

*Corso di Laurea in Ingegneria Gestionale
Sapienza - Università di Roma*

Corso di Basi di Dati

A.A. 2019/2020

2 – Il Modello Relazionale

Tiziana Catarci

Sistemi di Basi di Dati

- ▶ **Base di Dati** : *Collezione di dati*, che tipicamente descrive le informazioni di interesse di una o più organizzazioni correlate.
- ▶ **DBMS (*Database Management System*)** : *Sistema software* in grado di memorizzare, gestire e interrogare grandi collezioni di dati.



**Come vengono rappresentati
i dati in un DBMS ?**



- ▶ **Modello dei dati** : *Collezione di costrutti* utilizzati per organizzare i dati di interesse e descriverne la struttura in modo che risulti comprensibile ad un elaboratore. Il modello più diffuso è il **modello relazionale**.

Un po' di storia...

- ▶ **Modello reticolare (inizio anni '60):**
 - ▶ struttura a record e puntatori;
 - ▶ Es. (Integrated Data Store) IDS di Bachman (Turing Award nel 1973).
- ▶ **Modello gerarchico (fine anni '60):**
 - ▶ struttura ad albero;
 - ▶ Es. IMS dell'IBM (1968).
- ▶ **Modello relazionale (1970):**
 - ▶ proposto da Edgar F. Codd per favorire l'indipendenza dei dati;
 - ▶ Turing Award nel 1981.

Il Modello Relazionale

- ▶ È il modello più largamente usato.
 - ▶ Produttori: IBM, Microsoft, Oracle, etc.
 - ▶ Competitori recenti:
 - ▶ modello orientato agli oggetti;
 - ▶ altro modello emergente: XML.
- ▶ Disponibile come modello logico in DBMS reali a partire dal 1981 (non è facile realizzare l'indipendenza dei dati con efficienza e affidabilità!).
- ▶ Si basa sul concetto matematico di **relazione** (con una variante).
- ▶ Le relazioni provengono dalla teoria degli insiemi ed hanno una rappresentazione naturale per mezzo di **tabelle**.
 - ▶ La presenza contemporanea di questi due concetti, uno formale e l'altro intuitivo, è responsabile del grande successo ottenuto dal modello relazionale.
- ▶ Il modello è “**basato su valori**”: i riferimenti fra dati in strutture (relazioni) diverse sono rappresentati per mezzo dei valori stessi.

Relazione matematica

- ▶ D_1, D_2, \dots, D_n (n insiemi – detti **domini** della relazione - anche non distinti)
- ▶ Il **prodotto cartesiano** $D_1 \times D_2 \times \dots \times D_n$ è l'insieme di tutte le n-uple **ordinate** (d_1, d_2, \dots, d_n) tali che $d_1 \in D_1, d_2 \in D_2, \dots, d_n \in D_n$
- ▶ Una relazione matematica su D_1, D_2, \dots, D_n è un **sottoinsieme del prodotto cartesiano** $D_1 \times D_2 \times \dots \times D_n$

Insieme di tutte le possibile coppie in cui il primo elemento appartiene a D_1 e il secondo a D_2

$D_1 = \{a, b\}$
 $D_2 = \{x, y, z\}$

Prodotto Cartesiano : $D_1 \times D_2$

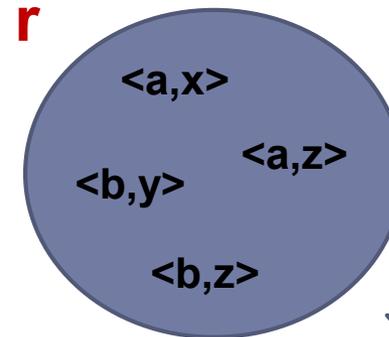
- Una relazione su n domini ha **grado** (o **arità**) **n**.
- Il numero di n-uple indica la **cardinalità** della relazione.

• **ESEMPIO** :

- ❖ grado = 2
- ❖ cardinalità = 6

D_1	D_2
a	x
a	y
a	z
b	x
b	y
b	z

Una relazione $r \subseteq D_1 \times D_2$



D_1	D_2
a	x
a	z
b	y
b	z

- Una relazione è un **insieme**:
- **Non è definito alcun ordinamento** fra le n-uple;
 - Le n-uple sono **distinte tra loro**;
 - **Ciascuna n-upla è ordinata**.

