# Corso di Basi di Dati

Normalizzazione (Cap. 9)

## Data la seguente relazione

Docente	Dipartimento	Facoltà	Preside	Corso
Verdi	Matematica	Ingegneria	Neri	Analisi
Verdi	Matematica	Ingegneria	Neri	Geometria
Rossi	Fisica	Ingegneria	Neri	Analisi
Rossi	Fisica	Scienze	Bruni	Analisi
Bruni	Fisica	Scienze	Bruni	Fisica

#### Individuare:

- Le proprietà della relazione
- Chiave della relazione
- Eventuali ridondanze o anomalie

### Data la seguente relazione

Docente	Dipartimento	Facoltà	Preside	Corso
Verdi	Matematica Ingegneria		Neri	Analisi
Verdi	Matematica	Ingegneria	Neri	Geometria
Rossi	Fisica	Ingegneria	Neri	Analisi
Rossi	Fisica	Scienze	Bruni	Analisi
Bruni	Fisica	Scienze	Bruni	Fisica

#### Individuare:

- Le proprietà della relazione
- Chiave della relazione --> Dipartimento, Facoltà, Corso
- Eventuali ridondanze o anomalie

### Data la seguente relazione

Docente	ente Dipartimento Facoltà Preside		Corso	
Verdi	Matematica	Ingegneria	Neri	Analisi
Verdi	Matematica	Ingegneria	Neri	Geometria
Rossi	Fisica	Ingegneria	Neri	Analisi
Rossi	Fisica	Scienze	Bruni	Analisi
Bruni	Fisica	Scienze	Bruni	Fisica

#### Ridondanze:

- o dipendenza funzionale Facoltà → Preside introduce una ridondanza
- per ogni Corso, il Preside deve essere ripetuto

#### Anomalie:

- anomalia di aggiornamento (preside cambia, dobbiamo aggiornare tutte le tuple)
- anomalia di cancellazione (cancellando un preside, si perdono informazioni sui docenti)

#### Data la relazione dell'esercizio 9.1

Docente	Dipartimento Facoltà Preside		Corso	
Verdi	Matematica	Ingegneria	Neri	Analisi
Verdi	Matematica	Ingegneria	Neri	Geometria
Rossi	Fisica	Ingegneria	Neri	Analisi
Rossi	Fisica	Scienze	Bruni	Analisi
Bruni	Fisica	Scienze	Bruni	Fisica

Individuare la decomposizione in forma normale di Boyce e Codd

#### Data la relazione dell'esercizio 9.1

Docente	nte Dipartimento Facoltà Preside		Corso	
Verdi	Matematica	Ingegneria	Neri	Analisi
Verdi	Matematica	Ingegneria	Neri	Geometria
Rossi	Fisica	Ingegneria	Neri	Analisi
Rossi	Fisica	Scienze	Bruni	Analisi
Bruni	Fisica	Scienze	Bruni	Fisica

Individuare la decomposizione in forma normale di Boyce e Codd

#### Soluzione:

- Una chiave per questa relazione è Dipartimento, Facoltà, Corso.
- Anche gli attributi Docente, Facoltà, Corso sembrano formare una chiave in questa relazione, ma generalmente parlando questo non è corretto perché lo stesso docente può insegnare lo stesso corso in differenti dipartimenti di una Facoltà.

### Decomposizione:

Docente	Dipartimento	Facoltà	Corso
Verdi	Matematica	Ingegneria	Analisi
Verdi	Matematica	Ingegneria	Geometria
Rossi	Fisica	Ingegneria	Analisi
Rossi	Fisica	Scienze	Analisi
Bruni	Fisica	Scienze	Fisica

Facoltà	Preside
Ingegneria	Neri
Scienze	Bruni

- Questa decomposizione è corretta perché, con un join tra le due relazioni, otteniamo tutte e sole le tuple della relazione originaria
- la decomposizione risolve il problema delle anomalie, perché è in forma normale di Boyce-Codd.

