Algoritmi e Strutture Dati

Alberi bilanciati

Sabrina De Capitani di Vimercati

Università degli Studi di Milano

Ricordate...

- ullet Gli alberi binari di ricerca permettono di trovare un elemento in tempo O(h)
 - nel caso peggiore h=O(n)
 - nel caso di albero bilanciato h=O(log n)
- Nozione intuitiva di bilanciamento: tutti i rami di un albero hanno approssimativamente la stessa lunghezza
- Caso ideale per un albero k-ario:
 - ciascun nodo ha 0 o k figli
 - la lunghezza di due rami qualsiasi differisce di al più una unità

Bilanciamento perfetto

- Un albero binario perfettamente bilanciato di n nodi ha altezza [log n] + 1
- \bullet Se ogni nodo ha 0 oppure 2 figli allora si ha $n_{_f}=n_{_i}$ +1 che dove $n_{_f}$ è il numero di foglie ed $n_{_i}$ il numero di nodi interni
- Le foglie sono circa il 50% dei nodi
- Generalizzando, per alberi di arità k si ha $n_f = (k-1)n_i + 1$
- Il costo delle operazioni di ricerca/inserimento/cancellazione è pertanto O(log n)
- Ripetuti inserimenti e cancellazioni possono distruggere il bilanciamento con un conseguente degrado delle prestazioni

2

Fattore di bilanciamento

- •Il fattore di bilanciamento $\beta(v)$ di un nodo v è la massima differenza di altezza fra i sottoalberi di v
 - es.,: l'albero perfetto è un albero dove $\beta(v)=0$ per ogni nodo v
- •In questo caso si parla di bilanciamento in altezza

Alberi bilanciati

Alberi AVL (Adel'son-Vel'skii e Landis, 1962)

- $\beta(v) \le 1$ per ogni nodo v
- Bilanciamento ottenuto tramite rotazioni

Alberi Rosso-Neri (Bayer, 1972)

Alberi 2-3 (Hopcroft, 1983)

- $\beta(v) = 0$ per ogni nodo v
- Bilanciamento ottenuto tramite merge/split, grado variabile

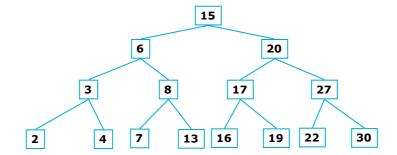
B-Alberi (Bayer, McCreight, 1972)

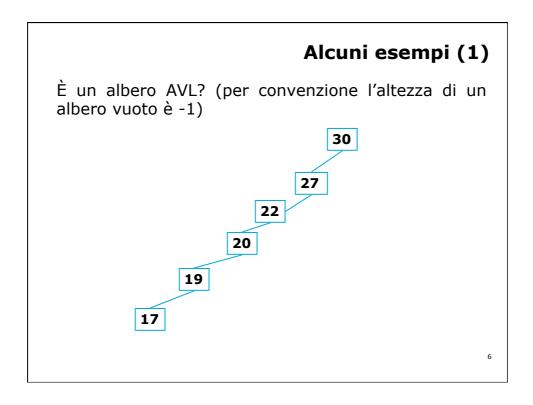
- $\beta(v) = 0$ per ogni nodo v
- Usati per strutture in memoria secondaria

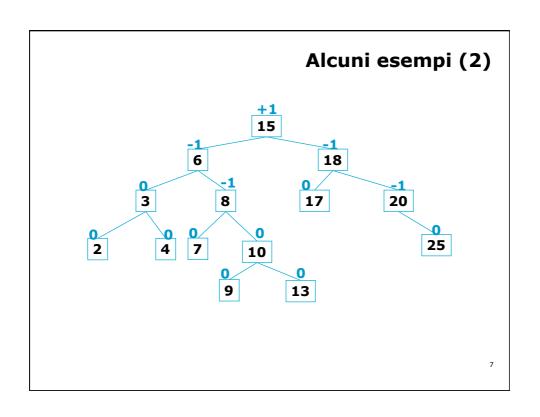
4

Alberi AVL

- Alberi binari di ricerca bilanciati in altezza
- $\bullet \beta(v)$ = altezza del sottoalbero sinistro di v altezza del sottoalbero destro di v







Operazioni su alberi AVL

- •La ricerca viene effettuata applicando la stessa procedura usata per gli alberi binari di ricerca
- •Inserimenti e cancellazioni possono sbilanciare l'albero
- •Il bilanciamento viene preservato attraverso opportune rotazioni

8

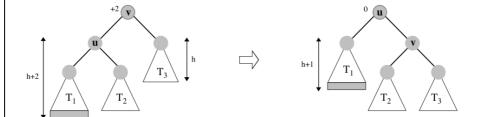
Ribilanciamento

- Si ricalcolano i fattori di bilanciamento solo nel ramo interessato dall'inserimento/cancellazione (gli altri fattori non possono mutare), dal basso verso l'alto
- •Se appare un fattore di bilanciamento pari a +/- 2 occorre ribilanciare per mezzo di rotazioni
- Si distinguono i seguenti casi:
 - DD inserimento nel sottoalbero destro di un figlio destro
 - SD inserimento nel sottoalbero destro di un figlio sinistro
 - DS inserimento nel sottoalbero sinistro di un figlio destro
 - SS inserimento nel sottoalbero sinistro di un figlio sinistro

10

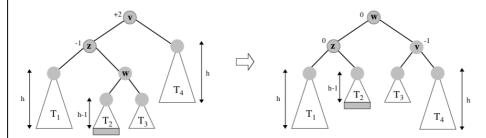
Rotazione SS

- Applicare una rotazione semplice verso destra su v
- •L'altezza dell'albero coinvolto nella rotazione passa da h+3 a h+2



Rotazione SD

- Applicare due rotazioni semplici: una sul figlio del nodo critico, l'altra sul nodo critico
- L'altezza dell'albero coinvolto nella rotazione passa da h+3 a h+2



12

Inserimento in un albero AVL (1)

- Si trova (come per alberi binari di ricerca) la posizione in cui inserire il nuovo elemento
 - durante tale operazione memorizzato il puntatore al più basso nodo v il cui fattore di bilanciamento è -1 o +1
- Viene effettuato l'inserimento
- Vengono aggiornati i fattori di bilanciamento dei nodi lungo il percorso tra il nodo appena inserito e v
 - tutti i nodi lungo questo percorso avevano un fattore di bilanciamento 0 e vengono portati a +1 o -1
 - l'unico nodo che potrebbe registrare uno sbilanciamento è proprio v (nodo critico)

Inserimento in un albero AVL (2)

- Se il fattore di bilanciamento di v è +1 (-1) ed il nodo è stato inserito nel sottoalbero destro (sinistro), il sottoalbero con radice v è automaticamente bilanciato
 - in caso contrario, il sottoalbero è sbilanciato e richiede un ribilanciamento
- Viene determinata ed effettuata la rotazione necessaria
- Tutti questi passi, rotazioni comprese, interessano solo i nodi lungo il cammino dalla radice alla nuova foglia inserita

