

Árvores B- Remoção

Prof.: Leonardo Tórtoro Pereira
leonardop@usp.br

*Material baseado em aulas dos professores: Elaine Parros Machado de Souza, Gustavo Batista, Robson Cordeiro, Moacir Ponti Jr., Maria Cristina Oliveira e Cristina Ciferri.

O que sabemos de Árvores B?

Relembrando...

→ Árvore-B

◆ Inserção

- Algoritmo

◆ Busca

- Complexidade
- Algoritmo

Remoção

Remoção

→ Na inserção:

- ◆ Split garante a manutenção das propriedades da árvore-B

→ Porém

- ◆ Propriedades precisam ser mantidas também durante a remoção de chaves
- ◆ Remoção sempre nas folhas!
- ◆ Vários casos a serem analisados

Remoção

→ Caso 1:

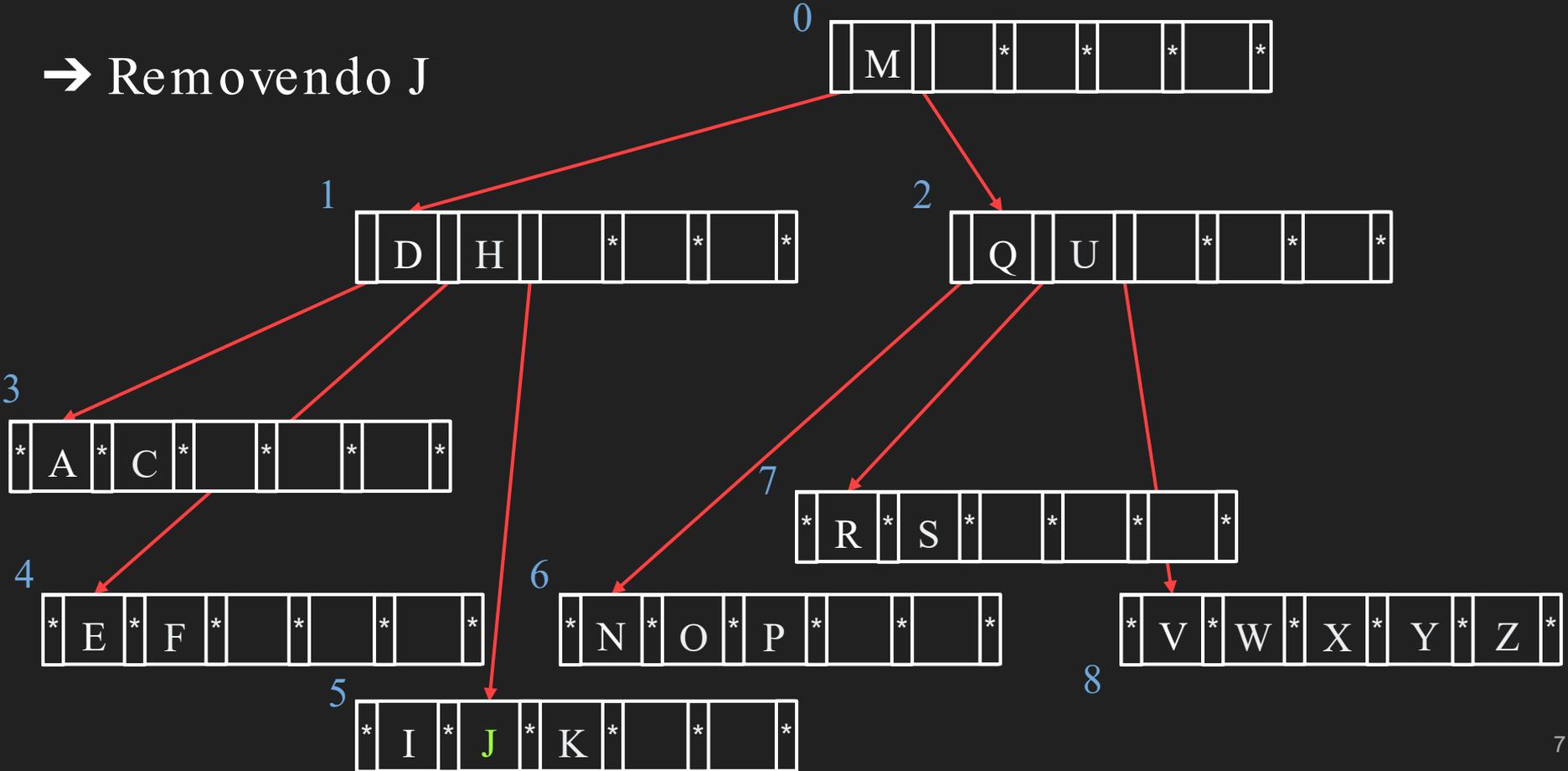
- ◆ Remoção de uma chave em uma página folha mantendo ocupação mínima da árvore
 - Número mínimo de chaves na página

→ Solução:

- ◆ Chave é retirada e demais chaves da página são reorganizadas

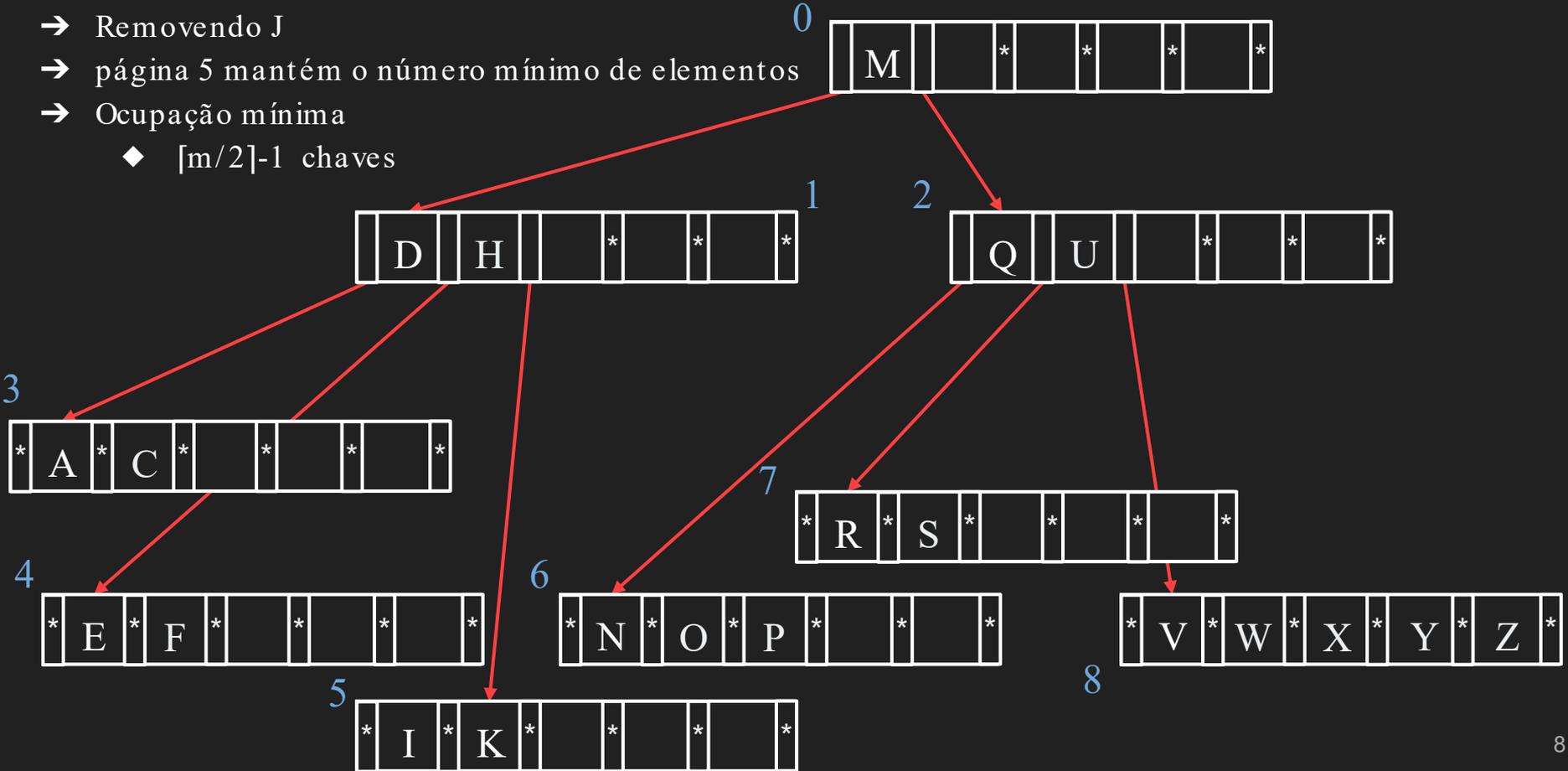
Remoção: Caso 1

→ Removendo J



Remoção: Caso 1

- Removendo J
- página 5 mantém o número mínimo de elementos
- Ocupação mínima
 - ◆ $\lfloor m/2 \rfloor - 1$ chaves



Remoção

→ Caso 2:

- ◆ Removendo uma chave de uma página não folha

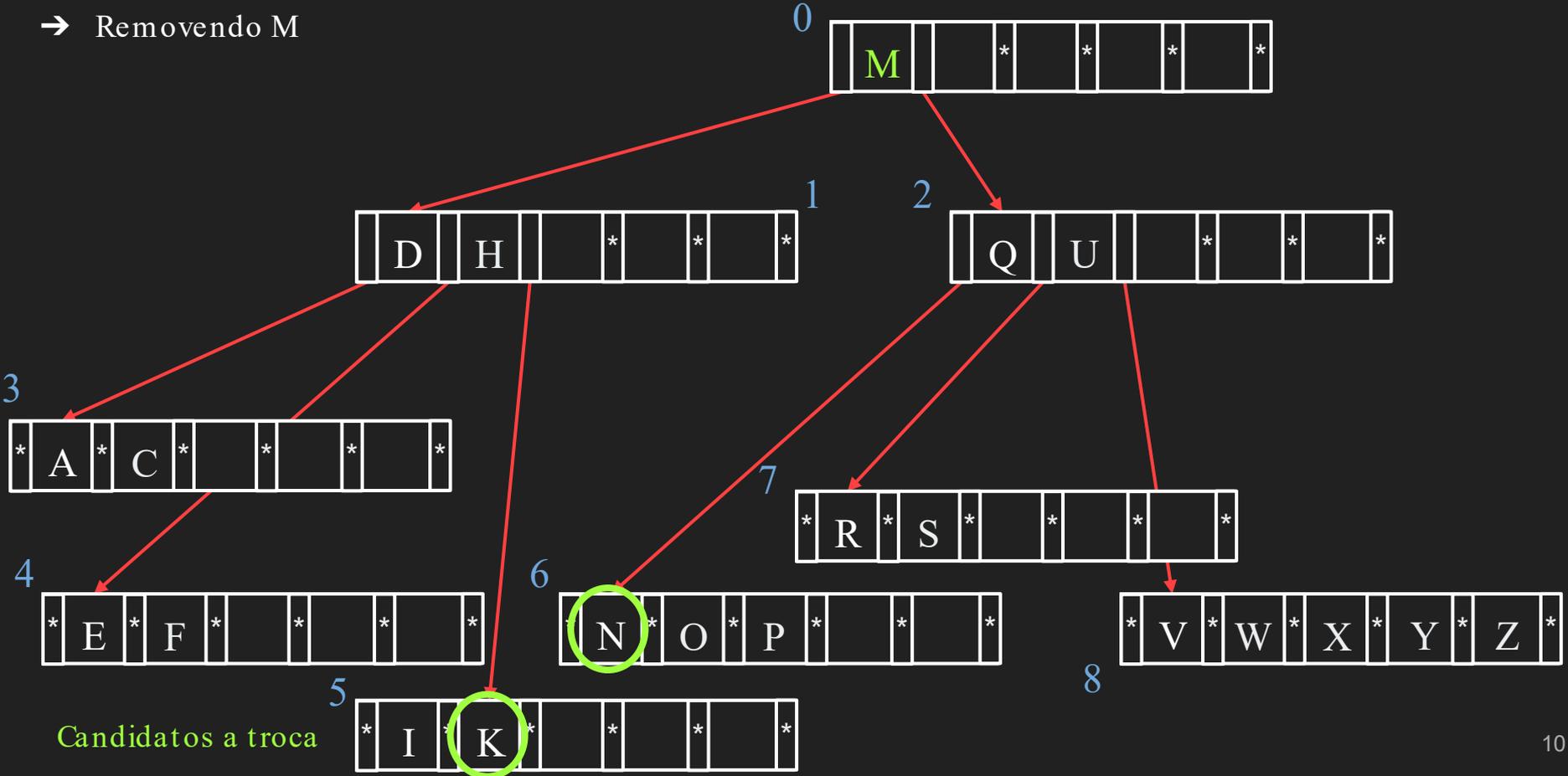
→ Solução: sempre remover de páginas folha

- ◆ Trocar a chave com sua sucessora imediata (ou com a predecessora imediata) que está numa folha

