

La progettazione delle Basi di Dati

Gli argomenti trattati in questi lucidi possono essere ritrovati nel cap. 6 del libro:

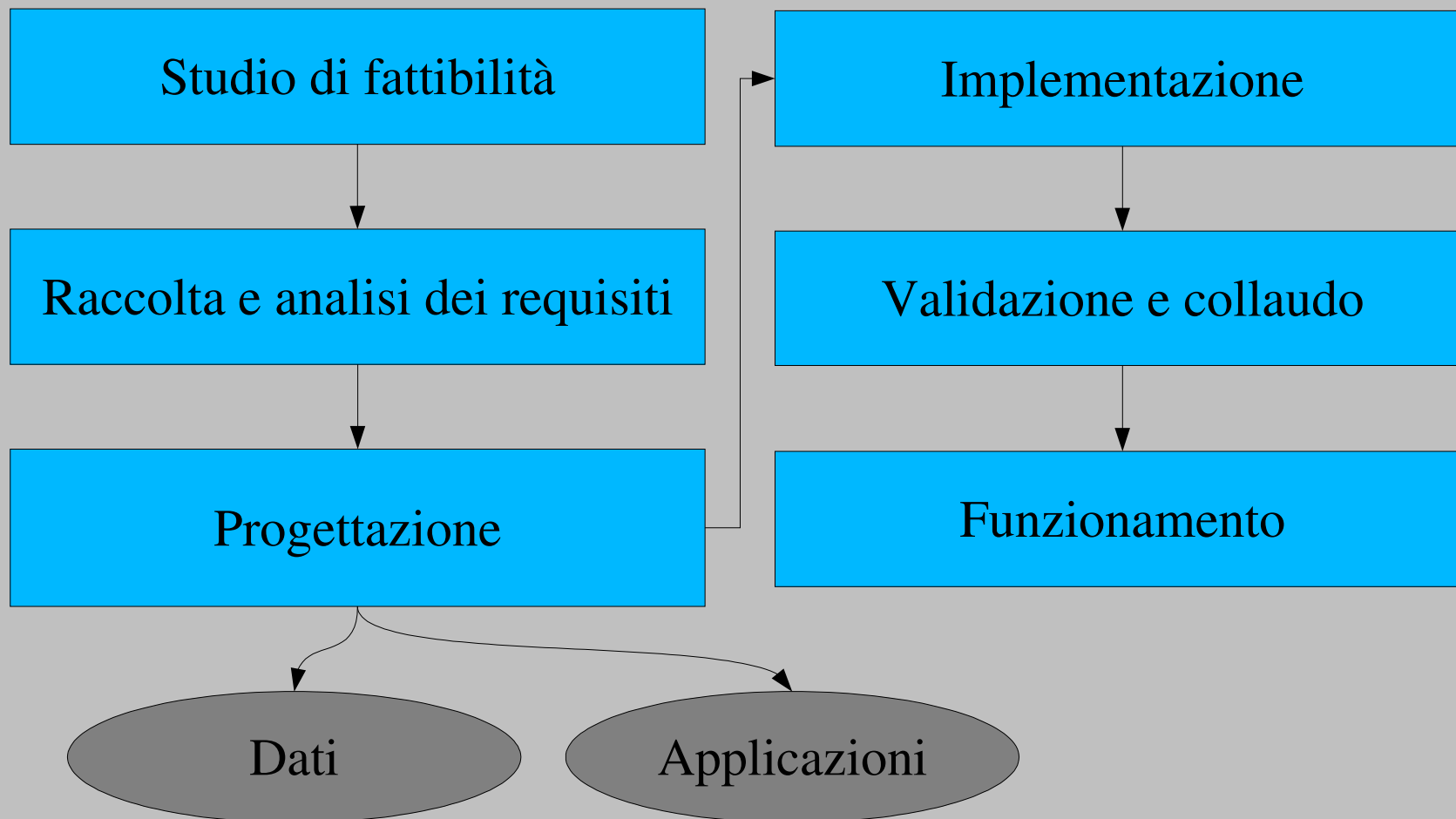
Atzeni, Ceri, Paraboschi, Torlone
Basi di dati: Modelli e linguaggi di interrogazione
McGraw-Hill

Cos'è la progettazione?

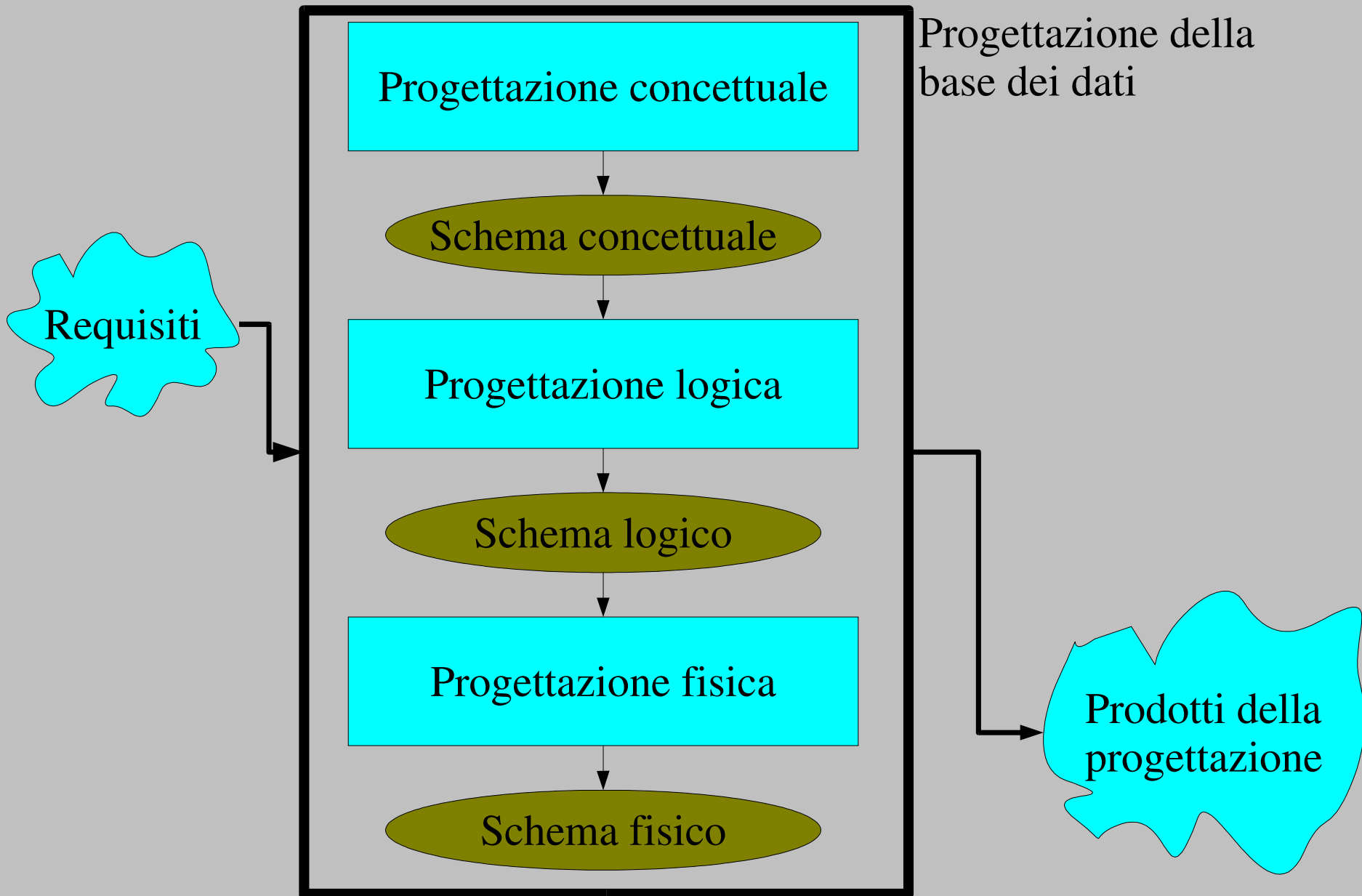
A partire dai requisiti, definire **struttura**, **caratteristiche** e **contenuto** della base di dati.

