

Basi si dati, II compito, 4/6/2018 - Versione (1)

1. Si consideri il seguente schema relazionale (l'insieme di dipendenze funzionali costituisce una copertura canonica):

$R \langle ABCDE, \{ AC \rightarrow D, BD \rightarrow A, BD \rightarrow E, A \rightarrow B \} \rangle$

1. Trovare almeno una chiave
 2. Sapendo che l'insieme degli attributi primi è $\{ A, B, C, D \}$, verificare quali forme normali sono rispettate dallo schema
 3. Applicare l'algoritmo di sintesi e dire se dati e dipendenze sono stati preservati.
 4. Applicare l'algoritmo di analisi e dire se dati e dipendenze sono stati preservati.
2. Si consideri il seguente schema relazionale $R(\text{IdR}, \text{IdS}^*, A, B)$, $S(\text{IdS}, I, J, K)$, $T(\text{IdT}, \text{IdS}^*, X, Y, W)$ e la seguente interrogazione

```
SELECT    S.J, T.X, count(*)
FROM      R, S, T
WHERE     R.A = T.W and S.IdS=T.IdS and S.K = 1
GROUP BY  S.J, T.X
```

1. Si scriva un piano di accesso logico
 2. Si scriva un piano di accesso fisico efficiente che non fa uso di indici
 3. Si scriva un piano di accesso fisico efficiente che fa uso di due indici posti su attributi a scelta dello studente
3. Si consideri un sistema gestito con protocollo disfare-rifare, si supponga che avvenga un fallimento di sistema, e si assuma che al momento della ripartenza questo sia il contenuto del log, dove un record $(W, T1, A, 1, 10)$ sta a indicare una scrittura (Write) della transazione T1 sulla variabile A con vecchio valore 1 e con nuovo valore 10:

```
(begin,T3) (W,T3,A,0,5) (begin,T2) (W,T2,B,0,20) (begin-ckp,{T3,T2}) (W,T3,A,5,10) (end-ckp)
(begin,T1) (commit,T3) (W,T1,C,0,20) (commit,T1) (begin,T4) (W,T4,C,20,50) (commit,T4)
```

Si risponda alle seguenti domande, nell'ordine in cui sono poste.

1. Al momento in cui il sistema aveva iniziato a prendere il log, qual era il valore di A, B e C?
2. Qual era il valore di A, B e C nel buffer al momento del checkpoint
3. Al termine del checkpoint qual è il valore di A, B e C sul disco?
4. Quali sono le transazioni che erano riuscite a dare il commit prima del fallimento di sistema e quali no?
5. Al momento del fallimento di sistema qual è il valore di A, B e C nel buffer?
6. Cosa possiamo dire riguardo al valore di A, B e C sul disco al momento del fallimento di sistema?
7. Quali sono le transazioni che al momento della ripartenza dovranno essere rifatte e quali sono quelle che dovranno essere disfatte?
8. L'operazione di scrittura $(W, T3, A, 0, 5)$ deve essere rifatta? Perché?
9. Si elenchino tutte le operazioni che devono essere rifatte.

Basi si dati, II compito, 4/6/2018 - Versione (2)

1. Si consideri il seguente schema relazionale (l'insieme di dipendenze funzionali costituisce una copertura canonica):

$R \langle ABCDE, \{ CE \rightarrow A, AD \rightarrow C, AD \rightarrow B, C \rightarrow D \} \rangle$

1. Trovare almeno una chiave
 2. Sapendo che l'insieme degli attributi primi è $\{ A, C, D, E \}$, verificare quali forme normali sono rispettate dallo schema
 3. Applicare l'algoritmo di sintesi e dire se dati e dipendenze sono stati preservati.
 4. Applicare l'algoritmo di analisi e dire se dati e dipendenze sono stati preservati.
2. Si consideri il seguente schema relazionale $R(\text{IdR}, \text{IdS}^*, A, B)$, $S(\text{IdS}, I, J, K)$, $T(\text{IdT}, \text{IdS}^*, X, Y, W)$ e la seguente interrogazione

```
SELECT    R.A, T.X, sum(R.B)
FROM      R, S, T
WHERE     R.IdS = S.IdS and R.A=T.Y and T.W = 100
GROUP BY  R.A, T.X
```

