

BD, Seconda prova di verifica, del 19/12/2012

1. Si consideri lo schema relazionale $R(A,B,C,D,E)$ con le seguenti DF, già in forma canonica:

$\{BCD \rightarrow A, AC \rightarrow B, AD \rightarrow B, ED \rightarrow B\}$

- Si trovino le chiavi di R .
- Dire se lo schema è in 3FN o in FNBC.
- Si applichi allo schema l'algoritmo di sintesi, e (opzionale) si dica se la decomposizione risultante è in FNBC.
- Si applichi allo schema l'algoritmo di analisi e si dica se la decomposizione risultante preserva le dipendenze.
- Si porti il seguente schema relazionale in forma canonica:

$R\langle(A, B, C, D, E), \{ABC \rightarrow E, CE \rightarrow B, AD \rightarrow E, AC \rightarrow D, E \rightarrow D, CD \rightarrow B\}\rangle$

2. Si consideri lo schema relazionale

Dirigenti (IdDir :int, Nome string, AnnoNascita :int, Sesso :string, Stipendio :int)

DirProg (IdDir* :int, IdProg* :int)

Progetti(IdProg :int, AnnoInizio :int, AnnoFine :int, Budget :int)

e l'interrogazione

```
SELECT      DISTINCT d.IdDir, d.Nome, d.Stipendio, COUNT(*) AS C
FROM        Dirigenti d, DirProg dp, Progetti p
WHERE       d.IdProg = dp.IdProg AND dp.IdDir = d.IdDir
            AND d.Stipendio > 150000
GROUP BY   d.IdDir, d.Nome, d.Stipendio, p.AnnoInizio
HAVING     SUM(p.Budget) = 1000000
```

- Disegnare l'albero di sintassi astratta di un'espressione algebrica (albero logico) per l'interrogazione e si dia il tipo del risultato.
- Disegnare un piano d'accesso efficiente supponendo che siano presenti tutti gli indici che si desidera.
- Specificare se la condizione su "d.Stipendio" può essere posizionata sia prima che dopo l'operazione fisica di Group By.

3. Correggere la seguente affermazione incompleta.

- In uno schema relazionale $R(T)$ un insieme di attributi $X \subseteq T$ è chiave se e solo se, per ogni istanza r di R , e $\forall t_1, t_2 \in r$, si ha che $t_1[X] = t_2[X] \Rightarrow t_1 = t_2$.

BD, Seconda prova di verifica, del 19/12/2012: Soluzione

1. Si consideri lo schema relazionale $R(A,B,C,D,E)$ con le seguenti DF, già in forma canonica:

$\{BCD \rightarrow A, AC \rightarrow B, AD \rightarrow B, ED \rightarrow B\}$

(a) Si trovino le chiavi di R .

Gli attributi CDE si trovano solo a sinistra, gli attributi AB si trovano in entrambe le posizioni, per cui tutte le chiavi hanno forma CDEX, dove X è un sottoinsieme di AB. Calcoliamo $CDE^+ = CDEBA$. Quindi CDE è chiave, e pertanto è l'unica chiave.

(b) Dire se lo schema è in 3FN o in FNBC.

La dipendenza $BCD \rightarrow A$ viola la 3FN e quindi anche la BCNF.

(c) Si applichi allo schema l'algoritmo di sintesi, e (opzionale) si dica se la decomposizione risultante è in FNBC.

Raggruppo le dipendenze:

$BCD \rightarrow A$

$AC \rightarrow B$

$AD \rightarrow B$

$ED \rightarrow B$

Genero le relazioni:

$R1(ABCD), R2(ABC), R3(ABD), R4(BDE)$

Elimino le relazioni incluse in altre:

$R1(ABCD), R4(BDE)$

Aggiungo una relazione che contenga una chiave:

$R1(ABCD), R4(BDE), R5(CDE)$

Per dire se la decomposizione risultante è in FNBC, occorre proiettare le dipendenze su ogni relazione:

$R1 \langle (ABCD), \{BCD \rightarrow A, AC \rightarrow B, AD \rightarrow B\} \rangle$

$R4 \langle (BDE), \{DE \rightarrow B\} \rangle$

$R5 \langle (CDE), \{ \} \rangle$

Le dipendenze $\{AC \rightarrow B, AD \rightarrow B\}$ violano la BCNF.