È una componente del “*ciclo di vita*” dei sistemi informativi.

Il ciclo di vita dei sistemi informativi



La progettazione dei dati



Progettazione concettuale

Rappresentare le specifiche **informali** della realtà di interesse tramite una descrizione **formale** completa ma indipendente dal sistema usato per rappresentare i dati.

Progettazione logica

Tradurre lo schema concettuale secondo il **modello dei dati** del DBMS. Nel caso di sistema relazionale, lo schema logico che si ottiene è l'insieme delle tabelle e dei vincoli relativi.

Questa fase è indipendente dal DBMS scelto, ma dipende dal modello logico dei dati.

Progettazione fisica

Completare lo schema logico con la specifica dei parametri fisici di memorizzazione (ad esempio con l'elenco degli indici da utilizzare).

Questa fase dipende dal DBMS scelto.

Specifiche dei requisiti

Distinguiamo tra specifiche dei **dati** e delle **operazioni**.

Le specifiche dei **dati** servono nella progettazione concettuale.

Le specifiche delle **operazioni** servono in tutte le fasi della progettazione per controllare la completezza (prog. concettuale) oppure per ottimizzare le prestazioni (prog. fisica e logica).

La Progettazione Concettuale

Il modello Entità-Relazione

È un modello **concettuale**, ovvero serve a descrivere la realtà di interesse.

È composto da **costrutti** che si combinano tra loro per formare degli **schemi concettuali**, i quali descrivono la struttura della realtà di interesse.

Le Entità /1

Ogni **entità** rappresenta un classe di oggetti (fatti, cose, persone, etc.) che hanno delle proprietà comuni ed una esistenza “**autonoma**”.

Le entità sono rappresentate da rettangoli che racchiudono il nome delle entità.



Aeroplano

Aeroporto

Volo

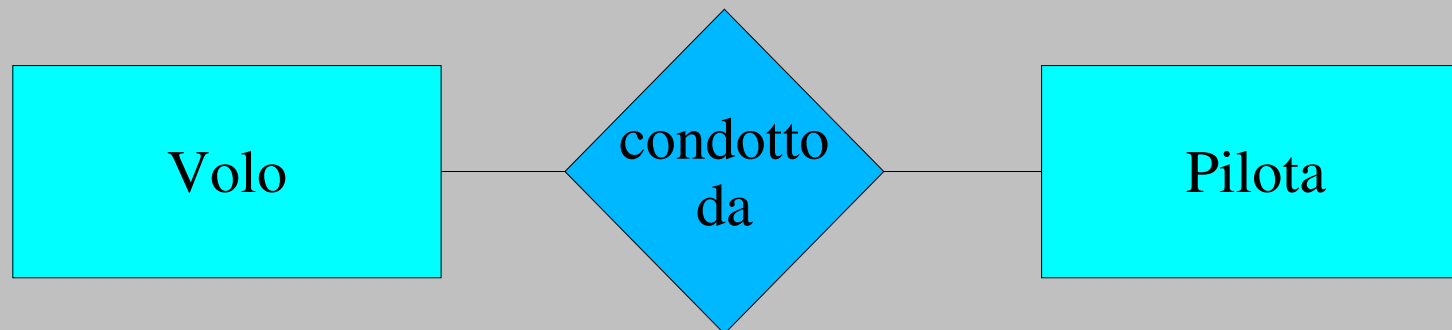
Le Entità /2

Una **istanza** di una entità è un oggetto della classe che l'entità rappresenta.

Notare che l'istanza **non** è un insieme di valori che identificano un oggetto, ma è proprio l'oggetto stesso.

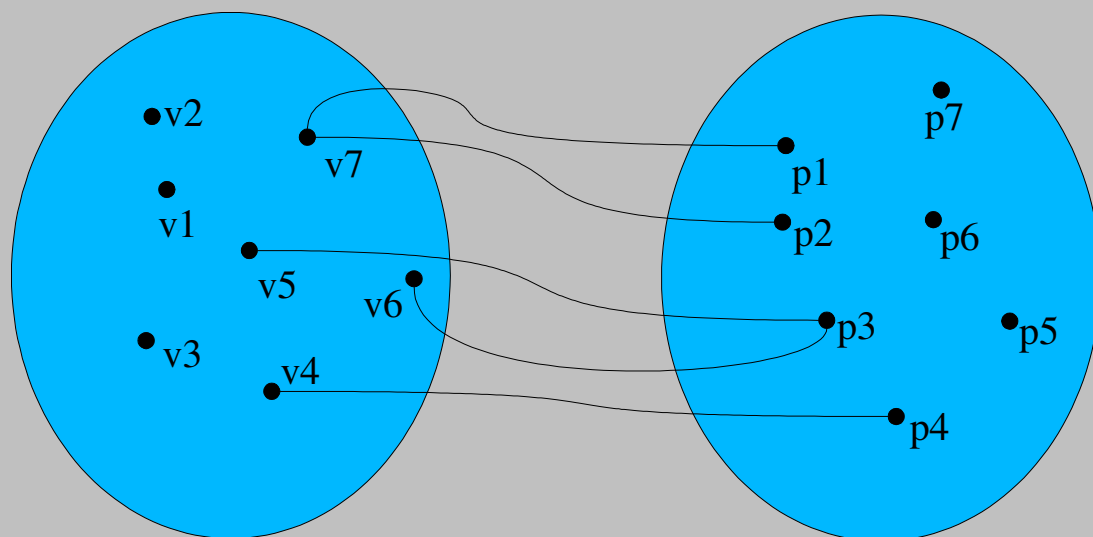
Le Relazioni /1

Una relazione rappresenta un legame logico, significativo per l'applicazione, tra due o più entità. Ad esempio, la relazione “condotto da” lega l'entità Volo all'entità Pilota:



Le Relazioni /2

Una **istanza** di una relazione è una ennupla (coppia per relazioni binarie) costituita da occorrenze di entità, una per ogni entità coinvolta.



sono istanze le coppie:

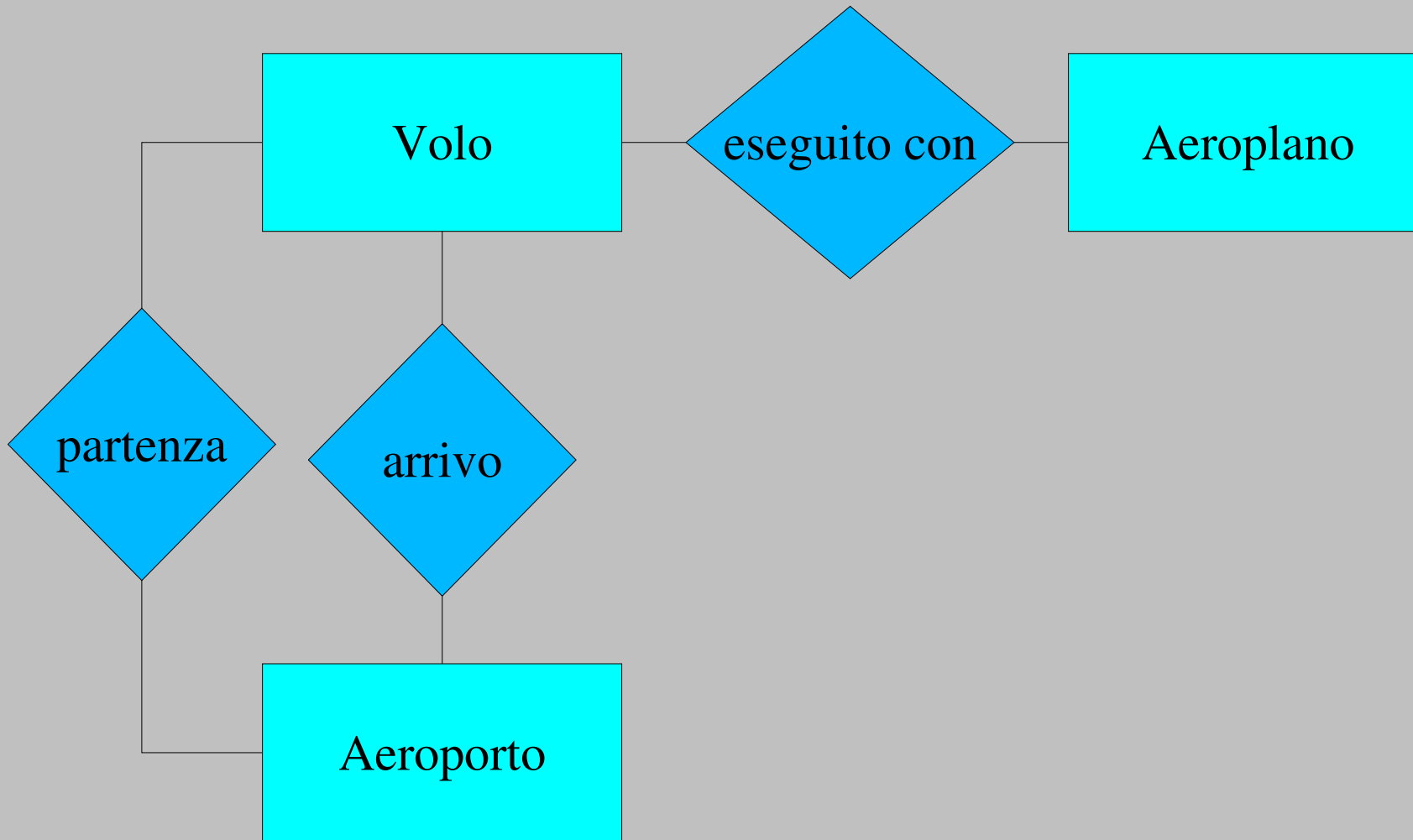
$(v7, p1)$

$(v7, p2)$

$(v4, p4)$.

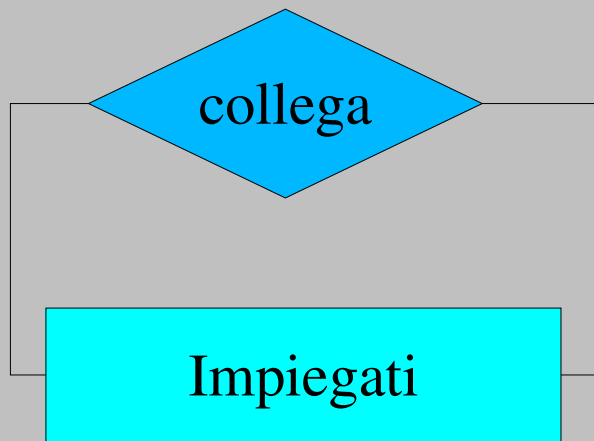
....

