



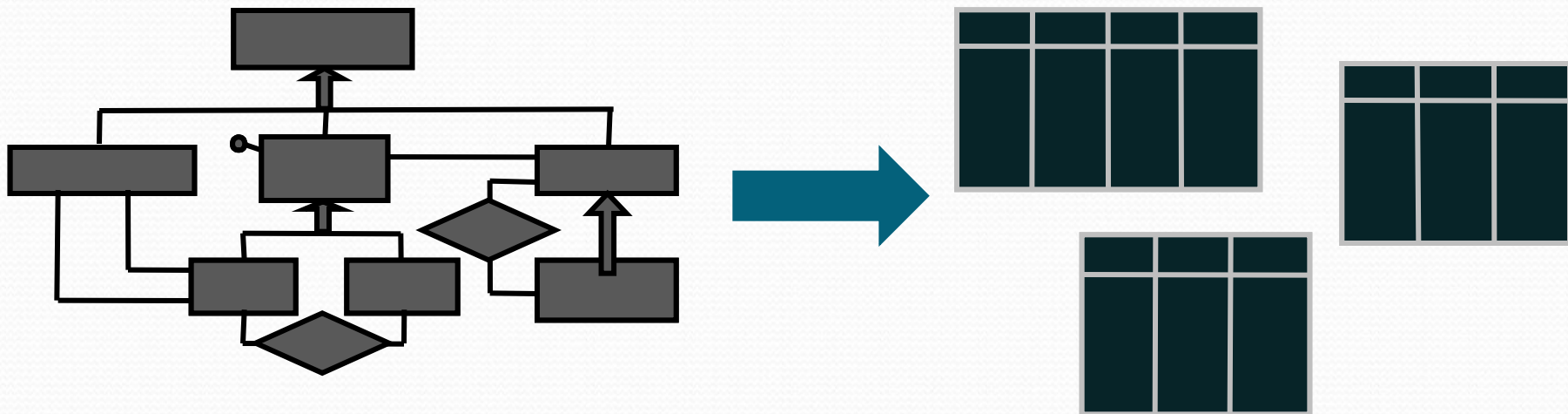
Basi Dati

Il data model Relazionale e i vincoli
per le basi di dati relazionali

a.a. 2021/2022
Prof.ssa G. Tortora

Da schema concettuale a logico

Terminata la fase di analisi concettuale del database e creato un modello di alto livello (tipicamente un ER o un EER) che descrive il *miniworld*, passiamo alla definizione di uno schema logico, più vicino al DBMS ma meno comprensibile ai “*non addetti ai lavori*”.



Il Data Model Relazionale

- Fu proposto da Codd nel 1970 per favorire l'indipendenza dei dati e reso disponibile come modello logico in DBMS reali nel 1981.
- È il modello più diffuso, sia a livello teorico che commerciale.
- La forza del modello relazionale è nella sua semplicità e nei solidi formalismi matematici su cui si poggia.

Il Data Model Relazionale (2)

- Si basa sul concetto matematico di **Relazione**.
- Le relazioni hanno una rappresentazione naturale per mezzo di **tabelle**:
 - Ciascuna riga rappresenta una collezione di valori di dati relati.
- Il database è rappresentato come una collezione di relazioni.

Il Data Model Relazionale - *Esempio*

EMPLOYEE	FNAME	MINIT	LNAME	<u>SSN</u>	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
	John	B	Smith	123456789	09-JAN-55	731 Fondren, Houston, TX	M	30000	333445555	5
	Franklin	T	Wong	333445555	08-DEC-45	638 Voss, Houston, TX	M	40000	888665555	5
	Alicia	J	Zelaya	999887777	19-JUL-58	3321 Castle, Spring, TX	F	25000	987654321	4
	Jennifer	S	Wallace	987654321	20-JUN-31	291 Berry, Bellaire, TX	F	43000	888665555	4
	Ramesh	K	Narayn	666884444	15-SEP-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
	Joyce	A	English	453453453	31-JUL-62	5631 Rice, Houston, TX	F	25000	333445555	5
	Ahmad	V	Jabbar	987987987	29-MAR-59	980 Dallas, Houston, TX	M	25000	987654321	4
	James	E	Borg	888665555	10-NOV-27	450 Stone, Houston, TX	M	55000	null	1

					DEPT_LOCATIONS	DNUMBER	DLOCATION
DEPARTMENT	DNAME	<u>DNUMBER</u>	MGRSSN	MGRSTARTDATE		1	Houston
	Research	5	333445555	22-MAY-78		4	Stafford
	Administration	4	987654321	01-JAN-85		5	Bellaire
	Headquarters	1	888665555	19-JUN-71		5	Sugarland
						5	Houston

Esempio di una parte del database “*Company*” nel data model Relazionale.

