

Basi di Dati: Strutture ed Algoritmi

Appelli del 2001

Appello del 15.1.2001

1. Si considerino la base di dati:

Studenti(Matricola, Nome, Area, Altro)
Frequenze(Matricola, Codice, Semestre)
Corsi(Codice, Titolo, Crediti)

le cui chiavi primarie sono Matricola per Studenti, (Matricola, Codice) per Frequenze e Codice per Corsi, e l'interrogazione:

```
SELECT      DISTINCT Nome
FROM        Studenti, Corsi, Frequenze
WHERE       Frequenze.Matricola = Studenti.Matricola
            AND Frequenze.Codice = Corsi.Codice
            AND Area = Software AND Crediti = 12 AND Semestre = 'Primo99'
```

Si supponga che esista un indice denso per ogni chiave primaria, di ordinamento per Studenti e Corsi.

- Dare l'albero logico iniziale dell'interrogazione.
- Dare l'albero fisico che secondo voi potrebbe essere prodotto dall'ottimizzatore usando solo i metodi di giunzione *nested loop* e *merge-join*.
- Dare l'albero fisico che secondo voi potrebbe essere prodotto dall'ottimizzatore usando solo i metodi di giunzione *nested loop* e *index nested loop*.
- Stimare il costo e la cardinalità del risultato di uno dei due piani di accesso a vostra scelta.

2. Si supponga che le transazioni siano gestite con l'algoritmo disfare/rifare, la concorrenza con il 2PL-stretto e il punto di allineamento con il metodo buffer consistent (versione 1). Si supponga che il giornale abbia il seguente contenuto quando si verifica un fallimento di sistema; i record di modifica sono (W, Tid, Variabile, VecchioValore, NuovoValore), supponendo per semplicità che le variabili siano pagine e il loro valore sia un intero:

```
( T1, begin )
( W, T1, X, 5, 10 )
  ( T2 begin )
    ( W, T2, X, 10, 20 )
  ( T1 commit )
    ( W, T2, Y, 30, 60 )
( CKP {T2} )
  ( W, T2, W, 35, 70 )
```

- a) Si dica se il giornale può avere il contenuto mostrato. Se la risposta è sì, mostrare la sequenza di azioni eseguite dalle transazioni. Se la risposta è no, spostare o rimuovere il primo record "impossibile" del giornale per avere un giornale corretto.
- b) Per il giornale considerato al punto (a), quali sono i valori di X, Y e W prima che inizi la procedura di ripristino.
3. Spiegate le differenze fra operatori logici, operatori fisici e operatori dei metodi di accesso, della macchina fisica.

Appello del 6.2.2001

1. Si considerino la relazione R(A, B, C, D) e l'interrogazione:

```
SELECT    A
FROM      R
WHERE     A > 5 AND B = 20
ORDER BY  A
```

Dare i piani di accesso e la stima dei loro costi nei seguenti casi:

- a) uso di un indice denso hash su B;
- b) uso di un indice denso hash su B e di un indice denso a B⁺-albero su A;
- c) uso di un indice denso a B⁺-albero su A;
- d) uso del solo indice denso a B⁺-albero su (A, C, B).

Nei casi (b) e (d) definite anche l'operatore fisico che occorre per eseguire il piano di accesso.

2. Per ognuno dei seguenti protocolli di blocco dei dati dire quali delle seguenti proprietà sono garantite: c-serializzabilità, ripristinabile, senza cascata di abort.

- a) Prima di una scrittura si prende un blocco esclusivo che si rilascia al commit. Non vengono mai presi blocchi condivisibili prima di una lettura;
- b) Come (a) ma si prendono i blocchi condivisibili che poi possono essere rilasciati in qualsiasi momento;
- c) Come (b), ma seguendo il protocollo 2PL;
- d) Come (b), ma seguendo il protocollo 2PL-stretto.

3. Spiegate le differenze fra gli operatori fisici usati nei piani di accesso per recuperare i record di una tabella e gli operatori dei metodi di accesso, della macchina fisica, ad una tabella.

Appello del 10.7.2001

1. Si consideri la giunzione naturale delle tabelle R(A, B, C), S(A, D, E) e T(D, F, G). Si supponga che le tre relazioni abbiano lo stesso N_{reg} e N_{pag} , con $N_{reg} \gg N_{pag}$ (ad es. $N_{reg} = 100 \cdot 000$, $N_{pag} = 1000$). Si considerino due piani di accesso per eseguire la giunzione, uno profondo a sinistra e l'altro profondo a destra usando solo l'operatore fisico *nested loop*. Si dia una stima del loro costo e si dica quale è di costo inferiore.

2. Si definiscano le nozioni di grafo di serializzabilità e di storia c-serializzabile. Si considerino le storie:

$$S_1 = \{r_1[a], w_1[b], r_2[b], w_2[c], r_3[c], w_3[a]\}$$

$$S_2 = \{r_1[a], r_2[a], w_1[b], w_2[b], r_1[b], r_2[b], w_2[c], w_1[d]\}$$

Si dica se sono c-serializzabili. Se lo sono, si dica quali sono le storie seriali equivalenti.

3. Spiegare come viene eseguito dal gestore del buffer il comando di lettura di una pagina. Quando il gestore del buffer scrive una pagina nella memoria permanente?

Appello del 19.9.2001

1. Si consideri la tabella:

Familiari(CapoFamiglia, Nome, AnnoNascita, Parentela), con indice denso a B⁺-albero sulla chiave.

Si dia un piano di accesso e la stima del costo per l'interrogazione:

```
SELECT    CapoFamiglia, Nome
FROM      Familiari
WHERE     CapoFamiglia <=3
ORDER BY  CapoFamiglia desc;
```

Si supponga di disporre dell'operatore fisico `IndexOnlyFilter(R, Idx, {Ai}, C)` per la restrizione di R con l'uso del solo indice `Idx`. L'operatore ritorna i valori degli `{Ai}` di ogni record di R che soddisfa `C`, con `{Ai}` un sottoinsieme degli attributi sui quali è definito l'indice. Si dia un piano di accesso per l'interrogazione che usi l'`IndexOnlyFilter` e la stima del costo.

Si dica come cambia il piano d'accesso se l'interrogazione fosse

```
SELECT    CapoFamiglia, Nome, AnnoNascita
FROM      Familiari
WHERE     CapoFamiglia <=3
ORDER BY  CapoFamiglia desc;
```

2. Si consideri le transazioni:

$$T_1 = \{r_1[A], w_1[B]\}$$

$$T_2 = \{w_2[A], r_2[B]\}$$

quali storie c-serializzabili si possono definire su T_1 e T_2 ?

3. Spiegare come vengono eseguite le operazioni di lettura, scrittura, abort e commit in un sistema in cui le transazioni sono gestite con tecnica *non disfare-non rifare*.