Compito di Complementi di Basi di Dati

4 aprile 2005

Anno Accademico 2004/2005

Esercizio 1:

Sia dato uno schema relazionale $R(A, B, C, D, E, F)$ con associato l'insieme di dipendenze funzionali $F = \{AC \rightarrow B, FB \rightarrow D, BD \rightarrow C, F \rightarrow E, E \rightarrow DF\}$.

(a) Determinare le chiavi candidate, gli attributi primi e gli attributi non primi di R.

(b) Stabilire se le scomposizioni $\rho = (ACDEF, ABCE)$ e $\rho' = (ABC, BCD, CDE, DEF, ABF)$ sono o meno scomposizioni lossless join che conservano le dipendenze.

(c) Stabilire se R è o meno in BCNF.

(d) Stabilire se R è o meno in 3NF.

(e) Stabilire se R è o meno in 2NF.

(f) Nel caso in cui R non sia in 3NF, fornire una scomposizione lossless join di R in 3NF che conservi le dipendenze.

Esercizio 2:

Con riferimento all'utilizzo del linguaggio SQL nei linguaggi di programmazione, si descrivano le principali caratteristiche dell'SQL statico e dell'SQL dinamico. In particolare, con riferimento all'SQL dinamico, si descrivano le modalità di esecuzione immediata e non immediata.

Esercizio 3:

Si descrivano le modalità di gestione della concorrenza in SQL:1999. In particolare, si descrivano i diversi possibili livelli di isolamento. Successivamente, si stabilisca se i seguenti schedule appartengono o meno a TS, 2PL e 2PL stretto:

1. $s_1 : r_3(y), r_1(x), w_1(y), w_5(x), w_1(t), w_4(x), r_2(z), r_3(z), w_5(z), w_2(z), r_5(t), r_4(t)$;
2. $s_2 : r_2(x), r_1(x), w_2(x), r_4(z), r_3(x), w_1(x), r_3(x), r_3(y), w_1(y), w_1(z), w_5(x), r_5(z), r_5(y)$.

Esercizio 4:

Si calcolino i profili dei risultati intermedi e del risultato finale (limitandosi alla cardinalità e alla dimensione delle tabelle e al numero di valori distinti di ciascuno dei loro attributi) della seguente interrogazione[1]:

[1] Si ricordi che la formula $col(n, m, k)$ che calcola il numero di valori distinti presenti in k oggetti estratti a partire da n oggetti diversi di m valori distinti, distribuiti omogeneamente, ammette la seguente approssimazione:

(a) $col(n, m, k) = k$ se $k \leq m/2$;

(b) $col(n, m, k) = (k + m)/3$ se $m/2 \leq k \leq 2m$;

(c) $col(n, m, k) = m$ se $k \geq 2m$.

$$\pi_{Nome,Voto}(Studenti \bowtie (\sigma_{Materia=\text{`}Genetica\text{'}}(ProveEsami))),$$

relativa alle tabelle:

$Studenti(Nome, \underline{Matricola}, Provincia, Nascita)$;
$ProveEsami(\underline{Materia}, \underline{Matricola}, Data, Voto)$,

aventi i seguenti profili:

$CARD(Studenti) = 750$, $SIZE(Studenti) = 35$,
$CARD(ProveEsami) = 45000$, $SIZE(ProveEsami) = 32$,
$SIZE(Nome, Studenti) = 18$, $SIZE(Matricola, Studenti) = 5$,
$SIZE(Provincia, Studenti) = 2$, $SIZE(Nascita, Studenti) = 10$,
$SIZE(Materia, ProveEsami) = 15$, $SIZE(Matricola, ProveEsami) = 5$,
$SIZE(Data, ProveEsami) = 10$, $SIZE(Voto, ProveEsami) = 2$,
$VAL(Nome, Studenti) = 700$, $VAL(Matricola, Studenti) = 750$,
$VAL(Provincia, Studenti) = 40$, $VAL(Nascita, Studenti) = 720$,
$VAL(Materia, ProveEsami) = 30$, $VAL(Matricola, ProveEsami) = 750$,
$VAL(Data, ProveEsami) = 250$, $VAL(Voto, ProveEsami) = 13$.

