

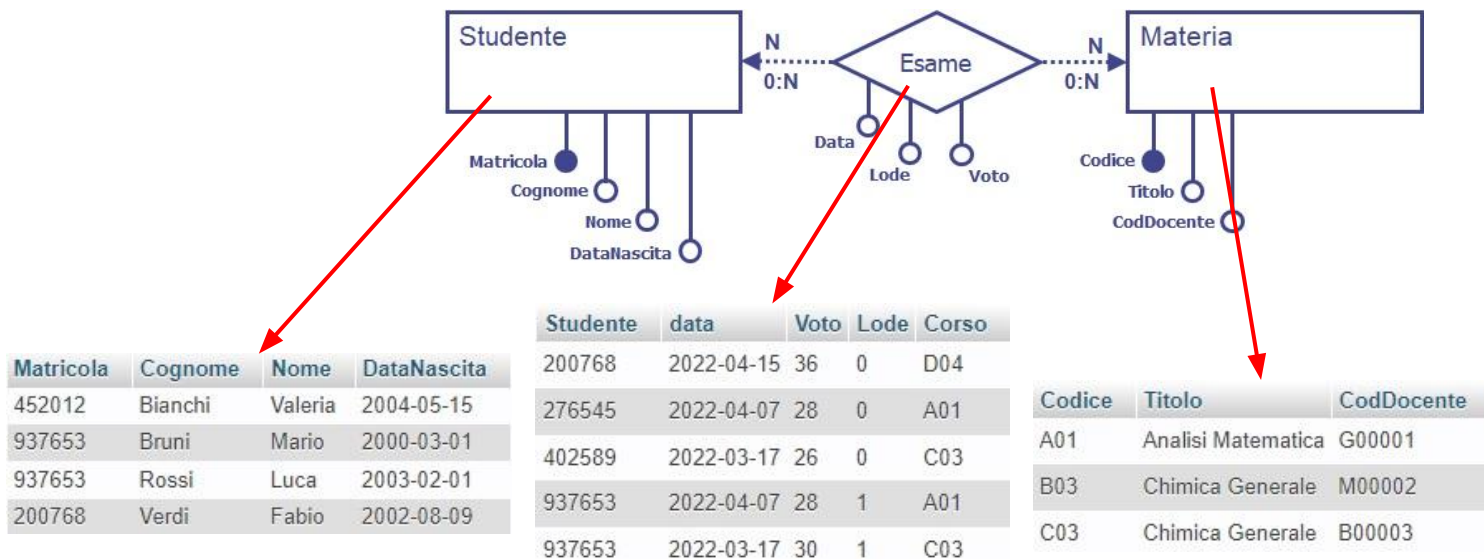
14. Il modello dei dati relazionale

vincoli di integrità

database

il modello dei dati relazionale - vincoli di integrità

Consideriamo l'esempio seguente (del quale riportiamo anche il diagramma E/R) per la registrazione degli esami sostenuti dagli studenti iscritti in un corso di laurea:



il modello dei dati relazionale - vincoli di integrità

In generale non sempre è vero che qualunque insieme di tuple coerenti con lo schema di relazione, rappresenti informazioni corrette:

Ci sono due studenti con lo stesso numero di matricola

Sono stati inseriti un corso (D04) ed una matricola (276545) inesistenti.

Nella tabella degli esami compare un voto pari a 36.

Sempre nella tabella degli esami è stata inserita la lode ad un voto pari a 28.

Matricola	Cognome	Nome	DataNascita
452012	Bianchi	Valeria	2004-05-15
937653	Bruni	Mario	2000-03-01
937653	Rossi	Luca	2003-02-01
200768	Verdi	Fabio	2002-08-09

Studente	data	voto	Lode	Corso
200768	2022-04-15	36	0	D04
276545	2022-04-07	28	0	A01
402589	2022-03-17	26	0	C03
937653	2022-04-07	28	1	A01
937653	2022-03-17	30	1	C03

Codice	Titolo	CodDocente
A01	Analisi Matematica	G00001
B03	Chimica Generale	M00002
C03	Chimica Generale	B00003

il modello dei dati relazionale - vincoli di integrità

In un database è opportuno evitare situazioni come quelle appena descritte; allo scopo è stato introdotto il concetto di **vincolo di integrità**, come **proprietà che deve essere soddisfatta dalle istanze di relazione che rappresentano informazioni corrette**.