Considerare uno schema di relazione R (E, N, L, C, S, D, M, P, A) con le seguenti dipendenze funzionali:

```
E \rightarrow NS,

NL \rightarrow EMD,

EN \rightarrow LCD,

C \rightarrow S,

D \rightarrow M,

M \rightarrow D,

EPD \rightarrow AE,

NLCP \rightarrow A.
```

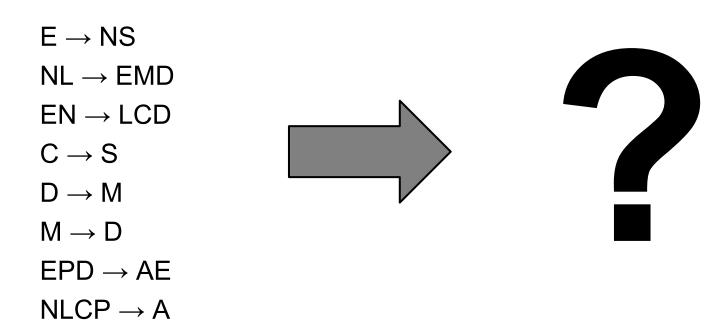
Calcolare una **copertura ridotta** per tale insieme e decomporre la relazione in **terza forma normale.** 

## Algoritmo per la copertura ridotta (Reminder)

I passi per calcolare la copertura ridotta di una relazione sono i seguenti:

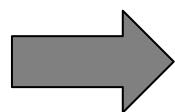
- 1. Sostituzione l'insieme di dipendenze funzionali con un <u>insieme</u> equivalente che ha i <u>secondi membri</u> costituiti da un <u>singolo attributo</u>
- Per ogni dipendenza verifica dell'esistenza di <u>attributi eliminabili dal</u> <u>primo membro</u>
- 3. Eliminazione delle dipendenze ridondanti

1. Sostituzione l'insieme di dipendenze funzionali con un <u>insieme equivalente</u> che ha i <u>secondi membri</u> costituiti da un <u>singolo</u> <u>attributo</u>



1. Sostituzione l'insieme di dipendenze funzionali con un <u>insieme equivalente</u> che ha i <u>secondi membri</u> costituiti da un <u>singolo</u> <u>attributo</u>

$$E \rightarrow NS$$
 $NL \rightarrow EMD$ 
 $EN \rightarrow LCD$ 
 $C \rightarrow S$ 
 $D \rightarrow M$ 
 $M \rightarrow D$ 
 $EPD \rightarrow AE$ 
 $NLCP \rightarrow A$ 



