

Normalizzazione

Definizioni e dipendenze funzionali



Le forme normali

Una **forma normale** è una **proprietà** che deve essere soddisfatta dai valori degli attributi di ogni schema di relazione di una base dati; tale proprietà garantisce un certo livello di "qualità" di ciascuno schema.

Quando uno "schema" non è normalizzato

- presenta ridondanze
- crea non poche difficoltà in operazioni di aggiornamento

NORMALIZZAZIONE



Applicazione forme normali

Le forme normali sono state definite sul modello relazionale

- Non tengono quindi conto delle metodologie di progettazione

L'applicazione delle metodologie E/R produce schemi relazionali già in forma normale

- anche se nei complessi processi di progetto reali possono prodursi situazioni che richiedono una verifica a posteriori della qualità dello schema prodotto

Non devono sostituirsi alle metodologie, ma possono essere applicate nel modello E-R nella fase di analisi di qualità della progettazione concettuale

NORMALIZZAZIONE



La normalizzazione

Procedura che permette di trasformare schemi di relazione 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

Non costituisce una metodologia di progettazione

NORMALIZZAZIONE



Un esempio

5

La relazione:

Progettazioni (Impiegato, Stipendio, Progetto, Bilancio, Funzione)

In cui:

- Lo stipendio di ciascun impiegato è unico ed è legato ad esso indipendentemente dai progetti a cui partecipa
- Il bilancio di ciascun progetto è unico e dipende dal solo progetto indipendentemente dagli impiegati che vi partecipano

NORMALIZZAZIONE

Una relazione con anomalie

6

PROGETTAZIONI

Impiegato	Stipendio	Progetto	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

NORMALIZZAZIONE

La causa delle anomalie

7

Uso di una sola relazione per rappresentare informazioni eterogenee

- gli impiegati con i relativi stipendi
- i progetti con i relativi bilanci
- le partecipazioni degli impiegati ai progetti con le relative funzioni

La fusione di concetti disomogenei in una unica relazione comporta:

- Ridondanza
- Anomalie di aggiornamento
- Anomalie di cancellazione
- Anomalie di inserimento

Applichiamo queste considerazioni al caso degli impiegati

NORMALIZZAZIONE

Anomalie

8

Lo stipendio di ciascun impiegato è ripetuto in tutte le enuple relative

- Ridondanza
 - presenza di dati ripetuti in diverse tuple, senza aggiunta di informazioni significative

Se lo stipendio di un impiegato varia, è necessario andarne a modificare il valore in diverse enuple

- anomalia di aggiornamento
 - necessità di estendere l'aggiornamento di un dato a tutte le tuple in cui esso compare con inutili perdite di tempo

NORMALIZZAZIONE

Anomalie

NORMALIZZAZIONE

Se un impiegato interrompe la partecipazione a tutti i progetti, dobbiamo cancellarlo

- anomalia di cancellazione
 - L'eliminazione di una tupla motivata dal fatto che non è più valido l'insieme dei concetti in essa espressi, può comportare l'eliminazione di dati che conservano la loro validità

Un nuovo impiegato senza progetto non può essere inserito

- anomalia di inserimento
 - l'inserimento di informazioni relative a uno solo dei concetti di pertinenza di una relazione è impossibile se non esiste un intero insieme di concetti in grado di costituire una tupla completa.



La dipendenza funzionale

NORMALIZZAZIONE

Per scoprire e rimuovere le anomalie in uno schema del modello logico si devono scoprire legami di tipo funzionale tra gli attributi di una relazione dette **dipendenze funzionali (DF)** che fissano ulteriori vincoli di integrità di uno schema di relazione R(X).

La notazione per indicare una dipendenza funzionale da Y a Z, con Y e Z attributi di una relazione, è

$$Y \rightarrow Z$$

e si legge *Y determina Z* o *Z dipende funzionalmente da Y*



Definizione di dipendenza funzionale

NORMALIZZAZIONE

Consideriamo

- una relazione r su R(X)
- due sottoinsiemi non vuoti Y e Z di X

∃ in r una **dipendenza funzionale (DF)** da Y a Z:
 $Y \rightarrow Z$

se, per ogni coppia di ennuple t_1 e t_2 di r con gli stessi valori su Y, risulta che t_1 e t_2 hanno gli stessi valori anche su Z, ovvero :

$$\forall t_1, t_2 \in r : t_1[Y]=t_2[Y] \rightarrow t_1[Z]=t_2[Z]$$



Dipendenze funzionali: esempi

NORMALIZZAZIONE

PROGETTAZIONI

Impiegato	Stipendio	Progetto	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



Esempi di Dipendenze Funzionali

13

NORMALIZZAZIONE

Ogni impiegato ha un solo stipendio (anche se partecipa a più progetti)

- Lo stipendio "dipende" dall'impiegato
 - Impiegato → Stipendio**

Ogni progetto ha un bilancio

- Il bilancio "dipende" dal progetto
 - Progetto → Bilancio**

Ogni impiegato in ciascun progetto ha una sola funzione (anche se può avere funzioni diverse in progetti diversi)