Concetti del modello Relazionale



Relazione dal punto di vista matematico

- Siano D_1, D_2, \dots, D_n n insiemi.
- 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 $D_1 \times D_2 \times \dots \times D_n$.
- D_1, D_2, \dots, D_n sono i **domini** della relazione. Una relazione su n domini ha grado n .
- Il numero di n -uple è la **cardinalità** della relazione. Nelle applicazioni reali, la cardinalità è sempre finita.

Relazione Matematica - *Esempio*

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

Prodotto cartesiano $D_1 \times D_2$

a	x
a	y
a	z
b	x
b	y
b	z

Una relazione $r \subseteq D_1 \times D_2$

a	x
a	y
b	y
b	z

Relazioni nel modello Relazionale

- A ogni dominio (attributo) è associato un nome, unico nella relazione, che “*descrive*” il ruolo del dominio.
- L'ordinamento fra gli attributi è irrilevante:
 - la struttura è **non posizionale**.
- Nella terminologia del modello relazionale, una riga è detta **tupla**:

Nome relazione

Attributi

STUDENT	Name	SSN	HomePhone	Address	OfficePhone	Age	GPA
<div>Tuple</div>	Benjamin Bayer	305-61-2435	373-1616	2918 BlueBonnet Lane	null	19	3.21
	Katherine Ashly	381-62-1245	375-4409	125 Kirby Road	null	18	2.89
	Dick Davidson	422-11-2420	null	3452 Elgin Road	749-1253	25	3.53
	Charles Cooper	489-22-1100	376-9821	265 Lark Lane	749-1253	28	3.93
	Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	null	19	3.25

Esempio di Relazione

Domini, attributi, tuple e relazioni

- Nel modello relazionale, un dominio D è un insieme di valori atomici, cioè indivisibili.
- Un metodo per specificare un dominio è specificare un tipo di dato da cui sono presi i dati che formano il dominio.

Esempi di domini

- **Usa_Phone_Numbers**: insieme di numeri a 10 cifre che rappresentano numeri telefonici validi negli stati uniti.
- **Social_Security_Number**: insieme di SSN validi, di 9 cifre.
- **Names**: insieme di nomi di persone.
- **Employee_Ages**: possibile età dei dipendenti, da 16 a 80 anni.
- **Academic_Department**: insieme di dipartimenti universitari (matematica e informatica, economia, ecc...).

Domini

- Per ogni dominio viene specificato un tipo di dato (o formato).
 - *Esempio:* `USA_PHONE_NUMBER` può essere dichiarato come una stringa *(ddd)ddd-dddd*.
- Potrebbe essere necessario specificare l'unità di misura per interpretare i valori di un dominio.
 - *Esempio:* `peso_persona` è espresso con l'unità di peso **kg**.

Schemi di relazione

- Uno **schema di relazione**, denotato da $R(A_1, A_2, \dots, A_n)$, descrive una relazione.
- Uno schema di relazione è formato da:
 - Un nome di relazione R ;
 - Una lista di attributi (A_1, A_2, \dots, A_n) .
- Ciascun A_i è il nome di un ruolo giocato da qualche dominio D nello schema R .
- D è detto dominio di A_i : $D = \text{Dom}(A_i)$.
- Il **grado** di una relazione è il numero di attributi, n , del suo schema di relazione.

Schemi di relazione - *Esempio*

Nome della relazione: Student

STUDENT						
Name	SSN	HomePhone	Address	OfficePhone	Age	GPA

- Grado 7
- $\text{Dom}(\text{Name}) = \text{Names}$
- $\text{Dom}(\text{SSN}) = \text{social_Security_Numbers}$

Istanze di Relazione

- Una relazione (o istanza di relazione) r dello schema $R(A_1, A_2, \dots, A_n)$, denotata $r(R)$ è un insieme di tuple $r = \{t_1, t_2, \dots, t_n\}$.
- Ogni t_i è una lista ordinata di n valori $t = \langle v_1, v_2, \dots, v_n \rangle$ dove ciascun $v_i \in \text{dom}(A_i) \cup \{\text{null}\}$
 - Intensione della relazione $\rightarrow R$ (schema)
 - Estensione della relazione $\rightarrow r(R)$ istanza di relazione

