

[Appunti Basi di Dati]

Appunti del corso "Basi di Dati" dell'Università degli Studi di Genova

da

Lorenzo Vaccarecci



**UNIVERSITÀ DEGLI STUDI
DI GENOVA**

**Dipartimento di Informatica
Università degli Studi di Genova
2024**

Indice

1	Modello Relazionale	2
1.1	Introduzione	2
1.2	Relazioni	2
1.2.1	Definizione: Prodotto cartesiano	2
1.2.2	Definizione: Relazione	2
1.2.3	Definizione: Schema di relazione	2
1.2.4	Definizione: Schema di base di dati	3
1.2.5	Definizione: Tuple e relazione	3
1.3	Valori nulli	3
1.4	Chiavi	3
1.4.1	Definizione: Chiave e super-chiave	3
1.4.2	Definizione: Chiave esterna	4

CAPITOLO 1

Modello Relazionale

1.1 Introduzione

Le interrogazioni sulle relazioni possono essere espresse in due formalismi:

- **Algebra relazionale:** le interrogazioni vengono espresse usando operatori specifici alle relazioni.
- **Calcolo relazionale:** le interrogazioni vengono espresse usando formule logiche.

1.2 Relazioni

Un dominio è un insieme (anche infinito) di valori. Indicheremo con \mathcal{D} l'insieme di tutti i domini.

1.2.1 Definizione: Prodotto cartesiano

Siano $D_1, D_2, \dots, D_k \in \mathcal{D}$ con k domini. Il prodotto cartesiano indicato con $D_1 \times D_2 \times \dots \times D_k$, è definito come l'insieme $\{(v_1, v_2, \dots, v_k) \mid v_1 \in D_1, \dots, v_k \in D_k\}$.

Gli elementi appartenenti al prodotto cartesiano sono detti **tuple**. Il prodotto cartesiano ha **grado** k .

1.2.2 Definizione: Relazione

Una relazione di k domini è un sottoinsieme finito del prodotto cartesiano, ha **grado** k quindi ogni tupla ha k componenti. La **cardinalità** di una relazione è il numero di tuple appartenenti alla relazione. Una relazione è sempre un insieme finito. **Non vi sono tuple duplicate**.

La coppia (nome di attributo, dominio) è detta **attributo**. L'uso di attributi permette di denotare le componenti di ogni tupla per nome piuttosto che per posizione.

1.2.3 Definizione: Schema di relazione

- R un nome di relazione
- $\{A_1, A_2, \dots, A_n\}$ un insieme di nomi di attributi
- $dom : \{A_1, A_2, \dots, A_n\} \rightarrow \mathcal{D}$ una funzione totale che associa ad ogni nome di attributo il corrispondente dominio.

La coppia $(R(A_1, A_2, \dots, A_n), dom)$ è uno schema di relazione. U_r denota l'insieme dei nomi di attributi di R , ovvero $U_r = \{A_1, A_2, \dots, A_n\}$.

1.2.4 Definizione: Schema di base di dati

Siano S_1, S_2, \dots, S_n schemi di relazioni distinti, $\mathcal{S} = \{S_1, S_2, \dots, S_n\}$ è detto schema di base di dati.

1.2.5 Definizione: Tuple e relazione

Una tupla t definita su una relazione R è un insieme di funzioni totali che associano all'attributo di nome A_i un valore del dominio di tale attributo. Una relazione definita su uno schema di relazione è un insieme finito di tuple definite su tale schema. Tale relazione è anche detta istanza dello schema. $t = [A_1 : v_1, A_2 : v_2, \dots, A_n : v_n]$ dove $v_i \in \text{dom}(A_i)$ con $i = 1, \dots, n$. Notazione: $t[A_i]$ indica il valore dell'attributo A_i (quindi v_i) nella tupla t .

1.3 Valori nulli

Un aspetto importante nella modellazione dei dati riguarda il fatto che non sempre sono disponibili tutte le informazioni sulle entità del dominio applicativo che vengono rappresentate nella base di dati. L'approccio adottato è quello di introdurre un valore speciale, detto **valore nullo**, il quale denota la mancanza di un valore.

Nella trattazione assumiamo di denotare il valore nullo con il simbolo '?'. Il valore nullo è un valore accettato in tutti i domini.

1.4 Chiavi

Una **chiave** di una relazione è un insieme di attributi che distingue fra loro le tuple della relazione. Più formalmente:

1.4.1 Definizione: Chiave e super-chiave

Sia R uno schema di relazione. Un insieme $X \subseteq U_R$ di attributi di R è chiave di R se verifica le seguenti proprietà:

1. **Univocità**: nella relazione non ci possono essere due tuple distinte che abbiano lo stesso valore per tutti gli attributi della chiave X .
2. Nessun **sottoinsieme proprio** di X verifica la proprietà (1).

Un insieme di attributi che verifica la proprietà (1) ma non la (2) è detto **super-chiave** di R . Una super-chiave può essere una chiave della relazione.

In una relazione ci possono essere più insiemi di attributi che soddisfano le due proprietà. In tal caso si parla di **chiavi candidate**. **Una relazione ha sicuramente almeno una chiave (sia primaria che super)**. Nel caso in cui ci sono più chiavi candidate, una di queste viene scelta come **chiave primaria** su cui il DBMS ottimizza le operazioni.

Un criterio per scegliere la chiave primaria è quello di scegliere la chiave più piccola in termini di numero di attributi o quella più usata nelle interrogazioni. Una chiave non può contenere valori nulli.

1.4.2 Definizione: Chiave esterna

Sia R_1 ed R_2 due relazioni, sia X una chiave per R_1 e Y una chiave per R_2 tale che Y e X contengano lo stesso numero di attributi e di dominio compatibile (*es. interi e reali sono compatibili*). X è una chiave esterna di R_1 su R_2 se per ogni tupla t di R_1 esiste una tupla t' di R_2 tale che $t[X] = t'[Y]$. R_1 viene detta relazione **referente** e R_2 relazione **riferita**.

Vincolo di integrità referenziale: se una tupla t di R_1 fa riferimento ai valori della chiave di una tupla t' di R_2 , allora t' deve esistere in R_2 . Può essere violata da inserimenti e modifiche nella relazione referente e da cancellazioni e modifiche nella relazione riferita. Una relazione può contenere più chiavi esterne, eventualmente anche sulla stessa relazione e possono assumere valori nulli.

Notazione:

$$R_1 (\dots, chiave_esterna^{R_2}, \dots)$$