- ◆ Remover a chave da folha

Remoção: Caso 2

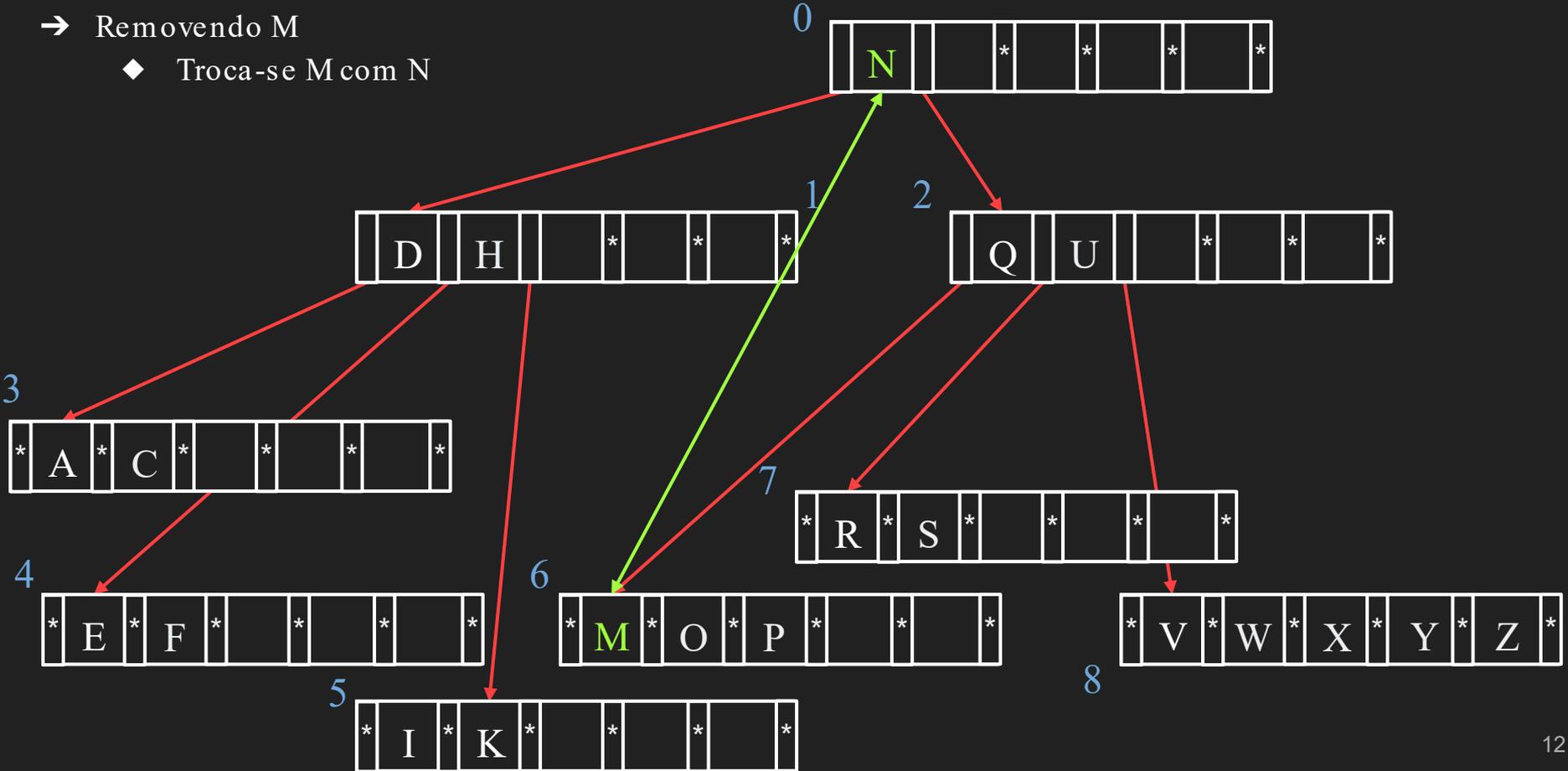
→ Removendo M



Remoção: Caso 2

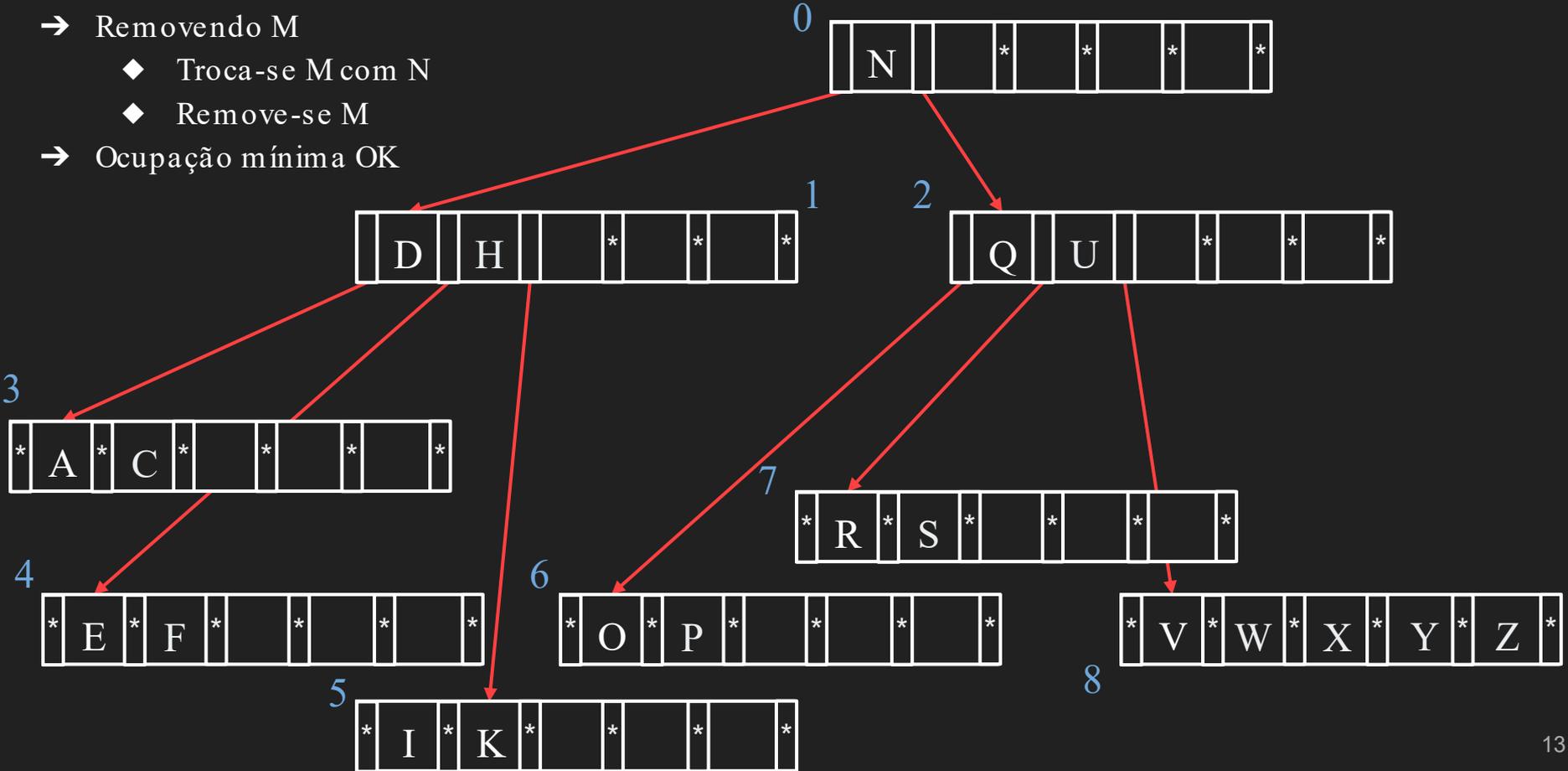
→ Removendo M

◆ Troca-se M com N



Remoção: Caso 2

- Removendo M
 - ◆ Troca-se M com N
 - ◆ Remove-se M
- Ocupação mínima OK



Remoção

→ Caso 3:

- ◆ Remoção causa underflow na página
 - Número de chaves abaixo da ocupação mínima

