

Corso di Informatica

Modulo T2

D2.1 - La normalizzazione 1

Mauro Malatesta D2.1- La normalizzazione 1-11

1
11/05/2013

Prerequisiti

E' necessario conoscere:

- Chiave
- Chiave candidata
- Chiave primaria
- Attributi semplici
- Attributi strutturati
- Attributi multivalore

Mauro Malatesta D2.1- La normalizzazione 1-11

2
11/05/2013

Introduzione

Le tabelle prodotte nella progettazione logica vanno esaminate con cura per vedere se espongano il rischio di fenomeni di **ridondanza**, **inconsistenza** e **incongruenza**. Per garantire l'assenza di determinati difetti sulle tabelle, occorre sottoporle al processo detto di **normalizzazione**.

In questa Unità illustriamo le forme normali **1FN** e **2FN**, riservandoci di trattare le ulteriori forme normali nell'Unità successiva.

Informazioni generali

Negli esempi che seguono, le relazioni sono rappresentate mediante tabelle.

Come convenzione grafica indichiamo:

- le chiavi primarie con la **sottolineatura in grassetto**
- le chiavi esterne in **grassetto**.

La normalizzazione

Una situazione reale può essere rappresentata con un insieme di relazioni del tipo:

$$R(\underline{A}_1, A_2, \dots, A_n)$$

Allo scopo di garantire i vantaggi della gestione di un data base (flessibilità, facilità di uso e di realizzazione) ed eliminare i problemi connessi (ridondanza, incongruenza e inconsistenza) occorre sottoporre la relazione al procedimento di **normalizzazione** che, operando una **decomposizione** delle tabelle, offre la garanzia che le tabelle siano esenti dai rischi suddetti.

La normalizzazione può essere effettuata a diversi livelli crescenti di restrizione chiamate **forme normali**: abbiamo la **1FN**, **2FN**, e così via con livelli via via crescenti di restrizione.

Mauro Malatesta D2.1- La normalizzazione 1-11

5
11/05/2013

Un esempio

Esempio: STUDENTI (Matricola, Nome, Materia, Interrogazione, Classe)

Matricola	Nome	Materia	Interrogazione	Classe
00151	Rossi Mario	Matematica	12/02/2003 6	V A
00152	Verdi Luigi	Italiano	20/03/2003 5	IV B
00153	Bianchi Carla	Italiano	21/02/2003 6 15/03/2003 7	VA
00154	Bianchi Ada	Storia	25/03/2003 8	IV A

Si nota che:

- **Nome** è strutturato e deve essere decomposto in *Cognome* e *Nome*
- **Classe** è strutturato e va decomposto in *Anno* e *Sezione*
- **Interrogazione** è strutturato e deve essere decomposto in *Data* e *Voto* ed è anche multivalore (può avere più valori per una medesima istanza)

Mauro Malatesta D2.1- La normalizzazione 1-11

6
11/05/2013

Un esempio

Effettuando le modifiche accennate, si ottiene la relazione seguente:

Matricola	Cognome	Nome	Materia	Data	Voto	Anno	Sez
00151	Rossi	Mario	Matematica	12/02/2003	6	V	A
00152	Verdi	Luigi	Italiano	20/03/2003	5	IV	B
00153	Bianchi	Carla	Italiano	21/02/2003	6	V	A
00153	Bianchi	Carla	Italiano	15/03/2003	7	V	A
00154	Bianchi	Ada	Storia	25/03/2003	8	IV	A

Lo schema corretto risulta quindi:

STUDENTI (Matricola, Cognome, Nome, Materia, Data, Voto,
Anno, Sez)

Mauro Malatesta D2.1- La normalizzazione 1-11

7
11/05/2013

La prima forma normale (1FN)

Definizione: data una relazione $R(A)$ essa si dice in **prima forma normale** o **1FN** se:

- tutte le righe hanno lo stesso numero di colonne;
- i dati contenuti in una colonna sono dello stesso tipo;
- l'ordine delle righe non è rilevante;
- ogni riga è diversa dalle altre;
- gli attributi sono informazioni atomiche.

Sono praticamente i requisiti del modello relazionale

Regola pratica per normalizzare in 1FN una relazione:

- 1) sostituire ogni attributo strutturato con quelli semplici che lo compongono;
- 2) replicare le tuple contenenti attributi multivalore con tante tuple, una per ogni valore dell'attributo multivalore medesimo.

