



# BASI DI DATI 2

*STRUTTURE DI INDICI PER I FILE*

# Concetto di indice

- Un file ha un'organizzazione *primaria* (*organizzazione non ordinata, ordinata, hash*).
- Gli **indici** sono strutture aggiuntive di accesso ai file, usate per velocizzare il reperimento delle informazioni.
  - ▣ Forniscono un modo alternativo per accedere ai record del file (**secondary access paths**).
- Concettualmente una struttura di accesso secondaria è simile all'indice di un libro:
  - ▣ Tutti i termini principali sono listati in ordine alfabetico.
  - ▣ Cercando un termine nell'indice, otteniamo una lista di indirizzi (*i numeri di pagina*), che ci permette di localizzare molto velocemente il testo richiesto all'interno del libro.

## Concetto di indice (2)

- Gli **indici**, quindi, permettono di accelerare notevolmente le operazioni di ricerca in particolari condizioni, senza modificare il **posizionamento fisico** dei record.
- Di norma sono basati su un singolo file ordinato (**indici a livello singolo**), o su strutture dati ad albero (**indici multilivello**, **B<sup>+</sup>-alberi**)



# Tipi di indice a livello singolo

# Indici a livello singolo

- In un file con più campi, in genere si definisce l'indice su un solo campo, detto *indexing field*.
- Un indice memorizza:
  - ▣ Il valore del campo *index*.
  - ▣ Una lista di puntatori a tutti i *disk block* che contengono record con quel valore di campo.
- I valori nell'indice sono *ordinati* così da consentire l'esecuzione di una ricerca binaria.

La ricerca è più efficiente poiché il file indice è più piccolo del file di dati.

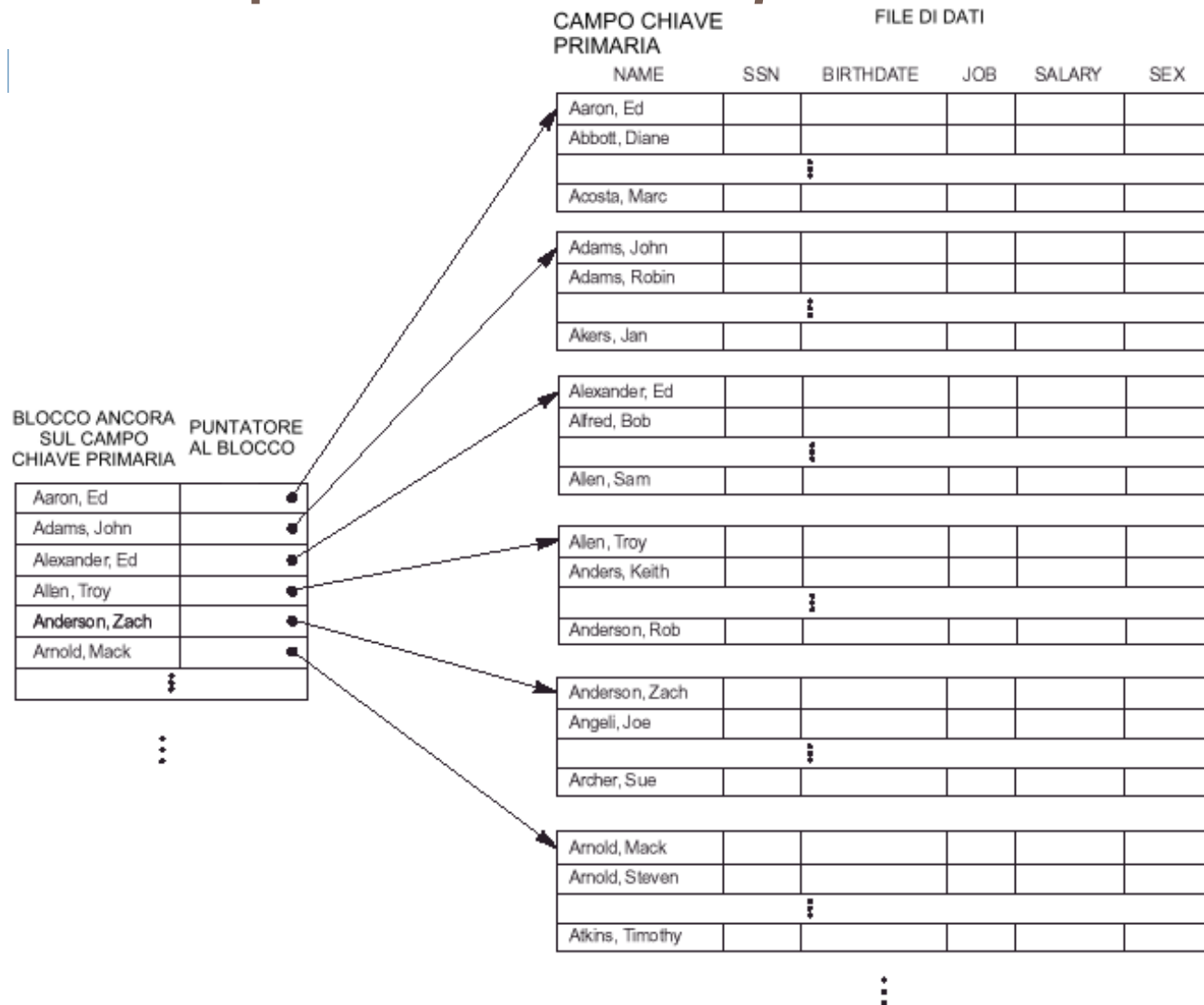
# Tipi di indici

- Esistono più tipi di indici:
  1. **Indice primario**: specificato su un **campo chiave** di **ordinamento** (*ordering key field*) di un file ordinato di record.
  2. **Indice clustering**: specificato su un **campo non chiave** di **ordinamento** (*ordering non-key field*) di un file ordinato di record.
  3. **Indice secondario**: specificato su un **campo non di ordinamento** (*non-ordering field*) di un file di record.

# Indici primari

- Un *primary index* è un file ordinato di record a lunghezza fissa con due campi. Ogni record, (detto *index entry*) ha la seguente struttura:
  - ▣ Il 1° campo, dello stesso tipo di dati della chiave primaria, contiene il valore della chiave primaria del primo record di un blocco (detto *record ancora*).
  - ▣ Il 2° campo: di tipo puntatore a un disk block contiene l'indirizzo del blocco contenente il record.
- I due valori dei campi dell'indice si indicano con **<K(i), P(i)>**
- Il numero di entry è pari al numero di blocchi del file dati ordinato.

# Indici primari: *Esempio*



- Indice primario su un campo di ordinamento chiave.



# Indici densi e sparsi

- Gli indici possono essere *densi* o *sparsi*:
  - ▣ È **denso** se contiene un'entry per ogni possibile valore del campo chiave.
  - ▣ È **sperso** se contiene meno entry rispetto al numero di valori del campo chiave.
- Gli indici primari sono sparsi: il file indice è molto più piccolo del file dati poiché:
  - ▣ Il numero di entry è inferiore a quello dei record del file dati.
  - ▣ La dimensione di un'entry è inferiore a quella di un record di dati (*ha solo due campi*).

# Retrieval con indici primari

- Un record con valore chiave **K** è posizionato nel blocco di indirizzo **P(i)**, dove  **$K(i) \leq K \leq K(i + 1)$** :
  - L'**i**-mo blocco contiene tutti i record che soddisfano tale formula, poiché i dati sono ordinati fisicamente sul campo chiave.
- Per trovare un record con valore di chiave primaria **K**:
  - Si effettua una ricerca binaria sul file indice per trovare l'entry **i** corrispondente.
  - Si recupera il blocco **P(i)** del file di dati.
- La ricerca binaria sul file indice è molto efficiente grazie alle sue ridotte dimensioni.

# Risparmio di accessi a blocco: *Esempio*

- Supponiamo di avere:
  - ▣ File ordinato di  $r = 30.000$  record
  - ▣ Dimensione di un blocco  $B = 1.024$  byte
  - ▣ Ogni record è a lunghezza fissa  $R=100$  byte
  - ▣ Il bfr  $\lfloor (B/R) \rfloor$  è  $\lfloor (1024/100) \rfloor \rightarrow 10$  record per blocco
  - ▣ Il numero di blocchi  $b \lceil (r/bfr) \rceil$  è  $\lceil (30000/10) \rceil \rightarrow 3000$  blocchi
  
- Effettuando una ricerca binaria sul file dati dovremmo effettuare:  
 $\lceil (\log_2 b) \rceil = \lceil (\log_2 3000) \rceil = 12$  accessi a blocco

## Risparmio di accessi a blocco: *Esempio* (2)

- Usando invece un indice primario:
  - ▣ Campo chiave di ordinamento  $V = 9$  byte
  - ▣ Campo puntatore a blocco  $P = 6$  byte
  - ▣ dimensione di ogni entrata  $R_i = (9+6) = 15$  byte
  - ▣ Num. entrate  $r_i = b$  (# blocchi file dati) = 3000
  - ▣  $bfr_i = \lfloor (B/R_i) \rfloor = \lfloor (1024/15) \rfloor = 68$  entrate a blocco
  - ▣  $b_i = \lceil (r_i/bfr_i) \rceil = \lceil (3000/68) \rceil = 45$  blocchi
- Una ricerca binaria sul file indice richiede circa:  
 $\lceil (\log_2 b_i) \rceil = \lceil (\log_2 45) \rceil = 6$  accessi a blocchi del file indice + 1 accesso al blocco del file dati = **7 accessi a blocco.**

# Inserimento e cancellazione

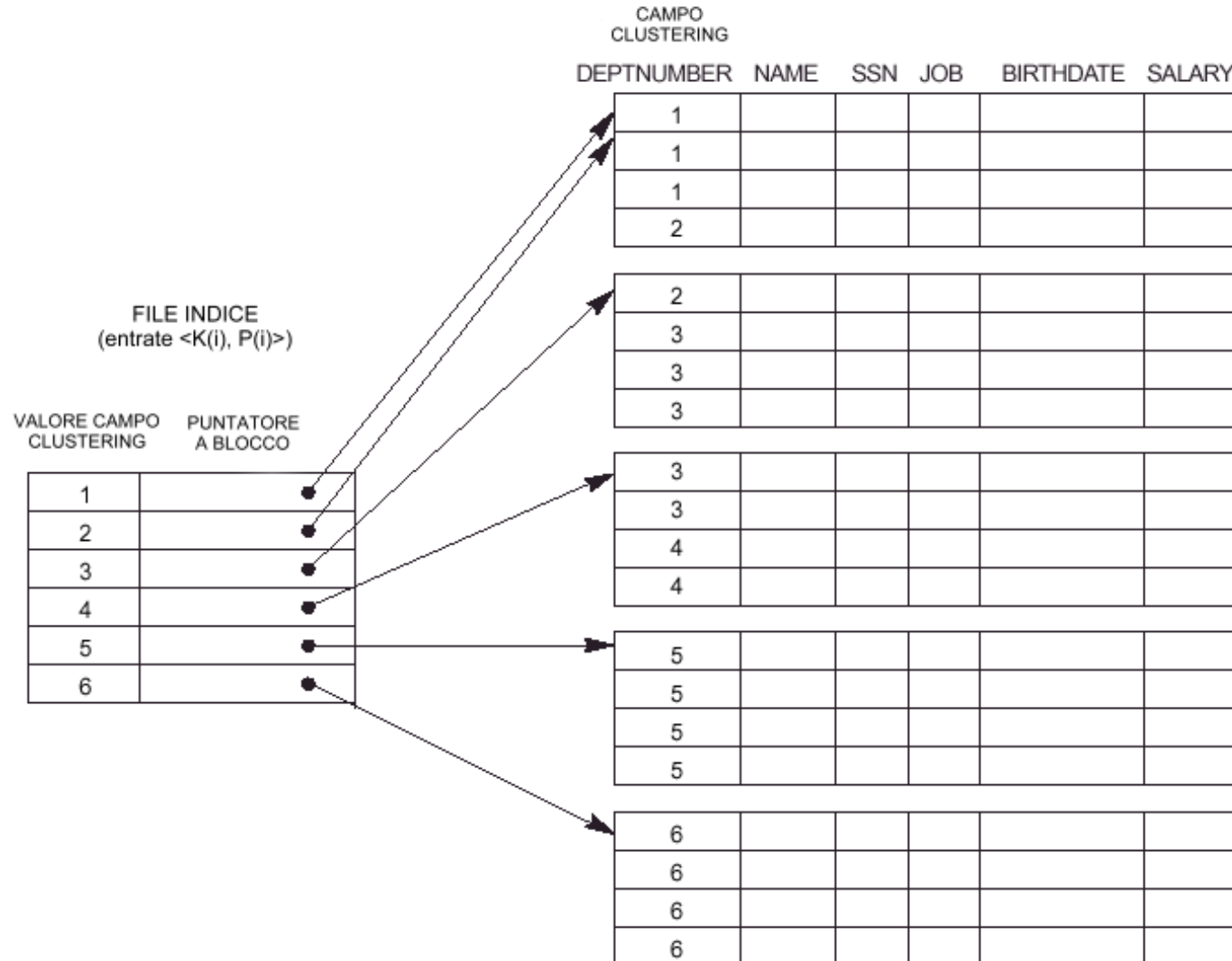
- L'inserimento e la cancellazione presentano difficoltà (*come per i file ordinati*):
  - ▣ Per inserire un nuovo record, bisogna:
    - spostare alcuni record per fare spazio al nuovo record e
    - cambiare alcune entry nel file indice.
  - ▣ Soluzioni possibili:
    1. utilizzo di un file di overflow non ordinato (*come per i sorted file*).
    2. utilizzo di una lista a puntatori di record di overflow per ciascun blocco nel file dati.
- Per cancellare un record, si usano i marcatori di cancellazione.