Le Relazioni /3

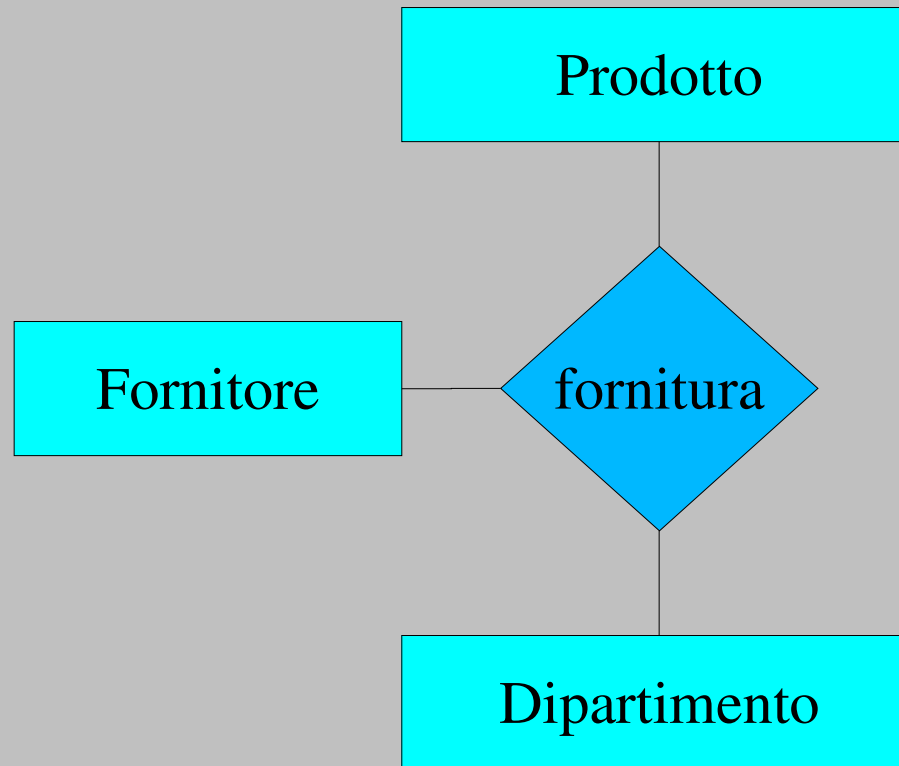


Le Relazioni /4

Relazioni ricorsive

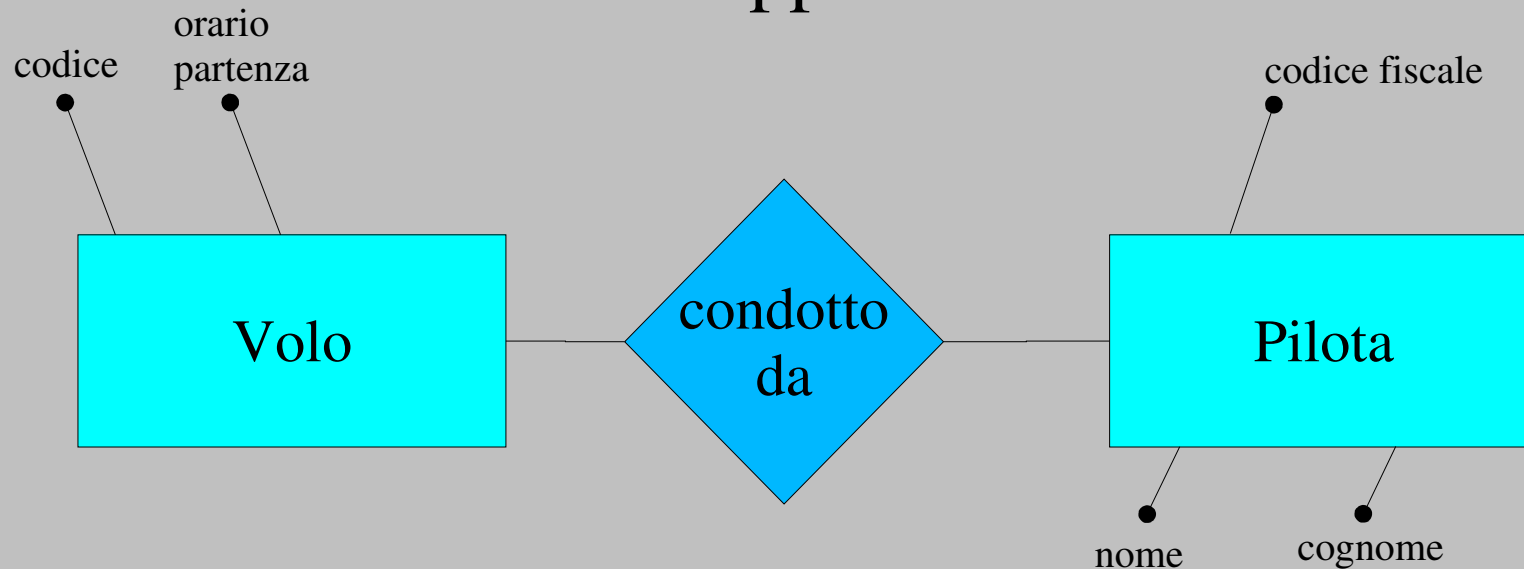


Relazioni ternarie

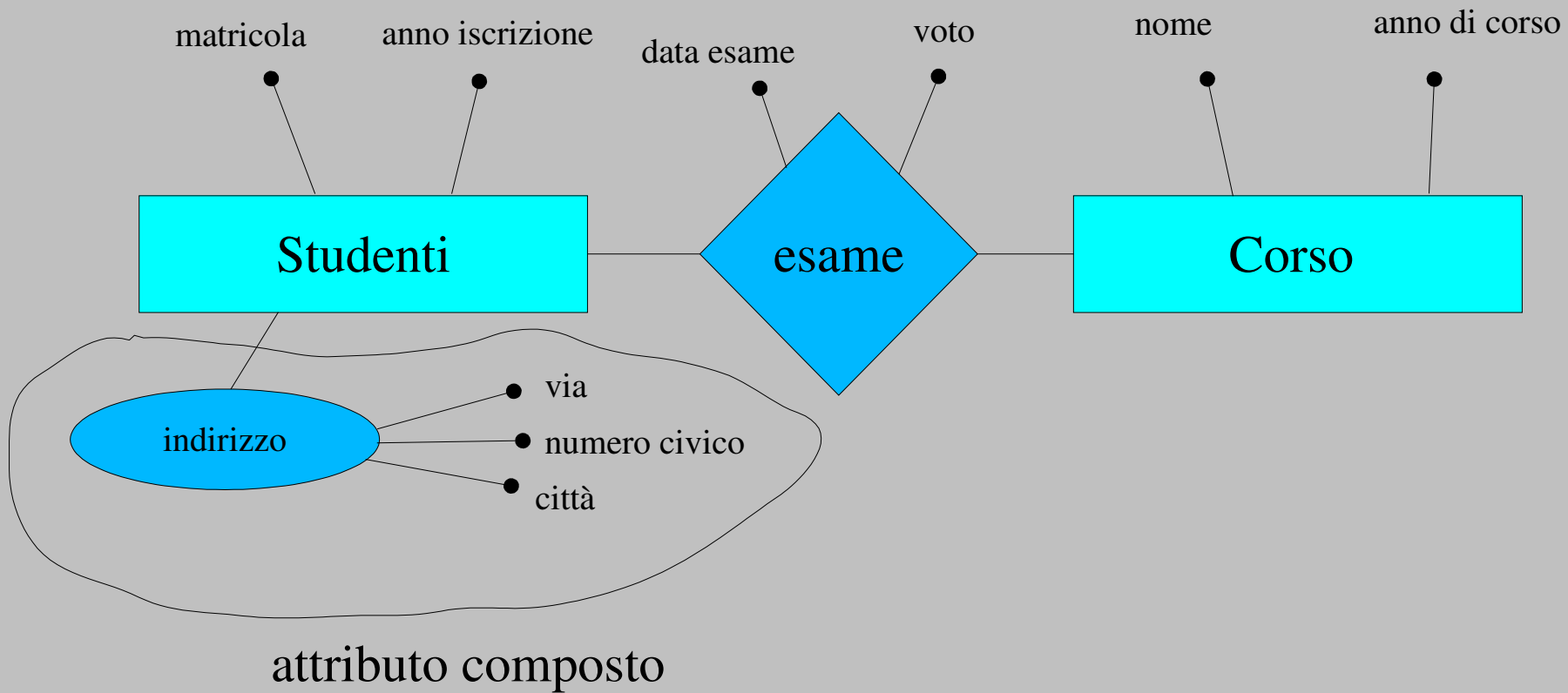


Gli Attributi

Descrivono le proprietà elementari di entità o relazioni che sono di interesse ai fini dell'applicazione.

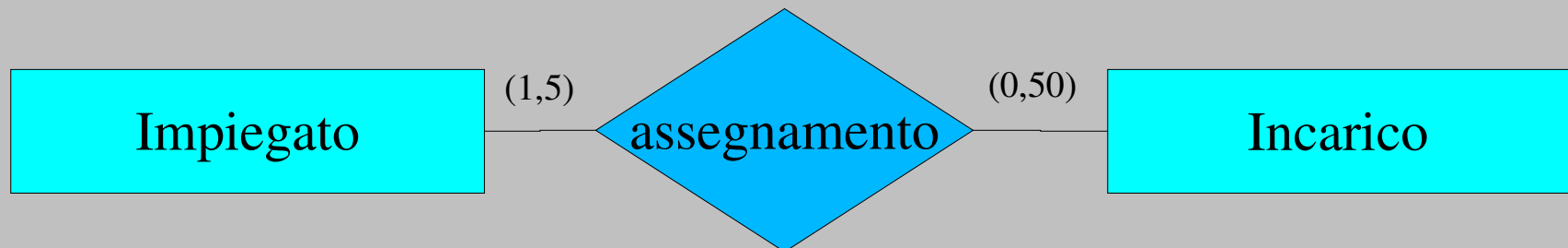


Gli Attributi /2



Cardinalità delle Relazioni

Dicono quante volte, in una relazione tra entità, una occorrenza di una di queste entità può essere legata a occorrenze della altre entità coinvolte.



Ad ogni impiegato possono essere assegnati da un minimo di 1 fino a un massimo di 5 incarichi. Un incarico può essere assegnato fino 50 impiegati.

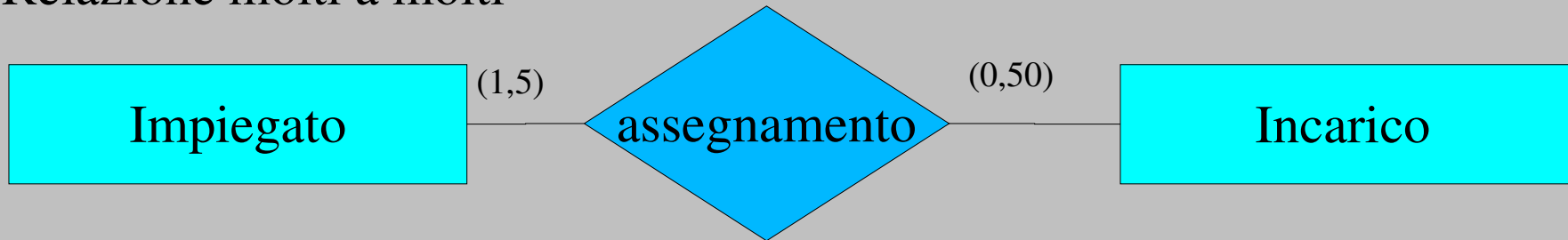
Cardinalità delle Relazioni /2

Per le relazioni binarie, si distinguono i seguenti tipi:

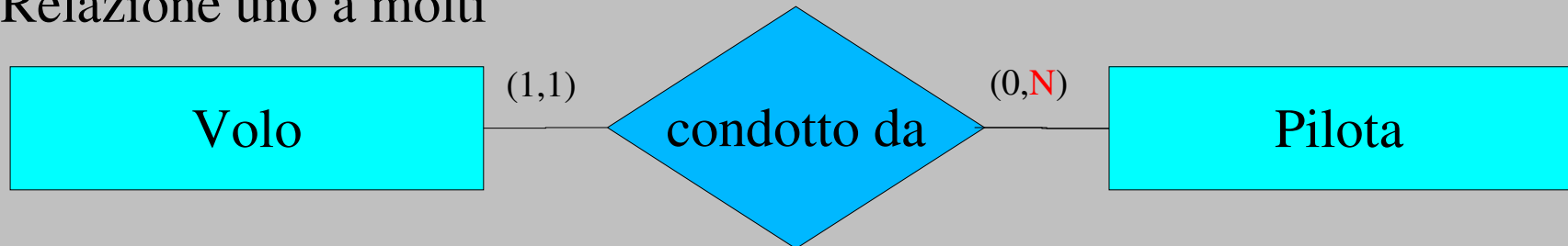
- **relazioni 1 a 1**: quando la cardinalità massima delle entità coinvolte è 1;
- **relazioni 1 a molti**: quando la cardinalità delle entità coinvolte è 1 per una entità e un numero maggiore di 1 per l'altra;
- **relazioni molti a molti**: quando la cardinalità massima delle entità coinvolte è maggiore di 1 in entrambi i casi.

