

Progettazione di basi di dati:
Metodologie e modelli
Il modello ER

Il problema

- Proviamo a modellare una applicazione definendo direttamente lo schema logico della base di dati:
 - da dove cominciamo?
 - rischiamo di perderci subito nei dettagli
 - dobbiamo pensare subito a come correlare le varie tabelle (chiavi etc.)
 - il modello relazionale è “rigido”

Rivediamo un esempio già discusso

Strutture nidificate

DA FILIPPO		
VIA ROMA 2, ROMA		
RICEVUTA FISCALE		
1235 DEL 12/10/2019		
3	Coperti	3,00
2	Antipasti	6,20
3	Primi	12,00
2	Bistecche	18,00
TOTALE		39,20

DA FILIPPO		
VIA ROMA 2, ROMA		
RICEVUTA FISCALE		
1240 DEL 13/10/2019		
2	Coperti	2,00
2	Antipasti	7,00
2	Primi	8,00
2	Orate	20,00
2	Caffè	2,00
TOTALE		39,00

Strutture nidificate

DA FILIPPO		
VIA ROMA 2, ROMA		
RICEVUTA FISCALE		
1235 DEL 12/10/2019		
3	Coperti	3,00
2	Antipasti	6,20
3	Primi	12,00
2	Bistecche	18,00
TOTALE		39,20

DA FILIPPO		
VIA ROMA 2, ROMA		
RICEVUTA FISCALE		
1240 DEL 13/10/2019		
2	Coperti	2,00
2	Antipasti	7,00
2	Primi	8,00
2	Orate	20,00
2	Caffè	2,00
TOTALE		39,00

Strutture nidificate

DA FILIPPO		
VIA ROMA 2, ROMA		
RICEVUTA FISCALE		
1235 DEL 12/10/2019		
3	Coperti	3,00
2	Antipasti	6,20
3	Primi	12,00
2	Bistecche	18,00
TOTALE		39,20

DA FILIPPO		
VIA ROMA 2, ROMA		
RICEVUTA FISCALE		
1240 DEL 13/10/2019		
2	Coperti	2,00
2	Antipasti	7,00
2	Primi	8,00
2	Orate	20,00
2	Caffè	2,00
TOTALE		39,00

"Soluzione"

- La abbiamo vista alla lavagna qualche settimana fa

Strutture nidificate

Ricevute

<u>Numero</u>	Data	Qtà	Descrizione	Importo	Totale
1235	12/10/2019	3	Coperti	3,00	39,20
		2	Antipasti	6,20	
		3	Primi	12,00	
		2	Bistecche	18,00	
1240	13/10/2019	2	Coperti	2,00	39,00
		...	...	...	
		...	...	...	

- Ma i valori debbono essere semplici, non relazioni!

Relazioni che rappresentano strutture nidificate

Ricevute

<u>Numero</u>	Data	Totale
1235	12/10/2019	39,20
1240	13/10/2019	39,00

Dettaglio

<u>Numero</u>	Qtà	<u>Descrizione</u>	Importo
1235	3	Coperti	3,00
1235	2	Antipasti	6,20
1235	3	Primi	12,00
1235	2	Bistecche	18,00
1240	2	Coperti	2,00
...	...	...	...

"Soluzione"

- Abbiamo definito due relazioni
- Perché?
 - Perché ci sono due tipi di concetti diversi, con livello di dettaglio differente
 - ricevuta
 - voce o riga della ricevuta

Strutture nidificate, riflessione

- Abbiamo rappresentato veramente tutti gli aspetti delle ricevute?
- Dipende da che cosa ci interessa!
 - possono esistere linee ripetute in una ricevuta?
 - l'ordine delle righe e' rilevante?
- Sono possibili rappresentazioni diverse

"Soluzione"

- Anche questa la abbiamo discussa, sempre alla lavagna

Rappresentazione alternativa per strutture nidificate

Ricevute

<u>Numero</u>	Data	Totale
1235	12/10/2019	39,20
1240	13/10/2019	39,00

Dettaglio

<u>Numero</u>	<u>Riga</u>	Qtà	Descrizione	Importo
1235	1	3	Coperti	3,00
1235	2	2	Antipasti	6,20
1235	3	3	Primi	12,00
1235	4	2	Bistecche	18,00
1240	1	2	Coperti	2,00
...	...	...	...	...

Commento

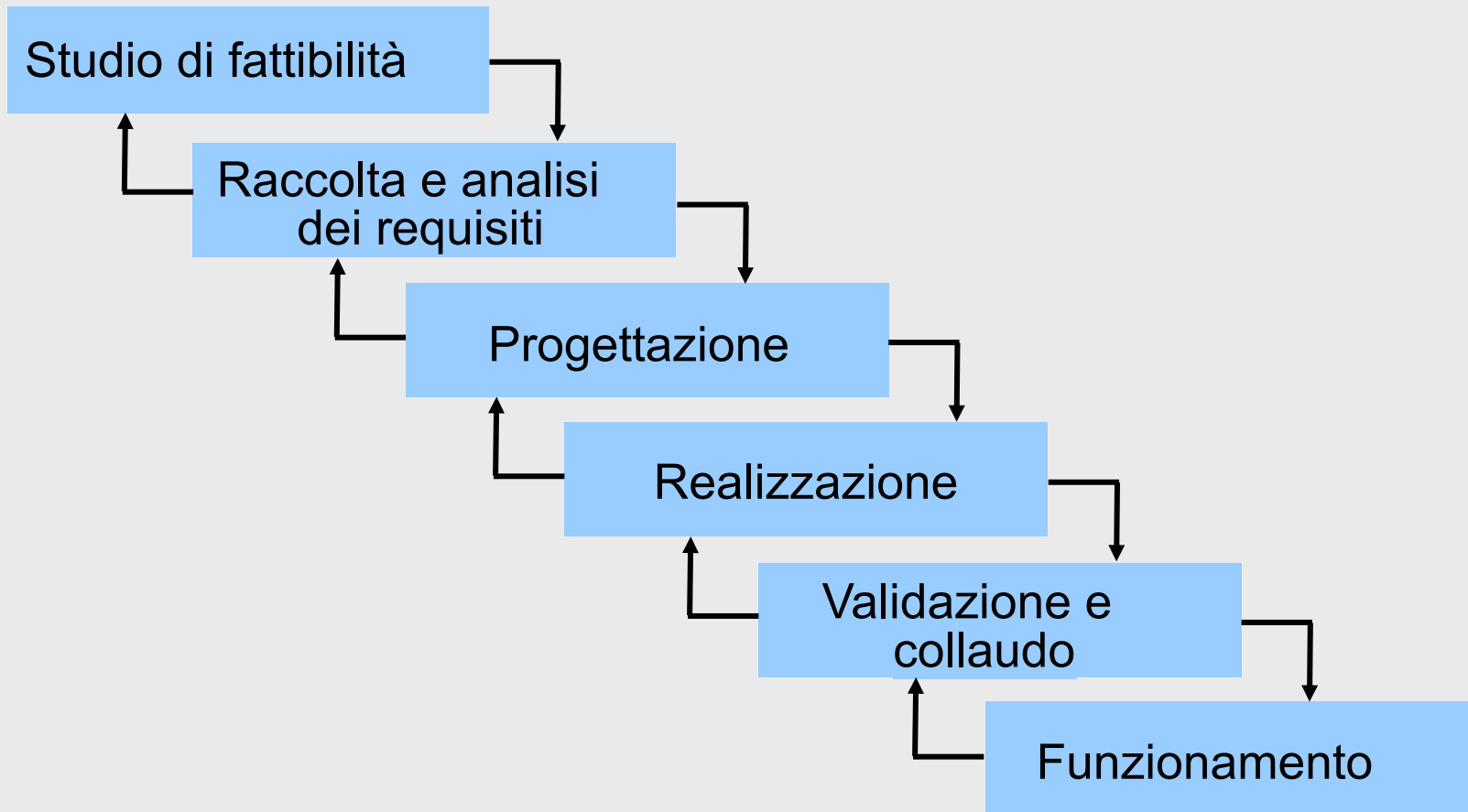
- I concetti sono gli stessi di prima, i dettagli cambiano

Il problema

- Proviamo a modellare una applicazione definendo direttamente lo schema logico della base di dati:
 - da dove cominciamo?
 - rischiamo di perderci subito nei dettagli
 - dobbiamo pensare subito a come correlare le varie tabelle (chiavi etc.)
 - il modello relazionale è “rigido”

Progettazione di basi di dati

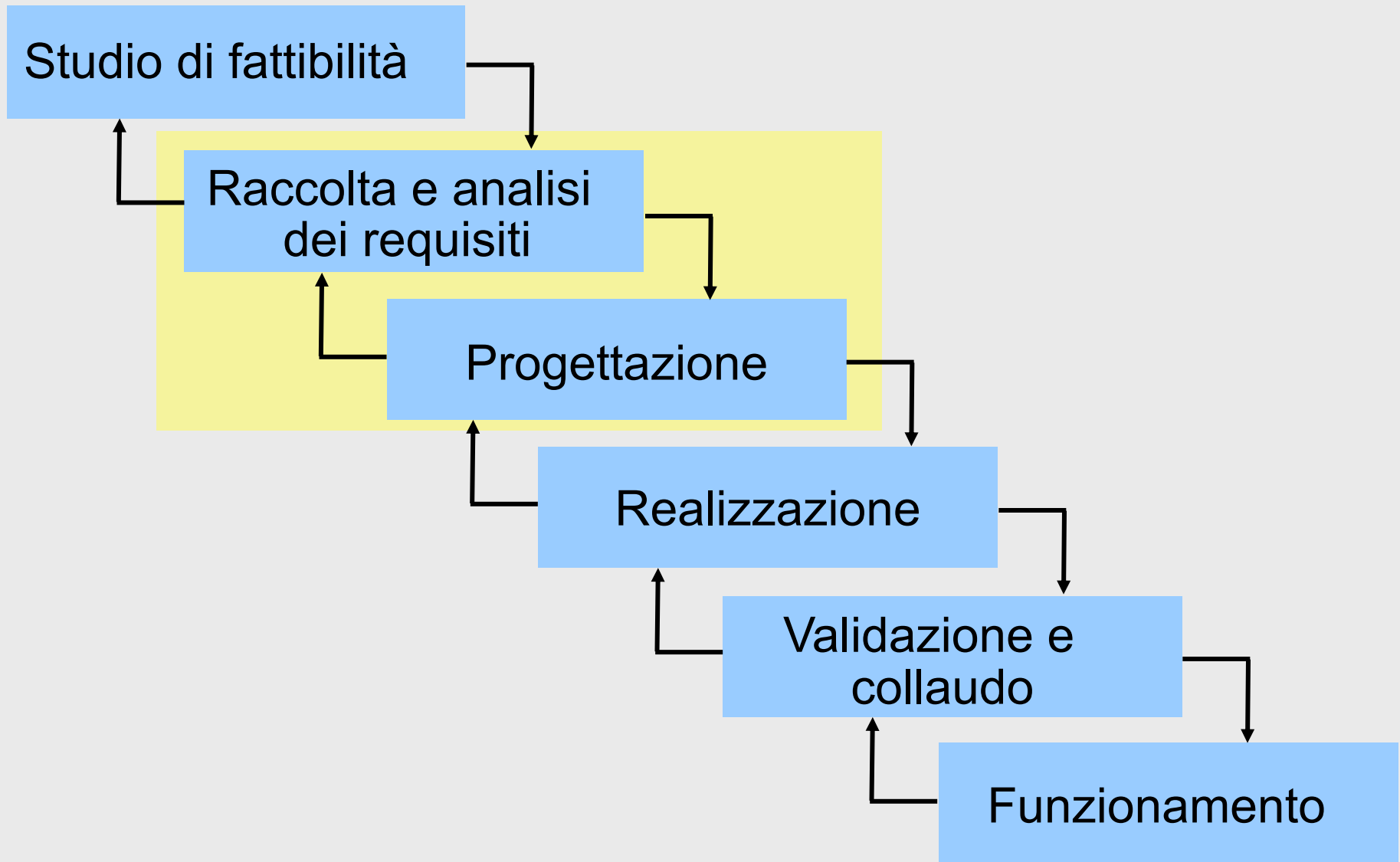
- È una delle attività del processo di sviluppo dei sistemi informativi
- va quindi inquadrata in un contesto più generale:
- il ciclo di vita dei sistemi informativi:
 - Insieme e sequenzializzazione delle attività svolte da analisti, progettisti, utenti, nello sviluppo e nell'uso dei sistemi informativi
 - attività iterativa, quindi ciclo



- Nota bene: le fasi non sono nettamente separate (vedere anche le frecce di ritorno), si può procedere in modo ciclico, graduale, ora non ci interessa

Fasi (tecniche) del ciclo di vita

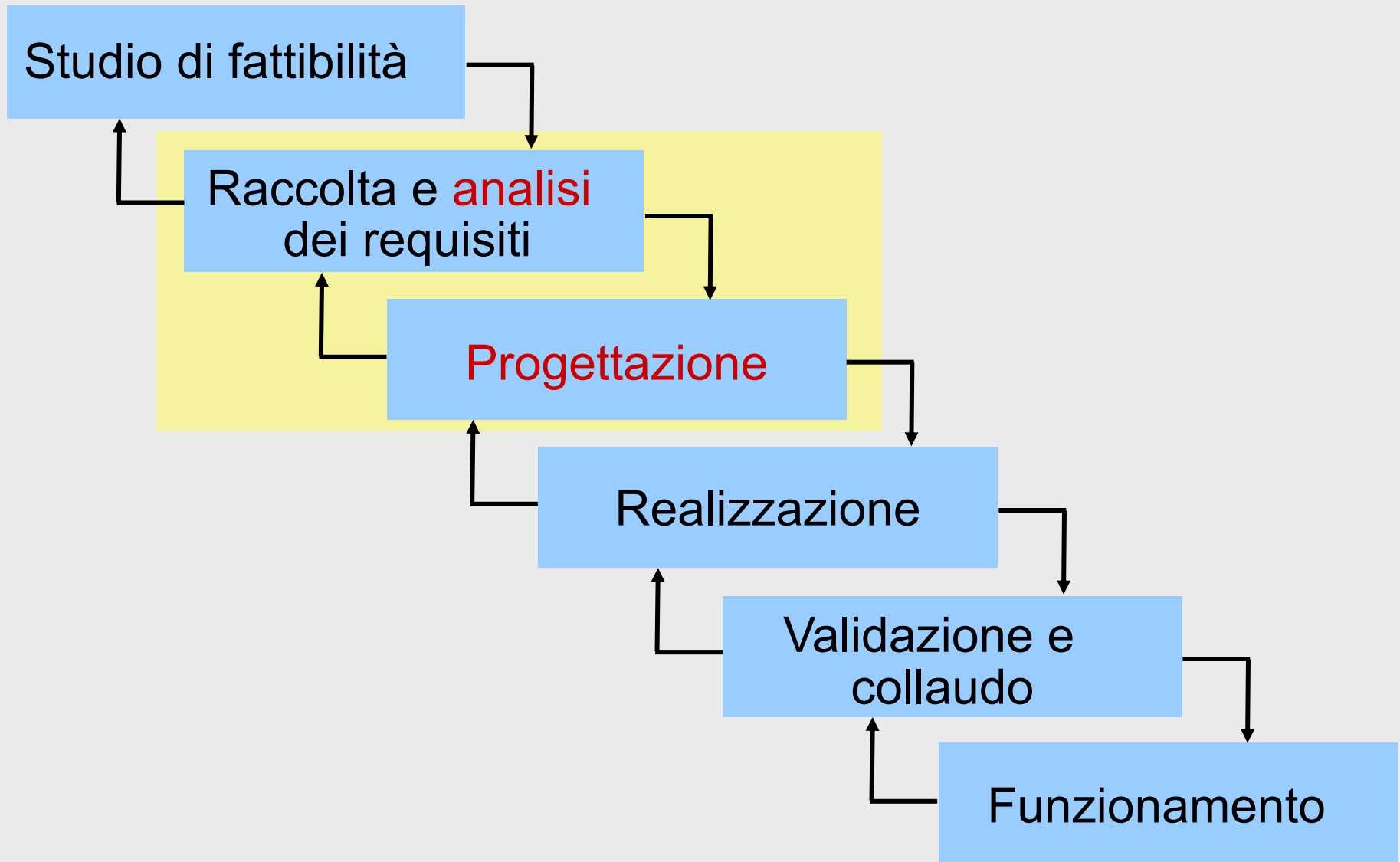
- Studio di fattibilità: definizione costi e priorità
- Raccolta e analisi dei requisiti: studio delle proprietà del sistema
- Progettazione: di dati e funzioni
- Realizzazione
- Validazione e collaudo: sperimentazione
- Funzionamento: il sistema diventa operativo



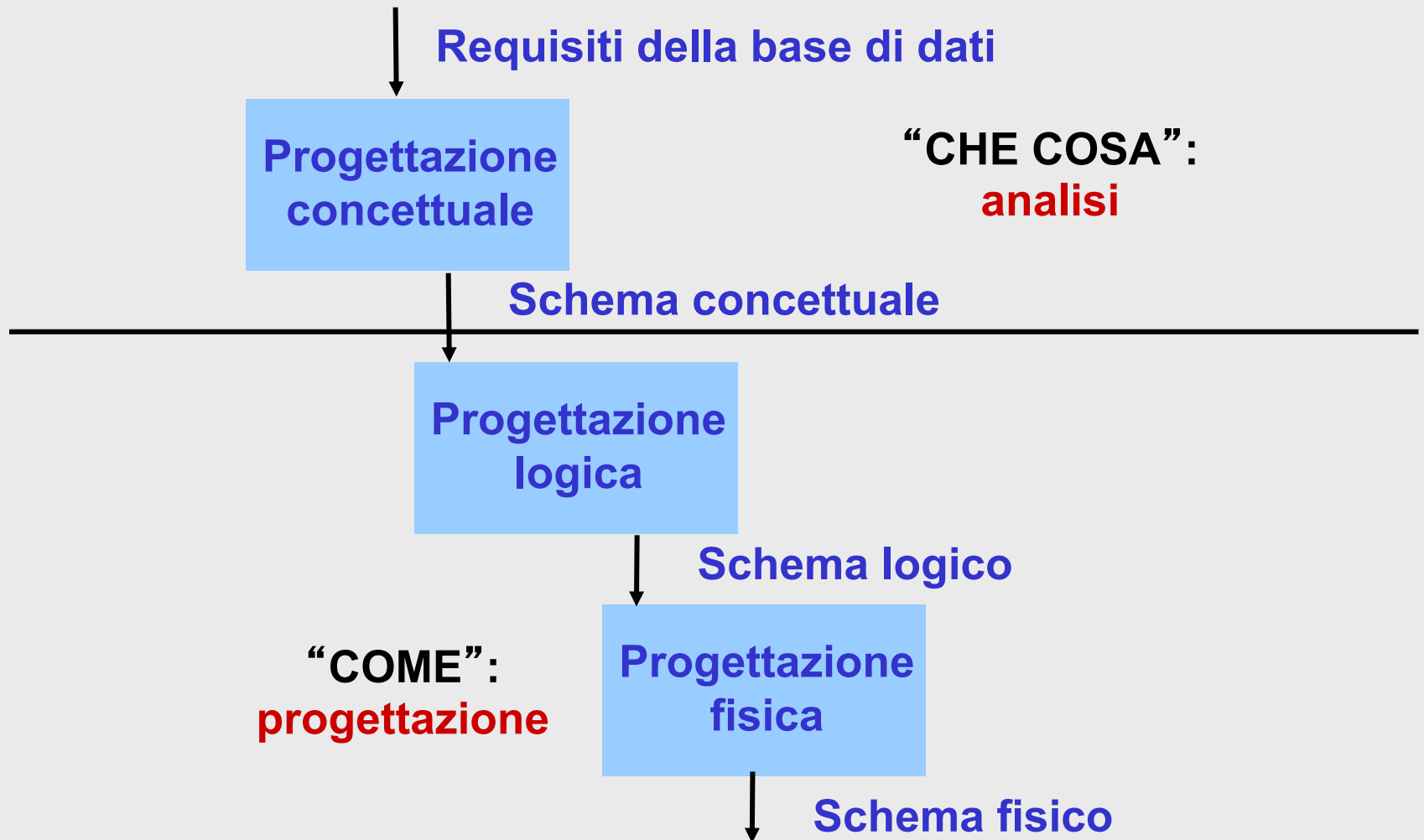
Aspetti di interesse

- La progettazione di un sistema informativo riguarda due aspetti:
 - progettazione dei dati
 - progettazione delle applicazioni
- Ma
 - questo è un corso di basi di dati
 - i dati hanno un ruolo centrale
 - i dati sono più stabili
 - ci concentriamo sui dati

- Per garantire prodotti di buona qualità è opportuno seguire una
 - **metodologia**, con:
 - articolazione delle attività in fasi
 - criteri di scelta
 - modelli di rappresentazione
 - generalità e facilità d'uso



Progettazione dei dati (Progettazione di base di dati)



I prodotti delle fasi sono schemi di alcuni **modelli di dati**:

- Schema concettuale
- Schema logico (ad esempio schema relazionale)
- Schema fisico

Modello dei dati

- insieme di costrutti ..
- ad esempio, il **modello relazionale** prevede la **relazione** (tabella)...

