

Basi di dati, primo modulo — 17 settembre 2003

Tempo a disposizione: un'ora e quarantacinque minuti. Libri chiusi.

Domanda 1 (35%) Considerare la seguente interrogazione in SQL:

```
SELECT DISTINCT R1.A, R3.L
FROM R1, R2, R3
WHERE R1.C = R2.E AND R2.D=R3.H AND R1.B=R3.G AND R1.C=10
```

con riferimento ad una base di dati su cui sono note le seguenti informazioni:

- la relazione $R_1(\underline{ABC})$ ha 100.000 ennuple, una struttura heap e un indice secondario sulla chiave A ; l'attributo C ha 100 valori diversi, distribuiti in modo uniforme; l'attributo B ha un vincolo di riferimento verso la chiave G di R_3
- la relazione $R_2(\underline{DEF})$ ha 300.000 ennuple, una struttura heap e un indice secondario sulla chiave D ;
- la relazione $R_3(\underline{GHL})$ ha 100.000 ennuple, una struttura heap e un indice secondario sulla chiave G ; l'attributo H ha un vincolo di riferimento verso la chiave D di R_2

In tale contesto:

1. Mostrare un possibile piano di esecuzione (in termini di operatori dell'algebra relazionale e loro realizzazioni; giustificare brevemente le scelte più significative.
2. Stimare, nella misura in cui ciò sia possibile, le dimensioni dei risultati intermedi e finale.
3. Commentare sulle esigenze implementative imposte, in questo caso, dalla parola chiave `DISTINCT`.

Domanda 2 (20%) Sia data una relazione `Vendite(Libro,Libreria,Data,Quantità)` su cui si eseguono interrogazioni volte a calcolare il totale delle vendite in un certo periodo. Indicare (con un breve commento, non più di tre righe) quale livello di isolamento (`READ UNCOMMITTED`, `READ COMMITTED`, `REPEATABLE READ` o `SERIALIZABLE`) si potrebbe scegliere in ciascuno dei seguenti casi

1. l'interrogazione è eseguita, con lo scopo di stilare la classifica delle vendite, mentre vengono corretti i dati di alcune librerie, ma senza aggiungerne di nuovi;
2. l'interrogazione è eseguita, con lo scopo di stilare la classifica delle vendite, mentre vengono inseriti i dati di alcune librerie;
3. l'interrogazione è eseguita con lo scopo di conoscere indicativamente il numero di copie vendute di ciascun libro, mentre vengono effettuati alcuni inserimenti, relativi ad un insieme di piccole librerie (cioè con poche vendite);
4. l'interrogazione è eseguita in un tempo morto (cioè in assenza di aggiornamenti).

Domanda 3 (25%) Illustrare brevemente un possibile uso di regole attive per la gestione di dati derivati, per l'aggiornamento del numero di prenotati per un esame a fronte dell'inserimento di nuove prenotazioni o di annullamenti di prenotazioni. Assumere che tanto le prenotazioni quanto gli annullamenti siano gestiti attraverso inserimenti di ennuple in opportune relazioni; si consideri cioè una base di dati con tre relazioni:

- *Prenotazioni(Studiante, Corso, DataPrenotazione)*
- *Annullamenti(Studiante, Corso, DataAnnullamento)*
- *Totali(Corso, NumeroPrenotati)*

Domanda 4 (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 uno strumento di middleware e per server locali che vogliono realizzare tale servizio.