# Indici di Clustering

- Si costruiscono su un file i cui record sono ordinati su un campo non chiave, detto *campo clustering*, che può avere ripetizioni di valori.
- L'indice di clustering è un file ordinato con due campi:
  - ▣ 1° campo: stesso tipo di dati del campo clustering, contiene un valore del campo clustering.
  - ▣ 2° campo: di tipo puntatore a un disk block, riferenzia il **primo blocco** del file dati che contiene un record con tale valore del campo clustering.
- È un indice sparso: c'è una entry per ogni valore distinto del campo clustering.

# Indici di Clustering: *Esempio*

FILE DI DATI

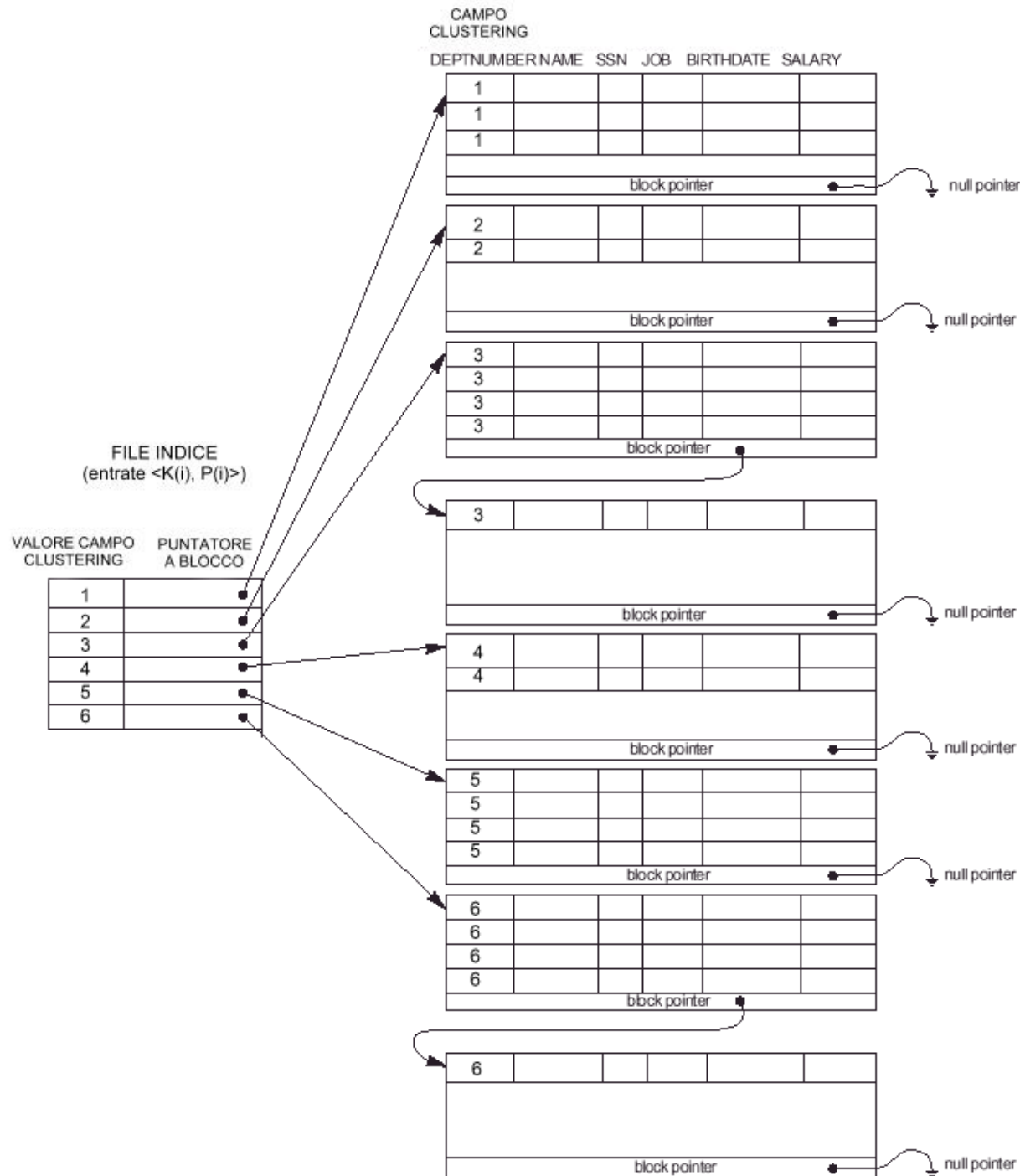


- Indice clustering su un campo di ordinamento non chiave.

# Inserimento e cancellazione

- Inserimento e cancellazione comportano delle difficoltà (i record sono ancora ordinati fisicamente).
- Per migliorare la situazione spesso si riserva un intero blocco per ogni valore del campo cluster.





- Indice clustering con un blocco separato per ogni gruppo di record con lo stesso valore di clustering.

# Indici primari vs. hashing

- Gli indici ricordano le directory usate per l'hashing:
  - ▣ Analogie
    - Su entrambi si esegue una ricerca di un puntatore a un blocco di dati che contiene il record desiderato.
  - ▣ Differenze
    - Una ricerca su indice usa i valori del campo di ricerca stesso.
    - Una ricerca su una directory hash usa il valore hash calcolato applicando la funzione hash al campo di ricerca.

# Indici Secondari

- A differenza dell'indice primario, **è possibile avere più indici secondari** associati allo stesso file:
  - ▣ In questo modo si velocizzano ricerche effettuate su campi non di ordinamento, creando nuove strutture di accesso.
- Un indice secondario è un file ordinato con due campi:
  - ▣ 1° campo (di indicizzazione): stesso tipo di dati di un campo non-ordering.
  - ▣ 2° campo: puntatore a un disk block o a un record.

# Indici Secondari (2)

- Esistono due tipi di indici secondari, in base al campo di indicizzazione:
  1. Il campo di indicizzazione è un campo chiave, detto **chiave secondaria**, contenente un valore **distinto** per ogni record nel file dati.
  2. Il campo di indicizzazione è un campo **non chiave**, in cui più record nel file dati possono avere lo stesso valore.

# Indici Secondari: Caso 1

- È un indice denso :
  - ▣ C'è un'entrata per ogni record nel file dati.
  - ▣ Le **block anchor** non si possono usare perché i record del file dati non sono ordinati fisicamente rispetto al campo chiave secondaria.
- Le entrate dell'indice  $\langle K(i), P(i) \rangle$  sono ordinate sul valore  $K(i)$ , quindi si può eseguire una ricerca binaria sull'indice.
- $P(i)$  è un puntatore al blocco.
  - ▣ Se non esistesse l'indice il tempo di ricerca sarebbe lineare poiché il file non è ordinato su quel campo.

CHIAVE  
SECONDARIAFILE INDICE  
(entrate  $\langle K(i), P(i) \rangle$ )CAMPO INDICE  
CHIAVE PUNTATORE  
A BLOCCO

1	
2	
3	
4	
5	
6	
7	
8	

9	
10	
11	
12	
13	
14	
15	
16	

17	
18	
19	
20	
21	
22	
23	
24	

9			
5			
13			
8			

6			
15			
3			
17			

21			
11			
16			
2			

24			
10			
20			
1			

4			
23			
18			
14			

12			
7			
19			
22			

- Un indice secondario denso, su un campo non di ordinamento.

# Indici Secondari: Caso 2

- Il file indice secondario è creato su un campo **non chiave** (*oltre che non-ordering*), quindi più record nel file dati possono avere lo stesso valore del campo di indicizzazione.
- Possibili implementazioni:
  1. Includere più entry con lo stesso valore  **$k(i)$** , una per ogni record (*indice denso*).
  2. Avere record di lunghezza variabile per le entry, con un campo puntatore:  
si crea una lista a puntatori  **$\langle P(i,1), \dots, P(i,k) \rangle$**   
nell'entrata per  **$k(i)$** , con un puntatore a ciascun blocco con valore di indexing field  **$k(i)$** .

## Indici Secondari: Caso 2 (2)

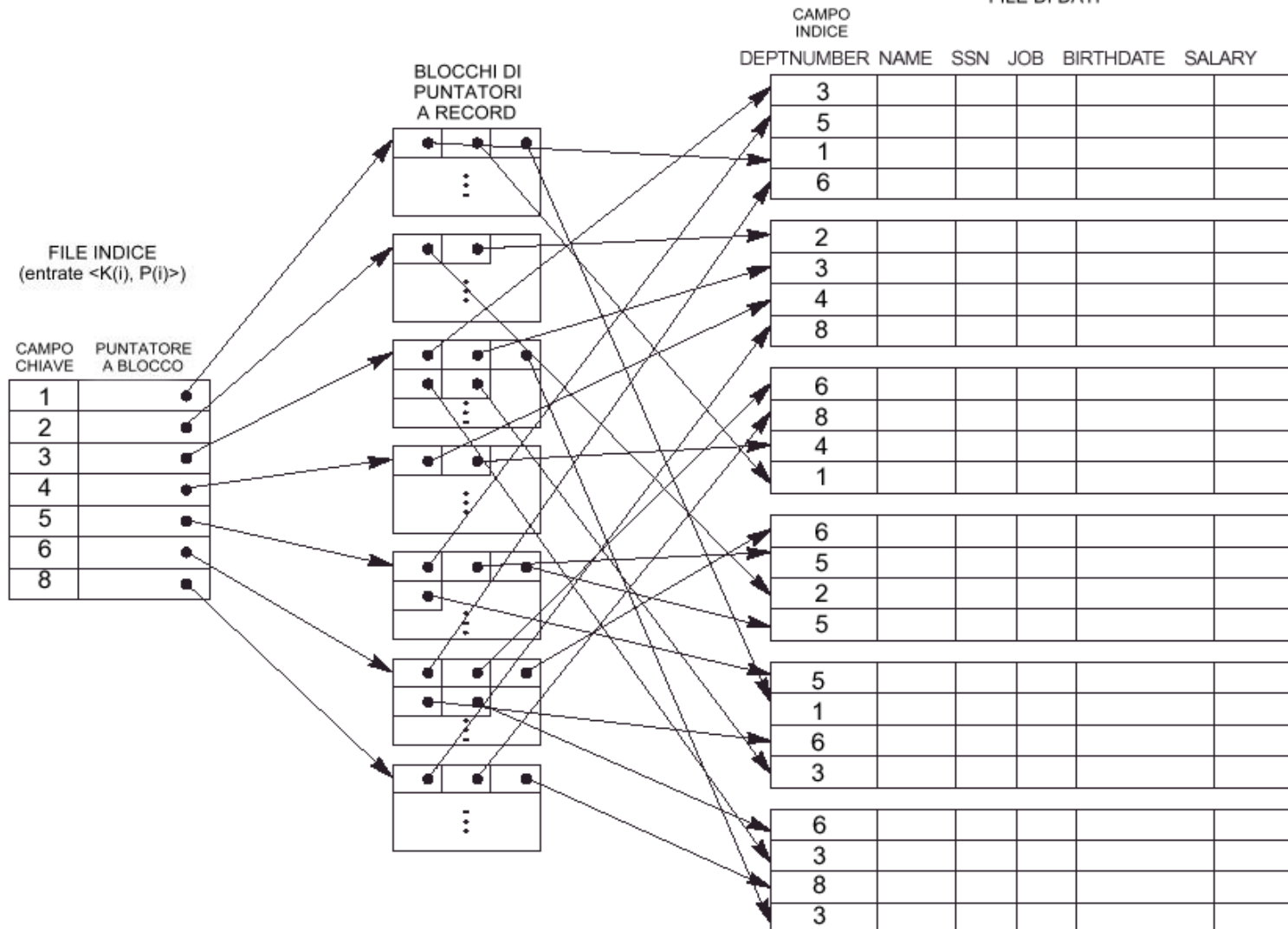
3. Si crea un nuovo livello di indirizione per gestire puntatori multipli:

- Le entrate restano a lunghezza fissa;
- Il secondo campo **P(i)** è un puntatore ad un blocco di puntatori a record.  
Ogni puntatore nel secondo blocco referencia un record nel file dati con valore del campo di indicizzazione **k(i)**.
- Se i record con valore **k(i)** sono tanti da riempire il secondo blocco, si crea una lista a puntatori di blocchi.