Il Modello Relazionale

- ▶ Il costrutto di base per la descrizione dei dati è la **relazione**. Una relazione è sostanzialmente una **tabella**.
 - ▶ A ciascun dominio è associato un **nome** (**attributo**), unico nella relazione, che **descrive il ruolo del dominio**. Gli attributi sono usati come intestazione delle colonne (il cui ordinamento è irrilevante).
 - ▶ Le righe della tabella rappresentano specifici **record** (o **tuple**) **diversi fra loro**.

Nome	Cognome	Posizione	Squadra	Età
Tommaso	Rocchi	Attaccante	S.S.Lazio	31
Alessandro	Del Piero	Attaccante	Juventus F.C.	34
Francesco	Totti	Attaccante	A.S. Roma	32

Tupla : l'i-esimo valore proviene dall'i-esimo dominio!

Tabelle e Relazioni

- ▶ Una tabella rappresenta una relazione se :
 - ▶ i valori di ciascuna colonna sono fra loro omogenei (appartengono allo stesso dominio);
 - ▶ le righe sono diverse fra loro;
 - ▶ le intestazioni delle colonne (attributi) sono diverse tra loro.
- ▶ Inoltre, in una tabella che rappresenta una relazione :
 - ▶ l'ordinamento tra le righe è irrilevante;
 - ▶ l'ordinamento tra le colonne è irrilevante (struttura non posizionale).
- ▶ **Il modello relazionale è basato su valori** : i riferimenti fra dati in relazioni diverse sono rappresentati per mezzo di valori dei domini che compaiono nelle tuple.

Struttura basata su valori

studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

Informazioni relative ad un insieme di studenti.

esami

Studente	Voto	Corso
3456	30	04
3456	24	02
9283	28	01
6554	26	01

Informazioni relative agli esami sostenuti da specifici studenti (individuati tramite il *numero di matricola*) per determinati corsi (rappresentati attraverso i relativi *codici*).

I riferimenti sono realizzati tramite valori.

corsi

Codice	Titolo	Docente
01	Analisi	Mario
02	Chimica	Bruni
04	Chimica	Verdi

Informazioni relative ai corsi frequentati.

Struttura basata su puntatori

Stessa base di dati del caso precedente, ma rappresentata in modo esplicito attraverso puntatori.

studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

esami

Studente	Voto	Corso
	30	
	24	
	28	
	26	

corsi

Codice	Titolo	Docente
01	Analisi	Mario
02	Chimica	Bruni
04	Chimica	Verdi

Struttura tipica dei modelli reticolare e gerarchico.

Vantaggi della struttura basata su valori

- ▶ **Indipendenza dalle strutture fisiche**, che possono cambiare anche dinamicamente.
- ▶ Si rappresenta solo ciò che è **rilevante dal punto di vista dell'applicazione** (dell'utente):
 - **i puntatori sono meno comprensibili** per l'utente finale, e sono legati ad aspetti realizzativi.
- ▶ I dati sono **portabili più facilmente da un sistema ad un altro**:
 - essendo l'informazione contenuta nei valori, è **relativamente semplice trasferire basi di dati da un calcolatore ad un altro**;
 - in presenza di puntatori, l'operazione è più complessa, perché i puntatori hanno un significato locale al singolo sistema.
- ▶ I valori consentono **bi-direzionalità**, mentre i puntatori sono direzionali.
- ❖ **Nota:** nel modello relazionale, **i puntatori possono essere usati a livello fisico** (ma non devono essere visibili a livello logico).

Relazione nel modello relazionale dei dati

Casa	Fuori	RetiCasa	RetiFuori
Juve	Lazio	3	1
Lazio	Milan	2	0
Juve	Roma	1	2
Roma	Milan	0	1

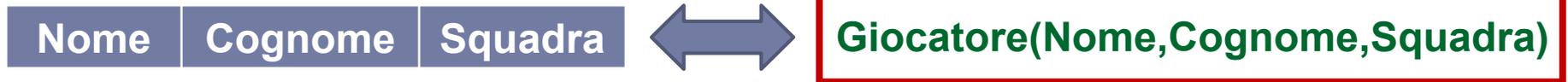
- ▶ I domini degli attributi sono **String** per *Casa* e *Fuori*, ed **Integer** per *RetiCasa* e *RetiFuori*. Essi non vengono mostrati nella rappresentazione tabellare.
- ▶ Sia **X** l'insieme degli attributi di una relazione **R**. Se **t** è una tupla di **R**, cioè una tupla su **X**, e **A** \in **X**, allora **t[A]** (oppure **t.A**) indica il valore che **t** ha in corrispondenza dell'attributo **A**.
 - ▶ Se **t** è la prima tupla della tabella, allora si ha che **t[Fuori] = Lazio**
 - ▶ **t[Fuori,RetiFuori]** indica una tupla sui due attributi **Fuori** e **RetiFuori**.
 - ▶ Riferendoci alla prima tupla **t** della tabella, si ha che **t[Fuori,RetiFuori] = <Lazio, 1>**