Cardinalità delle Relazioni /3

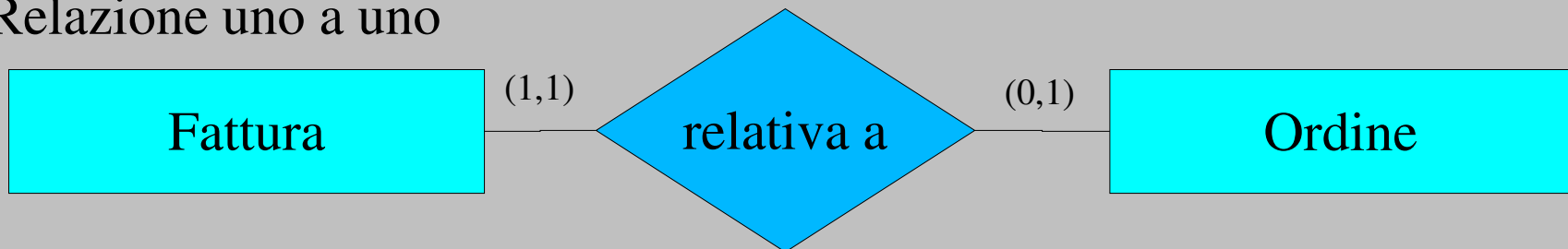
Relazione molti a molti



Relazione uno a molti



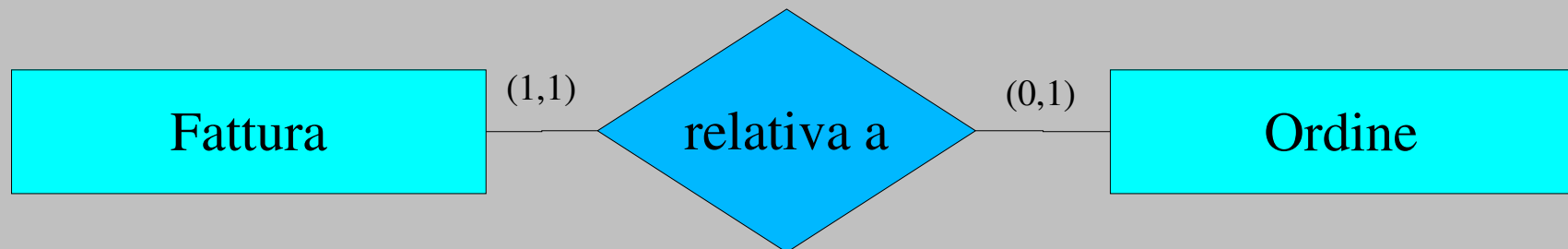
Relazione uno a uno



Cardinalità delle Relazioni /4

Si dice che la partecipazione di una entità ad una relazione è:

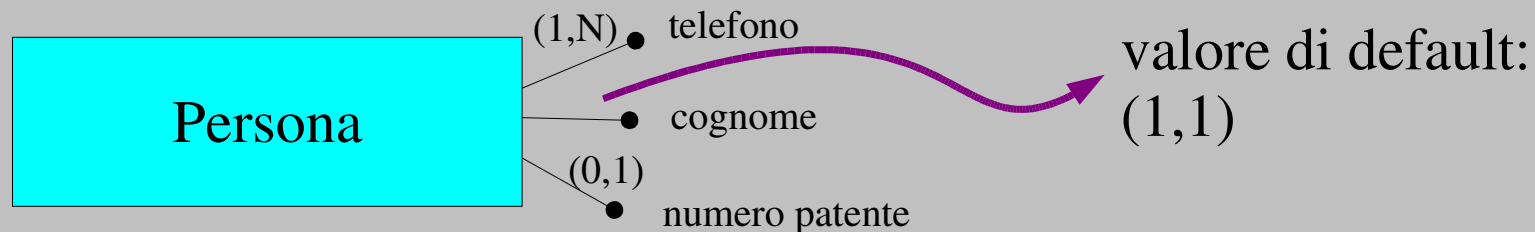
- **opzionale**: se la cardinalità minima è 0
- **obbligatoria**: se la cardinalità minima è almeno 1



La partecipazione di Fattura è obbligatoria, quella di Ordine è opzionale.

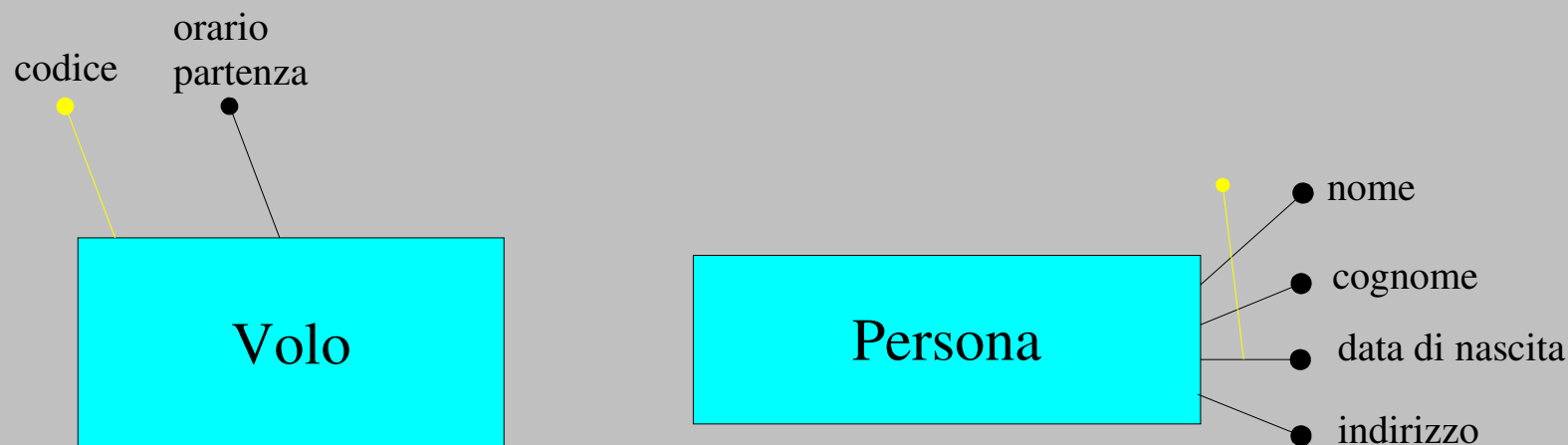
Cardinalità degli Attributi

Descrivono il numero minimo e massimo di attributi che possono essere associati ad una entità.



Identificatori delle Entità

Vengono specificati per ogni entità e descrivono i concetti (attributi o relazioni) che permettono di identificare in maniera univoca le occorrenza delle entità.