# Indici Secondari: *Esempio*

FILE DI DATI



- Un indice secondario su un campo non chiave, implementato con un livello di indirezione.

## Indici Secondari: Caso 2 (3)

- Il retrieval via indice richiede un accesso a blocco in più, a causa del livello extra.
- Il retrieval su condizioni di selezione complesse può essere gestito direttamente con i puntatori senza dover operare sul file.
- L'inserimento di nuovi record è immediato.
- *Osservazione:*
  - ▣ Un indice secondario fornisce un **ordinamento logico** sui record rispetto al campo di indicizzazione.

# Tipi di indice: *Sommario*

	Campo ordering	Campo non ordering
Campo chiave	Indice Primario	Indice Secondario (chiave)
Campo non chiave	Indice Clustering	Indice Secondario (non chiave)

# Proprietà dei tipi di indice

	Numero di entry di 1° livello	Denso o non denso	Ancoraggio dei blocchi del file dati
Primario	Numero di blocchi del file dati	Non denso	Sì
Clustering	Numero di valori distinti del campo indexing	Non denso	Sì/No*
Secondario (chiave)	Numero di record del file dati	Denso	No
Secondario (non chiave)	Num. record del file dati (caso 1) oppure num. valori distinti del campo indexing (casi 2 e 3)	Denso o non denso	No

\* Sì, se ogni valore distinto del campo ordering inizia un nuovo blocco; no altrimenti.



# Indici *Multilivello*

# Indici multilivello

- L'idea di base degli indici multilivello è di ridurre la dimensione dell'indice per velocizzare la ricerca binaria.

Per fare ciò, si creano vari livelli di indici, ognuno con poche entry.

- Definiamo **fan-out** (*fo*) di un indice multilivello il blocking factor  $bfr_i$  dell'indice:
  - ▣ Approssimativamente, una ricerca impiegherà  $\log_{fo} b_i$  accessi a blocco (per un indice con  $b_i$  entry).

# Struttura di un indice multilivello

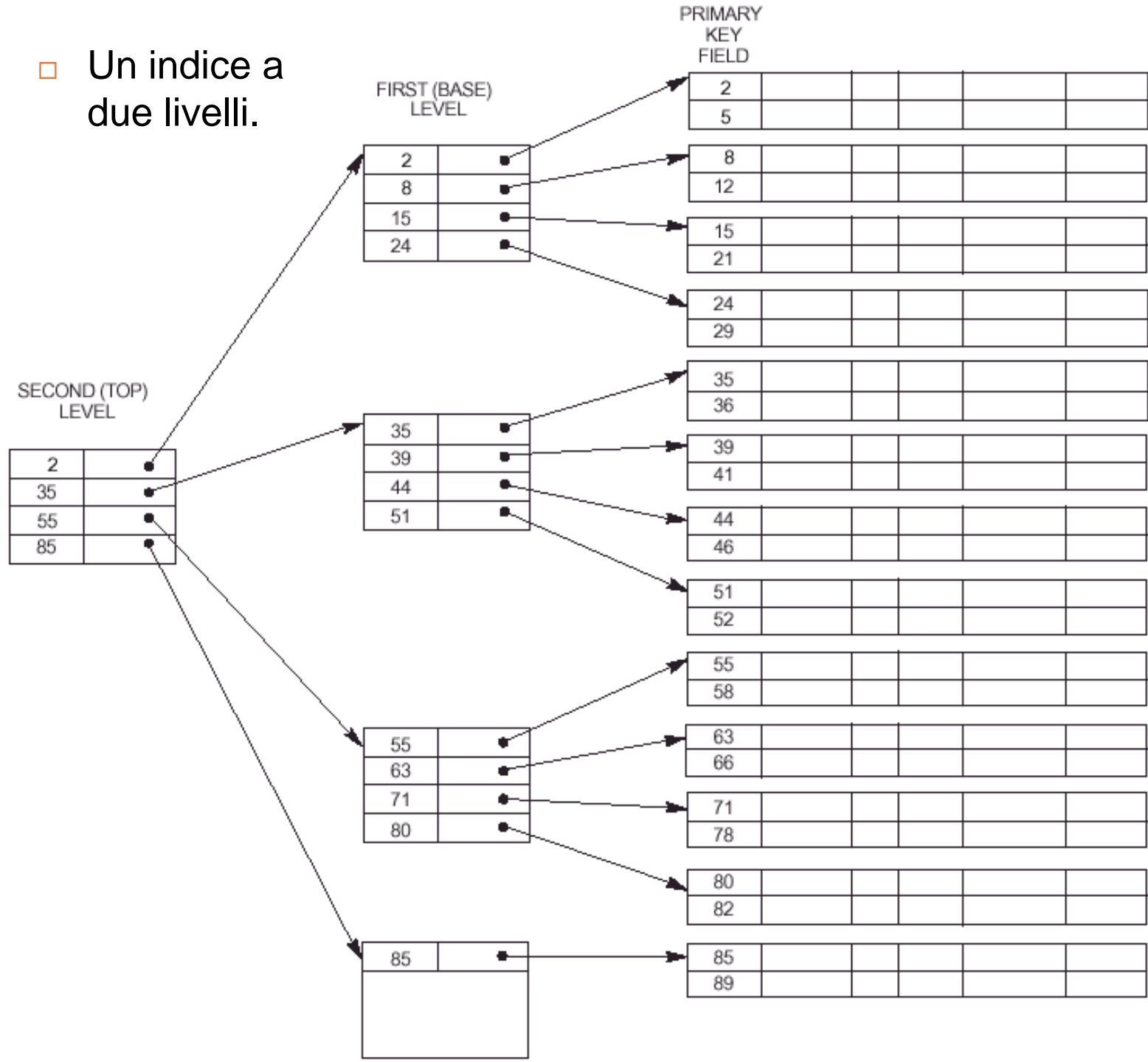
- **Primo livello (o livello di base):**  
è un file indice ordinato con un valore distinto per ogni  $k(i)$ .
- **Secondo livello:**  
è un indice primario sul primo livello.  
Utilizza le block anchor:
  - ▣ numero entry di 2° livello = numero blocchi del 1° livello.
- È possibile continuare questo processo, creando un indice per il secondo livello, ecc ...
  - ▣ È necessario creare un nuovo livello solo se il precedente occupa più di un blocco su disco.
- L'indice di livello superiore è detto *top index*.

# Struttura di un indice multilivello (2)

- Lo schema multilivello può essere usato su qualsiasi tipo di indice, purché l'indice del primo livello abbia valori distinti per  $k(i)$  ed entrate a lunghezza fissa.



- Un indice a due livelli.



# Blocking Factor

- I livelli hanno lo stesso blocking factor del primo livello, poiché tutte le entry hanno la stessa dimensione.
- Se il primo livello ha  $r_1$  entrate e blocking factor  $bfr_i = fo$ , allora richiederà  $\lceil (r_1/fo) \rceil$  blocchi, che è anche il numero di entry del secondo livello.
- Ciascun livello riduce il numero di entrate del livello precedente di un fattore  $fo$ , quindi vale la formula ( $t$  rappresenta il livello top):

$$(r_1/(fo)^t) \geq 1 \rightarrow t = \lceil (\log_{fo} r_1) \rceil$$

# Esempio

- Supponiamo di voler convertire l'indice denso secondario dell'esempio visto in precedenza, costruito su un file di  $r = 30.000$  record a lunghezza fissa:
  - ▣  $bfr_i (= fo) = 68$
  - ▣ Numero blocchi 1° livello  $b_1 = 442 (= 30.000/68)$
  - ▣ Numero blocchi 2° livello  $b_2 = \lceil (b_1 / fo) \rceil = 7$
  - ▣ Numero blocchi 3° livello  $b_3 = \lceil (b_2 / fo) \rceil = 1$
- Quindi il livello top è il terzo ( $t=3$ ) ed il **numero di accessi a blocco è  $t + 1 = 4$**  contro i 10 accessi dell'indice secondario (+1 per l'accesso al file dati).



# Indici multilivello dinamici con B-Tree e B<sup>+</sup>-Tree

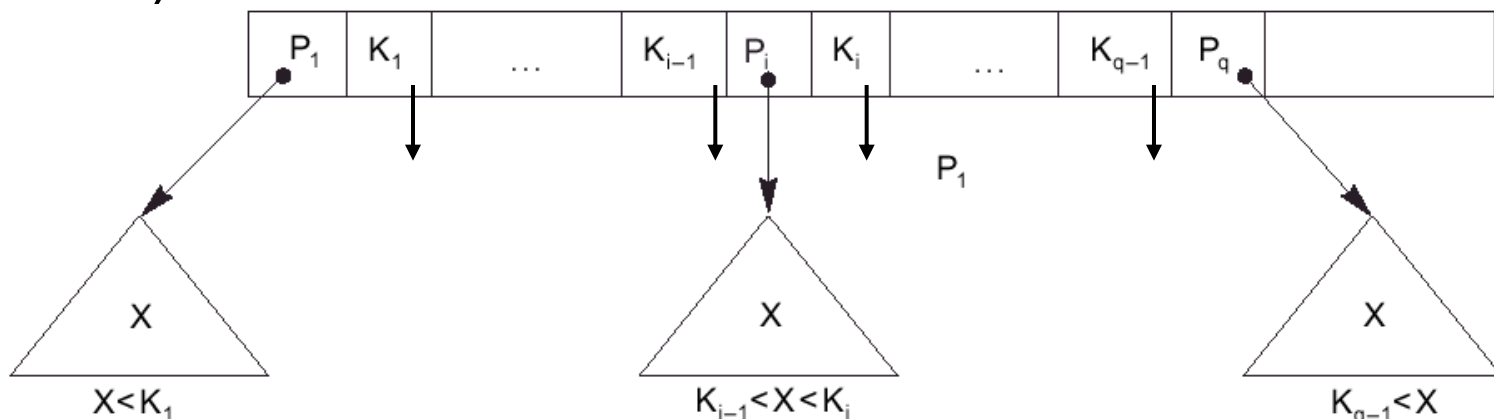
# Alberi di ricerca

- Un albero di ricerca è un particolare tipo di struttura dati ad albero, utilizzato per recuperare un record dato il valore di uno dei suoi campi.
- Gli indici multilivello sono una variante degli alberi di ricerca:

*il valore del campo indice in ogni nodo guida verso il prossimo nodo, finché non si raggiunge il blocco del data file contenente il record cercato.*

# Proprietà di un albero di ricerca

- Dato un albero di ricerca di *ordine*  $p$ :
  - ogni nodo contiene al più  $p - 1$  valori di ricerca e  $p$  puntatori, nell'ordine:  
 $\langle P_1, K_1, P_2, K_2, \dots, P_{q-1}, K_{q-1}, P_q \rangle$ , con  $q \leq p$
  - ogni  $P_i$  è un puntatore ad un nodo figlio (o un puntatore nullo).
  - ogni  $K_i$  è un valore di ricerca da qualche insieme ordinato (*si suppone che tutti i valori di ricerca siano distinti*).