Due tipi (principali) di modelli

- modelli logici:
 - esempi: **relazionale**, reticolare, a oggetti, XML, noSQL, ...
- modelli concettuali:
 - il più noto è il modello **Entity-Relationship**

Due tipi (principali) di modelli

- **modelli logici**: utilizzati nei DBMS esistenti per l'organizzazione dei dati
 - utilizzati dai programmi
 - indipendenti dalle strutture fisicheesempi: **relazionale**, a oggetti
- **modelli concettuali**: permettono di rappresentare i dati in modo indipendente da ogni sistema
 - cercano di descrivere i concetti del mondo reale
 - sono utilizzati nelle fasi preliminari di progettazioneil più noto è il modello **Entity-Relationship**

Modelli concettuali, perché?

- Proviamo a modellare una applicazione definendo direttamente lo schema logico della base di dati:
 - da dove cominciamo?

Modelli concettuali, perché?

- Proviamo a modellare una applicazione definendo direttamente lo schema logico della base di dati:
 - da dove cominciamo?
 - rischiamo di perderci subito nei dettagli
 - dobbiamo pensare subito a come correlare le varie tabelle (chiavi etc.)

Modelli concettuali, perché?

- servono per ragionare sulla realtà di interesse, indipendentemente dagli aspetti realizzativi
- permettono di rappresentare le classi di oggetti di interesse e le loro correlazioni
- prevedono efficaci rappresentazioni grafiche (utili anche per documentazione e comunicazione)

Modello Entity-Relationship (Entità-Relazione)

- Il più diffuso modello concettuale
 - Ne esistono molte versioni,
 - (più o meno) diverse l'una dall'altra

I costrutti del modello E-R

- Entità
- Relationship
- Attributo
- Identificatore
- Generalizzazione
-

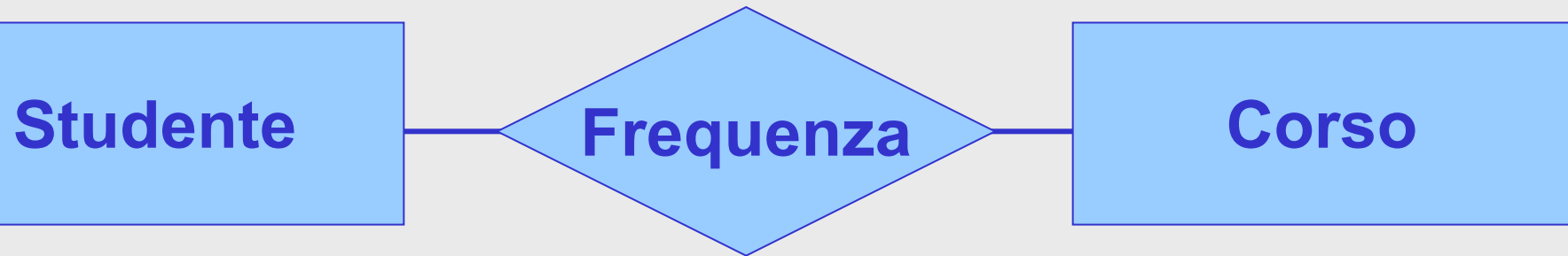
Entità

- Classe di oggetti (fatti, persone, cose) della realtà di interesse con proprietà comuni e con esistenza “autonoma” (nel senso che i singoli oggetti sono in qualche modo riconoscibili)
- Esempi:
 - impiegato, città, conto corrente, ordine, fattura

Relationship

- Legame logico fra due o più entità, rilevante nell'applicazione di interesse
- Esempi:
 - Residenza (fra persona e città)
 - Frequenza (fra studente e corso)

Uno schema E-R, graficamente



Entità

- Classe di oggetti (fatti, persone, cose) della realtà di interesse con proprietà comuni e con esistenza “autonoma” (nel senso che i singoli oggetti sono in qualche modo riconoscibili)
- Esempi:
 - impiegato, città, conto corrente, ordine, fattura

Entità: schema e istanza

- Entità:
 - classe di oggetti, persone, ... "omogenei"
- Occorrenza di entità:
 - elemento della classe (l'oggetto, la persona, ..., non i dati)
- nello schema concettuale rappresentiamo le entità, non le singole occorrenze (“astrazione”)

Rappresentazione grafica di entità

Impiegato

Dipartimento

Città

Vendita

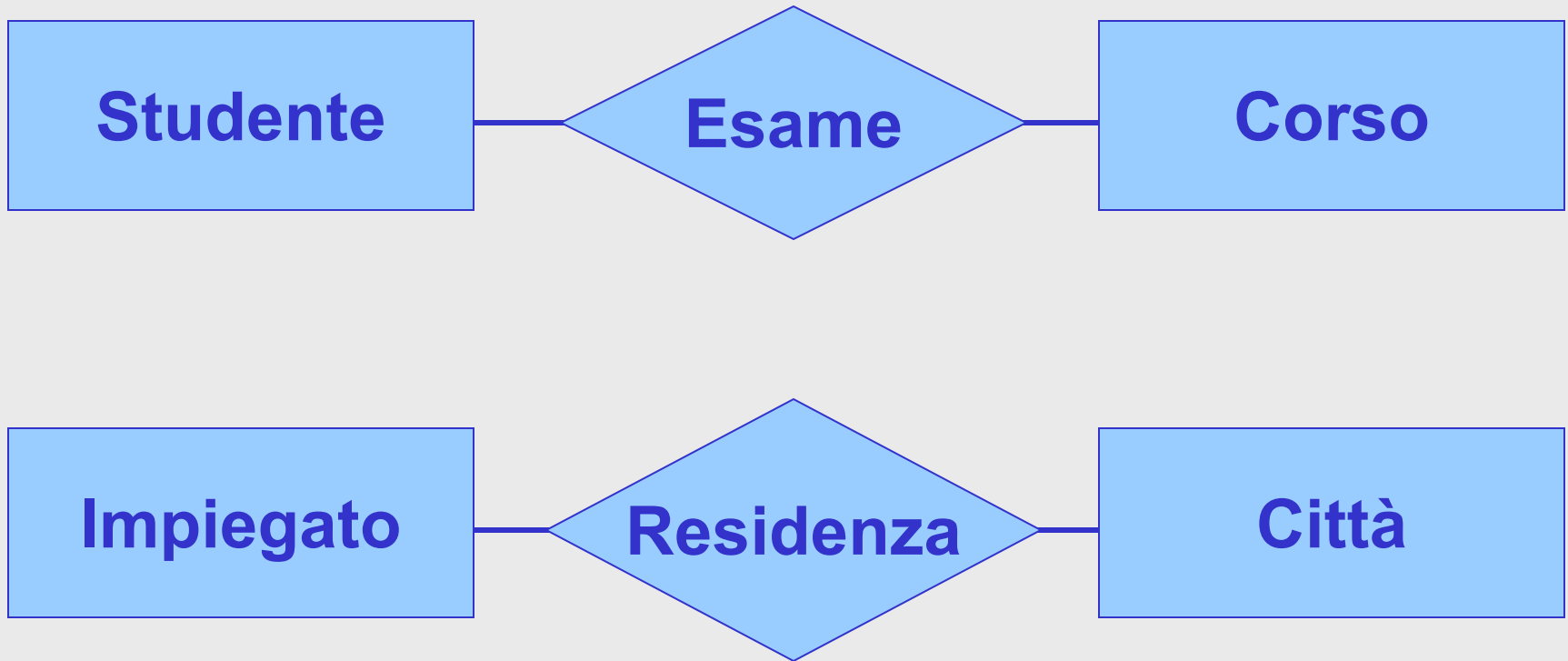
Entità, commenti

- Ogni entità ha un nome che la identifica univocamente nello schema:
 - nomi espressivi
 - opportune convenzioni
 - singolare

Relationship

- Legame logico fra due o più entità, rilevante nell'applicazione di interesse
- Esempi:
 - Residenza (fra persona e città)
 - Esame (fra studente e corso)
- Chiamata anche:
 - relazione, correlazione, associazione

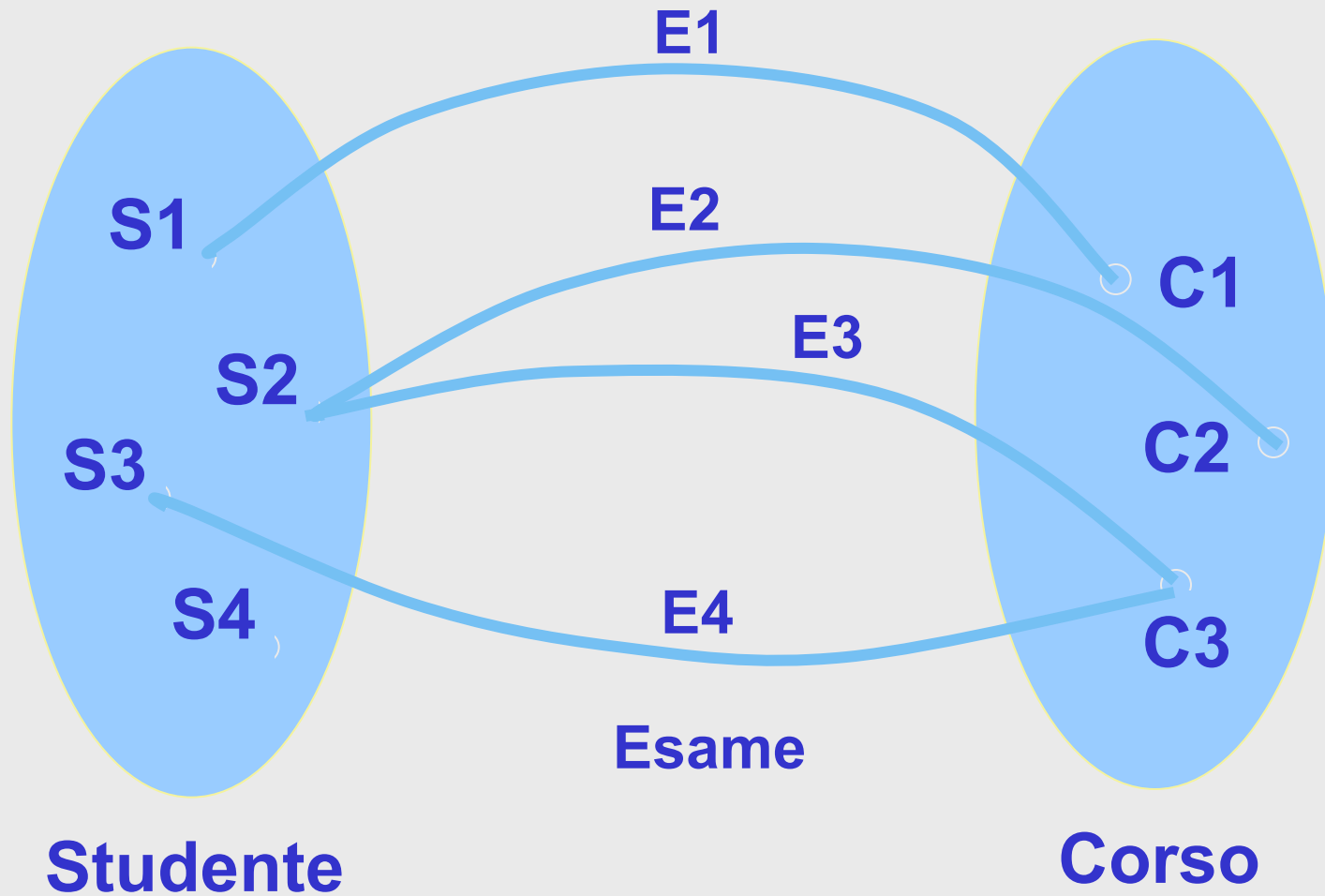
Rappresentazione grafica di relationship



Relationship, commenti

- Ogni relationship ha un nome che la identifica univocamente nello schema:
 - nomi espressivi
 - opportune convenzioni
 - singolare
 - sostantivi invece che verbi (se possibile)

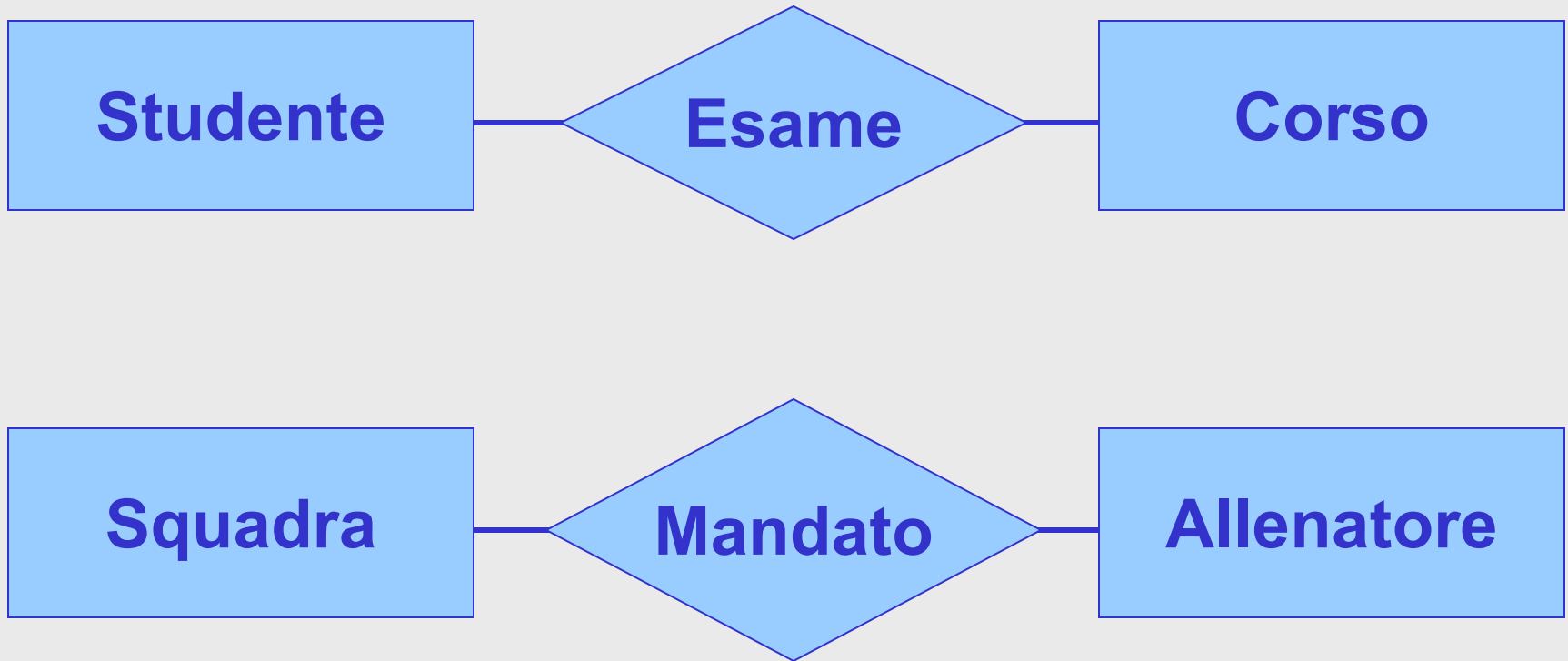
Esempi di occorrenze



Relationship, occorrenze

- Una occorrenza di una relationship binaria è coppia di occorrenze di entità, una per ciascuna entità coinvolta
- Una relationship ha un **insieme** di occorrenze e quindi
 - nell'ambito di una relationship non ci possono essere occorrenze (coppie) ripetute

Relationship corrette?

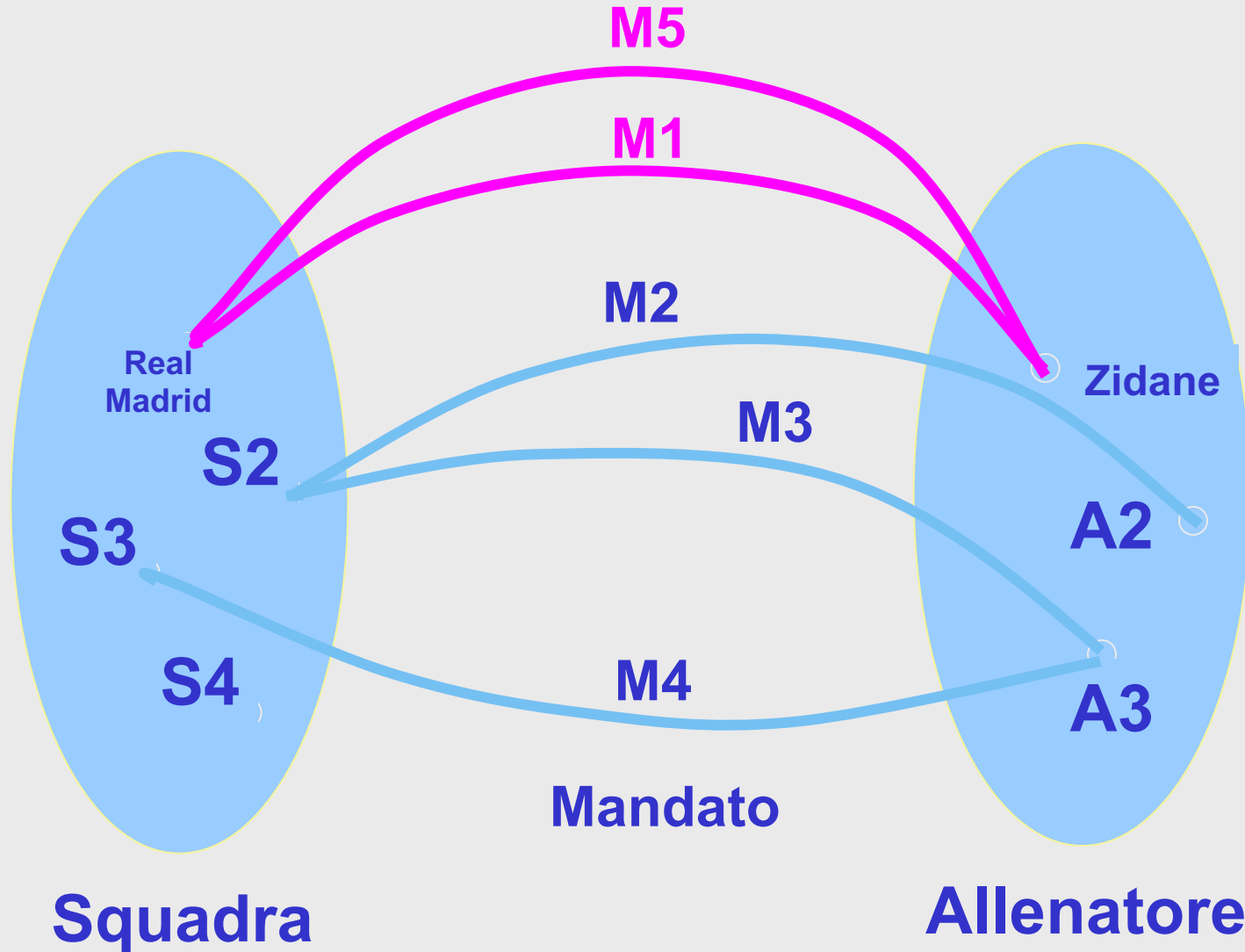


Gli allenatori del Real Madrid

https://en.wikipedia.org/wiki/List_of_Real_Madrid_CF_managers

											1 Supercopa de España	
Carlo Ancelotti		25 June 2013	25 May 2015	119	89	14	16	323	103	74.79	1 Copa del Rey 1 UEFA Champions League 1 UEFA Super Cup 1 FIFA Club World Cup	
Rafael Benítez		3 June 2015	4 January 2016	25	17	5	3	69	22	68.00		
Zinedine Zidane		4 January 2016	31 May 2018	149	104	29	16	393	160	69.80	1 La Liga 1 Supercopa de España 3 UEFA Champions League 2 UEFA Super Cup 2 FIFA Club World Cup	[C][FIFA]
Julen Lopetegui		12 June 2018 ^[3]	29 October 2018	14	6	2	6	21	20	42.86		[C]
Santiago Solari		30 October 2018	11 March 2019	32	22	2	8	71	37	68.75	1 FIFA Club World Cup	[C]
Zinedine Zidane		11 March 2019 ^[4]	27 May 2021	114	69	25	20	207	104	60.53	1 La Liga 1 Supercopa de España	[C]
Carlo Ancelotti		1 June 2021 ^[5]	present	12	8	2	2	31	12	66.67		