Remoção

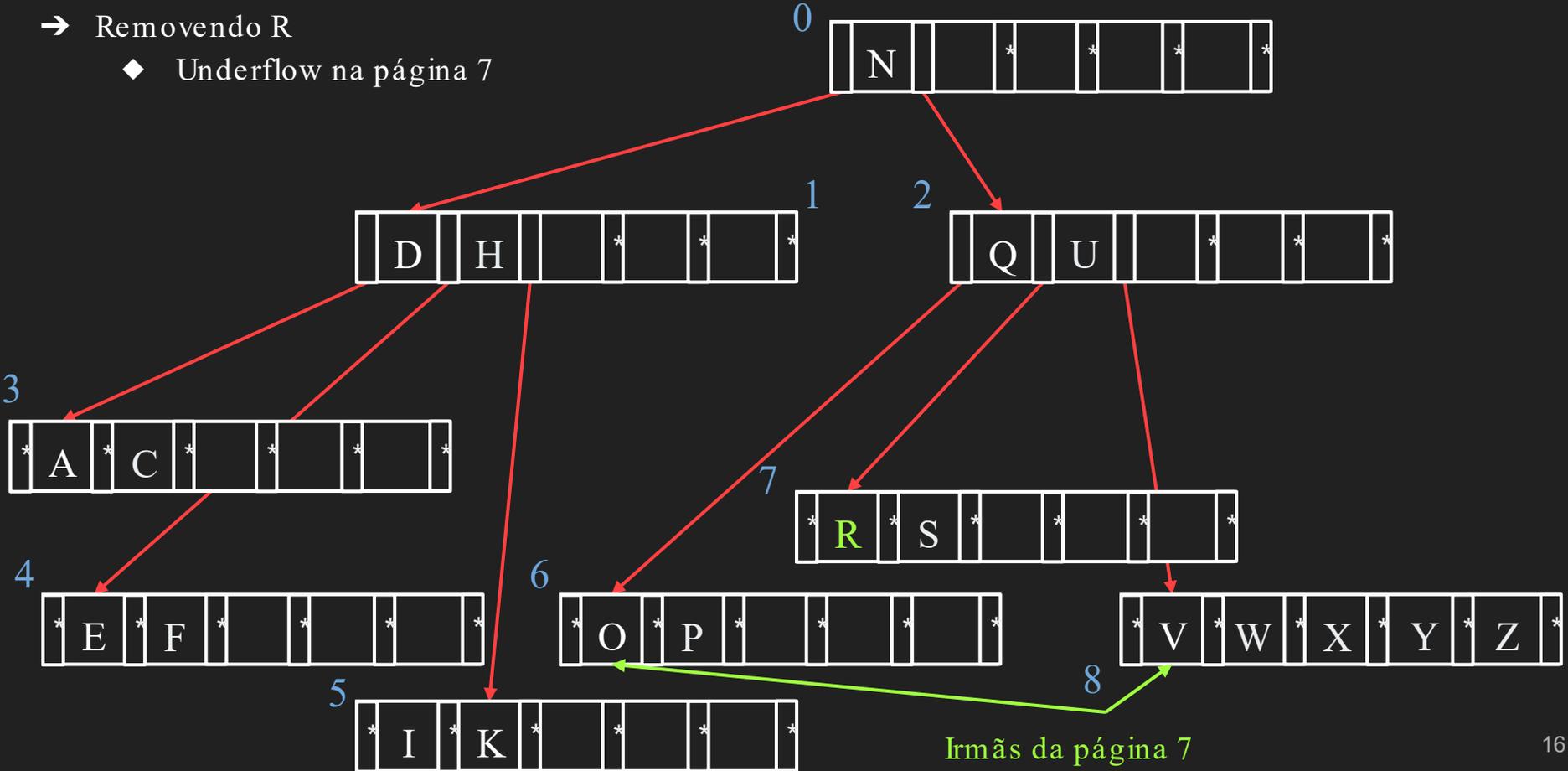
→ Solução: redistribuição

- ◆ Procura-se uma página irmã (com mesmo nó pai e chave separadora em comum) que contenha mais chaves do que o mínimo
 - Se existir, redistribui-se as chaves entre essas páginas
- ◆ Redistribuição provoca uma alteração na chave separadora que está no nó pai