$$\mathsf{E} \to \mathsf{S}$$

$$\mathsf{E}\to\mathsf{N}$$

$$NL \rightarrow E$$

$$\mathsf{NL}\to\mathsf{M}$$

$$NL \rightarrow D$$

$$EN \rightarrow L$$

$$EN \rightarrow C$$

$$EN \rightarrow D$$

$$C \rightarrow S$$

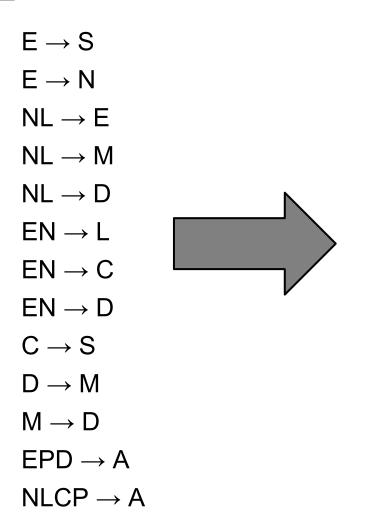
$$D \rightarrow M$$

$$\mathsf{M}\to\mathsf{D}$$

$$\mathsf{EPD} \to \mathsf{A}$$

$$NLCP \rightarrow A$$

2. Per ogni dipendenza verifica dell'esistenza di <u>attributi eliminabili dal primo</u> <u>membro</u>

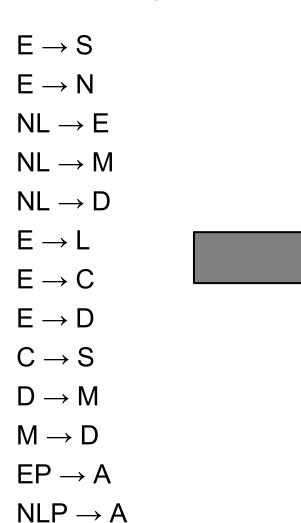




2. Per ogni dipendenza verifica dell'esistenza di <u>attributi eliminabili dal primo</u> membro

$$\begin{array}{l} E \rightarrow S \\ E \rightarrow N \\ NL \rightarrow E \\ NL \rightarrow M \\ NL \rightarrow D \\ EN \rightarrow L \\ EN \rightarrow C \\ EN \rightarrow D \\ C \rightarrow S \\ D \rightarrow M \\ M \rightarrow D \\ EPD \rightarrow A \\ NI \ CP \rightarrow A \\ \end{array} \qquad \begin{array}{l} E \rightarrow S \\ E \rightarrow N \\ NL \rightarrow M \\ NL \rightarrow E \\ NL \rightarrow M \\ NL \rightarrow D \\ E \rightarrow L \\ (EN \rightarrow L, E \rightarrow N) \\ (EN \rightarrow L, E \rightarrow N) \\ (EN \rightarrow C, E \rightarrow N) \\ (EN \rightarrow D, E \rightarrow N)$$

### 3. Eliminazione delle dipendenze ridondanti





3. Per il terzo passo, esaminiamo in dettaglio (a titolo di esempio) i passaggi con riferimento a due delle dipendenze funzionali.

Vediamo se la dipendenza  $\mathbf{E} \to \mathbf{N}$  è implicata dalle altre *calcolando la chiusura* iniziando dall'attributo E (considerando l'insieme delle D.F. ma senza la D.F.  $E \to N$ ).

$$N \notin X_{F}^{+} = \{E, S, L, C, D, M\}$$

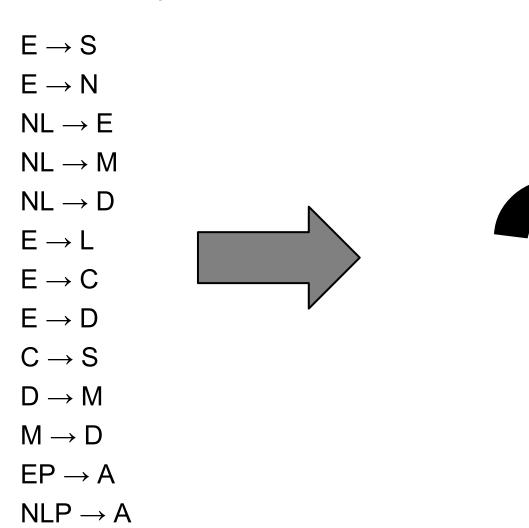
quindi la dipendenza non è implicata dalle altre e non può essere eliminata.

Considerando la dipendenza  $E \rightarrow S$ :

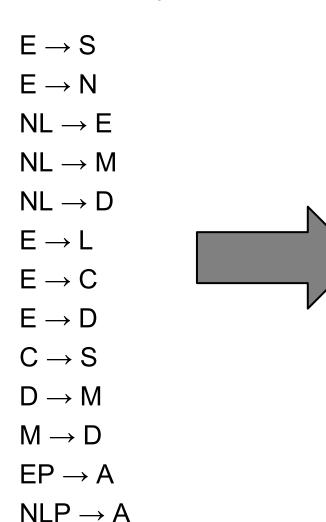
$$S \in X_F^+ = \{E, S, N, L, C, D, M\}$$

quindi la dipendenza è implicata dalle altre e può essere eliminata

### 3. Eliminazione delle dipendenze ridondanti



### 3. Eliminazione delle dipendenze ridondanti



$$\begin{array}{l} \mathsf{E} \to \mathsf{N} \\ \mathsf{NL} \to \mathsf{E} \\ \mathsf{NL} \to \mathsf{D} \\ \mathsf{E} \to \mathsf{L} \\ \mathsf{E} \to \mathsf{C} \\ \mathsf{C} \to \mathsf{S} \\ \mathsf{D} \to \mathsf{M} \\ \mathsf{M} \to \mathsf{D} \\ \mathsf{NLP} \to \mathsf{A} \end{array}$$