Mauro Malatesta D2.1- La normalizzazione 1-11

8
11/05/2013

La prima forma normale (1FN)

Ad esempio, la relazione

LIBRI (Autore, Titolo, Editore, Genere, Collocazione)

non è in 1FN. La sua trasformazione in 1FN è rappresentata da:

LIBRI (Cognome, Nome, Titolo, Editore, Genere, Armadio, Piano)

<u>Cognome</u>	<u>Nome</u>	<u>Titolo</u>	Editore	Genere	Armadio	Piano
Purdum	Jack	Il "C"	Jackson	Informatica	A	10
Attisani	Aldo	Sistemi	Sansoni	Sistemi	A	08
Zwirner	Giuseppe	Analisi I	CEDAM	Matematica	C	03
Zwirner	Giuseppe	Analisi II	CEDAM	Matematica	C	04

Chiave primaria composta

Mauro Malatesta D2.1- La normalizzazione 1-11

9
11/05/2013

La prima forma normale (1FN)

Si consideri ora la relazione

ORDINI (Forn, TelForn, IndForn, Articolo, Quantità,
IVA, Data, Prezzo)

Essa è in 1FN in quanto tutti i suoi attributi sono semplici.

<u>Forn</u>	Telforn	IndForn	<u>Articolo</u>	Q.tà	IVA	<u>Data</u>	Prezzo
Rossi	456456	v. Roma, 15	Cabinet	3	15	20/01/03	60
Bianchi	786788	v.Olmi, 45	Pentium	5	18	22/01/03	500
MATEL	765678	p.zza Dante	Modem	2	18	18/02/03	60
Rossi	456456	v. Roma, 15	Cabinet	1	15	20/02/03	10
Electra	894587	l.go Bari	RS232	6	18	23/02/03	12

Chiave primaria composta

Mauro Malatesta D2.1- La normalizzazione 1-11

10
11/05/2013

Anomalie di funzionamento

Gli schemi precedenti, sebbene in 1FN, non sono ancora in grado di rappresentare la realtà, poiché comportano problemi detti **anomalie di funzionamento** che sono:

- **ridondanza**: nella relazione STUDENTI i dati di alcuni studenti sono ripetuti in occasione di interrogazioni nella stessa materia;
- **anomalia da inserimento**: non è possibile inserire né un nominativo, né una classe se non c'è almeno un'interrogazione da registrare;
- **anomalia da modifica**: se, a seguito di un errore, occorre modificare il nome di uno studente, a causa della ridondanza, occorre modificare tutte le tuple relative a quel dato studente;
- **anomalia da eliminazione**: se si elimina un'interrogazione, si eliminano anche i dati dello studente associato.

Anomalie di funzionamento

Per eliminare le anomalie si deve **decomporre** la relazione iniziale in più relazioni, in modo da evitare la ridondanza e assicurandosi che ogni attributo sia rappresentato da valori atomici.

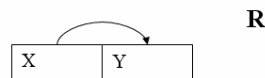
Per organizzare la decomposizione, occorre definire l'importante concetto di **dipendenza funzionale**; poi vedremo come decomporre le relazioni per ottenere la **seconda forma normale** o 2FN.

Dipendenza funzionale

Definizione: Data una relazione $R(A)$ e due sottoinsiemi degli attributi X ed Y , si dice che **Y è funzionalmente dipendente da X** , o che **X è un determinante di Y** , e si indica con

$$X \rightarrow Y$$

quando ad ogni insieme di valori di X corrisponde un unico insieme di valori di Y .



In particolare se $X \rightarrow Y$ e se $X = \{A_i\}$ e $Y = \{A_j\}$ significa che tutte le tuple che hanno lo stesso valore per A_i avranno anche lo stesso valore per A_j .
Se non sussiste la dipendenza $X \rightarrow Y$, scriviamo $X \not\rightarrow Y$

Dipendenza funzionale

Nella relazione FORNITORE (CodForn, TelForn, IndForn, Citta)

si ha, ad esempio:

$$X = \{\text{CodForn}\} \rightarrow Y = \{\text{TelForn}, \text{IndForn}, \text{Citta}\}$$

A valori diversi degli attributi di X corrispondono valori diversi degli altri attributi.

