

## Facoltà di Scienze Matematiche, Fisiche e Naturali

Laurea in Bioinformatica  
Basi di Dati  
Anno Accademico 2008/2009

## Progettazione logica



## Progettazione logica

- o La progettazione logica ha come obiettivo la costruzione di uno schema logico in grado di descrivere, in maniera corretta ed efficiente, tutte le informazioni contenute nello schema concettuale.

### Schema concettuale



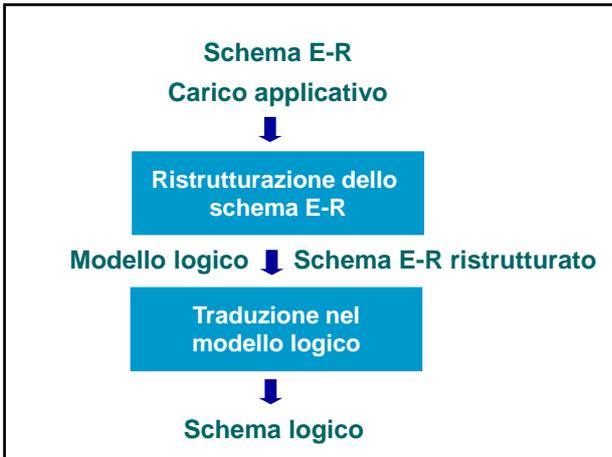
### Schema logico

## Obiettivo della progettazione logica

- o "tradurre" lo schema concettuale in uno schema logico che rappresenti gli stessi dati in maniera corretta ed efficiente.
- o Non si tratta di una pura e semplice traduzione:
  - alcuni aspetti non sono direttamente rappresentabili
  - è necessario considerare le prestazioni
- o Lo schema E-R va ristrutturato

## Progettazione logica

- o **Ristrutturazione dello schema E-R**: fase indipendente dal modello logico scelto. Si basa su:
  - criteri di ottimizzazione dello schema
  - semplificazione della fase successiva
- o **Traduzione verso il modello logico**: fa riferimento a uno specifico modello logico (modello relazionale) e può includere una ulteriore ottimizzazione che si basa sulle caratteristiche del modello stesso.



### Progettazione logica

---

- Ingresso:
  - schema concettuale
  - informazioni sul carico applicativo
  - modello logico
  
- Uscita:
  - schema logico
  - documentazione associata

### Ristrutturazione schema E-R

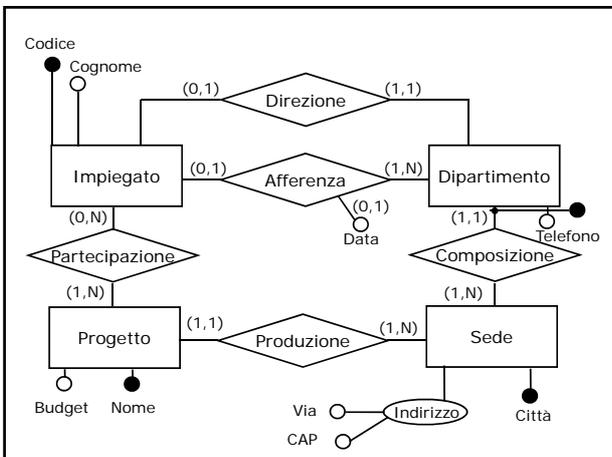
---

- Motivazioni:
  - semplificare la traduzione
  - "ottimizzare" le prestazioni
  
- Osservazione:
  - uno schema E-R ristrutturato non è (più) uno schema concettuale nel senso stretto del termine

### Ottimizzazione prestazioni

---

- Per ottimizzare il risultato abbiamo bisogno di analizzare le prestazioni a questo livello.
- Ma le prestazioni non sono valutabili con precisione su uno schema concettuale!
- Consideriamo "indicatori" dei parametri che regolano le prestazioni:
  - spazio: numero di occorrenze previste
  - tempo: numero di occorrenze (di entità e relazioni) visitate durante una operazione



