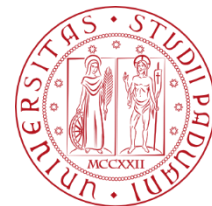


**Laurea in Informatica  
A.A. 2021-2022**

**Corso "Base di Dati"**

**Normalizzazione**

**Prof. Massimiliano de Leoni**



**UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA**

# Forme normali

- Una forma normale è una proprietà di una base di dati relazionale che ne garantisce la “qualità”, cioè l'assenza di determinati difetti
- Quando una relazione non è normalizzata:
  - presenta ridondanze;
  - si presta a comportamenti poco desiderabili durante gli aggiornamenti.

# Normalizzazione

- Procedura che permette di trasformare schemi non normalizzati in schemi che soddisfano una forma normale
- La normalizzazione va utilizzata come **tecnica di verifica** dei risultati della progettazione di una base di dati

# Una Relazione con Anomalie

<u>Impiegato</u>	Stipendio	<u>Progetto</u>	Bilancio	Funzione
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
Mori	48	Marte	2	direttore
Mori	48	Venere	15	progettista
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	direttore

# Anomalie / 1

- Lo stipendio di ciascun impiegato è ripetuto in tutte le ennuple relative
  - ridondanza
- Se lo stipendio di un impiegato varia, è necessario andarne a modificare il valore in diverse ennuple
  - anomalia di aggiornamento

<u>Impiegato</u>	<u>Stipendio</u>	<u>Progetto</u>	<u>Bilancio</u>	<u>Funzione</u>
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
Mori	48	Marte	2	direttore
Mori	48	Venere	15	progettista
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	direttore

# Anomalie / 2

- Se un impiegato interrompe la partecipazione a tutti i progetti, dobbiamo cancellarlo
  - **anomalia di cancellazione**
- Un nuovo impiegato senza progetto non può essere inserito
  - **anomalia di inserimento**

<u>Impiegato</u>	<u>Stipendio</u>	<u>Progetto</u>	<u>Bilancio</u>	<u>Funzione</u>
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
Mori	48	Marte	2	direttore
Mori	48	Venere	15	progettista
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	direttore

# Perché questi fenomeni indesiderabili?

Unica relazione per rappresentare informazioni eterogenee:

- Impiegati con i relativi stipendi
- Progetti con i relativi bilanci
- Partecipazione degli impiegati ai progetti con le relative funzioni

<u>Impiegato</u>	<u>Stipendio</u>	<u>Progetto</u>	<u>Bilancio</u>	<u>Funzione</u>
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
Mori	48	Marte	2	direttore
Mori	48	Venere	15	progettista
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	direttore

**Per studiare in maniera sistematica questi aspetti, è necessario introdurre un vincolo di integrità:  
la dipendenza funzionale**



# Proprietà

- Ogni impiegato ha un solo stipendio (anche se partecipa a più progetti)
- Ogni progetto ha un bilancio
- Ogni impiegato in ciascun progetto ha una sola funzione (anche se può avere funzioni diverse in progetti diversi)

<u>Impiegato</u>	<u>Stipendio</u>	<u>Progetto</u>	<u>Bilancio</u>	<u>Funzione</u>
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
Mori	48	Marte	2	direttore
Mori	48	Venere	15	progettista
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	direttore

# Dipendenza Funzionale

- Data una  $R(X)$  sugli attributi  $X$
- Dati due insiemi di attributi  $Y, Z \subset X$
- esiste in  $R$  una dipendenza funzionale (FD) da  $Y$  a  $Z$  se:
- Per ogni coppia di tuple  $t_1, t_2 \in R$ :  
se  $\pi_Y(t_1) = \pi_Y(t_2)$  allora  $\pi_Z(t_1) = \pi_Z(t_2)$

# Notazione

$$Y \rightarrow Z$$

Esempi:

Impiegato  $\rightarrow$  Stipendio

Progetto  $\rightarrow$  Bilancio

Impiegato Progetto  $\rightarrow$  Funzione

<u>Impiegato</u>	Stipendio	<u>Progetto</u>	Bilancio	Funzione
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
Mori	48	Marte	2	direttore
Mori	48	Venere	15	progettista
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	direttore

**Impiegato → Stipendio**

**Progetto → Bilancio**

**Impiegato Progetto → Funzione**

# Dipendenze Funzionali Banali

- Impiegato Progetto  $\rightarrow$  Progetto
- Si tratta però di una DF “banale”:
- $Y \rightarrow A$  è banale se  $A \subset Y$

# Anomalie e Dipendenze Funzionali

Le anomalie sono legate ad alcune dipendenze funzionali:

- Gli impiegati hanno un unico stipendio

Impiegato → Stipendio

- I progetti hanno un unico bilancio

Progetto → Bilancio

<u>Impiegato</u>	<u>Stipendio</u>	<u>Progetto</u>	<u>Bilancio</u>	<u>Funzione</u>
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
Mori	48	Marte	2	direttore
Mori	48	Venere	15	progettista
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	direttore

## Dipendenze Funzionali “Buone”

- Una dipendenza funzionale “buona” è tra una chiave e altri attributi. Per esempio:

Impiegato Progetto → Funzione

- Una dipendenza funzionale “buona” non causa anomalie

<u>Impiegato</u>	Stipendio	<u>Progetto</u>	Bilancio	Funzione
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
Mori	48	Marte	2	direttore
Mori	48	Venere	15	progettista
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	direttore

# Forma Normale di Boyce e Codd (BCNF)

- Una relazione  $R$  con chiavi  $K_1, \dots, K_n$ , è in forma normale di Boyce e Codd:  
se ogni dipendenza funzionale non banale  $X \rightarrow Y$  è «buona» cioè  $\exists i. K_i \subseteq X$  ( $X$  è superchiave)
- La forma normale richiede che i concetti in una relazione siano omogenei (solo proprietà direttamente associate alla chiave)



# Che facciamo se una relazione non soddisfa la BCNF?

- La rimpiazziamo con altre relazioni che soddisfano la BCNF

Come?

