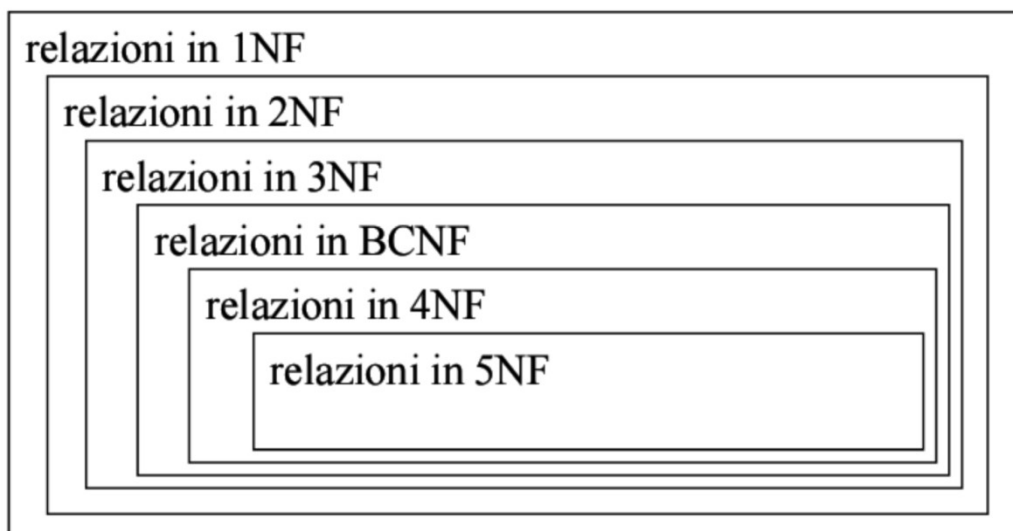


ITIS-LS “Francesco Giordani” Caserta

**prof. Ennio Ranucci
a.s. 2019-2020**

Normalizzazione



*La **normalizzazione** è un procedimento volto alla riduzione della ridondanza informativa, delle anomalie e del rischio di incoerenza dal database.*

Esistono vari livelli di normalizzazione (forme normali) che certificano la qualità dello schema del database.

Se una relazione non soddisfa una forma normale, viene eseguita una operazione di normalizzazione: una decomposizione in relazioni più piccole, una per ogni concetto.

Con **ridondanza** si intende la non necessaria ripetizione della stessa informazione con il conseguente spreco di spazio;

Con **anomalie da modifica** si intende la necessità di ripetere, in caso di cambiamento di una informazione, la modifica ovunque tale informazione è, duplicata;

Con **anomalia da inserzione** si intende la necessità di inserire ulteriori informazioni, non strettamente necessarie per poter inserire un nuovo dato;

Per **anomalie di cancellazione** si intende l'eliminazione di alcune informazioni in conseguenza della cancellazione di altre.

Ridondanze e anomalie: esempio

Impiegato	Stipendio	Progetto	Bilancio	Funzione
Rossi	20000000	Marte	2000	Tecnico
Verdi	35000000	Giove	15000	Progettista
Verdi	35000000	Venere	15000	Progettista
Neri	55000000	Venere	15000	Direttore
Neri	55000000	Giove	15000	Consulente
Neri	55000000	Marte	2000	Consulente
Mori	48000000	Marte	2000	Direttore
Mori	48000000	Venere	15000	Progettista
Bianchi	48000000	Venere	15000	Progettista
Bianchi	48000000	Giove	15000	Direttore

Lo Stipendio di ciascun Impiegato è **unico**: è **funzione** del solo Impiegato indipendentemente dai Progetti ai quali partecipa

Il Bilancio di ciascun progetto è **unico**: dipende dal solo progetto indipendentemente dai partecipanti

Ridondanze e anomalie

Impiegato	Stipendio	Progetto	Bilancio	Funzione
Neri	55000000	Venere	15000	Direttore
Neri	55000000	Giove	15000	Consulente
Neri	55000000	Marte	2000	Consulente



Ridondanza: lo Stipendio di ciascun impiegato è ripetuto per ogni partecipazione a un Progetto

Anomalia di modifica : se lo Stipendio di un Impiegato cambia occorre modificare il valore in tutte le tuple corrispondenti

Anomalia di cancellazione: se un Impiegato interrompe la partecipazione ai progetti e tutte le tuple sono eliminate, si perdono anche i suoi dati

Anomalia di inserimento: se si hanno i dati di un nuovo Impiegato non è possibile inserirli fino a che non si assegna ad un Progetto

LA PRIMA FORMA NORMALE

Ogni campo deve contenere un solo valore: (la 1NF è chiamata anche **FORMA ATOMICA**).

Un campo non atomico può essere strutturato o multivalore:

ogni attributo strutturato viene sostituito da tanti attributi quanti sono i valori che appaiono nell'attributo strutturato stesso;

ogni attributo multivalore diventa un attributo semplice spezzando la tupla che contiene il valore in tante tuple quanti sono i valori descritti nell'attributo;

LA SECONDA FORMA NORMALE

NORMALIZZAZIONE DELLE RELAZIONI

Una relazione è in **seconda forma normale (2FN)** quando è in prima forma normale e tutti i suoi attributi non-chiave dipendono dall'intera chiave, cioè non possiede attributi che dipendono soltanto da una parte della chiave.

