# Secondo Compitino di Basi di Dati Evoluzioni Recenti

30 Maggio 2003

NOME:

COGNOME:

---

| Esercizio | Punti previsti | Punti assegnati |
|---|---|---|
| 1 | 12 | |
| 2 | 7 | |
| 3 | 9 | |
| 4 | 5 | |
| Totale | 33 | |

**Esercizio 1 (Punti 12)**
Si consideri il seguente schema:

PERSONE(CodP, Nome, Cognome, Età)
GENITORI(Cod_G1,Cod_G2,Cod_Figlio)

La tabella GENITORI contiene per ogni persona i codici di entrambi i genitori.

Si richiede di:

1. Definire le seguenti relazioni in SQL-99:
   - FIGLI(Cod_G, Cod_Figlio) , che associa ad ogni persona i propri figli.
   - FRATELLI(Cod_F1,Cod_F2), che associa ogni persona ai propri fratelli/sorelle.
   - PRONIPOTI(Cod_Antenato,Cod_ProN), che associa ogni persona ai propri figli, ai propri nipoti, ai figli dei nipoti, e così via.
   - CUGINI(Cod_P1,Cod_P2) che contiene le coppie di persone che sono cugini, di qualunque grado.
2. Utilizzare le relazioni definite in precedenza per determinare il nome e il cognome dei cugini di Mario Rossi che non sono cugini di Anna Rossi.
3. Stabilire se, secondo SQL-99, l'interrogazione risultante viene eseguita dal sistema e perché.

**Esercizio 2 (Punti 7)**

Si consideri la relazione PERSONE dell'Esercizio 1, estesa con un campo CV, di tipo testo, che rappresenta il Curriculum Vitae et Studiorum della persona considerata. Si supponga che la tabella PERSONE contenga tre tuple a cui corrispondono i seguenti cv:

**d1** Mario Rossi ha una laurea in informatica. E' appassionato di libri gialli: la sua libreria ne è piena. Adora ascoltare la musica in compagnia di un buon libro, ma solo musica straniera.

**d2** Luigi Verdi ha studiato in conservatorio e ottenuto un diploma in pianoforte. Predilige la musica Jazz ma ama ascoltare anche musica sinfonica.

**d3** Anna Bianchi è appassionata di viaggi e di teatro. Durante i suoi viaggi, non manca mai di visitare i teatri locali. È un'esperta di musiche orientali.

Si supponga che l'insieme dei termini indice sia: diploma, laurea, libro, viaggio, musica, teatro. Si supponga inoltre che lo stemming venga applicato considerando tutte le parole con la stessa radice.

Si richiede di:
1. Presentare l'indice inverso per la base di dati sopra descritta.
2. Rappresentare i documenti nel modello booleano.
3. Rappresentare i documenti nel modello vettoriale.
4. Rappresentare nel modello booleano l'interrogazione per determinare quali documenti contengono ("laurea" o "diploma") e ("musica" ma non "teatro") . Stabilire quali documenti soddisfano l'interrogazione. Giustificare la risposta.
5. Rappresentare nel modello vettoriale l'interrogazione per determinare quali cv contengono la parola "musica". Stabilire quale documento è più simile all'interrogazione. Giustificare la risposta.

**Esercizio 3 (Punti 9)**

Si consideri il seguente schema relazionale:

PERSONE(<u>CodP</u>, Nome, Cognome, Età, CV, Foto, Foto_Sig)
FRATELLI(CodP1,CodP2)