- Decomponendo sulla base delle dipendenze funzionali, al fine di separare i concetti

<u>Impiegato</u>	Stipendio	<u>Progetto</u>	Bilancio	Funzione
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
Mori	48	Marte	2	direttore
Mori	48	Venere	15	progettista
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	direttore

**Impiegato → Stipendio**

**Progetto → Bilancio**

**Impiegato Progetto → Funzione**

<u>Impiegato</u>	Stipendio
Rossi	20
Verdi	35
Verdi	35
Neri	55
Neri	55
Neri	55
Mori	48
Mori	48
Bianchi	48
Bianchi	48

**Impiegato → Stipendio**

<u>Progetto</u>	Bilancio
Marte	2
Giove	15
Venere	15
Venere	15
Giove	15
Marte	2
Marte	2
Venere	15
Venere	15
Giove	15

**Progetto → Bilancio**

<u>Impiegato</u>	<u>Progetto</u>	Funzione
Rossi	Marte	tecnico
Verdi	Giove	progettista
Verdi	Venere	progettista
Neri	Venere	direttore
Neri	Giove	consulente
Neri	Marte	consulente
Mori	Marte	direttore
Mori	Venere	progettista
Bianchi	Venere	progettista
Bianchi	Giove	direttore

**Impiegato Progetto → Funzione**

# Procedura intuitiva di normalizzazione

Per ogni dipendenza  $X \rightarrow Y$  che viola la BCNF, definire una relazione su  $XY$  ed eliminare  $Y$  dalla relazione originaria

# Esercizio

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

1. Individuare le proprietà della relazione:
  - Chiave/i della relazione
  - Dipendenze funzionali
2. Identificare ridondanze o anomalie
3. Decomporre in BCNF

# Esercizio

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

1. Individuare le proprietà della relazione:
  - Chiave/i della relazione: **Dipart, Facoltà, Corso**
  - Dipendenze funzionali: **Facoltà → Preside**
2. Identificare ridondanze o anomalie



# Esercizio

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

## 2. Identificare ridondanze o anomalie

Ridondanza: Preside ripetuto ogni volta la facoltà è la stessa

Anomalia Aggiornamento: Se il preside cambia, occorre assicurarsi che cambia in tutte le tuple

Anomalia Cancellazione: Se tutti i corsi di Scienze vengono rimossi, si perde l'informazione che il preside è Bruni

# Esercizio

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

1. Individuare le proprietà della relazione:
  - Chiave/i della relazione: **Dipart, Facoltà, Corso**
  - Dipendenze funzionali: **Facoltà → Preside**
3. Decomporre in Forma Normale di Boyce Codd (BCNF)

# Esercizio

		<u>Docente</u>	<u>Dipartimento</u>	<u>Facoltà</u>	<u>Corso</u>
<u>Facoltà</u>	<u>Preside</u>	Verdi	Matematica	Ingegneria	Analisi
Ingegneria	Neri	Verdi	Matematica	Ingegneria	Geometria
Scienze	Bruni	Rossi	Fisica	Ingegneria	Analisi
		Rossi	Fisica	Scienze	Analisi
		Bruni	Fisica	Scienze	Fisica

- È in forma normale di **Boyce-Codd**  
 → Decomposizione **risolve le anomalie**

# Non sempre così facile

Impiegato	Progetto	Sede
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Venere	Milano
Neri	Saturno	Milano
Neri	Venere	Milano

**Impiegato → Sede**  
**Progetto → Sede**

# Decomposizione basata su dipendenze

Impiegato	Progetto	Sede
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Venere	Milano
Neri	Saturno	Milano
Neri	Venere	Milano

Impiegato	Sede
Rossi	Roma
Verdi	Milano
Neri	Milano

**Impiegato → Sede**

Progetto	Sede
Marte	Roma
Giove	Milano
Saturno	Milano
Venere	Milano

**Progetto → Sede**

## Proviamo a ricostruire

Impiegato	Sede
Rossi	Roma
Verdi	Milano
Neri	Milano

Progetto	Sede
Marte	Roma
Giove	Milano
Saturno	Milano
Venere	Milano

Impiegato	Progetto	Sede
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Venere	Milano
Neri	Saturno	Milano
Neri	Venere	Milano
Verdi	Saturno	Milano
Neri	Giove	Milano

**Diversa dalla relazione di partenza!**

# Decomposizione senza perdita

- Una relazione  $R(X)$  si **decompone senza perdita** su  $X_1, X_2 \subset X$  con  $X_1 \cup X_2 = X$  se  $\pi_{X_1}(X) \bowtie \pi_{X_2}(X) = X$
- La decomposizione senza perdita è garantita se gli attributi in  $X_1 \cap X_2$  sono una **chiave** della relazione  $\pi_{X_1}(X)$  e/o di  $\pi_{X_2}(X)$

# Decomposizione con Perdita

- Questo caso c'è perdita: l'attributo comune  $X1 \cap X2 = \{\text{Sede}\}$  non è chiave nè di R1 nè di R2  $\rightarrow$  Perdita

R1

<u>Impiegato</u>	Sede
Rossi	Roma
Verdi	Milano
Neri	Milano

**Impiegato  $\rightarrow$  Sede**

R2

<u>Progetto</u>	Sede
Marte	Roma
Giove	Milano
Saturno	Milano
Venere	Milano

**Progetto  $\rightarrow$  Sede**



# Decomposizione Senza Perdita

<u>Impiegato</u>	Stipendio
Rossi	20
Verdi	35
Verdi	35
Neri	55
Neri	55
Neri	55
Mori	48
Mori	48
Bianchi	48
Bianchi	48



<u>Impiegato</u>	<u>Progetto</u>	Funzione
Rossi	Marte	tecnico
Verdi	Giove	progettista
Verdi	Venere	progettista
Neri	Venere	direttore
Neri	Giove	consulente
Neri	Marte	consulente
Mori	Marte	direttore
Mori	Venere	progettista
Bianchi	Venere	progettista
Bianchi	Giove	direttore



<u>Progetto</u>	Bilancio
Marte	2
Giove	15
Venere	15
Venere	15
Giove	15
Marte	2
Marte	2
Venere	15
Venere	15
Giove	15

Impiegato → Stipendio

Impiegato Progetto  
→ Funzione

Progetto →  
Bilancio

Attributi Comuni:

**Impiegato**  
(Chiave di Relazione di Sinistra)

Attributi Comuni:

**Progetto**  
(Chiave di Relazione di Sinistra)

## Senza Perdità?

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

Sì, il join naturale ricostruisce la “relazione di partenza”:

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

# Proviamo a decomporre senza perdita sfruttando solo Impiegato → Sede

Impiegato	Progetto	Sede
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Venere	Milano
Neri	Saturno	Milano
Neri	Venere	Milano

<u>Impiegato</u>	Sede
Rossi	Roma
Verdi	Milano
Neri	Milano

<u>Impiegato</u>	<u>Progetto</u>
Rossi	Marte
Verdi	Giove
Verdi	Venere
Neri	Saturno
Neri	Venere