14

Cancellazione in un albero AVL

- •Si effettua la cancellazione di un elemento come descritto dall'algoritmo Tree-Delete
- Si ricalcolano i fattori di bilanciamento
 - osserviamo che i soli fattori di bilanciamento che possono cambiare sono quelli dei nodi lungo il cammino dalla radice al nodo eliminato e che questi possono essere ricalcolati risalendo l'albero dal basso verso l'alto
- •Eseguiamo una rotazione per ogni nodo il cui fattore di bilanciamento è ±2

Complessità operazioni su alberi AVL

Tutte le operazioni hanno costo $O(\log n)$ poiché l'altezza dell'albero è $O(\log n)$ e ciascuna rotazione richiede solo tempo costante

16

Alberi rosso-neri (1)

- •Gli alberi rosso-neri sono alberi binari di ricerca in cui le operazioni Tree-Insert e Tree-Delete sono opportunamente modificate in modo tale da garantire un'altezza dell'albero h = O(log n)
- A tale scopo si aggiunge un bit per ogni nodo: il colore che può essere rosso o nero
- Sono alberi binari di ricerca che soddisfano delle proprietà ulteriori

Alberi rosso-neri (2)

- 1. La radice è nera;
- 2. I nodi nil sono neri;
- 3. Se un nodo è rosso, allora entrambi i suoi figli sono neri;
- 4. Ogni percorso da un nodo interno ad una foglia ha lo stesso numero di nodi neri

18

Rappresentazione alberi rosso-neri

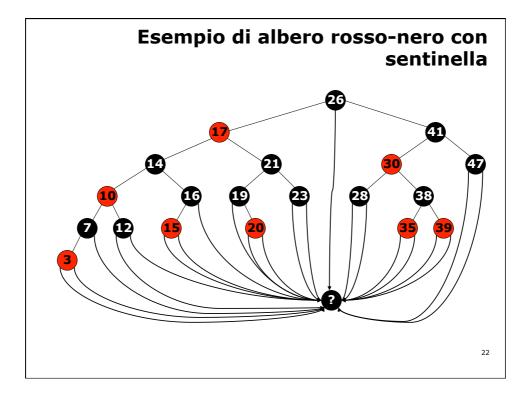
- •Ogni nodo mantiene:
 - parent: puntatore al genitore
 - left, right: puntatori ai figli sinistro/destro
 - color: colore
 - key, data: chiave e dati
- •Nodo nil
 - nodo sentinella che evita di trattare diversamente i puntatori ai nodi dai puntatori nil
 - al posto di un puntatore nil, si usa un puntatore ad un nodo nil
 - ne esiste solo uno, per risparmiare memoria
 - nodo con figli nil \rightarrow foglia nell'albero rosso-nero corrispondente

Alcune definizioni

- •Il numero di nodi neri lungo ogni percorso da un nodo v (escluso) ad una foglia è detto altezza nera di v, indicato bh(v)
 - ben definito perché tutti i percorsi hanno lo stesso numero di nodi neri (regola 4)
- •L'altezza nera di un albero RB è l'altezza nera della sua radice

20

Esempio di albero rosso-nero senza sentinella



Altezza massima di un albero rosso-nero

Un albero rosso-nero con n nodi interni ha altezza $h \le 2 \log_2(n+1)$

Perché?

- Il sottoalbero con radice in x contiene almeno 2^{bh(x)} 1 nodi interni (vediamo perché)
 - se altezza di x è 0 allora x è foglia ed il sottoalbero di cui x è radice contiene almeno $2^{bh(x)}$ $1=2^0$ 1=0 nodi interni
 - per il passo induttivo, sia x un nodo interno con due figli con altezza nera pari a bh(x) o bh(x)-1 a secondo del loro colore
 - Per ipotesi induttiva ogni figlio ha almeno $2^{bh(x)-1}$ 1 e quindi x ha almeno $2^{bh(x)}$ 1 nodi interni
- Se h è l'altezza dell'albero, per la proprietà 3 almeno metà dei nodi in qualsiasi percorso dalla radice a una foglia deve essere nera
 - L'altezza nera della radice deve essere almeno h/2 e quindi n $\geq 2^{h/2}$ 1 ovvero h ≤ 2 log₂(n+1)

Costo operazioni

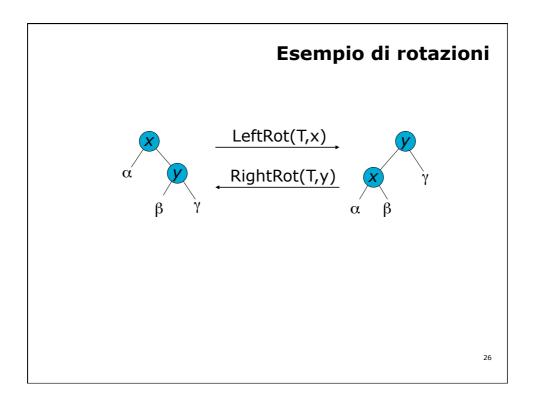
- Su di un albero rosso-nero con n nodi interni le operazioni Search, Minimum, Maximum, Successor e Predecessor richiedono tempo O(log n)
- •Le operazioni Insert e Delete su di un albero rossonero richiedono tempo O(log n) ma siccome esse modificano l'albero possono violare le proprietà degli alberi rosso-neri ed in tal caso occorre ripristinare tali proprietà

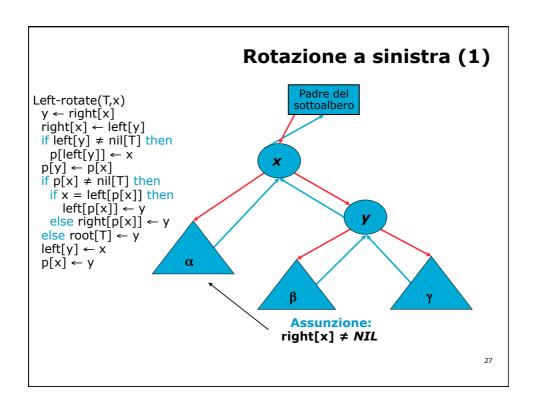
24

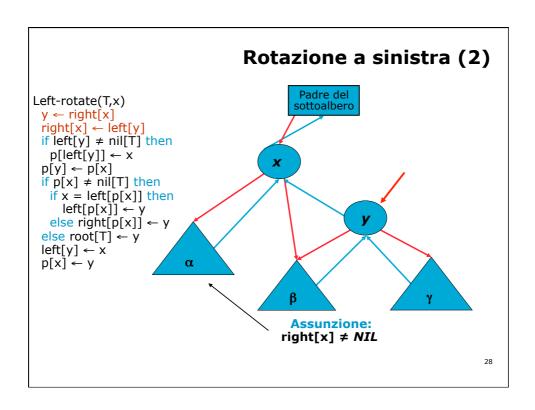
Rotazioni

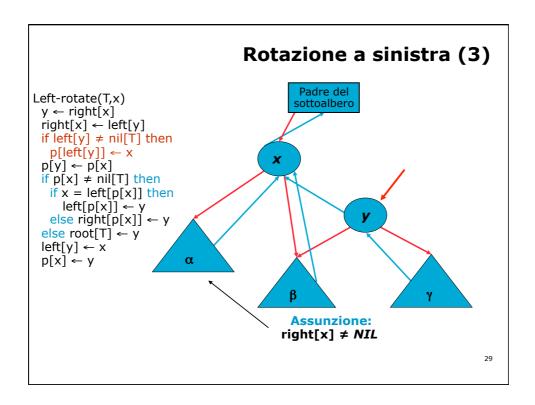
Per ripristinare le proprietà degli alberi rosso-neri si deve:

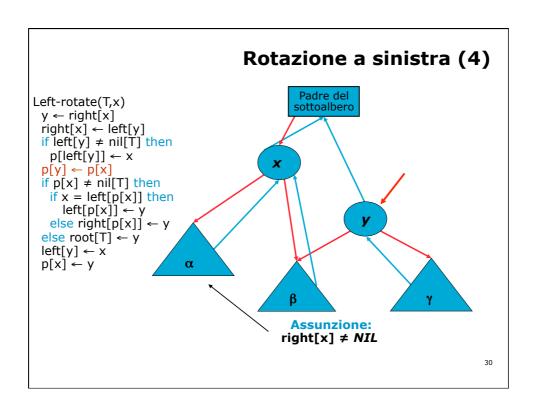
- modificando i colori nella zona della violazione operando dei ri-bilanciamenti dell'albero tramite rotazioni:
 - rotazione destra
 - rotazione sinistra

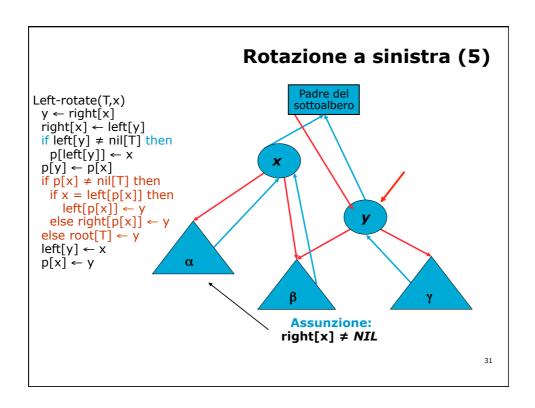


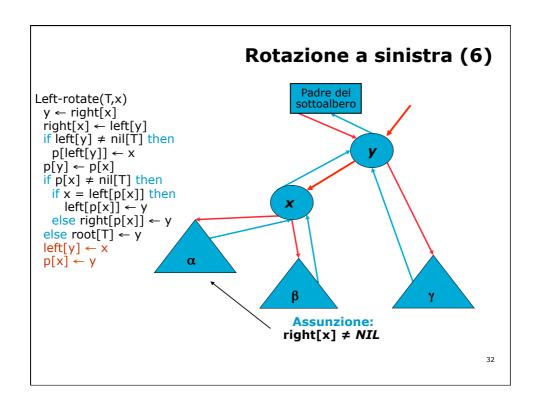












Rotazione a destra

```
\begin{aligned} & \text{Right-Rotate}(T,y) \\ & x \leftarrow \text{left}[y] \\ & \text{left}[y] \leftarrow \text{right}[x] \\ & p[\text{right}[x]] \leftarrow y \quad \triangleright \text{right}[x] \text{ può essere nil}[T] \\ & p[x] \leftarrow p[y] \\ & \text{if } p[y] = \text{nil}[T] \text{ then} \\ & \text{root}[T] \leftarrow x \\ & \text{else if } y = \text{left}[p[y]] \text{ then} \\ & \text{left}[p[y]] \leftarrow x \\ & \text{else} \\ & \text{right}[p[y]] \leftarrow x \\ & p[y] \leftarrow x \\ & \text{right}[x] \leftarrow y \end{aligned}
```

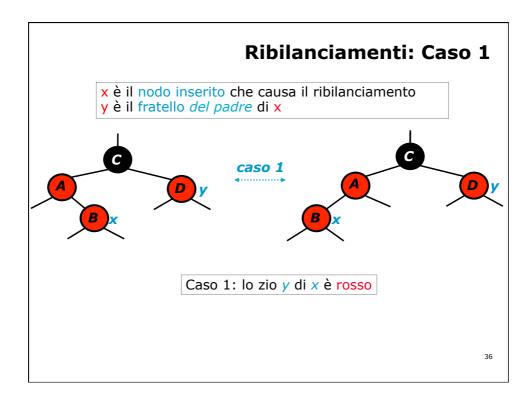
Inserimento

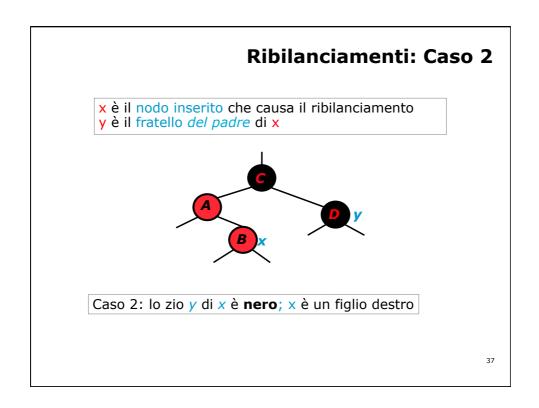
- Ricerca della posizione usando la stessa procedura usata per gli alberi binari di ricerca
- Coloriamo il nuovo nodo di rosso
- Se l'albero risultante non soddisfa le quattro proprietà che lo caratterizzano le si devono risistemare
 - RB-Insert-Fixup ricolora i nodi ed effettua delle rotazioni

34

Quali proprietà possono essere violate?

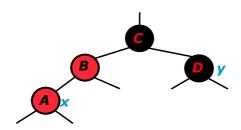
- •La proprietà 2 (i nodi nil sono neri) è soddisfatta perché il nodo inserito è rosso ed i due figli sono la sentinella nil[T] che è nera
- •La proprietà 4 (ogni percorso da un nodo interno ad una foglia ha lo stesso numero di nodi neri) è soddisfatta perché il nodo inserito sostituisce la sentinella (nera) ed esso è rosso con figli neri
- •Le proprietà che possono essere violate sono la 1 (la radice è nera) e la 3 (se un nodo è rosso entrambi i figli sono neri):
 - la 1 viene violata se il nodo inserito è la radice
 - la 3 viene violata se il padre del nodo inserito è rosso







x è il nodo inserito che causa il ribilanciamento y è il fratello *del padre* di x