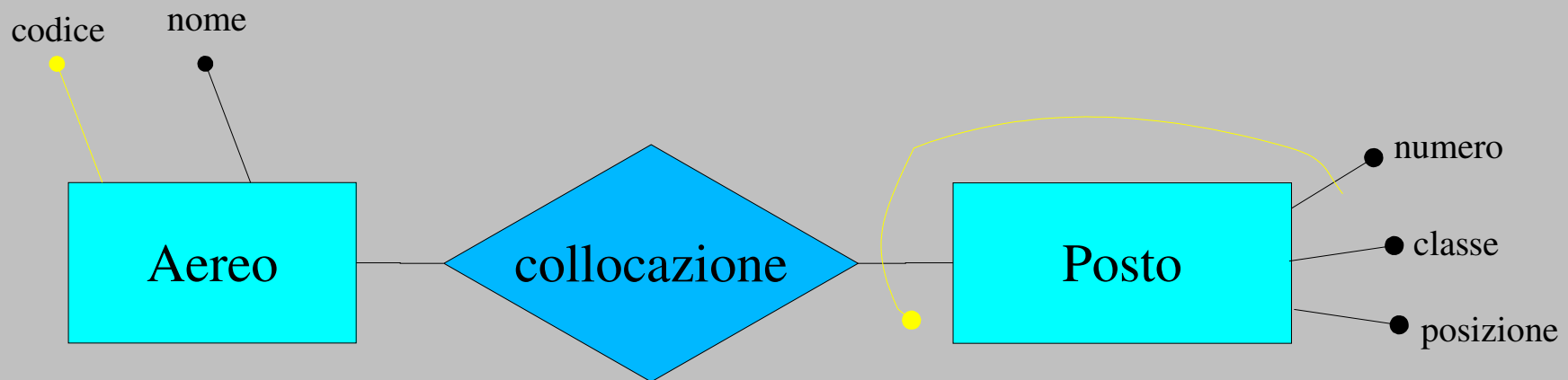
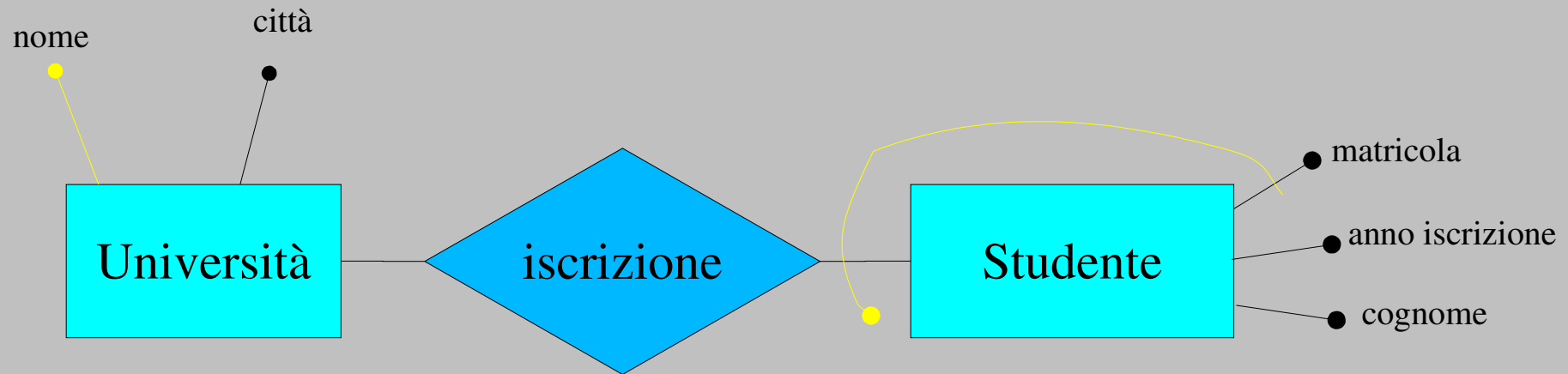


Identificatori delle Entità /2

Si parla di **identificatore interno** (detto anche **chiave**) quando esso è formato solo da attributi. Altrimenti, se entrano in gioco anche le relazioni, si parla di **identificatore esterno**.

Le entità che hanno identificatori esterni sono dette **entità deboli**.

Identificatori delle Entità /3

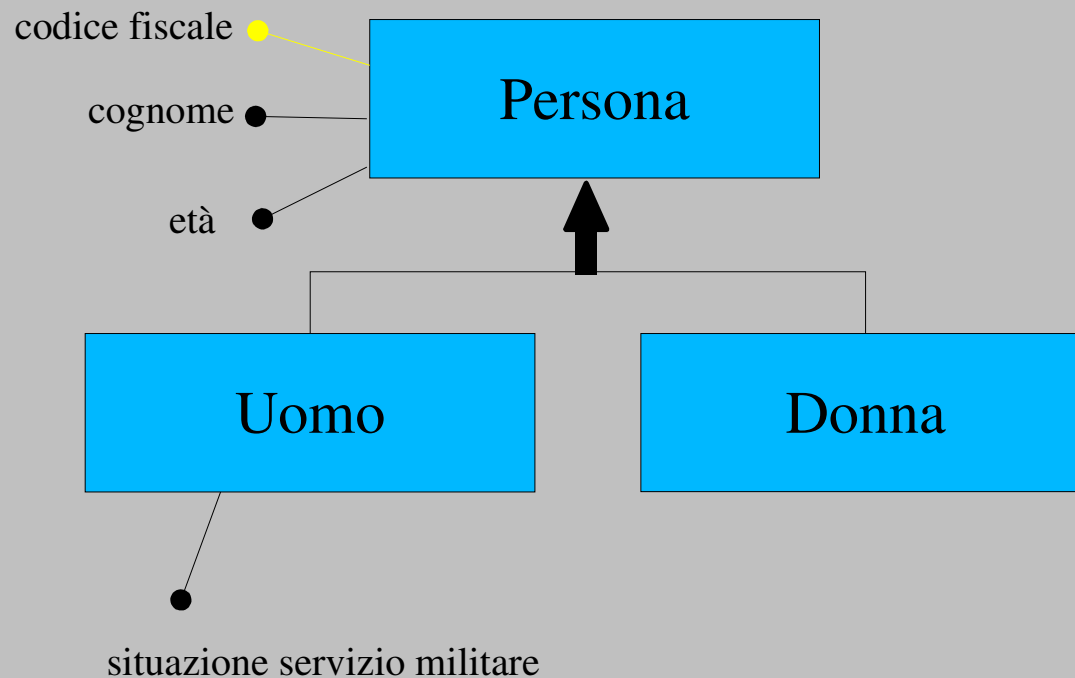


Generalizzazioni /1

Rappresentano legami logici tra una entità E , detta entità **padre**, e varie entità $E_1 \dots E_n$ dette entità **figlie**.

Si dice che E è una **generalizzazione** e che $E_1 \dots E_n$ sono **specializzazioni**.

Generalizzazioni /2



Caratteristiche delle generalizzazioni:

- ogni occorrenza delle entità figlie è anche una occorrenza del padre
- ogni attributo del padre è anche attributo di tutte le entità figlie (**ereditarietà**).

