

Domanda 1

Dire se il seguente schema relazionale

$\langle R(A,B,C,D), \{ AB \rightarrow C, C \rightarrow D, C \rightarrow A \} \rangle$

è in 3NF oppure no.

In caso R non sia in 3NF, mostrare una decomposizione di R in 3NF senza perdita di informazioni.

- Non è in 3NF (l'unica chiave è AB e D non è contenuta in nessuna chiave)
- Decomponiamo in
 - $R_1(A,B,C), \{ AB \rightarrow C, C \rightarrow A \}$
 - $R_2(C,D), \{ C \rightarrow D \}$
- AB è una chiave di R1, A è contenuta nella chiave di R1 e C è chiave di R2
- La decomposizione è senza perdita di informazioni perché l'attributo comune C è chiave in una delle due relazioni (R2)

Domanda 2

Due insiemi di dipendenze funzionali $F1$ e $F2$ definiti su uno schema S sono equivalenti se:

- Dopo aver applicato la chiusura transitiva a entrambi gli insiemi, otteniamo lo stesso insieme di dipendenze funzionali.
 - Non si dispone di elementi sufficienti per rispondere alla domanda, perché la nozione di equivalenza dipende dal numero di attributi dello schema.
 - La loro chiusura è la stessa.
 - Se essi sono esattamente lo stesso insieme di dipendenze funzionali.
- Osservazione: talvolta si usa l'espressione chiusura transitiva per riferirsi solo all'applicazione dell'assioma di transitività.
 - Qui l'ho considerata nel senso più ampio di applicazione di tutti gli assiomi fino a un punto fisso
 - Per questa ragione ho considerato valide entrambe le risposte

Domanda 3

- Si consideri lo schema $R(A,B,C,D)$. Quale dei seguenti insiemi di dipendenze funzionali è una copertura minimale?
 - A. $AB \rightarrow A, C \rightarrow BD$
 - B. $AB \rightarrow C, B \rightarrow D, B \rightarrow A$
 - C. $AB \rightarrow C, D \rightarrow BC$
 - D. $AB \rightarrow C, D \rightarrow B, D \rightarrow C$
- Le opzioni A e C non possono esserlo, perché hanno più di un attributo sul lato destro (Step 1 dell'algoritmo)
- L'opzione B non può esserlo, perché nella DF $AB \rightarrow C$ l'attributo A può essere eliminato, visto che A appartiene a $(B)^+$ (Step 3 dell'algoritmo)

Domanda 4

Cos'è una Chiave Esterna?

- A. Un trigger
- B. Una query SQL
- C. Un vincolo di integrità

Da definizione: è un il vincolo di integrità referenziale del modello relazionale.

Domanda 5

Si consideri il seguente schema relazionale:

$\langle R(A,B,C,D), \{ AB \twoheadrightarrow C, C \twoheadrightarrow D, C \twoheadrightarrow A \} \rangle$

e la seguente decomposizione di R:

- $R_1(A,B,C)$
- $R_2(C,D)$

La decomposizione R_1, R_2 è in 3NF?

NB: è l'inverso dell'esercizio 1,
quindi **ovviamente SI**

Domanda 6

L'insieme di dipendenze funzionali

- **$AB \rightarrow C, C \rightarrow D, C \rightarrow A$**

implica logicamente:

- A. **$AC \rightarrow D, AB \rightarrow D$**
- B. $AC \rightarrow D, C \rightarrow B$
- C. Nessuna delle altre risposte è corretta.
- D. $AB \rightarrow D, C \rightarrow B$

- $AB \rightarrow D$ si ottiene per transitività da $AB \rightarrow C$ e $C \rightarrow D$
- $AC \rightarrow D$ si ottiene come segue:
 - Assioma di aumento: $AB \rightarrow C$ e $C \subseteq AC$ implica $AB \rightarrow AC$
 - Assioma di aumento: $C \rightarrow D$ implica $AC \rightarrow AD$
 - Assioma di transitività: $AB \rightarrow AC$ e $AC \rightarrow AD$ implica $AB \rightarrow AD$
 - Regola di scomposizione: $AB \rightarrow AD$ implica $AB \rightarrow D$ e $AB \rightarrow A$ (banale)

Domanda 7

Si consideri lo schema R(A,B) e S(B,C). Quale delle seguenti è un'espressione SQL valida?