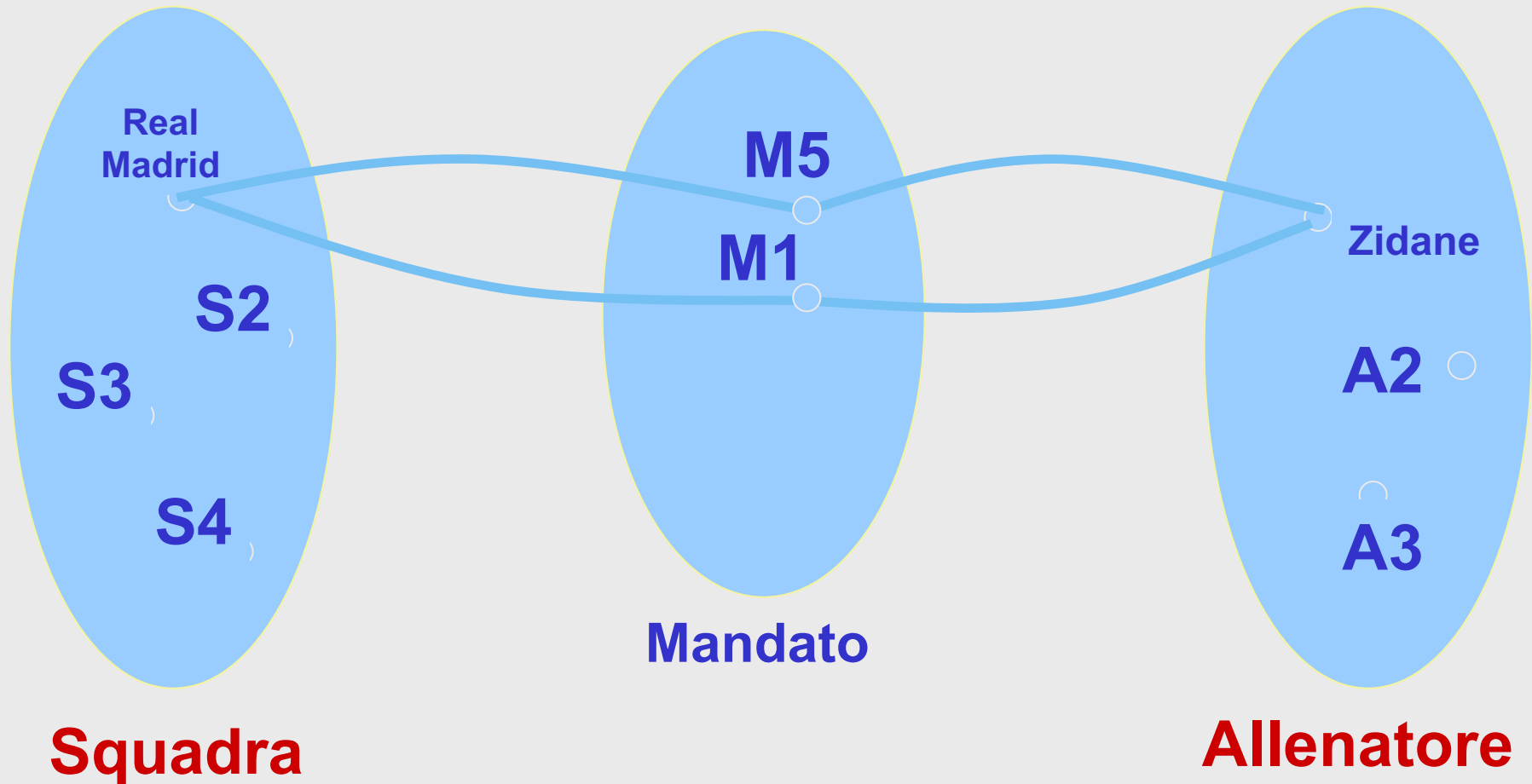
Attenzione



"Promuoviamo" la relationship



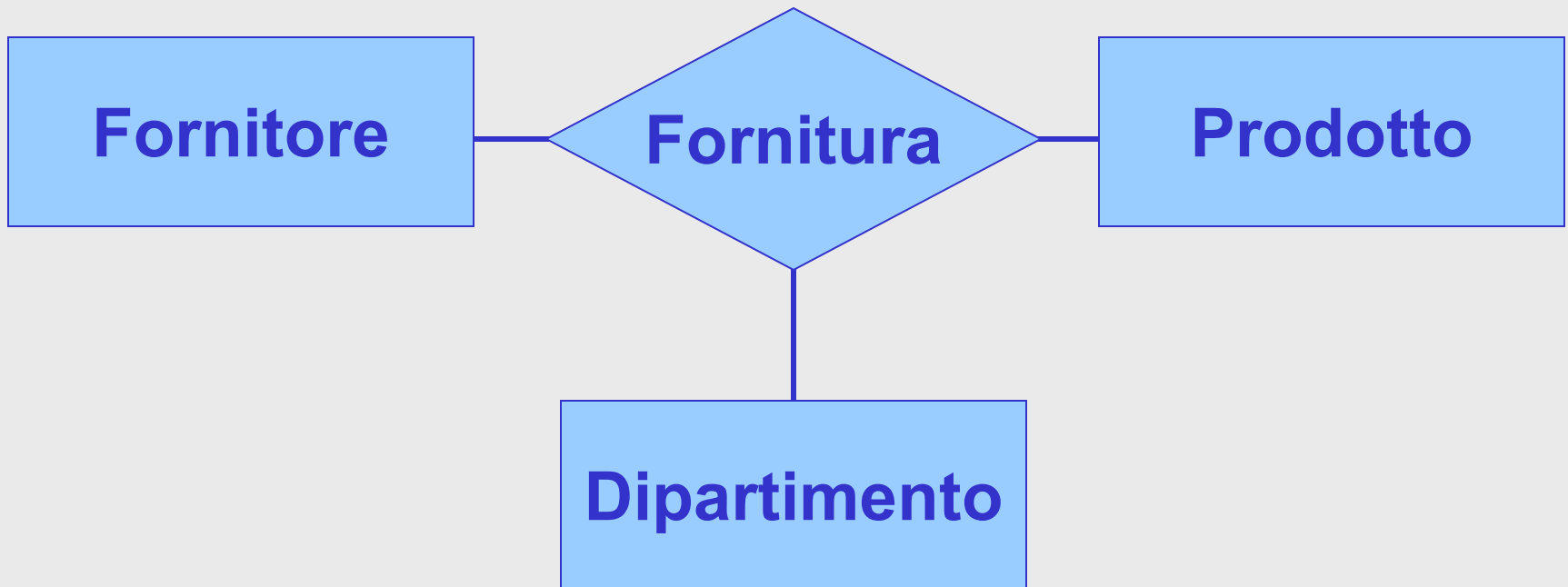
Con l'entità Mandato



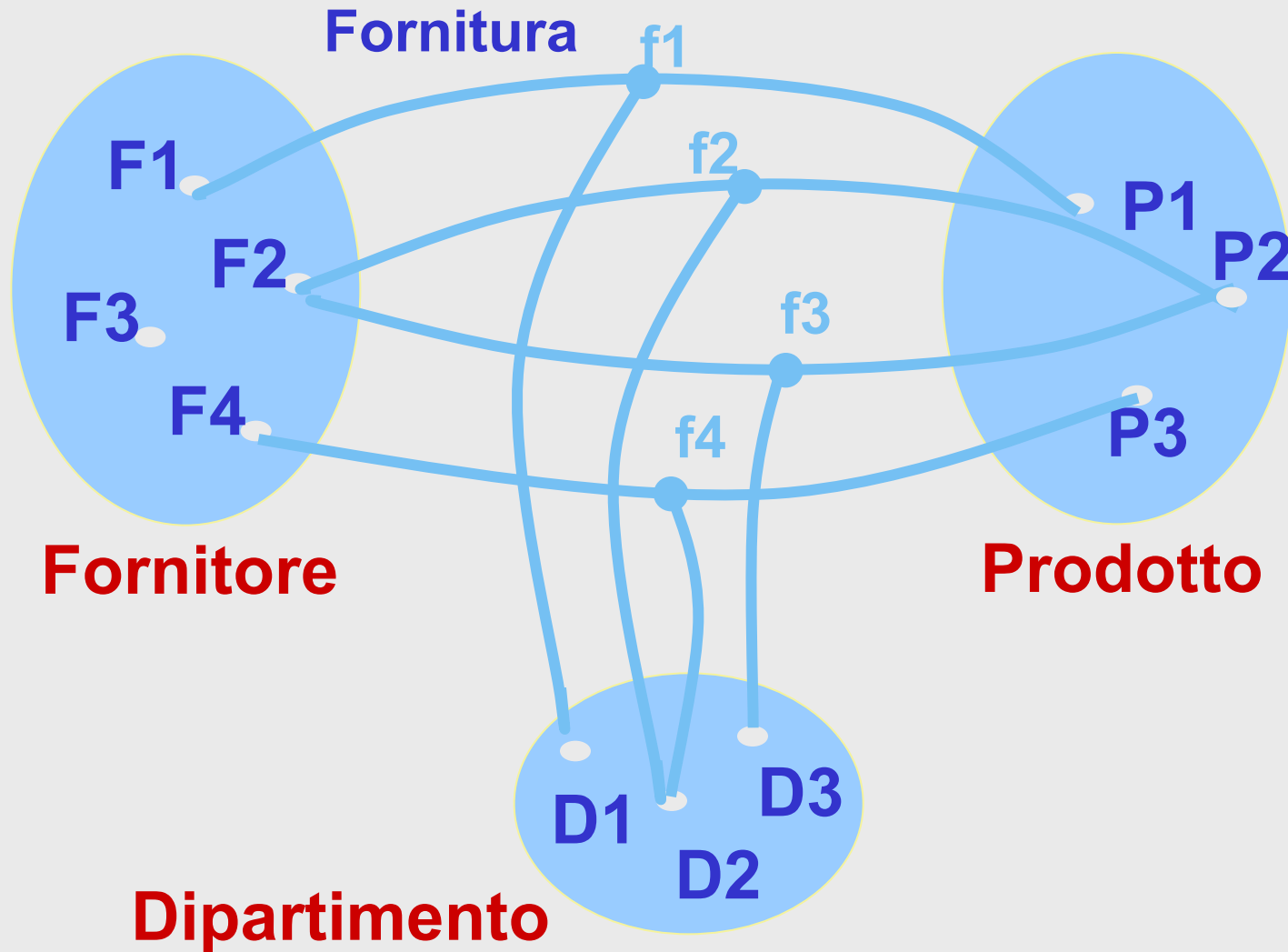
Relationship n-arie, occorrenze

- Una occorrenza di una relationship n-aria è una n-upla di occorrenze di entità, una per ciascuna entità coinvolta
- Nell'ambito di una relationship non ci possono essere occorrenze (n-uple) ripetute

Relationship n-aria



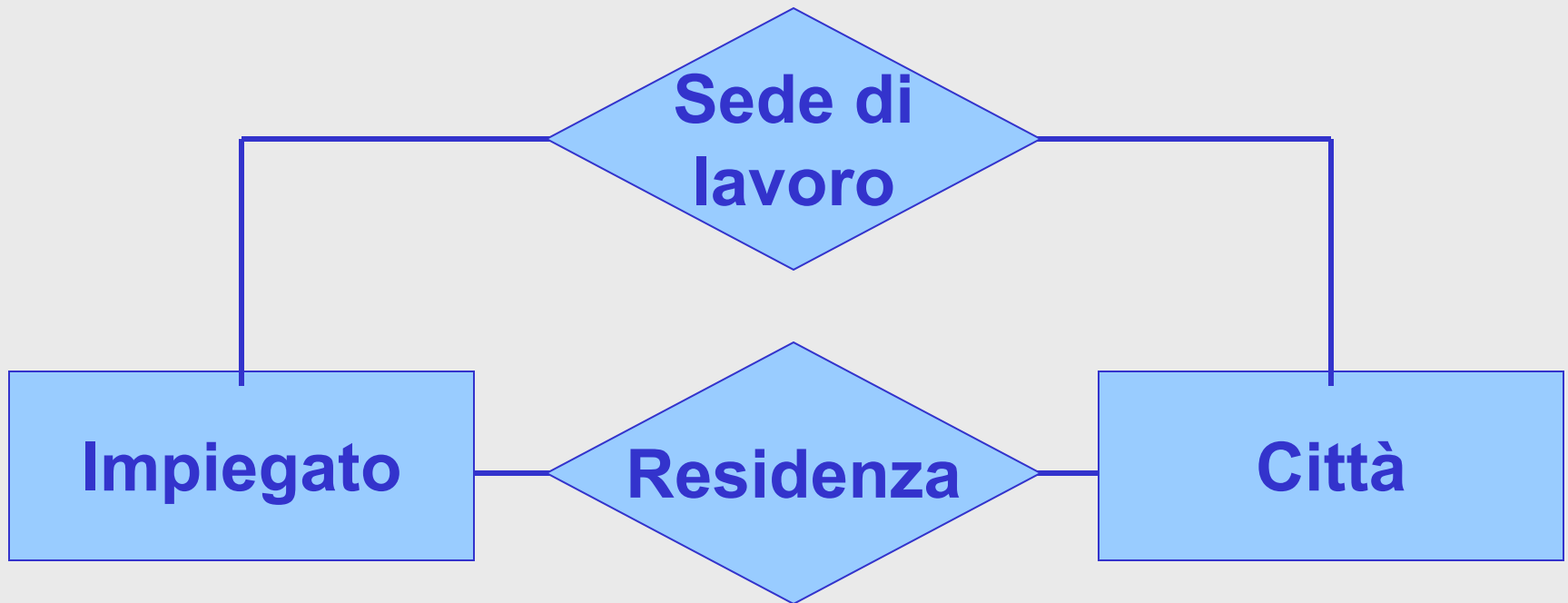
Esempi di occorrenze



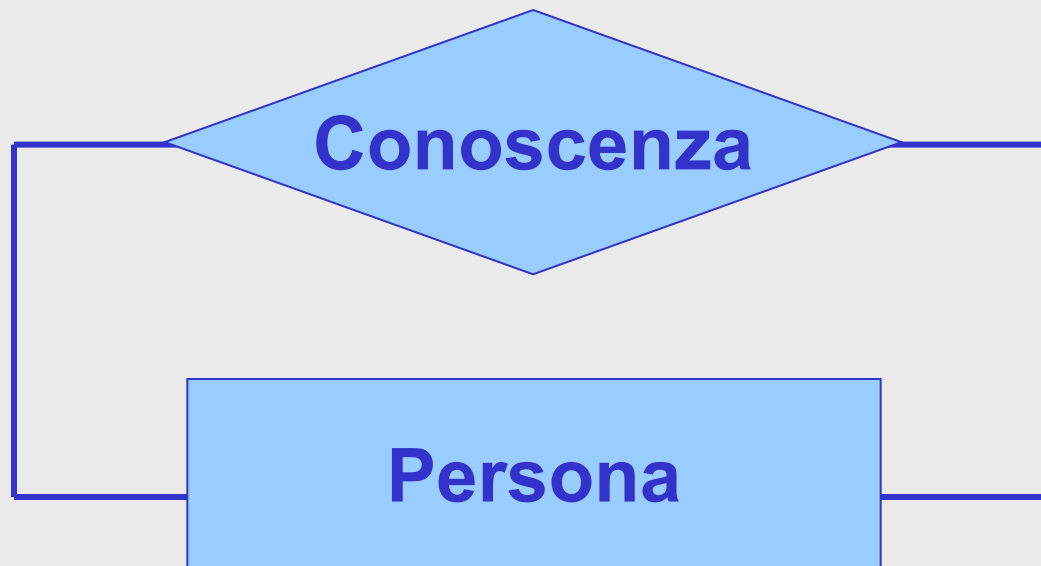
Relationship n-aria

- Quando è corretto lo schema precedente?
- Che cosa si intende per “fornitura”?

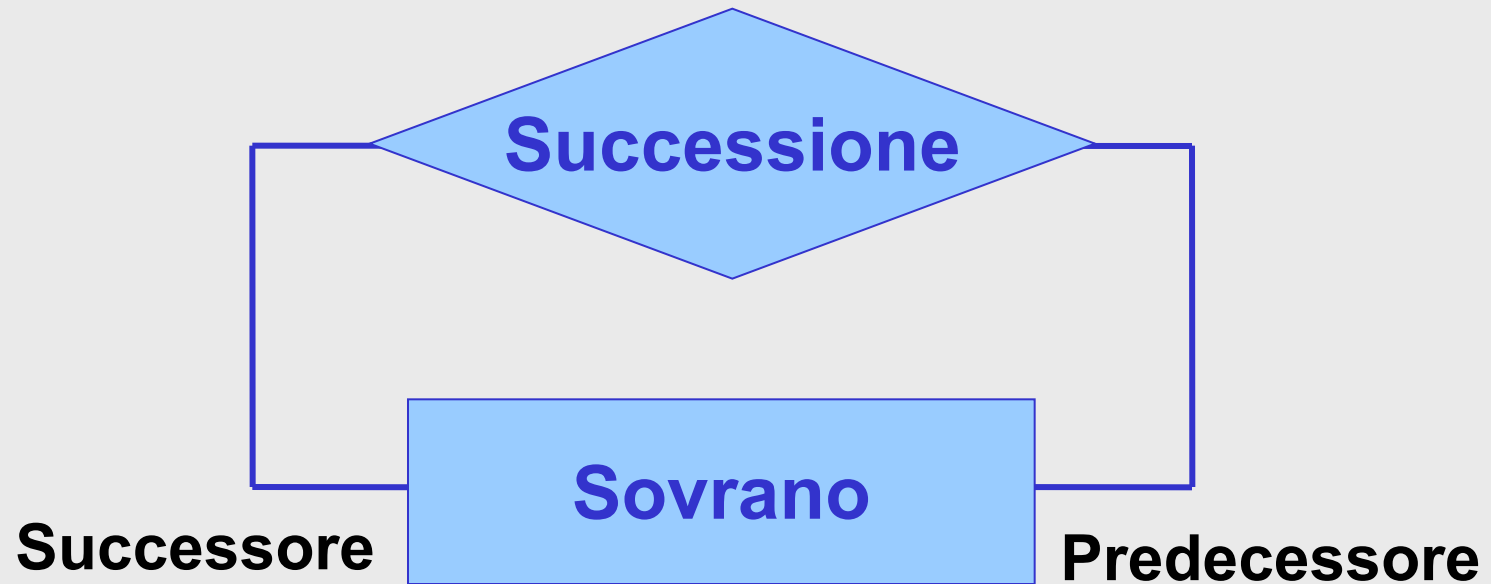
Due relationship sulle stesse entità



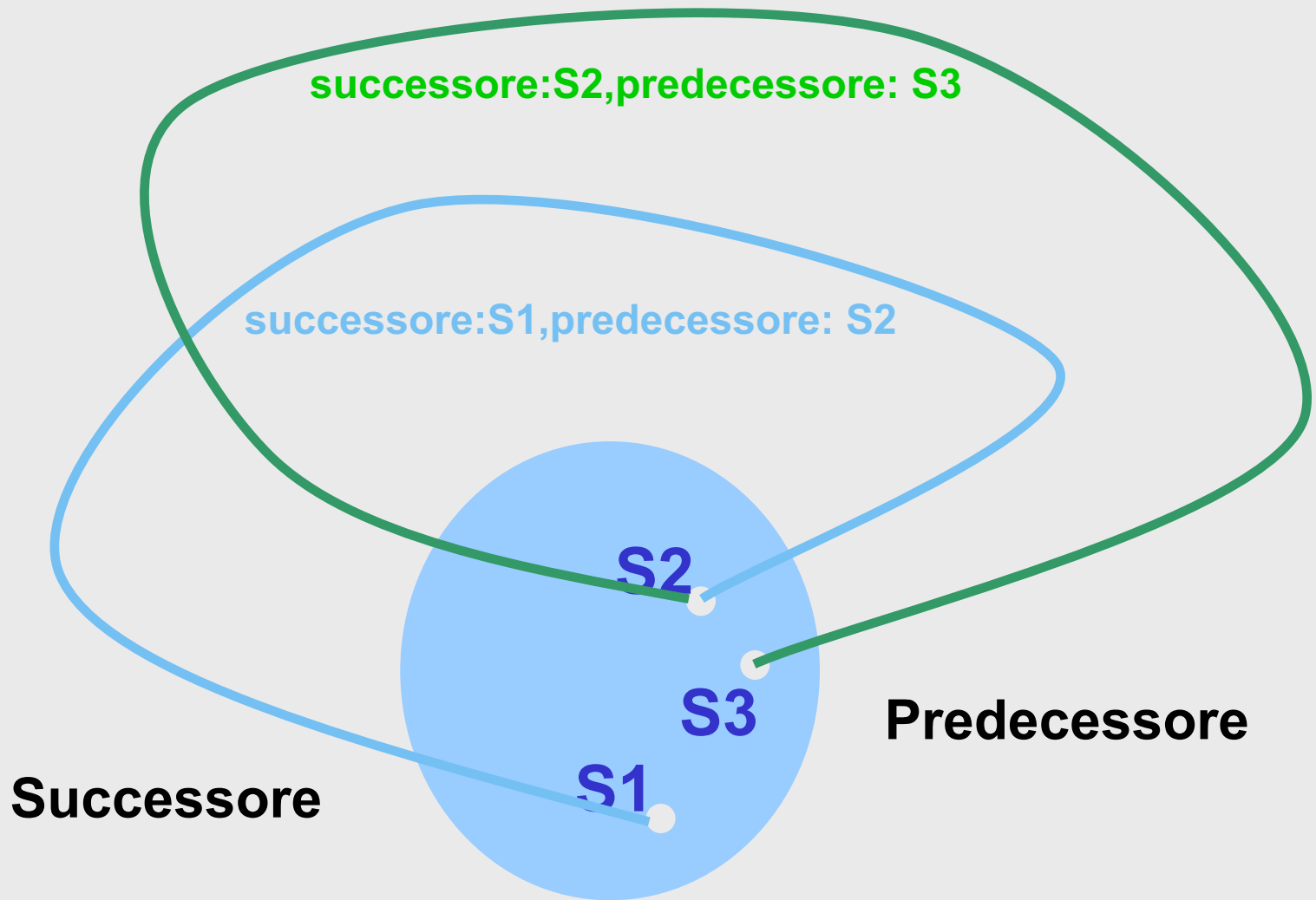
Relationship ricorsiva: coinvolge “due volte” la stessa entità



Relationship ricorsiva con “ruoli”



Esempi di occorrenze



Relationship ricorsiva

- “Successione” ... è corretta?