Remoção: Caso 2

→ Removendo R

◆ Underflow na página 7



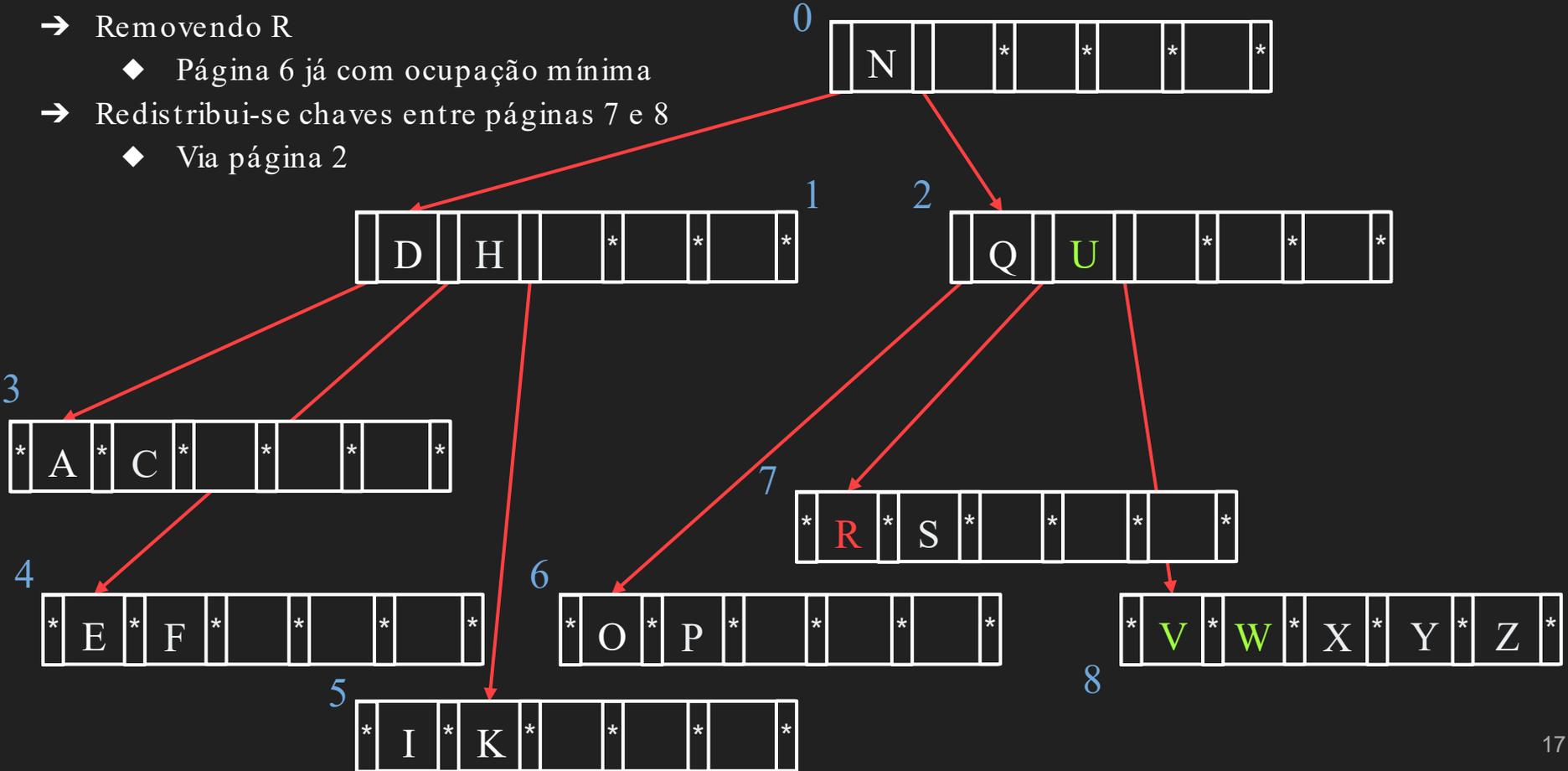
Remoção: Caso 2

→ Removendo R

◆ Página 6 já com ocupação mínima

→ Redistribui-se chaves entre páginas 7 e 8

◆ Via página 2



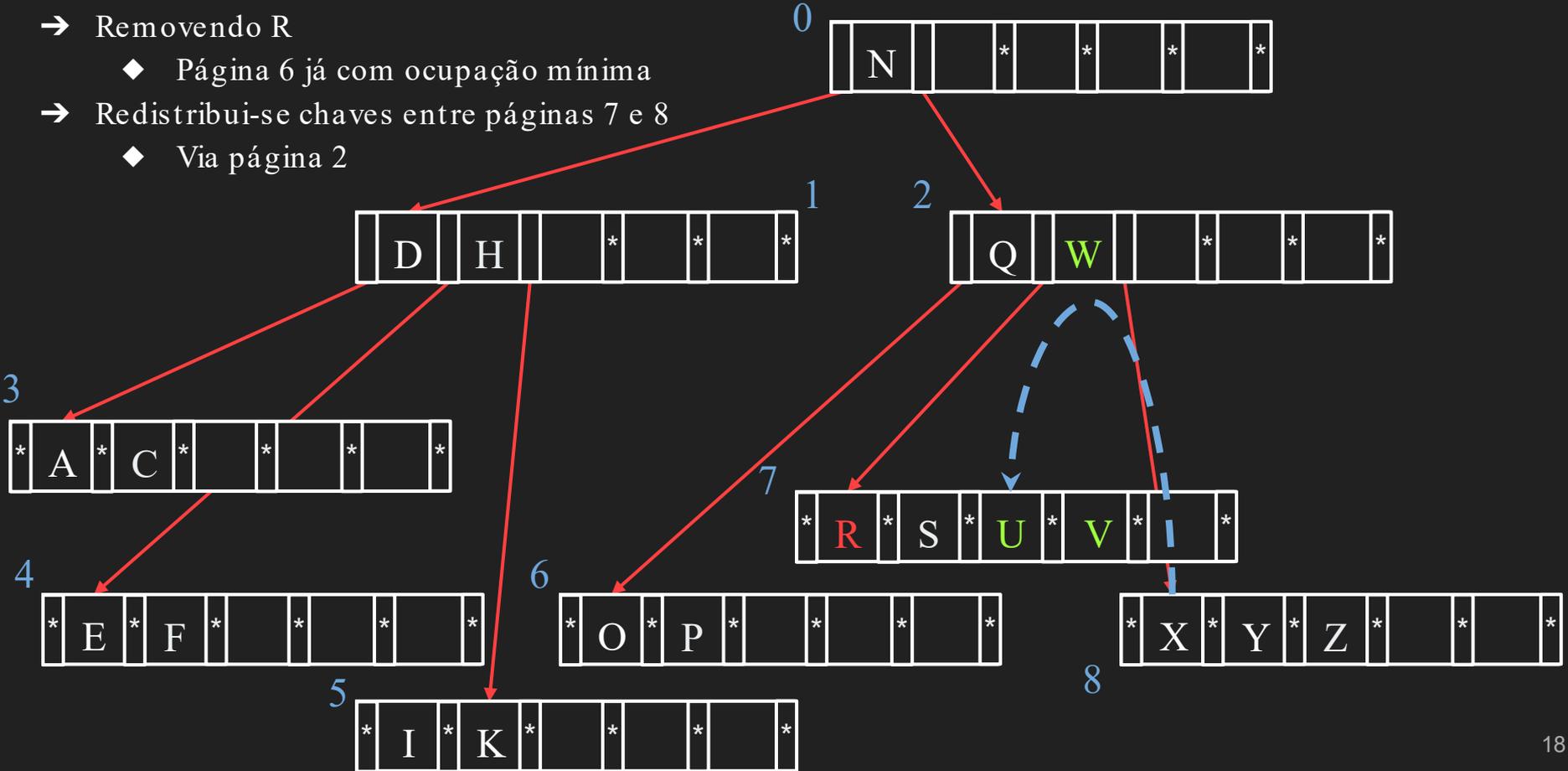
Remoção: Caso 2

→ Removendo R

◆ Página 6 já com ocupação mínima

→ Redistribui-se chaves entre páginas 7 e 8

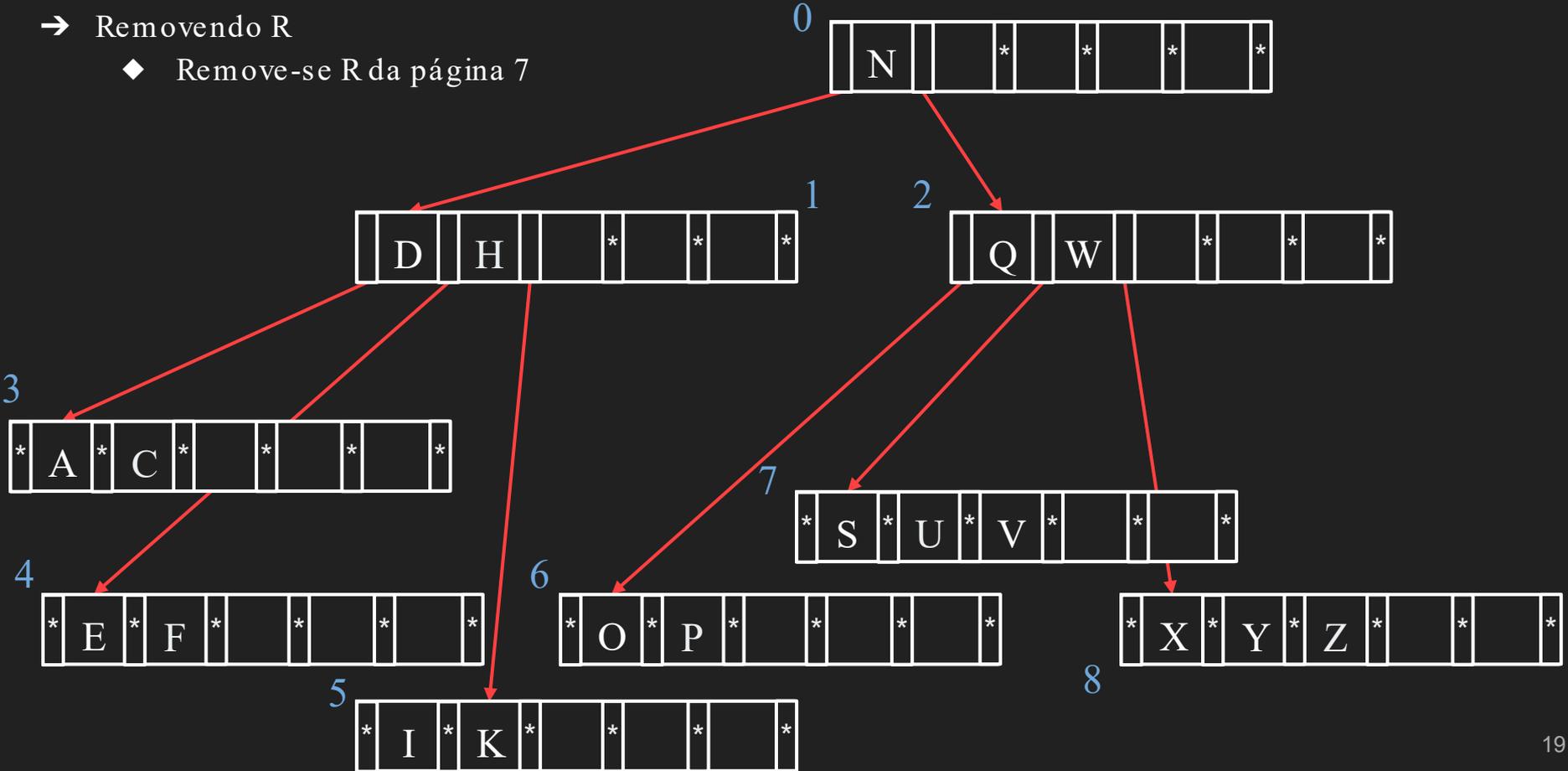
◆ Via página 2



Remoção: Caso 2

→ Removendo R

◆ Remove-se R da página 7



Remoção

→ Caso 3:

◆ Redistribuição:

- Ideia inovadora...
- Efeito local na árvore
 - Não se propaga para nós superiores

Remoção

→ Caso 3:

◆ Redistribuição:

● Necessário

- Mover apenas 1 chave para a página com underflow para restabelecer as propriedades da árvore-B

Remoção

→ Caso 3:

◆ Redistribuição:

- Estratégia usual

- Redistribuir as chaves de maneira equilibrada entre as páginas:
- “balanceamento” dos espaços disponíveis

Remoção

→ Caso 4:

- ◆ Ocorre underflow e a redistribuição não pode ser aplicada
 - Não há chaves suficientes para dividir entre duas páginas irmãs

