

Il modello Relazionale

c.vallati@iet.unipi.it

Sommario

- Il modello Relazionale per la progettazione logica
- Come passare dal modello ER al modello relazionale

Progettazione Logica

- Effettuata una progettazione *concettuale* della base di dati si passa alla progettazione *logica* della stessa
- A partire dalla descrizione dei dati che la base di dati dovrà memorizzare al suo interno, la progettazione logica definisce come questi dati verranno memorizzati usando le astrazioni fornite dalla tipologia di DBMS adottata
- *Non si deve conoscere il particolare tipo di DBMS che verrà usato ma si deve conoscerne la tipologia, in particolare il modello logico che questo adotterà*

Il Modello Relazionale

- Il modello Relazionale è il modello logico oggi più usato dai database commerciali comuni
- Tale modello si basa sul concetto matematico di **Relazione** (che va intesa in senso algebrico)
 - *Tale concetto di relazione non va confuso con il concetto di relazione tra i dati del modello concettuale*
- Ogni relazione è rappresentata graficamente sotto forma di **tabella**
- Un database viene rappresentato nel modello relazionale come una collezione di tabelle, a ciascuna delle quali è assegnato un nome unico

Costrutti fondamentali

- I costrutti fondamentali del modello relazionale sono i seguenti:
 - Relazione/Tabella
 - Attributo/Colonna
 - Dominio
 - Tupla/Riga

Relazione Matematica

Siano D_1, D_2, \dots, D_n (n insiemi anche non distinti)

Prodotto cartesiano $D_1 \times D_2 \times \dots \times D_n$:

- l'insieme di tutte le tuple ordinate (d_1, d_2, \dots, d_n) tali che $d_1 \in D_1, d_2 \in D_2, \dots, d_n \in D_n$

Relazione matematica R 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 ciascuno dei quali rappresenta.

Una relazione su n domini ha **grado** n .

il numero di tuple è la **cardinalità** della relazione.

Relazione - Esempio

$$D1 = \{0,1\}$$

$$D2 = \{2,3\}$$

Prodotto Cartesiano $D1 \times D2$

$D1 \times D2 =$

0	2
0	3
1	2
1	3

Relazione R $\subseteq D1 \times D2$

$R =$

0	2
1	3

- La rappresentazione grafica di una relazione avviene sotto forma tabellare
- Una tabella è un sottoinsieme del prodotto cartesiano fra i domini delle sue colonne
- I valori delle colonne di una tabella appartengono a dei **domini**

Schema di una relazione

- Nel modello logico relazionale la Relazione Matematica è adottata per rappresentare uno o più oggetti di una realtà di interesse
- Consiste in uno *Schema* e in una *Istanza*
- Lo Schema definisce **la struttura** di una relazione attraverso:
 - Il **nome della relazione** (il nome della tabella)
 - L'insieme dei **domini della relazione**
 - I **nomi assegnati a ciascun dominio** (detti attributi) che descrivono il ruolo giocato dal dominio stesso