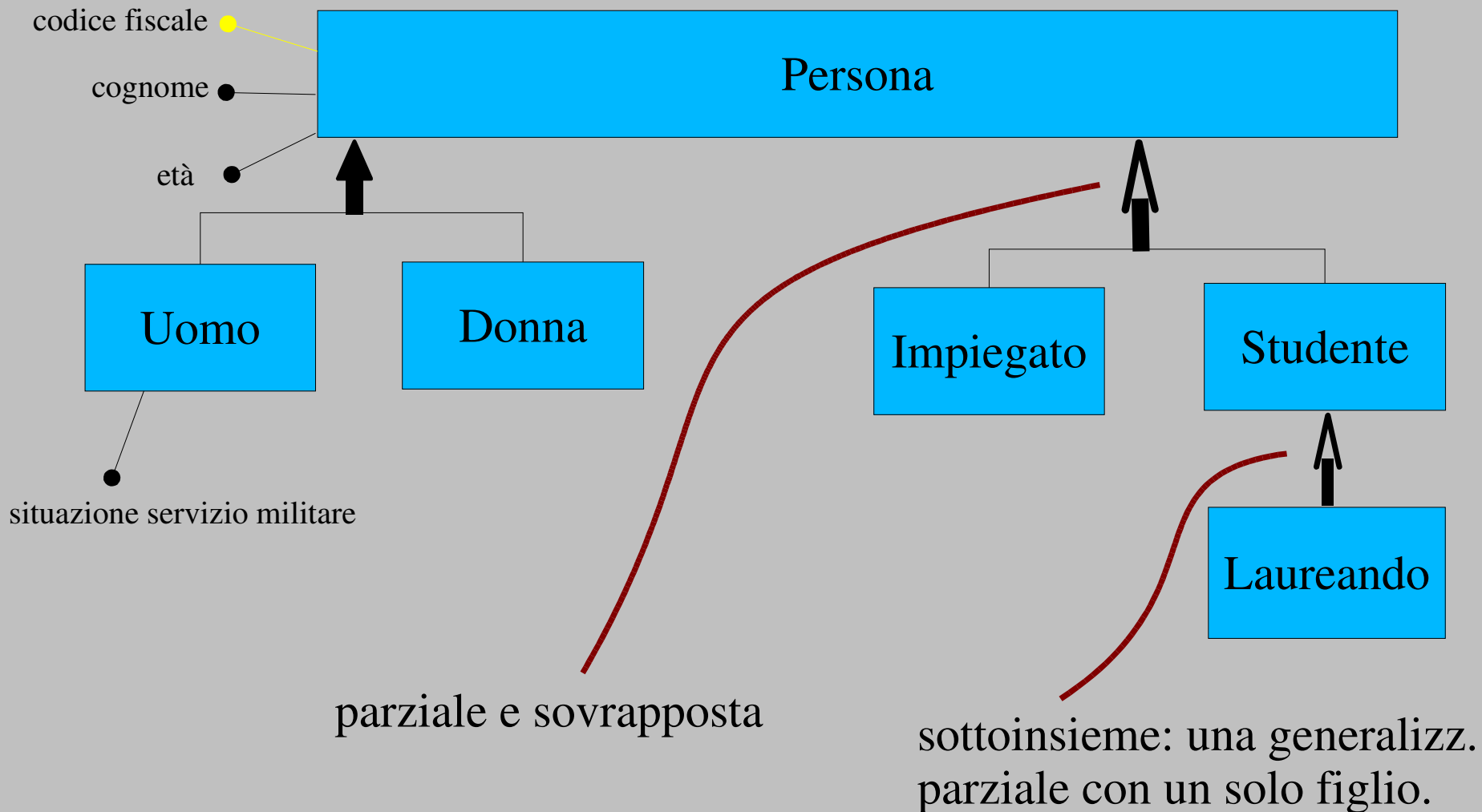
Generalizzazioni /3

Una generalizzazione è:

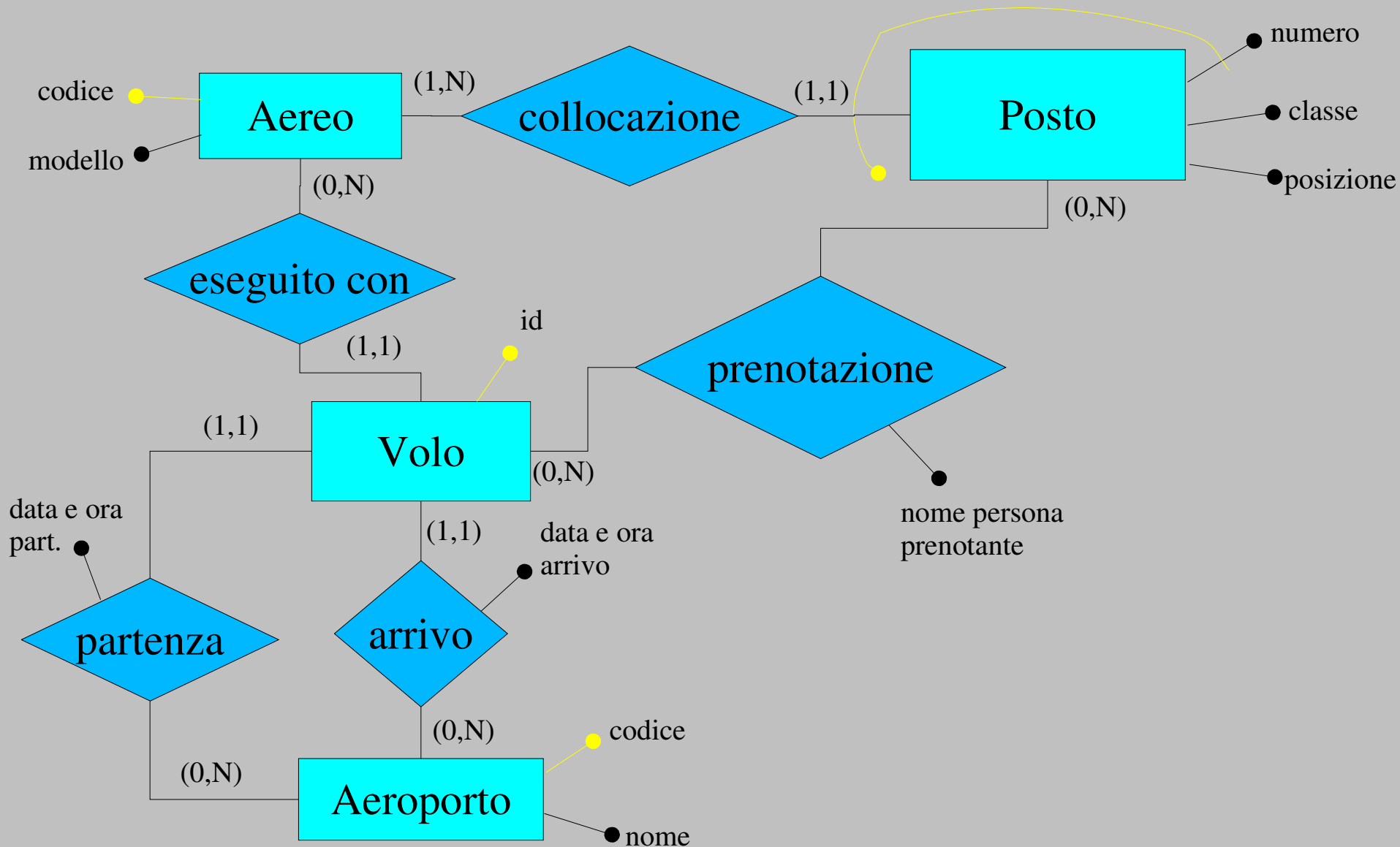
- **totale**: se ogni occorrenza della classe padre è occorrenza di una delle classi figlie.
- **parziale**: altrimenti

- **esclusiva**: se ogni occorrenza della classe padre è al più occorrenza di una delle classi figlie.
- **sovrapposta**: altrimenti