1. Si scriva un piano di accesso logico
 2. Si scriva un piano di accesso fisico efficiente che non fa uso di indici
 3. Si scriva un piano di accesso fisico efficiente che fa uso di due indici posti su attributi a scelta dello studente
3. Si consideri un sistema gestito con protocollo disfare-rifare, si supponga che avvenga un fallimento di sistema, e si assuma che al momento della ripartenza questo sia il contenuto del log, dove un record $(W, T1, A, 1, 10)$ sta a indicare una scrittura (Write) della transazione T1 sulla variabile A con vecchio valore 1 e con nuovo valore 10:

$(\text{begin}, T1) (W, T1, A, 0, 5) (\text{begin}, T4) (W, T4, B, 0, 20) (\text{begin-ckp}, \{T1, T4\}) (W, T1, A, 5, 10) (\text{end-ckp})$
 $(\text{begin}, T3) (\text{commit}, T1) (W, T3, C, 0, 20) (\text{commit}, T3) (\text{begin}, T2) (W, T2, C, 20, 50) (\text{commit}, T2)$

Si risponda alle seguenti domande, nell'ordine in cui sono poste.

1. Al momento in cui il sistema aveva iniziato a prendere il log, qual era il valore di A, B e C?
2. Qual era il valore di A, B e C nel buffer al momento del checkpoint
3. Al termine del checkpoint qual è il valore di A, B e C sul disco?
4. Quali sono le transazioni che erano riuscite a dare il commit prima del fallimento di sistema e quali no?
5. Al momento del fallimento di sistema qual è il valore di A, B e C nel buffer?
6. Cosa possiamo dire riguardo al valore di A, B e C sul disco al momento del fallimento di sistema?
7. Quali sono le transazioni che al momento della ripartenza dovranno essere rifatte e quali sono quelle che dovranno essere disfatte?
8. L'operazione di scrittura $(W, T1, A, 0, 5)$ deve essere rifatta? Perché?
9. Si elenchino tutte le operazioni che devono essere rifatte.

Basi si dati, II compito, 4/6/2018 - Versione 1 – soluzioni V1.0

1. Si consideri il seguente schema relazionale (l'insieme di dipendenze funzionali costituisce una copertura canonica):

$R \langle ABCDE, \{ AC \rightarrow D, BD \rightarrow A, BD \rightarrow E, A \rightarrow B \} \rangle$

1. Trovare almeno una chiave

$A \ BCDE^+ = BCDEA$ A può essere eliminato

$BCDE^+ = CDE$ B va tenuto

$BCDE^+ = BDE$ C va tenuto

$BCDE^+ = BCE$ D va tenuto

$BCDE^+ = BCDAE$ E può essere eliminato

BCD è una chiave

2. Sapendo che l'insieme degli attributi primi è $\{ A, B, C, D \}$, verificare quali forme normali sono rispettate dallo schema

Dato che $BD^+ = BDAE$, ed E non è primo, la dipendenza $BD \rightarrow E$ viola la BCNF e anche la terza forma normale.

3. Applicare l'algoritmo di sintesi e dire se dati e dipendenze sono stati preservati.

Raggruppiamo le dipendenze e definiamo le relazioni:

$AC \rightarrow D$ R1(ACD)

$BD \rightarrow A, BD \rightarrow E$ R2(ABDE)

$A \rightarrow B$ R3(AB)

Eliminiamo le relazioni contenute in altre: R1(ACD), R2(ABDE)

Verifichiamo se una relazione è superchiave:

$ACD^+ = ACDBE$ sì, R1 è superchiave, per cui la decomposizione è: $\{ R1(ACD), R2(ABDE) \}$

4. Applicare l'algoritmo di analisi e dire se dati e dipendenze sono stati preservati.

$AC^+ = ACDBE$, quindi $AC \rightarrow D$ ha a sinistra una superchiave.