Definendo uno schema di relazione, associamo ad esso i vincoli di integrità di interesse corrispondenti a proprietà del mondo reale oggetto del nostro sistema informativo.

```
CREATE TABLE Persone (  
    ID INT NOT NULL AUTO_INCREMENT,  
    Cognome VARCHAR(24) NOT NULL,  
    Nome VARCHAR(24),  
    Età INT,  
    Dipart CHAR(5) NOT NULL,  
    CONSTRAINT PK PRIMARY KEY (ID),  
    CONSTRAINT FK FOREIGN KEY (Dipart) REFERENCES Dipartimenti(CodDip),  
    CONSTRAINT CHK1 CHECK (Età >= 18)  
);
```

NB: I vincoli di integrità sono definiti al livello di schema di relazione con riferimento a tutte le istanze di relazione: consideriamo corrette solo quelle istanze che soddisfano tutti i vincoli.

il modello dei dati relazionale - vincoli di integrità

È possibile classificare i vincoli a seconda degli elementi di un database che ne sono coinvolti; in questo modo, distinguiamo due categorie di vincoli:

- vincolo *intrarelazionale*, se è definito su singole tabelle e può essere valutato su ciascuna riga di una tabella indipendentemente dalle altre, come nel caso dei voti registrati nella tabella degli esami.
- vincolo *interrelazionale*, se coinvolge invece più tabelle, come nel caso della registrazione di un esame ad uno studente non registrato nella tabella degli studenti, o di un corso non registrato nella tabella dei corsi.

il modello dei dati relazionale - vincoli di integrità

I **vincoli di tupla** sono vincoli intrarelazionali ed esprimono condizioni sui valori assegnati ai campi di ciascuna riga di una tabella indipendentemente dalle altre.

Ad es. possiamo impostare un vincolo nella definizione della tabella *Esame* per evitare l'inserimento di voti non compresi tra 18 e 30 o di mettere la lode se il voto è diverso da trenta, mediante il soddisfacimento della seguente condizione booleana composta:

```
CHECK (Voto=30 AND Lode=1) OR (Voto>=18 AND Voto<=30 AND Lode=0)
```

oppure equivalentemente:

```
CHECK (Voto>=18 AND Voto <=30) AND (NOT(Lode=1) OR (Voto=30))
```

NB: Il campo *Lode* è di tipo booleano; il valore `true` corrisponde a 1 (in realtà a qualunque valore diverso da 0), mentre `false` corrisponde a 0.

il modello dei dati relazionale - vincoli di integrità

I **vincoli di tupla** sono vincoli intrarelazionali ed esprimono condizioni sui valori assegnati ai campi di ciascuna riga di una tabella indipendentemente dalle altre.

La definizione che abbiamo dato ammette anche espressioni più complesse, purché definite sui valori delle singole righe; ad es. se abbiamo lo schema di relazione seguente relativo alla registrazione dei pagamenti:

Pagamenti(Data, Importo, Ritenute, Netto)

è possibile definire un vincolo che impone, come naturale, che il netto sia pari alla differenza tra l'importo originario e le ritenute, nel modo seguente:

CHECK (Netto = Importo - Ritenute)

il modello dei dati relazionale - vincoli di integrità

Per la definizione insiemistica di relazione che abbiamo dato, ogni istanza di relazione derivante da uno schema di relazione, è un insieme di tuple (sottoinsieme del prodotto cartesiano dei corrispondenti domini).

Quanto appena ribadito, introduce un cosiddetto **vincolo di chiave** per le tuple di una relazione, il quale stabilisce, coerentemente con la definizione insiemistica di relazione, che le tuple di una stessa istanza di relazione, devono per definizione essere tutte distinte tra loro.

Come dire che per ciascuna istanza di relazione esiste sempre una chiave primaria, al limite definita da tutti i campi dello schema di relazione di appartenenza (*univocità*).