Alcune Notazioni

- ▶ **Schema di relazione** : un nome di relazione **R** con un insieme di attributi $X = \{A_1, \dots, A_n\}$ e corrispondenti domini D_1, \dots, D_n si può denotare come :

$R(A_1:D_1, \dots, A_n:D_n)$ oppure $R(A_1, \dots, A_n)$

Giocatore



- ▶ **(Istanza di) relazione** su uno schema $R(X)$: insieme **r** di tuple su **X**.

Giocatore

Nome	Cognome	Squadra
Tommaso	Rocchi	S.S.Lazio
Alessandro	Del Piero	Juventus F.C.
Francesco	Totti	A.S. Roma

Istanza della relazione **Giocatore**

Informazione Incompleta – Valore NULL

Prefetture

Città	Prefettura
Roma	Via Novembre
Firenze	NULL
Tivoli	NULL
Prato	NULL

Di Firenze non si conosce l'indirizzo della prefettura.

Tivoli non è provincia: non ha prefettura.

Prato è “nuova” provincia: ha la prefettura?

Come vengono rappresentati questi valori?

- ▶ **VALORE NULLO:** può essere interpretato in tre modi differenti:
 - ▶ valore *sconosciuto* : esiste un valore del dominio, ma non è noto (Firenze).
 - ▶ valore *inesistente* : non esiste un valore del dominio (Tivoli).
 - ▶ valore *senza informazione* : non è noto se esista o meno un valore del dominio (Prato).

I DBMS NON DISTINGUONO I TIPI DI VALORE NULLO (implicitamente adottano l'interpretazione «SENZA INFORMAZIONE»).

Esercizio

- ▶ *Sia dato il seguente schema relazionale, il cui scopo è rappresentare una base di dati che memorizzi i dati relativi agli studenti di un'università.*
- ▶ **Indicare se i valori NULL inseriti sono ammissibili rispetto al contesto.**

Studenti

Matricola	Cognome	Nome	Data di Nascita
276545	Rossi	Maria	NULL
NULL	Verdi	Fabio	12/02/1972
345678	NULL	Fabio	12/02/1972

AMMISSIBILE :
l'informazione in questo
contesto non è essenziale.

NON AMMISSIBILE :
la matricola identifica uno
studente.

NON AMMISSIBILE :
non è possibile rappresentare
uno studente senza cognome.

ATTENZIONE :
le ultime due tuple della relazione
possono essere diverse o
addirittura coincidere.

Vincoli di Integrità

- ▶ Un DBMS **deve prevenire** l'immissione di informazioni non corrette.
- ▶ Ad uno schema di base di dati si può associare un insieme di **vincoli di integrità**.
- ▶ Un'istanza della base di dati che soddisfa tutti i vincoli di integrità specificati nello schema si dice **LEGALE**.
- ▶ ***Tipi di vincoli :***
 - ▶ **Intrarelazionali** : *il loro soddisfacimento è definito rispetto a singole relazioni di una base di dati.*
 - ▶ *Vincoli di **tupla**.*
 - ▶ *Vincoli di **chiave**.*
 - ▶ **Interrelazionali** : *vincoli che coinvolgono più relazioni della Base di Dati*
 - ▶ *Vincoli di **integrità referenziale** (o vincoli di **foreign key**).*

Vincoli di tupla

- ▶ Esprimono **condizioni sui valori di ciascuna tupla**, indipendentemente dalle altre tuple.
- ▶ Fondamentali per garantire che ciascun insieme di tuple sullo schema rappresenti informazioni corrette per l'applicazione.
- ▶ In particolare, un vincolo di tupla si definisce **vincolo di dominio** se coinvolge un solo attributo.

Studenti

Matricola	Voto	Lode
123456	36	NO
654321	30	NO
321654	30	SI
123456	25	SI

Nel sistema italiano i voti ammissibili vanno da 18 a 30...

...e la lode può essere assegnata solo se il voto è pari a 30....

Come esprimere queste condizioni sulla relazione?