(e) Si porti il seguente schema relazionale in forma canonica:

$R \langle (A, B, C, D, E), \{ABC \rightarrow E, CE \rightarrow B, AD \rightarrow E, AC \rightarrow D, E \rightarrow D, CD \rightarrow B\} \rangle$.

Verifichiamo la presenza di attributi estranei.

$A \in X$ è estraneo sse $F \vdash (X - \{A\}) \rightarrow Y$

In $ABC \rightarrow E$:

$$AB^+ = AB$$

$$AC^+ = ACDBE \text{ quindi } B \text{ è estraneo}$$

$$C^+ = C.$$

La dipendenza diventa $AC \rightarrow E$.

In $CE \rightarrow B$:

$$C^+ = C$$

$$E^+ = ED$$

Non ci sono quindi attributi estranei in $CE \rightarrow B$.

Le chiusure dei sottoinsiemi di AD, AC, CD sono tutte banali: $A^+ = A, D^+ = D, C^+ = C$

Non ci sono quindi altri attributi estranei, l'insieme è diventato:

$\{AC \rightarrow E, CE \rightarrow B, AD \rightarrow E, AC \rightarrow D, E \rightarrow D, CD \rightarrow B\}$

Verifichiamo la presenza di dipendenze ridondanti.

$X \rightarrow Y$ è ridondante sse $F - \{X \rightarrow Y\} \vdash X \rightarrow Y$

$AC^+_{F - \{AC \rightarrow E\}} = ACDBE$, quindi $AC \rightarrow E$ è ridondante e lo togliamo da F .

$CE^+_{F - \{CE \rightarrow B\}} = CEDB$, quindi $CE \rightarrow B$ è ridondante e lo togliamo da F .

$AD \rightarrow E$ è l'unica dipendenza ad avere E alla destra, quindi non è ridondante

$AC^+_{F - \{AC \rightarrow D\}} = AC$

$E^+_{F - \{E \rightarrow D\}} = E$

$CD \rightarrow B$ è l'unica dipendenza ad avere B alla destra, quindi non è ridondante.

Lo schema in forma canonica è quindi $\{AD \rightarrow E, AC \rightarrow D, E \rightarrow D, CD \rightarrow B\}$.

2. Si consideri lo schema relazionale

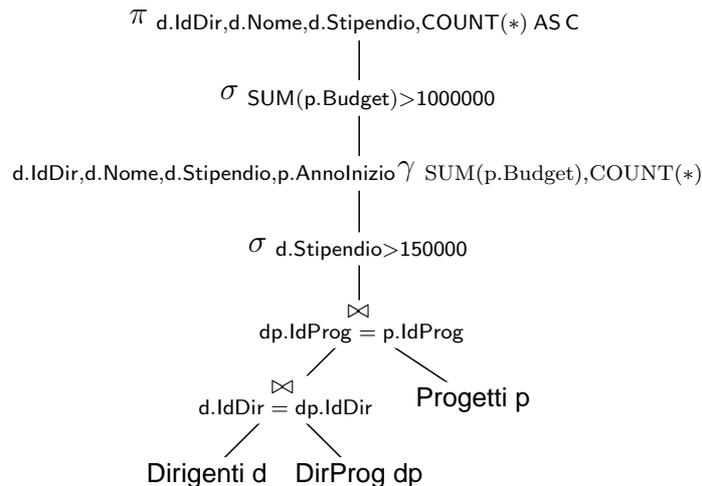
Dirigenti (IdDir :int, Nome string, AnnoNascita :int, Sesso :string, Stipendio :int)
 DirProg (IdDir * :int, IdProg * :int)
 Progetti(IdProg :int, AnnoInizio :int, AnnoFine :int, Budget :int)

e l'interrogazione

```

SELECT    DISTINCT d.IdDir, d.Nome, d.Stipendio, COUNT(*) AS C
FROM      Dirigenti d, DirProg dp, Progetti p
WHERE     d.IdProg = dp.IdProg AND dp.IdDir = d.IdDir
          AND d.Stipendio > 150000
GROUP BY  d.IdDir, d.Nome, d.Stipendio, p.AnnoInizio
HAVING    SUM(p.Budget) = 1000000
  