# Vincoli di un albero di ricerca

- In un albero di ricerca, devono valere sempre due vincoli:
  1. All'interno di ogni nodo,  $K_1 < K_2 < \dots < K_{q-1}$
  2. Per ogni valore  $X$  nel sottoalbero puntato da  $P_i$  si deve avere  $K_{i-1} < X < K_i$  per  $1 < i < q$ 
    - $X < K_i$  (per  $i = 1$ )
    - $K_{i-1} < X$  (per  $i = q$ )

# Ricerca di valori

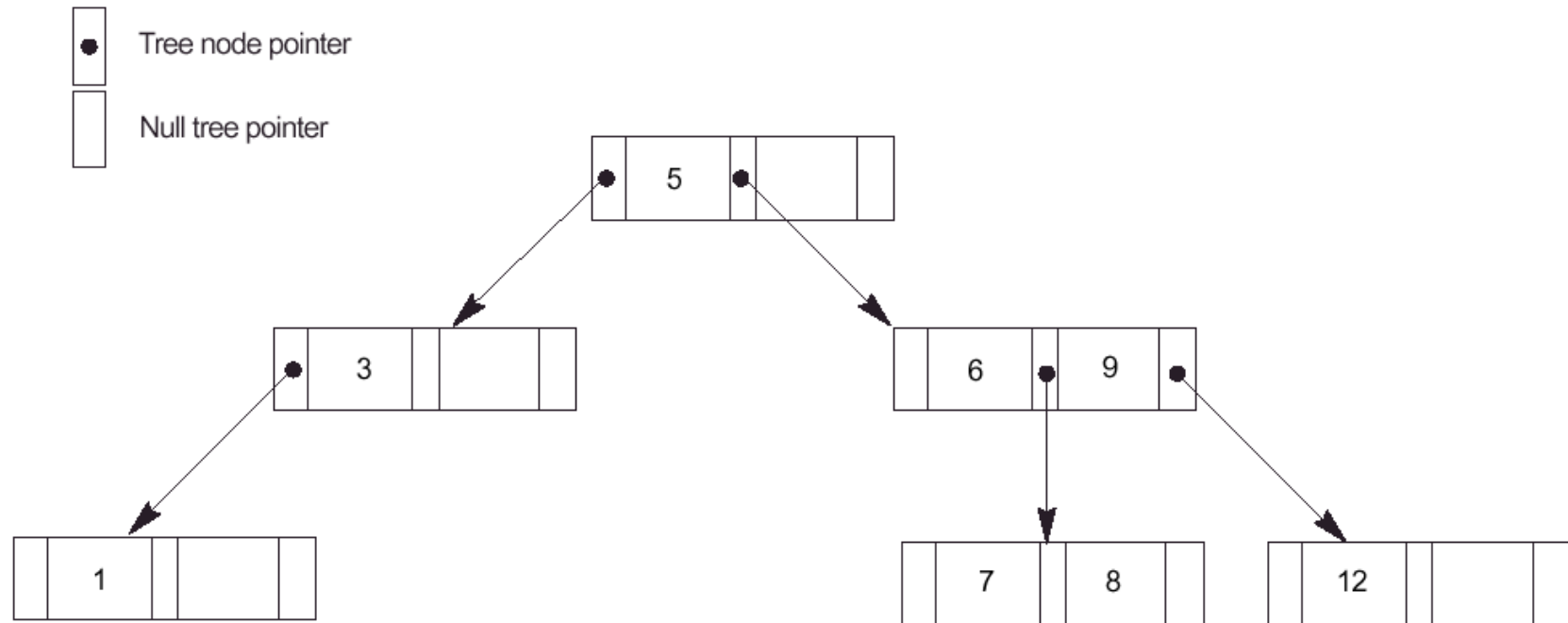
- Utilizzo di alberi di ricerca per recuperare un record memorizzato in un file su disco:
  - ▣ I valori nell'albero sono i valori di un campo del record, detto **campo di ricerca**.

Ad ogni valore nell'albero è associato un puntatore o al record nel file dati con tale valore o al disk block contenente il record.
  - ▣ Un albero di ricerca può essere esso stesso memorizzato su disco, assegnando ogni nodo ad un blocco su disco.



# Albero di ricerca: *Esempio*

- Un albero di ricerca di ordine  $p = 3$ :



# Problemi con alberi di ricerca

- Gli algoritmi per inserire e cancellare record dall'albero (*mantenendo sempre i due vincoli*) non garantiscono che l'albero rimanga **bilanciato**:
  - ▣ Mantenere l'albero di ricerca bilanciato è fondamentale per limitare il numero di accessi a blocco, poiché l'altezza di un albero bilanciato è sempre logaritmica sul numero di nodi.
- La cancellazione negli alberi di ricerca potrebbe lasciare dei nodi quasi vuoti, causando uno spreco di spazio ed un aumento del numero di livelli.

# I B-Tree

- I B-Tree ed i B<sup>+</sup>-Tree sono **alberi di ricerca bilanciati**, progettati per ottimizzare operazioni su dischi magnetici o qualunque altro tipo di memoria secondaria ad accesso diretto.
- Alcuni vincoli posti sulla struttura assicurano che:
  1. L'albero è sempre bilanciato.
  2. Lo spreco di spazio nei nodi è *abbastanza* limitato.

# Definizione formale di B-Tree

- Dato un **B-Tree di ordine  $p$** :
  1. Ogni nodo interno ha la forma:  
 $\langle P_1, \langle K_1, Pr_1 \rangle, P_2, \langle K_2, Pr_2 \rangle, \dots, P_{q-1}, \langle K_{q-1}, Pr_{q-1} \rangle, P_q \rangle$  con  $q \leq p$ ,
    - ▣ ogni  $P_i$  è un puntatore ad un albero (*un altro nodo nel B-Tree*).
    - ▣ ogni  $Pr_i$  è un *data pointer* (*puntatore al record con valore  $K_i$  del campo chiave di ricerca*).
  2. All'interno di ogni nodo deve valere che  
 $K_1 < K_2 < \dots < K_{q-1}$

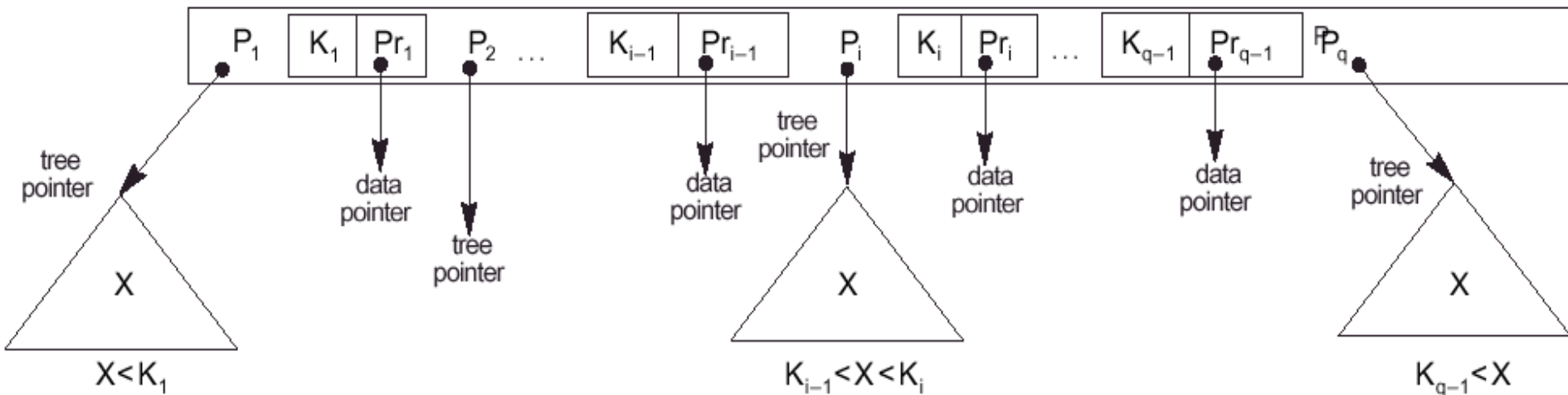
## Definizione formale di B-Tree (2)

3. Per ogni valore  $X$  del campo chiave di ricerca nel sottoalbero puntato da  $P_i$  vale:  
$$\begin{array}{ll} K_{i-1} < X < K_i & (\text{per } 1 < i < q), \\ X < K_i & (\text{per } i = 1), \\ K_{i-1} < X & (\text{per } i = q) \end{array}$$
4. Ogni nodo ha al più **p** puntatori ad albero.
5. Ogni nodo, tranne la radice ed i nodi foglia, ha almeno  $\lceil \mathbf{(p/2)} \rceil$  puntatori ad albero.  
Il nodo radice ne ha almeno **2**, a meno che non sia l'unico nodo nell'albero.

# Definizione formale di B-Tree (3)

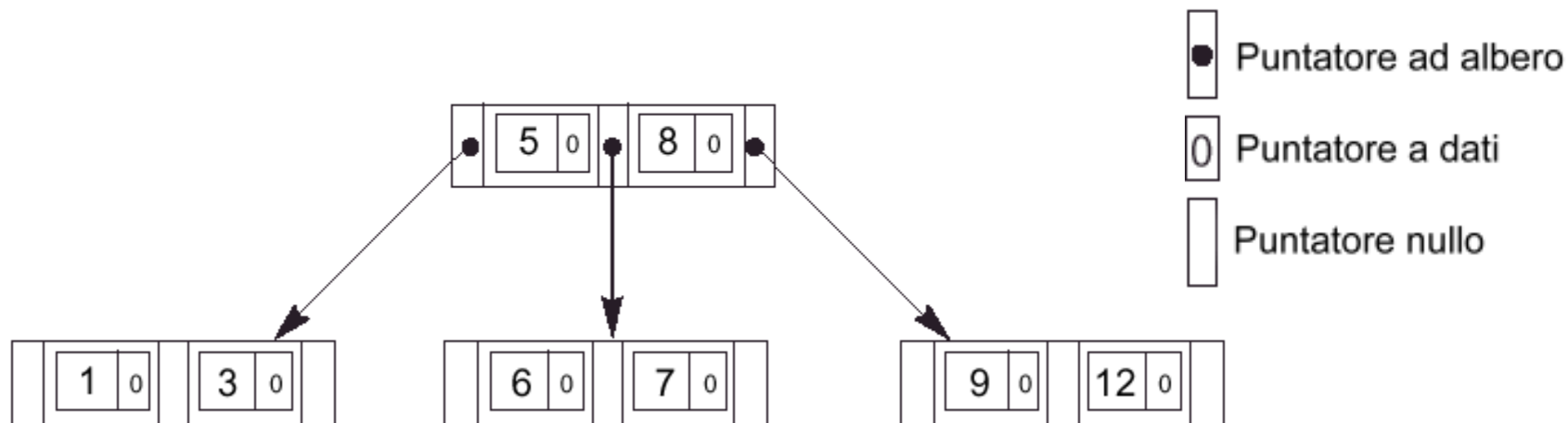
6. Un nodo con  $q$  ( $q \leq p$ ) puntatori ad albero ha  $q-1$  valori del campo chiave di ricerca (*e quindi  $q-1$  data pointer*).
7. Tutti i nodi foglia sono allo stesso livello.  
I nodi foglia hanno la stessa struttura dei nodi interni, tranne per i puntatori ad albero, che sono nulli.

# Esempio di nodo di B-Tree



- Un nodo di un B-Tree con  $q-1$  valori di ricerca.

# B-Tree: *Esempio*



- Un B-Tree di ordine  **$p = 3$** .
- Da osservare l'unicità del valore di ricerca poiché assumiamo che è una struttura di accesso sul campo chiave.
- Se il campo fosse non chiave si dovrebbe far puntare Pr ad una lista a puntatori aggiungendo un livello extra.



# Costruzione di un B-Tree

- Inizialmente un B-Tree ha un unico nodo radice (che è anche foglia).
- Se la radice contiene  $p-1$  valori di ricerca e si tenta di inserire un'altra entrata nell'albero, il nodo viene scisso in due nodi a livello 1.

*Nella radice resta solo il valore mediano mentre gli altri valori vengono distribuiti equamente tra i due nuovi nodi.*

# Costruzione di un B-Tree (2)

- Se un nodo non radice è pieno e vi si inserisce una nuova entrata, quel nodo è scisso in due nodi allo stesso livello e l'entrata mediana è spostata al nodo padre insieme a due puntatori ai nodi risultanti dalla scissione.

- ▣ Se anche il nodo padre è pieno, viene scisso anch'esso.

La scissione si può propagare fino alla radice, aggiungendo un nuovo livello ogni volta che la radice viene scissa.

# Costruzione di un B-Tree (3)

- Se la cancellazione di un valore fa sì che un nodo diventi pieno per meno della metà, esso è combinato con suoi vicini.

Ciò può propagarsi fino alla radice, riducendo il numero di livelli dell'albero.

- *Osservazione:*
  - ▣ È stato mostrato tramite analisi e simulazione, che, dopo numerosi inserimenti e cancellazioni casuali su un B-Tree, quando il numero di valori nell'albero si stabilizza, i nodi sono pieni circa al 69%:
    - La scissione e la combinazione dei nodi accadranno solo raramente, rendendo inserimento e cancellazione molto efficienti.

# Esempio 1

- Calcolo dell'ordine  $p$  di un B-Tree su disco:
  - ▣ Dimensione blocchi su disco  $B = 512$  byte
  - ▣ Campo di ricerca di  $V = 9$  byte
  - ▣ Puntatore a record di  $Pr = 7$  byte
  - ▣ Puntatore a blocco di  $P = 6$  byte
- Ogni nodo del B-Tree, contenuto in un blocco del disco, può avere al più  $p$  puntatori ad albero,  $p-1$  puntatori a record e  $p-1$  valori del campo chiave di ricerca.