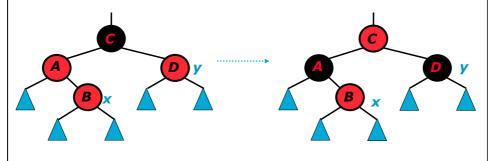


Caso 3: lo zio y di x è **nero**; x è un figlio sinistro

38

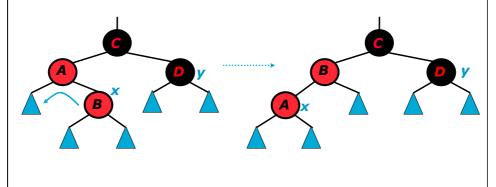
Come risolvere il caso 1

- •Tutti i ∕ sono sottoalberi di uguale altezza nera
- •Il colore del padre di x ed il nodo y sono posti a nero, ed il nonno di x (il nodo C) diventa rosso
- Poi si continua verso l'alto facendo del nonno di x il nuovo x



Come risolvere il caso 2

- •Trasformiano il caso 2 nel caso 3 con una rotazione sinistra
- •Continua a valere la proprietà 4: tutti i percorsi sotto x contengono lo stesso numero di nodi neri



•Cambiamo colori e si effettua una rotazione destra (rotazione per mantenere radice nera)

```
RB-Insert
RB-Insert(T,z)
y \leftarrow nil[T] \triangleright tiene traccia del genitore di z left[z] \leftarrow nil[T]
x←root[T]
                                              right[z] \leftarrow nil[T]
while x≠nil[T] do
                                              color[z]←red
                                             RB-Insert-Fixup(T,z)
  y←x
  if key[z] < key[x] then
    x \leftarrow left[x]
  else x←right[x]
p[z]←y
if y=nil[T] then
  root[T]←z
else if key[z] < key[y] then
        left[y]←z
     else right[y] ←z
                                                                                  42
```

RB-Insert-Fixup

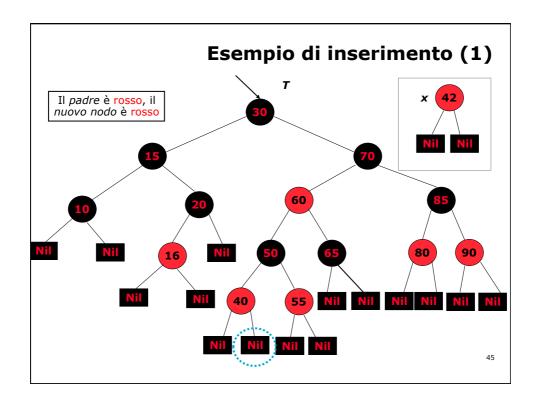
```
RB-INSERT-FIXUP(T, z)
1 while color[p[z]] = RED
 2
         do if p[z] = left[p[p[z]]]
 3
               then y \leftarrow right[p[p[z]]]
 4
                    if color[y] = RED
 5
                       then color[p[z]] \leftarrow BLACK
                                                                         Case 1
 6
                            color[y] \leftarrow BLACK
                                                                         ⊳ Case 1
 7
                            color[p[p[z]]] \leftarrow RED
                                                                         ⊳ Case 1
 8
                            z \leftarrow p[p[z]]
                                                                         ⊳ Case 1
 9
                       else if z = right[p[z]]
10
                             then z \leftarrow p[z]
                                                                         ⊳ Case 2
11
                                   LEFT-ROTATE(T, z)
                                                                         ⊳ Case 2
12
                            color[p[z]] \leftarrow BLACK
                                                                         ⊳ Case 3
13
                            color[p[p[z]]] \leftarrow \texttt{RED}
                                                                         ⊳ Case 3
14
                            RIGHT-ROTATE(T, p[p[z]])
                                                                         ⊳ Case 3
15
               else (same as then clause
                            with "right" and "left" exchanged)
16 color[root[T]] \leftarrow BLACK
```

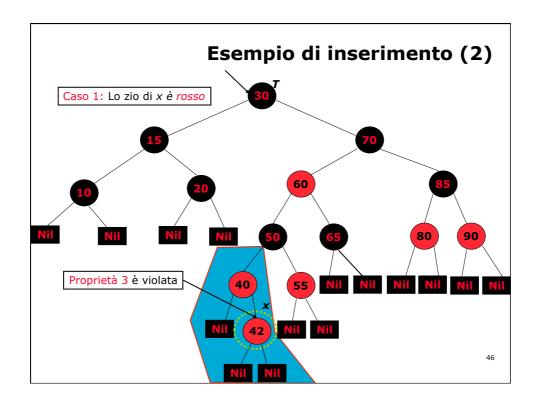
Complessità operazione di inserimento

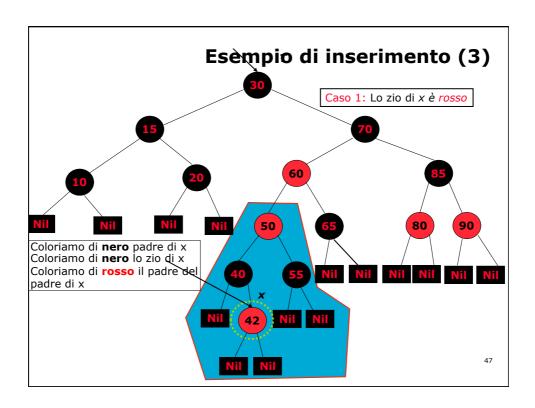
- •La complessità di RB-Insert richiede un tempo O(log n) pari all'altezza di un albero rosso-nero con n nodi
- •La complessità si RB-Insert-Fixup è in funzione del livello d_z in cui si trova il nodo z

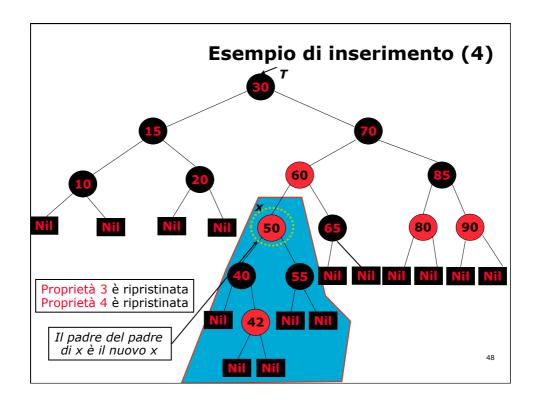
$$T(d_z) \le \begin{cases} c & \text{Casi } 0,2,3 \\ a + T(d_z - 2) & \text{Caso } 1 \end{cases}$$

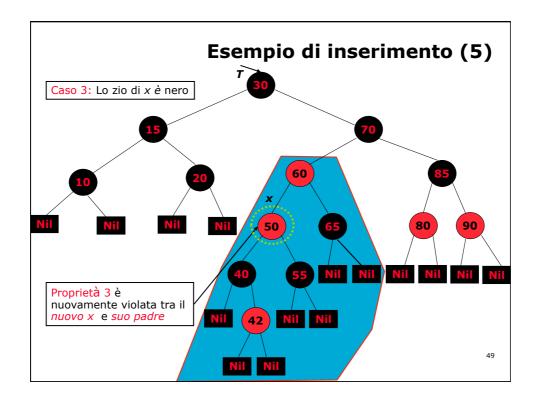
•È immediato concludere che la complessità di RB-Insert-Fixup è O(log n)