$BD^+ = BDAE$ per cui:

$R1 = R1(BDAE, \{ BD \rightarrow A, BD \rightarrow E, A \rightarrow B \})$

$R2 = R2(BDC, \{ \})$

Dato che una dipendenza è andata perduta, dovremmo calcolare la proiezione esatta delle dipendenze. Non lo facciamo per semplicità, ma proseguiamo. In R1 la dipendenza $A \rightarrow B$ viola la BCNF, per cui abbiamo:

$R11(AB, \{ A \rightarrow B \})$, $R12(ADE, \{ \})$.

La decomposizione è quindi $\{ R2(BDC, \{ \}), R11(AB, \{ A \rightarrow B \}), R12(ADE, \{ \}) \}$ che preserva i dati ma non le dipendenze.

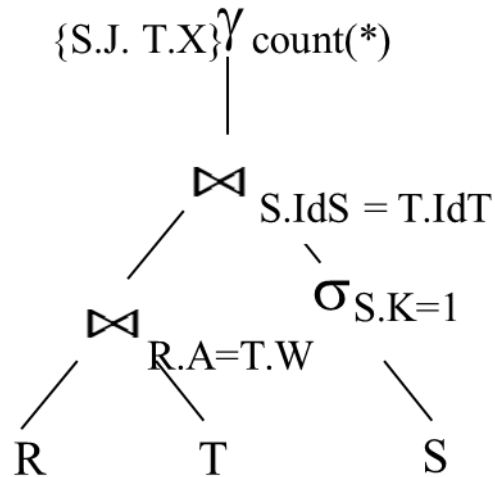
2. Si consideri il seguente schema relazionale $R(\text{IdR}, \text{IdS}^*, A, B)$, $S(\text{IdS}, I, J, K)$, $T(\text{IdT}, \text{IdS}^*, X, Y, W)$ e la seguente interrogazione

```

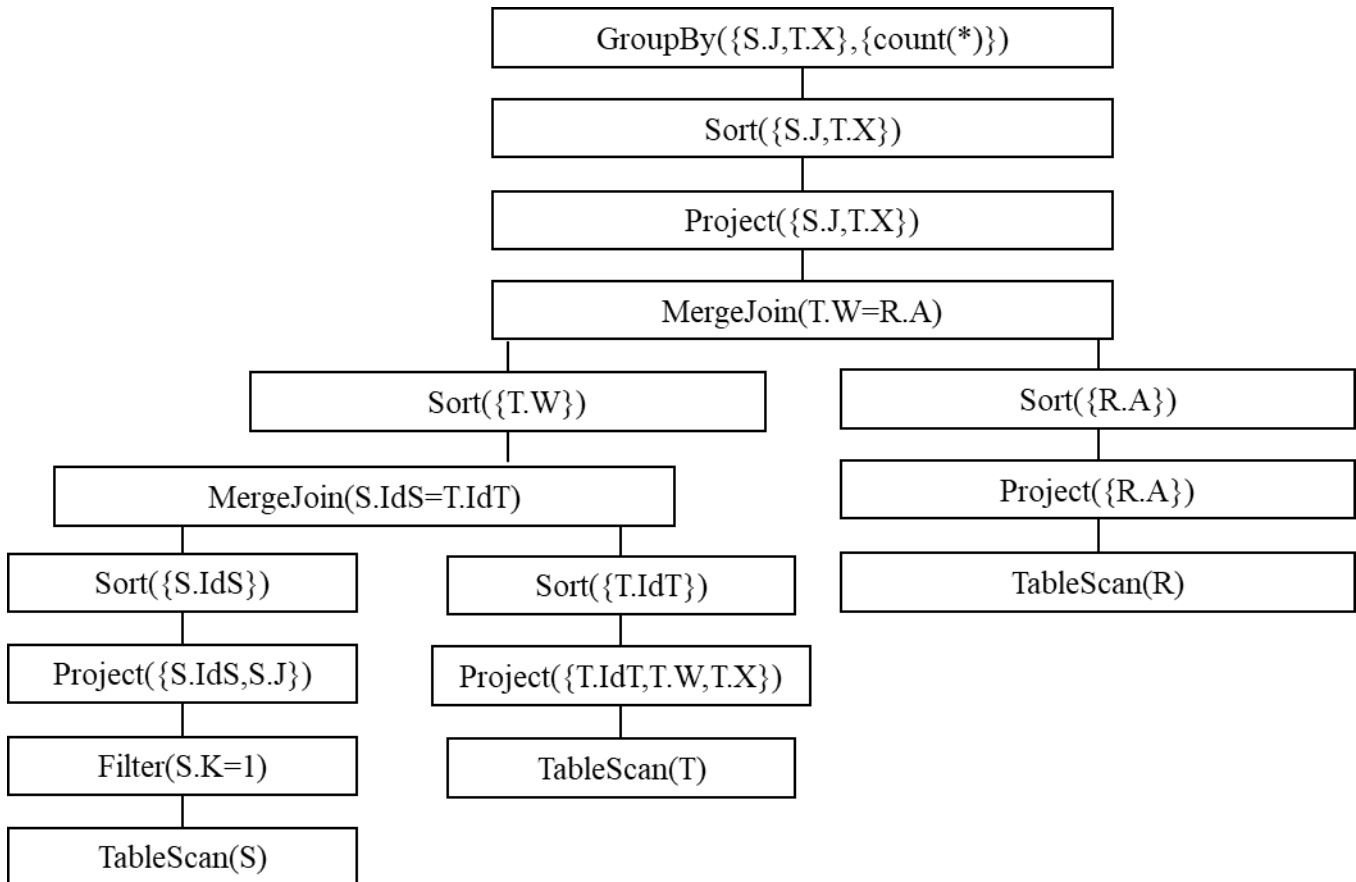
SELECT    S.J, T.X, count(*)
FROM      R, S, T
WHERE     R.A = T.W and S.IdS=T.IdT and S.K = 1
GROUP BY  S.J, T.X