```

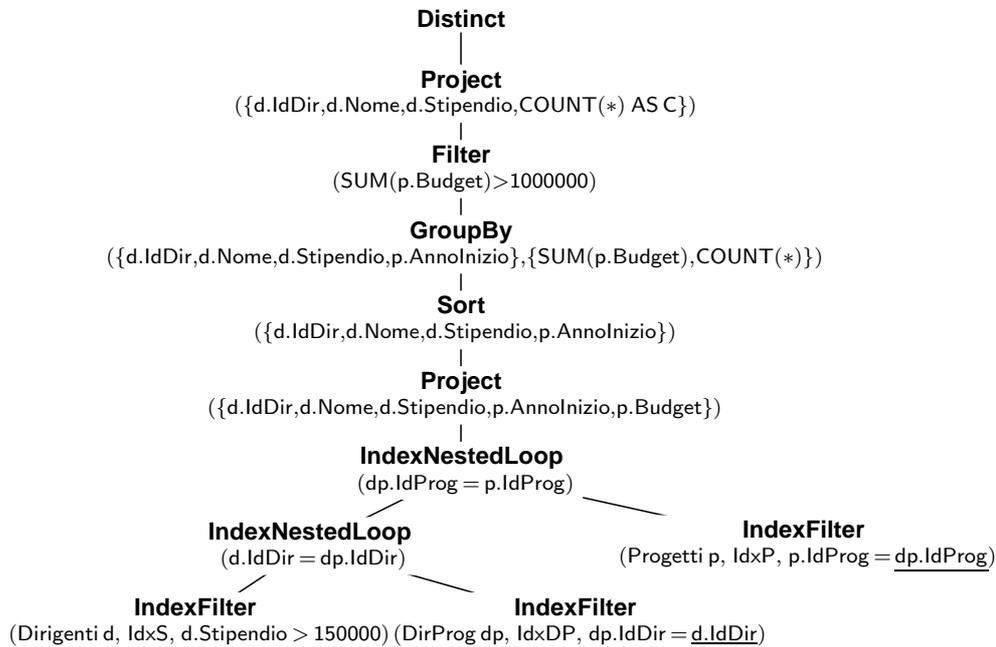
- (a) Disegnare l'albero di sintassi astratta di un'espressione algebrica (albero logico) per l'interrogazione e si dia il tipo del risultato.



Tipo risultato: {(d.IdDir :int, d.Nome :string, d.Stipendio :int, C :int)}

- (b) Disegnare un piano d'accesso efficiente supponendo che siano presenti tutti gli indici che si desidera.

Si suppone che esistano gli indici IdxS su Dirigenti.Stipendio, IdxP su Progetti.IdProg e IdxDP su (DirProg.IdDir):



(c) Specificare se la condizione su “d.Stipendio” può essere posizionata sia prima che dopo l’operazione fisica di Group By.

Si, ma si ottiene un piano migliore anticipandola

3. Correggere la seguente affermazione incompleta.

(a) In uno schema relazionale $R(T)$ un insieme di attributi $X \subseteq T$ è chiave se e solo se, per ogni istanza r di R , e $\forall t_1, t_2 \in r$, si ha che $t_1[X] = t_2[X] \Rightarrow t_1 = t_2$.

In uno schema relazionale $R(T)$ un insieme di attributi $X \subseteq T$ è chiave se e solo se, per ogni istanza valida r di R , e $\forall t_1, t_2 \in r$, si ha che $t_1[X] = t_2[X] \Rightarrow t_1 = t_2$ e per nessun sottoinsieme stretto Y di X valga $Y_F^+ = T$.