Ho ottenuto una copertura ridotta

$$\mathsf{E} \to \mathsf{N}$$

$$NL \rightarrow E$$

$$\mathsf{NL}\to\mathsf{D}$$

$$\mathsf{E} \to \mathsf{L}$$

$$\mathsf{E} \to \mathsf{C}$$

$$\mathsf{C}\to\mathsf{S}$$

$$\mathsf{D}\to\mathsf{M}$$

$$\mathsf{M} \to \mathsf{D}$$

$$NLP \rightarrow A$$

### Ho ottenuto una copertura ridotta

 $E \rightarrow N$ 

 $NL \rightarrow E$ 

 $NL \rightarrow D$ 

 $E \rightarrow L$ 

 $E \rightarrow C$ 

 $\boldsymbol{C} \to \boldsymbol{S}$ 

 $D \rightarrow M$ 

 $\mathbf{M} \to \mathbf{D}$ 

 $NLP \rightarrow A$ 

Ora possiamo procedere con la sintesi di schema in terza forma normale

## Sintesi di schema in 3NF (Reminder)

Dati uno schema R(U) e un insieme di dipendenze F su U

- 1. Viene calcolata una copertura ridotta G di F
- 2.G viene partizionato in sottoinsiemi tali che a ogni insieme appartengono dipendenze che hanno primi membri con la stessa chiusura
- 3. Viene costruito un insieme **U** di sottoinsiemi di U, uno per ciascuna partizione di dipendenze, con tutti gli attributi coinvolti nella partizione
- 4.Se un elemento di **U** è propriamente contenuto in un altro, allora esso viene eliminato da **U**
- 5.Viene costruito uno schema di relazione Ri(Ui) per ciascun elemento U<sub>i</sub> ∈ **U** con associate le dipendenze in G i cui attributi sono tutti contenuti in U<sub>i</sub>
- 6.Se nessuno degli Ui è chiave per R(U), allora viene calcolata una chiave K di R(U) e viene aggiunto allo schema generato uno schema di relazione sugli attributi K, senza dipendenze.

2. Individuo il partizionamento di G (primi membri con stessa chiusura)

$$E \rightarrow N$$

$$NL \rightarrow E$$

$$NL \rightarrow D$$

$$\textbf{E} \rightarrow \textbf{L}$$

$$\mathbf{E} \to \mathbf{C}$$

$$\boldsymbol{C} \to \boldsymbol{S}$$

$$D \rightarrow M$$

$$\mathbf{M} \to \mathbf{D}$$

$$NLP \rightarrow A$$

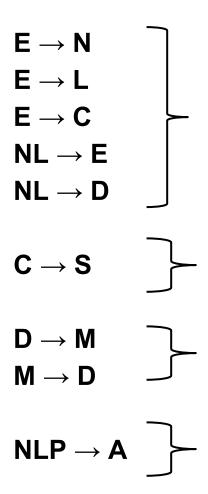
2. Individuo il partizionamento di G (primi membri con stessa chiusura)

$$\begin{array}{c} \textbf{E} \rightarrow \textbf{N} \\ \textbf{E} \rightarrow \textbf{L} \\ \textbf{E} \rightarrow \textbf{C} \end{array}$$
 
$$\begin{array}{c} \textbf{E}^{+} = \{ \, \textbf{E}, \, \textbf{N}, \, \textbf{L}, \, \textbf{C}, \, \textbf{D}, \, \textbf{C}, \, \textbf{S}, \, \textbf{M} \, \} \\ \textbf{NL} \rightarrow \textbf{E} \\ \textbf{NL} \rightarrow \textbf{D} \end{array}$$
 
$$\begin{array}{c} \textbf{NL}^{+} = \{ \, \textbf{E}, \, \textbf{N}, \, \textbf{L}, \, \textbf{C}, \, \textbf{D}, \, \textbf{C}, \, \textbf{S}, \, \textbf{M} \, \} \\ \textbf{C} \rightarrow \textbf{S} \end{array}$$
 
$$\begin{array}{c} \textbf{CS}^{+} = \{ \, \textbf{C}, \, \textbf{S} \, \} \\ \textbf{D} \rightarrow \textbf{M} \\ \textbf{M} \rightarrow \textbf{D} \end{array}$$
 
$$\begin{array}{c} \textbf{D}^{+} = \{ \, \textbf{D}, \, \textbf{M} \, \}; \, \, \textbf{M}^{+} = \{ \, \textbf{D}, \, \textbf{M} \, \} \\ \textbf{NLP} \rightarrow \textbf{A} \end{array}$$
 
$$\begin{array}{c} \textbf{NLP}^{+} = \{ \, \textbf{E}, \, \textbf{N}, \, \textbf{L}, \, \textbf{C}, \, \textbf{D}, \, \textbf{C}, \, \textbf{S}, \, \textbf{M}, \, \textbf{A} \, \} \end{array}$$

