Corso di Basi di dati

26 Maggio 2004 - Fila A

1. Si vogliono gestire gli appelli d'esame dei corsi di un corso di laurea e le relative iscrizioni. Ogni corso ha un codice, un docente titolare, un titolo, un anno di corso e un semestre in cui viene tenuto e prevede delle prove di esame (scritto, orale, laboratorio). Ogni corso deve prevedere un appello in ognuna delle sessioni (identificate da un numero e corrispondenti ognuna ad un periodo). Per ogni appello si vuole memorizzare data, ora di inizio e ora di fine di ogni prova d'esame del corso corrispondente. Uno studente é identificato da un numero di matricola e caratterizzato da nome, cognome, luogo e data di nascita. Uno studente si iscrive ad un appello di un certo corso specificando quali prove vuole sostenere (ad esempio é possibile sostenere solo la prova scritta o solo la prova di laboratorio di un certo corso) ed eventualmente delle note. Una settimana prima della data dell'appello le iscrizioni vengono chiuse e a seconda del numero di iscritti viene fissata l'aula. Per le prove di laboratorio, se il numero degli iscritti é maggiore di 15, vengono fissati piú turni, numerati progressivamente, ognuno in data e ora diverse, e gli studenti iscritti vengono assegnati ai vari turni. Non ci possono essere prove di corsi dello stesso anno contemporanee, né prove contemporanee nella stessa aula o laboratorio, inoltre uno studente non si puø' iscrivere a piú prove contemporanee.

   (a) Definire uno schema E-R per la situazione descritta. Precisare le cardinalitá delle associazioni e i vincoli di identificazione (chiavi).

   (b) Indicare in modo testuale i vincoli d'integritá non catturati graficamente.

   (c) Realizzare uno schema relazionale corrispondente allo schema E-R di (a), specificando chiavi e chiavi esterne (e relative relazioni riferite).

2. Si consideri il seguente schema relazionale, relativo a una base di dati per la gestione di accessi a pagine web e del loro contenuto.

   $PAGWEB(\underline{IDPag}, Indirizzo, Titolo, Testo)$
   $CONTIENE(\underline{IDPag^*, Keyword})$
   $VISITA(\underline{IDUt, IDPag^*, DataOra})$

   (a) Formulare le seguenti interrogazioni in algebra relazionale:

      i. Determinare gli utenti che hanno visitato nel Marzo 2004 pagine del dominio `www.disi.unige.it` contenenti la keyword basi di dati;

      ii. Determinare gli utenti che non hanno mai visitato pagine contenenti la keyword basi di dati;

      iii. Determinare gli utenti che hanno visitato tutte le pagine del dominio `www.disi.unige.it` contenenti la keyword basi di dati.

   (b) Formulare in calcolo relazionale l'interrogazione (i).

   (c) Formulare le seguenti interrogazioni in SQL:

      i. Determinare per ogni utente l'indirizzo dell'ultima pagina visitata;

      ii. Determinare per ogni pagina quante keyword diverse contiene, quante volte é stata visitata e da quanti utenti diversi é stata visitata;

      iii. Determinare, tra le pagine del dominio `www.disi.unige.it` contenenti la keyword basi di dati, la piú visitata e quella che é stata visitata per ultima.

3. Si consideri lo schema relazionale R(A,B,C,D,E) con le seguenti dipendenze funzionali:
   $A \rightarrow BC, CD \rightarrow E, AC \rightarrow E, B \rightarrow D, E \rightarrow AB$

   (a) Si mostri un esempio di relazione su tale schema che verifichi la dipendenza $CD \rightarrow E$ e uno che la violi
   (b) Si porti l'insieme di dipendenze in forma minimale (se non lo é giá)
   (c) Si determino le chiavi di R
   (d) Si specifichi se lo schema é in 3NF o in BCNF
   (e) Si mostri una scomposizione lossless join in 3NF dello schema, si specifichi se preserva le dipendenze e se é in BCNF
   (f) Si mostri una scomposizione lossless join in BCNF dello schema, si specifichi se preserva le dipendenze e se é in 3NF

4. Si considerino le relazioni R($\underline{\text{X}}$,Y,Z*) e S($\underline{\text{K}}$,L,W) aventi rispettivamente cardinalitá $T(R)$ e $T(S)$. Assumere definito un vincolo di integritá referenziale tra l'attributo Z di R e la chiave K di S.

   (a) Discutere in cosa consiste il clustering di R ed S, specificando su quali attributi, mostrando un esempio di interrogazione per cui é vantaggioso e uno per cui é svantaggioso.
   (b) Supponendo le relazioni clusterizzate come in (a), si supponga di avere per ogni relazione un indice sparso clusterizzato (si specifichi su quale attributo) e un indice denso non clusterizzato su ogni altro attributo. Indicare la cardinalitá di tali indici.
   (c) Per uno degli indici a scelta mostrare un esempio di operazione di aggiornamento velocizzata ed uno di operazione rallentata da tali indici.
   (d) Per uno degli indici a scelta mostrare un esempio di operazione per cui l'indice é utile solo se ad albero.
   (e) Per uno degli indici a scelta mostrare un esempio di operazione che potrebbe essere eseguita accedendo unicamente l'indice e non i dati.