Remoção

→ Caso 4:

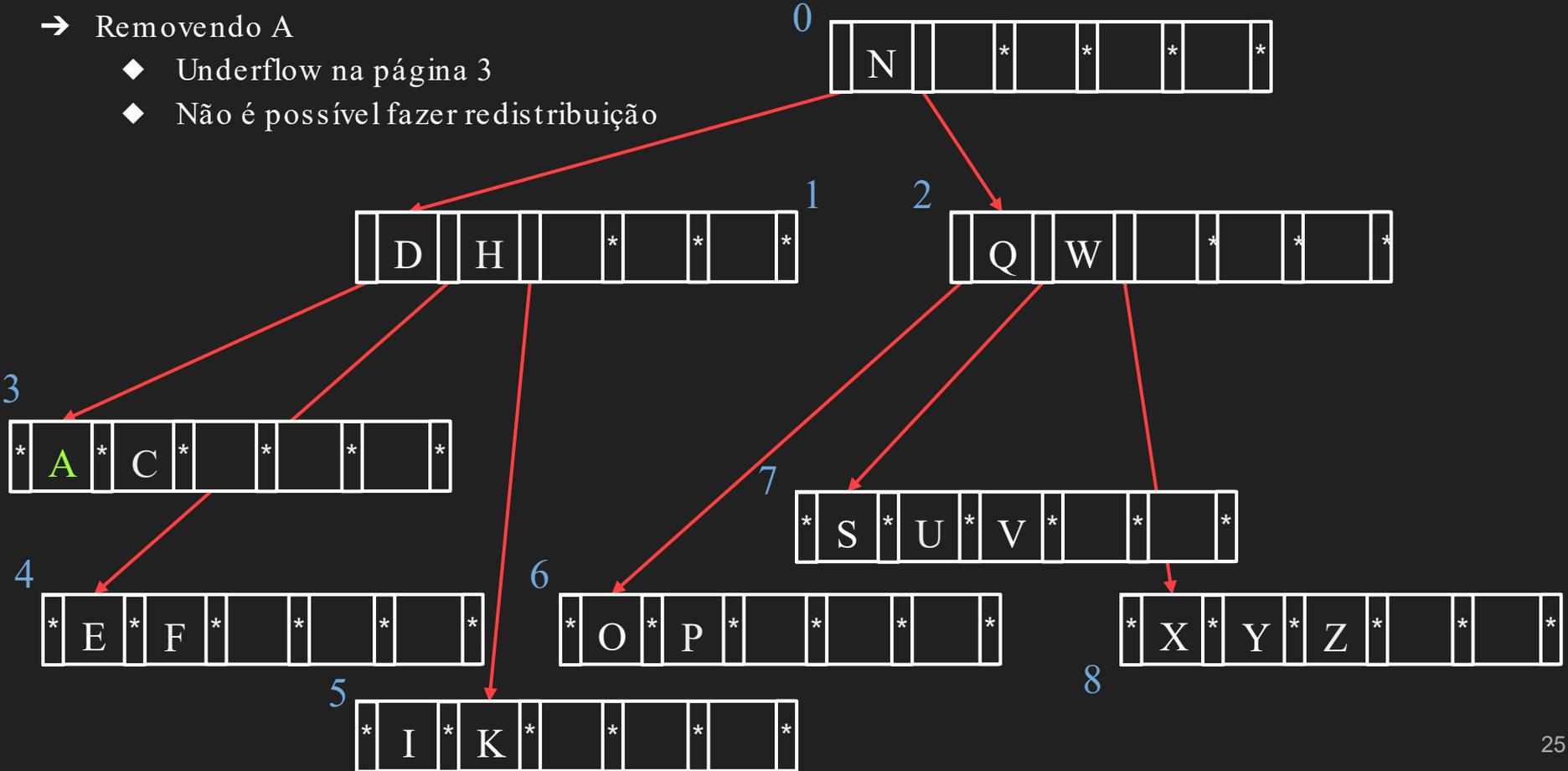
◆ Solução: concatenação

- Combina-se o conteúdo das duas páginas com a chave separadora da página pai para formar uma única página
- Pode ocorrer underflow da página pai
 - Propagação de underflow (caso 5)