## Esempio 1 (2)

- Quindi deve essere:

$$(p*P)+((p-1)*(Pr+V)) \leq B \rightarrow (p*6)+((p-1)*(7+9)) \leq 512 \rightarrow \\ (22 * p) \leq 528 \rightarrow p \leq 24$$

- Scegliamo  $p = 23$ , poiché in ogni nodo ci potrebbe essere la necessità di memorizzare informazioni addizionali, quali il numero di entrate  $q$  o un puntatore al nodo padre.

## Esempio 2

- Calcolo del numero di blocchi in un B-Tree
  - ▣ Supponiamo di costruire un B-Tree sul campo di ricerca dell'esempio precedente, non-ordering e chiave. Supponiamo che ogni nodo sia pieno per il 69%.
  - ▣ Ogni nodo avrà in media  $p * 0.69 = 23 * 0.69 =$  circa 16 puntatori (*fo medio*) e quindi 15 valori del campo chiave di ricerca.

Partendo dalla radice vediamo quanti valori e puntatori esistono in media a ogni livello:

Radice:	1 nodo	15 entry	16 ptr
livello 1:	16 nodi	240 (16*15) entry	256 (16*16) ptr
livello 2:	256 nodi	3840 (256*15) entry	4096 ptr
livello 3:	4096 nodi	61440 (4096 *15) entry	-

## Esempio 2 (2)

- Per ogni livello:

$$\# \text{entrate livello } i\text{-mo} = [\# \text{puntatori livello } (i-1)\text{-mo}] * 15$$

- Per le date dimensioni di blocco, puntatore e campo chiave di ricerca, un B-tree a 3 livelli contiene fino a  $3840 + 240 + 15 = 4095$  entrate.
- Un B-tree a 4 livelli contiene fino a  $4095 + 61440 = 65535$  entrate.

# I B<sup>+</sup>-Tree

- I B<sup>+</sup>-Tree sono usati nella maggior parte delle implementazioni di indici multilivello dinamici.
- I puntatori ai dati sono memorizzati solo nei **nodi foglia**, che hanno quindi una struttura diversa da quella dei nodi interni:
  - ▣ Per un campo di ricerca chiave, i nodi foglia hanno un'entrata per ogni valore del campo, insieme ad un puntatore al record (o al blocco che lo contiene).
  - ▣ Per un campo non chiave, il puntatore referencia un blocco che contiene puntatori ai record del file dati, creando un ulteriore livello di indirezione.



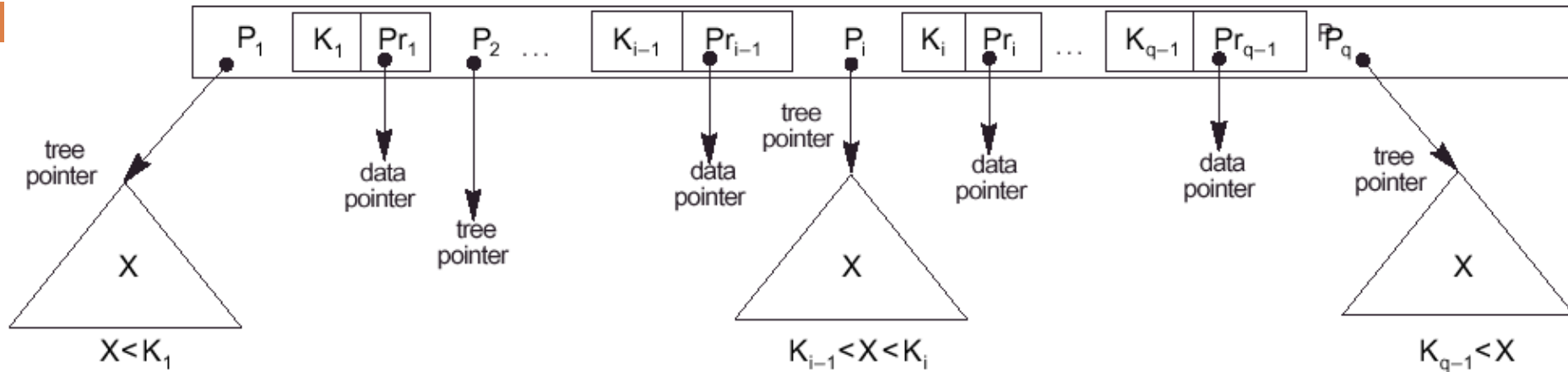
# I B<sup>+</sup>-Tree (2)

- I nodi foglia sono collegati da puntatori, così da fornire un accesso **ordinato** ai record sul campo chiave.

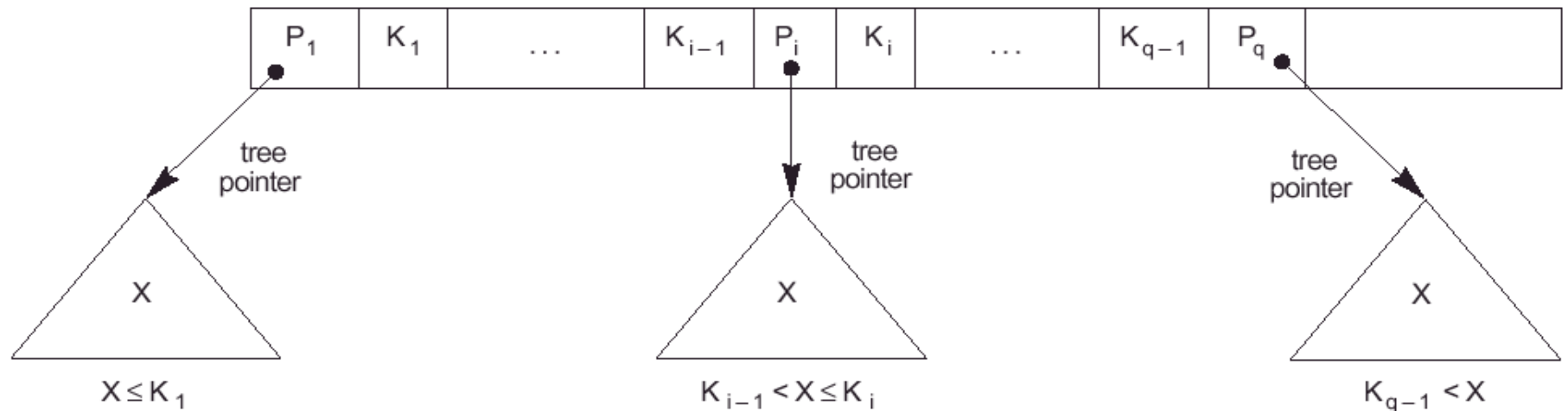
I nodi foglia sono simili al primo livello di un indice multi-livello.

- I nodi interni corrispondono agli altri livelli di un indice multi-livello.
  - ▣ Alcuni valori del campo di ricerca dei nodi foglia sono ripetuti in nodi interni del B<sup>+</sup>-Tree per guidare la ricerca.

# Confronto tra B-Tree e B<sup>+</sup>-Tree



a) Struttura di un nodo di un B-Tree



b) Struttura di un nodo di un B+-Tree

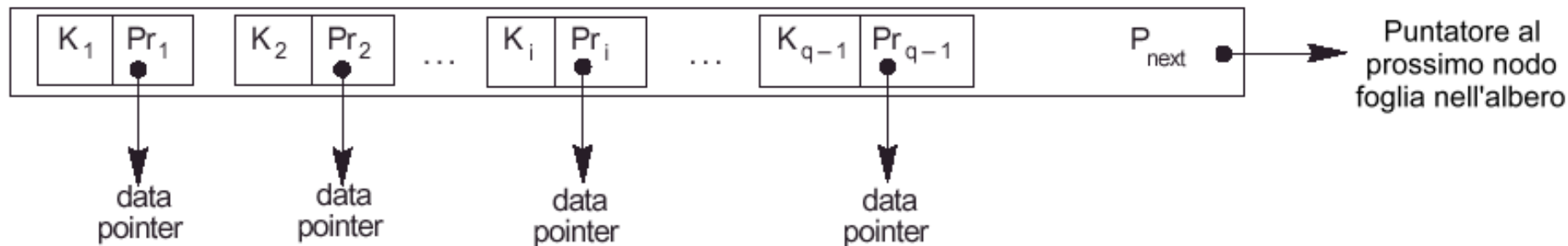
# Struttura dei nodi interni di un B<sup>+</sup>-Tree di ordine p

1. Ogni nodo interno ha la forma  $\langle P_1, K_1, P_2, K_2, \dots, P_{q-1}, K_{q-1}, P_q \rangle$ , con  $q \leq p$ 
  - ▣ ogni  $P_i$  è un puntatore ad albero.
2. In ogni nodo interno si ha  $K_1 < K_2 < \dots < K_{q-1}$
3. Per ogni valore  $X$  del campo di ricerca nel sottoalbero puntato da  $P_i$  vale che:
  - ▣  $K_{i-1} < X \leq K_i$  (per  $1 < i < q$ ),
  - ▣  $X \leq K_i$  (per  $i = 1$ ),
  - ▣  $K_{i-1} < X$  (per  $i = q$ ).

# Struttura dei nodi interni di un B<sup>+</sup>-Tree di ordine $p$ (2)

4. Ogni nodo interno ha al più  $p$  puntatori ad albero.
5. Ogni nodo interno, tranne la radice, ha almeno  $\lceil (p/2) \rceil$  puntatori ad albero. Il nodo radice ne ha almeno 2 se è un nodo interno.
6. Un nodo interno con  $q$  puntatori,  $q \leq p$ , ha  $q-1$  valori del campo di ricerca.

# Struttura dei nodi foglia di un B<sup>+</sup>-Tree di ordine p



1. Ogni nodo foglia ha la forma  
 $\langle \langle K_1, Pr_1 \rangle, \langle K_2, Pr_2 \rangle, \dots, \langle K_{q-1}, Pr_{q-1} \rangle, P_{next} \rangle$ ,  
con  $q \leq p$ :
  - ▣ Ogni  $Pr_i$  è un data pointer.
  - ▣  $P_{next}$  è un puntatore al prossimo nodo foglia del B<sup>+</sup>-Tree.
2. In ogni nodo foglia si ha  $K_1 < K_2 < \dots < K_{q-1}$ ,  
con  $q \leq p$ .

# Struttura dei nodi foglia di un B<sup>+</sup>-Tree di ordine p (2)

3. Ogni  $Pr_i$  è un data pointer, che referencia:
  - ▣ Il record con valore  $K_i$  nel campo di ricerca oppure
  - ▣ il blocco del file contenente tale record, oppure
  - ▣ un blocco di puntatori a record che referenziano i record con valore del campo di ricerca  $K_i$ , se il campo di ricerca è non chiave.
4. Ogni nodo foglia ha almeno  $\lceil (P_{leaf} / 2) \rceil$  valori.
5. Tutti i nodi foglia sono allo stesso livello.

# Visita di un B<sup>+</sup>-Tree

- Partendo dalla foglia più a sinistra, è possibile visitare tutti i nodi foglia come una lista a puntatori, usando **P<sub>next</sub>**.  
Ciò fornisce un accesso ai record di dati ordinato sul campo di indicizzazione.
- Se il campo di ricerca non è chiave, è necessario un ulteriore livello di indirizzione, simile a quello della struttura ad indice secondario con campo non chiave.

# Ordine di B<sup>+</sup>-Tree e B-Tree

- Un nodo interno di un B<sup>+</sup>-Tree contiene più entrate di un nodo interno di un B-Tree, poiché nel B<sup>+</sup>-Tree tali nodi hanno solo valori di ricerca e puntatori ad albero.
- Di conseguenza, a parità di dimensione di blocco, l'ordine **p** sarà maggiore per un B<sup>+</sup>-Tree che per un B-Tree.
- Poiché le strutture di nodi interni e di nodi foglia sono diverse, l'ordine **p** può essere diverso nei due tipi di nodo:
  - **P<sub>leaf</sub>** indica il numero massimo di data pointers in un nodo foglia.



# Esempio

- Calcolo dell'ordine  $p$  di un B<sup>+</sup>-Tree su disco:
  - ▣ Dimensione blocchi su disco  $B = 512$  byte
  - ▣ Campo di ricerca di  $V = 9$  byte
  - ▣ Puntatore a record di  $P_r = 7$  byte
  - ▣ Puntatore a blocco di  $P = 6$  byte
- Un nodo interno di un B<sup>+</sup>-Tree, contenuto in un singolo blocco, può avere fino a  $p$  puntatori ad albero e  $p-1$  valori del campo di ricerca.

Quindi deve essere:

$$(p \cdot P) + ((p-1) \cdot V) \leq B \rightarrow (p \cdot 6) + ((p-1) \cdot 9) \leq 512 \rightarrow (15 \cdot p) \leq 512$$

## Esempio (2)

- Il massimo intero che soddisfa la disuguaglianza è  $p=34$ , che è maggiore di 23 (ordine del B-Tree corrispondente).