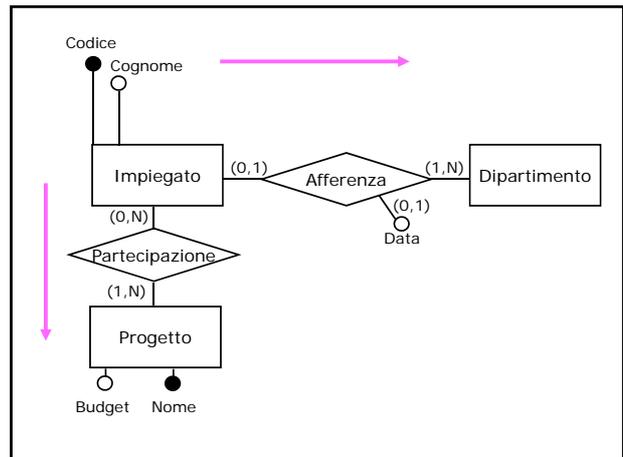
### Tavola dei volumi

---

Concetto	Tipo	Volume
Sede	E	10
Dipartimento	E	80
Impiegato	E	2000
Progetto	E	500
Composizione	R	80
Afferenza	R	1900
Direzione	R	80
Partecipazione	R	6000

## Valutazione costo: Esempio

- o Operazione:
  - trova tutti i dati di un impiegato, del dipartimento nel quale lavora e dei progetti ai quali partecipa.
- o Si costruisce una **tavola degli accessi** basata su uno **schema di navigazione**.



## Tavola degli accessi

- o Accediamo ad una occorrenza di Impiegato per poi accedere ad una occorrenza di Dipartimento attraverso Afferenza (un Impiegato afferisce ad un solo Dipartimento).
- o Accediamo a tre occorrenze di Progetto attraverso tre occorrenze di Partecipazione (un impiegato partecipa in media a tre progetti).

Concetto	Costrutto	Accesso	Tipo
Impiegato	E	1	L
Afferenza	R	1	L
Dipartimento	E	1	L
Partecipazione	R	3	L
Progetto	E	3	L

## Ristrutturazione schema E-R

- o **Analisi delle ridondanze**: si decide se eliminare o mantenere eventuali ridondanze presenti nello schema.
- o **Eliminazione delle generalizzazioni**: tutte le generalizzazioni presenti nello schema vengono analizzate e sostituite da altri costrutti.
- o **Partizionamento/accorpamento di entità e relazioni**: si decide se partizionare concetti dello schema in più concetti (entità e/o relazioni) o accorpate concetti separati in un unico concetto.
- o **Scelta degli identificatori primari**: si seleziona un identificatore per quelle entità che ne hanno più di uno.

## Ristrutturazione schema E-R

- o **Analisi delle ridondanze**: si decide se eliminare o mantenere eventuali ridondanze presenti nello schema.
- o **Eliminazione delle generalizzazioni**: tutte le generalizzazioni presenti nello schema vengono analizzate e sostituite da altri costrutti.
- o **Partizionamento/accorpamento di entità e relazioni**: si decide se partizionare concetti dello schema in più concetti (entità e/o relazioni) o accorpate concetti separati in un unico concetto.
- o **Scelta degli identificatori primari**: si seleziona un identificatore per quelle entità che ne hanno più di uno.

## Analisi delle ridondanze

- o Una ridondanza in uno schema E-R è una informazione significativa ma derivabile da altre.
- o **In questa fase si decide se eliminare le ridondanze eventualmente presenti o mantenerle.**

## Ridondanze

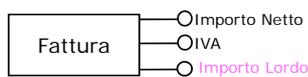
- **Vantaggi**
  - semplificazione delle interrogazioni
- **Svantaggi**
  - appesantimento degli aggiornamenti
  - maggiore occupazione di spazio

## Forme di ridondanza in uno schema E-R

- **Attributi derivabili:**
  - da altri attributi della stessa entità (o relazione)
  - da attributi di altre entità (o relazioni)
  - da operazioni di conteggio di occorrenze
- **Relazioni derivabili dalla composizione di altre relazioni in presenza di cicli.**

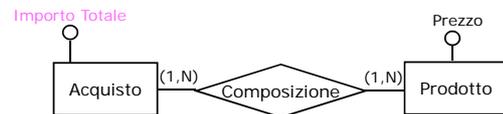
## Attributo derivabile

- L'**Importo Lordo** è derivabile dalla somma di **Importo Netto** e **IVA**.



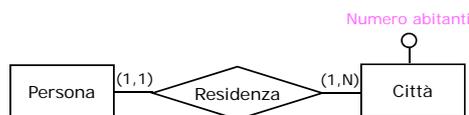
## Attributo derivabile da altra entità

- L'**Importo Totale** è derivabile, attraverso la relazione **Composizione**, dall'attributo **Prezzo** dell'entità **Prodotto**, sommando i prezzi dei prodotti di cui un acquisto è composto.