Inventario (Prodotto, Magazzino, Qtà, IndirizzoMagazzino)

NO

IndirizzoMagazzino dipende funzionalmente da Magazzino, che rappresenta una porzione di chiave !



Essendo Magazzino → IndirizzoMagazzino

- **R1** (Magazzino, IndirizzoMagazzino)
- **R2** (Prodotto, Magazzino, Qtà)

2FN

LA TERZA FORMA NORMALE

Definizione di schema in 3NF

Uno schema $R(A)$, con A insieme di attributi, è in terza forma normale (3NF) se è in 2NF e ogni attributo non primo di A non dipende in modo transitivo da ogni chiave di A .

Sia dato uno schema di relazione $R(A)$, due dipendenze funzionali di Y da X e Z da Y , dove X, Y, Z sono sottosistemi dell'insieme di attributi di A .

Si dice che Z dipende transitivamente da X se si verificano le seguenti condizioni:

- Y dipende da X e X non dipende da Y ,
- Z dipende da Y e Y non dipende da Z ,
- Z non appartiene né ad X né ad Y .

Il seguente schema non è in 3FN:

AUTOMOBILI (TARGA, MARCA, DATA, COLORE, MODELLO, CILINDRATA)

L'attributo che non fa parte di alcuna chiave (**non primo**) **CILINDRATA**, dipende transitivamente dall'unica chiave della relazione **TARGA** in quanto:

- **MODELLO** dipende da **TARGA**
- **CILINDRATA** dipende da **MODELLO**

Bisogna eliminare la dipendenza transitiva:

- si definisce un nuovo schema **CARATTERISTICHE (MODELLO, CILINDRATA)**
- si **cancella** l'attributo **CILINDRATA** dallo schema **AUTOMOBILI**.

AUTOMOBILI (TARGA, MARCA, DATA, COLORE, MODELLO) CARATTERISTICHE (MODELLO, CILINDRATA)

Tabella STUDENTI:

(#Matricola,nome,cognome,via,cap,città,#isee,tassa)

- Con #isee si intende il codice dello scaglione di reddito che determina la tassa scolastica.

Tabella STUDENTI:

(#MATRICOLA, nome, cognome, via, cap, città, #isee)

Tabella TASSA:

(#ISEE, tassa)

Forma Normale di Boyce e Codd

Uno schema di relazione $R(X)$ è in forma normale di Boyce e Codd se:

- per ogni dipendenza funzionale (non banale)

$$Y \rightarrow Z$$

definita su $R(X)$,

Y contiene una chiave K di $R(X)$

cioè Y è una superchiave di $R(X)$.

AllievoMateriaInsegnante

Allievo	Materia	Insegnante
Carlo	Matematica	Prof. Bianchi
Carlo	Fisica	Prof. Verdi
Grazia	Matematica	Prof. Bianchi
Grazia	Fisica	Prof. Rossi

dipendenze funzionali:

- Allievo, Materia \rightarrow Insegnante
- Insegnante \rightarrow Materia

chiavi candidate:

- Allievo, Materia
- Allievo, Insegnante

insegnante
non è chiave
candidata!

- AllievoMateriaInsegnante non è in Boyce-Codd NF
- ha una ridondanza (la materia di un insegnante è ripetuta su più tuple)

Interventi

Paziente	DataIntervento	OraIntervento	Chirurgo	Sala
Bianchi	25/10/2005	8.00	De Bakey	Sala1
Rossi	25/10/2005	8.00	Romano	Sala2
Negri	26/10/2005	9.30	Veronesi	Sala1
Viola	25/10/2005	10.30	De Bakey	Sala1
Verdi	25/10/2005	11.30	Romano	Sala2

Dipendenze funzionali:

- $\{Paziente, DataIntervento\} \rightarrow OraIntervento, Chirurgo, Sala$
- $\{Chirurgo, DataIntervento, OraIntervento\} \rightarrow Paziente, Sala$
- $\{Sala, DataIntervento, OraIntervento\} \rightarrow Paziente, Chirurgo$
- $\{Chirurgo, DataIntervento\} \rightarrow Sala$

Ci sono tre insiemi di attributi che possono svolgere la funzione di chiave: $\{Paziente, DataIntervento\}$, $\{Chirurgo, DataIntervento, OraIntervento\}$, $\{Sala, DataIntervento, OraIntervento\}$. Scegliamo come chiave primaria la coppia di attributi: $\{Paziente, DataIntervento\}$.

Interventi

Paziente	DataIntervento	OraIntervento	Chirurgo
Bianchi	25/10/2005	8.00	De Bakey
Rossi	25/10/2005	8.00	Romano
Negri	26/10/2005	9.30	Veronesi
Viola	25/10/2005	10.30	De Bakey
Verdi	25/10/2005	11.30	Romano

OccupazioneSale

Chirurgo	DataIntervento	Sala
De Bakey	25/10/2005	Sala1
Romano	25/10/2005	Sala2
Veronesi	26/10/2005	Sala1