# Un problema: Conservazione Dipendenze /1

Aggiungiamo tupla **in verde** circa l'impiegato Neri, che opera a Milano, al progetto Marte:

Impiegato	Sede
Rossi	Roma
Verdi	Milano
Neri	Milano

Impiegato	Progetto
Rossi	Marte
Verdi	Giove
Verdi	Venere
Neri	Saturno
Neri	Venere
Neri	Marte

Impiegato → Sede

~~Progetto~~ → Sede

# Un problema: Conservazione Dipendenze /2

Ricostruiamo la relazione di partenza con il join:

Impiegato	Progetto	Sede
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Venere	Milano
Neri	Saturno	Milano
Neri	Venere	Milano
Neri	Marte	Milano

Impiegato → Sede

~~Progetto → Sede~~

# Conservazione delle Dipendenze

- Una decomposizione **conserva le dipendenze** se ciascuna delle dipendenze funzionali dello schema originario coinvolge attributi che compaiono tutti insieme in uno degli schemi decomposti
- **Progetto** → **Sede** non è conservata

# Esercizio

Assumendo la tabella di cui sotto (con una chiave quella in rosso):

1. Trovare le Dipendenze Funzionali che violano la forma normale di Boyce Codd e non sono banali
2. Decomporre in Forma Normale di Boyce-Codd

<u>Prodotto</u>	<u>Componente</u>	<u>Tipo</u>	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

**Q = Quantità**

**PC = Prezzo Unitario** (per tutte le quantità insieme e non per unità di quantità)

**PT = Prezzo Totale**

# Soluzione: Dipendenze Funzionali

- Dipendenze Funzionali che violano BCNF:
  - Prodotto  $\rightarrow$  PT
  - Componente, Prodotto  $\rightarrow$  Fornitore



# Soluzione: Decomposizione per Prodotto → PT

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
Libreria	Legno	Noce	50	10.000	Forrest
Libreria	Bulloni	B212	200	100	Bolt
Libreria	Vetro	Cristal	3	5.000	Clean
Scaffale	Legno	Mogano	5	15.000	Forrest
Scaffale	Bulloni	B212	250	100	Bolt
Scaffale	Bulloni	B412	150	300	Bolt
Scrivania	Legno	Noce	10	8.000	Wood
Scrivania	Maniglie	H621	10	20.000	Bolt
Tavolo	Legno	Noce	4	10.000	Forrest

<u>Prodotto</u>	PT
Libreria	400.000
Scaffale	300.000
Scrivania	250.000
Tavolo	200.000

# Soluzione: Decomposizione per Componente, Prodotto → Fornitore

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

Prodotto	Componente	Tipo	Q	PC
Libreria	Legno	Noce	50	10.000
Libreria	Bulloni	B212	200	100
Libreria	Vetro	Cristal	3	5.000
Scaffale	Legno	Mogano	5	15.000
Scaffale	Bulloni	B212	250	100
Scaffale	Bulloni	B412	150	300
Scrivania	Legno	Noce	10	8.000
Scrivania	Maniglie	H621	10	20.000
Tavolo	Legno	Noce	4	10.000

Prodotto	Componente	Fornitore
Libreria	Legno	Forrest
Libreria	Bulloni	Bolt
Libreria	Vetro	Clean
Scaffale	Legno	Forrest
Scaffale	Bulloni	Bolt
Scrivania	Legno	Wood
Scrivania	Maniglie	Bolt
Tavolo	Legno	Forrest

# Soluzione Finale

<u>Prodotto</u>	<u>Componente</u>	<u>Tipo</u>	<u>Q</u>	<u>PC</u>	<u>Fornitore</u>	<u>PT</u>
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

<u>Prodotto</u>	<u>Componente</u>	<u>Tipo</u>	<u>Q</u>	<u>PC</u>
Libreria	Legno	Noce	50	10.000
Libreria	Bulloni	B212	200	100
Libreria	Vetro	Cristal	3	5.000
Scaffale	Legno	Mogano	5	15.000
Scaffale	Bulloni	B212	250	100
Scaffale	Bulloni	B412	150	300
Scrivania	Legno	Noce	10	8.000
Scrivania	Maniglie	H621	10	20.000
Tavolo	Legno	Noce	4	10.000

<u>Prodotto</u>	<u>Componente</u>	<u>Fornitore</u>
Libreria	Legno	Forrest
Libreria	Bulloni	Bolt
Libreria	Vetro	Clean
Scaffale	Legno	Forrest
Scaffale	Bulloni	Bolt
Scrivania	Legno	Wood
Scrivania	Maniglie	Bolt
Tavolo	Legno	Forrest

<u>Prodotto</u>	<u>PT</u>
Libreria	400.000
Scaffale	300.000
Scrivania	250.000
Tavolo	200.000

# Qualità delle Decomposizioni

Una decomposizione dovrebbe sempre soddisfare:

- **Decomposizione senza perdita:**  
possibile ricreare informazioni originarie  
(attributi comuni nel join di A e B sono chiave di A e/o B)
- **Conservazione delle dipendenze:**  
vincoli di integrità originari sono mantenuti  
(per ogni dipendenza funzionale, esiste una relazione con tutti i suoi attributi)

# Relazioni non normalizzabile in Forma Normale di Boyce-Codd

<b>Dirigente</b>	<b><u>Progetto</u></b>	<b><u>Sede</u></b>
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Marte	Milano
Neri	Saturno	Milano
Neri	Venere	Milano

**Progetto Sede → Dirigente**  
**Dirigente → Sede**

**Progetto Sede → Dirigente**

coinvolge tutti gli attributi e  
quindi nessuna  
decomposizione può  
preservare tale dipendenza

➔ la BCNF non è  
sempre raggiungibile  
raggiungibile”

# Una possibile riorganizzazione

<b>Dirigente</b>	<b><u>Progetto</u></b>	<b><u>Sede</u></b>	<b>Reparto</b>
Rossi	Marte	Roma	1
Verdi	Giove	Milano	1
Verdi	Marte	Milano	1
Neri	Saturno	Milano	2
Neri	Venere	Milano	2

**Sede Reparto → Dirigente**  
**Dirigente → Sede Reparto**  
**Progetto Sede → Reparto**

# Decomposizione in BCNF

Chiavi:

- Dirigente
- Sede, Reparto

Dirigente	Sede	Reparto
Rossi	Roma	1
Verdi	Milano	1
Neri	Milano	2

Chiavi:

- Sede, Progetto

Progetto	Sede	Reparto
Marte	Roma	1
Giove	Milano	1
Marte	Milano	1
Saturno	Milano	2
Venere	Milano	2

- **Decomposizione senza perdita:** gli attributi comuni {Sede, Reparto} sono una chiave della relazione sx
- **Tutte le dipendenze sono rispettate**

# Terza Forma Normale

Una relazione  $R$  con chiavi  $K_1, \dots, K_n$  è in Terza Forma Normale se:

Per ogni dipendenza funzionale non banale  $X \rightarrow Y$ , almeno una delle seguenti condizioni sono valide:

- $X$  è superchiave (BCNF)
- ogni attributo in  $Y$  è contenuto in almeno una tra le chiavi  $K_1, \dots, K_n$ .