Edoardo IV



Ritratto di Edoardo IV

Re d'Inghilterra e Signore d'Irlanda



In carica 4 marzo 1461 -
30 ottobre 1470

Incoronazione 28 giugno 1461

Predecessore Enrico VI

Successore Enrico VI

Re d'Inghilterra e Signore d'Irlanda

In carica 11 aprile 1471 -
9 aprile 1483

Predecessore Enrico VI

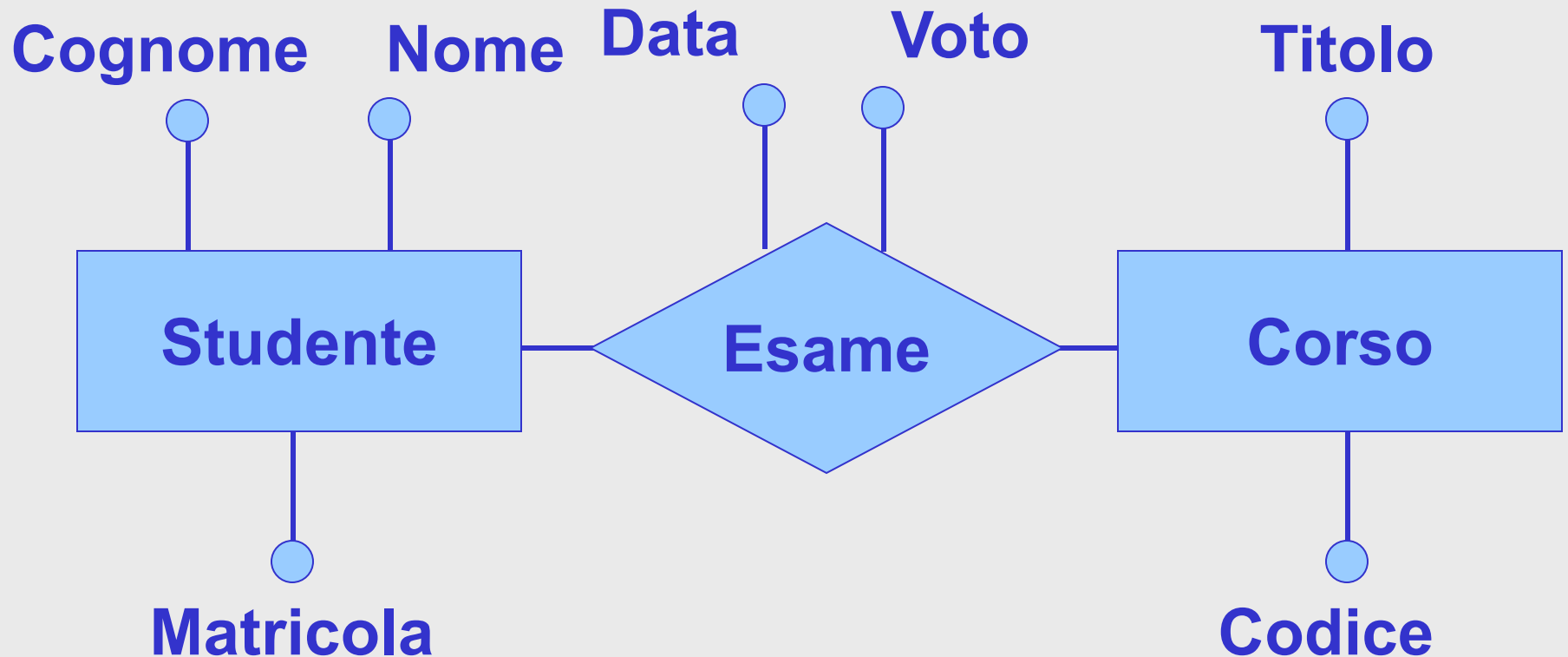
Successore Edoardo V



Attributo

- Proprietà elementare di un' entità o di una relationship, di interesse ai fini dell' applicazione
- Associa ad ogni occorrenza di entità o relationship un valore appartenente a un insieme detto **dominio** dell' attributo

Attributi, rappresentazione grafica

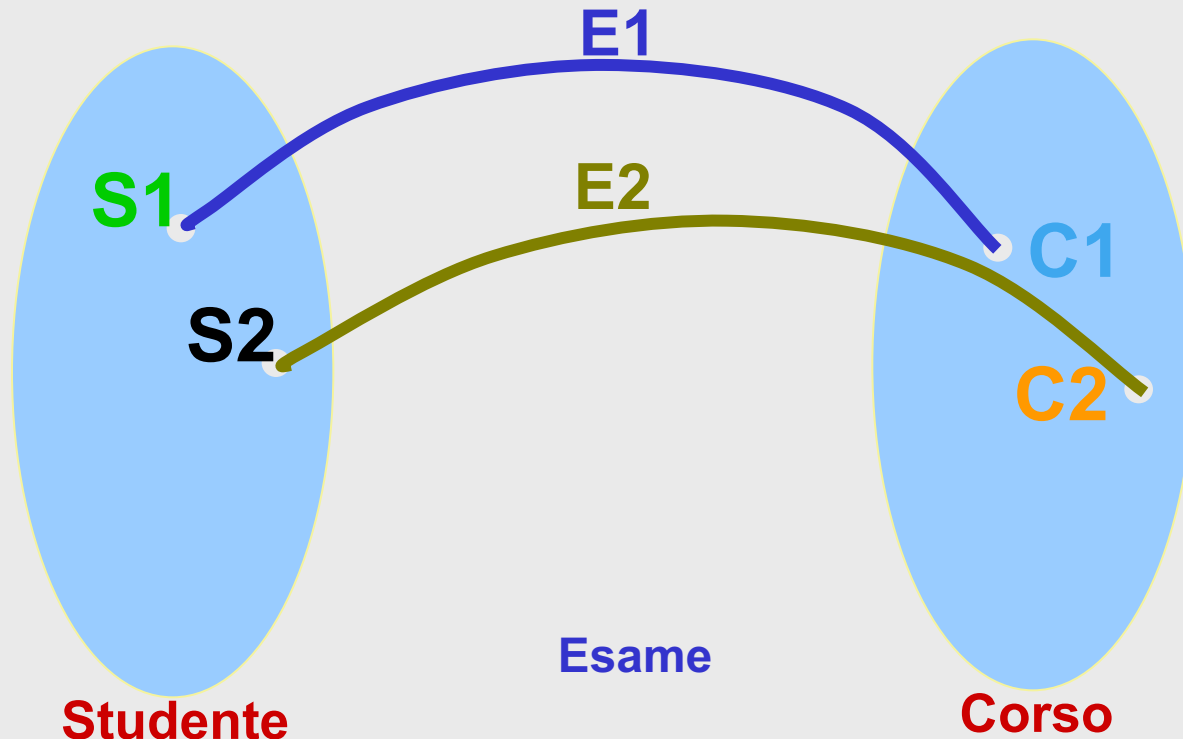


Esempi di occorrenze

Matricola: 34567
Cognome: Rossi
Nome: Mario

Data: 25/07/2004
Voto: 26

Codice: Inf205
Titolo: Basi di dati

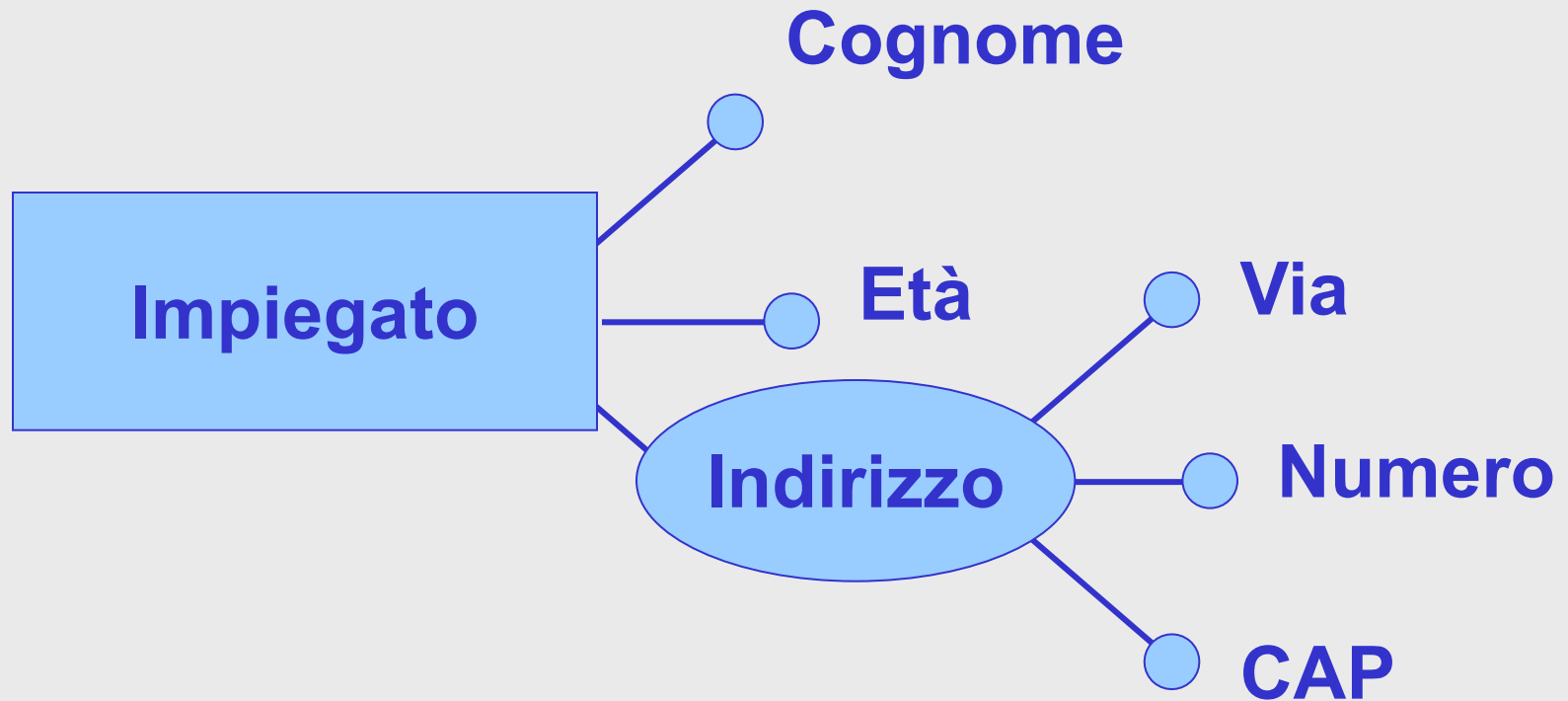


Matricola: 46742
Cognome: Neri
Nome: Piero

Attributi composti

- Raggruppano attributi di una medesima entità o relationship che presentano affinità nel loro significato o uso
- Esempio:
 - Via, Numero civico e CAP formano un Indirizzo
- Nota bene: in questo corso li utilizziamo poco (quasi per niente)

Rappresentazione grafica





Altri costrutti del modello E-R

- Cardinalità
 - di relationship
 - di attributo
- Identificatore
 - interno
 - esterno
- Generalizzazione

Cardinalità di relationship

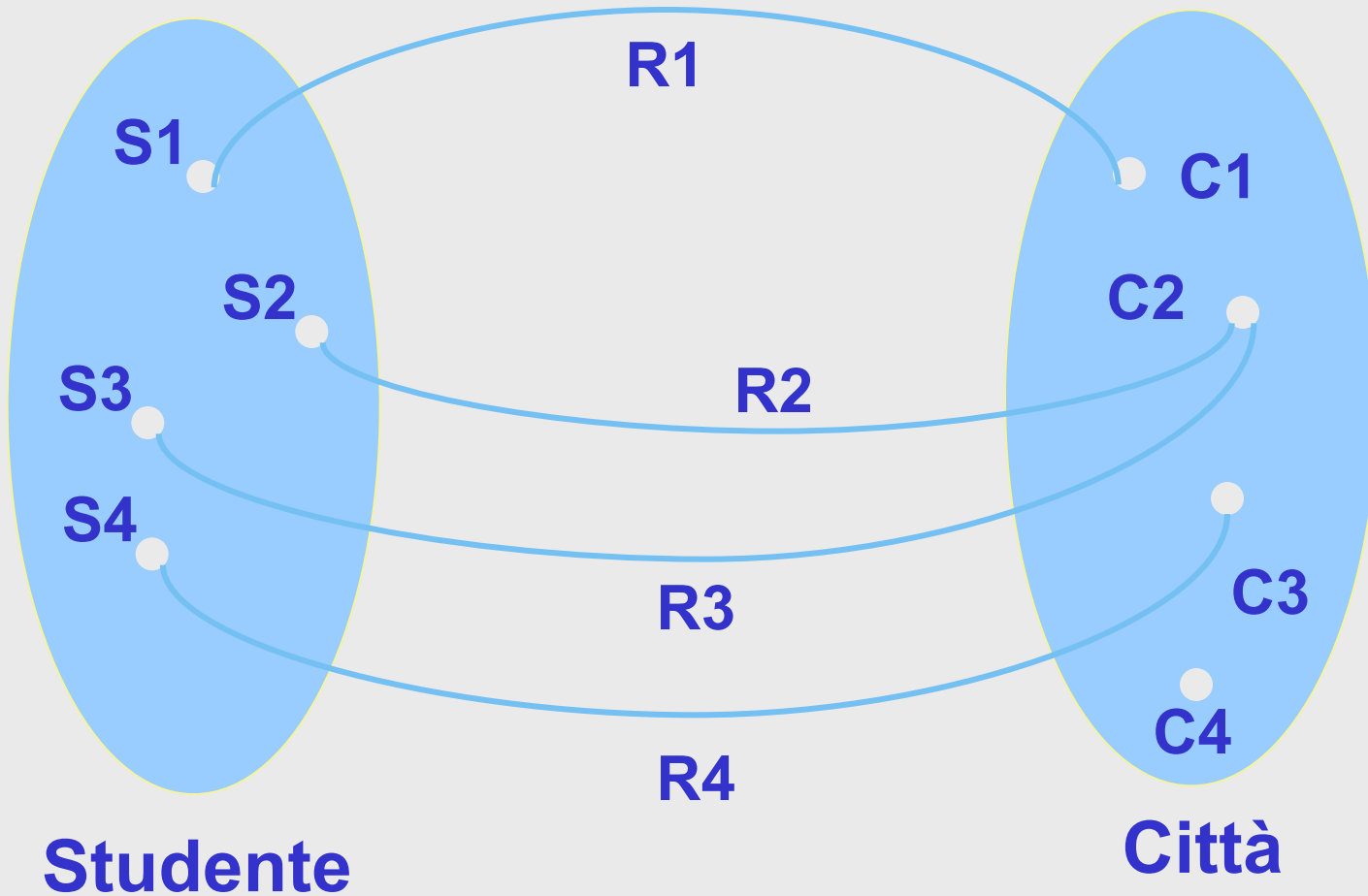
- Coppia di valori associati a ogni entità che partecipa a una relationship
- specificano il numero minimo e massimo di occorrenze delle relationship cui ciascuna occorrenza di una entità può partecipare

Esempio di cardinalità

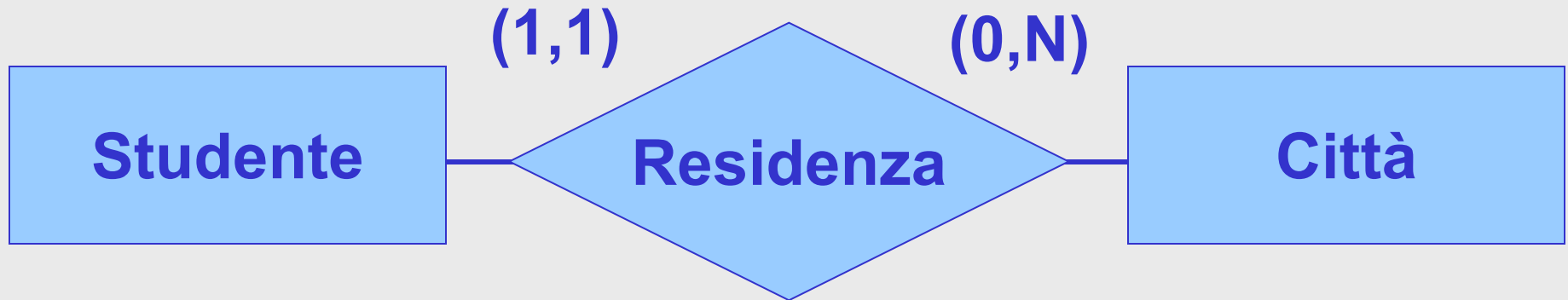


- per semplicità usiamo solo tre simboli:
- 0 e 1 per la cardinalità minima:
 - 0 = “partecipazione **opzionale**”
 - 1 = “partecipazione **obbligatoria**”
- 1 e “N” per la massima:
 - “N” non pone alcun limite

Occorrenze di Residenza

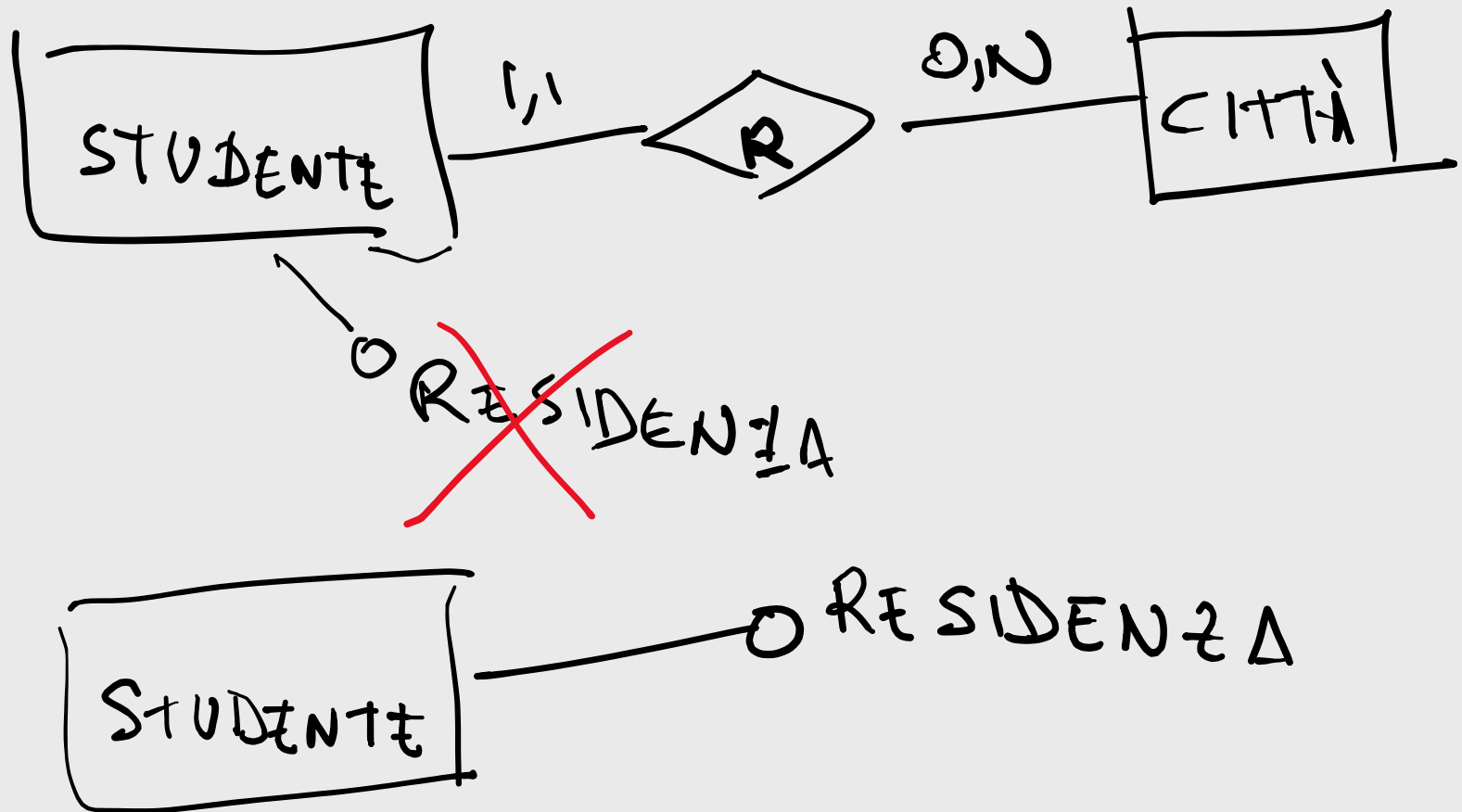


Cardinalità di Residenza



Relationship "uno a molti"

