Corso di Basi di Dati 2

Primo Compitino

20 Marzo 1998

1. (Indicativamente punti 10) Dato il seguente insieme di chiavi:

$$6, 7, 8, 5, 4, 14, 16, 23, 1, 2, 32, 18, 17, 20, 12, 13, 30, 10$$

   (a) mostrare il B-albero di ordine 5 ottenuto inserendo un elemento dopo l'altro nell'ordine dato;

   (b) mostrare il B-albero ottenuto dalla cancellazione dell'elemento 12 e poi dell'elemento 16 dal B-albero in (a);

   (c) mostrare il B^+-albero di ordine 4 ottenuto inserendo un elemento dopo l'altro nell'ordine dato;

   (d) mostrare il B^+-albero ottenuto dalla cancellazione dell'elemento 5 e poi dell'elemento 6 dal B^+-albero in (c);

   (e) qual è l'altezza minima e quale la massima di un B-albero di ordine 5 che contiene tali chiavi?

2. (Indicativamente punti 10)

   (a) Dato il seguente insieme di chiavi:

$$24, 18, 39, 15, 7, 13, 5, 69, 12, 8, 10, 23$$

      i. mostrare la struttura di hash virtuale, con funzioni $H_i(k) = k \ mod \ 5 * 2^i$ e capacità di ogni pagina 2, ottenuta inserendo gli elementi nell'ordine dato;

      ii. mostrare le operazioni effettuate per ricercare in tale struttura le chiavi 7 e 36.

   (b) Mostrare la struttura di hash estensibile, con capacità di ogni pagina 2, supponendo di inserire nell'ordine i record le cui pseudochiavi hanno i seguenti valori:

$$h(r_1) = 0001000 \quad h(r_2) = 0011000 \quad h(r_3) = 0010100 \quad h(r_4) = 1000000 \quad h(r_5) = 0100000 \quad h(r_6) = 1111000$$

3. (Indicativamente, punti 6) Sia dato il seguente schema relazionale: $R(A, B, C)$ dove A è un attributo di tipo intero. Sia data inoltre una relazione r di schema R contenente 3000 tuple memorizzate in 200 pagine. Si supponga che il range dei valori assunti dall'attributo A nelle tuple di r sia [2500...7500]. Si supponga inoltre di avere i seguenti indici sulla relazione r:

   - un indice I_1 clusterizzato sull'attributo A contenente 100 entrate e con 10 foglie;
   - un indice I_2 non clusterizzato sull'attributo B contenente 150 entrate e con 20 foglie.

   Data l'interrogazione

```
SELECT * FROM R WHERE A BETWEEN (2500,5000) AND B IN ("a","b")
```

   si valutino le possibili strategie per l'esecuzione dell'interrogazione (indicandone i relativi costi) e si determini la strategia ottima.

4. (Indicativamente punti 4) Siano date le relazioni R(A,B,C) ed S(D,E,F); si supponga inoltre che sia allocato un indice sull'attributo A della relazione R e un indice sull'attributo D della relazione S. Data l'interrogazione SELECT * FROM R, S WHERE D=d AND A=E si richiede di determinare tutte le possibili strategie di esecuzione (sulla base dell'ottimizzatore del System R). Per ogni strategia di esecuzione determinare inoltre se le tuple restituite sono ordinate e, se sono ordinate, in base a quali attributi.