



UNIVERSITÀ  
DEGLI STUDI  
DI MILANO

LA STATALE

# Normalizzazione di schemi

Dipendenze funzionali e forme normali

Prof. Stefano Montanelli

# Dipendenza funzionale

Una **dipendenza funzionale**  $X \rightarrow Y$  tra due sottinsiemi di attributi  $X$  e  $Y$  di una relazione  $R$  stabilisce un vincolo sulle ennuple che possono formare uno stato di relazione  $r$  di  $R$

Il vincolo stabilisce che, per ogni coppia di ennuple  $t_1$  e  $t_2$  in  $r$  per cui  $t_1[X] = t_2[X]$ , si ha  $t_1[Y] = t_2[Y]$ , ovvero  $t_1[X] = t_2[X] \rightarrow t_1[Y] = t_2[Y]$

Se  $X$  è una chiave di  $R$ , allora  $X \rightarrow Y$  vale per ogni sottinsieme  $Y$  di attributi di  $R$

$X \rightarrow Y$  **NON implica**  $Y \rightarrow X$

# Regole di inferenza

1. Regola riflessiva:  $X \supseteq Y \models X \rightarrow Y$
2. Regola di arricchimento:  $\{X \rightarrow Y\} \models XZ \rightarrow YZ$
3. Regola transitiva:  $\{X \rightarrow Y, Y \rightarrow Z\} \models X \rightarrow Z$
4. Regola di decomposizione:  $\{X \rightarrow YZ\} \models X \rightarrow Y, X \rightarrow Z$
5. Regola di unione:  $\{X \rightarrow Y, X \rightarrow Z\} \models X \rightarrow YZ$
6. Regola pseudo-transitiva:  $\{X \rightarrow Y, WY \rightarrow Z\} \models WX \rightarrow Z$

# Individuazione delle dipendenze funzionali

In genere, è il progettista della base di dati a individuare le dipendenze funzionali in base alla propria conoscenza del dominio

Dato un insieme iniziale di dipendenze funzionali  $F$ , è possibile dedurne altre in base a specifiche **regole di inferenza**

L'insieme  $F^+$  delle dipendenze funzionali individuate dal progettista unito all'insieme delle dipendenze inferite prende il nome di **chiusura** di  $F$

# Normalizzazione di relazioni

Le **forme normali** sono **proprietà** delle relazioni definite con riferimento alle dipendenze funzionali che sono soddisfatte quando non sussistono anomalie

Se una relazione non è compatibile con una forma normale, la si **decompon**e in relazioni più piccole che rispettino la forma normale desiderata

L'obiettivo è ottenere uno schema che soddisfi le seguenti proprietà:

- **Garantire join senza perdita:** se ricostruiamo una relazione dalle sue parti decomposte non dobbiamo generare ennuple non inizialmente presenti
- **Garantire la conservazione delle dipendenze:** ogni dipendenza funzionale deve essere rispettata nello schema normalizzato

# Forma normale di Boyce-Codd (BCNF)

Uno schema di relazione  $R$  è in forma normale di Boyce-Codd (BCNF) se, ogni volta che sussiste in  $R$  una dipendenza funzionale non banale  $X \rightarrow A$ ,  $X$  è una superchiave di  $R$

**Normalizzazione:** decomporre  $R$  in modo che esista una diversa relazione per ogni dipendenza. La chiave di ciascuna relazione sarà il componente di sinistra della dipendenza

La BCNF **non** è sempre raggiungibile

# Terza forma normale (3NF)

Uno schema di relazione  $R$  è in terza forma normale (3NF) se, per ogni dipendenza funzionale non banale  $X \rightarrow A$  di  $R$ , è soddisfatta almeno una delle seguenti condizioni:

- $X$  contiene una chiave di  $R$  ( $X$  è superchiave)
- $A$  appartiene ad almeno una chiave di  $R$

**Normalizzazione:** decomporre  $R$  in modo che esista una diversa relazione per ogni dipendenza. Mantenere una relazione che contenga la chiave della relazione di partenza

E' dimostrato che la 3NF è sempre raggiungibile senza perdita e conservando le dipendenze

# Seconda forma normale (2NF)

La seconda forma normale interessa le relazioni che hanno chiave composta e si basa sul concetto di **dipendenza funzionale completa**:

Una dipendenza funzionale  $X \rightarrow Y$  è completa se la rimozione di qualsiasi attributo A da X comporta che la dipendenza non sia più valida

Una dipendenza  $X \rightarrow Y$  è parziale se  $\exists A \in X: (X-A) \rightarrow Y$

# Seconda forma normale (2NF)

Uno schema di relazione  $R$  è in seconda forma normale (2NF) se ogni *attributo non primo*  $A$  di  $R$  dipende funzionalmente in modo completo dalla chiave primaria di  $R$  (anche transitivamente)

Un attributo  $A$  dello schema  $R$  è primo se e solo se fa parte di almeno una chiave di  $R$ . In caso contrario  $A$  è detto *non primo*

**Normalizzazione:** data una chiave primaria composta  $X$ , decomporre  $R$  realizzando una relazione che conservi  $X$  e, per ogni dipendenza parziale  $(X - A) \rightarrow Y$ , una distinta relazione con schema  $(X - A) \cup Y$  e chiave primaria  $X - A$

Le relazioni con chiave composta da un solo attributo sono sempre in 2FN

# Prima forma normale (1NF)

Uno schema di relazione  $R(X)$  è detto in prima forma normale (1NF o flat) se ogni attributo appartenente a  $X$  è un attributo semplice, cioè atomico

La prima forma normale esclude attributi multivалore e attributi strutturati