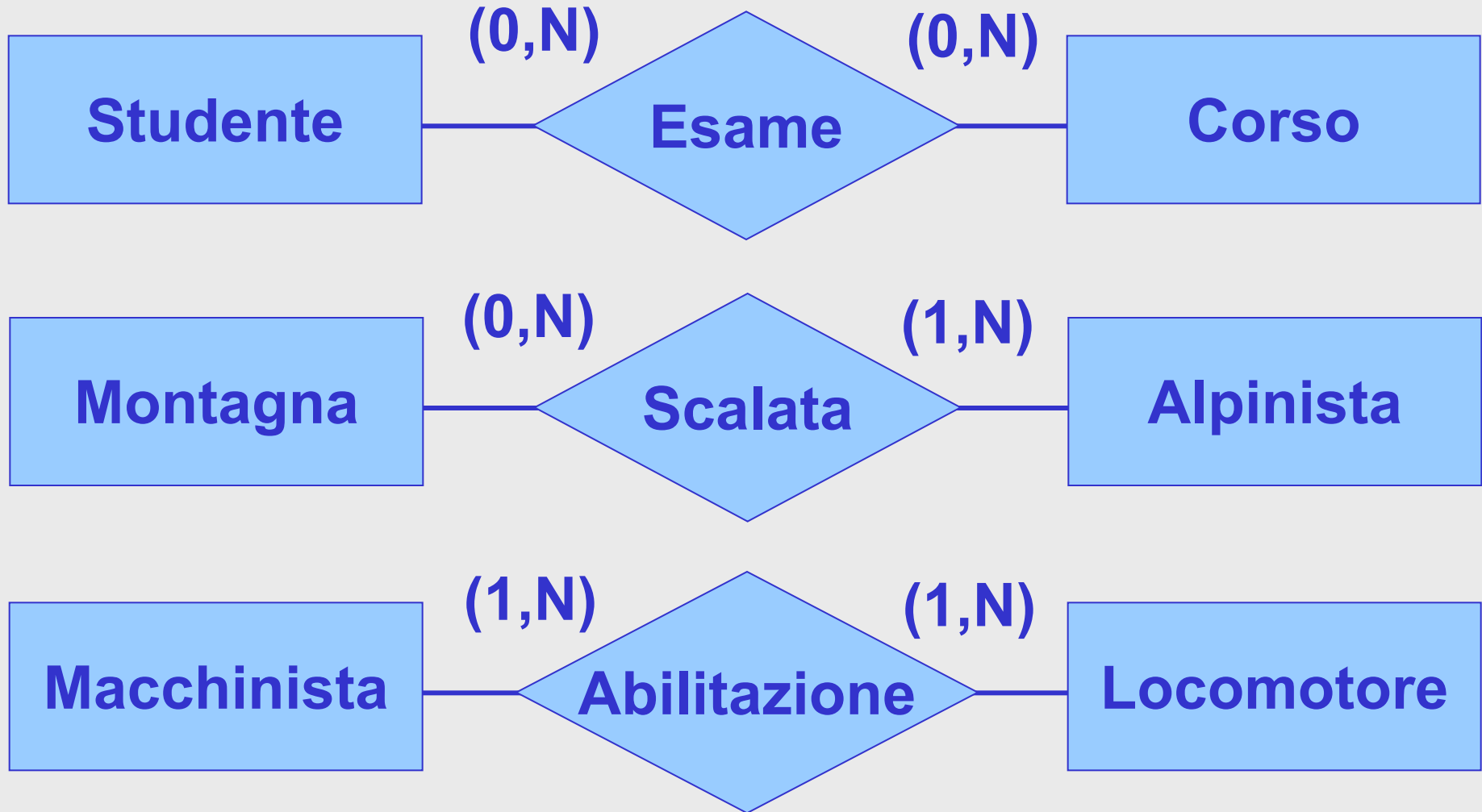
- La città di residenza è una "proprietà" dello studente
- Un attributo "reificato", cioè promosso a entità perché "ci interessa" come concetto
- Probabilmente perché ha, a sua volta, proprietà (nome, provincia, ...)
- Se c'è la relationship, allora non c'è l'entità



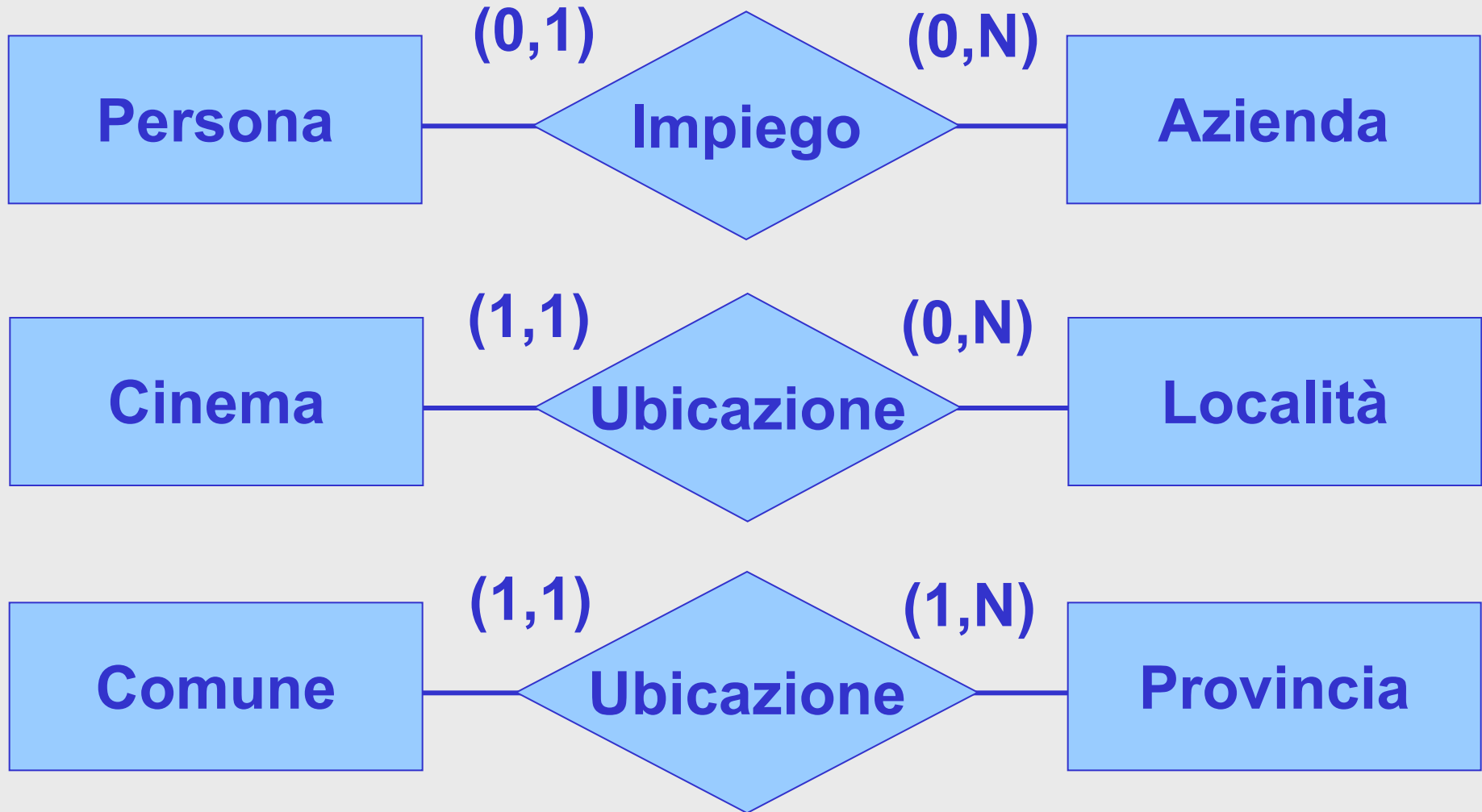
Tipi di relationship

- Con riferimento alle cardinalità **massime**, abbiamo relationship:
 - uno a uno
 - uno a molti
 - molti a molti

Relationship “molti a molti”

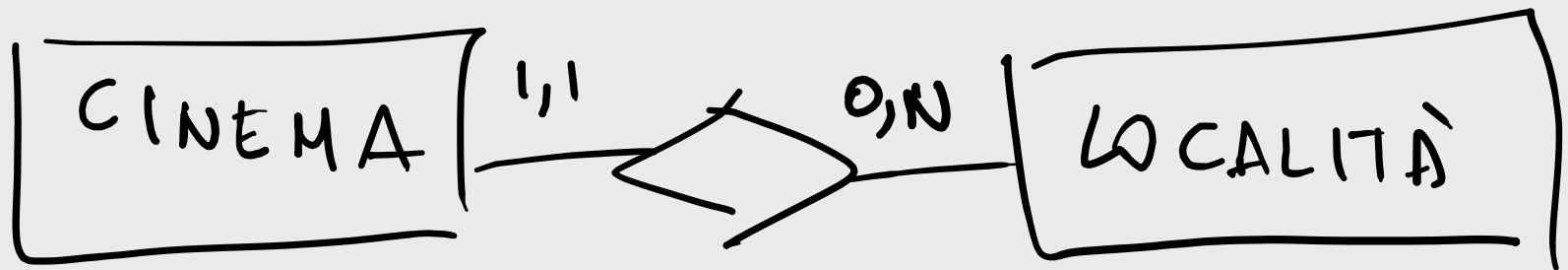


Relationship “uno a molti”



Due avvertenze

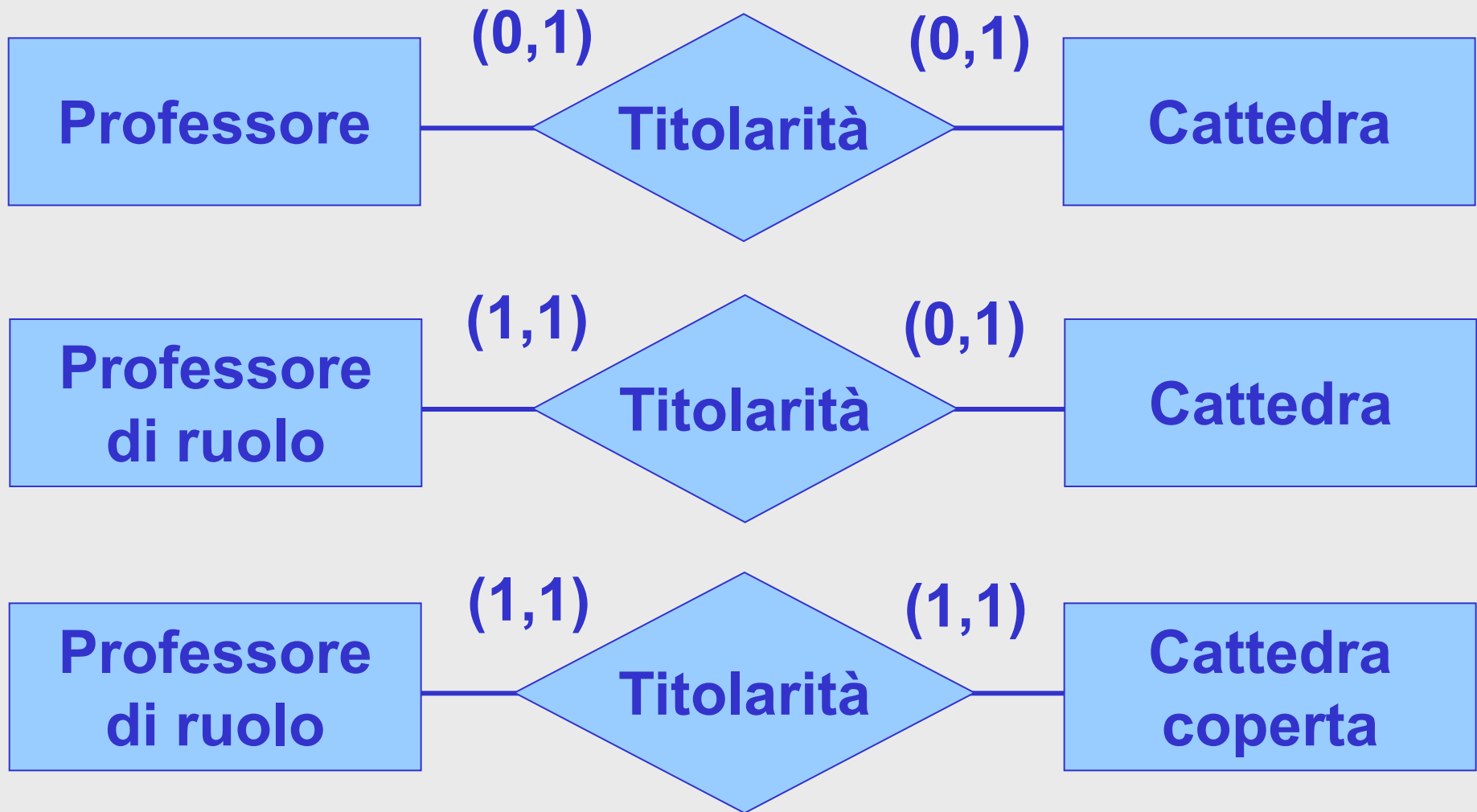
- Le relationship obbligatorie-obbligatorie sono molto rare
- Attenzione al "verso" nelle relationship uno a molti
 - esiste anche un'altra notazione



ALTRA
NOTAZIONE (MEGLIO NON
UTILIZZARLA)



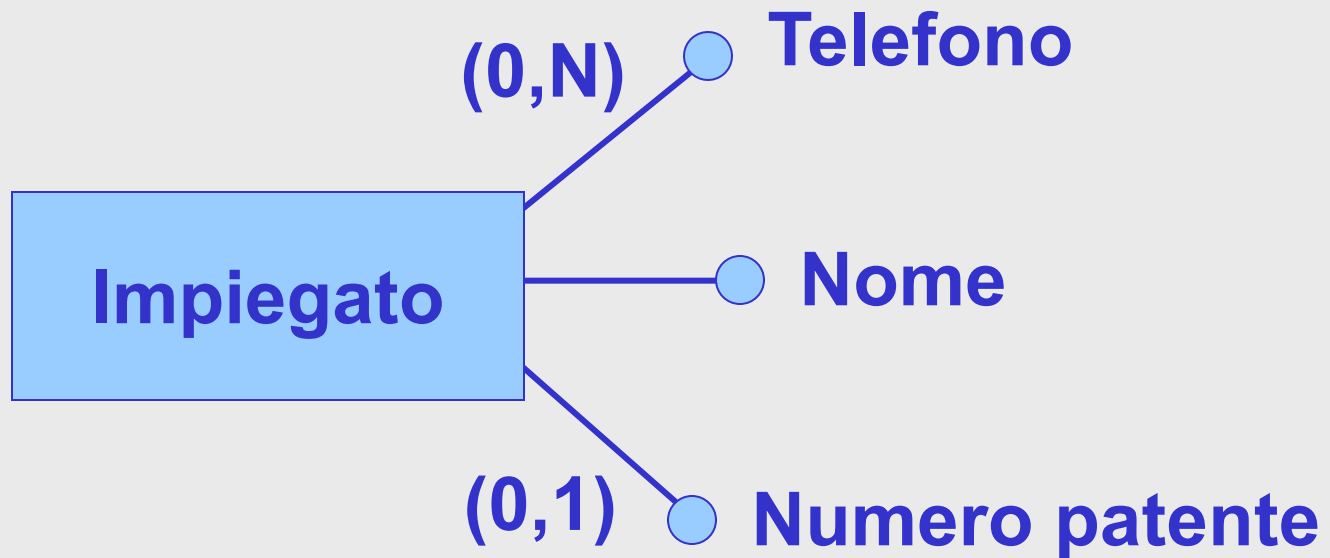
Relationship “uno a uno”



Cardinalità di attributi

- E' possibile associare delle cardinalità anche agli attributi, con due scopi:
 - indicare opzionalità ("informazione incompleta")
 - indicare attributi multivalore

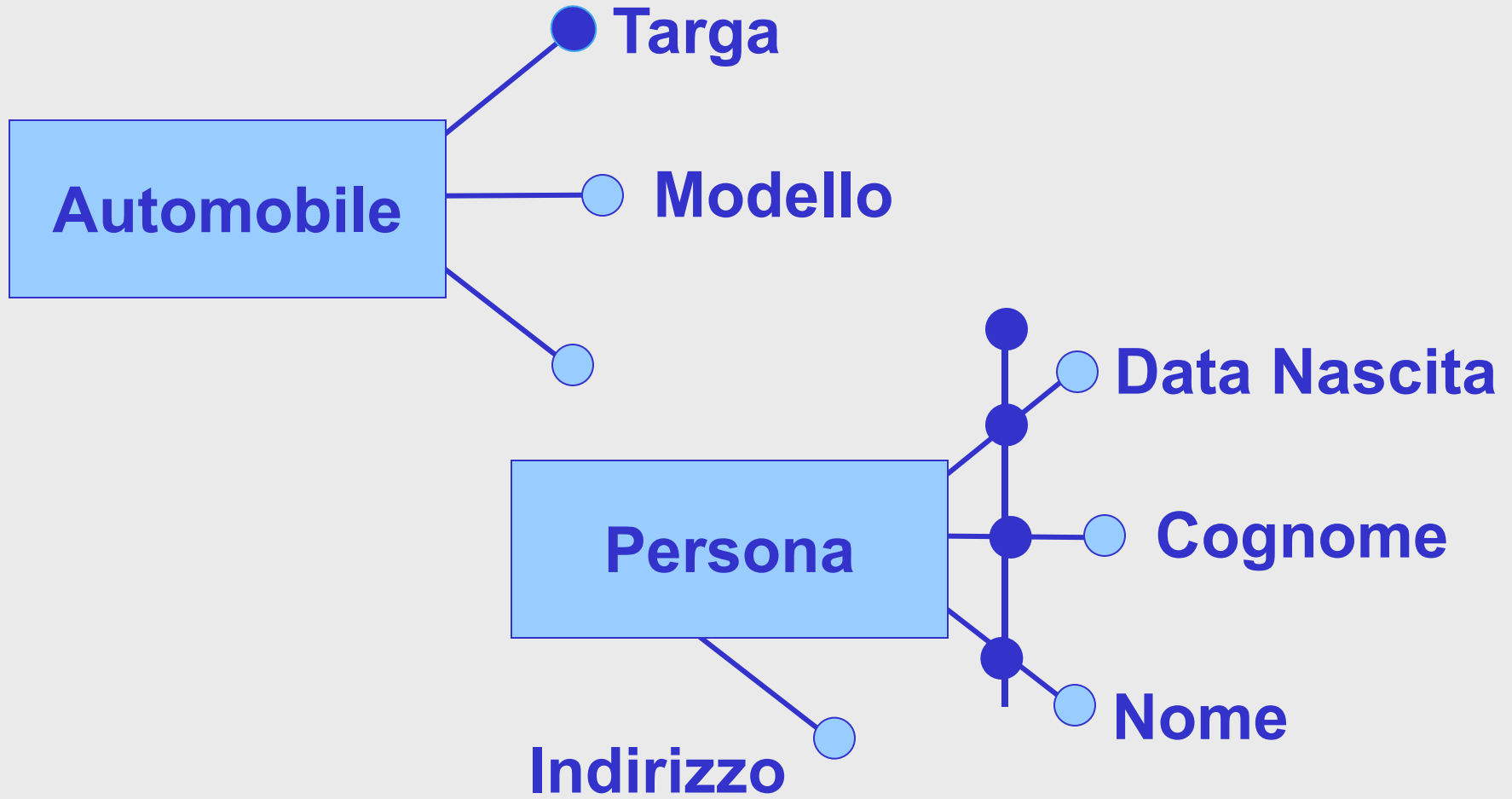
Rappresentazione grafica



Identificatore di una entità

- “strumento” per l’ identificazione univoca delle occorrenze di un’ entità
- costituito da:
 - attributi dell’ entità
 - identificatore interno
 - (attributi +) entità esterne attraverso relationship
 - identificatore esterno