NB: Vedremo che in realtà questo vincolo viene superato dalla maggior parte dei DBMS relazionali oggi in commercio.

il modello dei dati relazionale - vincoli di integrità

Per rispettare il vincolo di chiave può essere inserito un campo di tipo intero che viene incrementato automaticamente dal DBMS ed impostato come chiave primaria.

```
CREATE TABLE Persone (  
    ID int NOT NULL AUTO_INCREMENT,  
    LastName varchar(255) NOT NULL,  
    FirstName varchar(255),  
    Age int,  
    PRIMARY KEY (ID)  
);
```

Autoincremento e chiave primaria

Nome	Tipo	Lunghezza/Valori	Predefinito	Codifica caratteri	Attributi	Null	Indice	A...
ID	INT		Nessuno			<input type="checkbox"/>	PRIMARY	<input checked="" type="checkbox"/>
Cognome	VARCHAR	25	Nessuno			<input type="checkbox"/>	---	<input type="checkbox"/>
Nome	VARCHAR	25	Nessuno			<input type="checkbox"/>	---	<input type="checkbox"/>
Età	INT		Nessuno			<input type="checkbox"/>	---	<input type="checkbox"/>

Scegli dai Campi Centrali

il modello dei dati relazionale - vincoli di integrità

Può capitare che i valori di alcuni campi all'interno di una specifica tupla possano essere sconosciuti o non pertinenti per quella tupla; in questi casi si utilizza un valore speciale `null`: si tratta di un valore aggiuntivo rispetto a quelli del corrispondente dominio, e ben distinto da essi.

La data di morte non ha senso per quei musicisti contemporanei che sono ancora in vita.

Nominativo	DataNascita	DataMorte	Nazionalità
Antonio Vivaldi	1678-03-04	1741-07-28	Italia
Arvo Part	1935-09-11	NULL	Estonia
Claude Debussy	1862-08-22	1918-03-25	Francia
Johann Sebastian Bach	1685-03-31	1750-07-28	Germania
Stefano Bollani	1972-12-05	NULL	Italia
Wolfgang Amadeus Mozart	1756-01-27	1791-12-05	Austria

Il modello di dati relazionale introduce un ulteriore vincolo intrarelazionale, definito **vincolo di integrità sull'entità**, in base al quale si stabilisce che nessun campo facente parte della chiave primaria possa assumere valore `null`.

il modello dei dati relazionale - vincoli di integrità

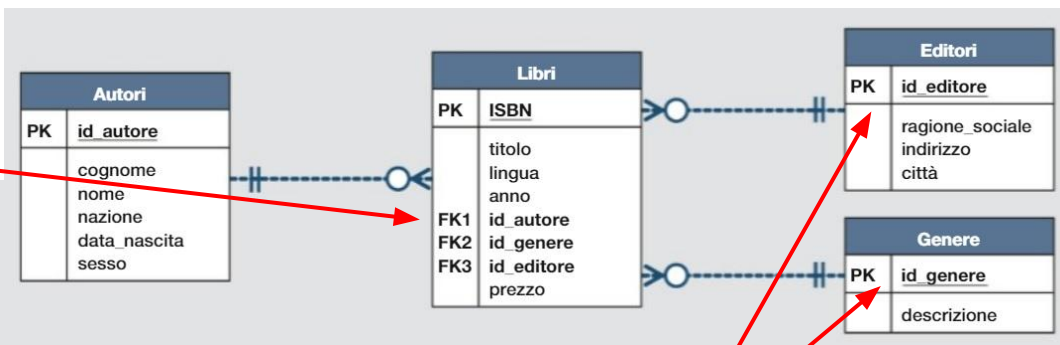
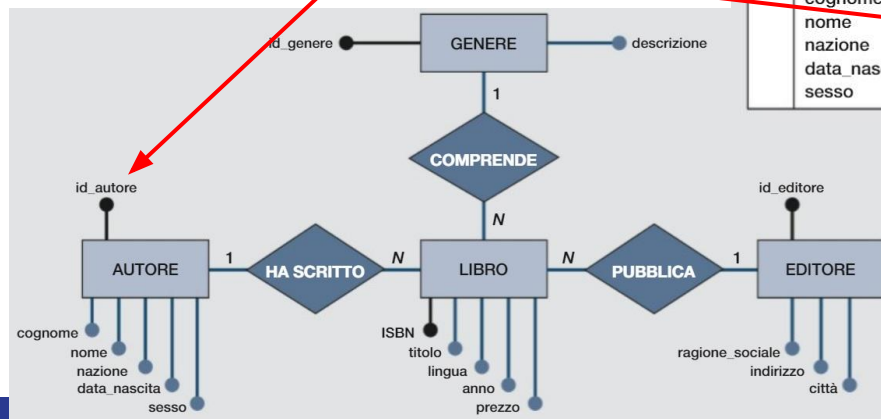
Quelli di chiave primaria e chiave esterna sono concetti relativi alla progettazione logica; cmq già nel diagramma E/R, per ogni entità, vengono evidenziati gli attributi chiave che presumibilmente nel modello logico diventeranno altrettante chiavi primarie, mentre non viene inserito alcun riferimento alle chiavi esterne.