Istanza di una relazione

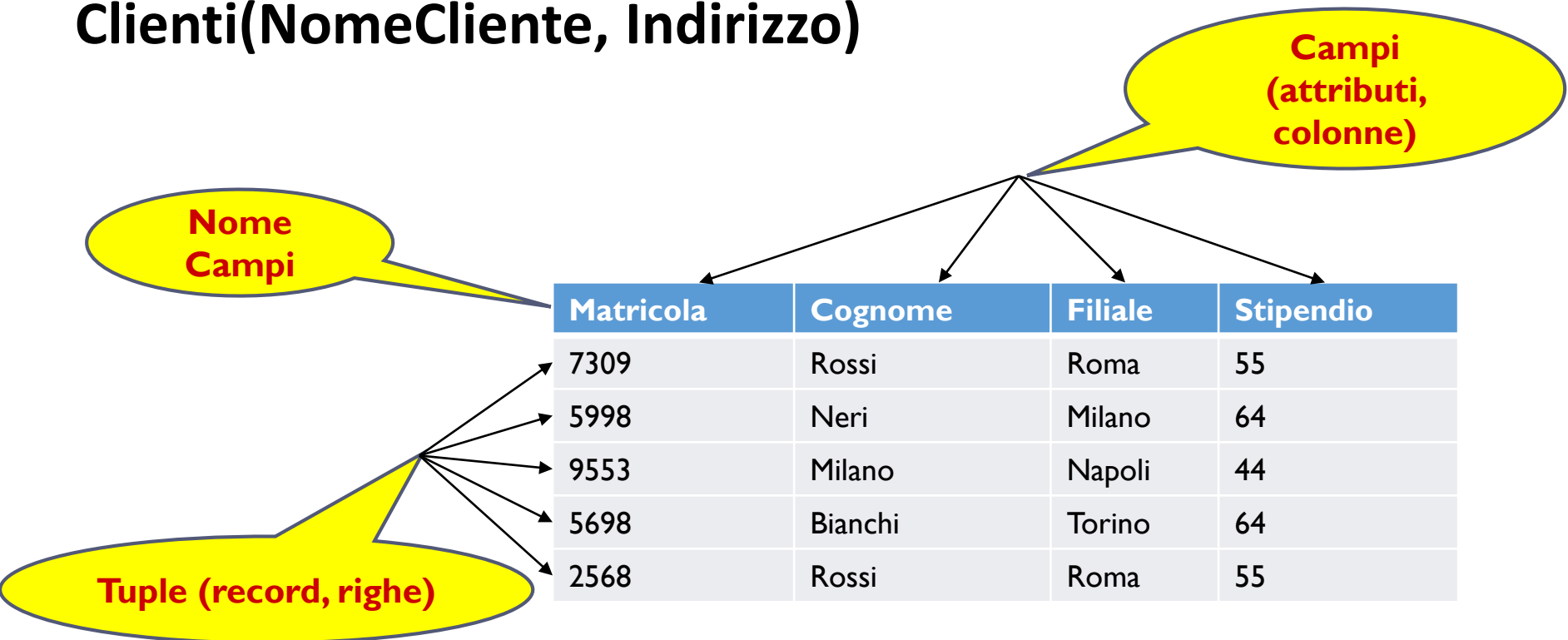
- *L'istanza di una relazione è una tabella riempita secondo il suo schema*
- Le righe di una tabella si chiamano *tuple* o record
- Le colonne della tabella corrispondono agli *attributi*
- *Il contenuto di un'istanza di una relazione NON dipendono dall'ordinamento delle tuple stesse*

Rappresentazione di uno schema

Per rappresentare in **forma compatta** uno schema si utilizza la seguente notazione:

Impiegati (Matricola, Cognome, Filiale, Stipendio)

Clienti(NomeCliente, Indirizzo)



Esempio

Un'istanza di basi di dati e un insieme di istanza di relazioni

Studenti

Matricola	Cognome	Nome	Data di Nascita
276545	Rossi	Marco	NULL
253065	Rossi	Luca	23/04/1972
243064	Di Martino	Gennaro	12/02/1972

Esami

Studente	Voto	Corso
276545	25	01
253065	26	03
243064	30	05

Corsi

Codice	Titolo	Docente
01	Analisi	Giani
03	Chimica	Fantoni
05	Chimica	Baggio

Informazioni incomplete

- L'informazione all'interno di una tabella devono essere rappresentate in maniera omogenea attenendosi strettamente al dominio scelto per ciascun attributo
- In alcune casistiche si ha però il bisogno di rappresentare tuple incomplete, in cui alcuni attributi non hanno un valore specificato
- A tal proposito è introdotto un particolare valore per indicare l'assenza di informazione associato a tale attributo
- Tale valore, detto valore NULL, si aggiunge a quelli di dominio

Vincoli di Integrità

- Le strutture del modello permettono di definire come i dati verranno organizzati
- In molti casi però non è vero che qualsiasi insieme di tuple rappresentano informazioni corrette
- Al fine di evitare l'inserimento di tuple composte da informazioni non corrette il modello relazionale permette di definire dei vincoli di integrità sui dati di una relazione, cioè delle *proprietà che devono essere soddisfatte dalle istanze*
- Un vincolo può essere visto come un predicato che associa ad ogni istanza un valore vero o falso in modo da verificare che ciascuna istanza sia corretta e ammissibile

Esempio

Un'istanza di basi di dati con istanze non corrette

Studenti

Matricola	Cognome	Nome	Data di Nascita
276545	NULL	Marco	NULL
243064	Rossi	Luca	23/04/1972
243064	Di Martino	Gennaro	12/02/1972