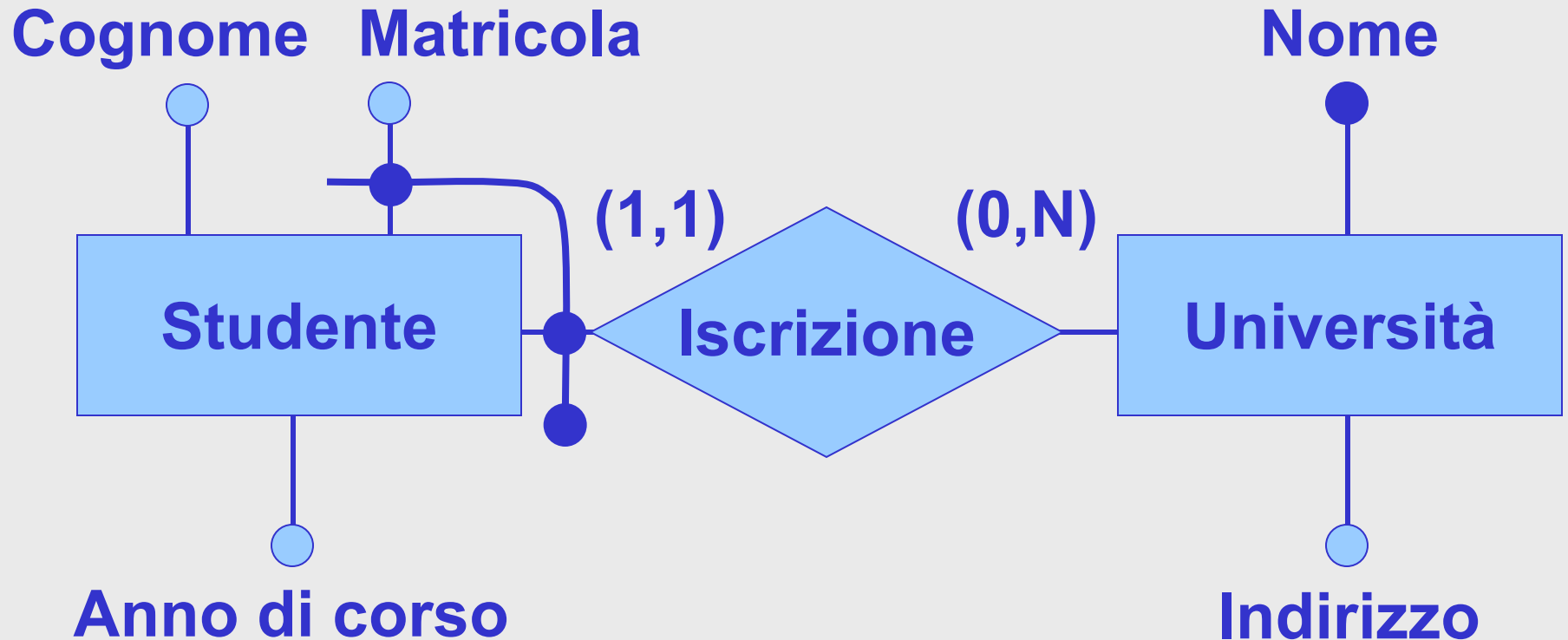
Identificatori interni



L'identificatore interno può non bastare

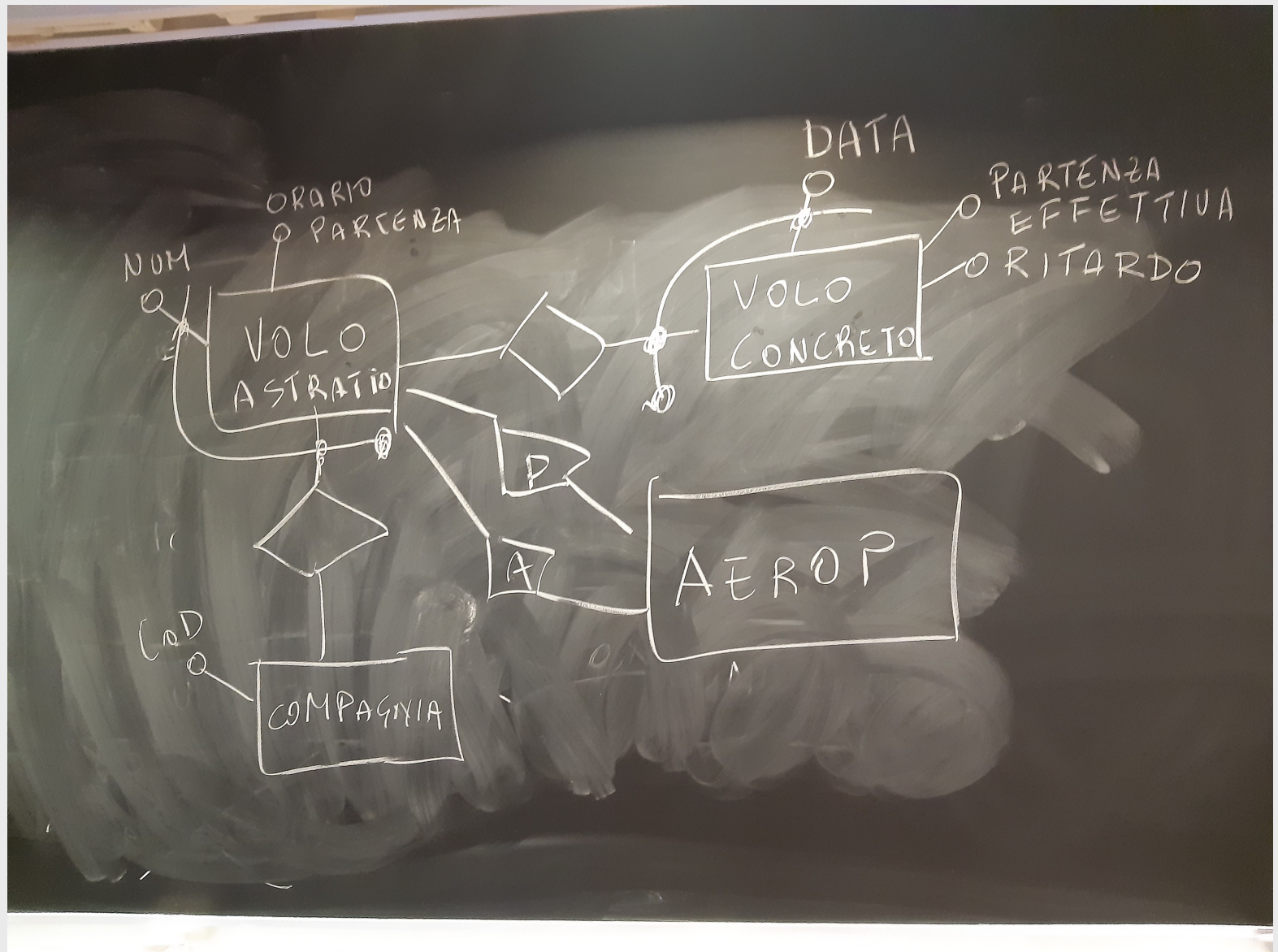
- Anagrafe nazionale degli studenti
- Ogni studente ha una matricola
- Basta la matricola per identificare?

Identificatore esterno



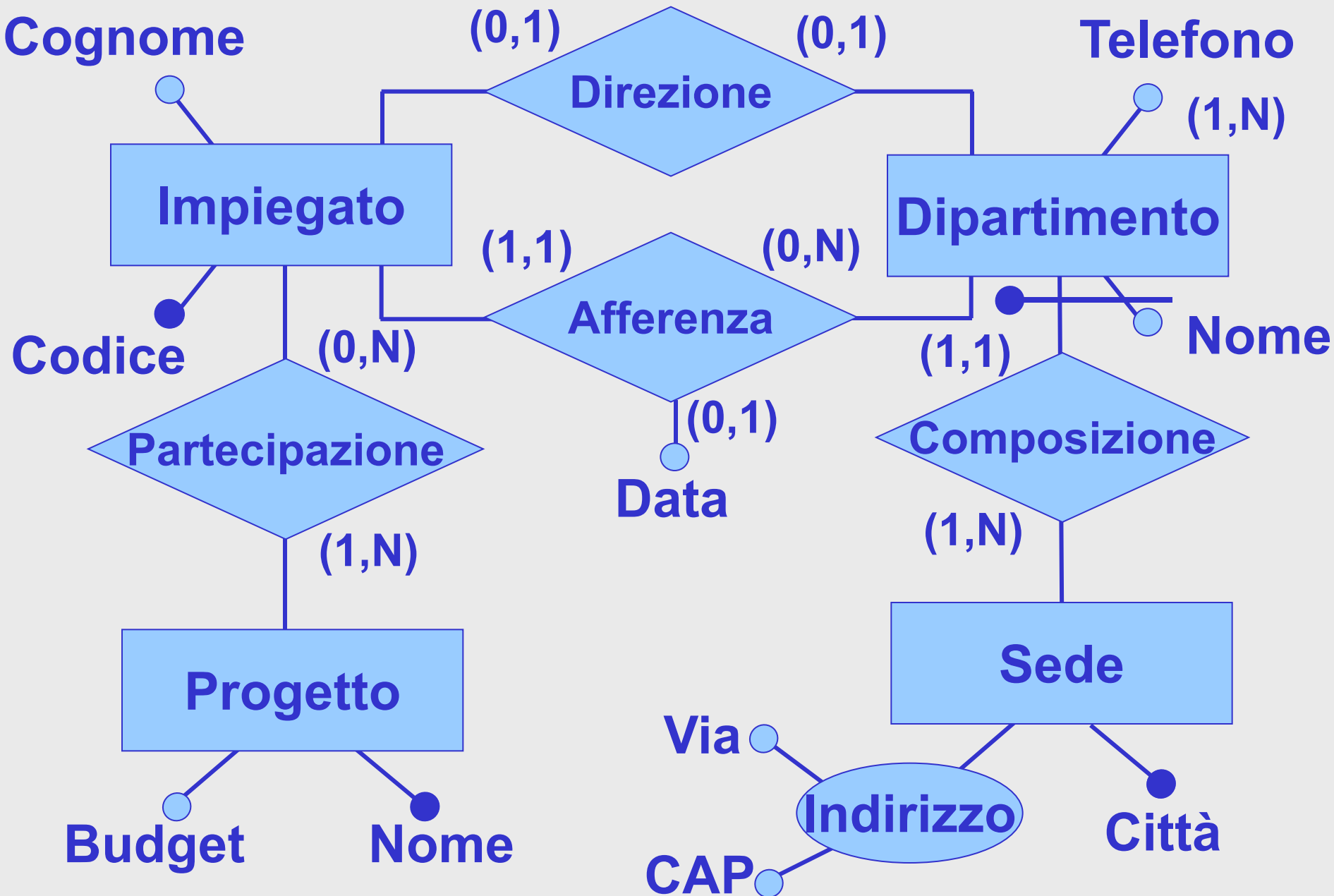
Identificatore esterno molto comune

- Volo aereo
- Identificato da?



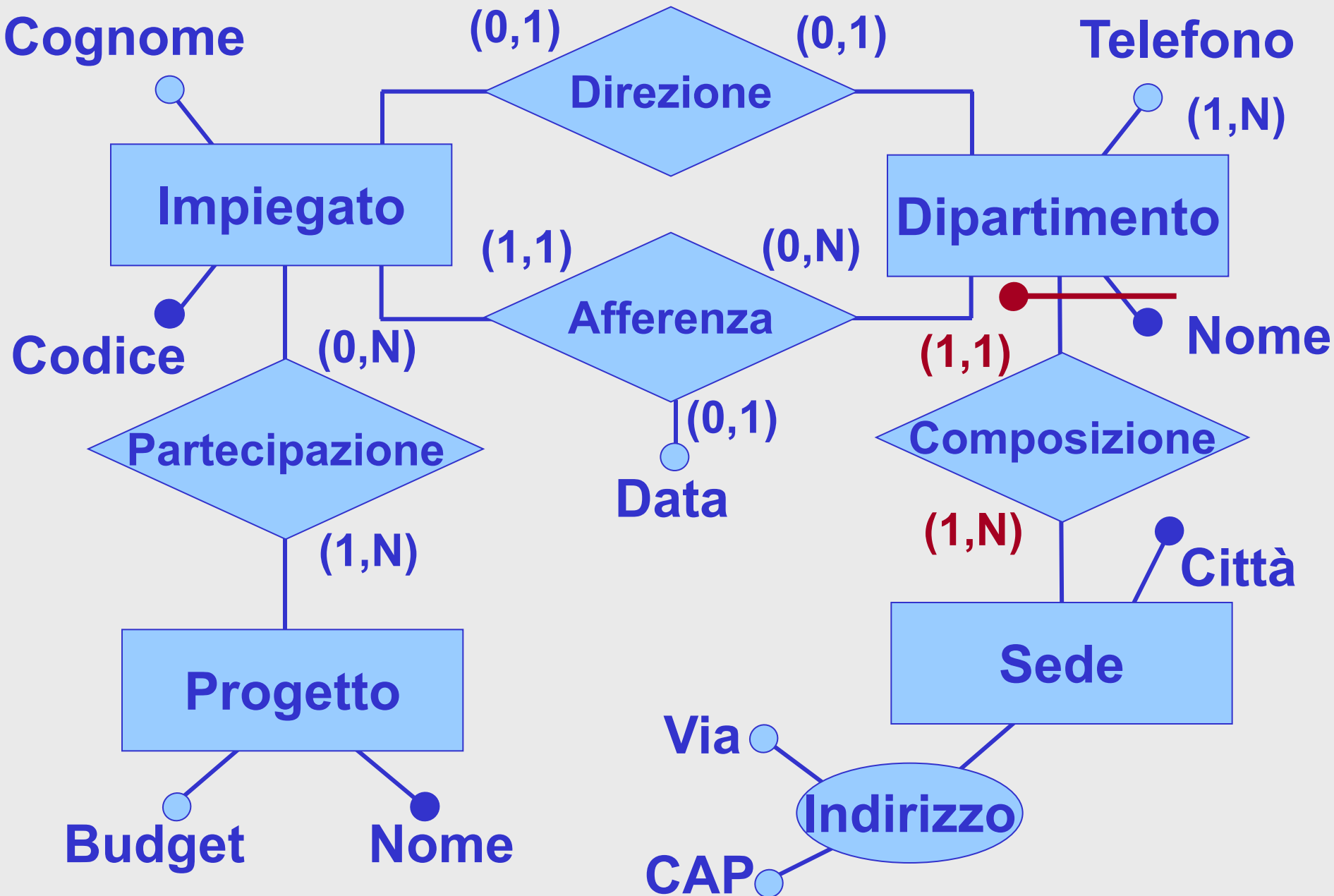
Alcune osservazioni

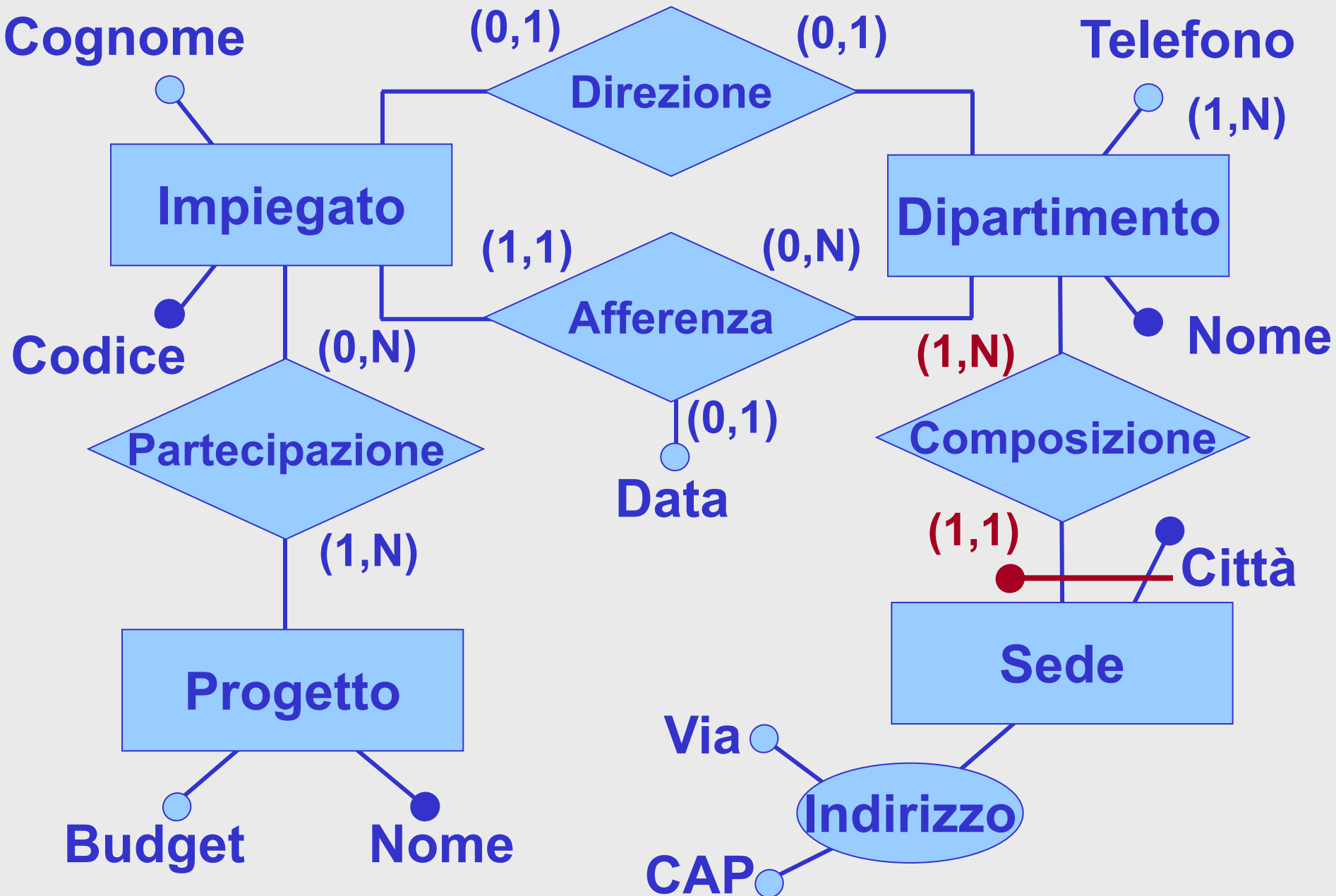
- per ogni entità, almeno un identificatore
- identificazione esterna solo attraverso relationship a cui l'entità partecipa con cardinalità (1,1)
 - in teoria, se ci sono altri identificatori, anche (0,1) potrebbe andare
- perché non parliamo degli identificatori delle relationship?