# BCNF e Terza Forma Normale (3NF)

<b>Dirigente</b>	<b><u>Progetto</u></b>	<b><u>Sede</u></b>
Rossi	Marte	Roma
Verdi	Giove	Milano
Verdi	Marte	Milano
Neri	Saturno	Milano
Neri	Venere	Milano

**Progetto Sede → Dirigente**  
**Dirigente → Sede**

- Pro: Sempre «raggiungibile»
- Contro: Meno restrittiva della BCNF

Alcune Dipendenze  
Escluse → Ammette  
alcune anomalie

Esempio:

- Ridondanza tra Dirigente e Sede
- Possibilità di un Dirigente di «apparire» in più sedi

# Decomposizione in Terza Forma Normale

- Si crea una relazione per ogni gruppo di attributi coinvolti in una dipendenza funzionale
- Si verifica che alla fine una relazione contenga una chiave della relazione originaria

# Una possibile strategia

- Se la relazione non è normalizzata si decompone in terza forma normale
- alla fine si verifica se lo schema ottenuto è anche in BCNF

# Teoria della normalizzazione

Data una relazione  $R$  e un insieme di dipendenze funzionali definite su  $R$ , generare una decomposizione di  $R$  che:

- Senza perdita
- Conservi le dipendenze
- Contenga solo relazioni normalizzate (possibilmente in BCNF, sicuramente in 3NF)

# Implicazione dipendenze funzionali

- Un insieme  $F$  di FD **implica** un'altra FD  $f$  se ogni relazione che soddisfa tutte le FD in  $F$  soddisfa anche  $f$ .

Esempio:  $R(\text{Impiegato}, \text{Categoria}, \text{Stipendio})$

*$\text{Impiegato} \rightarrow \text{Categoria}$*

*$\text{Categoria} \rightarrow \text{Stipendio}$*

implicano

*$\text{Impiegato} \rightarrow \text{Stipendio}$* .

# Chiusura di un insieme di attributi

- Dati uno schema di Relazione  $R(U)$ ,
- Un insieme  $F$  di Dip. Funz. definite su  $U$
- Un insieme di attributi  $X \subseteq U$
- La **chiusura**  $X_F^+$  di  $X$  rispetto ad  $F$ , è l'insieme degli attributi che dipendono funzionalmente da  $X$ :
$$X_F^+ = \{ A \mid A \in U \text{ e } F \text{ implica } X \rightarrow A \}$$
- Si noti che se  $\{ X_1, X_2 \} \rightarrow A$  allora  $\{ X_1, X_2, \dots \} \rightarrow A$
- Se  $A$  appartiene a  $X_F^+$  allora  $X \rightarrow A$  è implicata da  $F$
- Se  $X_F^+ = U$ , allora  $X \rightarrow U$ . Quindi,  **$X$  è candidata ad essere chiave**

# Calcolo di $X_F^+$

<u>Impiegato</u>	<u>Stipendio</u>	<u>Progetto</u>	<u>Bilancio</u>	<u>Funzione</u>
Rossi	20	Marte	2	tecnico
Verdi	35	Giove	15	progettista
Verdi	35	Venere	15	progettista
Neri	55	Venere	15	direttore
Neri	55	Giove	15	consulente
Neri	55	Marte	2	consulente
Mori	48	Marte	2	direttore
Mori	48	Venere	15	progettista
Bianchi	48	Venere	15	progettista
Bianchi	48	Giove	15	direttore

Impiegato →  
Stipendio

Progetto →  
Bilancio

Impiegato Progetto →  
Funzione

$X = \{ \text{Impiegato, Progetto} \}$

$X_p = \{ \text{Impiegato, Progetto} \}$

Da **Impiegato → Stipendio**:

$X_p = \{ \text{Impiegato, Progetto, Stipendio} \}$

Da **Progetto → Bilancio** aggiungiamo Bilancio:

$X_p = \{ \text{Impiegato, Progetto, Stipendio, Bilancio} \}$

Da **Impiegato Progetto → Funzione** aggiungiamo Funzione:

$X_p = \{ \text{Impiegato, Progetto, Stipendio, Bilancio, Funzione} \}$

## Calcolo di $X_F^+$

**Input:** un insieme  $X$  di attributi e un insieme  $F$  di dipendenze funzionali

**Output:** un insieme  $X_P$  di attributi.

1. Inizializziamo  $X_P$  con l'insieme di input  $X$ .
2. Se esiste una FD  $Y \rightarrow A$  in  $F$  con  $Y \subseteq X_P$  e  $A \notin X_P$ , allora aggiungiamo  $A$  a  $X_P$ .
3. Ripetiamo il passo (2) fino a quando non ci sono ulteriori attributi che possono essere aggiunti a  $X_P$ .



# Coperture di dipendenze funzionali

- Due insiemi di dipendenze funzionali  $F_1$  ed  $F_2$  sono **equivalenti** se  $F_1$  implica ciascuna dipendenza in  $F_2$  e viceversa.  
Per es.  $\{A \rightarrow B; AB \rightarrow C; A \rightarrow C\}$  e  $\{A \rightarrow B; AB \rightarrow C\}$  sono equivalenti
- Se due insiemi sono equivalenti diciamo anche che ognuno è una **copertura** dell'altro.
- Tra due equivalenti, meglio uno più semplice

# Proprietà desiderabili di FD

- Un insieme di dipendenze  $F$  è:
  - **non ridondante** se non esiste dipendenza  $f \in F$  tale che  $F - \{f\}$  implica  $f$ ;
  - **ridotto** se (i) è non ridondante e (ii) non esiste un insieme  $F'$  equivalente a  $F$  ottenuto eliminando attributi dai primi membri di una o più dipendenze di  $F$ .
- Esempio (parte in rosso rimovibile):
  - $\{A \rightarrow B; AB \rightarrow C; A \rightarrow C\}$  è ridondante;
  - $\{A \rightarrow B; AB \rightarrow C\}$  non è ridondante né ridotto;
  - $\{A \rightarrow B; A \rightarrow C\}$  è ridotto

# Calcolo copertura ridotta

F:  $M \rightarrow RSDG$ ,  $MS \rightarrow CD$ ,  $G \rightarrow R$ ,  $D \rightarrow S$ ,  $S \rightarrow D$ ,  
 $MPD \rightarrow AM$ .