Esami

Studente	Voto	Corso
276545	25	01
200010	26	03
243064	36	05

Corsi

Codice	Titolo	Docente
01	Analisi	Giani
03	Chimica	Fantoni
05	Chimica	Baggio

Vincoli di integrità

- Esistono diverse categorie dei vincoli di integrità. Questi possono essere distinti in due categorie a seconda degli elementi di una base di dati coinvolti:
 - **Vincolo Intrarelazionale:** sono vincoli il cui soddisfacimento è definito rispetto alle singole relazioni della base di dati, questi possono riguardare valori di una singola tupla o di più tuple appartenenti alla stessa relazione
 - **Vincolo di Tupla:** è un vincolo che può essere valutato su ciascuna tupla indipendentemente dalle altre
 - **Vincolo di Valore o di Dominio:** è un vincolo che impone una restrizione sul dominio di un singolo attributo
 - **Vincolo Interrelazionale:** se il vincolo coinvolge più relazioni

Vincolo di Tupla

- Esprimono condizioni sui valori di ciascuna tupla indipendentemente dalle altre
- Possono essere espressi tramite espressioni booleane e aritmetiche

Esempio:

nella tabella Esami(STUDENTE,MATERIA,DATA,VOTO,LODE)

la LODE può essere assegnata solo se il VOTO è 30

Vincolo: $(LODE = L) \text{ AND } (VOTO = 30)$

nella tabella Pagamenti(DATA,IMPORTO,RITENUTE,NETTO)

Il netto è pari alla differenza fra l'importo e le ritenute

Vincolo: $NETTO = IMPORTO - RITENUTE$

Vincolo di Dominio

E' un **Vincolo di base**:

- va specificato per ogni attributo
- riguarda il suo dominio, ovvero che tipo di dati “ha senso” per quell’attributo
- i domini che si possono usare dipendono dal DBMS (Domini predefiniti)
- è possibile definire nuovi Domini

Vincolo NOT NULL

- Impone alle tuple di non avere valori nulli su uno o più attributi

Chiavi

- Le tuple di una tabella hanno bisogno di un meccanismo per essere identificate univocamente
- A tal fine per ogni tabella viene definita una **Chiave**
- Una Chiave è un insieme di attributi utilizzato per *identificare univocamente* le tuple di una relazione
- Un insieme degli attributi è definito **superchiave** di una relazione se non esistono due tuple distinte con gli stessi valori degli attributi nella tabella
 - Tutti gli attributi di una tabella possono essere considerati super-chiave in quanto non sono inserite due tuple con valori identici degli attributi

Vincolo di chiave

- Una superchiave è **chiave** se essa è una *superchiave minimale*, cioè non esiste un'altra superchiave composta da un sottoinsieme dei suoi attributi che sia ancora superchiave
- Quando tale proprietà è imposta come vincolo alle tuple di uno schema di relazione, tale vincolo è detto **vincolo di chiave**
- Tra tutte le chiavi di una relazione solitamente viene selezionata la **chiave primaria** che è una chiave su cui si mette un vincolo NOT NULL al fine di poter garantire l'identificazione univoca di tutte le tuple, tale vincolo è detto **vincolo di chiave primaria**

Esempio - Chiave

- L'insieme {Matricola, Nome, Cognome} è una superchiave
- L'insieme {Matricola} è una superchiave. Essa è anche una superchiave minimale che può essere candidata a chiave primaria in quanto la Matricola difficilmente sarà sconosciuta
- L'insieme {Nome, Corso} NON è una superchiave
- L'insieme {Cognome, Corso} è una superchiave minimale, **però lo è solo in relazione a questa specifica istanza di tuple**, in generale è possibile avere due persone con lo stesso cognome che frequentano lo stesso corso di studi

Studenti

Matricola	Cognome	Nome	Data di Nascita	Corso
276545	Rossi	Marco	NULL	Ing. Gestionale
253065	Rossi	Luca	23/04/1972	Ing. Informatica
243064	Di Martino	Luca	12/02/1972	Ing. Informatica

Riferimenti tra relazioni

Una base di dati solitamente è composta da diverse relazioni in cui le varie tuple sono collegate tra di loro attraverso riferimenti. In questo caso la relazione Studenti e Esami sono collegate attraverso gli attributi Studente e Matricola, in particolare le tuple della tabella esami hanno come valore dell'attributo Studente valori presenti nelle tuple della tabella Studenti nell'attributo Matricola. Essendo questi riferimenti validi (non esiste nessuna tupla della tabella Esami che abbia un campo studente con un valore non presente nel campo Matricola in una delle tuple della tabella Studenti) questi sono effettivamente utilizzabili