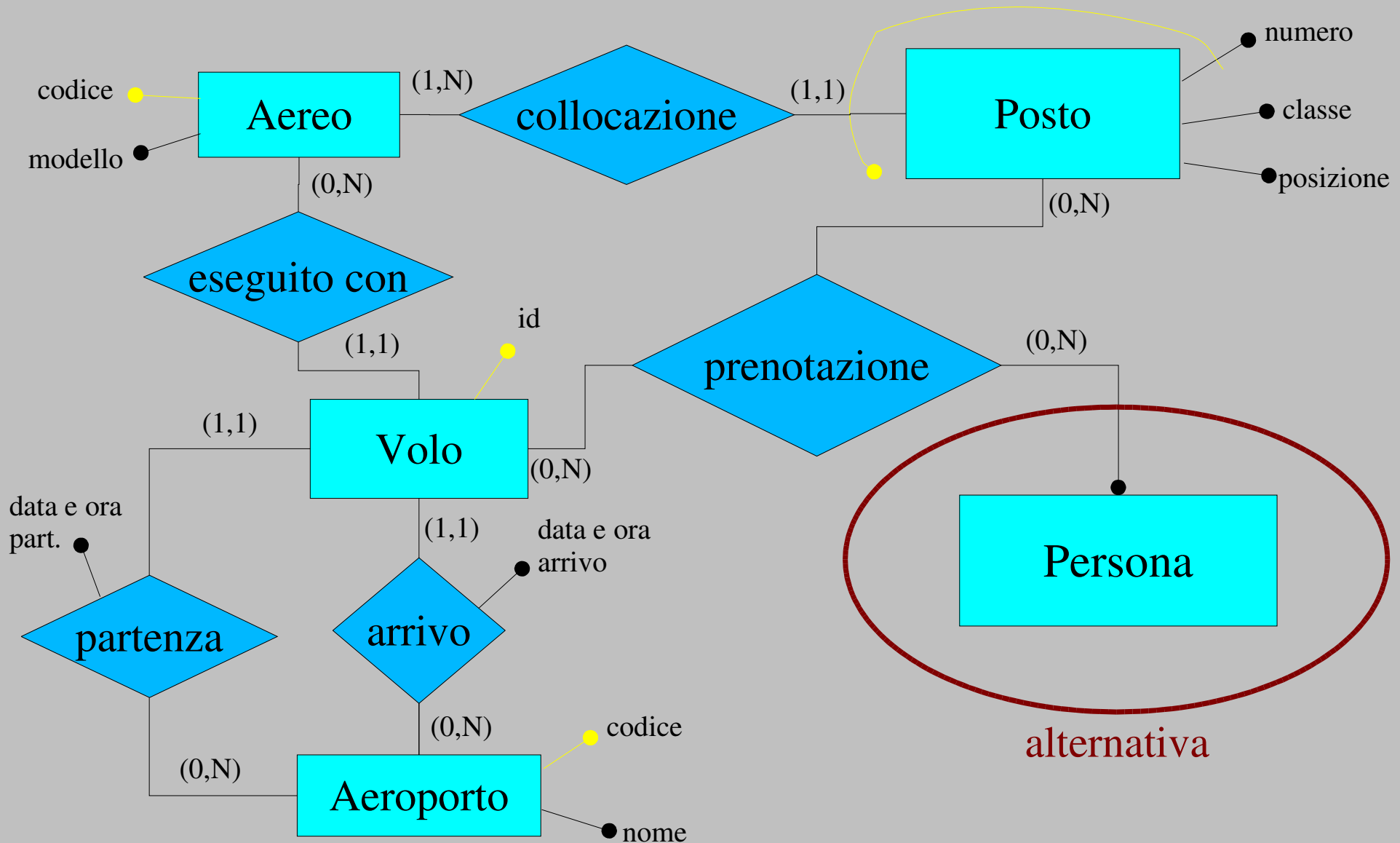
Generalizzazioni /4



Uno schema complessivo



Uno schema complessivo



Applicazioni del modello ER

I diagrammi ER possono essere usati in vari modi:

- come ausilio alla progettazione delle basi di dati
- a scopo documentativo, in quanto facilmente comprensibili anche per chi non è esperto di basi di dati
- per descrivere la struttura di una base di dati già esistente

Insufficienza degli schemi ER

Gli schemi ER non sono quasi mai sufficienti per rappresentare tutte le informazioni che si desidera. Infatti:

- talvolta è necessario tralasciare qualche attributo per rendere più chiaro lo schema
- lo schema indica il nome degli attributi ma non i valori che questi possono assumere (dominio)
- alcuni vincoli tra i dati non possono essere espressi

Le Regole Aziendali

Occorre accompagnare il diagramma ER con ulteriore documentazione. Le informazioni supplementari allo schema ER sono spesso chiamate “**business rules**” (regole aziendali).

Le Regole Aziendali /2

Una regola aziendale può essere:

- la **descrizione di un concetto** rilevante di una applicazione (definizione precisa di una entità o di un attributo,...)

“una istanza della entità volo identifica un singolo volo, compiuto tra due aeroporti, in un determinato giorno e a una determinata ora di quel giorno”

Le Regole Aziendali /3

- un **vincolo di integrità** sui dati, sia esso la documentazione di un vincolo espresso nello schema ER che un vincolo nuovo non esprimibile nello schema.
“un impiegato deve avere uno stipendio maggiore del direttore del dipartimento a cui afferisce”
- una **derivazione**, ovvero una regola matematica per ottenere alcuni attributi da altri attributi dello schema
“l'attributo Costo di ottiene sommando all'attributo Costo Netto l'attributo Tasse”

Il dizionario dei dati

Le regole aziendali di tipo descrittivo vengono di solito raccolte nel “**dizionario dei dati**”.

Esso è composto da un tabella per le entità:

<i>Entità</i>	<i>Descrizione</i>	<i>Attributi</i>	<i>Identificatore</i>
Volo	Un singolo volo compiuto in una data e ora ben precisa.	Codice	Codice
Aeroporto	Un aeroporto	Codice, Nome	Codice
Aereo	Un singolo aeroplano	Codice, Nome, Data primo volo	Codice

Il dizionario dei dati /2

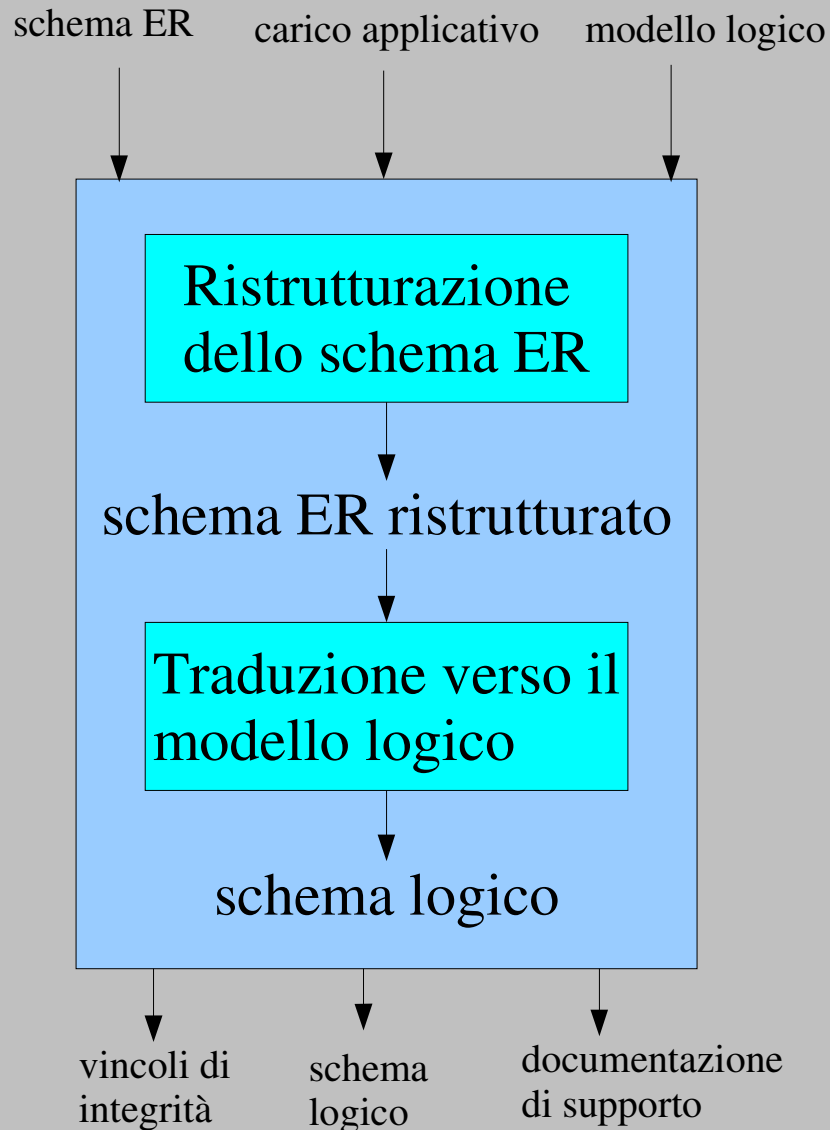
...e da una tabella per le relazioni.

<i>Relazione</i>	<i>Descrizione</i>	<i>Entità coinvolte</i>	<i>Attributi</i>
Partenza	Associa un volo all'aeroporto di partenza	Aereoporto, Volo	Data e Ora di partenza
Eseguito con	Associa un volo all'aereo che lo esegue.	Aereo, Volo	

La Progettazione Logica

(non ancora terminato)

La Progettazione Logica



Si tratta di costruire lo schema logico a partire dai risultati della prog. concettuale e dal **carico applicativo**.

Occorre ottimizzare la traduzione secondo il tipo di operazioni che ci aspettiamo.

Fasi della progettazione logica

- **Ristrutturazione dello schema ER:** si ottiene un nuovo schema ottimizzato a seconda del carico applicativo. Per questo, non è più un vero e proprio schema concettuale.
- **Traduzione verso il modello logico:** produce lo schema logico desiderato. Nel caso del modello relazionale, si ottengono un insieme di tabelle.

Analisi delle Prestazioni

Serve per effettuare degli studi di massima su due parametri che regolano le prestazioni del software:

costo delle operazioni

occupazione di memoria

Il carico applicativo

- Per studiare questi parametri occorrono le seguenti informazioni:
 - volume dei dati
 - numero delle istanze per ogni entità e per ogni relazione
 - dimensioni di ogni attributo
 - caratteristiche delle operazioni
 - tipo dell'operazione (**interattiva** o **batch**)
 - frequenza dell'operazione
 - dati coinvolti (sia entità che relazioni)