Remoção: Caso 4

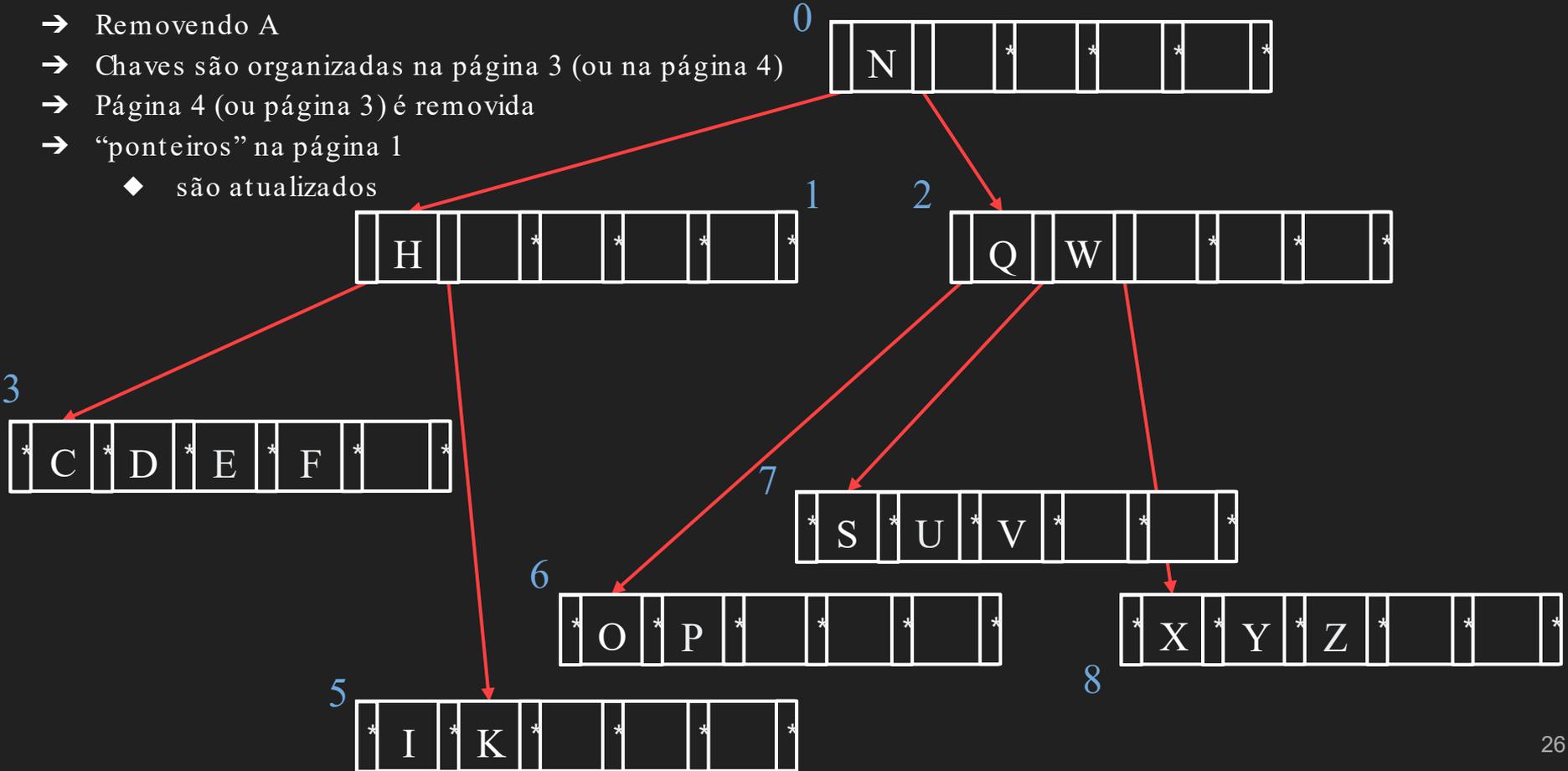
→ Removendo A

- ◆ Underflow na página 3
- ◆ Não é possível fazer redistribuição



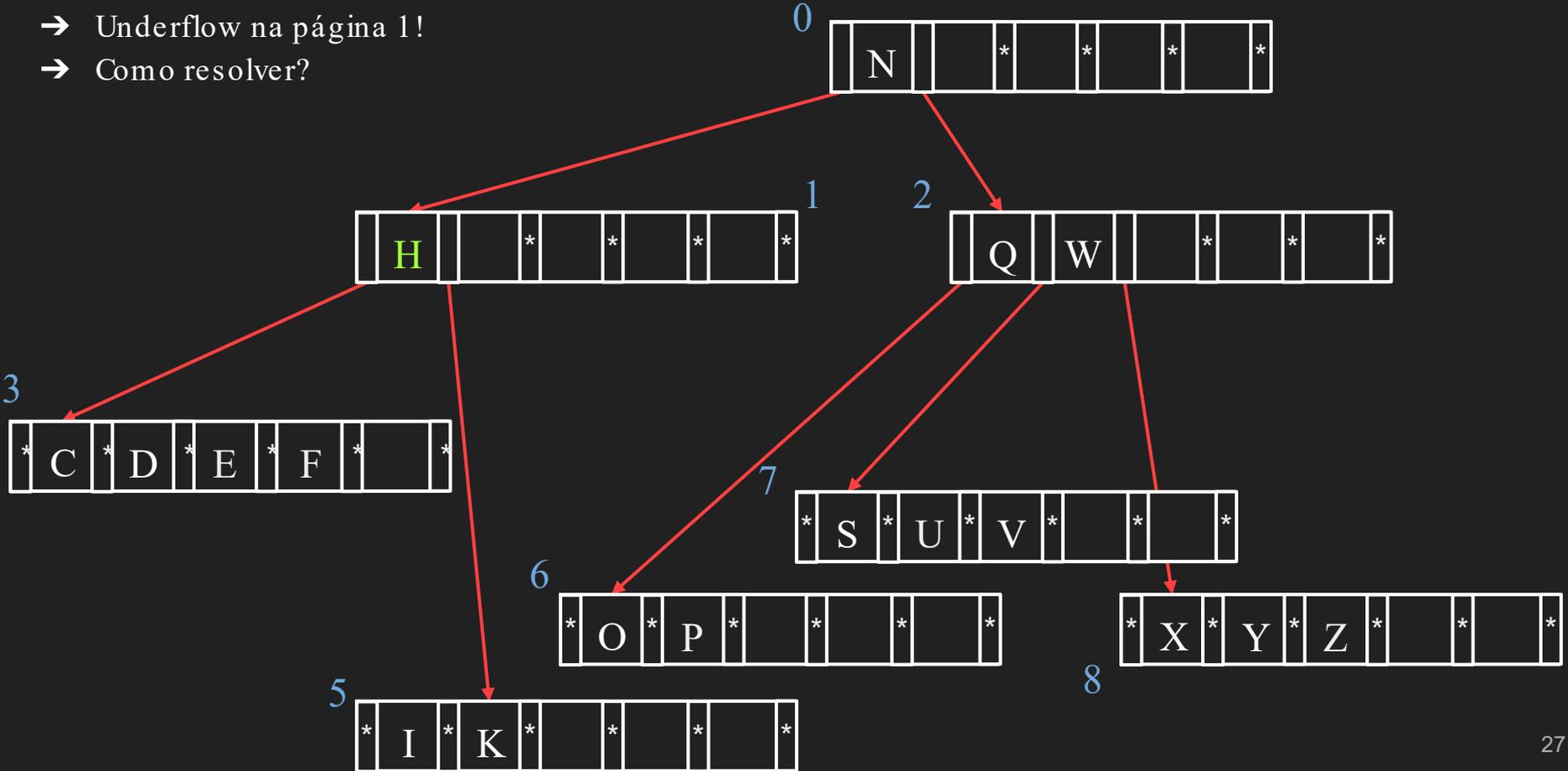
Remoção: Caso 4

- Removendo A
- Chaves são organizadas na página 3 (ou na página 4)
- Página 4 (ou página 3) é removida
- “ponteiros” na página 1
 - ◆ são atualizados



Remoção: Caso 4

- Underflow na página 1!
- Como resolver?

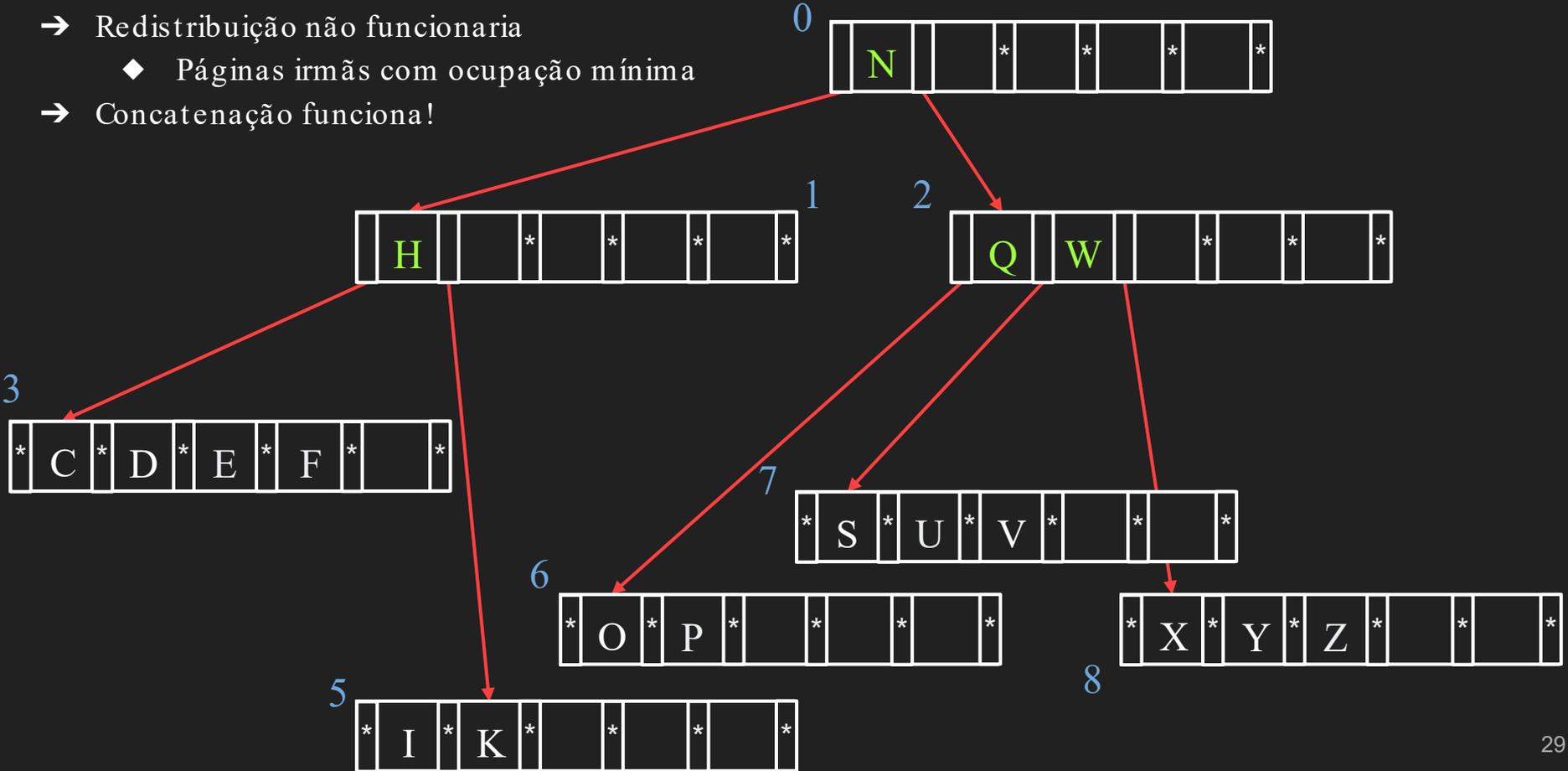


Remoção

- Caso 5: Underflow da página pai
 - ◆ Propagação de Underflow
 - ◆ Solução: Redistribuição ou Concatenação

