Prova Scritta di Basi di Dati (I modulo) del 16 giugno 2015

Esercizio 1

Data una base di dati contenente le due variabili relazionali

Corso	con schema	{corso_di_laurea, facoltà}
Laurea	con schema	{studente, sesso, corso_di_laurea, voto}

si vuole sapere per ogni facoltà chi sono gli studenti (maschi) e le studentesse che, tra gli studenti della stessa facoltà e dello stesso sesso, hanno avuto il voto più alto.

Si esprima la domanda sia nell'Algebra Relazionale che nel Calcolo Relazionale.

Nel Calcolo Relazionale: $\{x(\text{facoltà, studente}) \mid f\}$

$f = \exists\, y(\text{studente, sesso, corso\_di\_laurea, voto})\ \text{Laurea}(y) \wedge y(\text{studente}) = x(\text{studente}) \wedge \exists\, z(\text{corso\_di\_laurea, facoltà})\ \text{Corso}(z) \wedge z(\text{corso\_di\_laurea}) = y(\text{corso\_di\_laurea}) \wedge z(\text{facoltà}) = x(\text{facoltà}) \wedge \forall\, u(\text{studente, sesso, corso\_di\_laurea, voto})\ \neg\text{Laurea}(u) \vee u(\text{voto}) \leq y(\text{voto}) \vee u(\text{sesso}) \neq y(\text{sesso}) \vee \exists\, v(\text{corso\_di\_laurea, facoltà})\ \text{Corso}(v) \wedge v(\text{corso\_di\_laurea}) = u(\text{corso\_di\_laurea}) \wedge v(\text{facoltà}) \neq x(\text{facoltà})$

Nell'Algebra Relazionale:

$E_1 = \text{Corso} \bowtie \text{Laurea}$

$E_2 = \pi_{\text{studente, sesso, facoltà, voto}}\ (E_1)$

$E_3 = \pi_{\text{sesso, facoltà, voto}}\ (E_2)$

$E_4 = \rho_{\text{s:= sesso, f:= facoltà, u:= voto}}\ (E_3)$

$E_5 = E_2 \bowtie E_4$

$E_6 = \sigma_{\text{s} \neq \text{sesso} \vee \text{f} \neq \text{facoltà} \vee \text{u} \leq \text{voto}}\ (E_5)$

$E_7 = E_6 \div E_4$

$E_8 = \pi_{\text{studente, sesso, facoltà}}\ (E_6)$

Esercizio 2

Di un modello $M = [R, F]$ si sa che

— $R = \{A, B, C\}$
— $F = \{A \rightarrow B, f\}$ dove f è una dipendenza funzionale semplice (incognita) il cui determinante è ridotto ai minimi termini
— A è una criticità
— B è un attributo primario
— M è in Prima Forma Normale

Trovare tutti i possibili valori per f. Per ciascuno dei valori di f, dire se M è in Terza Forma Normale.

Siccome B deve essere un attributo primario, $B \in det(f)$. Siccome $det(f)$ deve essere ridotto ai minimi termini, $A \notin det(f)$ perché, altrimenti, $det(f)$ sarebbe riducibile a $det(f)–B$. Dunque, abbiamo o $det(f) = B$ oppure $det(f) = BC$.

Nel primo caso, f non può essere $B \rightarrow C$ perché, altrimenti, A sarebbe una chiave mentre per ipotesi A è una criticità. Quindi, tolta la dipendenza funzionale banale $B \rightarrow B$ che farebbe di B un attributo secondario, non resta che la dipendenza funzionale $B \rightarrow A$.

Nel secondo, tolte le dipendenze funzionali banali $BC \rightarrow B$ e $BC \rightarrow C$ che farebbero di B un attributo secondario, non resta che la dipendenza funzionale $BC \rightarrow A$.

Dunque si hanno due possibili valori per f: $B \rightarrow A$ oppure $BC \rightarrow A$. In entrambi i casi, i tre attributi A, B, C sono tutti primari e quindi il modello M è in Terza Forma Normale.

Esercizio 3

Si consideri il modello $M = [R, F]$ dove

$$R = ABCDE \qquad F = \{AB \to CD, AC \to D, BCD \to E\}$$

(3*a*) Trovare tutte le chiavi di M.

(3*b*) Trovare un modello M' in *forma semplice* che sia equivalente ad M.
Applicare ad M' l'algoritmo di Scomposizione in Terza Forma Normale.

(3*c*) Sia $\boldsymbol{S}$ il ricoprimento di R che si ottiene al punto (3*b*).
Applicando il Test di Conservazione dei Dati con input M' ed $\boldsymbol{S}$, verificare che la scomposizione di M' indotta da $\boldsymbol{S}$ conserva i dati.

(3*a*) Le sovrachiavi esplicite sono ABE, $ABCE$ e $ABCD$ e la loro intersezione AB è una sovrachiave visto che $\langle AB \rangle_F = R$. Per il criterio di unicità della chiave, AB è l'unica chiave del modello M.

(3*b*) In termini di dipendenze funzionali semplici F è equivalente a

$$\{AB \to C, AB \to D, AC \to D, BCD \to E\}$$

Se esaminiamo queste dipendenze funzionali (per esempio da sinistra a destra), troviamo che la dipendenza funzionale $AB \to D$ è ridondante. A questo punto abbiamo

$$F' = \{AB \to C, AC \to D, BCD \to E\}$$

e nessun determinante delle tre dipendenze funzionali è contraibile.
Dunque, $M' = [R, F']$.

L'applicazione ad M' dell'algoritmo di Scomposizione in Terza Forma Normale genera il ricoprimento

$$\boldsymbol{S} = \{ABC, ACD, BCDE\}.$$

(Si osservi che la chiave è contenuta in ABC.)

(3*c*) Applicando il Test di Conservazione dei Dati al tableau associato ad $\boldsymbol{S}$, si ha che il tableau finale contiene la riga (1, 2, 3, 4 5), la qual cosa dimostra che la scomposizione di M' indotta da $\boldsymbol{S}$ conserva dati.