Vincoli di tupla

- ▶ Una possibile sintassi per esprimere vincoli di questo tipo è quella che permette di definire espressioni booleane (cioè, con connettivi AND, OR, e NOT) confrontando *valori di attributo* o *espressioni aritmetiche* su valori di attributo.

Voto \geq 18 AND Voto \leq 30
(NOT(LODE = 'SI')) OR (Voto = 30)

Studenti

Matricola	Voto	Lode
123456	36	NO
654321	30	NO
321654	30	SI
123456	25	SI

Grazie ai vincoli di tupla, gli inserimenti errati non sono più consentiti.

Vincoli di tupla

- ▶ La tabella seguente soddisfa il vincolo:

(SeggiAperti>400) AND (SeggiAperti<600) ?

Regione	SeggiAperti
Piemonte	567
Lombardia	670
Sicilia	594

NO!!!!
La Lombardia ha
670 seggi aperti

- ▶ Esprimere il seguente vincolo : “La regione Lazio non deve avere esattamente 500 seggi”.

NOT(SeggiAperti=500 AND Regione=“Lazio”)

Vincoli di Chiave

- ▶ Una *chiave* è un insieme di attributi utilizzato per identificare univocamente le tuple di una relazione.

Studenti

Matricola	Cognome	Nome	Data di Nascita
276545	Rossi	Maria	03/05/1975
345678	Verdi	Fabio	12/02/1972
745989	Rossi	Fabio	12/02/1972

Matricola è **CHIAVE** per la relazione Studenti.

<Matricola,Cognome,Nome,Data di Nascita> è **SUPERCHIAVE** per la relazione Studenti, ma NON E' CHIAVE.

<Cognome,Nome> è **SUPERCHIAVE** per la relazione Studenti; dato che nessuno dei suoi sottoinsiemi è superchiave, allora è anche **CHIAVE**.

- ▶ un insieme **K** di attributi è **superchiave per una** istanza di relazione **r** se **r** non contiene due tuple distinte **t1** e **t2** tali che **t1[K] = t2[K]**.
- ▶ Una *superchiave* è un insieme di campi che contiene una chiave.
- ▶ **K** è **chiave per r** se è una **superchiave minimale** (cioè non contiene un'altra superchiave) per **r**.

Vincoli di Chiave

- ▶ Un **vincolo di chiave** è un'asserzione che specifica che un insieme di attributi formano una chiave per una relazione.
- ▶ In altre parole, se in una relazione $\mathbf{R(A,B,C,D)}$ dichiaro un vincolo di chiave su $\{\mathbf{A,B}\}$, sto asserendo che in tutte le istanze della basi di dati, non esistono due tuple della relazione \mathbf{R} che coincidono negli attributi \mathbf{A} e \mathbf{B} e sto anche asserendo che nessun sottoinsieme proprio di $\{\mathbf{A,B}\}$ è una chiave.
- ▶ Non ci sono limitazioni per il numero di vincoli di chiave che si definiscono per una relazione (a parte il limite derivante dal numero di attributi).

Esistenza delle chiavi

- ▶ Poiché le **relazioni sono insiemi**, una relazione non può contenere tuple uguali fra loro.
 - Ne segue che **ogni relazione ha come superchiave l'insieme degli attributi su cui è definita**.
- ▶ Poiché l'insieme di tutti gli attributi è una superchiave per ogni relazione, ogni schema di relazione ha almeno una superchiave.
- ▶ Ne segue che **ogni schema di relazione ha (almeno) una chiave!**

Importanza delle chiavi

- ▶ L'esistenza delle chiavi garantisce l'accessibilità a ciascun dato della base di dati.
- ▶ Ogni singolo valore è univocamente accessibile tramite:
 - nome della relazione;
 - valore della chiave (che indica al massimo una tupla della relazione);
 - nome dell'attributo in corrispondenza del quale è presente il valore da accedere.
- ▶ Come vedremo più avanti, le chiavi sono lo strumento principale attraverso il quale vengono correlati i dati in relazioni diverse (“il modello relazionale è basato su valori”).

Chiavi e Valori nulli

- ▶ In presenza di valori nulli, i valori degli attributi che formano la chiave:
 - non permettono di identificare le tuple come desiderato;
 - non permettono di realizzare facilmente i riferimenti da altre relazioni.

Matricola	Cognome	Nome	Corso	Nascita
NULL	NULL	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Civile	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	NULL
NULL	Neri	Mario	NULL	5/12/78

Questa tupla non è identificabile in alcun modo. Potenzialmente potrebbe essere uguale alle ultime due.