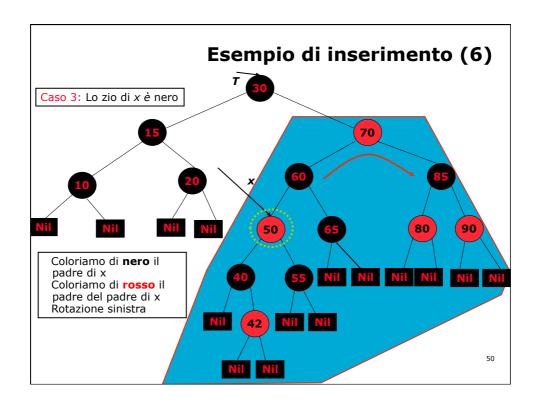


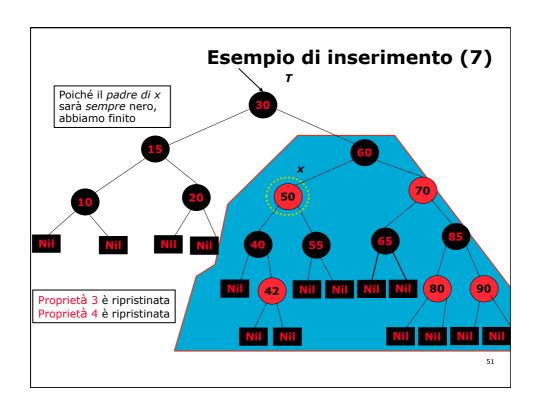


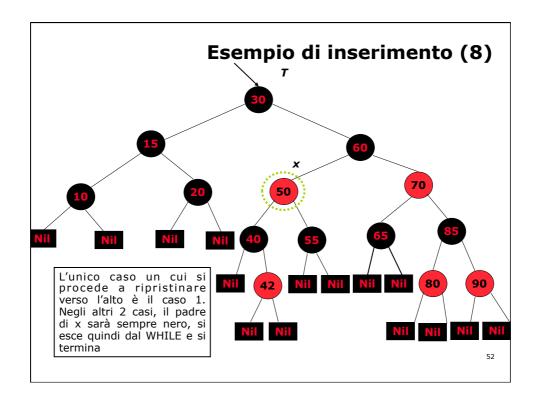












Cancellazione di un elemento (1)

- L'algoritmo di cancellazione per alberi rosso-neri è costruito sull'algoritmo di cancellazione per gli alberi binari di ricerca
- Dopo la cancellazione si deve decidere se è necessario ribilanciare o meno
- •Se il nodo cancellato è rosso non sono necessarie operazioni di ribilanciamento
 - le altezza nere non sono cambiate
 - non sono stati creati nodi rossi consecutivi
 - la radice resta nera perché se il nodo cancellato è rosso, allora non può essere la radice

Cancellazione di un elemento (2)

- •Se il nodo cancellato è nero:
 - possiamo violare la proprietà 1: la radice può essere un nodo rosso
 - possiamo violare la proprietà 3: se il genitore e uno dei figli del nodo cancellato erano rossi
 - abbiamo violato la proprietà 4: altezza nera cambiata

54

Cancellazione di un elemento (3)

```
RB-Delete(T,z) \triangleright z \neq nil[T]; nodo y da cancellare
   if left[z] = nil[T] or right[z] = nil[T] then
      y \leftarrow z
   else
      y \leftarrow \text{Tree-Successor}(z)
   if left[y] = nil[T] then \triangleright elimino y che ha un sottoalbero vuoto
      x \leftarrow right[y]
      x \leftarrow left[y]
   p[x] \leftarrow p[y] \triangleright x sottoalbero di y, l'altro è vuoto if p[y] = nil[T] then
      root[T] \leftarrow x
   else if y = left[p[y]] then
            left[p[y]] \leftarrow x
   RB-Delete-FixUp(T,x)
   return y
                                                                                55
```

Sabrina De Capitani di Vimercati -Algoritmi e Strutture Dati

Cancellazione di un elemento (4)

- •Le differenze con la procedura Tree-Delete sono:
 - i riferimenti a nil sono sostituiti con i riferimenti alla sentinella nil[T]
 - è stato eliminato il controllo per verificare se x è nil e p[x] ← p[y] viene fatto in modo incondizionato
 - se x è la sentinella nil[T], il suo puntatore p punta al padre del nodo y cancellato
 - se y è nero si fa una chiamata alla RB-Delete-Fixup

56

RB-Delete-Fixup

```
RB-DELETE-FIXUP(T, x)
 1 while x \neq root[T] and color[x] = BLACK
          do if x = left[p[x]]
               then w \leftarrow right[p[x]]
if color[w] = RED
                       then color[w] \leftarrow BLACK

color[p[x]] \leftarrow RED
                                                                              Case 1
                                                                              ⊳ Case 1
                             LEFT-ROTATE (T, p[x])
                                                                              ⊳ Case 1
                              w \leftarrow right[p[x]]
                                                                              ⊳ Case 1
                     if color[left[w]] = BLACK and color[right[w]] = BLACK
10
                     then color[w] \leftarrow RED
                                                                             ⊳ Case 2
11
                             x \leftarrow p[x]
                                                                              ⊳ Case 2
                       else if color[right[w]] = BLACK
12
13
                               then color[left[w]] \leftarrow \texttt{BLACK}
                                                                             > Case 3
                                     color[w] \leftarrow RED
                                                                              > Case 3
                                      RIGHT-ROTATE(T, w)

    Case 3

                             w \leftarrow right[p[x]]color[w] \leftarrow color[p[x]]
                                                                              ⊳ Case 3
17
                                                                              ⊳ Case 4
18
                             color[p[x]] \leftarrow \text{BLACK}
19
                             color[right[w]] \leftarrow \texttt{BLACK}
                                                                              ⊳ Case 4
20
                             Left-Rotate(T, p[x])

    Case 4

                             x \leftarrow root[T]

    Case 4

                else (same as then clause with "right" and "left" exchanged)
23 color[x] \leftarrow BLACK
```

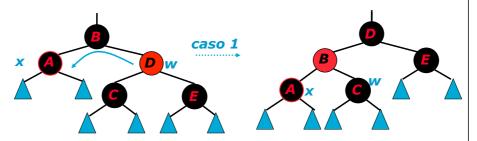
RB-Delete-Fixup

- •Riceve come input un nodo x che può essere:
 - il nodo che era l'unico figlio del nodo cancellato y prima della sua rimozione
 - la sentinella nil[T] se y non aveva figli
- •Ripristina la proprietà rosso-nero con cambi di colore e rotazioni
- Deve gestire quattro casi diversi (e quattro casi simmetrici)

58

Delete: Caso 1

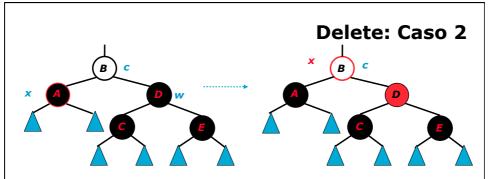
Fratello rosso, padre nero



Il fratello w di x è rosso

w deve avere figli neri

cambiamo i colori di w e del padre di x e li ruotiamo tra loro Non violiamo né la proprietà $\frac{3}{4}$ né la $\frac{4}{4}$ e ci riduciamo ad uno degli altri casi