Schemi e istanze di Relazione - *Esempio*

STUDENT

Name	SSN	HomePhone	Address	OfficePhone	Age	GPA
------	-----	-----------	---------	-------------	-----	-----

Schema di relazione

Benjamin Bayer	305-61-2435	373-1616	2918 BlueBonnet Lane	null	19	3.21
Katherine Ashly	381-62-1245	375-4409	125 Kirby Road	null	18	2.89
Dick Davidson	422-11-2420	null	3452 Elgin Road	749-1253	25	3.53
Charles Cooper	489-22-1100	376-9821	265 Lark Lane	749-1253	28	3.93

Istanza di relazione

Caratteristiche di una Relazione

- L'ordinamento delle tuple di una relazione non è parte della definizione.
La definizione non specifica alcun ordine.
- Una definizione alternativa di relazione considera non significativo anche l'ordine degli attributi;
In accordo a tale definizione una tupla può essere considerata come un insieme di coppie (**<attributo>**, **<valore>**).

Caratteristiche di una Relazione (2)

- Relazioni equivalenti, con diversi ordinamenti di righe e colonne:

STUDENT

Name	SSN	Home-Phone	Address	Office-Phone	Age	GPA
Benjamin Bayer	305-61-2435	373-1616	2918 BlueBonnet Lane	null	19	3.21
Katherine Ashly	381-62-1245	375-4409	125 Kirby Road	null	18	2.89
Dick Davidson	422-11-2420	null	3452 Elgin Road	749-1253	25	3.53
Charles Cooper	489-22-1100	376-9821	265 Lark Lane	749-1253	28	3.93

STUDENT

Name	Home-Phone	SSN	Age	Address	Office-Phone	GPA
Charles Cooper	376-9821	489-22-1100	28	265 Lark Lane	749-1253	3.93
Katherine Ashly	375-4409	381-62-1245	18	125 Kirby Road	null	2.89
Dick Davidson	null	422-11-2420	25	3452 Elgin Road	749-1253	3.53
Benjamin Bayer	373-1616	305-61-2435	19	2918 BlueBonnet Lane	null	3.21

Schema di database relazionale

- Uno **schema** di database relazionale è un insieme di schemi di relazione:

$$S_o\{R_1, R_2, \dots, R_n\}$$

- Una **istanza** di database relazionale DB di S è un insieme di istanze di relazione

$$DB = \{r_1, r_2, \dots, r_m\} \text{ tale che } r_i \text{ è una istanza di } R_i.$$

Schema di database relazionale (2)

EMPLOYEE

FNAME	MINIT	LNAME	<u>SSN</u>	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
-------	-------	-------	------------	-------	---------	-----	--------	----------	-----

DEPARTMENT

DNAME	<u>DNUMBER</u>	MGRSSN	MGRSTARTDATE
-------	----------------	--------	--------------

DEPT_LOCATION

<u>DNUMBER</u>	<u>DLOCATION</u>
----------------	------------------

PROJECT

PNAME	<u>PNUMBER</u>	PLOCATION	DNUM
-------	----------------	-----------	------

WORKS_ON

<u>ESSN</u>	<u>PNO</u>	HOURS
-------------	------------	-------

DEPENDENT

<u>ESSN</u>	<u>DEPARTMENT NAME</u>	SEX	BDATE	RELATIONSHIP
-------------	------------------------	-----	-------	--------------

Lo schema del database 'Company'

Istanza di database relazionale

EMPLOYEE	FNAME	MINIT	LNAME	<u>SSN</u>	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
	John		Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
	Franklin		Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
	Alicia		Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
	Jennifer		Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
	Ramesh		Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
	Joyce		English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
	Ahmad		Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
	James		Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	null	1

DEPT_LOCATIONS				<u>DNUMBER</u>	<u>DLOCATION</u>
					Houston
					Stafford
					Bellaire
					Sugarland

DEPARTMENT	DNAME	<u>DNUMBER</u>	<u>MGRSSN</u>	<u>MGRSTARTDATE</u>
	Research	5	333445555	1968-05-22
	Administration	4	987654321	1995-01-01
	Headquarters	1	888665555	1961-06-19

