Corso di Basi di dati

Secondo Compitino - Fila B

19 maggio 2004

1. Si consideri il seguente schema relazionale, relativo a una base di dati per la gestione degli spot pubblicitari proposti da varie emittenti televisive.

   $EMITTENTE(\underline{CodE}, NomeE, Tipo)$
   $SPOT(\underline{CodS}, NomeS, Prodotto, Durata)$
   $TRASMETTE(\underline{CodE^*, CodS^*, DataIn}, DataF, Freq)$

   (a) Fomulare in SQL le seguenti interrogazioni
      i. Determinare i prodotti di cui esistono almeno due spot.
      ii. Determinare gli spot di durata superiore alla media degli spot dello stesso prodotto.
      iii. Determinare tra gli spot della pasta Barilla che sono stati trasmessi almeno 1000 volte quello piú lungo.
      iv. Determinare i prodotti che sono stati reclamizzati almeno 10 volte al giorno su tutte le emittenti regionali.

   (b) Definire una vista che associ ad ogni emittente il suo tipo ed alcune informazioni riepilogative relative al 2003, in particolare il numero di messaggi pubblicitari trasmessi, la durata totale di tali messaggi, il numero di prodotti diversi reclamizzati e il prodotto reclamizzato nello spot piú lungo tramesso dall'emittente.
      i. Quali operazioni si possono effettuare sulla vista?
      ii. É possibile e come autorizzare un utente Bob
         A. ad accedere a tutte le informazioni contenute nella vista tranne il prodotto reclamizzato nello spot piú lungo
         B. ad accedere a tutte le informazioni contenute nella vista solo relativamente alle emittenti regionali
         C. ad inserire le informazioni statistiche relative a una nuova emittente regionale

2. Relativamente allo schema relazionale di 1. si consideri la relazione

   $SPOT(\underline{CodS}, NomeS, Prodotto, Durata)$

   in cui vengono memorizzati gli spot pubblicitari. Ogni spot, identificato da un certo nome, reclamizza un solo prodotto. Di ogni spot possono esistere piú versioni, identificate da un codice, ognuna con una durata diversa. Possono esistere piú spot dello stesso prodotto, e piú spot con la stessa durata, ma non due versioni dello stesso spot con la stessa durata.

   (a) Formalizzare la situazione sopra descritta tramite dipendenze funzionali.
   (b) Individuare le chiavi della relazione.
   (c) Discutere se la relazione é in 3NF e in BCNF.

3. Relativamente allo schema relazionale $R(X, Y, Z, W, S, T)$ con le seguenti dipendenze funzionali

   $XYZ \to SW, XZW \to T, XW \to Y, Y \to Z, Z \to W$

   (a) Mostrare una relazione su tale schema che verifichi la dipendenza $XYZ \to SW$ e una che non la verifichi.

(b) Determinare le chiavi di R.

(c) Specificare se lo schema é in 3NF o in BCNF.

(d) Mostrare una scomposizione lossless join dello schema in BCNF e dire se preserva o meno le dipendenze.

(e) Mostrare una scomposizione non lossless join dello schema.

(f) Mostrare una scomposizione lossless join dello schema che preserva le dipendenze.

4. Relativamente allo schema relazionale $R(A, B, C, D)$ si supponga che l'attributo A sia chiave primaria della relazione e che la relazione sia clusterizzata sull'attributo C. La relazione contiene 5000 tuple su 50 blocchi e i valori distinti per gli attributi B, C, D sono rispettivamente 2, 50 e 1000.

   (a) Discutere su quali attributi sarebbe vantaggioso (in assenza di informazioni sul carico applicativo) definire degli indici e di che tipo.

   (b) Si supponga ora definito un indice sull'attributo C. Per ognuna delle seguenti operazioni specificare se la presenza dell'indice sarebbe utile/indifferente/dannosa per l'esecuzione dell'operazione, motivando brevemente la risposta. Se la risposta dipende dal tipo dell'indice, specificarlo.

      i. `INSERT INTO R VALUES (1,2,10,2)`

      ii. `UPDATE R SET D = 2 WHERE C = 10`

      iii. `UPDATE R SET C = C+10 WHERE D = 2`

      iv. `SELECT * FROM R WHERE C = 10`

      v. `SELECT * FROM R WHERE 0 < C < 30`

      vi. `SELECT * FROM R WHERE C <> 10`

      vii. `SELECT * FROM R WHERE C = 10 OR D = 2`

      viii. `SELECT * FROM R WHERE C = 10 AND D = 2`

      ix. `SELECT * FROM R WHERE D = 2 ORDER BY C`

      x. `SELECT DISTINCT C FROM R WHERE D = 2`