Anno Accademico 2003/2004

Esercizio 1:

Sia dato il seguente schema relazionale:

$Dipendente(CodiceDipendente, Nome, Cognome, CittàResidenza)$;

$Lavora(Dipendente, Banca, Stipendio)$;

$Banca(CodiceBanca, CittàFiliale)$;

$Supervisione(Supervisore, Dipendente)$.

Si assuma che ogni dipendente lavori per un'unica banca e abbia al più un supervisore, che ogni banca possa avere più filiali in città diverse (ma che in una specifica città una banca abbia al più una filiale) e che la relazione *Supervisione* leghi un dipendente al suo supervisore (a sua volta dipendente della banca).

Definire preliminarmente le chiavi primarie ed esterne delle relazioni date. Successivamente, formulare opportune interrogazioni in SQL che permettano di determinare (senza usare l'operatore CONTAINS e usando solo se necessario le funzioni aggregate):

(a) nome, cognome e città di residenza di tutti i dipendenti delle banche che hanno una filiale a Udine, ma non hanno una filiale a Gorizia;

(b) nome e cognome delle persone che risiedono in una città diversa da quella in cui risiede il loro supervisore;

(c) le banche che hanno solo dipendenti che risiedono in una delle città in cui è presente una loro filiale;

(d) i dipendenti il cui supervisore non ha un supervisore;

(e) per ogni città, la banca (le banche) con una filiale in città che ha (hanno) il maggior numero di dipendenti.

Esercizio 2:

Con riferimento all'utilizzo del linguaggio SQL nei linguaggi di programmazione, si descrivano le principali caratteristiche dell'SQL statico e dell'SQL dinamico. In particolare, con riferimento all'SQL dinamico, si descrivano le modalità di esecuzione immediata e non immediata.

Esercizio 3:

Sia dato uno schema Entità/Relazioni contenente due entità (Studente e Corso) e una relazione che le lega (Esame). L'entità Studente possieda quattro attributi (Matricola, NomeStudente, CognomeStudente, NumeroEsami), dove l'attributo NumeroEsami rappresenta il numero di esami superati da un certo studente, e l'entità Corso possieda tre attributi (CodiceCorso, NomeCorso, NumeroStudenti), dove l'attributo NumeroStudenti rappresenta il numero di studenti che hanno superato un certo esame. Si assuma che ogni studente abbia superato almeno un esame e che ogni corso sia stato superato da almeno uno studente. Il volume dei dati sia il seguente: le istanze dell'entità Studente siano 3000, quelle dell'entità Corsi 50 e quelle della relazione Esame 45000. Si considerino le seguenti operazioni, di natura interattiva: la registrazione di un esame, svolta mediamente 180 volte al mese, la lettura del numero di esami sostenuti da un dato studente, 10 volte al giorno, e la lettura del numero di studenti che hanno superato un certo esame, 5 volte al giorno. Si esegua un'analisi delle ridondanze per stabilire se conviene o meno mantenere gli attributi ridondanti NumeroEsami e NumeroStudenti.

Esercizio 4:

Si descrivano le modalità di gestione della concorrenza in SQL:1999. In particolare, si descrivano i diversi possibili livelli di isolamento. Successivamente, si stabilisca se i seguenti schedule appartengono o meno a TS, 2PL e 2PL stretto:

1. $s_1 : r_3(y), r_1(x), w_1(y), w_5(x), w_1(t), w_4(x), r_2(z), r_3(z), w_5(z), w_2(z), r_5(t), r_4(t);$
2. $s_2 : r_2(x), r_1(x), w_2(x), r_4(z), r_3(x), w_1(x), r_3(x), r_3(y), w_1(y), w_1(z), w_5(x), r_5(z), r_5(y).$