In ogni caso, quando due tabelle derivano da due entità dello schema E/R legate da un'associazione *uno-a-molti*, durante la progettazione logica è necessario copiare i campi della chiave primaria del dominio (la tabella dal lato uno dell'associazione) nel codominio (quella dal lato molti) in modo che ciascuna riga della tabella codominio riporti le stesse informazioni della chiave primaria della riga della tabella dominio a cui risulta essere associata.

il modello dei dati relazionale - vincoli di integrità

Ad es. tra gli autori ed i libri scritti da questi esiste un'associazione *uno-a-molti* in cui l'entità *Autore* è il dominio, mentre l'entità *Libri* è il codominio, visto che ogni autore può scrivere più libri, mentre ogni libro può essere scritto da un unico autore.

La chiave primaria del dominio viene riportata come chiave esterna nel codominio.



Lo stesso discorso vale anche per le chiavi primarie delle entità *Editori* e *Genere*.

il modello dei dati relazionale - vincoli di integrità

Oltre al *vincolo di chiave* e al *vincolo di integrità sull'entità*, esiste un ulteriore importante vincolo *interrelazionale* detto di **integrità referenziale** che solitamente viene gestito al livello fisico dal DBMS in uso e che regola la consistenza e l'integrità dei collegamenti (riferimenti) tra i dati in presenza di un'associazione che lega tra loro due o più tabelle.

Si tratta di un insieme di regole del modello relazionale specifiche per rendere valide le associazioni tra le tabelle ed eliminare gli eventuali errori possibili in operazioni di inserimento modifica o cancellazione di dati collegati tra loro mediante chiave esterna.

In pratica, **per ogni valore non nullo** (diverso da `null`) **della chiave esterna deve necessariamente esistere un valore corrispondente della chiave primaria nella tabella associata.**

il modello dei dati relazionale - vincoli di integrità

Quello mostrato è lo schema della relazione *Libri*, corrispondente all'omonima entità dell'esempio della biblioteca:

Libri(ISBN, titolo, lingua, anno, prezzo, id_autore, id_genere, **id_editore**)

Nello schema l'attributo *id_editore* è una chiave esterna che riferisce la chiave primaria dello schema *Editori*:

Editori(**id_editore**, ragione_sociale, indirizzo, città)

Alcune possibili configurazioni delle tabelle che estendono questi schemi di relazione sono le seguenti:

ISBN	titolo	lingua	anno	prezzo	id_genere	id_editore
123-456-789	Pinocchio	Italiano	1990	10,00€	G01	E100
987-654-321	The Hobbit	Inglese	1999	10,00€	G01	E101

id_editore	ragione_sociale	indirizzo	città
E099	Zanichelli	Via Irnerio, 34	Bologna
E100	Mondadori	Via Bianca di Savoia, 12	Milano
E101	RCS	Via Rizzoli, 8	Milano

il modello dei dati relazionale - vincoli di integrità

Il valore della chiave esterna del libro «Pinocchio» («E100») corrisponde al valore della chiave primaria dell'editore «Mondadori», mentre il valore della chiave esterna del libro «The Hobbit» («E101») corrisponde al valore della chiave primaria dell'editore «RCS».