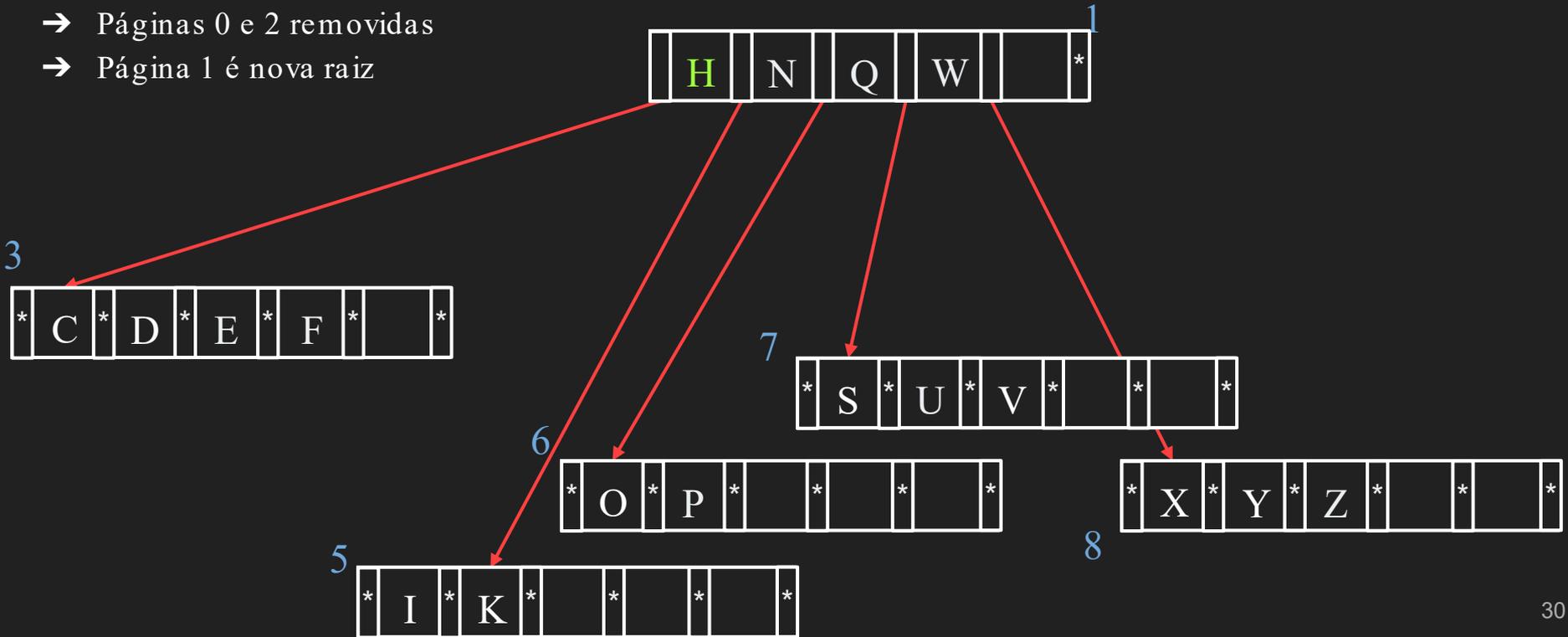
Remoção: Caso 4

- Redistribuição não funcionaria
 - ◆ Páginas irmãs com ocupação mínima
- Concatenação funciona!



Remoção: Caso 4

- Concatenação
- Organização das chaves na página 1
 - ◆ Atualiza ponteiros
- Páginas 0 e 2 removidas
- Página 1 é nova raiz



Remoção

→ Caso 6: diminuição da altura da árvore

- ◆ Ocorre quando o nó raiz tem uma única chave e aplica-se a concatenação de seus nós filhos
- ◆ Como ocorreu no exemplo anterior...
- ◆ Requer atualização da identificação do nó raiz no cabeçalho do arquivo

Resumo

Remoção - Resumo

1. Se a chave não estiver numa folha, troque-a com sua sucessora (ou antecessora)
2. Remova a chave da folha
3. Se a folha continuar com o número mínimo de chaves,
FIM

Remoção - Resumo

4. Senão (*underflow* na página)
 - a. se uma das páginas irmãs diretas (esquerda ou direita) tiver mais que o mínimo de chaves, aplique redistribuição e FIM
 - b. senão
 - i. concatene a pág. com uma das irmãs e a chave separadora do nó pai
 - ii. se nó pai for raiz e sua última chave foi rebaixada, elimine a raiz e FIM

Remoção - Resumo

4. Senão (*underflow*)

b. senão

- i. senão, se nó pai continuar com o mínimo de chaves, FIM
- ii. senão (*underflow* no pai), volte ao item 4.a para o nó pai

Desempenho Árvores-B

(Medido em números de acesso a disco)

Desempenho Árvores-B

→ Busca de chaves...

→ No pior caso, a altura é dada por:

$$d \leq 1 + \log_{\lfloor m/2 \rfloor} \lceil (N+1)/2 \rceil$$

→ Logo...

◆ Altura é $O(\log_{\lfloor m/2 \rfloor} N)$

◆ Busca no pior caso

• $O(\log_{\lfloor m/2 \rfloor} N)$

Desempenho Árvores-B

- Inserção de chaves...
- Toda inserção realiza busca
 - ◆ $O(\log_{\lfloor m/2 \rfloor} N)$
- Além disso pode realizar split
 - ◆ Cada split opera sobre um número fixo de páginas e é portanto constante
 - $O(1)$

Desempenho Árvores-B

- Inserção de chaves...
- No pior caso, overflows se propagam até a raiz e são realizados splits em tempo constante
 - ◆ $O(\log_{[m/2]} N)$ splits com $O(1)$ acessos cada
- Logo... inserção no pior caso
 - ◆ $O(\log_{[m/2]} N)$

Desempenho Árvores-B

- Remoção de chaves...
- Toda remoção realiza busca
 - ◆ $O(\log_{\lfloor m/2 \rfloor} N)$
- Além disso, pode realizar concatenação/redistribuição
- Cada concatenação/redistribuição opera sobre um número fixo de páginas e é portanto constante
 - ◆ $O(1)$

Desempenho Árvores-B

- Remoção de chaves...
- No pior caso, underflows se propagam até a raiz e são realizadas concatenações/redistribuições em tempo constante
 - ◆ $O(\log_{\lfloor m/2 \rfloor} N)$ operações com $O(1)$ acessos cada
- Logo, remoção no pior caso
 - ◆ $O(\log_{\lfloor m/2 \rfloor} N)$

Referências

- M. J. Folk and B. Zoellick, File Structures: A Conceptual Toolkit, Addison Wesley, 1987.