- A. `select A, AVG(C) as AV from R, S where R.B=S.B group by C`
- B. `select A, AVG(C) as AV from R, S where R.B=S.B group by B`
- C. `select A, AVG(C) as AV from R, S where R.B=S.B group by A`

Le colonne non aggregate che sono elencate nel SELECT devono essere incluse nel GROUP BY.

Questo è il caso della query 3, mentre non è il caso delle query 1 e 2.

Domanda 8

Date le tabelle mostrate in basso, e senza assumere altro, è possibile concludere che (B,C,A) sia una chiave esterna per S che fa riferimento a T?

- A. Potrebbe essere, ma non possiamo saperlo con certezza
- B. No, non può essere
- C. Sì, lo è di sicuro

- La risposta corretta è A, perché non sappiamo cosa potrebbe apparire in successive colonne della tabella S
- Infatti potrebbero essere aggiunte righe che violano il vincolo di integrità referenziale)

Domanda 9

Si consideri la tabella $R[A,B]$, dove PR denota le pagine di R e $|R|$ le tuple di R . Si consideri anche la tabella $S[B,C]$ con PS le pagine e $|S|$ le tuple. C'è un hash index definito su $S.B$ con un lookup cost pari a 2. Dobbiamo aggiornare ciascuna tupla di R che può andare in join con qualche tupla di S . In particolare, vogliamo eseguire il seguente update:

update R,S

set $R.A = 10$ where $R.A=S.B$

Il costo dell'update è:

- A. $PR + PS + |R| * 2$
- B. $PS + |S| * PR * 2$
- C. $PR + PS + |S| * 2$
- D. $PR + PR + |R| * 2$
- E. $PR + |R| * PS * 2$

La risposta corretta è D. Infatti:

- Devo aprire tutte le pagine di R (PR)
- Per ogni sua tupla, devo verificare se in S ci sono tuple di S in cui B ha lo stesso valore
- Il costo di questo controllo è solo $|R| * 2$. Infatti, per ipotesi, c'è un hash index su $S.B$ e il controllo dell'esistenza del valore può essere fatta *index-only* (senza neanche aprire il data file di S) e il costo del lookup sull'indice è 2
- Ogni pagina di R deve poi essere scritta su disco dopo aver eventualmente modificato il valore di $R.A$ a 10, quindi costo PR

Domanda 10

Abbiamo un indice B+-tree unclustered sull'attributo CITY di una tabella PERSON. Il selectivity factor è pari a 0.1 e sono presenti 1000 tuple nella tabella. Se ogni tupla occupa 500 KByte e ogni pagina occupa 600 KByte su disco, qual è il costo di trovare, accedere e mostrare su schermo le tuple che hanno CITY="Trento"?

- A. $3 * 0.1 + 1000$
- B. $3 * (1 + 0.1 * 1000)$
- C. $(3 + 1) * 0.1 * 1000$
- D. $3 + 0.1 * 1000$
- E. $3 * 0.1 * 1000d$

Il risultato è il costo del lookup sull'indice (3) sommato al numero di pagine corrispondente al numero di tuple che hanno valore «Trento» per l'attributo CITY.

Non sapendo il numero esatto di tali tuple, si usa il *selectivity factor*, che ci porta al valore $0,1 * 1000$

Domanda 11

Qual è lo scopo di un indice in un database server engine SQL?

- A. Per poter effettuare ricerche rapide.
- B. Mettere a disposizione un indice su un record.
- C. Tutte le motivazioni menzionate dalle altre risposte.
- D. Migliorare le prestazioni di una query.

Segue dalla nozione stessa di indice.

Domanda 12

Il costo di $\sigma_{A=1}(R \bowtie S)$ è sempre lo stesso del costo di $\sigma_{A=1}(R) \bowtie S$.

- A. Soltanto se A è un attributo in comune a R e S.
- B. **No**
- C. Soltanto se l'attributo A è una chiave esterna che fa riferimento a S.
- D. Soltanto se l'attributo A è una chiave di R o di S.
- E. Sì

Non essendoci indici, posso usare il Nested Loop Join. Quindi:

- il costo di $\sigma_{A=1}(R \bowtie S)$ è

$$P_R + P_R * P_S$$

[la selezione avviene in memoria al volo e non ha costi aggiuntivi].