Dovendo aggiungere un nuovo libro alla tabella *Libri*, questo potrà essere pubblicato solo da uno dei tre editori presenti nella tabella *Editori*, perché il contenuto della chiave esterna *id_editore* deve coincidere con una delle chiavi primarie della tabella *Editori* (*id_editore*), altrimenti bisogna aggiungere prima il nuovo editore nella tabella *Editori*.

ISBN	titolo	lingua	anno	prezzo	id_genere	id_editore
123-456-789	Pinocchio	Italiano	1990	10,00€	G01	E100
987-654-321	The Hobbit	Inglese	1999	10,00€	G01	E101

id_editore	ragione_sociale	indirizzo	città
E099	Zanichelli	Via Irnerio, 34	Bologna
E100	Mondadori	Via Bianca di Savoia, 12	Milano
E101	RCS	Via Rizzoli, 8	Milano

il modello dei dati relazionale - vincoli di integrità

Quando viene applicata l'integrità referenziale, è dunque necessario osservare le seguenti regole pratiche:

- Non è possibile inserire un valore nella chiave esterna della tabella correlata (*codominio* dell'associazione) se tale valore non esiste tra quelli già registrati nella chiave primaria della tabella principale (*dominio*);
NB: è invece possibile immettere un valore `null`, se il record in questione non è correlato con nessun altro.
- Non si può eliminare un record dalla tabella principale, se esistono righe legate ad essa attraverso la chiave esterna nella tabella correlata.
- Non si può modificare un valore della chiave primaria nella tabella principale se ad essa corrispondono righe nella tabella correlata.

il modello dei dati relazionale - vincoli di integrità

Supponiamo di avere una tabella per la registrazione dei dati anagrafici dei clienti ed una che riporta le informazioni sugli agenti che li seguono; ogni agente può seguire più clienti contemporaneamente, mentre ogni cliente può essere seguito da un solo agente per cui resta stabilita un'associazione *uno-a-molti* tra *Agenti* e *Clienti* (dove *CodAgente* è chiave esterna per *Clienti*).

L'agente Carlo Bressi segue contemporaneamente tre clienti: Antonia Rossi, Giacomo Alfieri e Maria Petrocelli.

NB: Il cliente Gianfranco Verdi non è attualmente seguito da nessun agente, per cui il campo *CodAgente* ha valore *null*.

CodAgente	NomeAgente	CodZona
AK 0098	Carlo Bressi	AB 2
PM 0753	Pietro Conti	PP 5
JP 1123	Alba Sergi	KR 1
Agenti		

CodAgente	CodCli	Nominativo	Indirizzo	Prov
AK 0098	C 002	Antonia Rossi	Via Roma, 126	RM
AK 0098	A 011	Giacomo Alfieri	Via Napoli, 29	NA
JP 1123	A 004	Franco Pisori	Via Venezia, 38	VE
AK 0098	C 014	Maria Petrocelli	Via Bologna, 78	BO
PM 0753	B 003	Piero Giovanelli	Via Palermo, 23	NA
null	D 147	Gianfranco Verdi	Via Cosenza, 18	CS
JP 1123	A 012	Carla Vecchi	Via Salerno, 88	MI
Clienti				

il modello dei dati relazionale - vincoli di integrità

In virtù del vincolo di integrità referenziale, modificando ad es. il CodAgente “AK 0098” in “BB 0003”, questo verrà modificato automaticamente anche in tutti i record correlati della tabella *Clienti*, oppure non sarà permesso eseguire tale operazione.

Come accennato prima, è lo specifico DBMS che al livello fisico si fa carico di far rispettare le regole imposte dall'integrità referenziale.

In alcuni casi, bisogna impostarla esplicitamente tramite qualche opzione del DBMS.



il modello dei dati relazionale - vincoli di integrità

Se invece cancelliamo il record “BB 0003” dalla tabella Agenti, verranno eliminati dalla tabella Clienti anche quelli con lo stesso valore nella chiave esterna (cioè i clienti seguiti da quell’agente), oppure, ancora una volta, il DBMS non consentirà l’esecuzione di questa operazione.

