de acessos a disco
      - » Acesso feito em blocos

anaeliza.moura@unicap.br

4

## Situação Problema

### □ Necessidade

- Reduzir o número de acessos a disco

### □ Solução

- Agrupar várias chaves dentro de um nó
  - Obter com o mesmo acesso vários registros
  - Reduzir o número de acessos
  - Diminuir o tempo necessário para inserções, remoções e pesquisas.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

5

## Árvores Multivias ou M-Vias

### □ Definição

- Uma árvore de busca multivias de ordem **M** é uma árvore n-ária na qual todos os nós têm grau menor ou igual a **M**.
- Um nó com **M** descendentes contém **M-1** valores de chave.

### □ Lembrete:

O **grau de um nó** é o número de subárvores (filhos) do nó.

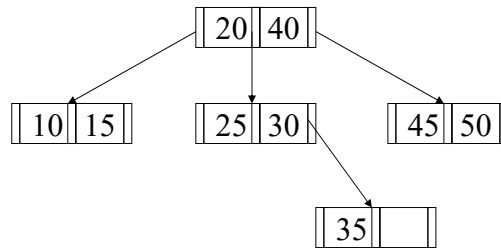
O **grau de uma árvore** (n-aridade) é a quantidade máxima de filhos que os nós de uma árvore podem ter.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

6

## Árvores Multivias ou M-Vias

### □ Exemplo: Ordem $M = 3$



anaeliza.moura@unicap.br

7

## Árvores Multivias ou M-Vias

### □ Desempenho da Busca

- Árvores multivias com **N** chaves no total e fator de ramificação **S**.
- Caminho médio de busca:  $O(\log_S N)$
- Se  $N = 10^6$  e  $S = 100$ , então uma busca requer, em média,  $\log_{100} 10^6 = 3$  passos.
- Árvore de busca binária:  $\log_2 10^6 = 20$  passos.

anaeliza.moura@unicap.br

8



## Árvores Multivias ou M-Vias

- Problema
  - Inserções aleatórias de maneira irrestrita
  - Aumento do caminho de busca
- Solução
  - Balanceamento

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

9



## Árvores B

- Definição
  - Bayer e McCreight em 1970.
  - Uma árvore B de ordem **M** é uma árvore de busca multivias **balanceada**.
  - Uma árvore B ou está vazia ou possui nós com **K** chaves e **K+1** apontadores.

OBS: Um nó de uma árvore B é chamado de **página**.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

10



## Árvores B

### □ Uso

- Árvores B são utilizadas como forma de armazenamento em diversos sistemas de BD comerciais.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

11



## Árvores B

### □ Características estruturais:

- Na raiz, **K** deve ser, no mínimo, 1. Ou seja, a raiz possui no mínimo **uma** chave e **dois** filhos;
- Nos demais nós, **K** deve ser, no mínimo, **M/2**. Ou seja, os demais nós possuem, no mínimo, **M/2** chaves e **M/2 + 1** filhos;
  - Exceção: Folhas não têm filhos.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

12

## Árvores B

### □ Características (cont.):

- O valor máximo de **K** é **M-1**, ou seja, todos os nós têm, no máximo, **M-1** chaves e **M** filhos;
- Todas as folhas estão no mesmo nível (balanceamento).

**OBS:** **M** deve ser escolhido de forma que o número máximo de chaves nos nós da árvore seja uma potência de 2.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

13

## Árvores B

### □ Características (cont.)

- Formato do nó:

$N, (C_1A_1), (C_2A_2), \dots, (C_{M-1}A_{M-1}), A_M$  onde:

- $N, M/2+1 \leq N \leq M$ , é o número de entrada ativas (ocupadas) de um nó em um dado momento;
- $A_i, 1 \leq i \leq M$ , é um apontador para uma subárvore;
- $C_i, 1 \leq i \leq M-1$ , é um valor de chave e  $C_i < C_{i+1}$ ;
- O par  $(C_iA_i)$  é chamado de entrada;
- O apontador  $A_M$  também é definido como entrada.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

14

## Árvores B

### □ Características (cont.)

– Definição do nó:

```
class NodeB <T> {  
    private int n;  
    private T[] chv;  
    private NodeB<T>[] pont;  
    NodeB (int m) { // construtor  
        chv = new T[m-1];  
        pont = new NodeB[m];  
    }  
    ...  
}
```

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

15

## Árvores B

### □ Características (cont.):

– Seja uma página com **N** chaves:

- Para qualquer chave **y**, pertencente à página apontada por **A<sub>1</sub>**, **y < C<sub>1</sub>**;
- Para qualquer chave **y**, pertencente à página apontada por **A<sub>i</sub>**,  $2 \leq i \leq N$ , **C<sub>i-1</sub> < y < C<sub>i</sub>**;
- Para qualquer chave **y**, pertencente à página apontada por **A<sub>N+1</sub>**, **y > C<sub>N</sub>**.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

16



## Árvores B

- Comparação em termos de nós e chaves por nível entre uma árvore binária e uma árvore B de ordem **M** de mesma altura.