La presenza di valori NULL rende impossibile capire se le due tuple fanno riferimento allo stesso studente o a due omonimi.

Chiave Primaria

- ▶ I valori delle chiavi permettono di identificare univocamente le tuple delle relazioni e di stabilire riferimenti fra tuple di relazioni diverse.
- ▶ I valori NULL rendono difficile tale identificazione...è necessario un meccanismo che ne limiti la presenza in almeno una chiave della relazione.

Studenti

<u>Matricola</u>	Cognome	Nome	Data di Nascita
276545	Rossi	Maria	NULL
234567	Verdi	Fabio	12/02/1972
345678	NULL	Fabio	12/02/1972

La presenza di valori **NULL** sull'attributo Matricola renderebbe impossibile capire se le tuple fanno riferimento allo stesso studente.

- ▶ **SOLUZIONE** : su una delle chiavi “candidate” si **vieta** la presenza di valori NULL ; tale chiave viene chiamata **CHIAVE PRIMARIA** (graficamente, gli attributi facenti parte della chiave primaria vengono **sottolineati**).

Chiave Primaria

- ▶ Un **vincolo di chiave primaria** è un'asserzione che specifica che :
 - ▶ un insieme di attributi formano una chiave per una relazione e
 - ▶ non si ammettono per tali attributi i valori nulli.
- ▶ La chiave primaria viene scelta tra le chiavi disponibili nella relazione.
- ▶ In quasi tutti i casi reali si dispone sempre di attributi i cui valori sono identificanti (matricola, targa, codice fiscale...).
- ▶ Quando ciò non accade, **è necessario introdurre un attributo aggiuntivo** probabilmente non significativo dal punto di vista dell'applicazione (ad esempio, un *codice*) che viene attribuito a ciascuna tupla all'atto dell'inserimento.
- ▶ **Un solo vincolo di chiave primaria** è ammesso per ciascuna relazione (mentre vi possono essere più chiavi).

Esercizio

- ▶ Definire uno schema di basi di dati che organizzi i dati necessari a generare la pagina dei programmi radiofonici di un quotidiano, con stazioni, ore e titoli dei programmi; per ogni stazione sono memorizzati, oltre al nome, anche la frequenza di trasmissione e la sede.

Vincoli di Integrità Referenziale

- ▶ Informazioni in relazioni diverse sono collegati attraverso valori comuni, in particolare **attraverso i valori delle chiave primarie**.
- ▶ Un vincolo di **integrità referenziale**, detto anche vincolo di **foreign key** (o **di chiave esterna**), fra un insieme di attributi **X** di una relazione **R₁** e un'altra relazione **R₂** impone ai valori su **X** di ciascuna tupla dell'istanza di **R₁** di comparire come valori della chiave primaria dell'istanza di **R₂**.

Ricoveri

<u>ID</u>	Reparto	Paziente
12	A	A102
13	B	A102
14	B	B372

Relazione referenziante

Pazienti

<u>Cod</u>	Cognome
A102	Necchi
B372	Rossini
B543	Missoni

Relazione referenziata

foreign key : Ricoveri(Paziente) \subseteq Pazienti(Cod)

Vincoli di Integrità Referenziale

Infrazioni

<u>Codice</u>	<u>Data</u>	<u>Vigile</u>	<u>Prov</u>	<u>Numero</u>
34321	1/2/95	3987	MI	39548K
53524	4/3/95	3295	TO	E39548
64521	5/4/96	3295	PR	839548
73321	5/2/98	9345	PR	839548

Vigili

<u>Matricola</u>	<u>Cognome</u>	<u>Nome</u>
3987	Rossi	Luca
3295	Neri	Piero
9345	Neri	Mario
7543	Mori	Gino

foreign key : $\text{Infrazioni}(\text{Vigile}) \subseteq \text{Vigili}(\text{Matricola})$

ATTENZIONE

La chiave esterna nella relazione referenziante deve avere lo stesso numero di colonne e tipi di dati compatibili della chiave primaria nella relazione referenziata.

Vincolo di foreign key tra l'attributo *Vigile* della relazione *Infrazioni* e la relazione *Vigili*.