Quindi ne risulta un fan-out maggiore e più entrate in ogni nodo interno.

- L'ordine  $P_{\text{leaf}}$  dei nodi foglia è:

$$(P_{\text{leaf}} * (Pr + V)) + P \text{ (per } P_{\text{next}}) \leq B \rightarrow$$

$$(P_{\text{leaf}} * (7 + 9)) + 6 \leq 512 \rightarrow$$

$$(P_{\text{leaf}} * 16) \leq 506$$

- Quindi ogni nodo foglia può tenere fino a

$P_{\text{leaf}} = 31$  combinazioni valore chiave/data pointer, supponendo che i data pointer siano puntatori a record.

## Esempio 2

- Calcolo del numero di blocchi e di livelli in un B<sup>+</sup>-Tree:
  - ▣ Supponiamo che ogni nodo del B<sup>+</sup>-Tree sia pieno per il 69%.
  - ▣ Ogni nodo interno avrà in media  
 $p * 0.69 = 34 * 0.69 = \text{circa } 23 \text{ puntatori (fo medio) e}$   
quindi 22 valori del campo di ricerca.
  - Ogni nodo foglia avrà in media  
 $P_{\text{leaf}} * 0.69 = 31 * 0.69 = \text{circa } 21 \text{ puntatori a record dati.}$

## Esempio 2 (2)

- Partendo dalla radice vediamo quanti valori e puntatori esistono in media a ogni livello:

Radice:	1 nodo	22 entry	23 ptr
livello 1:	23 nodi	506 (22*23) entry	529 (23*23) ptr
livello 2:	529 nodi	11638 (22*529) entry	12167 ptr
livello foglie:	12167 nodi	255507 (21 *12167) puntatori a record	

- Quindi un B<sup>+</sup>-Tree di quattro livelli tiene fino a 255.507 puntatori a record, contro i 65.535 calcolati per il B-Tree corrispondente.

# Inserimento in un B<sup>+</sup>-Tree

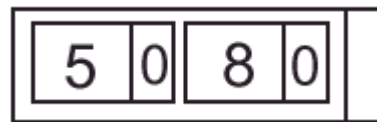
- Inizialmente la radice è l'unico nodo dell'albero:
  - ▣ essendo quindi un nodo foglia conterrà anche i puntatori ai dati.
- Se si cerca di inserire un'entrata in un nodo foglia pieno, il nodo va in *overflow* e deve essere scisso:
  - ▣ le prime  $j = \lceil (P_{\text{leaf}} + 1)/2 \rceil$  entrate sono mantenute nel nodo, mentre le rimanenti sono spostate in un nuovo nodo foglia.
  - ▣ Il j-esimo valore di ricerca (quello mediano) è replicato nel nodo padre.
  - ▣ Nel padre viene creato un puntatore al nuovo nodo.

# Inserimento in un B<sup>+</sup>-Tree (2)

- Se il nodo padre (interno) è pieno si ha un altro *overflow*:
  - ▣ Il nodo viene diviso, creando un nuovo nodo interno.
  - ▣ Le entrate nel nodo interno fino a  $P_j$  (con  $j = \lfloor ((p+1)/2) \rfloor$ ) sono mantenute nel nodo originario, mentre il  $j$ -mo valore di ricerca è spostato al padre, non replicato.
  - ▣ Il nuovo nodo interno conterrà le entrate dalla  $P_{j+1}$  alla fine delle entrate del nodo originario.
- Tale scissione si può propagare verso l'alto fino a creare un nuovo livello per il B<sup>+</sup>-Tree.

# Esempio di inserimento

- Vogliamo inserire dei record in un B<sup>+</sup>-Tree di ordine  $P = 3$  e  $P_{leaf} = 2$ , nella sequenza: 8, 5, 1, 7, 3, 12, 9, 6.



- Tree node pointer

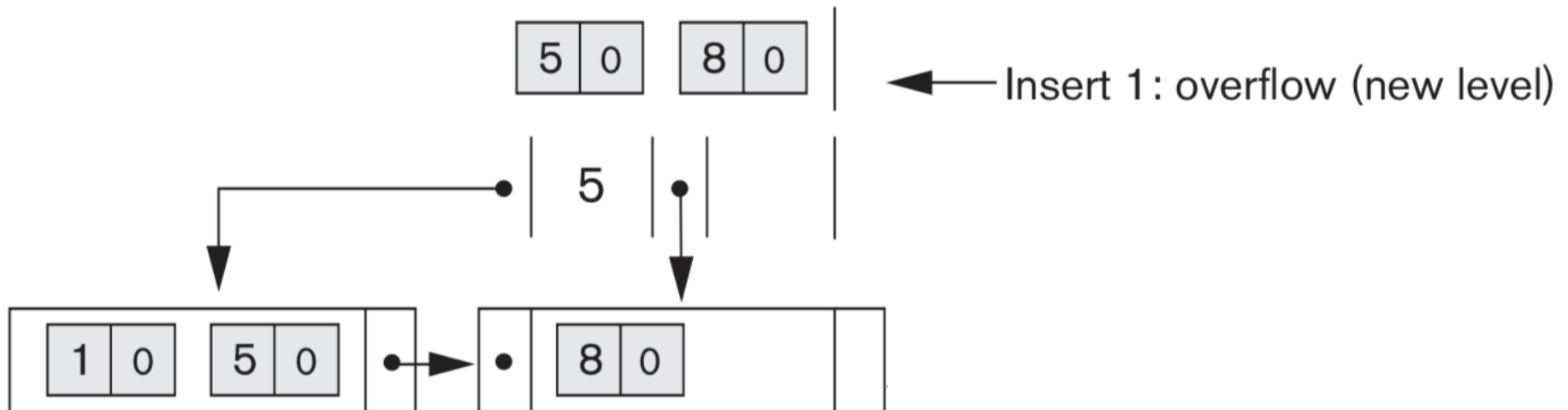
0 Data pointer

 Null tree pointer

- L'inserimento dei valori 8 e 5 non provoca overflow: sono entrambi inseriti nella radice.

## Esempio di inserimento (2)

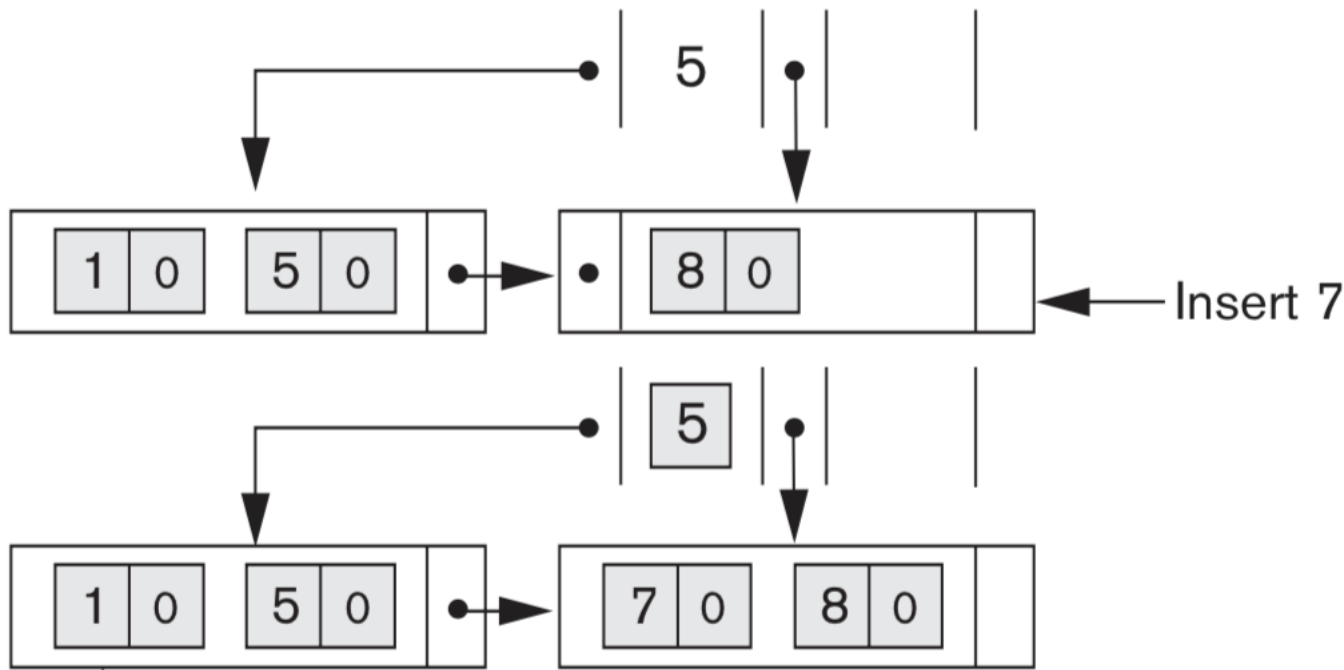
- L'inserimento di 1 provoca overflow. Il nodo è scisso ed il valore mediano è ripetuto in un nuovo nodo radice.
- Alcuni valori sono replicati nei nodi interni per guidare la ricerca.



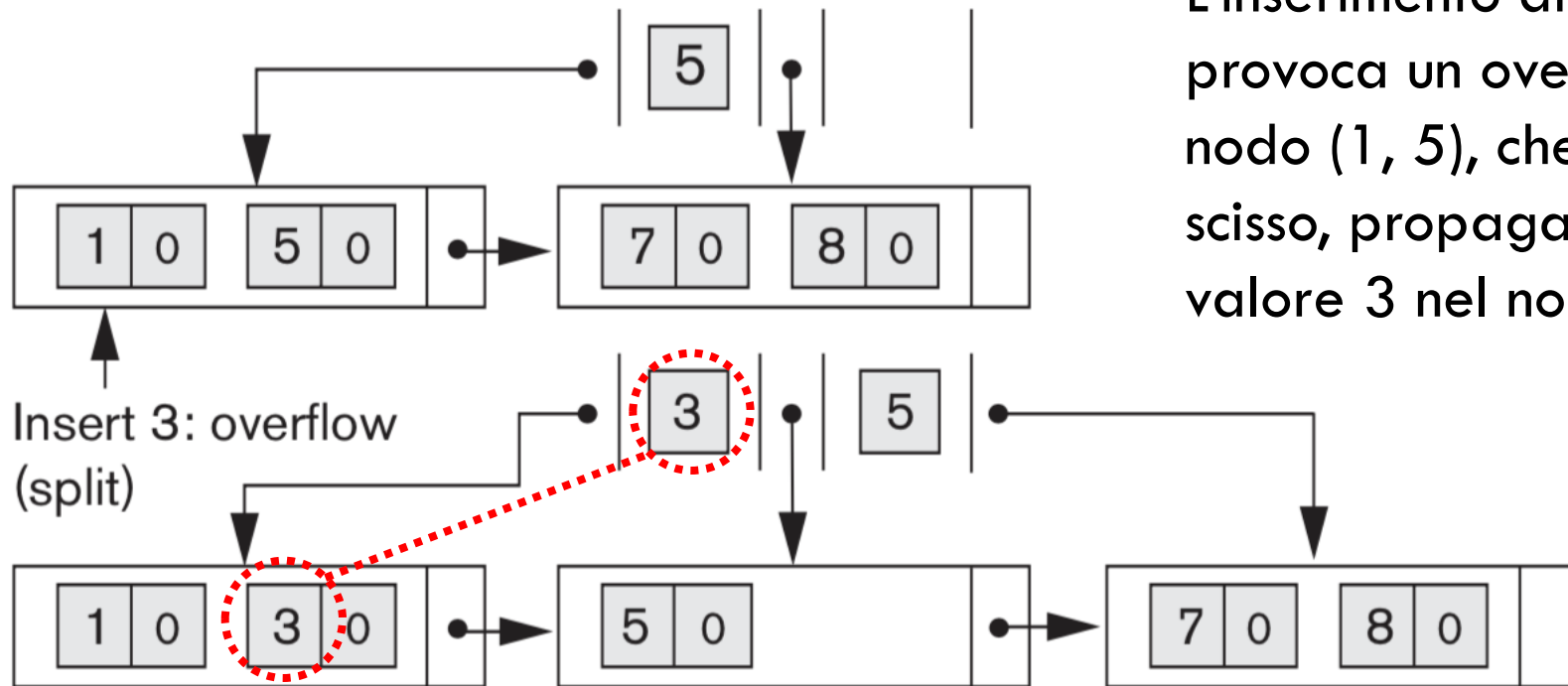


## Esempio di inserimento (3)

- L'inserimento di 7 non provoca overflow.
- Si noti che tutti i valori sono a livello foglia, perché i data pointer sono tutti a quel livello.