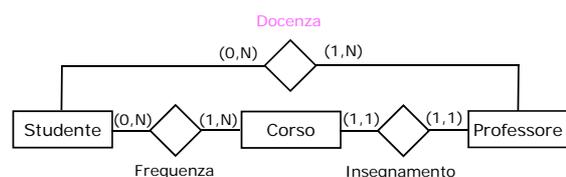
## Attributo derivabile da operazioni di conteggio

- Il **Numero abitanti** di una **Città** è derivabile contando le occorrenze della relazione **Residenza** a cui la **Città** partecipa.



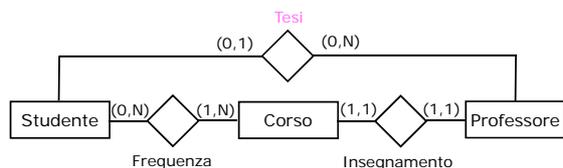
## Ridondanza dovuta a ciclo

- La relazione **Docenza** tra **Studenti** e **Professori** è derivabile dalle relazioni **Frequenza** e **Insegnamento**.



## La presenza di cicli non genera necessariamente Ridondanza

- La relazione **Tesi** tra **Studenti** e **Professori** NON è derivabile dalle relazioni **Frequenza** e **Insegnamento**.



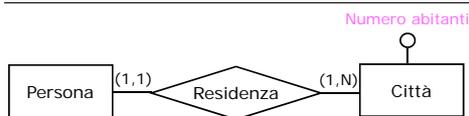
## Analisi delle ridondanze

- Una ridondanza in uno schema concettuale corrisponde alla presenza di un dato che può essere derivato da altri dati.
- Varie forme di ridondanza:
  - Attributi derivabili, occorrenza per occorrenza, da altri attributi della stessa entità (relazione).
  - Attributi derivabili da attributi di altre entità (relazione).
  - Attributi derivabili da operazioni di conteggio di occorrenza.
  - Relazioni derivabili dalla composizione di altre relazioni in presenza di cicli.

## Dato derivato in uno schema E-R

- La presenza di un dato derivato in uno schema E-R ha:
  - Vantaggio**
    - Riduzione degli accessi necessari per calcolare il dato derivato
  - Svantaggi**
    - Maggiore occupazione di memoria
    - Necessità di effettuare operazioni aggiuntive per mantenere il dato derivato aggiornato
- La decisione di mantenere o eliminare una ridondanza va presa confrontando:
  - Costo di esecuzione delle operazioni
  - Occupazione di memoria
 nel caso di presenza o assenza della ridondanza.

## Analisi di una ridondanza: Esempio



Concetto	Tipo	Volume
Città	E	200
Persona	E	1000000
Residenza	R	1000000

## Analisi di una ridondanza: Esempio

- Operazione 1:** memorizza una nuova persona con la relativa città di residenza. (500 volte al giorno)
- Operazione 2:** stampa tutti i dati di una città (incluso il numero di abitanti). (2 volte al giorno)

## Analisi di una ridondanza: Esempio

Presenza di ridondanza

- Operazione 1**

Concetto	Costrutto	Accesso	Tipo
Persona	E	1	S
Residenza	R	1	S
Città	E	1	L
Città	E	1	S

- Operazione 2**

Concetto	Costrutto	Accesso	Tipo
Città	E	1	L

### Analisi di una ridondanza: Esempio

- **Presenza di ridondanza: costi**
  - **Operazione 1:** 1500 accessi in scrittura e 500 accessi in lettura al giorno
  - **Operazione 2:** trascurabile.
- Contiamo doppi gli accessi in scrittura
  - Totale di 3500 accessi al giorno

### Analisi di una ridondanza: Esempio

#### Assenza di ridondanza

- **Operazione 1**

Concetto	Costrutto	Accesso	Tipo
Persona	E	1	S
Residenza	R	1	S

- **Operazione 2**

Concetto	Costrutto	Accesso	Tipo
Città	E	1	L
Residenza	R	5000	L

### Analisi di una ridondanza: Esempio

- **Assenza di ridondanza: costi**
  - **Operazione 1:** 1000 accessi in scrittura
  - **Operazione 2:** 10000 accessi in lettura al giorno
- Contiamo doppi gli accessi in scrittura
  - Totale di 12000 accessi al giorno

### Ristrutturazione schema E-R

- **Analisi delle ridondanze:** si decide se eliminare o mantenere eventuali ridondanze presenti nello schema.
- **Eliminazione delle generalizzazioni:** tutte le generalizzazioni presenti nello schema vengono analizzate e sostituite da altri costrutti.
- **Partizionamento/accorpamento di entità e relazioni:** si decide se partizionare concetti dello schema in più concetti (entità e/o relazioni) o accorpere concetti separati in un unico concetto.
- **Scelta degli identificatori primari:** si seleziona un identificatore per quelle entità che ne hanno più di uno.