Vincoli di Integrità Referenziale

Infrazioni

<u>Codice</u>	Data	Vigile	Prov	Numero
34321	1/2/95	3987	MI	39548K
53524	4/3/95	3295	TO	E39548
64521	5/4/96	2468	PR	839548
73321	5/2/98	9345	PR	839548

Vigili

<u>Matricola</u>	Cognome	Nome
3987	Rossi	Luca
3295	Neri	Piero
9345	Neri	Mario
7543	Mori	Gino

VIOLAZIONE del vincolo di foreign key tra l'attributo *Vigile* della relazione *Infrazioni* e la relazione *Vigili*.

Vincoli di Integrità Referenziale

Infrazioni

<u>Codice</u>	<u>Data</u>	<u>Vigile</u>	<u>Prov</u>	<u>Numero</u>
34321	1/2/95	3987	MI	39548K
53524	4/3/95	3295	TO	E39548
64521	5/4/96	3295	PR	839548
73321	5/2/98	9345	PR	839548

Auto

<u>Prov</u>	<u>Numero</u>	<u>Cognome</u>	<u>Nome</u>
MI	39548K	Rossi	Mario
TO	E39548	Rossi	Mario
PR	839548	Neri	Luca

Vincolo di foreign key tra gli attributi *Prov* e *Numero* della relazione *Infrazioni* e la relazione *Auto*.

foreign key : $\text{Infrazioni}(\text{Prov}, \text{Numero}) \subseteq \text{Auto}(\text{Prov}, \text{Numero})$

Vincoli di Integrità Referenziale

Infrazioni

<u>Codice</u>	<u>Data</u>	<u>Vigile</u>	<u>Prov</u>	<u>Numero</u>
34321	1/2/95	3987	MI	39548K
53524	4/3/95	3295	TO	39548K
64521	5/4/96	3295	PR	839548
73321	5/2/98	9345	PR	839548

Auto

<u>Prov</u>	<u>Numero</u>	<u>Cognome</u>	<u>Nome</u>
MI	39548K	Rossi	Mario
TO	E39548	Rossi	Mario
PR	839548	Neri	Luca

VIOLAZIONE del vincolo di foreign key tra gli attributi *Prov* e *Numero* della relazione *Infrazioni* e la relazione *Auto*.

Vincoli di Integrità Referenziale

Incidenti

Codice	Data	ProvA	NumeroA	ProvB	NumeroB
34321	1/2/95	TO	E39548	MI	39548K
64521	5/4/96	PR	839548	TO	E39548

Auto

Prov	Numero	Cognome	Nome
MI	39548K	Rossi	Mario
TO	E39548	Rossi	Mario
PR	839548	Neri	Luca

Vincoli multipli su più attributi.

Sono presenti 2 vincoli di foreign key:

foreign key : Incidenti(ProvA,NumeroA) \subseteq Auto(Prov,Numero)

foreign key : Incidenti(ProvB,NumeroB) \subseteq Auto(Prov,Numero)

ESERCIZIO

Individuare quanti e quali vincoli di foreign key sono presenti tra la relazione *Incidenti* e la relazione *Auto*



Vincoli di Integrità Referenziale

Impiegati

<u>Matricola</u>	Cognome	Progetto
34321	Rossi	IDEA
53524	Neri	XYZ
64521	Verdi	NULL
73032	Bianchi	IDEA

La presenza di un valore **NULL** in una chiave esterna ***non viola il vincolo.***

Progetti

<u>Codice</u>	Inizio	Durata	Costo
IDEA	01/2000	36	200
XYZ	07/2001	24	120
BOH	09/2001	24	150

In presenza di valori **NULL** i vincoli possono essere resi meno restrittivi...

Vincoli di Integrità: Considerazioni

- ▶ Usato a sproposito, un vincolo di integrità può impedire la memorizzazione di istanze della base di dati che si presentano nella pratica.
- ▶ I vincoli di integrità sono basati sulla semantica dell'organizzazione del mondo reale che viene descritta nelle relazioni della base di dati.
 - ▶ Possiamo controllare l'istanza di una base di dati per vedere se un vincolo di integrità viene violato, ma non possiamo MAI inferire che un vincolo di integrità è vero solo esaminando un'istanza.
 - ▶ Un vincolo di integrità è una asserzione su tutte le possibili istanze!
- ▶ Vincoli di integrità su chiavi e su chiavi esterne sono i più comuni; sono supportati anche vincoli di integrità più generali.