1. Sostituiamo l'insieme dato con quello equivalente dove i secondi membri sono singoli attributi;

$M \rightarrow R$ ,  $M \rightarrow S$ ,  $M \rightarrow D$ ,  $M \rightarrow G$ ,  $MS \rightarrow C$ ,  $MS \rightarrow D$ ,  $G \rightarrow R$ ,  $D \rightarrow S$ ,  $S \rightarrow D$ ,  
 $MPD \rightarrow A$ ,  $MPD \rightarrow M$

2. Eliminiamo le dipendenze ridondanti;

$M \rightarrow D$ ,  $M \rightarrow G$ ,  $MS \rightarrow C$ ,  $MS \rightarrow D$ ,  $G \rightarrow R$ ,  $D \rightarrow S$ ,  $S \rightarrow D$ ,  $MPD \rightarrow A$

3. Per ogni dipendenza verifichiamo se esistono attributi eliminabili dal primo membro

$M \rightarrow D$ ,  $M \rightarrow G$ ,  $MS \rightarrow C$ ,  $MS \rightarrow D$ ,  $G \rightarrow R$ ,  $D \rightarrow S$ ,  $S \rightarrow D$ ,  $MPD \rightarrow A$

# Sintesi di schemi in 3NF / 1

Dati uno schema  $R(U)$  e un insieme di dipendenze  $F$  su  $U$ , con chiavi  $K_1, \dots, K_n$ .

$R(\underline{M}\underline{C}\underline{G}\underline{R}\underline{D}\underline{S}\underline{P}\underline{A})$

$F: M \rightarrow RSDG, MS \rightarrow CD, G \rightarrow R, D \rightarrow S, S \rightarrow D, MPD \rightarrow AM.$

1. Viene calcolata una copertura ridotta  $G$  di  $F$ :

$M \rightarrow D, M \rightarrow G, M \rightarrow C, G \rightarrow R, D \rightarrow S, S \rightarrow D, MP \rightarrow A$

## Sintesi di schemi in 3NF / 2

Dati uno schema  $R(U)$  e un insieme di dipendenze  $F$  su  $U$ , con chiavi  $K_1, \dots, K_n$ .

$R(\underline{M}CGRDSPA)$

$F: M \rightarrow RSDG, MS \rightarrow CD, G \rightarrow R, D \rightarrow S, S \rightarrow D, MPD \rightarrow AM.$

1. Viene calcolata una copertura ridotta  $G$  di  $F$ :

$M \rightarrow D, M \rightarrow G, M \rightarrow C, G \rightarrow R, D \rightarrow S, S \rightarrow D, MP \rightarrow A$

2.  $G$  viene partizionato in sottoinsiemi tali che due dip. funzion.  $X \rightarrow A$  e  $Y \rightarrow B$  sono insieme se  $X^+_G = Y^+_G$ :

$\{M \rightarrow D, M \rightarrow G, M \rightarrow C\}, \{G \rightarrow R\}, \{D \rightarrow S, S \rightarrow D\},$   
 $\{MP \rightarrow A\}$

# Sintesi di schemi in 3NF / 3

Dati uno schema  $R(U)$  e un insieme di dipendenze  $F$  su  $U$ ,  
con chiavi  $K_1, \dots, K_n$ .

$R(\underline{M}CGRDSPA)$

$F: M \rightarrow RSDG, MS \rightarrow CD, G \rightarrow R, D \rightarrow S, S \rightarrow D, MPD \rightarrow AM.$

2.  $G$  viene partizionato in sottoinsiemi tali che due dip. funzion.  $X \rightarrow A$  e  $Y \rightarrow B$  sono insieme se  $X^+_G = Y^+_G$ :

$\{M \rightarrow D, M \rightarrow G, M \rightarrow C\}, \{G \rightarrow R\}, \{D \rightarrow S, S \rightarrow D\},$   
 $\{MP \rightarrow A\}$

3. Viene costruita una relazione per ogni sotto-insieme:

$R_1(M, D, C, G), R_2(G, R), R_3(D, S), R_4(M, P, A)$

# Sintesi di schemi in 3NF / 4

Dati uno schema  $R(U)$  e un insieme di dipendenze  $F$  su  $U$ ,  
con chiavi  $K_1, \dots, K_n$ .

$R(\underline{M}CGRDSPA)$

$F: M \rightarrow RSDG, MS \rightarrow CD, G \rightarrow R, D \rightarrow S, S \rightarrow D, MPD \rightarrow AM.$

- Viene costruita una relazione per ogni sotto-insieme :

$R_1(M, D, C, G), R_2(G, R), R_3(D, S), R_4(M, P, A)$

- Se esistono due relazione  $S(X)$  and  $T(Y)$  con  $X \subseteq Y$ ,  $S$  viene eliminata :

Non accede. Quindi stesse relazioni:

$R_1(M, D, C, G), R_2(G, R), R_3(D, S), R_4(M, P, A)$

# Sintesi di schemi in 3NF / 5

Dati uno schema  $R(U)$  e un insieme di dipendenze  $F$  su  $U$ ,  
con chiavi  $K_1, \dots, K_n$ .

$R(\underline{M}CGRDSPA)$

$F: M \rightarrow RSDG, MS \rightarrow CD, G \rightarrow R, D \rightarrow S, S \rightarrow D, MPD \rightarrow AM.$

4. Se esistono due relazione  $S(X)$  and  $T(Y)$  con  $X \subseteq Y$ ,  
 $S$  viene eliminata :

Non accede. Quindi stesse relazioni:

$R_1(M, D, C, G), R_2(G, R), R_3(D, S), R_4(M, P, A)$

5. Se, per qualche  $i$ , non esiste una relazione  $S(X)$  con  $K_i \subseteq X$ , viene aggiunta una relazione  $T(K_i)$  :