Attenzione

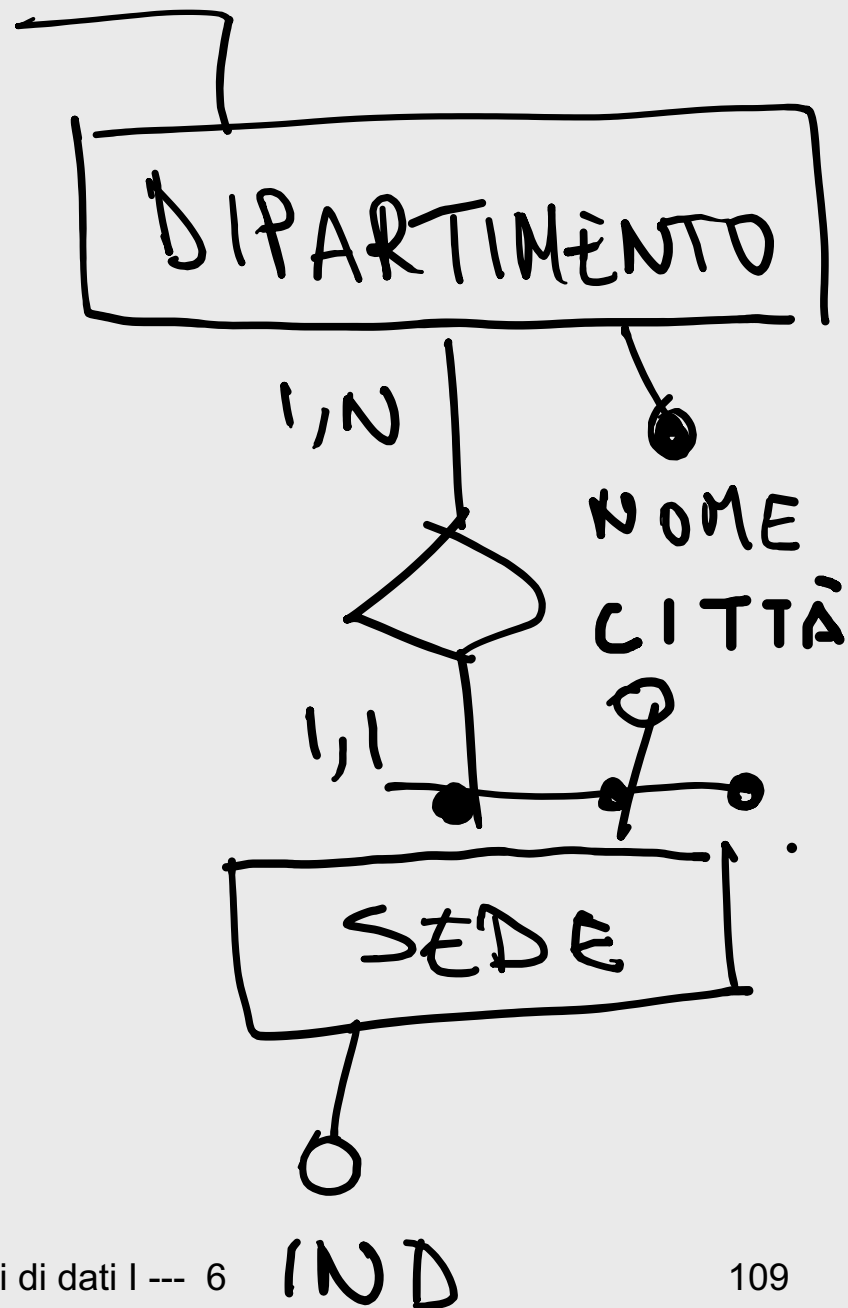
- Differenze apparentemente piccole in cardinalità e identificatori possono cambiare di molto il significato ...







23/11/2323



MILANO VIAPO

VENDITE

PROD

ROMA CORSO

NEGOZIO

PROD

VENDITE

ROMA CORSO

MILANO GALLERIA

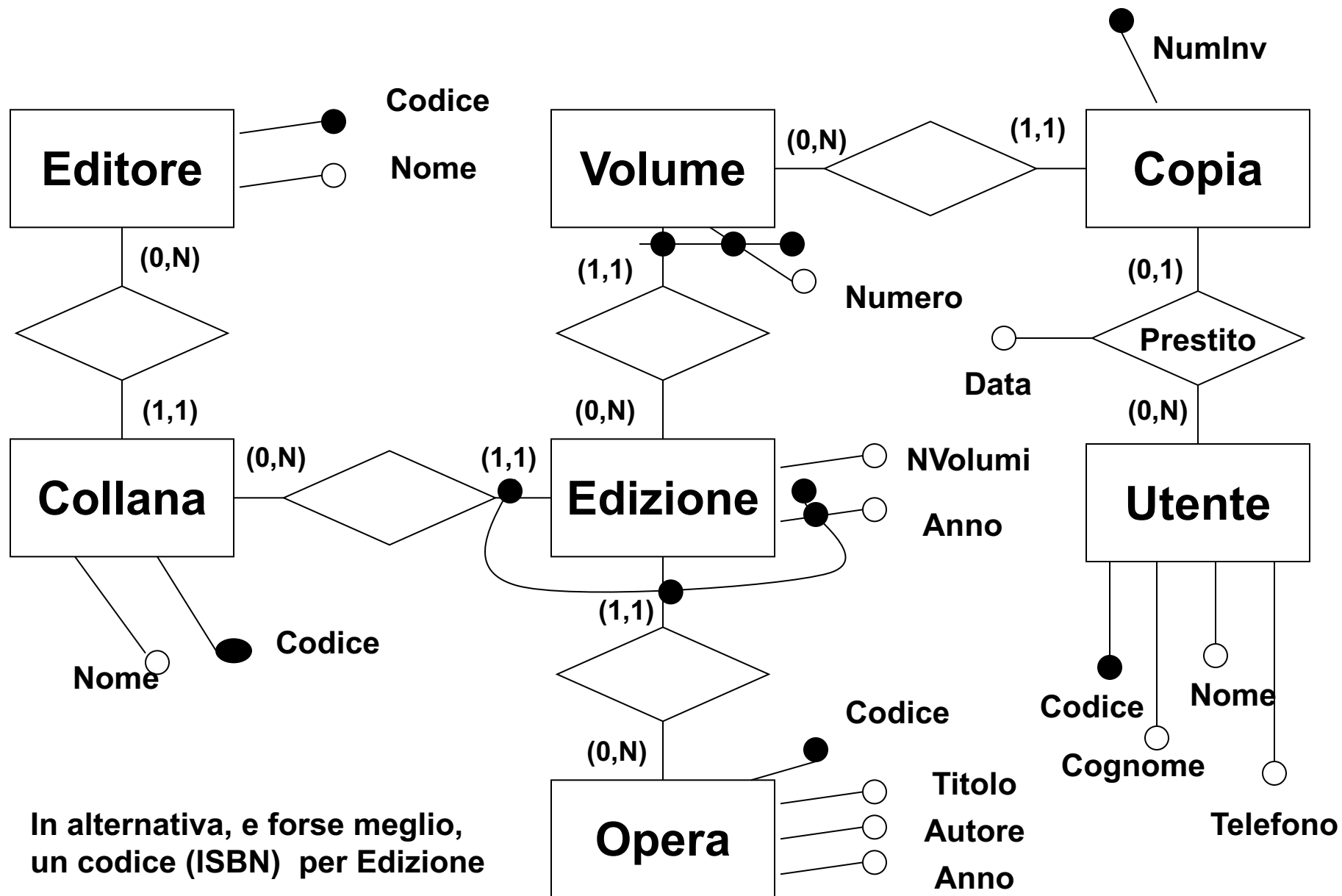
PRODUZIONE

MILANO BOVISA

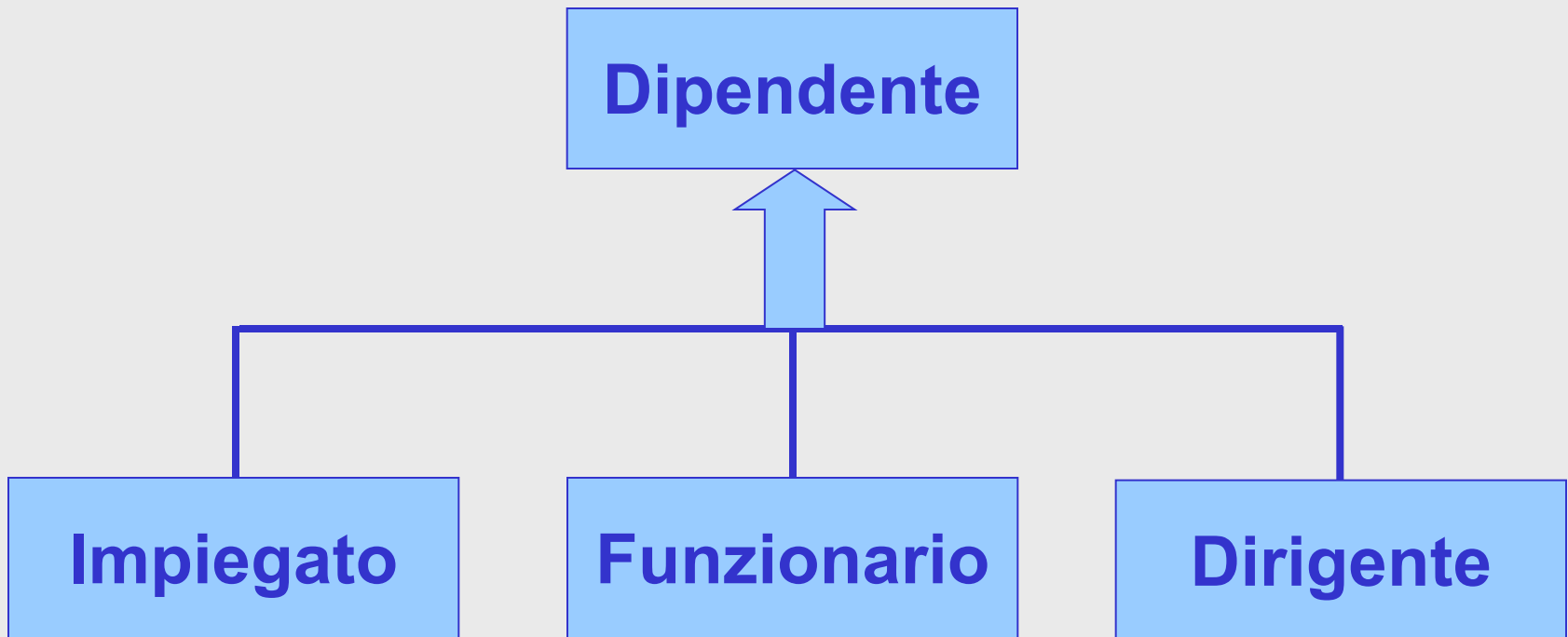
Esercizio (esame del 13/09/2002)

Definire uno schema E-R per una biblioteca, con le seguenti specifiche:

- oggetto dei prestiti sono esemplari (detti anche copie) di singoli volumi, identificati attraverso un numero di inventario;
ogni volume è relativo ad una specifica edizione (che può essere articolata in più volumi, anche in modo diverso dalle altre edizioni) di un'opera
- un volume può essere presente in più copie
- una edizione è caratterizzata dall'opera, dalla collana e dall'anno
- riassumendo ed esemplificando, è possibile prendere in prestito la seconda copia del terzo volume de “I Miserabili”, edizione Mondadori, collana Oscar, del 1975
- ogni collana ha un nome e un codice e un editore
- ogni editore ha un nome e un codice
- ogni opera ha un titolo, un autore e un anno di prima pubblicazione
- per ogni prestito in corso (quelli conclusi non interessano), sono rilevanti la data prevista di restituzione e l'utente (che può avere più volumi in prestito contemporaneamente), con codice identificativo, nome, cognome e recapito telefonico



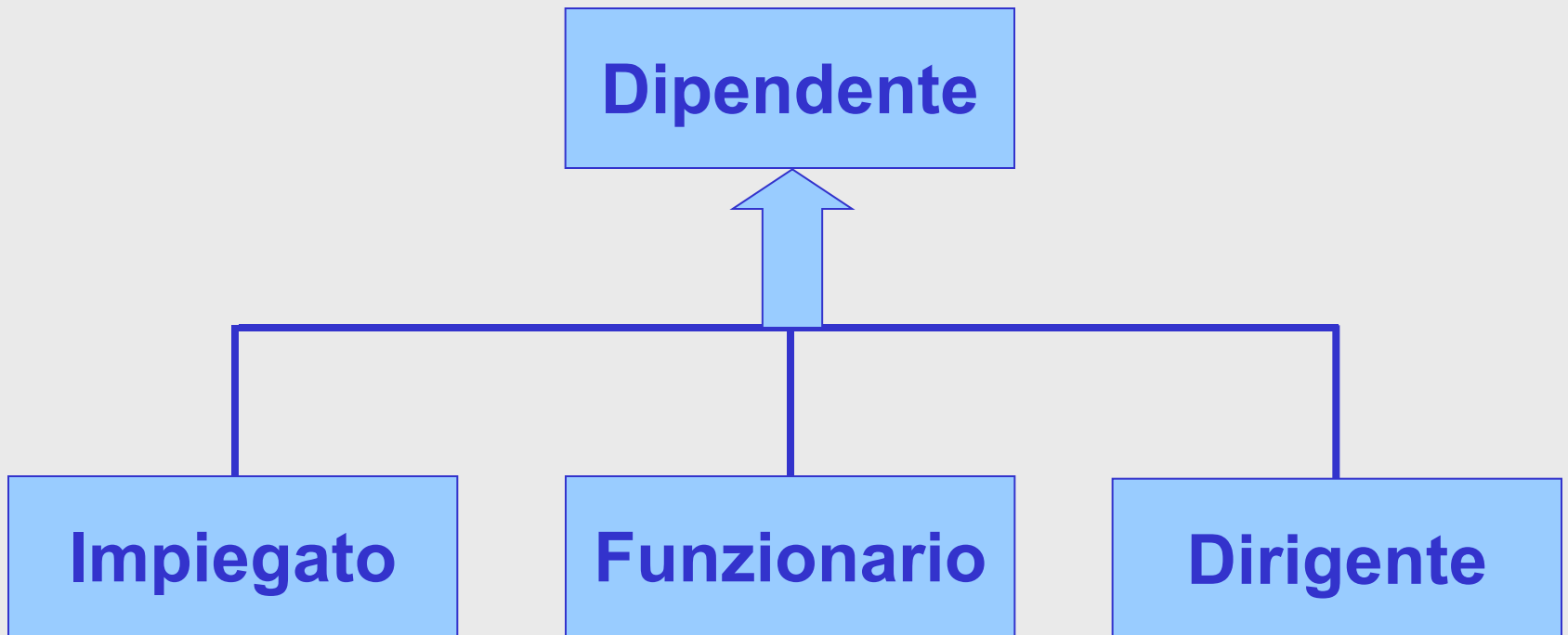
Un ultimo costrutto, la generalizzazione



Generalizzazione

- mette in relazione una o più entità $E_1, E_2, \dots, E_n$ con una entità E , che le comprende come casi particolari
 - E è **generalizzazione** di $E_1, E_2, \dots, E_n$
 - $E_1, E_2, \dots, E_n$ sono **specializzazioni** (o sottotipi) di E

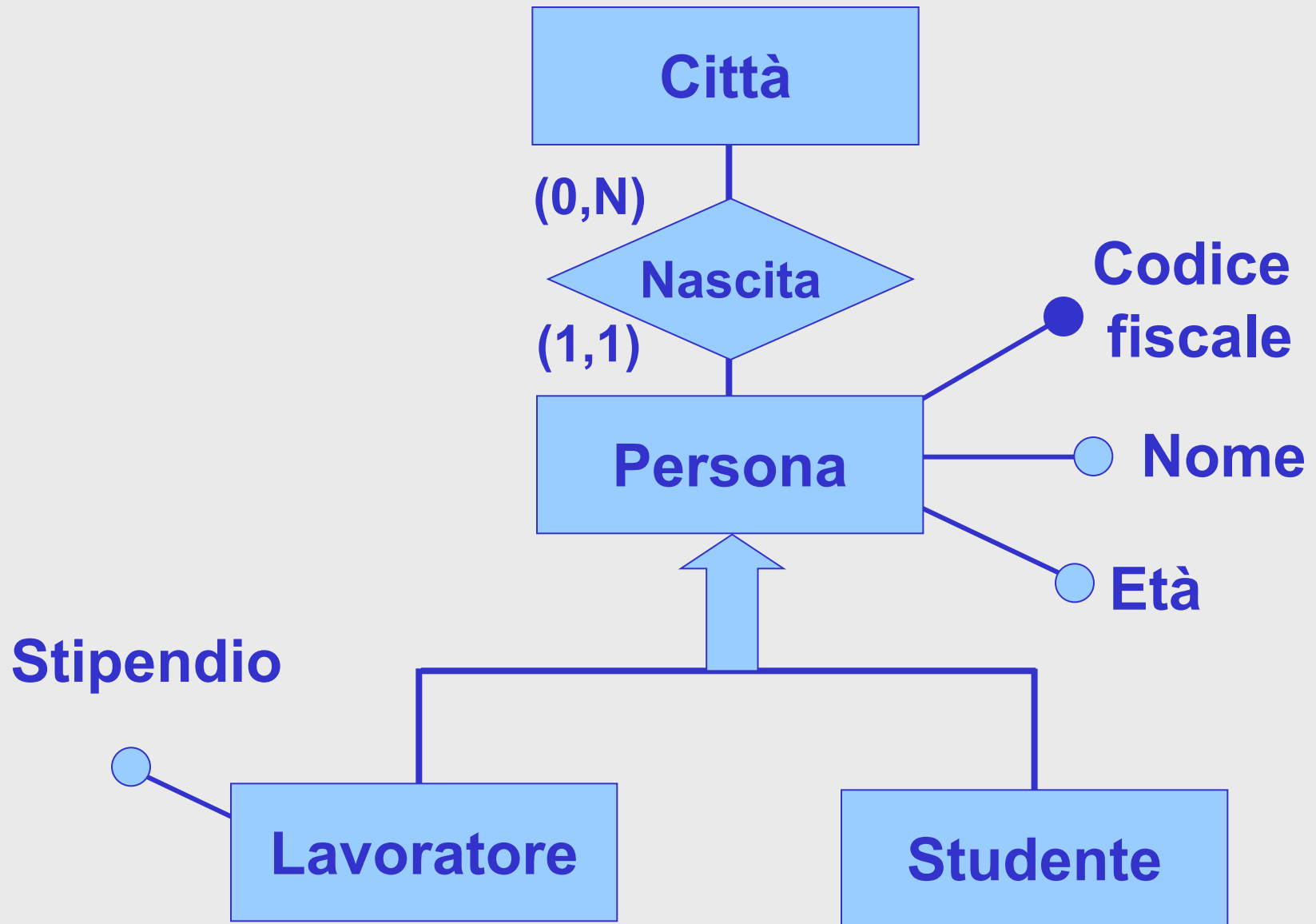
Rappresentazione grafica



Caratteristiche delle generalizzazioni

Se E (genitore) è generalizzazione di $E_1, E_2, \dots, E_n$ (figlie):

- ogni occorrenza di $E_1, E_2, \dots, E_n$ è occorrenza anche di E
- ogni proprietà di E è significativa per $E_1, E_2, \dots, E_n$



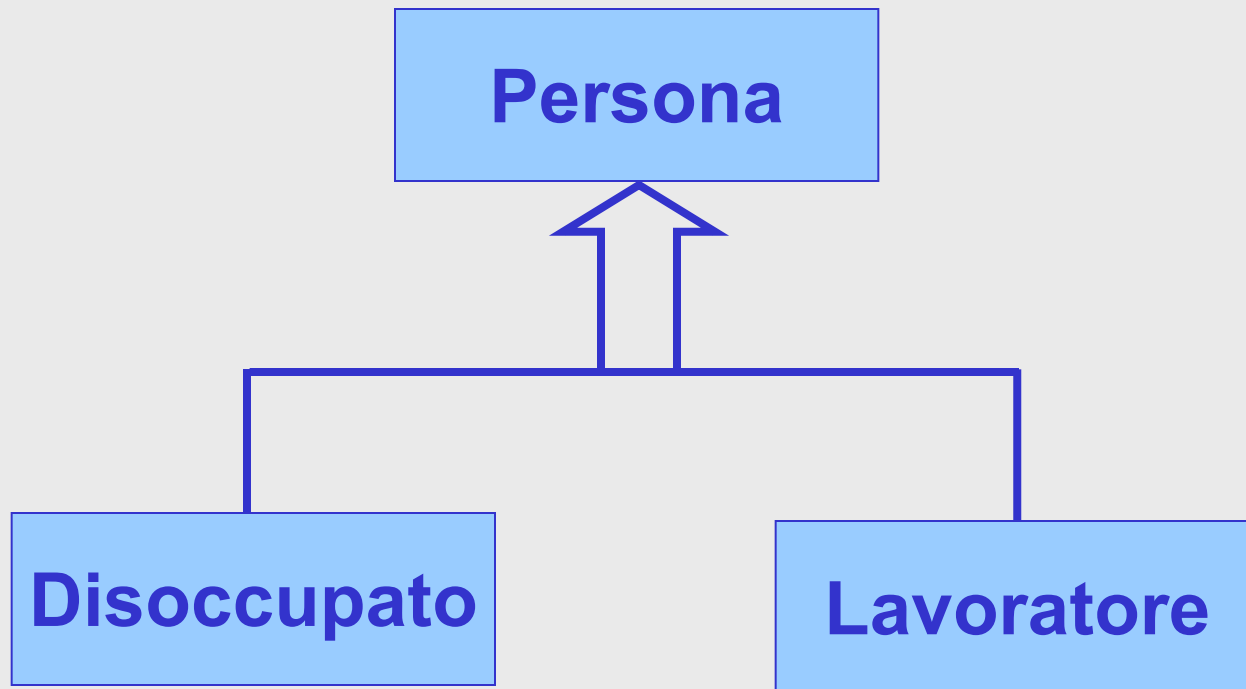
Ereditarietà

- tutte le proprietà (attributi, relationship, altre generalizzazioni) dell' entità genitore vengono **ereditate** dalle entità figlie e non rappresentate esplicitamente