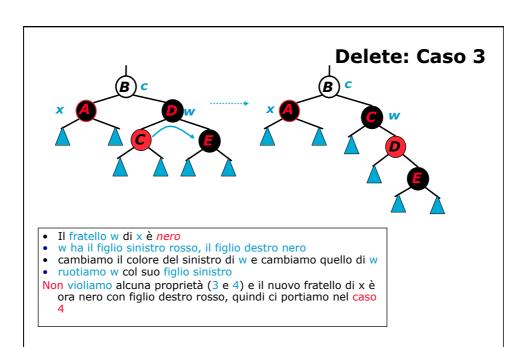


- Il fratello w di x è nero
- w ha in questo caso entrambi i figli neri
- cambiamo il colore di w e il nuovo x diventa il padre

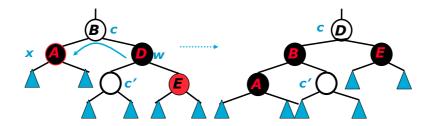
Spostiamo il $\frac{nero\ in\ più}{da\ x}$ al nuovo \underline{x} (il padre) e togliamo il nero da w per rispettare proprietà 4

Si ripristina se è il caso (se il padre era nero) il bilanciamento, altrimenti si termina

Se si arriva dal caso 1, B è sicuramente rosso, quindi dopo il caso 2 non c'è più bisogno di ribilanciare, perché ora B ha un solo nero (il nero in più) e può essere semplicemente colorato di nero.



Delete: Caso 4

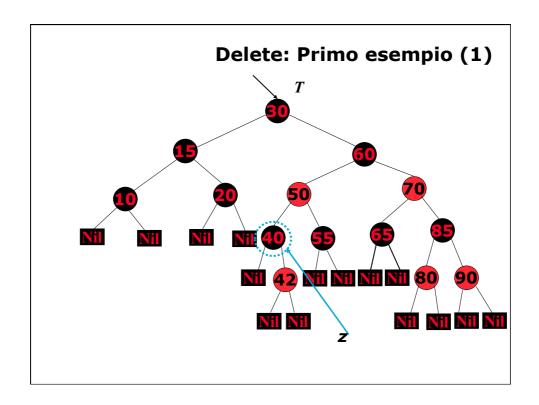


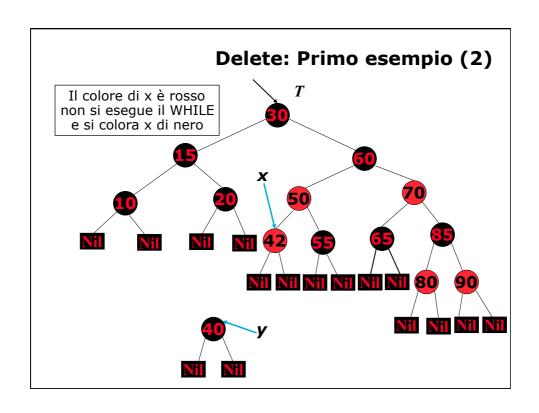
- Il *fratello w* di *x* è *nero*
- w ha in questo caso solo il figlio destro rosso
- cambiamo i colori opportunamente e con una rotazione del padre di x con w si elimina il nero in più su x

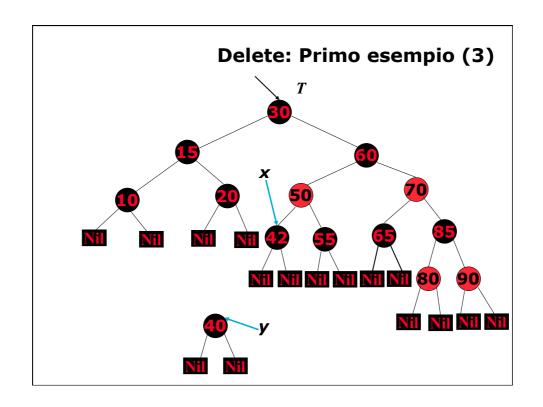
Non violiamo alcuna proprietà (3 e 4) e abbiamo finito!

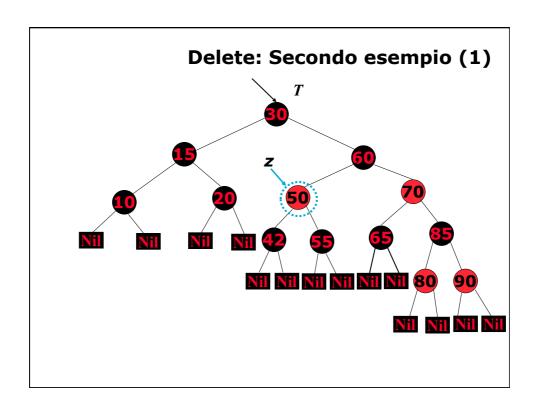
Delete: Simmetria dei casi

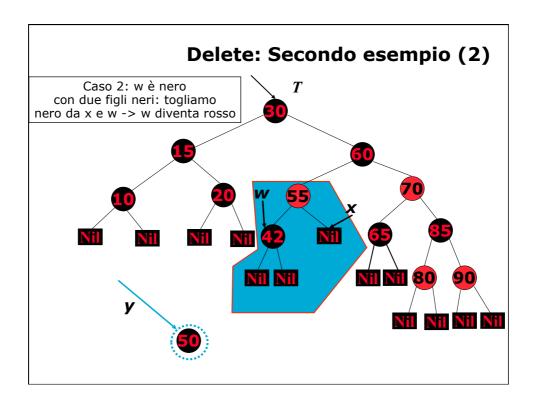
- Abbiamo visto i 4 casi possibili quando il nodo x che sostituisce y (cancellato) è un figlio sinistro
- Esistono anche i 4 casi simmetrici (con destro e sinistro scambiati) quando x è figlio destro