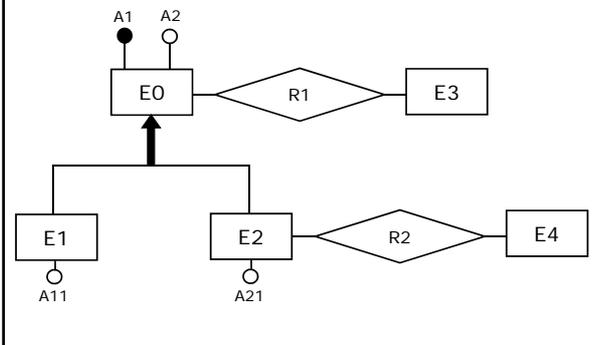
### Eliminazione delle gerarchie

- Il modello relazionale non può rappresentare direttamente le generalizzazioni.
- Entità e relazioni sono invece direttamente rappresentabili.
- Si eliminano perciò le gerarchie, sostituendole con entità e relazioni

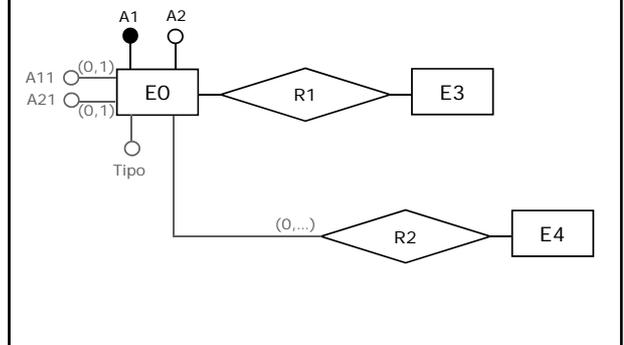
### Rappresentazione di gerarchie mediante entità e relazioni

- Accorpamento delle figlie della generalizzazione nel padre.
- Accorpamento del padre della generalizzazione nelle figlie.
- Sostituzione della generalizzazione con associazioni.

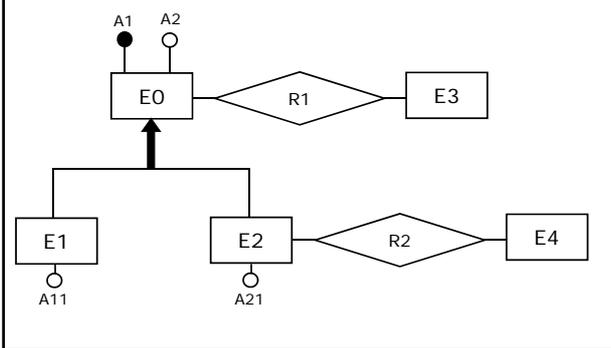
Accorpamento delle figlie della generalizzazione nel padre



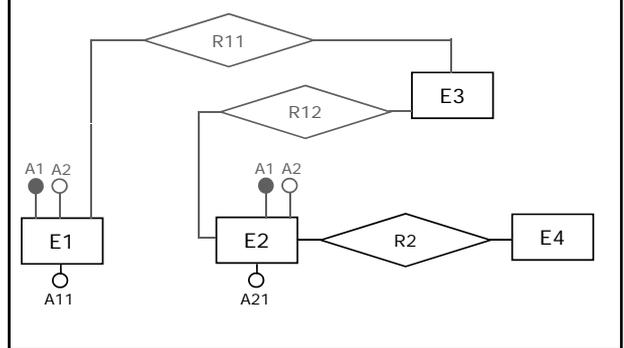
Accorpamento delle figlie della generalizzazione nel padre