Modello Relazionale: Conclusioni

- ▶ Una rappresentazione tabellare dei dati:
 - ▶ semplice e intuitiva, attualmente la più usata.
 - ▶ Il DBA può specificare vincoli di integrità basati sulla semantica dell'applicazione. Il DBMS controlla eventuali violazioni.
 - ▶ Due vincoli di integrità importanti: chiavi primarie e chiavi esterne.
 - ▶ Inoltre abbiamo *sempre* vincoli di dominio.
- ▶ Supporta un semplice ma potente *linguaggio di interrogazione* dei dati.
 - ▶ Le interrogazioni possono essere scritte in maniera intuitiva, e il DBMS è responsabile di una valutazione efficiente.
 - ▶ Fattore chiave: semantica precisa per le interrogazioni relazionali.
 - ▶ Permette all'ottimizzatore di riordinare estensivamente le operazioni, garantendo che la risposta non cambi.

ESERCIZIO 1

Medici

Matricola	Cognome	Nome	Reparto
203	Neri	Piero	A
574	Bisi	Mario	B
461	Bargio	Sergio	B
530	Belli	Nicola	C
405	Mizzi	Nicola	A
501	Monti	Mario	A

Reparti

Cod	Nome	Primario
A	Chirurgia	203
B	Pediatria	574
C	Medicina	530

ESERCIZIO

1) Individuare le Chiavi primarie delle due relazioni e gli eventuali vincoli di foreign key

2) L'insieme $\langle \text{Matricola}, \text{Cognome}, \text{Nome} \rangle$ è chiave?

Soluzione ESERCIZIO 1

Medici

<u>Matricola</u>	Cognome	Nome	Reparto
203	Neri	Piero	A
574	Bisi	Mario	B
461	Bargio	Sergio	B
530	Belli	Nicola	C
405	Mizzi	Nicola	A
501	Monti	Mario	A

Medici(Matricola, Cognome, Nome, Reparto)
primary Key : **Matricola**
foreign key :
Medici(Reparto) \subseteq Reparti(Cod)

2) L'insieme
<Matricola,Cognome,Nome> è
chiave?

NO!!!!!!!!!!

E' SUPERCHIAVE

Reparti

<u>Cod</u>	Nome	Primario
A	Chirurgia	203
B	Pediatria	574
C	Medicina	530

Reparti(Cod, Nome, Primario)
primary Key : **Cod**
foreign key : **Reparti(Primario) \subseteq Medici(Matricola)**

ESERCIZIO 2

Considerare la base di dati relazionale in figura, relativa a impiegati, progetti e partecipazioni di impiegati a progetti. Indicare quali possano essere, per questa base di dati, ragionevoli chiavi primarie e vincoli di integrità referenziale. Giustificare brevemente la risposta, con riferimento alla realtà di interesse (cioè perchè si può immaginare che tali vincoli sussistano) e all'istanza mostrata (verificando che sono soddisfatti).

IMPIEGATI

Matricola	Cognome	Nome	Età
101	Rossi	Mario	35
102	Rossi	Anna	42
103	Gialli	Mario	34
104	Neri	Gino	45

PROGETTI

ID	Titolo	Costo
A	Luna	70
B	Marte	60
C	Giove	90

PARTECIPAZIONE

Impiegato	Progetto
101	A
101	B
103	A
102	B

ESERCIZIO 3

Si vuole realizzare una base di dati per la comunità scientifica di ricerca paleontologica. Si devono memorizzare i dati riguardanti i reperti fossili di vertebrati custoditi in diversi musei, tenendo conto delle seguenti informazioni:

- ▶ I reperti sono caratterizzati dal luogo e dall'anno di ritrovamento, dal ricercatore responsabile della scoperta, dal museo e dalla sala in cui sono custoditi.
- ▶ I musei hanno un nome, un direttore (che assumiamo essere anche un ricercatore), un indirizzo, una città e un paese.
- ▶ Le sale dei musei hanno un identificatore, un nome e una dimensione.
- ▶ I ricercatori sono caratterizzati da un codice identificativo, un nome, un cognome e una data di nascita.

Produrre uno o più schemi di relazione per tale base di dati adoperando il modello relazionale. Si evidenzino le chiavi ed i vincoli di integrità referenziale dello schema.

