

# Tecnologia delle basi di dati (ex Basi di dati, primo modulo)

## 14 settembre 2007

**Tempo a disposizione:** un'ora e quarantacinque minuti.

**Nota:** è richiesta una “bella copia” comprensibile e ordinata.

**Domanda 1** (20%) Dimostrare che due schedule conflict-equivalenti hanno lo stesso grafo dei conflitti, mentre il viceversa non è in generale vero.

**Domanda 2** (20%) Si considerino un sistema con blocchi di dimensione  $B = 1000$  byte e puntatori ai blocchi di  $p = 5$  byte e una base di dati sulle seguenti relazioni, ognuna delle quali (i) ha una struttura heap (ii) ha un indice secondario sulla chiave (iii) non ammette valori nulli

- $R_1(\underline{ABC})$ , contenente  $S_1 = 2.000.000$  ennuple di  $l_1 = 40$  byte, di cui  $l_A = 5$  per il campo chiave  $A$ ; su questa relazione è definito un vincolo di integrità referenziale fra l'attributo  $B$  e la chiave  $D$  della relazione  $R_2$ ;
- $R_2(\underline{DEF})$ , contenente  $S_2 = 800.000$  ennuple di  $l_2 = 100$  byte, di cui  $l_D = 4$  per il campo chiave  $D$
- $R_3(\underline{GHL})$ , contenente  $S_3 = 500.000$  ennuple di  $l_3 = 25$  byte, di cui  $l_G = 5$  per il campo chiave  $G$

e con una vista definita come segue:

- `CREATE VIEW V AS SELECT * FROM (R1 JOIN R2 ON B=D) JOIN R3 ON C=G`

In tale contesto,

- mostrare un possibile piano di esecuzione (in termini di operatori dell'algebra relazionale e loro realizzazioni) per ciascuna delle seguenti interrogazioni
  1. `SELECT A, L FROM V`
  2. `SELECT A FROM V`
  3. `SELECT A, E FROM V`
- stimare il costo, in termini di numero di accessi a memoria secondaria (ignorando la presenza di eventuali buffer) per l'operazione 1.

**Domanda 3** (20%) Si supponga di dover realizzare un data mart relativo a squadre di calcio e giocatori. Fra i vari dati di interesse vi è l'età dei giocatori. Spiegare quali differenze possono sussistere nel data mart fra il caso in cui, per l'età dei giocatori, interessa ragionare sull'età media e quello in cui interessa ragionare sulle distribuzioni per fasce di età (cioè, ad esempio, quanti giocatori fra i 18 e i 20 anni, quanti fra i 21 e 23 e così via).

**Domanda 4** (20%) Si consideri un B-tree con nodi intermedi che contengono due chiavi e tre puntatori e foglie con due chiavi. Mostrare un possibile contenuto della struttura a seguito di inserimenti delle chiavi nel seguente ordine (a partire dall'albero vuoto): 2, 7, 15, 20, 17, 19, 16, 18, 5, 1. Mostrare anche i passi salienti che portano a tale contenuto.

**Domanda 5** (20%) Spiegare perché, scrivendo un programma che accede a due basi di dati diverse utilizzando JDBC, non è possibile garantire il commit a due fasi. Indicare quali funzionalità aggiuntive sono necessarie per i server locali che vogliono realizzare tale servizio e come si può procedere con un sistema che non disponga di tali funzionalità.