Studenti

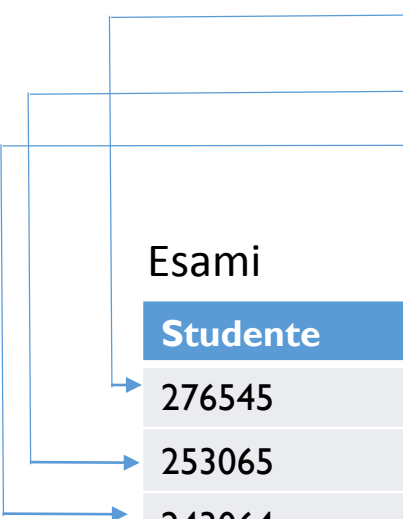
Matricola	Cognome	Nome	Data di Nascita
276545	Rossi	Marco	NULL
253065	Rossi	Luca	23/04/1972
243064	Di Martino	Gennaro	12/02/1972

Esami

Studente	Voto	Corso
276545	25	01
253065	26	03
243064	30	05

Corsi

Codice	Titolo	Docente
01	Analisi	Giani
03	Chimica	Fantoni
05	Chimica	Baggio



Vincoli di integrità referenziale

- Un vincolo di integrità referenziale o foreign key è un vincolo interrelazionale che permette di imporre che tali riferimenti siano in effetti utilizzabili. In particolare:

Un vincolo di **integrità referenziale** (detto anche vincolo di (“**foreign key**”)) fra un insieme di attributi X di una relazione $R1$ e un'altra relazione $R2$ impone ai valori su X di ciascuna ennupla dell'istanza di $R1$ di comparire come valori della chiave (primaria) dell'istanza di $R2$.

Conversione modello E-R in Modello Relazionale

- Il Modello Relazionale che rappresenta la realtà di interesse può essere ricavato direttamente dal Modello ER attraverso una sequenza di operazioni di conversione.

Entità

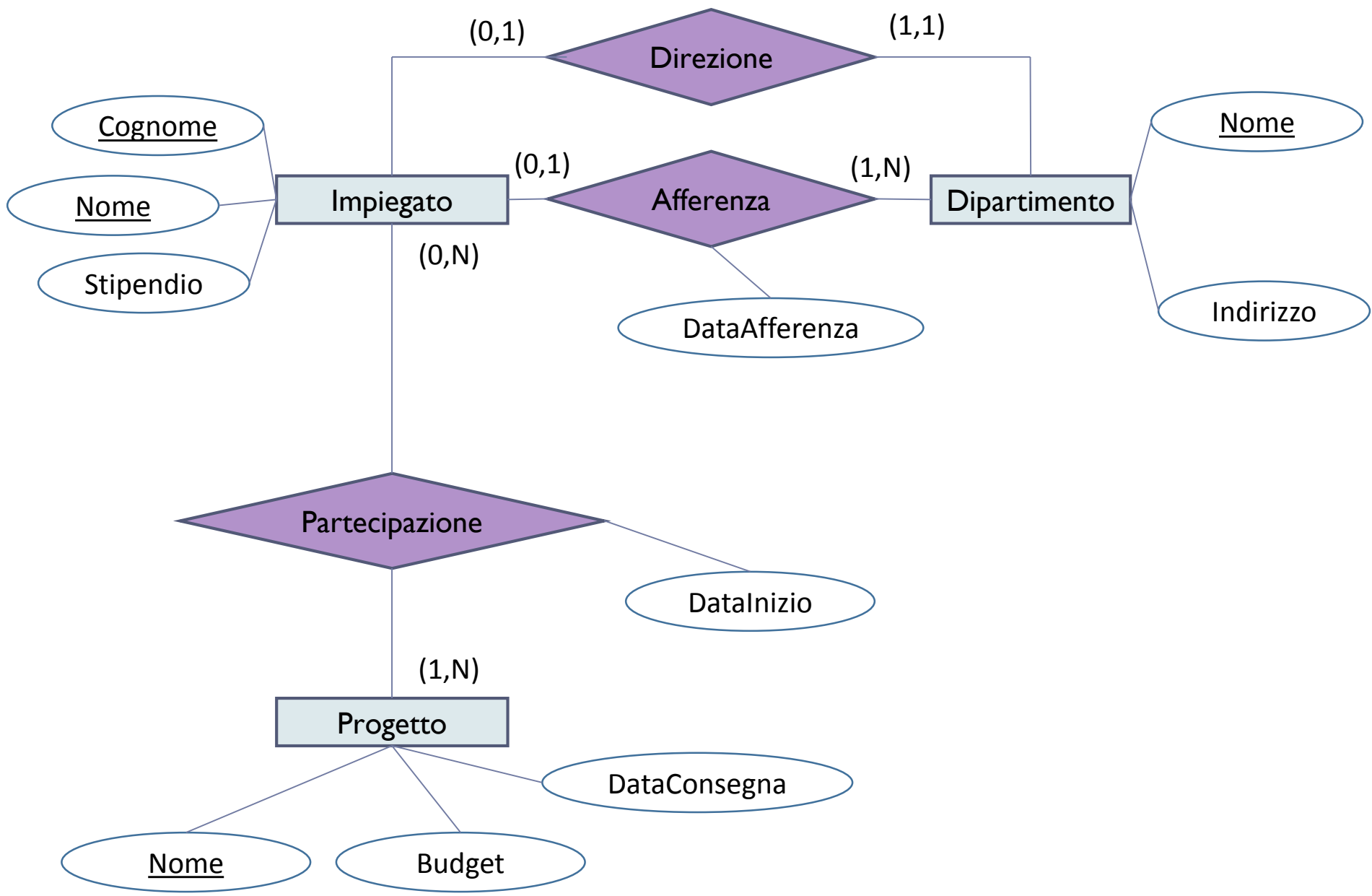
- Una relazione (tabella) per ciascuna entità del Modello ER
- utilizzando per i campi gli stessi attributi semplici dell'entità di provenienza
- chiave primaria scelta tra le chiavi candidate