FORNITORE

Chiave primaria

<u>CodForn</u>	TelForn	IndForn	Citta
001	456456	v. Roma, 15	Roma
002	786788	v.Olmi, 45	Roma
003	765438	p.zza Dante	Milano
004	54556	C.so Napoli	Ferrara
005	45345	L.go Bari	Ragusa

Dipendenza funzionale parziale

Definizione: Data una relazione $R(A)$ e due sottoinsiemi degli attributi X ed Y per cui valga la relazione $X \rightarrow Y$, si dice che **Y dipende parzialmente da X** se esiste un sottoinsieme **proprio** Z di X tale che $Z \rightarrow Y$

In altre parole se gli attributi di Y dipendono soltanto da una parte degli attributi di X , la dipendenza di Y da X è parziale.

Mauro Malatesta D2.1- La normalizzazione 1-11

15
11/05/2013

Dipendenza funzionale parziale

In questo caso, si ha, ad esempio, $X = \{\text{Forn, Articolo, Data}\}$, perciò dato:

$Z = \{\text{Forn}\} \rightarrow Y = \{\text{TelForn, IndForn, IVA, Prezzo}\}$

si ha dipendenza parziale tra la chiave e gli attributi indicati.

<u>Forn</u>	<u>TelForn</u>	<u>IndForn</u>	<u>Articolo</u>	<u>Q.tà</u>	<u>IVA</u>	<u>Data</u>	<u>Prezzo</u>
Rossi	456456	v. Roma, 15	Cabinet	3	15	20/01/03	60
Bianchi	786788	v.Olmi, 45	Pentium	5	18	22/01/03	500
MATEL	765678	p.zza Dante	Modem	2	18	18/02/03	60
Rossi	456456	v. Roma, 15	Cabinet	1	15	20/02/03	10
Electra	894587	l.go Bari	RS232	6	18	23/02/03	12

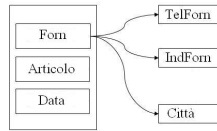
Chiave primaria

Mauro Malatesta D2.1- La normalizzazione 1-11

16
11/05/2013

Dipendenza funzionale parziale

Gli attributi di Y dipendono soltanto da una parte degli attributi di X; quindi, la dipendenza di Y da X è parziale.



<u>Forn</u>	TelForn	IndForn	<u>Articolo</u>	Q.tà	IVA	<u>Data</u>	Prezzo
Rossi	456456	v. Roma, 15	Cabinet	3	15	20/01/03	60
Bianchi	786788	v.Olmi, 45	Pentium	5	18	22/01/03	500
MATEL	765678	p.zza Dante	Modem	2	18	18/02/03	60
Rossi	456456	v. Roma, 15	Cabinet	1	15	20/02/03	10
Electra	894587	l.go Bari	RS232	6	18	23/02/03	12

Chiave primaria

Mauro Malatesta D2.1- La normalizzazione 1-11

17
11/05/2013

Dipendenza funzionale completa

Definizione: Data una relazione $R(A)$ e una dipendenza funzionale $X \rightarrow Y$ tra due sottoinsiemi degli attributi di $R(A)$ si dice che $X \rightarrow Y$ è una **dipendenza funzionale completa** se non esiste nessuna dipendenza funzionale $Z \rightarrow Y$, con Z sottoinsieme proprio di X .

In sostanza significa che gli attributi di qualunque sottoinsieme di attributi non chiave dipendono dall'intera chiave.

Ovviamente, se la chiave è formata da un solo attributo, tutte le dipendenze funzionali su qualunque sottoinsieme di attributi non chiave sono complete.

Mauro Malatesta D2.1- La normalizzazione 1-11

18
11/05/2013

Dipendenza funzionale completa

Data la relazione:

VEETURE (**Targa**, Marca, Data, Colore, Modello, Cilindrata, Prezzo)

si vede subito che $X=\{Targa\}$ è chiave principale poiché determina tutti gli altri attributi. Quindi sono assenti dipendenze parziali.