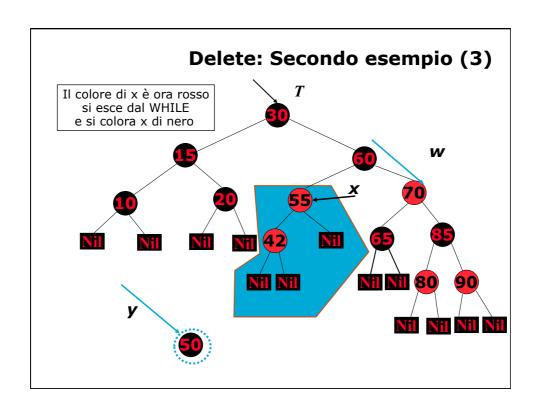


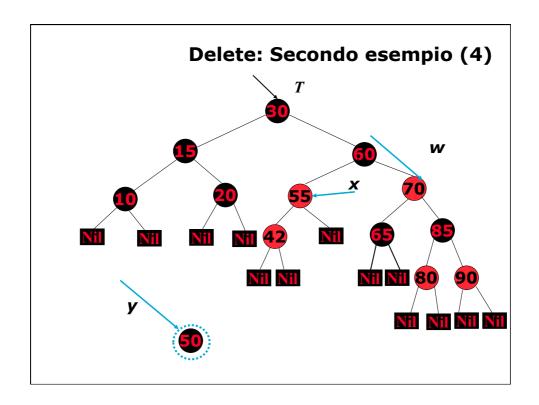


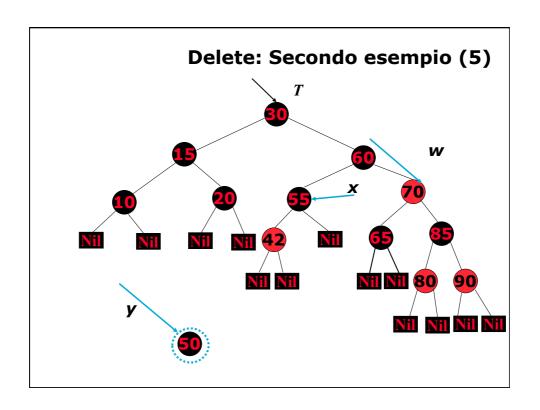


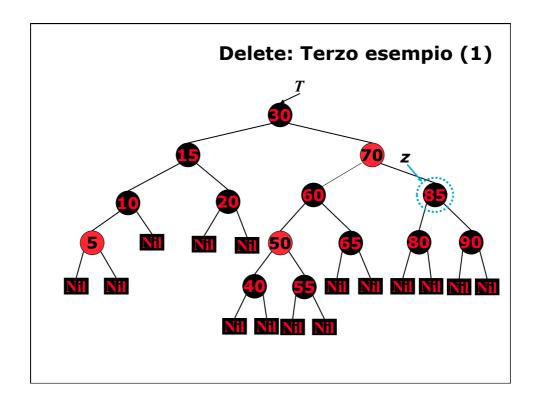


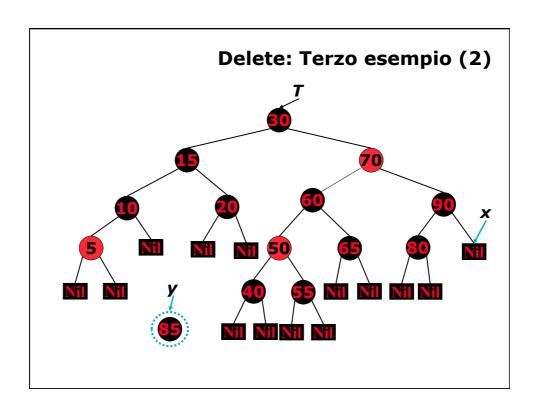


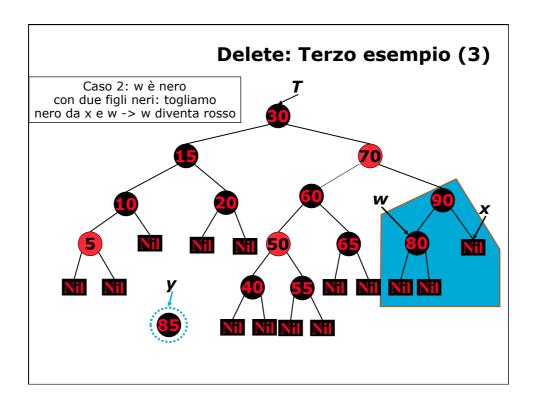


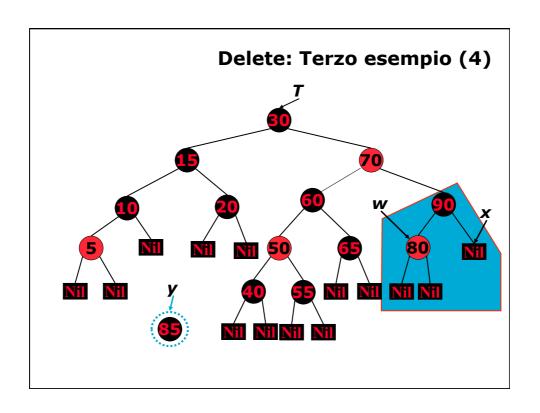


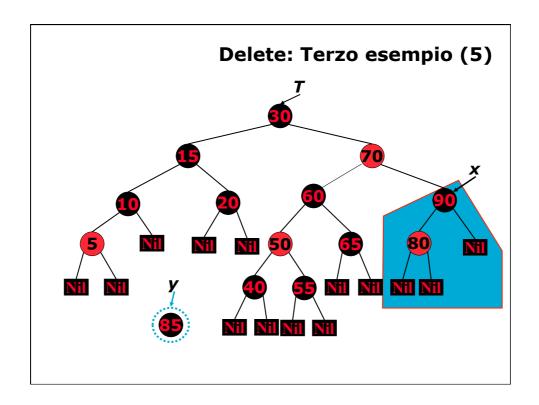


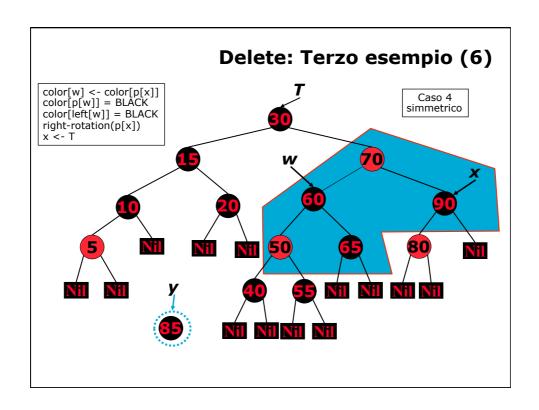


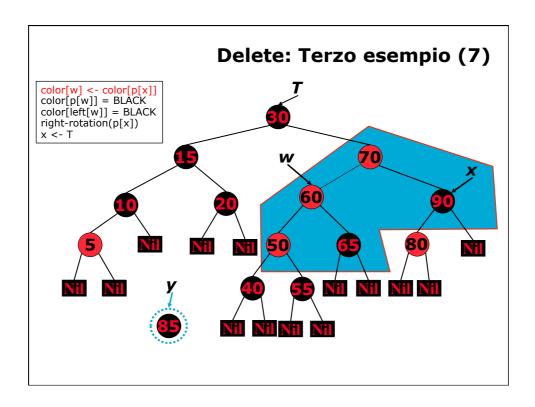


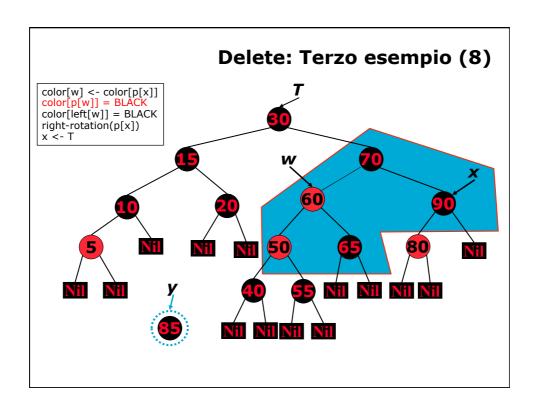


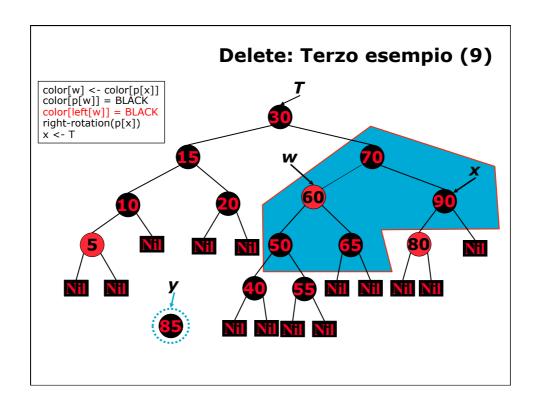


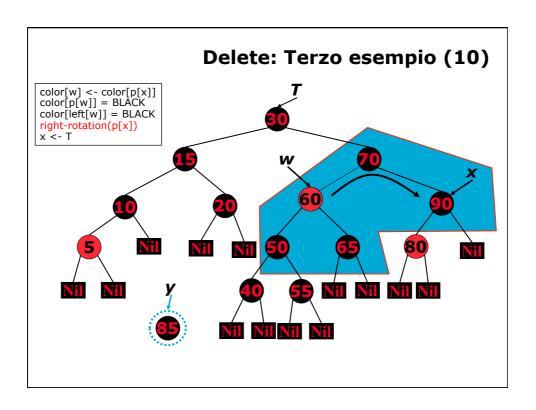


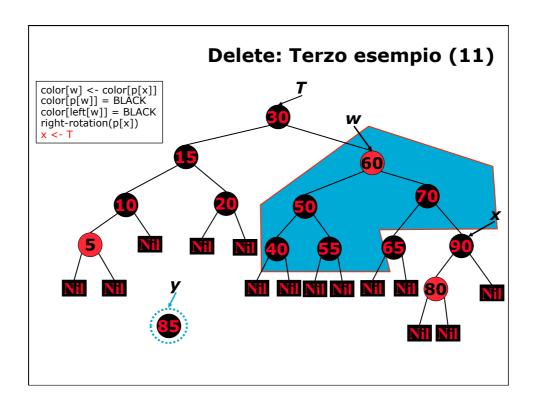


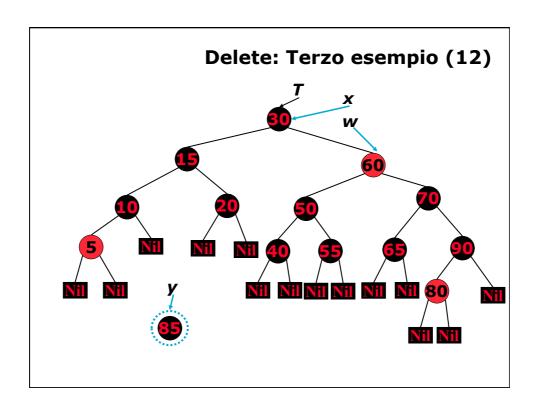












Cancellazione

L'operazione di cancellazione è concettualmente complicata!

Ma efficiente:

- •il caso 4 è risolutivo e applica una sola rotazione
- •il caso 3 applica una rotazione e passa nel caso 4
- •il caso 2 non fa rotazioni e passa in uno qualsiasi dei casi *ma salendo lungo il percorso di cancellazione*
- •il caso 1 fa una rotazione e passa in uno degli altri casi (ma se va nel caso 2, il caso 2 termina)

 Quindi

al massimo vengono eseguite 3 rotazioni per iterazione del ciclo while

Alberi 2-3 (1)

- Introduciamo prima la nozione di nodo 2-3
- •Un nodo 2-3 è un insieme ordinato contenente due chiavi k_1 e k_2 (in ordine crescente) e tre puntatori p_0 , p_1 e p_2 tali che:
 - p₀ punta al sottoalbero contenente chiavi k tali che k < k₁
 - p₁ punta al sottoalbero contenente chiavi k tali che k₁ < k < k₂
 - p₂ punta al sottoalbero contenente chiavi k tali che k > k₂

Alberi 2-3 (2)

Un albero 2-3 è un insieme di nodi 2-3 che soddisfa la definizione di albero e che verificano le seguenti proprietà:

- •le foglie sono tutte sullo stesso livello
- •ogni nodo ha al più tre figli

86

Ricerca in un albero 2-3