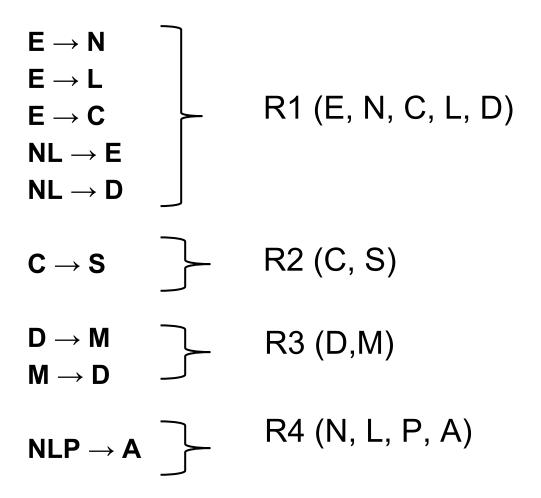
2. Individuo il partizionamento di G (primi membri con stessa chiusura)

$$\begin{array}{c} \mathsf{E} \to \mathsf{N} \\ \mathsf{E} \to \mathsf{L} \\ \mathsf{E} \to \mathsf{C} \end{array} \qquad \begin{array}{c} \mathsf{E}^+ = \{\, \mathsf{E},\, \mathsf{N},\, \mathsf{L},\, \mathsf{C},\, \mathsf{D},\, \mathsf{C},\, \mathsf{S},\, \mathsf{M}\,\,\} \\ \mathsf{C} \to \mathsf{E} \\ \mathsf{NL} \to \mathsf{D} \end{array} \qquad \begin{array}{c} \mathsf{NL}^+ = \{\, \mathsf{E},\, \mathsf{N},\, \mathsf{L},\, \mathsf{C},\, \mathsf{D},\, \mathsf{C},\, \mathsf{S},\, \mathsf{M}\,\,\} \\ \mathsf{C} \to \mathsf{S} \end{array} \qquad \begin{array}{c} \mathsf{CS}^+ = \{\, \mathsf{C},\, \mathsf{S}\,\,\} \\ \mathsf{D} \to \mathsf{M} \\ \mathsf{M} \to \mathsf{D} \end{array} \qquad \begin{array}{c} \mathsf{D}^+ = \{\, \mathsf{D},\, \mathsf{M}\,\,\}\,\,;\,\,\, \mathsf{M}^+ = \{\, \mathsf{D},\, \mathsf{M}\,\,\} \\ \mathsf{NLP} \to \mathsf{A} \end{array} \qquad \begin{array}{c} \mathsf{NLP}^+ = \{\, \mathsf{E},\, \mathsf{N},\, \mathsf{L},\, \mathsf{C},\, \mathsf{D},\, \mathsf{C},\, \mathsf{S},\, \mathsf{M},\, \mathsf{A}\,\,\} \end{array}$$

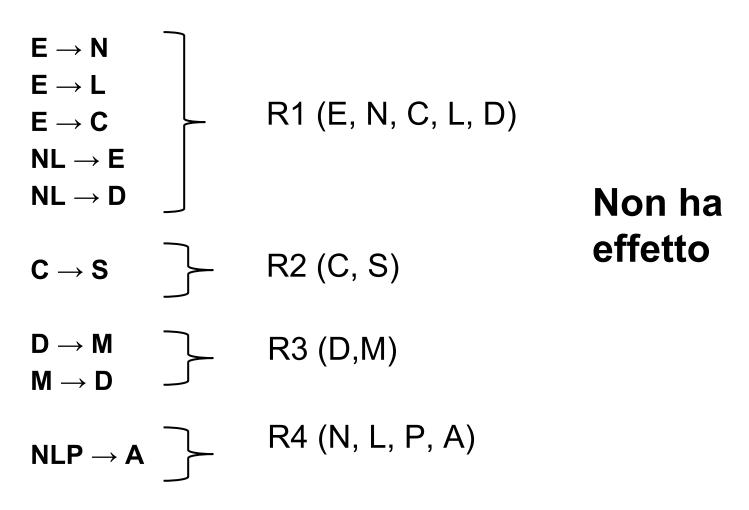
3. Viene costruito un insieme **U** di sottoinsiemi di U, uno per ciascuna partizione di dipendenze, con tutti gli attributi coinvolti nella partizione