# Esempio di inserimento (4)

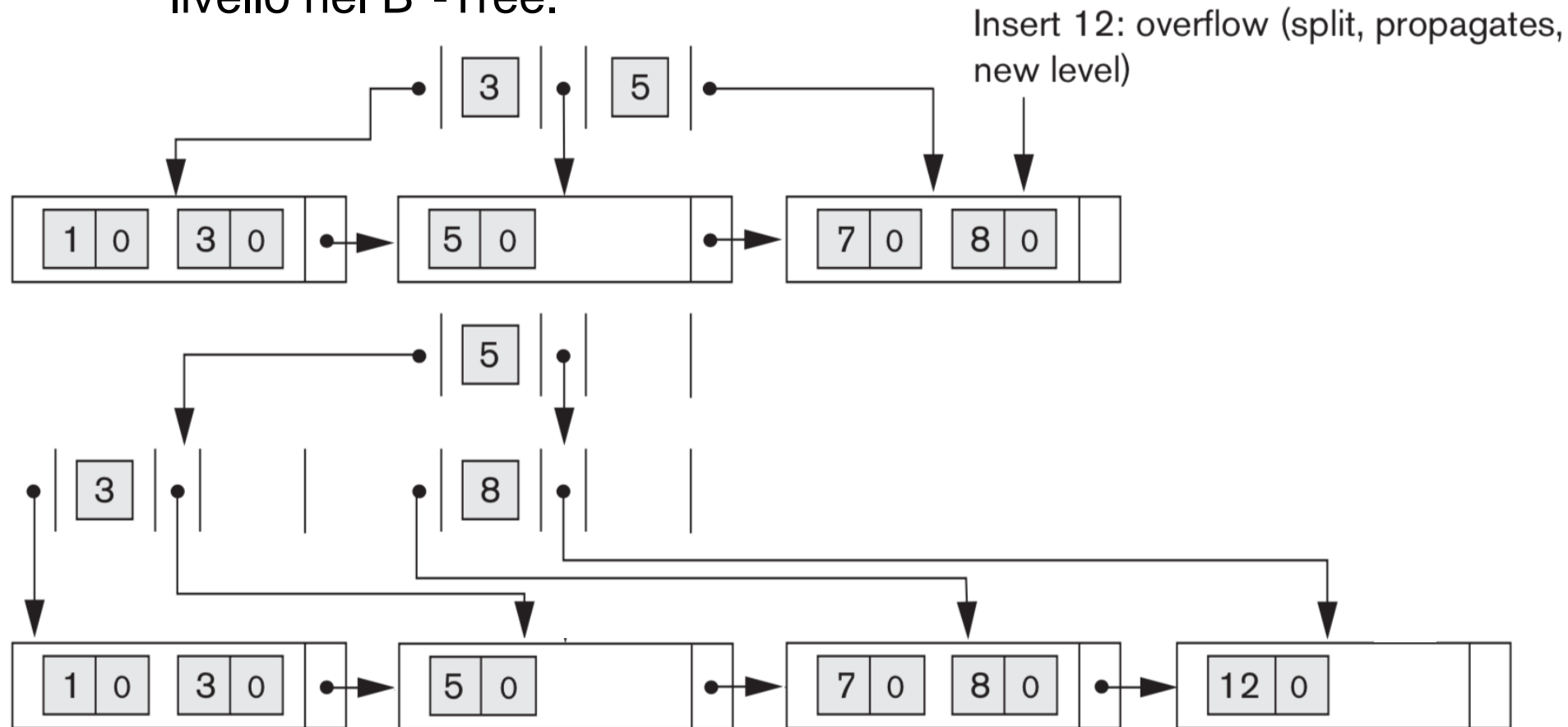


L'inserimento di 3  
provoca un overflow nel  
nodo (1, 5), che viene  
scisso, propagando il  
valore 3 nel nodo padre.

- Si noti che un valore che compare nel nodo interno, compare anche come valore più a destra del sottoalbero referenziato dal puntatore alla sinistra di tale valore.

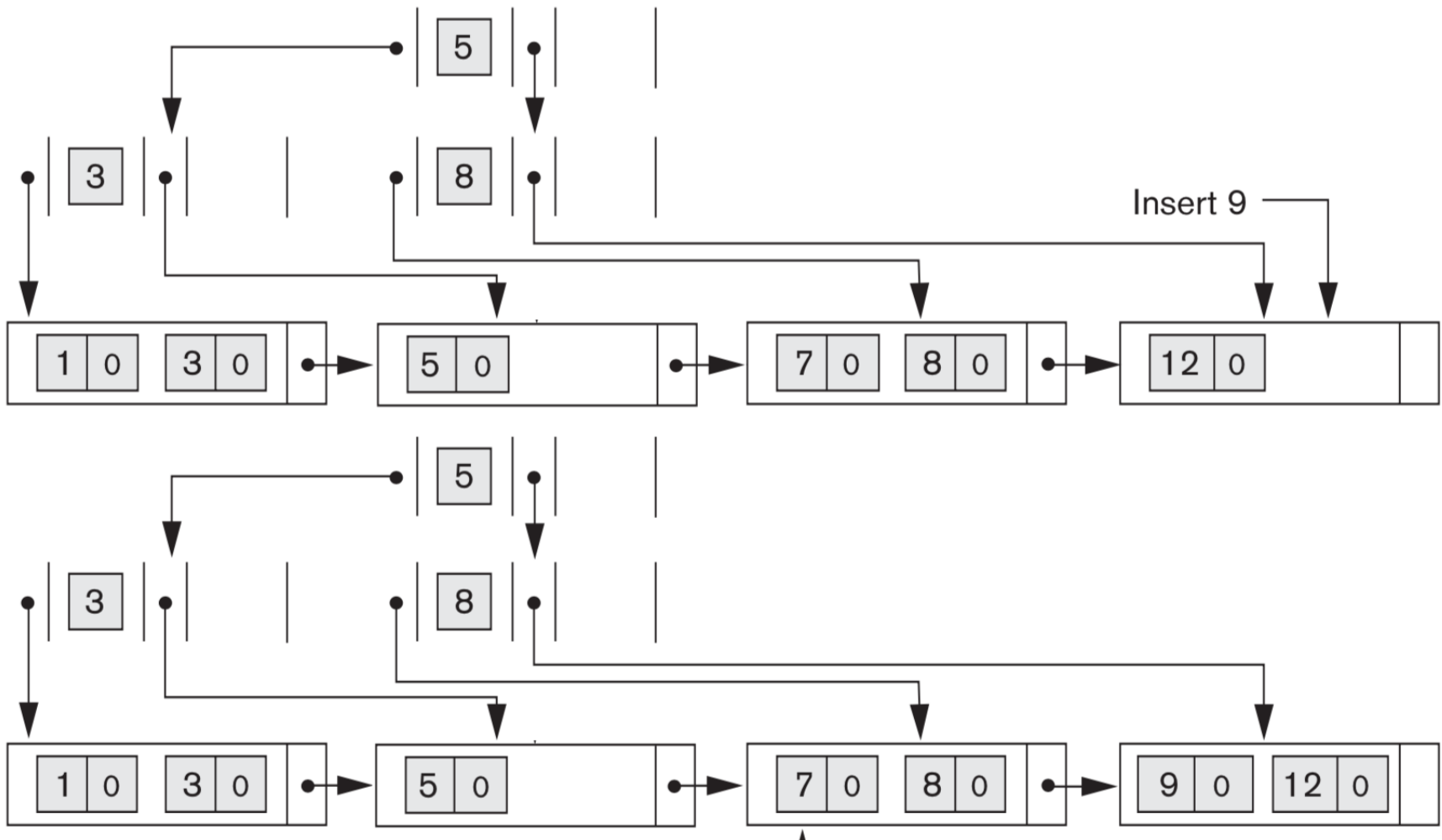
# Esempio di inserimento (5)

- L'inserimento di 12 provoca un overflow nel nodo (7, 8), che viene scisso, propagando il valore 8 nel nodo padre.
- La scissione si propaga fino alla radice, creando un nuovo livello nel B<sup>+</sup>-Tree.



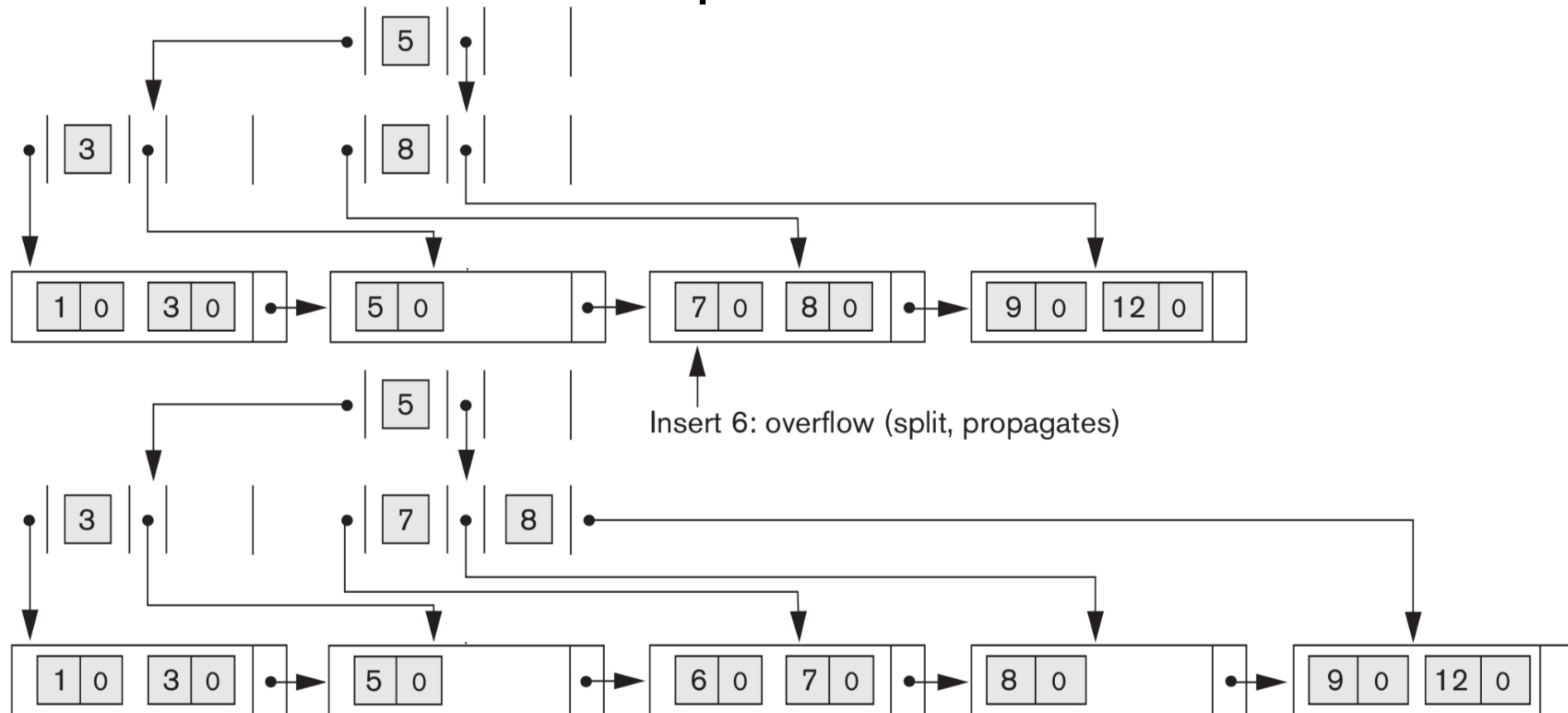
# Esempio di inserimento (6)

- L'inserimento del 9 non provoca overflow.



# Esempio di inserimento (7)

- L'inserimento di 6 provoca un overflow nel nodo (7, 8), che viene scisso, propagando il valore 7 nel nodo padre.



# Cancellazione in un B<sup>+</sup>-Tree

- Un'entrata è cancellata sempre a livello foglie.  
Se essa ricorre anche in un nodo interno, allora è rimossa anche da quello e al suo posto si inserisce il valore immediatamente alla sinistra del valore rimosso nel nodo foglia.
- La cancellazione di valori può causare l'**underflow** di un nodo, riducendo il numero di entrate in un nodo foglia per meno del minimo consentito.
  - ▣ In tal caso si cerca un fratello del nodo foglia in modo da ridistribuire le entrate, cosicché entrambi i nodi siano pieni almeno per metà;
  - ▣ Se ciò non è possibile si effettua il merge del nodo con i suoi fratelli, riducendo di uno il numero di nodi foglia.

