

Le Forme Normali

Nella strutturazione degli archivi, per eliminare del tutto la ridondanza, si utilizza un processo di normalizzazione dei Data Base; è questo un processo graduale che permette di portare un relazione non normalizzata fino ad un certo livello di normalizzazione, più o meno alto.

Prima Forma Normale

Una relazione è in Prima Forma Normale (1NF) se e solo se tutti i suoi attributi sono valori atomici.

Nel Modello Relazionale tutte le relazioni, per essere tali, devono rispettare la 1NF (1 Normal Form); questa impone che gli attributi non devono poter essere suddivisi ulteriormente in valori più piccoli.

Un classico esempio è un archivio dei clienti di un'azienda in cui vengono riportati, oltre al nome e all'indirizzo, anche città di residenza e numero telefonico: quest'ultimo è formato da prefisso e numero, che sono due cose separate. In particolare, se viene riportata la città non sarà necessario riportarne anche il prefisso telefonico in quanto l'una implica l'altro; è quindi meglio, da un punto di vista logico-strutturale, creare una relazione nella quale si faccia corrispondere al nome di una particolare città il rispettivo prefisso telefonico o viceversa.

Si noti che la 1NF permette solo di definire una relazione, ma non ne elimina la ridondanza.

Seconda Forma Normale

Per comprendere la 2NF è necessario spiegare prima cosa s'intende per dipendenza funzionale; dato un attributo X ed una chiave composta dagli attributi ABC...Z, possiamo dire che:

Si ha dipendenza funzionale tra un attributo X ed una chiave se, e solo se, ad un particolare valore di X corrisponde una particolare successione di valori degli attributi ABC...Z che formano la chiave, quindi se esiste una relazione biunivoca che unisce X alla chiave.

Immaginiamo una chiave composta dagli attributi AB; avremo dipendenza funzionale quando ai valori a1,b1 corrisponde x1, ad a1,b2 corrisponde x2 e così via. Se un attributo Y della relazione è dipendenza funzionale di un sottoinsieme degli attributi che formano la chiave, ad esempio se Y dipende soltanto da A (dipendenza funzionale parziale), allora la 2NF non è rispettata:

Una relazione è in Seconda Forma Normale (2NF) se e solo se è in 1NF e tutti i suoi attributi che non fanno parte della chiave dipendono funzionalmente da una qualunque chiave, sia essa primaria o candidata.

Per normalizzare la relazione è quindi necessario individuare gli attributi che dipendono funzionalmente da un sottoinsieme degli attributi della chiave e creare con essi una nuova relazione.

Da questa definizione si capisce come il processo di normalizzazione avvenga per gradi: le forme normali superiori, infatti, impongono che siano rispettate anche quelle di livello precedente.

Terza Forma Normale

La 2NF elimina le dipendenze funzionali parziali dirette fra gli attributi, tuttavia non risolve completamente il problema della ridondanza. Restano infatti le dipendenze funzionali transitive, che si hanno quando un attributo dipende indirettamente da un altro; ad esempio:

$$A \rightarrow B \rightarrow C$$

In cui B dipende da A, che è chiave primaria o candidata e, C dipende da B, che non è chiave: quindi C dipende transitivamente dalla chiave A. La 3NF si occupa proprio di eliminare tali dipendenze indirette:

Una relazione è in Terza Forma Normale (3NF) se e solo se è in 2NF e tutti i suoi attributi che non fanno parte della chiave non dipendono transitivamente da una qualunque chiave, sia essa primaria o candidata .

Come per il caso precedente, la normalizzazione della relazione avverrà individuando le dipendenze indirette e creando con queste nuove relazioni, che risulteranno normalizzate. Tuttavia quando si ha a che fare con molte chiavi candidate e molte dipendenze transitive, risulta complesso effettuare la scomposizione. Per superare questo problema, Boyce e Codd hanno dato una definizione più approfondita di Terza Forma Normale.

Forma Normale di Boyce e Codd

Definendo come determinante un attributo o gruppo di attributi che individuano in modo univoco un altro sottoinsieme di attributi della relazione, creando così un rapporto biunivoco, possiamo dire che:

Una relazione è nella Forma Normale di Boyce e Codd (BCNF) se, e solo se, ciascun determinante è una chiave candidata.

Essendo chiave candidata, significa che il determinante ha tutti i requisiti per essere chiave primaria, quindi non dipende da nessun altro attributo che non sia chiave primaria o candidata: vengono così eliminate tutte le dipendenze funzionali, sia quelle parziali sia quelle transitive.

È da notare, in fine, che la BCNF non impone che la relazione soddisfi nessun'altra forma normale di grado inferiore, ma è da sola sufficiente ad eliminare la ridondanza. Tuttavia per arrivare ad un simile livello di normalizzazione bisognerà, dal punto di vista pratico, passare comunque attraverso i livelli descritti in precedenza.

Conclusioni

I livelli di normalizzazione permettono di eliminare del tutto la ridondanza da una relazione, componendola in una serie di relazioni di più piccole dimensioni e, di conseguenza, più facili di gestire: si noti quindi l'attinenza tra il processo di normalizzazione ed i Modelli ER e Relazionale.