

NORMALIZZAZIONE DELLE BASI DI DATI E LE TRE FORME NORMALI

La normalizzazione è una tecnica di progettazione dei database, mediante la quale si elimina la ridondanza dei dati al fine di evitare anomalie nella loro consistenza in seguito a operazioni di inserimento, cancellazione o modifica; la normalizzazione è infatti 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 ha i seguenti difetti:

- Presenta ridondanze,
- Si presta a comportamenti poco desiderabili durante gli aggiornamenti

Le forme normali sono di solito definite sul modello relazionale, ma hanno senso in altri contesti, ad esempio il modello E-R

La normalizzazione va utilizzata come tecnica di verifica dei risultati della progettazione di una base di dati

La normalizzazione non costituisce una metodologia di progettazione

La normalizzazione viene eseguita in varie fasi. Al termine di ciascuna fase il database si trova in uno degli stati di normalizzazione, o come si dice di solito, è in una delle “forme normali”.

La normalizzazione dei dati di un database, infatti, si suddivide in tre forme importanti: prima, seconda e terza forma normale.

- **Prima forma normale:**

Una tabella è in prima forma normale se non contiene colonne ripetute, che vengono usate per descrivere proprietà dell'entità della tabella che possono comparire in un numero variabile. L'esempio classico è quello di una tabella Persona, che contiene le colonne Foglio1, Foglio2, Foglio3, le quali descrivono i figli della persona. È evidente, che, nel caso in cui la persona abbia meno di tre figli, alcune colonne non saranno utilizzate. Inoltre sarà impossibile rappresentare una persona con più di tre figli. Per portare tale tabella in forma normale occorre scomporla in due tabelle: una per la persona, senza le colonne dei figli, l'altra per i figli, contenente una chiave esterna verso la tabella delle persone, indicante la riga che rappresenta il genitore.

Condizione: in una tabella gli elementi delle colonne devono essere ad un solo valore, non possono essere array.

Normalizzazione: nel caso di presenza di un array, normalizzare la tabella vuol dire produrre tante righe quanti sono gli elementi del vettore.

N. protocollo	Uffici interessati				

NO

N. protocollo	Uffici interessati

SI

▪ **Seconda forma normale:**

Una tabella è in seconda forma normale se è in prima forma normale e tutti i suoi attributi sono funzionalmente dipendenti dall'intera chiave primaria.

Per portare una tabella ad essere in seconda forma normale bisogna eliminare tutti gli attributi che non dipendono dalla chiave primaria.

Condizione: in una tabella con chiave a più attributi, ogni colonna non appartenente alla chiave deve dipendere dall'insieme delle colonne-chiave e non solo da una parte di esse.

Normalizzazione: Normalizzare vuol dire produrre tante tabelle che soddisfino la condizione nel caso in cui in quella originaria essa non sia soddisfatta.

<u>Codice città</u>	<u>Codice via</u>	Città	Via
01	01	Roma	Verdi
01	02	Roma	Bianchi
01	03	Roma	Rossi
02	01	Milano	Gialli

NO

Città dipende soltanto da 'Codice città'

<u>Codice città</u>	<u>Codice via</u>	Via
01	01	Verdi
01	02	Bianchi
01	03	Rossi
02	01	Gialli

SI

<u>Codice città</u>	Città
01	Roma
01	Roma
01	Roma
02	Milano

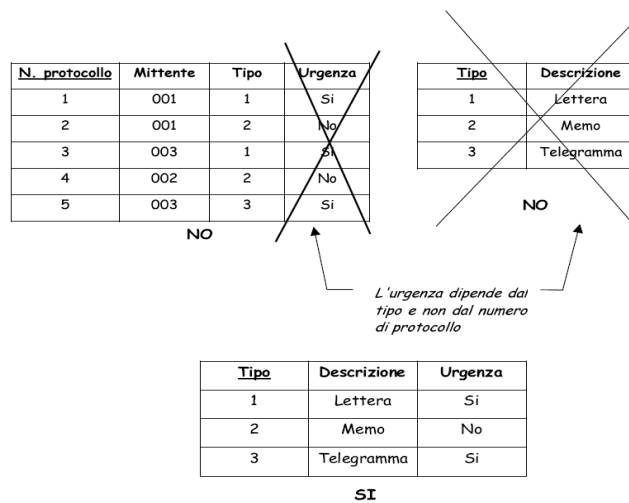
SI

▪ Terza forma normale

Una tabella è in terza forma normale se è in seconda forma normale e nessuno dei suoi attributi è transitivamente dipendente dalla chiave primaria. Avere una dipendenza transitiva significa che una colonna dipende da un'altra colonna descrittore, la quale a sua volta dipende dalla chiave primaria. La terza forma normale si ottiene eliminando dalla tabella tutti gli attributi che dipendono dal descrittore.

Condizione: in una tabella la dipendenza fra le colonne deve essere basata soltanto sulla chiave primaria.

Normalizzazione: Normalizzare vuol dire produrre tante tabelle che soddisfino la condizione nel caso in cui in quella originaria essa non sia soddisfatta.



Questo tipo di relazione è in forma normale di Boyce e Codd (BCNF) se tutti i suoi determinanti sono candidati chiave.

Ogni relazione in 3NF è anche in BCNF, a meno che:

- abbia 2 o più chiavi candidate,
- le chiavi candidate siano composte,
- abbiano almeno un attributo in comune.

▪ Forma normale di Boyce e Codd (BCNF)

Una relazione r è in forma normale di Boyce e Codd se, per ogni dipendenza funzionale (non banale) $X \rightarrow Y$ definita su di essa, X contiene una chiave K di r .

La forma normale richiede che i concetti in una relazione siano omogenei (solo proprietà direttamente associate alla chiave).