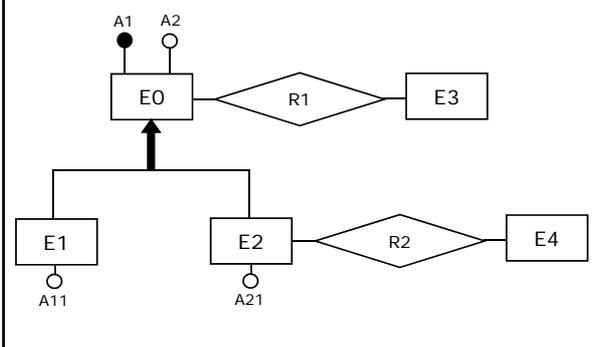
Accorpamento del padre della generalizzazione nelle figlie



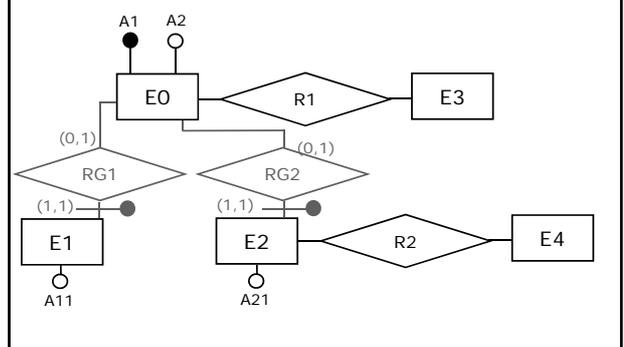
Accorpamento del padre della generalizzazione nelle figlie



Sostituzione della generalizzazione con associazioni



Sostituzione della generalizzazione con associazioni



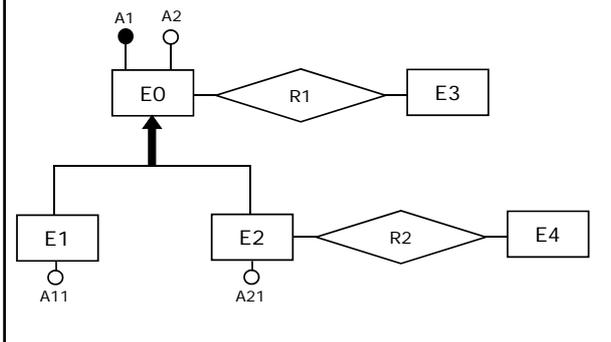
## Scelta alternativa

- La scelta fra le alternative si può fare con metodo simile a quello visto per l'analisi delle ridondanze
  - però non basato solo sul numero degli accessi.
- È possibile seguire alcune semplici regole generali.

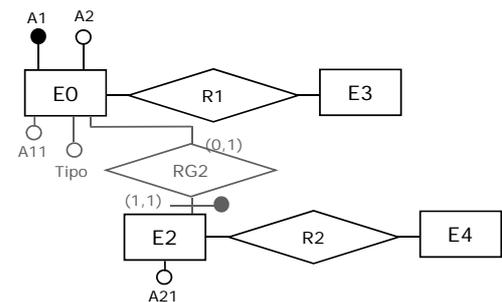
## Scelta alternativa

- Accorpamento delle figlie della generalizzazione nel padre.
  - Conviene se gli accessi al padre e alle figlie sono contestuali.
- Accorpamento del padre della generalizzazione nelle figlie.
  - Conviene se gli accessi alle figlie sono distinti.
- Sostituzione della generalizzazione con associazioni.
  - Conviene se gli accessi alle entità figlie sono separati dagli accessi al padre.
- Sono anche possibili soluzioni "ibride", soprattutto in gerarchie a più livelli.

## Soluzione ibrida: Esempio



## Soluzione ibrida: Esempio



## Ristrutturazione schema E-R

- Analisi delle ridondanze: si decide se eliminare o mantenere eventuali ridondanze presenti nello schema.
- Eliminazione delle generalizzazioni: tutte le generalizzazioni presenti nello schema vengono analizzate e sostituite da altri costrutti.
- **Partizionamento/accorpamento di entità e relazioni:** si decide se partizionare concetti dello schema in più concetti (entità e/o relazioni) o accorpate concetti separati in un unico concetto.
- Scelta degli identificatori primari: si seleziona un identificatore per quelle entità che ne hanno più di uno.