Relazioni

- Ad ogni relazione del modello E-R corrisponde una tabella nel modello relazionale, cui è assegnato il nome della corrispondente relazione del modello E-R

Per le relazioni si deve prima di tutto esaminare la **cardinalità**:

- Le **relazioni M:N** devono essere tradotte in tabelle che hanno come attributi gli eventuali attributi della relazione e (da non dimenticare!) le chiavi di entrambe le entità collegate dalla relazione.
- Le **relazioni 1:N** si possono conglobare nella tabella relativa all'entità dalla parte dell' N. La chiave della tabella dalla parte dell' 1 viene inglobata nella tabella dalla parte dell' N.
- Le **relazioni 1:1** si possono conglobare come le 1:N, a scelta nell'una o nell'altra tabella.



Relazione Partecipazione M:N che lega Impiegato e Progetto

Impiegato

Cognome	Nome	Stipendio
Rossi	Marco	60000
Rossi	Luca	50000
Di Martino	Luca	20000

Progetto

Nome	Budget	DataConsegna
Progetto I	10000	10/10/2016
Progettoll	20000	02/10/2015
Progettolll	1000	02/03/2017

Partecipazione

Cognome	Nome	NomeProgetto
Rossi	Marco	Progettoll
Rossi	Luca	Progettoll
Rossi	Marco	Progettolll

Relazione Afferenza 1:N che lega Impiegato e Dipartimento

Dipartimento

Nome	Budget
Aereonautica	1000000
Informatica	2000000
Chimica	100000

Impiegato

Cognome	Nome	Stipendio	NomeDip	DataAfferenza
Rossi	Marco	60000	Aereonautica	10/02/1995
Rossi	Luca	50000	Chimica	10/05/1996
Di Martino	Luca	20000	Aereonautica	22/02/1998

Entità deboli

- Per ogni entità debole bisogna creare una relazione R ereditandone gli attributi.
- Bisogna inoltre includere, come chiave esterna, la chiave primaria dell'entità "proprietaria".
- La **chiave primaria** di R sarà l'insieme della chiave esterna più la chiave parziale.

Gerarchie

- Il modello relazionale è “piatto” e non può rappresentare le gerarchie che sono sostituite da entità e associazioni.
- Viene creata una tabella per ogni entità coinvolta nella specializzazione.
- Le tabelle delle entità figlie ereditano dall'entità genitore la chiave.