<u>Targa</u>	Marca	Data	Colore	Modello	Cilindrata	Prezzo
AG-345-GH	Fiat	10/03/03	Rosso	Punto	1200	8000
BK-112-AB	Opel	20/09/94	Blu	Tigra 1.4	1400	4000
AF-643-GI	Fiat	04/01/02	Nero	Bravo	1600	7000
BK-378-AD	Opel	20/07/92	Nero	Tigra 1.6	1600	2500
BC-198-AF	Alfa	16/05/02	Grigio	165	1800	12000
KJ-492-AK	Renault	08/08/00	Grigio	Clio	1200	7000

Mauro Malatesta D2.1- La normalizzazione 1-11

19
11/05/2013

Esempi di dipendenza

Negli esempi precedenti si ha che:

VEETURE (**Targa**, Marca, Data, Colore, Modello, Cilindrata, Prezzo)
presenta dipendenza funzionale completa rispetto alla chiave *Targa*.

ORDINI (**Forn**, TelForn, IndForn, **Articolo**, Q.tà, IVA, **Data**, Prezzo)

presenta le dipendenze funzionali parziali (e quindi non complete) seguenti:

$\{Forn,Articolo,Data\} \rightarrow \{IVA, Prezzo\}$

$\{Forn,Articolo,Data\} \rightarrow \{TelForn,IndForn\}$

Mauro Malatesta D2.1- La normalizzazione 1-11

20
11/05/2013

La seconda forma normale (2FN)

Definizione: Una relazione $R(A)$ si dice in **seconda forma normale**, e si indica con **2FN**, quando è in 1FN e tutte le dipendenze funzionali dalla chiave di ogni sottoinsieme di attributi non chiave sono complete.

Regola pratica per normalizzare in 2FN

- 1) Verificare la 1FN
- 2) Individuare tutti gli attributi Y che dipendono parzialmente dal sottoinsieme di attributi X' della chiave X .
- 3) Costruire una nuova relazione avente X' come chiave primaria e Y come attributo
- 4) Togliere Y dalla relazione iniziale

Mauro Malatesta D2.1- La normalizzazione 1-11

21
11/05/2013

Esempi di 2FN

Pertanto:

VETTURE (Targa, Marca, Data, Colore, Modello, Cilindrata, Prezzo)

è in 2FN poiché *Targa* è l'unica chiave possibile.

ORDINI (Forn, TelForn, IndForn, **Articolo**, Q.tà, IVA, **Data**, Prezzo)

non è in 2FN poiché ci sono dipendenze parziali:

{Forn, Articolo, Data} \rightarrow {TelForn, IndForn}

{Forn, Articolo, Data} \rightarrow {IVA, Prezzo}

Per trasformare ORDINI(...) in 2FN occorre operare una **decomposizione**.

Mauro Malatesta D2.1- La normalizzazione 1-11

22
11/05/2013

Esempi di 2FN

<u>Forn</u>	Telforn	IndForn
Rossi	456456	v. Roma, 15
Bianchi	786788	v.Olmi, 45
MATEL	765678	p.zza Dante
Electra	894587	l.go Bari

<u>Articolo</u>	Prezzo	IVA
Cabinet	15	15
Pentium	100	18
Modem	25	18
RS232	3	18

<u>Forn</u>	<u>Articolo</u>	<u>Data</u>	Q.tà
Rossi	Cabinet	20/01/03	3
Bianchi	Pentium	22/01/03	5
MATEL	Modem	18/02/03	2
Rossi	Cabinet	20/02/03	1
Electra	RS232	23/02/03	6

La **decomposizione** genera le tre relazioni:

FORNITORE (**Forn**, TelForn, IndForn)

ARTICOLI (**Articolo**, Prezzo, IVA)

ORDINI (**Forn,Articolo,Data**, Q.tà)

Abbiamo eliminato le dipendenze parziali dalla chiave.

Mauro Malatesta D2.1- La normalizzazione 1-11

23
11/05/2013

Argomenti

- La normalizzazione
- Un esempio
- La prima forma normale (1FN)
- Anomalie di funzionamento
- Dipendenza funzionale
- Dipendenza funzionale parziale
- Dipendenza funzionale completa
- Esempi di dipendenza
- La seconda forma normale (2FN)
- Esempi di 2NF

Mauro Malatesta D2.1- La normalizzazione 1-11

24
11/05/2013

Altre fonti di informazione

- A.Lorenzi-D.Rossi, Le basi di dati e il linguaggio SQL – ed. ATLAS
- F.Cesarini-F.Pippolini-G.Soda, Informatica 3- ed. Cremonese
- A.Garavaglia-N.Mazzucchelli-F.Petracchi, Informatica-ed. Zanichelli