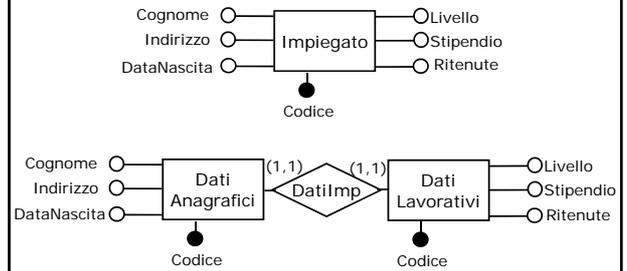
## Partizionamento/Accorpamento

- Ristrutturazioni effettuate per rendere più efficienti le operazioni in base a un semplice principio.
- Gli accessi si riducono:
  - separando attributi di un concetto che vengono acceduti separatamente
  - raggruppando attributi di concetti diversi acceduti insieme

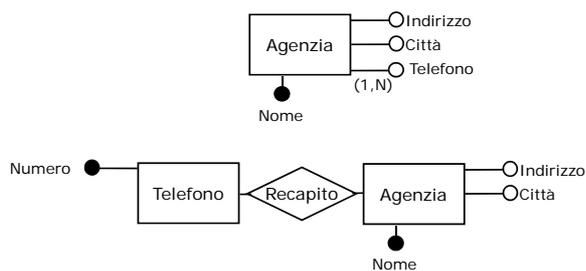
## Casi principali

- Partizionamento verticale di entità.
- Partizionamento orizzontale di relazioni.
- Eliminazione di attributi multivalore.
- Accorpamento di entità o relazioni.

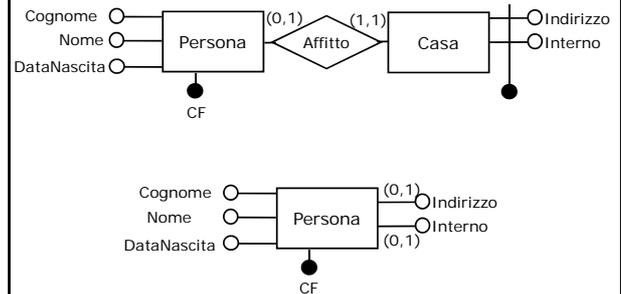
## Decomposizione verticale di entità



## Eliminazione attributi multivalore



## Accorpamento di entità



## Ristrutturazione schema E-R

- Analisi delle ridondanze: si decide se eliminare o mantenere eventuali ridondanze presenti nello schema.
- Eliminazione delle generalizzazioni: tutte le generalizzazioni presenti nello schema vengono analizzate e sostituite da altri costrutti.
- Partizionamento/accorpamento di entità e relazioni: si decide se partizionare concetti dello schema in più concetti (entità e/o relazioni) o accorpate concetti separati in un unico concetto.
- Scelta degli identificatori primari: si seleziona un identificatore per quelle entità che ne hanno più di uno.

## Scelta degli identificatori primari

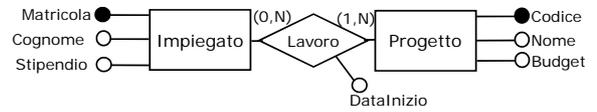
- Operazione indispensabile per la traduzione nel modello relazionale.
- Criteri
  - assenza di opzionalità
  - semplicità
  - utilizzo nelle operazioni più frequenti o importanti

## Traduzione verso il modello relazionale

### o Idea di base:

- Le entità diventano relazioni sugli stessi attributi.
- Le associazioni (ovvero le relazioni E-R) diventano relazioni sugli identificatori delle entità coinvolte (più gli attributi propri).

## Relazioni molti a molti



Impiegato(Matricola, Cognome, Stipendio)

Progetto(Codice, Nome, Budget)

Lavoro(Matricola, Codice, DataInizio)

## Relazioni molti a molti

Impiegato(Matricola, Cognome, Stipendio)

Progetto(Codice, Nome, Budget)

Lavoro(Matricola, Codice, DataInizio)

### o Con vincoli di integrità referenziale fra:

- Matricola in Lavoro e (la chiave di) Impiegato
- Codice in Lavoro e (la chiave di) Progetto

## Modifica Nomi

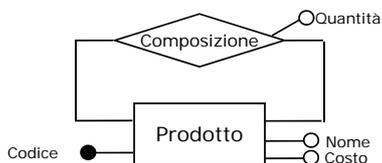
Impiegato(Matricola, Cognome, Stipendio)

Progetto(Codice, Nome, Budget)

Lavoro(Matricola, Codice, DataInizio)