Ridondanze e anomalie

- Il progetto logico di una base di dati basato su uno schema concettuale in cui le entità e associazioni reali sono ben separate in maniera corretta non dovrebbe contenere ridondanze o anomalie
- In ogni modo lo schema logico prodotto può essere verificato a posteriori per verificare la presenza di ridondanze e anomalie che possono essere eliminate o corrette

Esempio di schema con ridondanze e anomalie

<u>Impiegato</u>	<u>Stipendio</u>	<u>Progetto</u>	<u>Bilancio</u>	<u>Funzione</u>
Rossi	20000	Marte	2000	Tecnico
Verdi	35000	Giove	15000	progettista
Verdi	35000	Venere	15000	Progettista
Neri	55000	Venere	15000	Direttore
Neri	55000	Giove	15000	Consulente
Neri	55000	Marte	2000	Consulente
Mori	48000	Marte	2000	Direttore
Mori	48000	Venere	15000	Progettista
Bianchi	48000	Venere	15000	Progettista
Bianchi	48000	Giove	15000	Direttore

- Lo stipendio di ciascun impiegato è unico e funzione sola dell'impiegato stesso, indipendentemente dai progetti
- Il bilancio di ciascun progetto è unico e dipende solo dal progetto indipendentemente dagli impiegati

Anomalie

- Tale struttura particolare presenta le seguenti anomalie:
 - **Ridondanza:** lo stipendio di ciascun impiegato è replicato per ogni progetto a cui partecipa
 - **Anomalia di aggiornamento:** se lo stipendio di un impiegato cambia tutte le tuple corrispondenti devono essere aggiornate
 - **Anomalia di cancellazione:** se un impiegato interrompe la partecipazione a tutti i progetti non si può conservare traccia del suo nome e del suo stipendio
 - **Anomalia di inserimento:** analogamente un nuovo impiegato non può essere inserito fino a quando questo non partecipa ad almeno un progetto

Dipendenze funzionali

- Al fine di verificare la presenza di anomalie in uno schema logico si definisce un ulteriore vincolo di integrità tra gli attributi di una stessa relazione, la dipendenza funzionale
- Dati due attributi Y e Z di uno schema R(X) (o due sottoinsiemi Y e Z degli attributi X di R), si dice che esiste su R una dipendenza funzionale tra Y e Z, se per ogni coppia di tuple di R aventi gli stessi valori degli attributi di Y, risulta che hanno gli stessi valori anche sugli attributi Z
- In altre parole esiste un vincolo di dipendenza funzionale tra Y e Z in R se in ogni tupla il valore di Y determina il valore di Z (e.g. Impiegato determina Stipendio, Progetto determina Bilancio)
- Le dipendenze funzionali sono indicate con $Y \rightarrow Z$ in R

Dipendenze funzionali non banali

- Su una relazione possono essere verificate dipendenze funzionali banali, in particolare quando alcuni attributi di Z compaiono in Y (e.g. Impiegato, Progetto \rightarrow Progetto)
- Si dicono dipendenze funzionali non banali quelle in cui nessun attributo di Z compare tra gli attributi di Y
- Il concetto di dipendenza funzionale generalizza quello di chiave in quanto tra la chiave e gli altri attributi sussiste una dipendenza funzionale

Forma Normale

- Data uno schema logico di una relazione, la verifica della sua struttura rispetto alle sue dipendenze funzionali da indicazione della presenza di anomalie e ridondanze
- A tal fine viene definita una particolare forma dello schema logico di una base di dati, la forma normale
- *Una relazione R è in forma normale se per ogni dipendenza funzionale non banale $Y \rightarrow Z$ su di essa, Y contiene una chiave K di R , cioè y è superchiave per R*
- Se una relazione non soddisfa la forma normale, può essere applicata una decomposizione in forma normale, al fine di normalizzarla

Esempio Normalizzazione

Dipendenze Formali:

Impiegato, Progetto → Funzione

Impiegato → Stipendio

Progetto → Bilancio

<u>Impiegato</u>	<u>Stipendio</u>
Rossi	20000
Verdi	35000
Neri	55000
Mori	48000
Bianchi	48000

<u>Impiegato</u>	<u>Progetto</u>	<u>Funzione</u>
Rossi	Marte	Tecnico
Verdi	Giove	progettista
Verdi	Venere	Progettista
Neri	Venere	Direttore
Neri	Giove	Consulente
Neri	Marte	Consulente
Mori	Marte	Direttore
Mori	Venere	Progettista
Bianchi	Venere	Progettista
Bianchi	Giove	Direttore

<u>Progetto</u>	<u>Bilancio</u>
Marte	2000
Giove	15000
Venere	15000