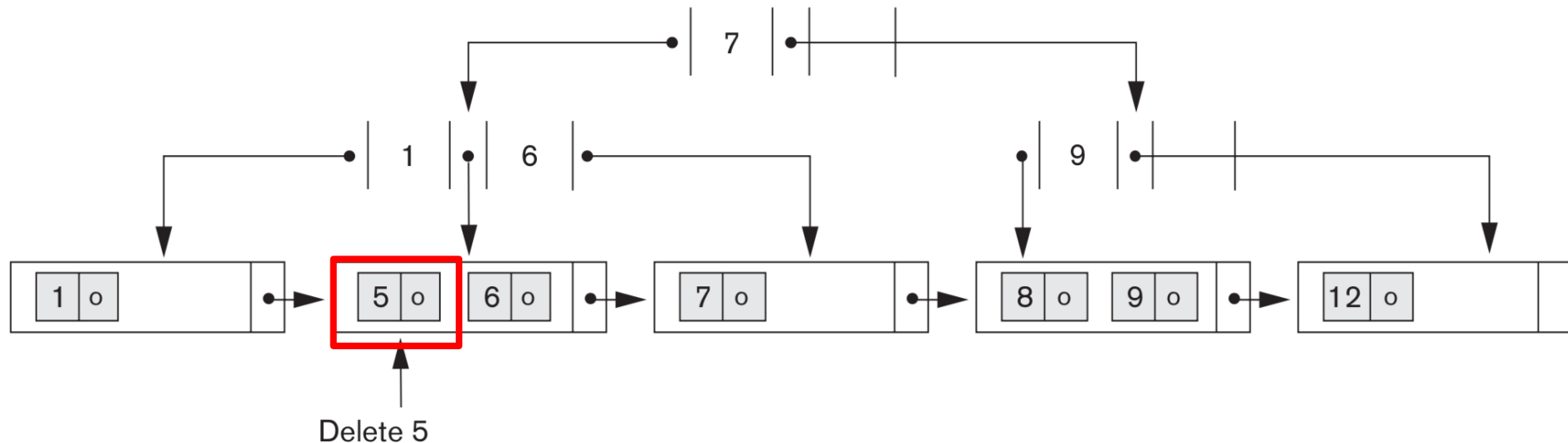
# Cancellazione in un B<sup>+</sup>-Tree (2)

## □ Approccio comune:

- ▣ Si tenta di ridistribuire le entrate con il fratello sinistro;
- ▣ Se non è possibile, si tenta di ridistribuire con il fratello destro.
- ▣ Altrimenti si fondono i tre nodi in due nodi foglia.  
L'*underflow* in questo caso si può trasmettere ai nodi interni, poiché sono necessari un puntatore ad albero ed un valore di ricerca in meno. Se la propagazione arriva alla radice, si riduce il numero di livelli dell'albero.

# Esempio di cancellazione

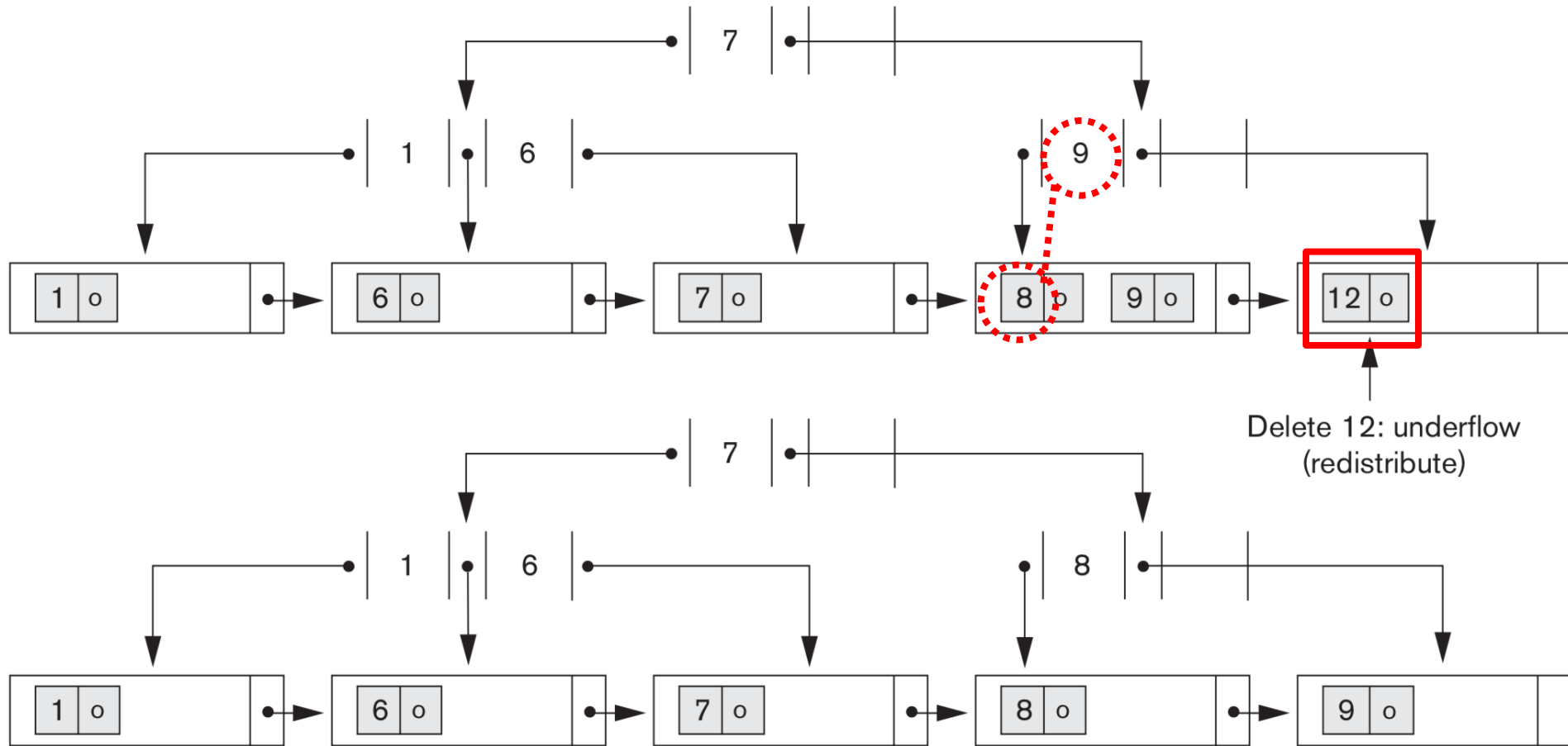
- Dato il seguente B<sup>+</sup>-Tree, di ordine  $P = 3$  e  $P_{leaf} = 2$ , vogliamo cancellare i record 5, 12, 9 e 6.



- La cancellazione di 5 non pone alcun problema.



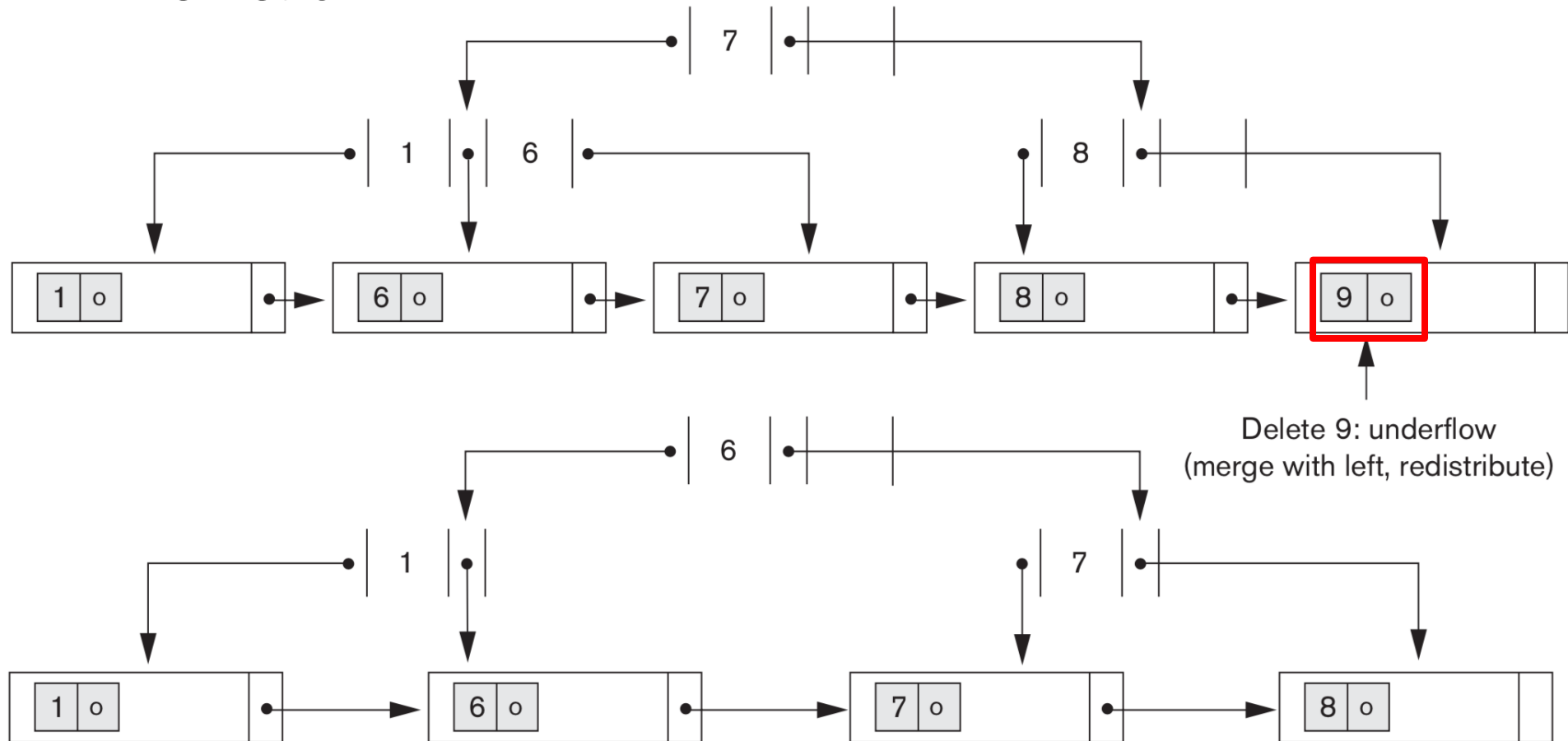
# Esempio di cancellazione (2)



- La cancellazione di 12 viene risolta con una redistribuzione, e l'aggiornamento del valore del nodo interno da 9 a 8.

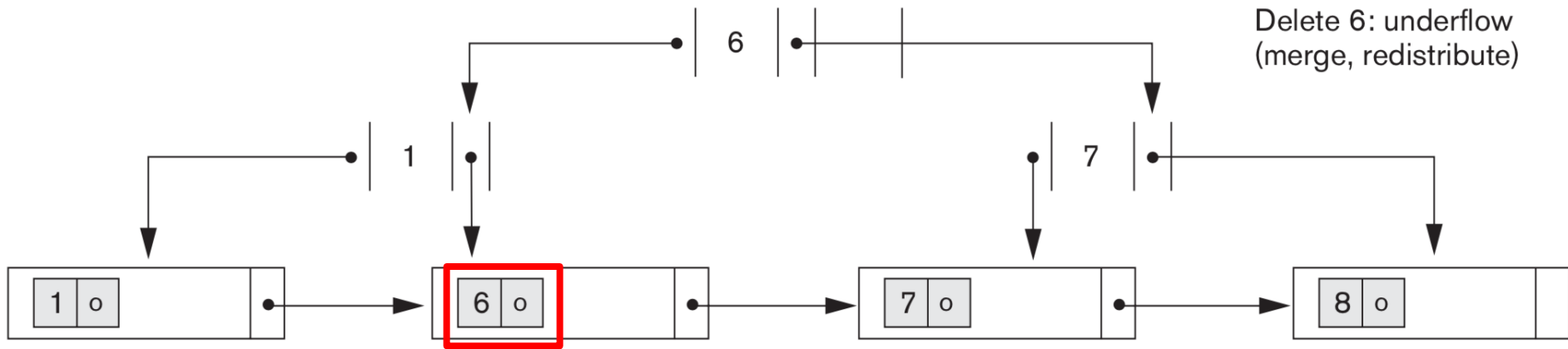
# Esempio di cancellazione (3)

- La cancellazione di 9 causa un *underflow*. I puntatori del nodo padre vengono redistribuiti con quelli del fratello sinistro.



# Esempio di cancellazione (4)

- La cancellazione di 6 causa un *underflow*. La redistribuzione dei nodi utilizzando i fratelli a destra e a sinistra causa un altro *underflow* nel nodo padre, che causa un *underflow* nel nodo radice, portando ad una riduzione del numero di livelli.



//  $B^+$ -Tree risultante