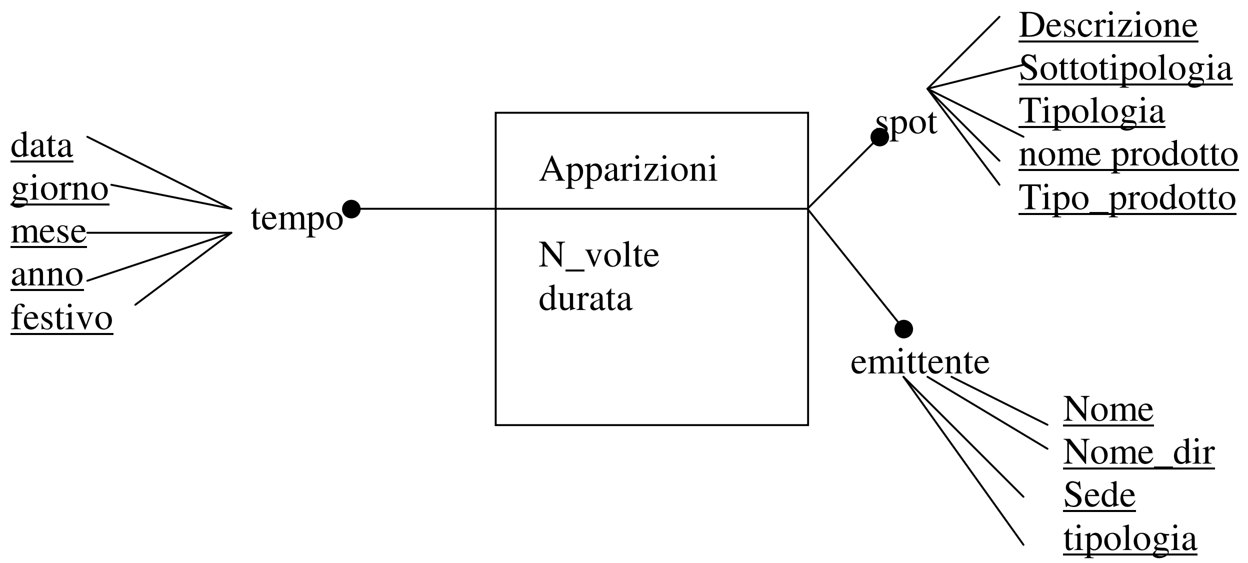
ottenuto dallo schema considerato nell'Esercizio 1 aggiungendo i campi per la rappresentazione delle fotografie delle persone e dei loro CV.

Si richiede di:

1. Presentare i comandi PL-SQL per inserire le informazioni di Mario Rossi, avente codice 12, età 25 anni, cv come quello presentato nell'Esercizio 2. Si supponga che la foto sia contenuta in un file rossi.gif, nella directory "Fotografie" e che NON si voglia copiare all'interno del DBMS. A partire dalla fotografia, si vogliono inoltre generare il maggior numero di informazioni per garantire un'altra flessibilità nella fase di interrogazione.
2. Si supponga che le variabili mia_foto e mia_foto_sig contengano i dati relativi ad una fotografia di esempio. Presentare gli statement SQL in Oracle 9i per rispondere alle seguenti interrogazioni:
   a) Determinare l'età delle persone il cui cv contiene "informatica" e/o "musica" (o una loro variante con la stessa radice), privilegiando chi contiene entrambi.
   b) Restituire i cv che contengono "viaggio" e "libro", all'interno della stessa frase. Discutere sotto quali ipotesi questa interrogazione può essere eseguita
   c) Restituire il cv che maggiormente rivela un temperamento artistico (quindi, il cui tema principale sia l'arte). Discutere sotto quali ipotesi questa interrogazione può essere eseguita.
   d) Quale interrogazione permette di determinare con buona approssimazione i laureati in informatica? Giustificare le scelte effettuate
   e) Determinare il nome e il cognome delle persone simili (in qualunque misura) alla foto in input, rispetto a colore e shape, ordinati dal più simile al meno simile.
   f) Determinare la coppia di persone che più si assomigliano tra di loro (pesando in modo omogeneo le features) che non sono fratelli, insieme al valore di distanza calcolato.

**Esercizio 4 (Punti 5)**

Si consideri il seguente Data Mart sviluppato per un gruppo di emittenti televisive che intendono analizzare e comparare le loro politiche pubblicitarie. A questo proposito si intende monitorare giornalmente ogni spot pubblicitario, in relazione al numero di volte in cui lo spot appare in ciascuna emittente e alla durata totale (cioe` al tempo complessivo occupato nel contesto di uno stesso giorno dallo stesso spot nel palinsesto dell'emittente) .
Ogni spot è almeno caratterizzato dalla descrizione, durata, sottotipologia, tipologia.
Esempi di tipologia sono: alimentare, telecomunicazioni, abbigliamento, ecc. Esempi di sottotipologia per la tipologia "telecomunicazioni" sono: telefonia fissa, telefonia mobile, satelliti, ecc. Ogni emittente è almeno caratterizzata dal nome, nome del direttore, indirizzo della sede, tipologia (nazionale, regionale). Lo spot contiene inoltre informazioni relative al prodotto pubblicizzato (nome prodotto, tipologia prodotto).




Si richiede di:

1. Modificare lo schema introducendo adeguate gerarchie ed introducendo la misura share medio. Presentare soltanto lo schema di fatto risultante.
2. Discutere quali oggetti devono essere creati in Oracle per rappresentare a livello logico lo schema generato (non è necessario presentare codice SQL).
3. Si supponga di determinare che la maggior parte delle interrogazioni richiede di analizzare gli spot non giornalmente ma a livello mensile. Qual'e' l'impatto di questa considerazione a livello concettuale e logico? Discutere eventuali soluzioni alternative. Utilizzare al più 6 righe per la risposta.