Nível	Binária	Árvore B
0	1 nó	1 nó x M-1 chaves
1	2 nós	M nós x (M-1) chaves
2	4 nós	M x M nós x (M-1) chaves
3	8 nós	M x M x M nós x (M-1) chaves
...	...	...
n	$2^n$ nós	$M^n$ nós x (M-1) chaves

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

17

## Árvores B

### □ Observações:

- A ordem **M** determina as quantidades máximas e mínimas de chaves dentro de cada nó.
- O número mínimo de chaves é estabelecido para determinar o percentual mínimo de ocupação dentro de um nó. Na árvore B esse percentual é de 50% (não considerando a raiz).

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

18

## Árvore B

### □ Inserção

- Em uma árvore B, a inserção de uma nova chave ocorre sempre em um nó folha.
- Passos:
  - Localizar a folha dentro da qual a chave deve ser inserida;
  - Se a folha não estiver completa, inserir chave na ordem correta;
  - Se a folha estiver completa, realizar a **cisão da página**.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

19

## Árvore B

### □ Inserção (Exemplo)

- Inserir chave 85 

85			
----	--	--	--
- Inserir chave 60 

60		85		
----	--	----	--	--
- Inserir chave 52 

52		60		85	
----	--	----	--	----	--
- Inserir chave 70 

52		60		70		85
----	--	----	--	----	--	----
- Inserir chave 58 ⇐ **Realizar cisão**

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

20



## Árvore B

### □ Inserção -> Cisão de Página

- O processo de cisão consiste em separar a folha completa em duas: folha esquerda e folha direita.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

21



## Árvore B

### □ Inserção -> Cisão de Página

- As **M** chaves serão divididas em três grupos:
  - as  $(M \div 2)$  chaves menores ficam na folha esquerda;
  - as  $(M \div 2)$  chaves maiores ficam na folha direita;
  - a chave do meio é colocada no nó pai, se possível.

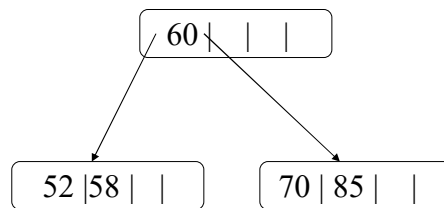
[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

22

## Árvore B

### □ Inserção (Exemplo - cont.)

- Inserir chave 58 (antes) 52 | 60 | 70 | 85
- Inserir chave 58 (depois)



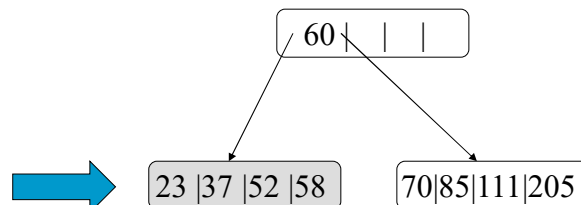
anaeliza.moura@unicap.br

23

## Árvore B

### □ Inserção (Exemplo - cont.)

- Inserir chaves 37, 111, 23, 205



- Inserir chave 5 ⇐ **Realizar cisão**

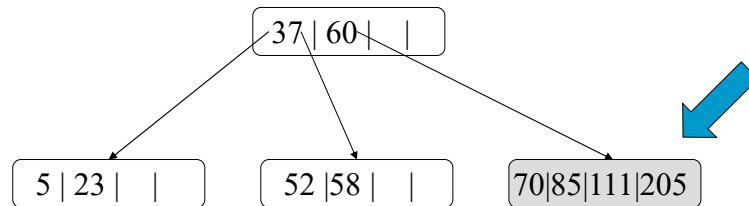
anaeliza.moura@unicap.br

24

## Árvore B

### □ Inserção (Exemplo - cont.)

– Inserir chave 5 (depois)