3. Viene costruito un insieme **U** di sottoinsiemi di U, uno per ciascuna partizione di dipendenze, con tutti gli attributi coinvolti nella partizione



4. Se un elemento di **U** è propriamente contenuto in un altro, allora esso viene eliminato da **U** 



5. Viene costruito uno schema di relazione Ri(Ui) per ciascun elemento U<sub>i</sub> ∈ **U** con associate le dipendenze in G i cui attributi sono tutti contenuti in U<sub>i</sub>

R1 (E, N, C, L, D) con 
$$E \rightarrow N$$
,  $E \rightarrow L$ ,  $E \rightarrow C$   
 $NL \rightarrow E$ ,  $NL \rightarrow D$ 

R2 (C, S) con 
$$C \rightarrow S$$

R3 (D,M) con D 
$$\rightarrow$$
 M, M  $\rightarrow$  D

R4 (N, L, P, A) con NLP 
$$\rightarrow$$
 A

6. Se nessuno degli Ui è chiave per R(U), allora viene calcolata una chiave K di R(U) e viene aggiunto allo schema generato uno schema di relazione sugli attributi K, senza dipendenze.

R1 (E, N, C, L, D) con 
$$E \rightarrow N$$
,  $E \rightarrow L$ ,  $E \rightarrow C$   
 $NL \rightarrow E$ ,  $NL \rightarrow D$ 

R2 (C, S) con  $C \rightarrow S$ 

R3 (D,M) con D  $\rightarrow$  M, M  $\rightarrow$  D

R4 (N, L, P, A) con NLP  $\rightarrow$  A

R4 contiene già la chiave della relazione originale R(U)

- Trovare copertura ridotta G
- 2. Trovare tutte le chiavi
- 3. Dire se ci sono e quali dipendenze violano 3NF
- 4. Normalizzare lo schema in 3NF

- Trovare copertura ridotta G
   C →A, C→D, A→B, D→B, D→E, B→F, AD→C
- 2. Trovare tutte le chiavi
- 3. Dire se ci sono e quali dipendenze violano 3NF
- 4. Normalizzare lo schema in 3NF

- Trovare copertura ridotta G
   C →A, C→D, A→B, D→B, D→E, B→F, AD→C
- 2. Trovare tutte le chiavi AD, C
- 3. Dire se ci sono e quali dipendenze violano 3NF
- 4. Normalizzare lo schema in 3NF

- Trovare copertura ridotta G
   C →A, C→D, A→B, D→B, D→E, B→F, AD→C
- 2. Trovare tutte le chiavi AD, C
- 3. Dire se ci sono (ed eventualmente quali) dipendenze che violano 3NF.
- 4. Normalizzare lo schema in 3NF

- 3) Ho:  $C \rightarrow A$ ,
  - C→D,
  - A→B,
  - D→B,
  - D→E,
  - B→F,
  - $AD \rightarrow C$

Con chiavi: AD, C

Next step: Dire se ci sono e quali dipendenze violano 3NF.

Condizione perchè una dipendenza funzionale X→Y sia in 3NF

- a. X contiene chiave K di R oppure
- b. ogni attributo in Y è contenuto in almeno una chiave di R

Ho:  $C \rightarrow A$ , non viola 3NF

C→D, non viola 3NF

A→B, viola (A non super chiave e B non presente in chiave)

D→B, viola (D non super chiave e B non presente in chiave)

D→E, viola (D non super chiave e E non presente in chiave)

B→F, viola (B non super chiave e F non presente in chiave)

AD→C non viola

Con chiavi: AD, C

Non 3NF

4. Normalizzare lo schema in 3NF (vedi step dell'algoritmo di sintesi 3NF)

 $C \rightarrow A$ ,  $C \rightarrow D$ ,  $A \rightarrow B$ ,  $D \rightarrow B$ ,  $D \rightarrow E$ ,  $B \rightarrow F$ ,  $AD \rightarrow C$ 