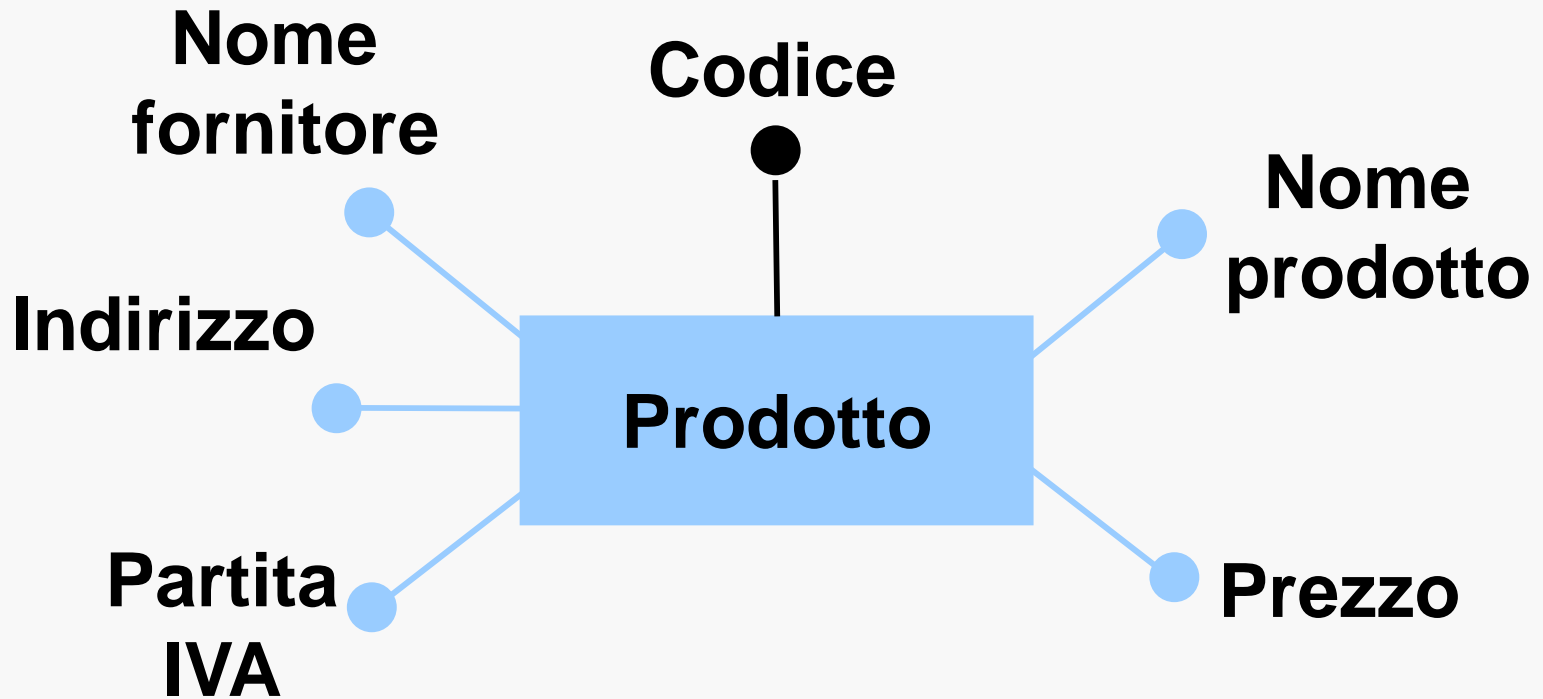
Non accede. Perché  $R_4$  contiene  $M$  e  $P$ :

$R_1(M, D, C, G), R_2(G, R), R_3(D, S), R_4(M, P, A)$



# Progettazione e normalizzazione

- la teoria della normalizzazione può essere usata nella progettazione logica per verificare lo schema relazionale finale
- si può usare anche durante la progettazione concettuale per verificare la qualità dello schema concettuale



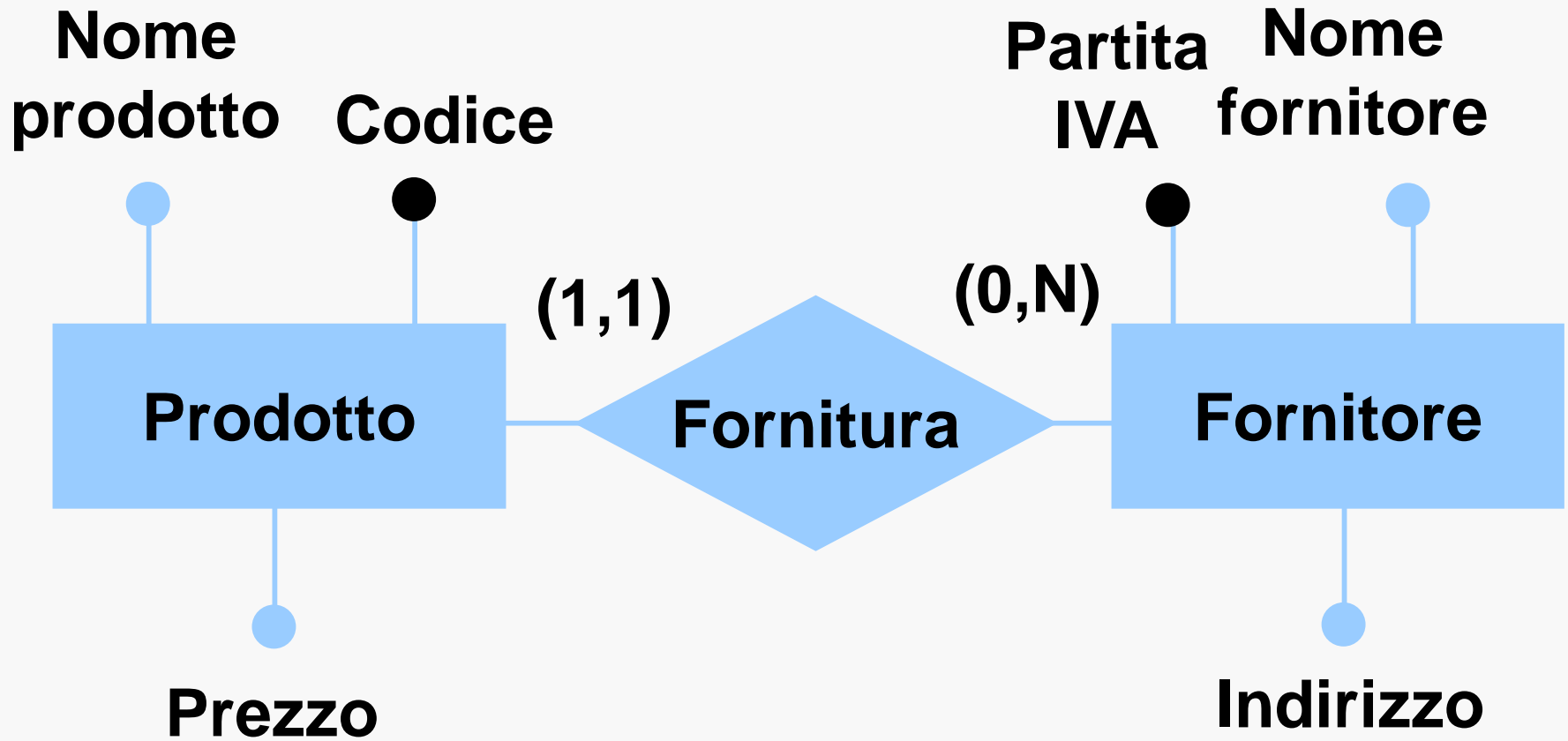
**PartitaIVA → NomeFornitore Indirizzo**