- La ricerca di una chiave k avviene confrontando la chiave da ricercare con il contenuto di un nodo a partire dal nodo radice
 - se $k < k_1$ la ricerca prosegue nel sottoalbero indicato da p_0
 - se $k = k_1$ la ricerca termina con successo
 - \bullet se nel nodo c'è una sola chiave la ricerca prosegue nel sottoalbero indicato da \textbf{p}_1
 - se nel nodo ci sono due chiavi è necessario eseguire ulteriori confronti:
 - se $k_1 < k < k_2$ la ricerca prosegue nel sottoalbero indicato da p_1
 - se $k = k_2$ la ricerca termina con successo
 - se $k > k_2$, la ricerca prosegue nel sottoalbero indicato da p_2

Inserimento in un albero 2-3

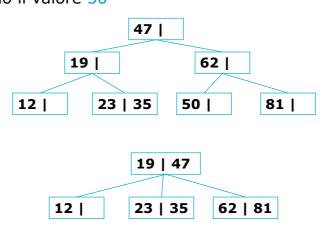
- L'inserimento di un nuovo elemento avviene dopo aver effettuato una ricerca senza successo che si arresta su un nodo v dal quale dovrebbe proseguire tramite un puntatore vuoto
- Se v contiene una sola chiave, la nuova chiave viene inserita nel nodo stesso, ordinatamente con quella preesistente
- Se \vee contiene già due chiavi, si considera l'insieme ordinato (a_1,a_2,a_3) costituito dalle due chiavi del nodo e dalla nuova chiave
 - v viene suddiviso (split) in due nodi contenenti rispettivamente a₁ e a₃, mentre a₂ viene inserito (ricorsivamente) nel padre del nodo d'arresto assieme a due puntatori ai nodi contenenti a₁ e a₃
 - questa operazione può ripetersi a ritroso fino alla radice e, nel caso in cui sia necessario suddividere anche la radice, l'altezza dell'albero viene incrementata di uno

88

Cancellazione in un albero 2-3

Può essere eseguita ri-compattando (operazione di merge) eventualmente i nodi interessati

Cancello il valore 50



Complessità operazioni su alberi 2-3

Anche in questo caso tutte le operazioni hanno costo $O(\log n)$ poiché l'altezza dell'albero è $O(\log n)$ e le operazioni di merge e split richiedono tempo costante