WORKS_ON	<u>ESSN</u>	<u>PNO</u>	HOURS
	123456789	1	32.5
	123456789	2	7.5
	666884444	3	40.0
	453453453	1	20.0
	453453453	2	20.0
	333445555	2	10.0
	333445555	3	10.0
	333445555	10	10.0
	333445555	20	10.0
	999887777	30	30.0
	999887777	10	10.0
	987987987	10	35.0
	987987987	30	5.0
	987654321	30	20.0
	987654321	20	15.0
	888665555	20	null

PROJECT	PNAME	<u>PNUMBER</u>	<u>PLOCATION</u>	<u>DNUM</u>
	ProductX	1	Bellaire	5
	ProductY	2	Sugarland	5
	ProductZ	3	Houston	5
	Computerization	10	Stafford	4
	Reorganization	20	Houston	1
	Newbenefits	30	Stafford	4

DEPENDENT	<u>ESSN</u>	<u>DEPENDENT_NAME</u>	SEX	BDATE	RELATIONSHIP
	333445555	Alice	F	1986-04-05	DAUGHTER
	333445555	Theodore	M	1983-10-25	SON
	333445555	Joy	F	1958-05-03	SPOUSE
	987654321	Abner	M	1942-02-28	SPOUSE
	123456789	Michael	M	1988-01-04	SON
	123456789	Alice	F	1988-12-30	DAUGHTER
	123456789	Elizabeth	F	1967-05-05	SPOUSE

Un'istanza del database "Company"

Notazioni del modello Relazionale

- Uno schema di relazione R , di grado n , è denotato da $R(A_1, A_2, \dots, A_n)$.
- Una tupla t in una relazione $r(R)$ è denotata da $t = \langle v_1, v_2, \dots, v_n \rangle$ dove v_i è il valore per l'attributo A_i
 - $t[A_i]$ si riferisce al valore v_i per l'attributo A_i ;
 - $t[A_u, A_w, \dots, A_z]$ dove A_u, A_w, \dots, A_z è una lista di attributi di R e riferisce alle sottotuple di valori v_u, v_w, \dots, v_z .
- Le lettere **Q, R, S** denotano nomi di relazioni.
- Le lettere **q, r, s** denotano stati di relazioni.
- Le lettere **t, u, v** denotano tuple.

Vincoli nel modello Relazionale



Vincoli del modello Relazionale

- Nel modello relazionale, i valori presenti in un'istanza di relazione devono soddisfare una serie di vincoli:
 1. Vincoli di dominio
 2. Vincoli di chiave
 3. Vincoli di integrità di entità
 4. Vincoli di integrità referenziale

Vincoli di dominio

- Il valore di ciascun attributo di A deve essere un valore atomico

{carattere, stringa a lunghezza fissa e variabile, data, ora, valuta, ecc...}

appartenente a $dom(A)$.

Vincoli di chiave – definizione di superchiave

- Una relazione è definita come un insieme di tuple. Per definizione tutti gli elementi di un insieme sono distinti, quindi tutte le tuple devono essere distinte.
- Devono allora esistere dei sottoinsiemi di attributi con la proprietà di non avere la stessa combinazione di valori in più tuple.
Sia sk un tale sottoinsieme di attributi di R :
$$t_1[sk] \neq t_2[sk]$$
- L'insieme di attributi sk è detto superchiave di R .

Vincoli di chiave – definizione di chiave

- Formalmente, una **chiave** k di uno schema di relazione R è una superchiave tale che, rimuovendo uno dei suoi attributi, non è più una superchiave.
 - k è detta anche **superchiave minimale**.
- Informalmente, una chiave k è un insieme di attributi minimale che permette di identificare univocamente una tupla.

Vincoli di chiave – definizione di chiave (2)

- In una relazione possono esistere più chiavi, dette **chiavi candidate**:
 - in tal caso se ne sceglie una, detta **chiave primaria**.
- Una chiave deve godere anche delle proprietà di **invarianza nel tempo**.
- **Esempio** di chiave:

Name	SSN	HomePhone	Address	OfficePhone	Age	GPA
Benjamin Bayer	305-61-2435	373-1616	2918 BlueBonnet Lane	null	19	3.21
Katherine Ashly	381-62-1245	375-4409	125 Kirby Road	null	18	2.89
Dick Davidson	422-11-2420	null	3452 Elgin Road	749-1253	25	3.53
Charles Cooper	489-22-1100	376-9821	265 Lark Lane	749-1253	28	3.93
Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	null	19	3.25

- $\{SSN\}$ è una chiave. Ogni insieme di attributi che include SSN è una superchiave.

Vincoli di chiave

- In una relazione R , non possono esistere valori duplicati per attributi chiave \mathbf{k} .

