

Basi di Dati, 20/12/2004, secondo compitino, compito A

1) Si consideri lo schema relazionale $R(A,B,C,D,E)$ con le seguenti DF

$CD \rightarrow EA, \quad ED \rightarrow A, \quad A \rightarrow C$

- Si trovino le chiavi di R .
- Dire se lo schema è in 3FN, e in ogni caso decomporlo con l'algoritmo di sintesi.
- Dire se lo schema è in FNBC, decomporlo con l'algoritmo di analisi, e specificare se il risultato preserva dati e dipendenze

Osservazione: l'insieme di dipendenze non è una copertura canonica.

- Date la definizione di decomposizione che preserva i dati
- Enunciate il criterio per stabilire quando una decomposizione binaria preserva i dati

3) Si consideri lo schema relazionale:

Studenti(Matricola, SNome, CorsoLaurea, Residenza)

Esami(Matricola*, CodMateria, CFU, Data, Voto)

e si consideri l'interrogazione

```
SELECT s.Matricola, s.Nome
```

```
FROM Studenti s, Esami e
```

```
WHERE s.Matricola=e.Matricola and s.CorsoLaurea='Informatica'
```

```
GROUP BY s.Matricola, s.Nome
```

```
HAVING sum(e.CFU)>60
```

- Disegnare l'albero di sintassi astratta di un'espressione algebrica ("albero logico") per l'interrogazione
- Disegnare un piano d'accesso ("albero fisico"), nell'ipotesi che non siano disponibili indici
- Disegnare un piano d'accesso ("albero fisico"), nell'ipotesi che siano disponibili indici a piacere

4) Si consideri il seguente schema, che rappresenta informazioni relative ai prestiti di una biblioteca

Prestiti(CodLibro, Titolo, Autore, CodUtente, Nome, DataPrestito, DataScadenza, CategoriaPrestito)

Un'ennupla della relazione rappresenta il fatto che un libro identificato da CodLibro è dato in prestito all'utente identificato da CodUtente; valgono inoltre le seguenti regole: (a) per ciascuna CategoriaPrestito la durata del prestito è fissa; (b) un utente non può avere in prestito più di un libro dello stesso autore, (c) un utente non può avere in prestito più di 5 libri

Per ciascuno tra (a), (b) e (c) dite se implica qualche dipendenza funzionale, e, se sì, quale.

5) (Opzionale)

- Enunciare ed illustrare *molto brevemente* le proprietà di una transazione
- Descrivere sinteticamente i tipi di fallimenti, e gli strumenti per la protezione dei dati dai fallimenti

Basi di Dati, 20/12/2004, secondo compito, compito A, soluzioni, versione 0.1

1) Si consideri lo schema relazionale $R(A,B,C,D,E)$ con le seguenti DF

$$CD \rightarrow EA, \quad ED \rightarrow A, \quad A \rightarrow C$$

a) Si trovino le chiavi di R .

Portiamo le dipendenze in forma canonica.

Riscriviamo le dipendenze come: $CD \rightarrow E$, $CD \rightarrow A$, $ED \rightarrow A$, $A \rightarrow C$.

Abbiamo che: $C_F^+ = C$, $D_F^+ = D$, $E_F^+ = E$, quindi nessun attributo è estraneo.

La $CD_{\{CD \rightarrow E, ED \rightarrow A, A \rightarrow C\}}^+ = CDEA$, quindi $CD \rightarrow A$ è ridondante.

Dipendenze in forma canonica: $\{CD \rightarrow E, ED \rightarrow A, A \rightarrow C\}$.

Calcolo della chiave.

Gli attributi BD non sono a destra di nessuna dipendenza, per cui appartengono ad ogni chiave.

$BD^+ = BD$. Applico l'algoritmo visto a lezione:

$$BDA^+ = BDACE$$

$$BDC^+ = BDCEA$$

$$BDE^+ = BDEAC$$

Quindi le chiavi sono: BDA, BDC, BDE.

b) Dire se lo schema è in 3FN, e in ogni caso decomporlo con l'algoritmo di sintesi.

Dato che tutti gli attributi sono primi, lo schema è in 3FN. Questo non impedisce di applicarvi l'algoritmo di sintesi.