```

1. Si scriva un piano di accesso logico

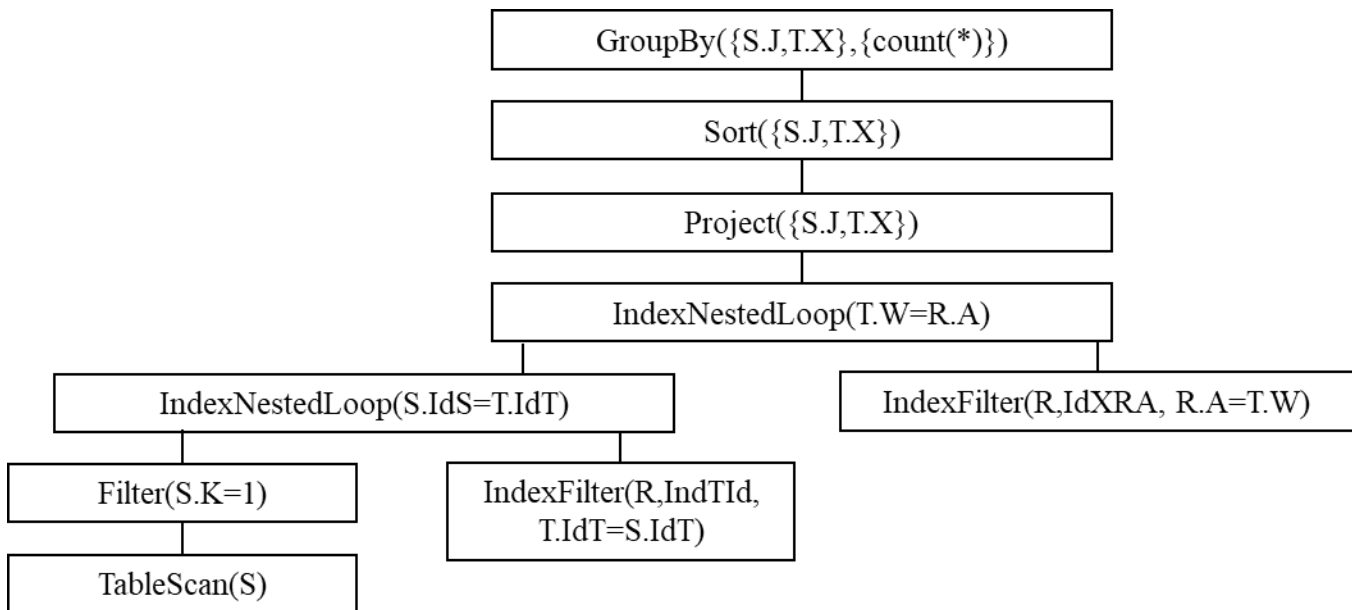


2. Si scriva un piano di accesso fisico efficiente che non fa uso di indici



3. Si scriva un piano di accesso fisico efficiente che fa uso di due indici posti su attributi a scelta dello studente

Ad esempio questo, ma sono possibili altre scelte:



3. Si consideri un sistema gestito con protocollo disfare-rifare, si supponga che avvenga un fallimento di sistema, e si assuma che al momento della ripartenza questo sia il contenuto del log, dove un record (W,T1,A,1,10) sta a indicare una scrittura (Write) della transazione T1 sulla variabile A con vecchio valore 1 e con nuovo valore 10:

(begin,T1) (W,T1,A,0,5) (begin,T2) (W,T2,B,0,20) (begin-ckp,{T1,T2}) (W,T1,A,5,10) (end-ckp)
 (begin,T3) (commit,T1) (W,T3,C,0,20) (commit,T3) (begin,T4) (W,T4,C,20,50) (commit,T4)

Si risponda alle seguenti domande, nell'ordine in cui sono poste.

- Al momento in cui il sistema aveva iniziato a prendere il log, qual era il valore di A, B e C?
- Qual era il valore di A, B e C nel buffer al momento del checkpoint
- Al termine del checkpoint qual è il valore di A, B e C sul disco?
- Quali sono le transazioni che erano riuscite a dare il commit prima del fallimento di sistema e quali no?
- Al momento del fallimento di sistema qual è il valore di A, B e C nel buffer?
- Cosa possiamo dire riguardo al valore di A, B e C sul disco al momento del fallimento di sistema?
- Quali sono le transazioni che al momento della ripartenza dovranno essere rifatte e quali sono quelle che dovranno essere disfatte?
- L'operazione di scrittura (W,T1,A,0,5) deve essere rifatta? Perché?
- Si elenchino tutte le operazioni che devono essere rifatte.

- (A=0,B=0,C=0)
- Alla fine del checkpoint: (A=10,B=20,C indefinito)
- Sul disco: (A=5 o 10,B=20,C=0)
- Riuscite: T1,T3,T4. Non riuscite : T2
- Nel buffer: (A=10, B=20, C=50)
- Sul disco: (A=5 o 10, B=20, C=0 o 20 o 50)
- Vedi domanda 4

8. No perché eseguita prima del checkpoint
9. (W,T1,A,5,10) (W,T3,C,0,20) (W,T4,C,20,50)

Appello di BD del Corso di laurea in Informatica del 20/6/2017

First intermediate test:

1. We represent information about bank accounts and money transfers among them. For each account we record the bank, the balance, the set of clients who hold (together) that account, and the associated transfers. Of a bank we record the name and the address of the headquarters, besides the list of all associated accounts. Every client may hold many accounts, and the client may be a person or a society – in both cases we keep a client code and a fiscal code. For a person, we only record name, family name, and birthdate. For a society, we keep the official name and the VAT code. Every transfer is associated to a source account and a destination account, and every account may be source and/or destination of many transfers. Each account may be associated to many credit cards. Each credit card is associated to one specific account and to one specific client, who is one of the clients who hold that account.
 - a. Draw a graphical conceptual schema for the database
 - b. Draw a graphical relational schema for the database
 - c. List all the attributes for all the tables in the relational schema
 - d.
2. Consider the following schema, which represents movies and actors

Movies(MId, Title, Director, ProductionHouse, Budget, Year)

MoviesActors(MId*, AId*)

Actors(AId, Name, BirthYear, Nationality)

- a. Draw the graphical relational schema.

Write the following queries:

- b. For each movie where some actor was born after 2000, give the title of the movie
- c. For each movie where some actor was born after 2000, give the title of the movie and the number of such actors
- d. For each movie where every actor was born after 2000, give the title of the movie
- e. For every two actors where the first acted in any movie where the second acted, give the AId and the Name of the one and of the other
- f. For each actor, give the name, the number of movies where she/he acted, and the average budget of these movies
- g. For each actor and each year, give the name of the actor, the year, the number of movies where she/he acted in that year, and the average budget of these movies