– Inserir chave 97 ⇐ **Realizar cisão**

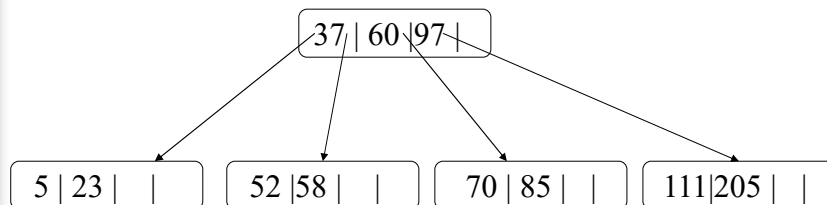
[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

25

## Árvore B

### □ Inserção (Exemplo - cont.)

– Inserir chave 97 (depois)



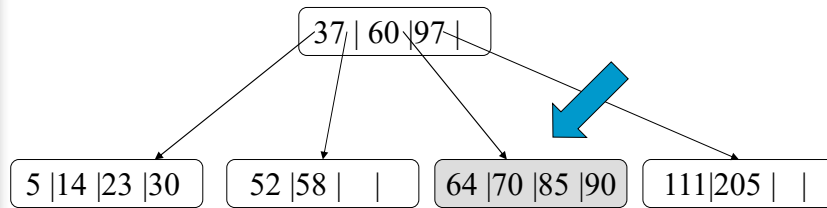
[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

26

## Árvore B

### □ Inserção (Exemplo - cont.)

– Inserir chaves 64, 14, 90, 30



– Inserir chave 75 ⇐ **Realizar cisão**

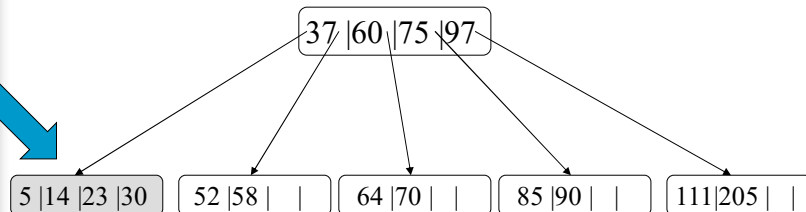
[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

27

## Árvore B

### □ Inserção (Exemplo - cont.)

– Inserir chave 75 (depois)



– Inserir chave 25 ⇐ **Realizar cisão**

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

28

## Árvore B

### □ Inserção

- A inserção da nova entrada no nó pai pode acarretar a necessidade de uma nova cisão;
- A cisão de páginas é propagável, podendo atingir até mesmo a raiz da árvore.
- Neste caso, surge uma nova raiz, o que implica em alteração da altura da árvore.
- Após o processo de inserção, a árvore permanece balanceada.

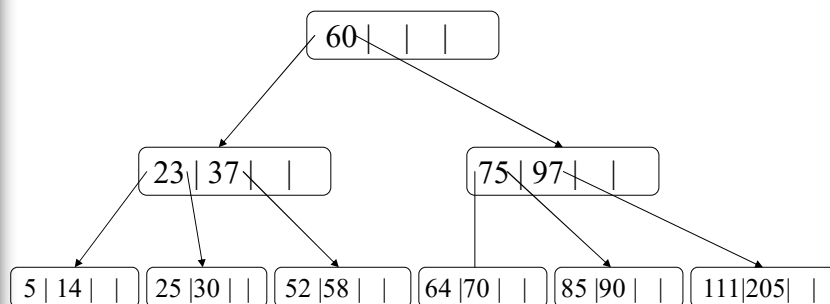
[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

29

## Árvore B

### □ Inserção (Exemplo - cont.)

- Inserir chave 25 (depois)



[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

30

## Árvores B

### □ Consulta

- Verifica se a chave procurada está na raiz;
- Caso não esteja, se a chave procurada for menor que a chave  $C_i$ ,  $1 \leq i \leq N-1$ , então repetir a pesquisa na subárvore  $A_i$ ;
- Se a chave procurada for maior que a chave  $C_{N-1}$ , então repetir a pesquisa na subárvore  $A_N$ ;
- A pesquisa termina quando encontramos a chave ou um apontador  $A_i$  igual a nulo.