Algoritmo di sintesi:

Raggruppo le dipendenze: $\{CD \rightarrow E\}$, $\{ED \rightarrow A\}$, $\{A \rightarrow C\}$.

Creo le relazioni: $R1(CDE)$, $R2(EDA)$, $R3(AC)$

Elimino le relazioni incluse: $R1(CDE)$, $R2(EDA)$, $R3(AC)$

Aggiungo una chiave: $R1(CDE)$, $R2(EDA)$, $R3(AC)$, $R4(ABD)$

(Si osservi che lo schema ottenuto è in BCNF, e la decomposizione preserva dati e dipendenze.)

a) Dire se lo schema è in FNBC, decomporlo con l'algoritmo di analisi, e specificare se il risultato preserva dati e dipendenze

Lo schema non è in FNBC, dato che $CD \rightarrow E$ è non banale ed il determinante non è una superchiave (la stessa cosa vale per le altre dipendenze in F).

Algoritmo di analisi:

$R1(ABCDE)$, decompongo rispetto a $CD \rightarrow E$:

$R2(CDEA, \{CD \rightarrow E, ED \rightarrow A, A \rightarrow C\})$, $R3(CDB, \{\})$; decompongo ancora $R2$ su $A \rightarrow C$:

$R4(AC)$, $R5(ADE)$, $R3(CDB, \{\})$

Calcolo la proiezione delle dipendenze su AC e su ADE. In particolare, mostriamo il calcolo della proiezione su ADE:

$$A_F^+ \cap ADE = AC \cap ADE = A,$$

$$D_F^+ \cap ADE = D,$$

$$E_F^+ \cap ADE = E,$$

$$AD_F^+ \cap ADE = ADCE \cap ADE = ADE,$$

$$AE_F^+ \cap ADE = AEC \cap ADE = AE,$$

$$DE_F^+ \cap ADE = DEA \cap ADE = DEA$$

per cui la proiezione di F su ADE è coperta da $\{AD \rightarrow E, ED \rightarrow A\}$:

$R_4(AC, \{A \rightarrow C\})$, $R_5(ADE, \{AD \rightarrow E, ED \rightarrow A\})$, $R_3(CDB, \{\})$.

Lo schema è ora in BCNF. I dati sono preservati per costruzione.

$CD \rightarrow E$ non è preservata, dato che ora $CD^+ = CD$.

2) a) *Date la definizione di decomposizione che preserva i dati*

Per definizione, dato uno schema $R(T, F)$, una decomposizione $\rho = \{R_1(T_1), \dots, R_n(T_n)\}$ preserva i dati se, e solo se, per ogni istanza valida r di R si ha che:

$$r = \Pi_{T_1} r \bowtie \dots \bowtie \Pi_{T_n} r$$

b) *Enunciate il criterio per stabilire quando una decomposizione binaria preserva i dati*

Una decomposizione binaria $R_1(T_1), R_2(T_2)$ di una relazione $R(T, F)$ preserva i dati se, e solo se, $F \vdash T_1 \cap T_2 \rightarrow T_1$ oppure $F \vdash T_1 \cap T_2 \rightarrow T_2$

3) *Si consideri lo schema relazionale:*

Studenti(Matricola, SNome, CorsoLaurea, Residenza)

Esami(Matricola*, CodMateria, CFU, Data, Voto)

e si consideri l'interrogazione

SELECT s.Matricola, s.Nome

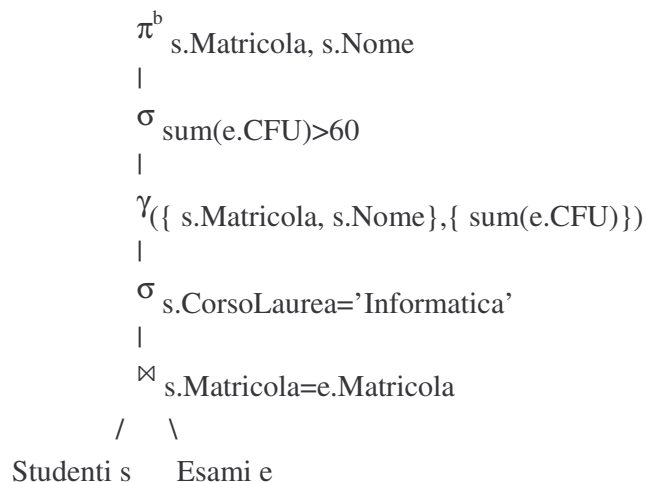
FROM Studenti s, Esami e

WHERE s.Matricola=e.Matricola and s.CorsoLaurea='Informatica'