Vincoli di integrità di entità

- Nessun valore di chiave primaria può essere “*null*”.
- Questo perché:
 - Se ciò fosse permesso, non si avrebbe modo di identificare l'entità descritta nella tupla.
 - Non si vogliono memorizzare informazioni su entità non identificabili.

Vincoli di integrità referenziale

- Specificati tra due relazioni, sono usati per **mantenere consistenza** tra tuple delle due relazioni.
- *Informalmente*: una tupla di una relazione, che riferisce ad una tupla di un'altra relazione, deve riferire ad una tupla esistente.
 - *È il concetto portante del modello relazionale!*

Vincoli di integrità referenziale - Esempio

- L'attributo *DNO* di Employee deve riferire ad un *DNUMBER* esistente nella relazione Department.
- La relazione Employee è detta essere **relata** a Departement tramite l'attributo *DNO*.

EMPLOYEE	FNAME	MINIT	LNAME	<u>SSN</u>	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
	John	B	Smith	123456789	09-JAN-55	731 Fondren, Houston, TX	M	30000	333445555	5
	Franklin	T	Wong	333445555	08-DEC-45	638 Voss, Houston, TX	M	40000	888665555	5
	Alicia	J	Zelaya	999887777	19-JUL-58	3321 Castle, Spring, TX	F	25000	987654321	4
	Jennifer	S	Wallace	987654321	20-JUN-31	291 Berry, Bellaire, TX	F	43000	888665555	4
	Ramesh	K	Narayn	666884444	15-SEP-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
	Joyce	A	English	453453453	31-JUL-62	5631 Rice, Houston, TX	F	25000	333445555	5
	Ahmad	V	Jabbar	987987987	29-MAR-59	980 Dallas, Houston, TX	M	25000	987654321	4
	James	E	Borg	888665555	10-NOV-27	450 Stone, Houston, TX	M	55000	null	1

DEPARTMENT	DNAME	<u>DNUMBER</u>	MGRSSN	MGRSTARTDATE
	Research	5	333445555	22-MAY-78
	Administration	4	987654321	01-JAN-85
	Headquarters	1	888665555	19-JUN-71

Vincoli di integrità referenziale – definizione di chiave esterna

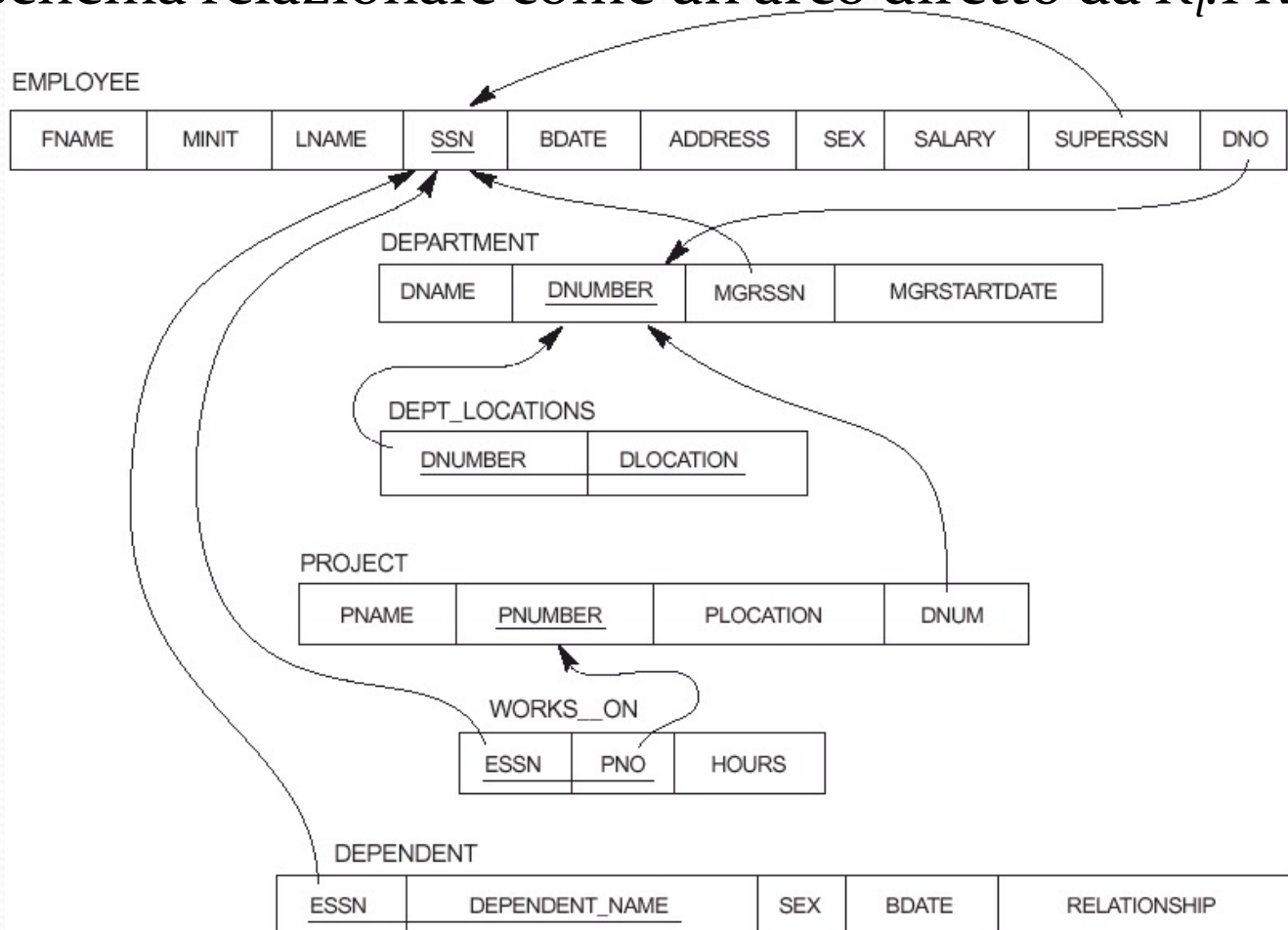
- *Formalmente:* un insieme di attributi FK in uno schema di relazione R_i è una **chiave esterna** se vale:
 1. gli attributi in FK hanno lo stesso dominio degli attributi della chiave primaria PK di un altro schema di relazione R_j (*gli attributi in FK riferiscono alla relazione R_j*)
 2. Un valore di FK in una tupla t_i di R_i o occorre come un valore di PK per qualche tupla t_j di R_j o è **null**.
 $t_i[FK] = t_j[PK]$ oppure $t_i[FK] = null$