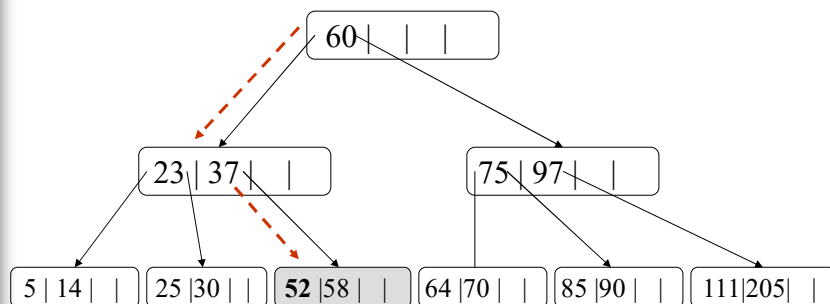
anaeliza.moura@unicap.br

31

## Árvore B

### □ Consulta(Exemplo)

- Procurar chave 52



anaeliza.moura@unicap.br

32





## Árvore B

### □ Consulta (Algoritmo)

- A pesquisa dentro de um nó é seqüencial.
- Se a ordem da árvore for maior que 10, devemos considerar a utilização de pesquisa binária.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

33



## Árvore B

### □ Remoção de uma chave X

- **Caso 1:** A chave **X** não se encontra em uma folha
  - **X** é substituída pela chave **Y**, imediatamente maior;
  - **Y** necessariamente pertence a uma folha.

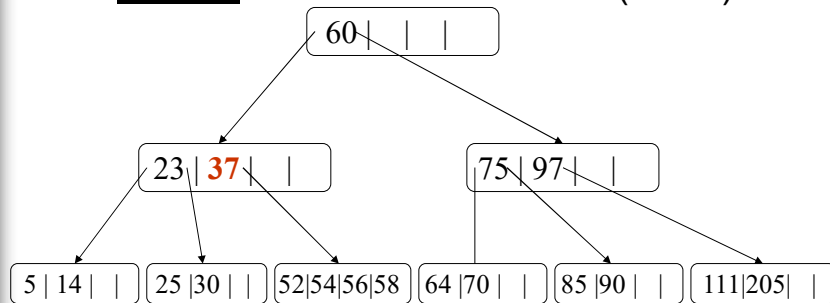
[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

34

## Árvore B

### Remoção (Exemplo)

– **Caso 1:** Remover a chave 37 (antes)



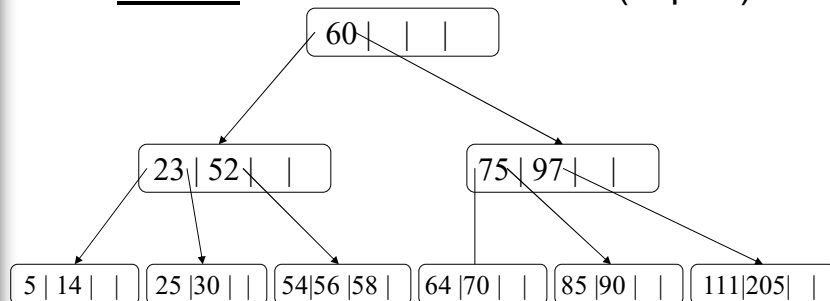
anaeliza.moura@unicap.br

35

## Árvore B

### Remoção (Exemplo)

– **Caso 1:** Remover a chave 37 (depois)



anaeliza.moura@unicap.br

36

## Árvore B

### □ Remoção

- **Caso 2:** A chave X se encontra em uma folha
  - A chave é simplesmente removida.

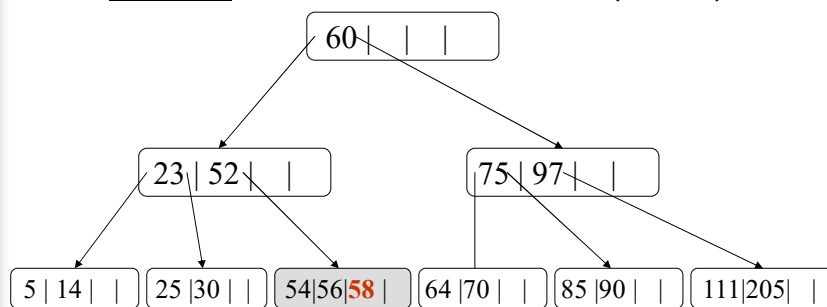
anaeliza.moura@unicap.br

37

## Árvore B

### □ Remoção (Exemplo)

- **Caso 2:** Remover a chave 58 (antes)