Soluzione ESERCIZIO 3

Reperti

<u>Codice</u>	Luogo	Anno	Ricercatore	Sala
---------------	-------	------	-------------	------

Reperti(Codice,Luogo,Anno,Ricercatore,Museo,Sala)

primary key : **Codice**

foreign key : Reperti(Ricercatore) \subseteq Ricercatori(Codice)

foreign key : Reperti(Sala) \subseteq Sale(Id)

Musei

<u>Codice</u>	Nome	Direttore	Indirizzo	Città	Paese
---------------	------	-----------	-----------	-------	-------

Musei(Codice,Nome,Direttore,Indirizzo,Città,Paese)

primary key : **Codice**

foreign key : Musei(Direttore) \subseteq Ricercatori(Codice)

Ricercatori

<u>Codice</u>	Nome	Cognome	Data di Nascita
---------------	------	---------	-----------------

Ricercatori(Codice,Nome,Cognome,
Data di Nascita)

primary key : **Codice**

Sale

<u>Id</u>	Nome	Dimensione	Museo
-----------	------	------------	-------

Sala(Id,Nome,Dimensione,Museo)

primary key : **Id**

foreign key : Sale(Museo) \subseteq Musei(Codice)



ESERCIZIO 4

Sia data la seguente base di dati :

Squadra (Nome, Città, Sede, Colori)

Calciatore (Codice, Nome, Cognome, Ruolo, Nazionalità)

Ingaggio (Calciatore, Anno, Squadra, Stipendio)

Incontro (Data, SquadraInCasa, SquadraFuoriCasa, GolSqCasa, GolSqFuori, Arbitro)

Arbitro (Codice, Nome, Cognome)

- ▶ **Descrivere in linguaggio naturale le informazioni organizzate nella base di dati.**
- ▶ **Individuare le chiavi primarie, i vincoli di integrità referenziale e gli attributi sui quali è sensato ammettere valori nulli.**

Soluzione ESERCIZIO 4

- ▶ **Descrivere in linguaggio naturale le informazioni organizzate nella base di dati**

La base di dati descrive le informazioni inerenti ad un campionato di calcio. La relazione **Squadra** specifica *Nome*, *Città*, *Sede* e *Colori sociali* di ciascuna squadra. La relazione **Calciatore** descrive i singoli calciatori specificandone un *Codice*, il *Nome*, il *Cognome*, il *Ruolo* e la *Nazionalità*. La relazione **Ingaggio** specifica l'ingaggio di un *Calciatore* da parte di una *Squadra* indicandone lo *Stipendio* percepito a l'*Anno*. La relazione **Incontro** rappresenta i singoli incontri di Calcio indicando, per ciascuno, *Data*, *Squadre coinvolte*, *Risultato* e *Arbitro*. La relazione **Arbitro** infine descrive i singoli arbitri indicando un *Codice*, il *Nome* e il *Cognome*.

Soluzione ESERCIZIO 4

- ▶ **Individuare le chiavi primarie, i vincoli di integrità referenziale e gli attributi sui quali è sensato ammettere valori nulli**

Chiavi primarie:

- ▶ *Nome* per **Squadra**
- ▶ *Codice* per **Calciatore**
- ▶ *Calciatore* e *Anno* per **Ingaggio**
- ▶ *Data* e *SquadraInCasa* (o anche *Data* e *SquadraFuoriCasa*) per **Incontro**
- ▶ *Codice* per **Arbitro**

Vincoli di foreign key

- ▶ tra *Calciatore* in **Ingaggio** e la relazione **Calciatore**
- ▶ tra *Squadra* in **Ingaggio** e la relazione **Squadra**
- ▶ tra *SquadraInCasa* in **Incontro** e la relazione **Squadra**
- ▶ tra *SquadraFuoriCasa* in **Incontro** e la relazione **Squadra**
- ▶ tra *Arbitro* in **Incontro** e la relazione **Arbitro**

Soluzione ESERCIZIO 4

Possibili valori NULL

- ▶ I valori **NULL** possono essere ammessi in tutti quei campi che non sono chiavi primarie. Tra questi, ad esempio, potrebbe essere ragionevole ammettere valori nulli sugli attributi *Sede* e *Colori* di Squadra.

ESERCIZIO 5

- **Indicare quali tra le seguenti affermazioni sono vere in una definizione rigorosa del modello relazionale:**
1. ogni relazione ha almeno una chiave.
 2. ogni relazione ha esattamente una chiave.
 3. ogni attributo appartiene al massimo ad una chiave.
 4. possono esistere attributi che non appartengono a nessuna chiave.
 5. una chiave può essere sottoinsieme di un'altra chiave.
 6. può esistere una chiave che coinvolge tutti gli attributi.
 7. può succedere che esistano più chiavi e che una di esse coinvolga tutti gli attributi.
 8. ogni relazione ha almeno una superchiave.
 9. ogni relazione ha esattamente una superchiave.
 10. può succedere che esistano più superchiavi e che una di esse coinvolga tutti gli attributi.