Con chiavi: AD, C

#### 4. Normalizzare lo schema in 3NF

Calcolo la chiusura dei primi membri:

$$C \rightarrow A$$
,  $C^+=\{A, B, C, D, E, F\}$   
 $C \rightarrow D$ ,

$$A \rightarrow B$$
,  $A^+=\{A, B, F\}$ 

$$D\rightarrow B$$
,  $D^+=\{D, B, F, E\}$   
 $D\rightarrow E$ ,

$$B \rightarrow F$$
,  $B^+=\{B, F\}$ 

$$AD \rightarrow C$$
  $AD^+=\{A, B, C, D, E, F\}$ 

#### 4. Normalizzare lo schema in 3NF

$$C \rightarrow A$$
,  $C^{+} = \{ A, B, C, D, E, F \}$   
 $C \rightarrow D$ ,  $A \rightarrow B$ ,  $A^{+} = \{ A, B, F \}$   
 $D \rightarrow B$ ,  $D^{+} = \{ D, B, F, E \}$   
 $D \rightarrow E$ ,  $B \rightarrow F$ ,  $B^{+} = \{ B, F \}$   
 $AD \rightarrow C$   $AD^{+} = \{ A, B, C, D, E, F \}$ 

Le chiusure di C e AD coincidono, quindi vanno considerati nella stessa partizione

#### 4. Normalizzare lo schema in 3NF

$$C \rightarrow A$$
, R1 (C,A, D) key C, AD  $C \rightarrow D$ , AD $\rightarrow C$ 
 $A \rightarrow B$ , R2 (A, B) key A

 $D \rightarrow B$ , R3 (B, E, D)key D

 $D \rightarrow E$ , R4 (B, F) key B

## 4. Normalizzare lo schema in 3NF

$C \rightarrow A$ , $C \rightarrow D$ , $AD \rightarrow C$	R1 (C,A, D) CONTIENE LE	key C, AD  CHIAVI DELLA RELAZIONE ORIGINALE
А→В,	R2 (A, B)	key A
D→B, D→E,	R3 (B, E, D) ł	key D
B→F,	R4 (B, F)	key B

Si consideri la relazione riportata sotto, la quale rappresenta alcune informazioni sui prodotti di una falegnameria e i relativi componenti. Vengono indicati:

- il tipo del componente di un prodotto (Tipo),
- la quantità del componente necessaria per un certo prodotto (Q),
- il prezzo unitario del componente di un certo prodotto (PC),
- il fornitore del componente (Fornitore),
- il prezzo totale del singolo prodotto (PT).

Individuare le dipendenze funzionali e la chiave di questa relazione.

Prodotto	Componente	Tipo	Q	PC	Fornitore	PT
Libreria	Legno	Noce	50	10.000	Forrest	400.000
Libreria	Bulloni	B212	200	100	Bolt	400.000
Libreria	Vetro	Cristal	3	5.000	Clean	400.000
Scaffale	Legno	Mogano	5	15.000	Forrest	300.000
Scaffale	Bulloni	B212	250	100	Bolt	300.000
Scaffale	Bulloni	B412	150	300	Bolt	300.000
Scrivania	Legno	Noce	10	8.000	Wood	250.000
Scrivania	Maniglie	H621	10	20.000	Bolt	250.000
Tavolo	Legno	Noce	4	10.000	Forrest	200.000

Prodotto	Componente	Tipo	Q	PC	Fornitore	PT
Libreria	Legno	Noce	50	10.000	Forrest	400.000
Libreria	Bulloni	B212	200	100	Bolt	400.000
Libreria	Vetro	Cristal	3	5.000	Clean	400.000
Scaffale	Legno	Mogano	5	15.000	Forrest	300.000
Scaffale	Bulloni	B212	250	100	Bolt	300.000
Scaffale	Bulloni	B412	150	300	Bolt	300.000
Scrivania	Legno	Noce	10	8.000	Wood	250.000
Scrivania	Maniglie	H621	10	20.000	Bolt	250.000
Tavolo	Legno	Noce	4	10.000	Forrest	200.000

In relazione all'esercizio precedente, consideriamo le seguenti operazioni:

- Inserimento di un nuovo prodotto;
- 2. Cancellazione di un prodotto;
- 3. Aggiunta di una componente a un prodotto;
- 4. Modifica del prezzo di un prodotto.

Discutere i tipi di anomalia che possono essere causati da tali operazioni.

R (A, B, C, D, E, F, G)

con:

AF→BE

EF→BCD

 $A \rightarrow F$ 

 $B \rightarrow C$ 

- 1. Trovare copertura ridotta
- Trovare tutte le chiavi
- 3. FD che violano 3NF
- 4. Normalizzare in 3NF