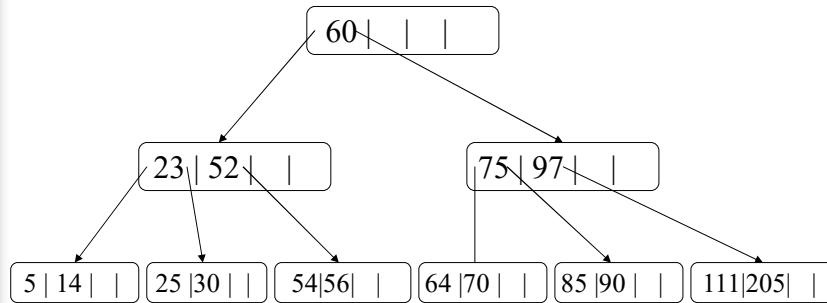
anaeliza.moura@unicap.br

38

## Árvore B

### □ Remoção (Exemplo)

– **Caso 2:** Remover a chave 58 (depois)



[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

39

## Árvore B

### □ Remoção

– Quando uma chave é retirada de um nó folha, o número de chaves restantes pode ser menor que  $M/2$ .

– Tratamentos:

- Concatenação
- Redistribuição

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

40

## Árvore B

### □ Remoção com Concatenação

- Duas páginas **P** e **Q** são chamadas **irmãos adjacentes** se têm o mesmo pai **W** e são apontadas por ponteiros adjacentes em **W**.
- **P** e **Q** podem ser concatenadas se são irmãos adjacentes e juntas possuem menos de **M-1** chaves.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

41

## Árvore B

### □ Remoção com Concatenação

- A concatenação agrupa as entradas de duas páginas em uma só;
- No nó pai deixa de existir uma entrada: aquela da chave que se encontra entre os ponteiros para **P** e **Q**.
- Essa chave passa a fazer parte do nó concatenado e seu ponteiro desaparece.

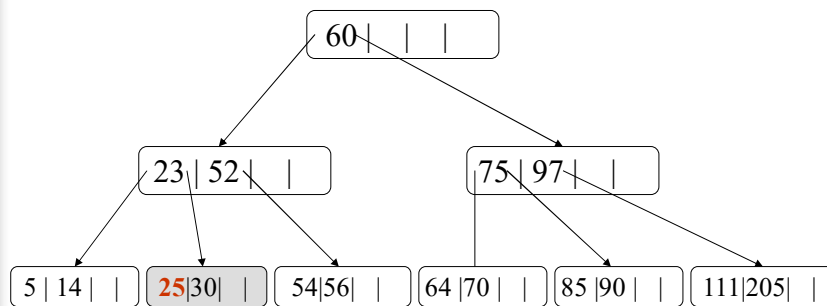
[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)

42

## Árvore B

### Remoção com Concatenação

– Exemplo: Remover a chave 25 (antes)



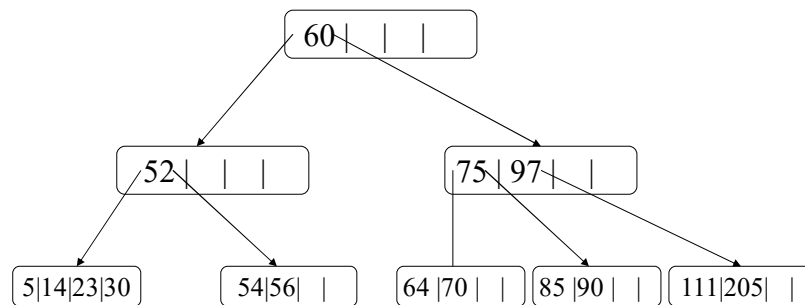
anaeliza.moura@unicap.br

43

## Árvore B

### Remoção com Concatenação

– Exemplo: Remover a chave 25 (depois)



anaeliza.moura@unicap.br

44

## Árvore B

### Remoção com Concatenação

- Como foi retirada uma chave do nó **W**, caso ele passe a ter menos de  **$M/2$**  chaves, o processo se repete;
- Ou seja, a concatenação é um processo propagável;
- Se a propagação atingir a raiz, a árvore diminuirá de altura.