Lavoro(MatrImpiegato, CodProgetto, DataInizio)

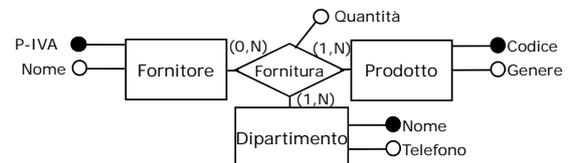
## Relazioni ricorsive



Prodotto(Codice, Nome, Costo)

Composizione(Composto, Componente, Quantità)

## Relazioni n-arie



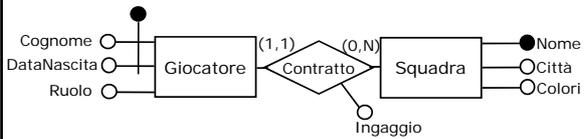
Fornitore(P-IVA, Nome)

Prodotto(Codice, Genere)

Dipartimento(Nome, Telefono)

Fornitura(Fornitore, Prodotto, Dipartimento, Quantità)

## Relazioni uno a molti



Giocatore(Cognome, DataNascita, Ruolo)  
 ? Contratto(CognGiocatore, DataNascGiocatore,  
Squadra, Ingaggio)  
 Squadra(Nome, Città, Colori)

## Relazioni uno a molti

Giocatore(Cognome, DataNascita, Ruolo)  
 Contratto(CognGiocatore, DataNascGiocatore,  
Squadra, Ingaggio)  
 Squadra(Nome, Città, Colori)

### Soluzione più compatta:

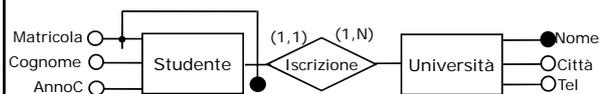
Giocatore(Cognome, DataNascita, Ruolo,  
 Squadra, Ingaggio)  
 Squadra(Nome, Città, ColoriSociali)

## Relazioni uno a molti

Giocatore(Cognome, DataNascita, Ruolo,  
 Squadra, Ingaggio)  
 Squadra(Nome, Città, ColoriSociali)

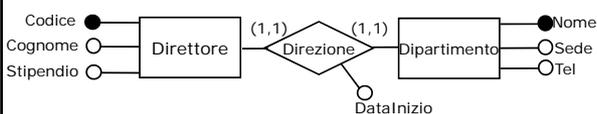
- o Con vincolo di integrità referenziale fra **Squadra** in **Giocatore** e la chiave di **Squadra**.
- o Se la cardinalità minima della relazione è 0, allora **Squadra** in **Giocatore** deve ammettere valore nullo.

## Entità con identificazione esterna



Studente(Matricola, NomeUniversità,  
 Cognome, AnnoDiCorso)  
 Università(Nome, Città, Telefono)

## Relazioni uno a uno

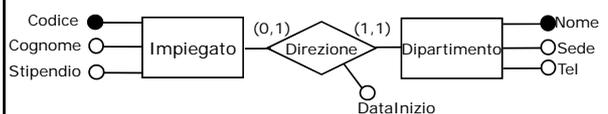


### Due possibilità:

Direttore (Codice, Cognome, Stipendio,  
 NomeDipDiretto, InizioDirezione)  
 Dipartimento (Nome, Sede, Telefono)

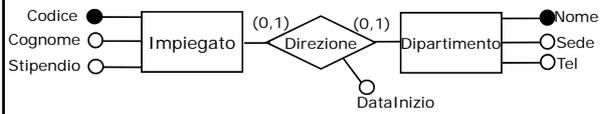
Direttore (Codice, Cognome, Stipendio)  
 Dipartimento (Nome, Sede, Telefono ,  
 Direttore, InizioDirezione)

## Relazioni uno a uno con partecipazione opzionale



Impiegato (Codice, Cognome, Stipendio)  
 Dipartimento (Nome, Sede, Telefono,  
 Direttore, InizioDirezione)

## Relazioni uno a uno con partecipazione opzionale



Impiegato (Codice, Cognome, Stipendio)  
Dipartimento (Nome, Sede, Telefono)  
Direzione (NomeDirettore, Dipartimento,  
DataInizio)

## Riferimenti

- o Paolo Atzeni, Stefano Ceri, Stefano Paraboschi, Riccardo Torlone. Basi di Dati. Modelli e linguaggi di interrogazione. McGraw-Hill, 2002.