- il costo di $\sigma_{A=1}(R) \bowtie S$ è

$$P_R + |R_{A=1}| * P_S$$

Questo non cambia se A è in comune a R e S, se A è chiave esterna, se A è chiave di R o di S

Domanda 13

Si considerino le due tabelle nella figura sotto e si supponga che ENROLLED.Sid abbia un vincolo di chiave esterna (Foreign Key) che fa riferimento a STUDENT.Sid. Quali dei seguenti comandi verrà eseguito con successo e inserirà un record nella tabella ENROLLED?

Student

Sid	Name	Login	Age	GPA
53666	Kayne	A@cs	28	4.2
53655	Tupac	B@cs	26	3.5
53688	Bieber	C@cs	22	3.9

Enrolled

Sid	Cid	Grade
53666	15-415	C
53688	15-721	A
53688	15-826	B
53655	15-415	C
53666	15-721	C

1. INSERT INTO ENROLLED values(53667, '15-420', 'C');
2. INSERT INTO ENROLLED values(53666, '15-421', 'C');
3. INSERT INTO ENROLLED values(53667, '15-415', 'C');
4. INSERT INTO ENROLLED values(53666, '15-415', 'C');

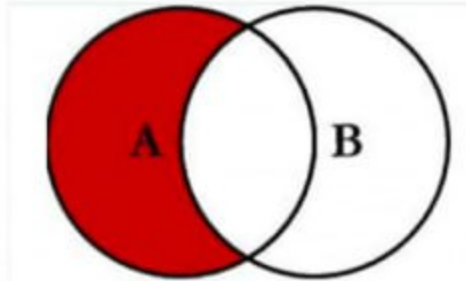
Opzioni:

- Solo 4
- Solo 2 e 4
- Solo 1 e 3
- Solo 3

- La 1 e la 3 vanno a violare il vincolo di integrità referenziale
- La 2 non ha nessun problema
- La 4 non viola alcun vincolo, perché - per violare una chiave - è necessario che ci siano due tuple diverse con stesso valore di chiave e qui non è il caso.

Domanda 14

Si considerino due relazioni (tabelle) A e B che hanno il medesimo schema. L'attributo C è chiave per entrambe. Ogni cerchio nella figura in basso corrisponde all'insieme delle tuple di ciascuna delle relazioni. Indicare quale delle query riportate in basso restituisce le tuple della parte evidenziata in rosso.



- A. $A - (A \cup B)$
- B. $B - A$
- C. $A - (A \bowtie B)$
- D. $B - (A \cap B)$

- Avendo le due relazioni lo stesso schema e C come chiave, nel join potranno esserci solo le tuple di A e B che coincidono sul valore di C (e quindi anche sui valori degli altri attributi, essendo C chiave).
- Se da A tolgo queste tuple, ottengo tutte le tuple di A che non sono in B

Domanda 15

Siano $R(A,B)$ e $S(B)$ due relazioni. La divisione R/S in algebra relazionale è definita come:

- A. $\pi_A(R) - \pi_A((\pi_A(R) \times S) - R)$
- B. $\pi_B(R) - \pi_A((\pi_B(R) \times S) - R)$
- C. $\pi_B(R) - \pi_B(S - (\pi_B(R) \times S))$
- D. $\pi_A(R) - \pi_A(R - (\pi_A(R) \times S))$

- La risposta corretta è A.
- La divisione R/S restituisce i valori di A per i quali tali per ogni valore di B in S ci sia una tupla $\langle A,B \rangle$ in R .
- Quindi:
 - La proiezione su A lascia solo la colonna A
 - Metto il risultato in join con S e produco tutte le combinazioni tra valori di A e valori di B in S
 - Sottraendo le tuple di R a quest'ultimo ottengo i valori di A per cui non ho un valore corrispondente
 - La proiezione successiva mi restituisce i solo valori di A
 - La differenza tra la proiezione su A di R e il risultato del passo precedente mi produce gli unici valori di A in R che sono associati con tutti i valori di B in S



Domanda 16

Un'entità debole:

Da definizione.

- A. è in relazione di ISA con un'altra entità
- B. non partecipa ad alcuna relazione
- C. ha la propria chiave contenente attributi di altre entità
- D. Non ha una chiave