anaeliza.moura@unicap.br

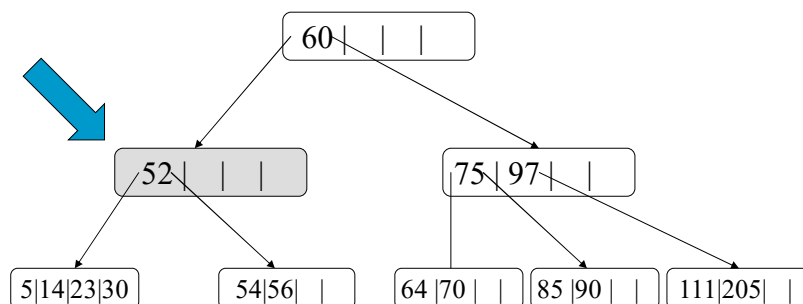
45

## Árvore B

### Remoção com Concatenação

- Exemplo: Remover a chave 25 (cont.)

⇒ **Propagação**



anaeliza.moura@unicap.br

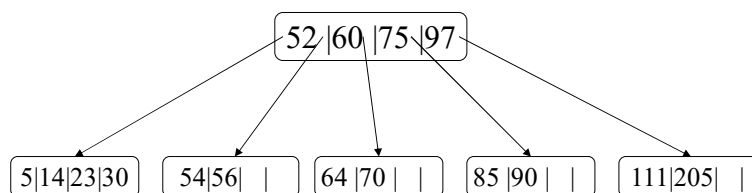
46

## Árvore B

### □ Remoção com Concatenação

– Exemplo: Remover a chave 25 (cont.)

⇒ **Propagação**



anaeliza.moura@unicap.br

47

## Árvore B

### □ Remoção com Redistribuição

– Se a página **P** e seu irmão adjacente **Q** possuem em conjunto **M-1** ou mais chaves, estas podem ser equilibradamente distribuídas:

- Concatena-se **P** e **Q**;
- Efetua-se a cisão da página resultante.

anaeliza.moura@unicap.br

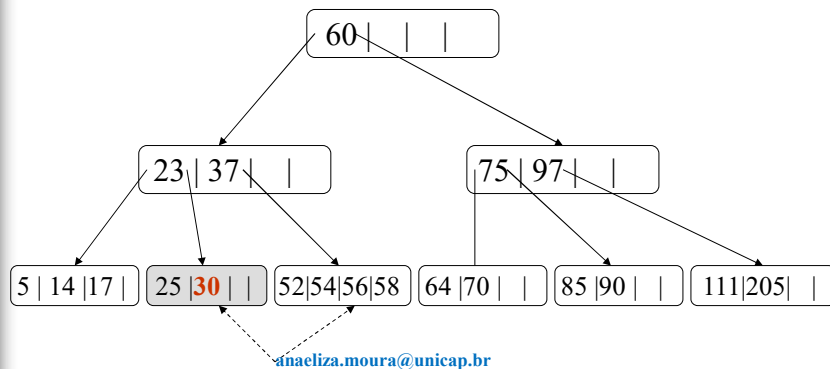
48



## Árvore B

### Remoção com Redistribuição

– Exemplo: Remoção da chave 30 (antes)

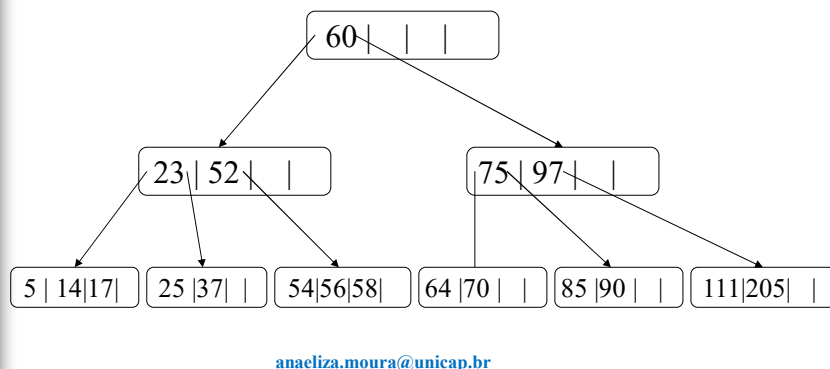


49

## Árvore B

### Remoção com Redistribuição

– Exemplo: Remoção da chave 30 (depois)



50



## Árvore B

- Remoção com Redistribuição
  - A redistribuição não é propagável;
  - A página **W**, pai de **P** e **Q**, é modificada, mas seu número de chaves permanece o mesmo.

[anaeliza.moura@unicap.br](mailto:anaeliza.moura@unicap.br)