- La funzione "dipende" dal progetto e dall'impiegato
 - Impiegato Progetto → Funzione**



Dipendenze funzionali banali (1)

14

NORMALIZZAZIONE

Sulla relazione PROGETTAZIONI anche:

Impiegato Progetto → Progetto

è una DF ma:

- Si tratta di una DF "banale", ossia sempre soddisfatta per definizione

Le DF devono descrivere proprietà significative e quindi

$Y \rightarrow A$ è non banale

- se A non appartiene a Y

$Y \rightarrow Z$ è non banale

- se nessun attributo in Z appartiene a Y



Dipendenze funzionali banali (2)

15

NORMALIZZAZIONE

Si osservi la DF banale:

Progetto → Progetto Bilancio

è riconducibile ad una DF non banale eliminando gli attributi del secondo membro che compaiono anche al primo

- Infatti si può dimostrare che se vale

$X \rightarrow Z$

Allora vale anche

$X \rightarrow W$

Con W sottoinsieme di Z



Vincoli e dipendenze funzionali

16

NORMALIZZAZIONE

Una dipendenza funzionale è una caratteristica dello schema, aspetto **intensionale**, e non della particolare istanza dello schema, aspetto **estensionale**

Una dipendenza funzionale è dettata dalla semantica degli attributi di una relazione e non può essere inferita da una particolare istanza dello schema

Poiché la DF è un **vincolo**, una relazione è corretta quando soddisfa la DF



DF: istanze legali

17

Una istanza di uno schema $R(X)$ che rispetti tutte le dipendenze funzionali definite sullo schema viene detta **istanza legale**.

Normalizzazione



Chiavi e attributi

18

Data una relazione $R(X)$

- un attributo che partecipa almeno ad una chiave di $R(X)$ dicesi **attributo primo**
- un attributo che non partecipa ad alcuna chiave di $R(X)$ dicesi **attributo non primo**

Normalizzazione



Vincoli di chiave e DF

19

Se K è una chiave in uno schema $R(X)$ allora ogni attributo non primo o sottoinsieme di attributi non primi di $R(X)$ dipende funzionalmente da K

Quindi il vincolo di DF

$$K \rightarrow Z$$

con $X \supseteq K \cup Z$

generalizza il vincolo di chiave $K \rightarrow X$

Normalizzazione



Esempio

20

Articolo	Magazzino	Quantità	Indirizzo
scarpe	NA1	2500	v. Leopardi 17, Napoli
scarpe	RM1	4500	v. S. Maria Maggiore 3, Napoli
pantaloni	NA1	3000	v. Leopardi 17, Napoli

DF0: vincolo di chiave (banale)

DF1: Articolo, Magazzino \rightarrow Quantità, Indirizzo

DF2: Articolo, Magazzino \rightarrow Quantità

DF3: Articolo, Magazzino \rightarrow Indirizzo

Normalizzazione



Vincoli di chiave e DF biunivoche

21

Se $Y \rightarrow Z$ e risulta anche $Z \rightarrow Y$ allora se Y è una chiave per uno schema R allora lo è anche Z .

Normalizzazione



Vincoli di chiave e DF: dimostrazione

22

Data una relazione r e una sua chiave K si può verificare che esiste una DF tra K e un qualunque sottoinsieme di attributi T dello schema di r che ovviamente non contenga K .

Dimostriamo che :

K è superchiave di $R(X)$ e $T = X - K \Leftrightarrow K \rightarrow T$

Dimostrazione

- (\Leftarrow) Se $K \rightarrow T$ allora per ogni istanza legale r si ha che $\forall t_1, t_2 \in r : t_1[K] = t_2[K] \Rightarrow t_1[T] = t_2[T]$, ovvero $t_1 = t_2$. Ciò equivale a dire che non possono esistere due tuple *distinte* con lo stesso valore di K .
- (\Rightarrow) Se K è superchiave di $R(T)$, dalla definizione di superchiave si ha che $t_1[K] = t_2[K] \Rightarrow t_1 = t_2$, e quindi $t_1[T] = t_2[T]$

Normalizzazione



Dipendenza funzionale completa

23

Una dipendenza funzionale

$Y \rightarrow Z$

è completa se Z non dipende da nessun altro sottoinsieme di Y

Se Y è una chiave, se la dipendenza funzionale è completa, Z deve dipendere da tutta la chiave e non **anche** da parte di essa.

Normalizzazione



Dipendenza funzionale non completa

24

Magazzini

Articolo	Magazzino	Quantità	Indirizzo
scarpe	NA1	2500	v. Leopardi 17, Napoli
scarpe	RM1	4500	v. S. Maria Maggiore 3, Napoli
pantaloni	NA1	3000	v. Leopardi 17, Napoli

DF1: Articolo, Magazzino \rightarrow Quantità, Indirizzo

DF2: Magazzino \rightarrow Indirizzo

Indirizzo dipende anche dalla sola parte Magazzino della chiave

Normalizzazione

