

[Appunti Basi di Dati]

Appunti del corso "Basi di Dati" dell'Università degli Studi di Genova
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
2	Modello ER	5
2.1	Introduzione	5
2.2	Vincoli di integrità	6
2.2.1	Vincoli di cardinalità	6
2.2.2	Vincoli di identificazione	6
2.3	Gerarchie di generalizzazione	7

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:

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

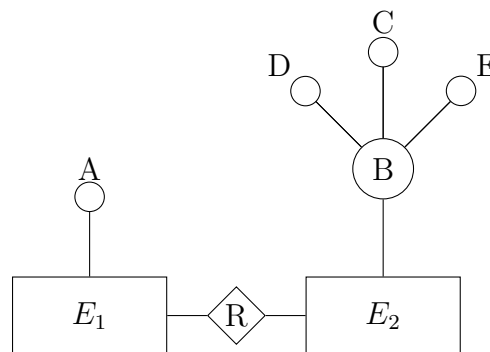
CAPITOLO 2

Modello ER

2.1 Introduzione

Il termine **diagramma ER** indica la rappresentazione grafica di uno schema concettuale ER. Gli elementi principali di un diagramma ER sono:

- **Entità**: una collezione di oggetti della realtà che vogliamo modellare che possiedono caratteristiche comuni, graficamente viene rappresentata da un rettangolo
- **Relazioni** (o associazione): rappresenta un legame logico tra più entità, graficamente viene rappresentata da un rombo
- **Attributi**: rappresenta una proprietà posseduta da un'entità o da una relazione, è **mono-valore** se può assumere al più un valore (*es. la matricola di uno studente*), è **multi-valore** altrimenti (*es. i numeri di telefono di un'azienda*). Può essere a sua volta essere formato da *sotto-attributi* rendendolo **composto**, un attributo composto può avere domini di diverso tipo (**domini composti**)



Le istanze di un'associazione sono un sottoinsieme del prodotto cartesiano delle entità coinvolte, non possono esistere duplicati. Le associazioni sono classificate in base al loro **grado** (il numero di entità in relazione):

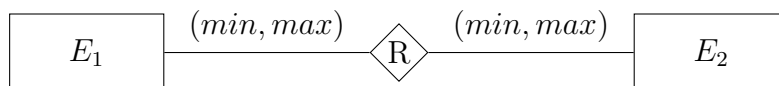
- **Unaria**: un'entità è in relazione con se stessa
- **Binaria**: due entità sono in relazione
- **n -aria**: n entità sono in relazione

Le informazioni sui domini degli attributi non sono direttamente rappresentabili in un diagramma ER, ma possono essere specificate nella documentazione a corredo dei diagrammi ER.

2.2 Vincoli di integrità

2.2.1 Vincoli di cardinalità

I vincoli di cardinalità per associazioni stabiliscono il numero minimo e massimo di istanze di un'associazione a cui le istanze delle entità coinvolte nella associazione possono partecipare. Graficamente, vengono rappresentati come coppia di valori interi (min, max) collocata vicino alla linea che connette l'associazione con ciascuna entità che mette in relazione.



(min, max):

- $min = 0$: l'entità non è obbligata a partecipare all'associazione
- $min = 1$: l'entità è obbligata a partecipare all'associazione
- $max = 1$: l'entità può partecipare al massimo una volta all'associazione
- $max = n$: l'entità può partecipare al massimo n volte all'associazione

Se in un diagramma ER non è specificata la cardinalità di un'associazione, si assume che il vincolo sia $(0, n)$.

Date due entità E_1 e E_2 se:

- $max_{E_1} = max_{E_2} = 1$ allora l'associazione è **uno a uno**
- $max_{E_1} = 1, max_{E_2} = n$ allora l'associazione è **uno a molti**
- $max_{E_1} = max_{E_2} = n$ allora l'associazione è **molti a molti**

Le cardinalità possono essere specificate anche per gli attributi, permettendo di specificare se un attributo è mono o multi-valore, e se è obbligatorio o meno. Se non viene specificata la cardinalità di un attributo, si assume che sia $(1, 1)$.

2.2.2 Vincoli di identificazione

Definire un vincolo d'identificazione per un'entità significa specificare un insieme di attributi e/o entità che posseggono la proprietà di identificare univocamente le istanze dell'entità (**identificatori o chaivi**).

- **Identificatore interno**: è costituito da uno o più attributi dell'entità stessa
- **Identificatore misto**: è costituito da attributi dell'entità e da attributi di altre entità
- **Identificatore esterno**: è costituito da attributi di altre entità
- **Identificatore semplice**: è costituito da un solo attributo
- **Identificatore composto**: è costituito da più attributi

Un'entità le cui istanze vengono identificate mediante l'associazione con altre entità viene chiamata **entità debole**.

2.3 Gerarchie di generalizzazione

Nel modello ER è possibile organizzare le entità in gerarchia di generalizzazione definendo un insieme di entità dette **figlie** come specializzazione di un'altra entità detta **padre** rappresentante le proprietà in comune a tutte le entità figlie.

- **Generalizzazione totale:** ogni istanza dell'entità padre è anche un'istanza di **almeno** una delle entità figlie
- **Generalizzazione parziale:** se un'istanza dell'entità padre non è un'istanza di nessuna delle entità figlie
- **Generalizzazione esclusiva:** se un'istanza dell'entità padre è un'istanza di **al più** una delle entità figlie
- **Generalizzazione condivise:** se un'istanza dell'entità padre può essere un'istanza di **più** entità figlie

Le generalizzazioni possono essere di quattro tipi diversi: totali esclusive, totali condivise, parziali esclusive e parziali condivise.