Vincoli di integrità referenziale (2)

- Le chiavi esterne sono simili al concetto dei “*puntatori*” in C, che o riferiscono ad una variabile allocata o sono *null*.
- L'attributo *DNO* di Employee è una chiave esterna, poiché rispetta le condizioni appena elencate.
- Una tupla t_i di una relazione R_i è detta **referenziare** una tupla t_j di una relazione R_j se vale
$$t_i[FK] = t_j[PK].$$

Vincoli di integrità referenziale - *Esempio*

- Un vincolo di integrità referenziale può essere mostrato in uno schema relazionale come un arco diretto da $R_i.FK$ a R_j .



Vincoli e operazioni di aggiornamento su relazioni (Insert)

- Insert può violare tutti e quattro i tipi di vincoli:
 - Dominio
 - Un valore di un attributo può non apparire nel corrispondente dominio.
 - Chiave
 - Il valore della chiave nella nuova tupla già esistente nella relazione $r(R)$.
 - Integrità di entità
 - La chiave primaria è inserita a *null*.
 - Integrità referenziale
 - Il valore di una chiave esterna riferisce ad una tupla che non esiste nella relazione referenziata.

Vincoli e operazioni di aggiornamento su relazioni (Insert) (2)

- Come gestire la violazione?
 - Forzare l'inserimento completo (della relazione riferita);
 - Rifiutare l'inserimento.
- Nel primo caso la violazione può riguardare in cascata l'inserimento su altre relazioni.

Vincoli e operazioni di aggiornamento su relazioni (Delete)

- La delete può violare solo l'integrità referenziale.
- Gestione violazione:
 - Rigettare la cancellazione.
 - Tentare di propagare la cancellazione.
 - Modificare i valori dell'attributo referenziante (posto a *null*).
 - Combinazioni delle tre (ed il DBMS dovrebbe permettere all'utente di specificare la gestione).

Vincoli e operazioni di aggiornamento su relazioni (Modify)

- Nessun problema per attributi che non sono né chiave primaria né chiave esterna.
- Modifica chiave primaria:
 - Analogo a cancellare una tupla e inserirne un'altra.
- Modifica chiave esterna:
 - Il dbms deve verificare che riferisca ad una tupla esistente.