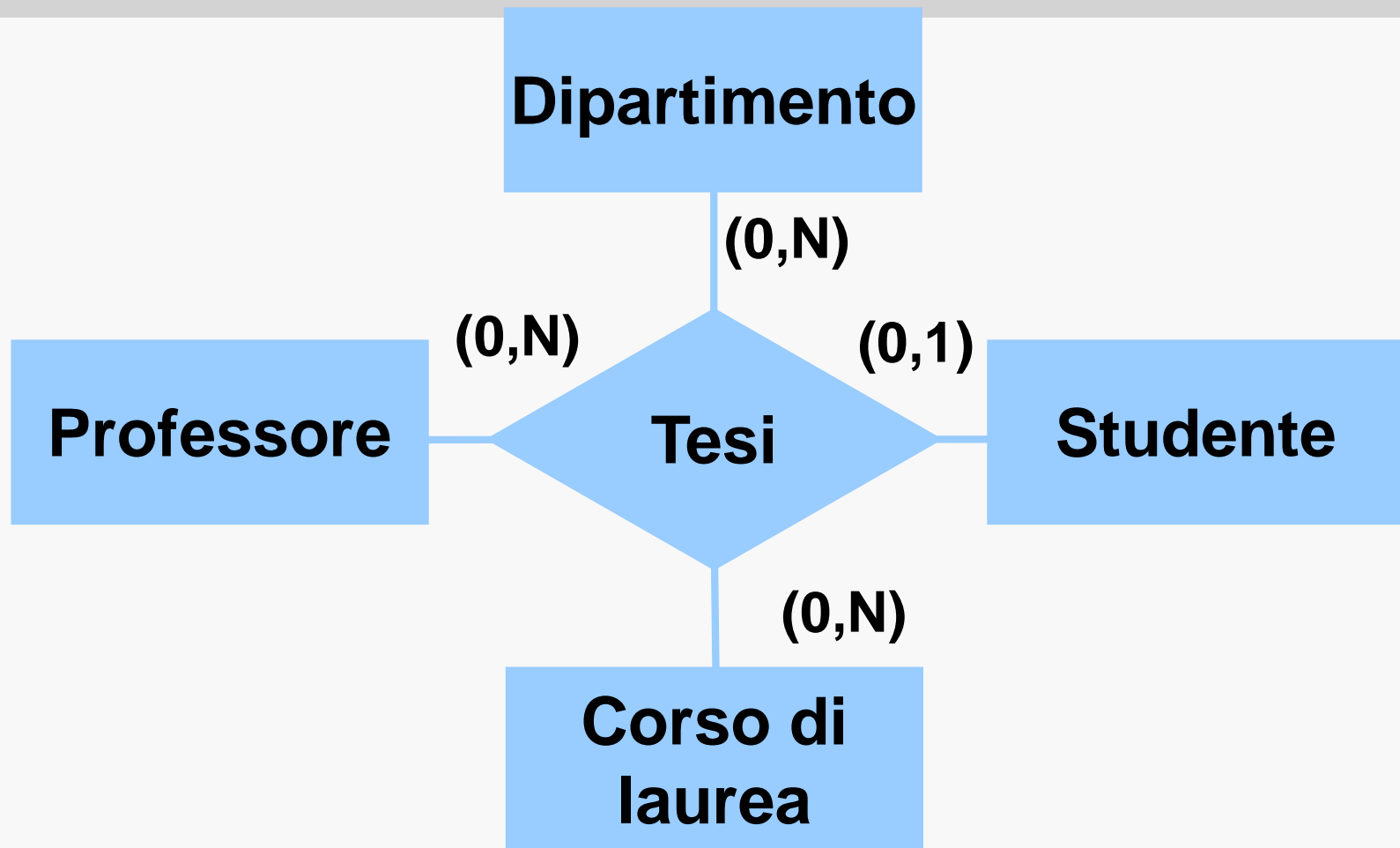
## Analisi dell'entità

- L'entità viola la forma normale a causa della dipendenza:

PartitaIVA → NomeFornitore Indirizzo

- Possiamo decomporre sulla base di questa dipendenza





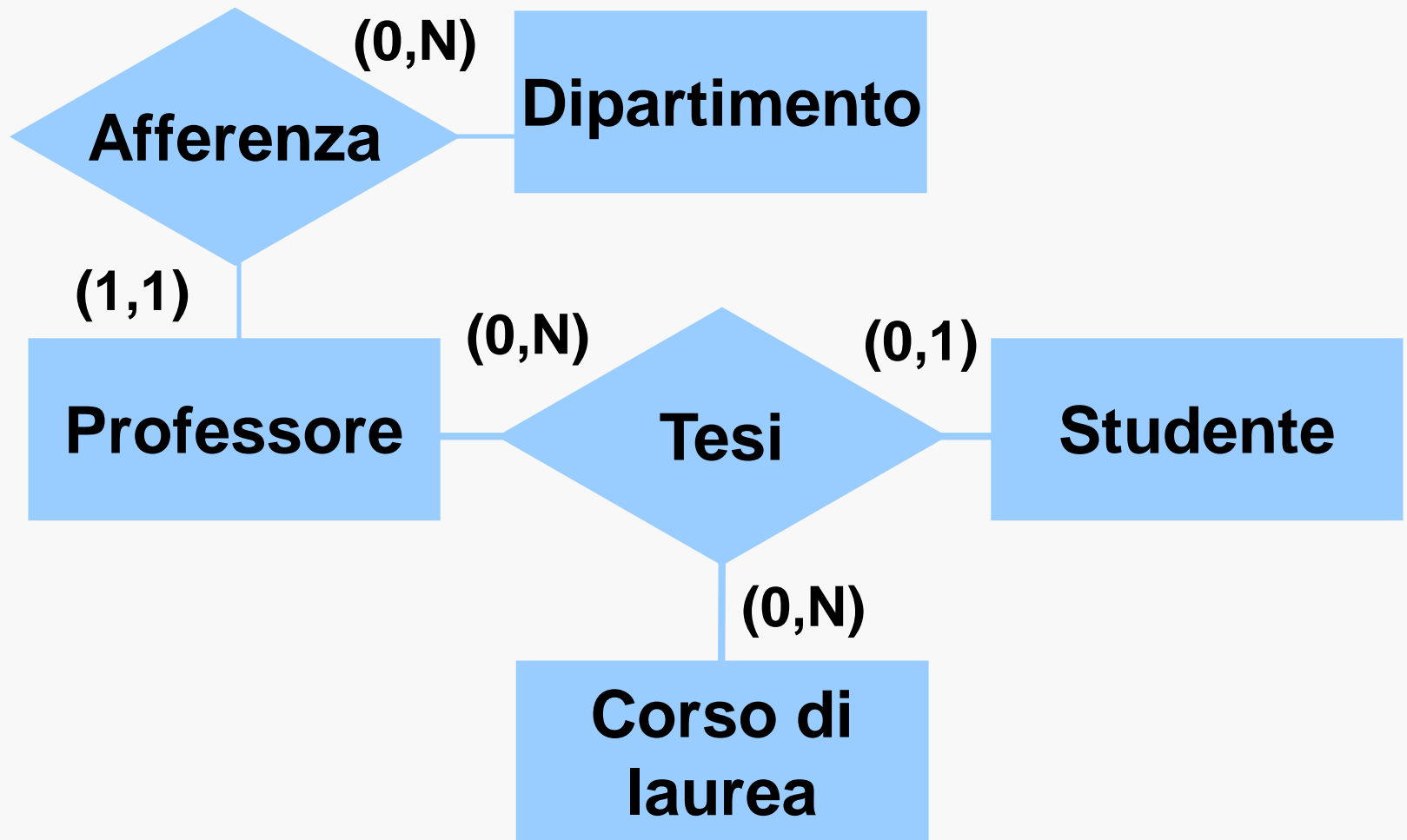
Studente → Corso di laurea  
Studente → Professore  
Professore → Dipartimento

# Analisi della relationship

- La relationship viola la terza forma normale a causa della dipendenza:

Professore → Dipartimento

- Possiamo decomporre sulla base di questa dipendenza



# Ulteriore analisi sulla base delle dipendenze

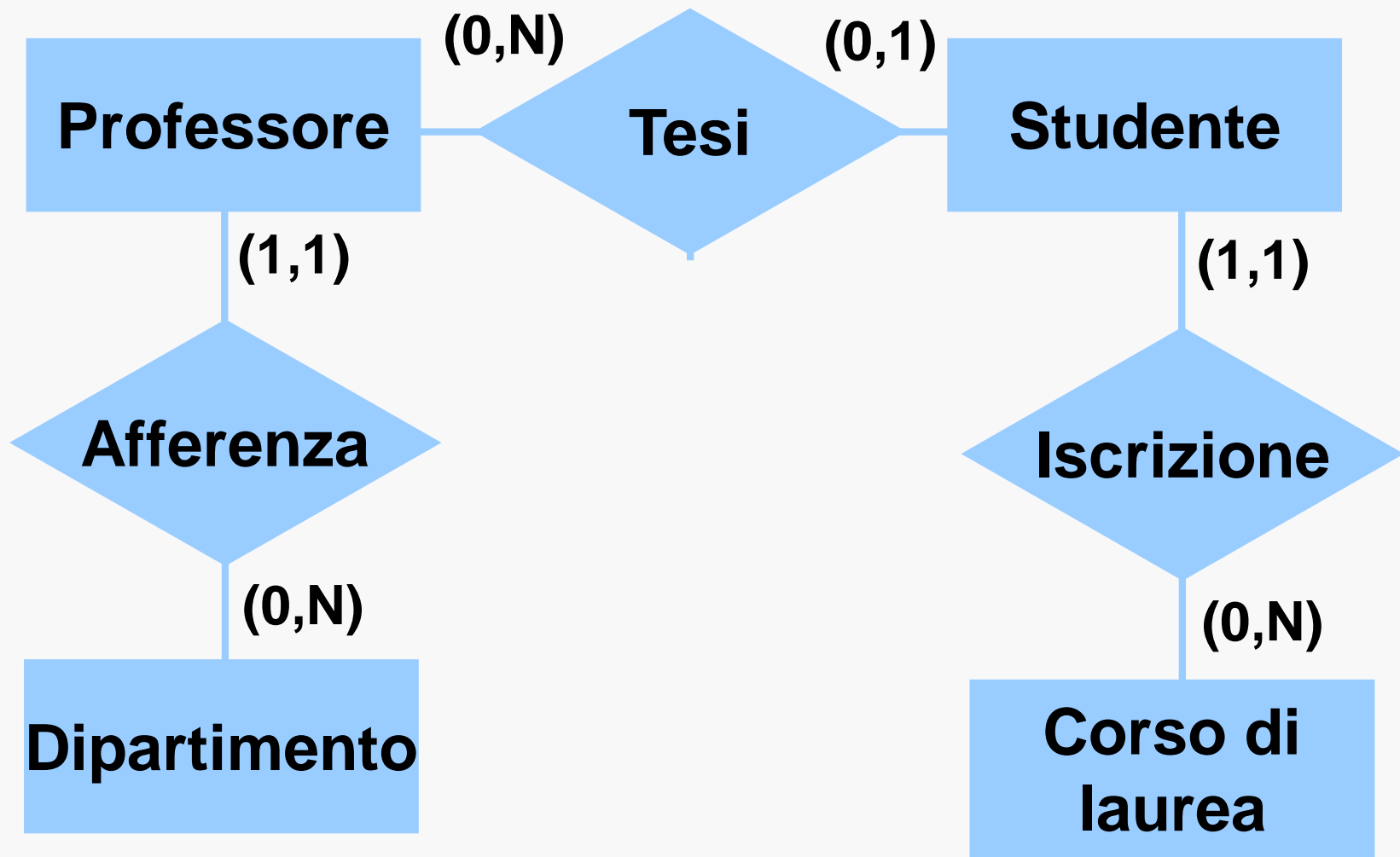
- La relationship **Tesi** è in BCNF sulla base delle dipendenze

**Studiante** → **CorsoDiLaurea**

**Studiante** → **Professore**

- le due proprietà sono indipendenti
- questo suggerisce una ulteriore decomposizione





# Riferimenti

Capitolo 9, escludendo:

- Dimostrazione della terminazione alle pagine 344-345