GROUP BY s.Matricola, s.Nome

HAVING sum(e.CFU)>60

d) *Disegnare l'albero di sintassi astratta di un'espressione algebrica ("albero logico") per l'interrogazione*



oppure:

```

      πb s.Matricola, s.Nome
      |
      σ sum(e.CFU)>60
      |
      γ({ s.Matricola, s.Nome},{ sum(e.CFU)})
      |
      ⋈ s.Matricola=e.Matricola
    /
σ s.CorsoLaurea='Informatica'
\
      Esami e
      |
      Studenti s

```

e) Disegnare un piano d'accesso ("albero fisico"), nell'ipotesi che non siano disponibili indici

```

      Project(s.Matricola, s.Nome)
      |
      Filter(sum(e.CFU)>60)
      |
      GroupBy({ s.Matricola, s.Nome},{ sum(e.CFU)})
      |
      MergeSort(s.Matricola=e.Matricola)
    /
Sort(s.Matricola)
\
      Sort(e.Matricola)
      |
      Filter(s.CorsoLaurea='Informatica')
      |
      TableScan(Students s)
      |
      TableScan(Esami e)

```

f) Disegnare un piano d'accesso ("albero fisico"), nell'ipotesi che siano disponibili indici a piacere

```

      Project(s.Matricola, s.Nome)
      |
      Filter(sum(e.CFU)>60)
      |
      GroupBy({ s.Matricola, s.Nome},{ sum(e.CFU)})
      |
      IndexNestedLoop(s.Matricola=e.Matricola)
    /
Sort(s.Matricola)
\
      IndexFilter(Esami e,IdxEsamiMatr,
                  s.Matricola=e.Matricola)
      |
      IndexFilter(Studenti s,IdxStCL,
                  CorsoLaurea='Informatica')

```

4) Si consideri il seguente schema, che rappresenta informazioni relative ai prestiti di una biblioteca

Prestiti(CodLibro, Titolo, Autore, CodUtente, Nome, DataPrestito, DataScadenza, CategoriaPrestito)

Un'ennupla della relazione rappresenta il fatto che un libro identificato da CodLibro è dato in prestito all'utente identificato da CodUtente; valgono inoltre le seguenti regole: (a) per ciascuna CategoriaPrestito la durata del prestito è fissa; (b) un utente non può avere in prestito più di un libro dello stesso autore, (c) un utente non può avere in prestito più di 5 libri

Per ciascuno tra (a), (b) e (c) dite se implica qualche dipendenza funzionale, e, se sì, quale.

(a) CategoriaPrestito, DataPrestito, $\rightarrow$ DataScadenza

CategoriaPrestito, DataScadenza $\rightarrow$ DataPrestito

(b) $\text{Not}(\text{CodUtente}_= \text{ and } \text{Autore}_= \text{ and } \text{Prestito}_\neq) \Rightarrow$

CodUtente, Autore $\rightarrow$ CodLibro, Titolo, Nome, DataPrestito, DataScadenza, CategoriaPrestito

(c) questo vincolo non implica nessuna dipendenza funzionale

5) (Opzionale)

c) *Enunciare ed illustrare molto brevemente le proprietà di una transazione*

d) *Descrivere sinteticamente i tipi di fallimenti, e gli strumenti per la protezione dei dati dai fallimenti*

c) Vedete definizione 1.4 a pagina 19 del testo.

d) Tipi di fallimenti:

Fallimenti di transazione: solo lo stato della transazione coinvolta è compromesso, è necessario disfare gli effetti utilizzando il giornale.

Fallimenti di sistema: è compromesso il contenuto della memoria temporanea, è necessario cancellare gli effetti delle transazioni non terminate e rieseguire gli effetti delle transazioni terminate con successo dopo l'ultimo checkpoint, utilizzando il giornale.

Disastri: è compromesso anche lo stato della memoria persistente, è necessario riportare in linea una copia della base di dati salvata su di un diverso dispositivo, e riapplicare tutti gli effetti di tutte le transazioni terminate con successo dopo la creazione della copia, utilizzando il giornale.