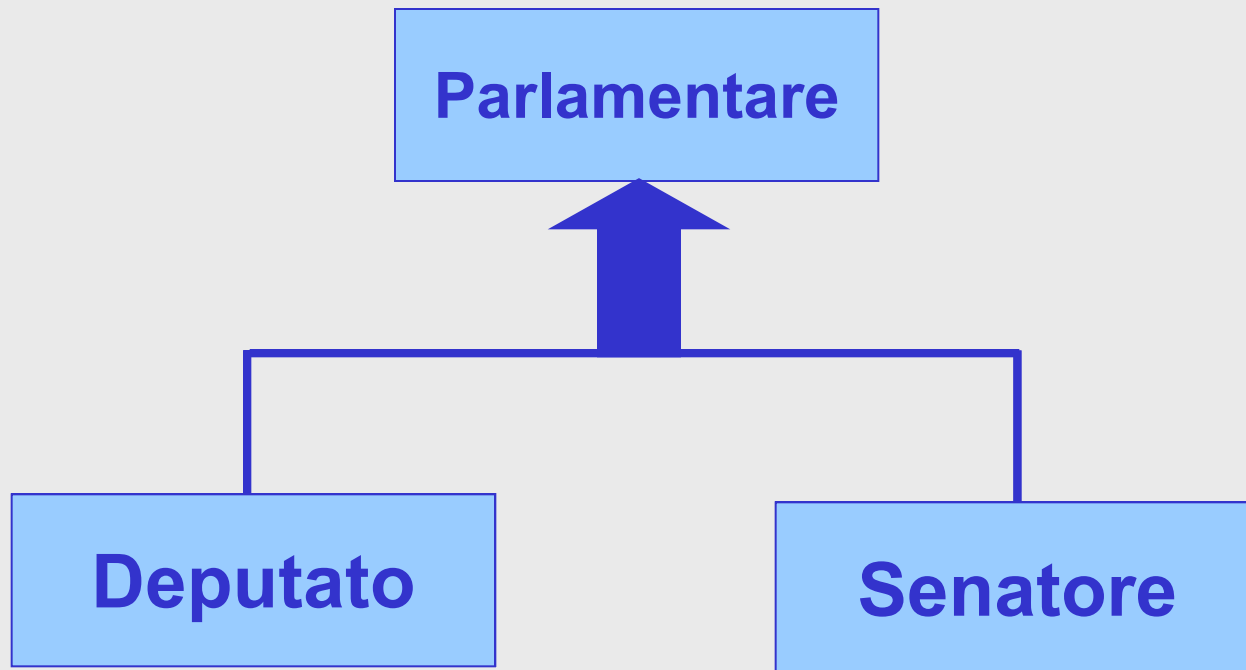
Tipi di generalizzazioni

- **totale** se ogni occorrenza dell'entità genitore è occorrenza di almeno una delle entità figlie, altrimenti è **parziale**
- **esclusiva** se ogni occorrenza dell'entità genitore è occorrenza di al più una delle entità figlie, altrimenti è **sovrapposta**
- consideriamo (senza perdita di generalità) solo generalizzazioni esclusive e distinguiamo fra totali e parziali

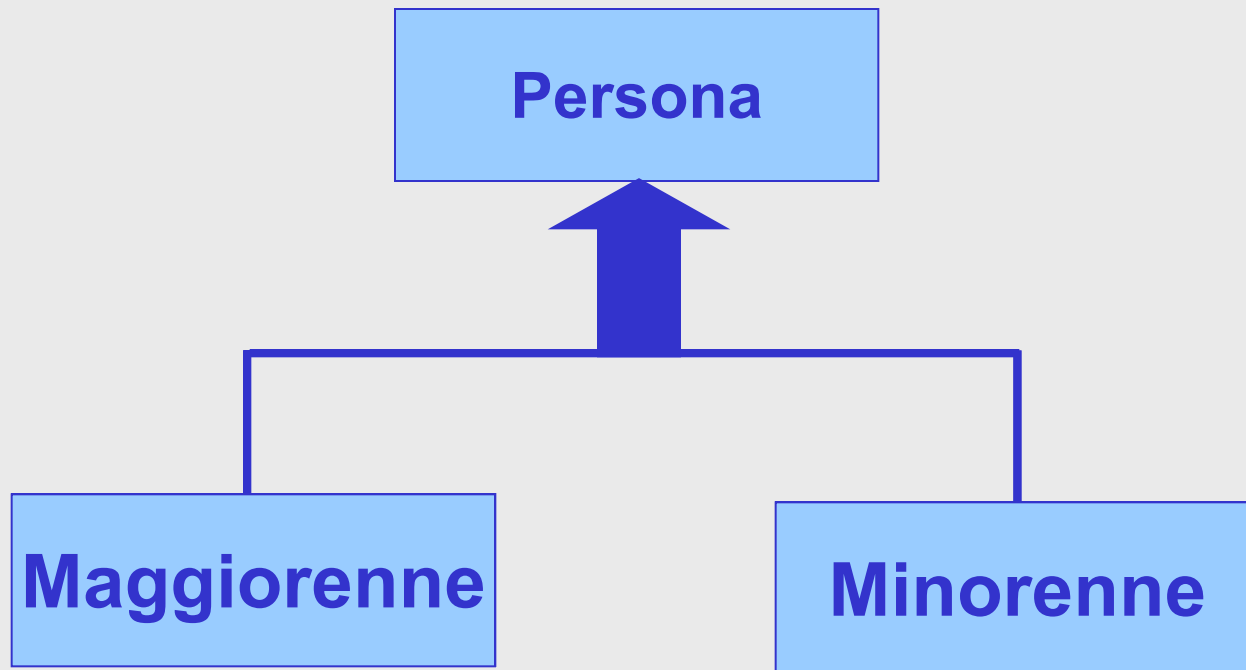
Generalizzazione parziale



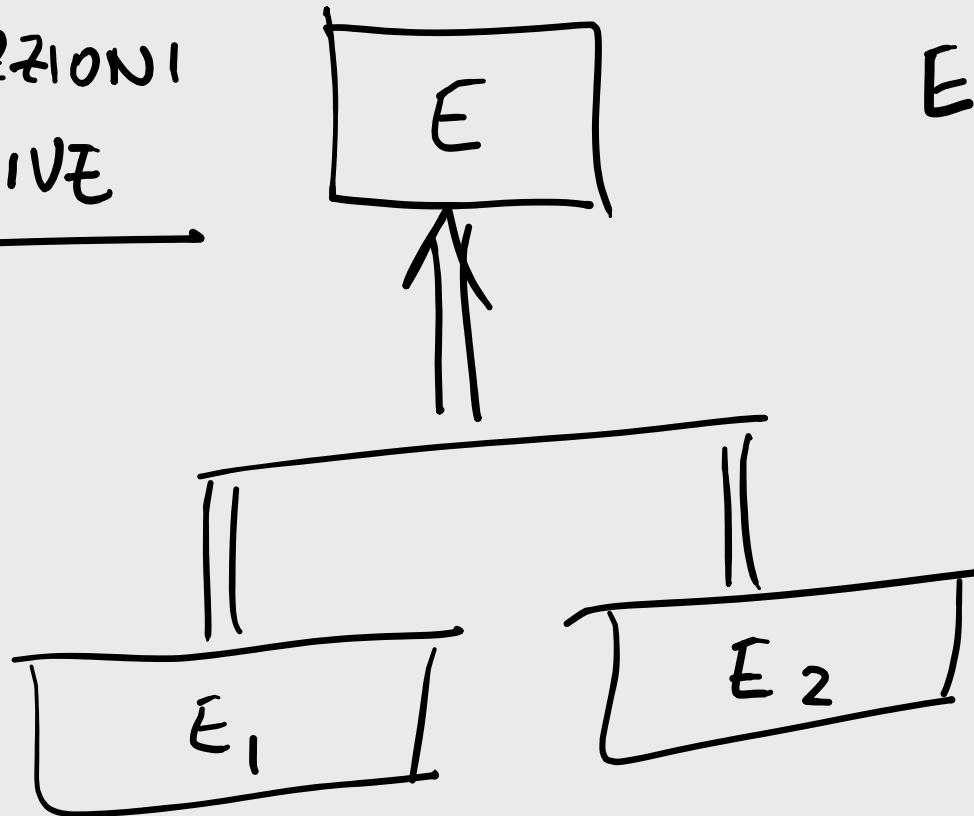
Generalizzazione totale



Generalizzazione totale



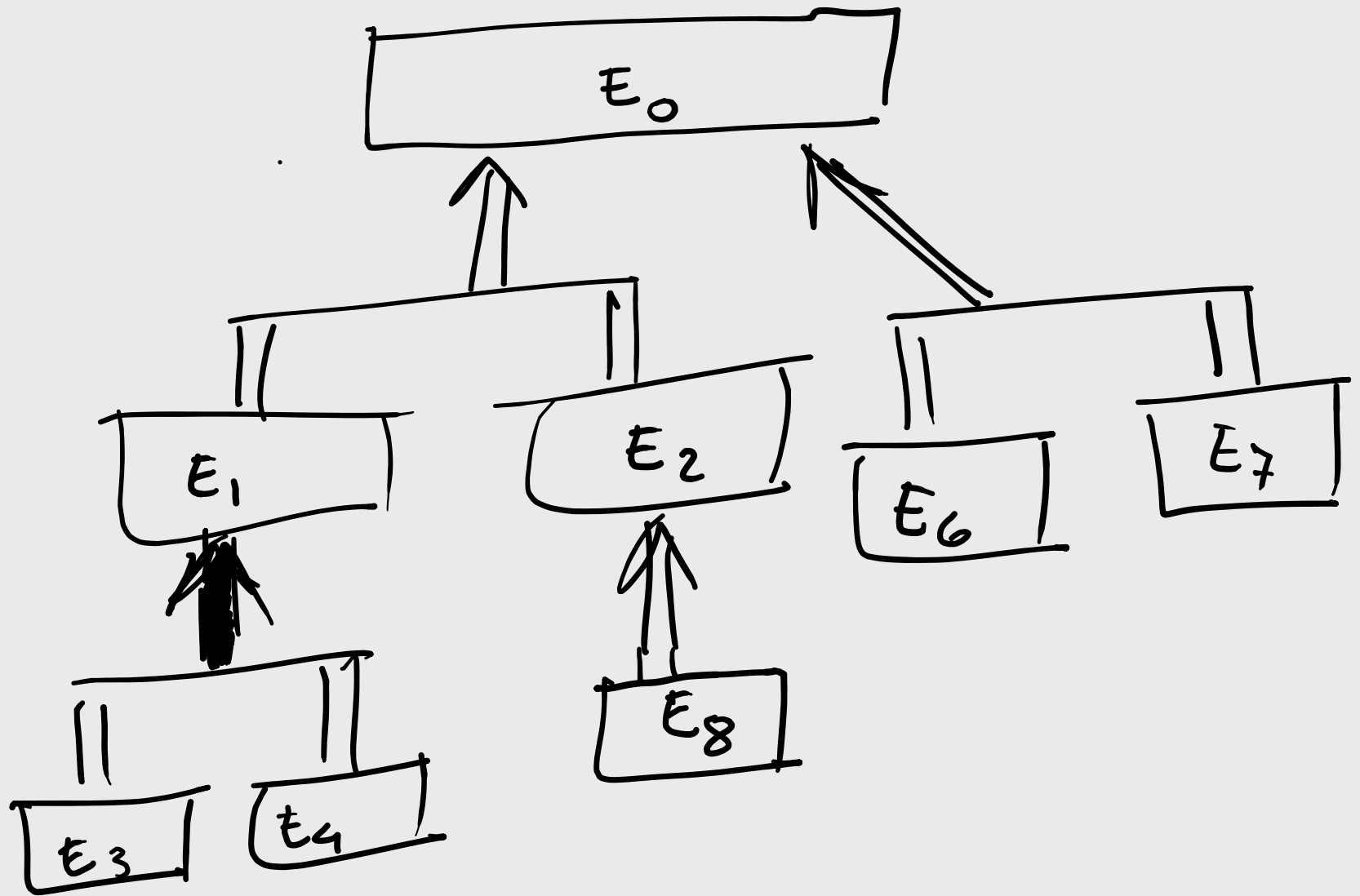
GENERALIZZAZIONI
ESCLUSIVE



$$E_1 \cap E_2 = \emptyset$$

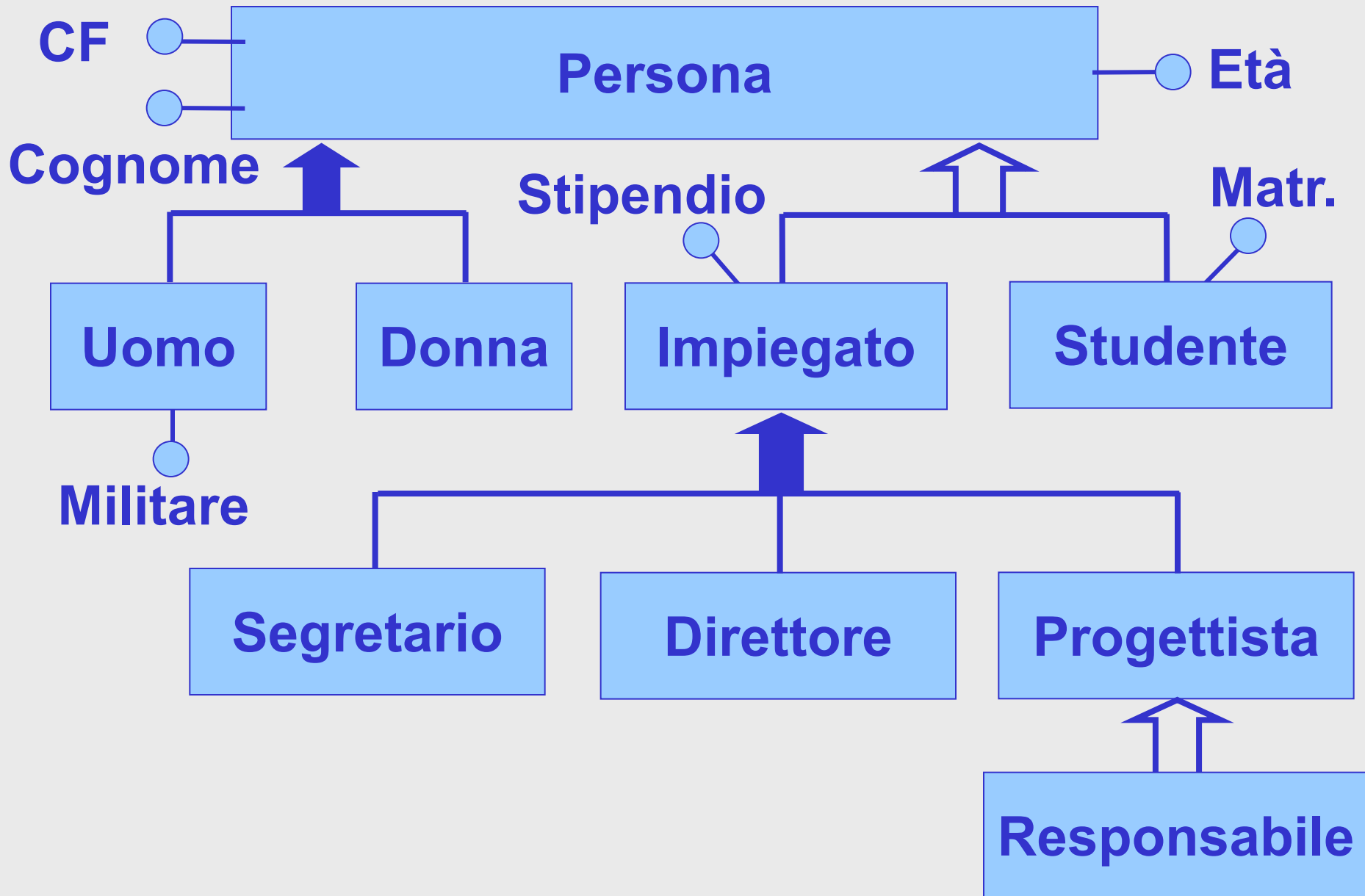
Altre proprietà

- possono esistere gerarchie a più livelli e multiple generalizzazioni allo stesso livello
- un'entità può essere inclusa in più gerarchie, come genitore e/o come figlia
- se una generalizzazione ha solo un' entità figlia si parla di **sottoinsieme**
- alcune configurazioni non hanno senso – vediamo esempio
- il genitore di una generalizzazione totale può non avere identificatore, purché ...



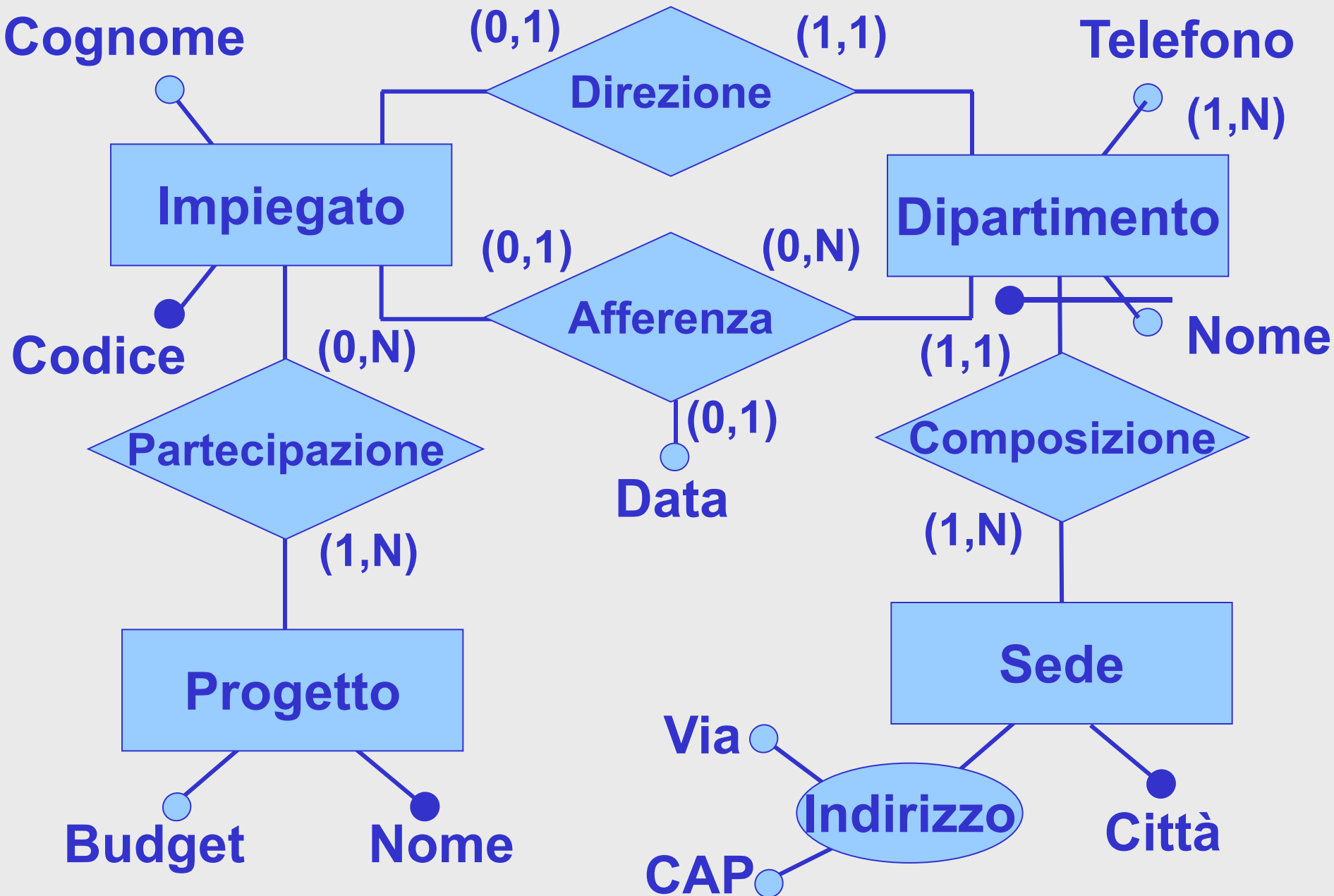
Esercizio

- Le persone hanno CF, cognome ed età; gli uomini anche la posizione militare; gli impiegati hanno lo stipendio e possono essere segretari, direttori o progettisti (un progettista può essere anche responsabile di progetto); gli studenti (che non possono essere impiegati) un numero di matricola; esistono persone che non sono né impiegati né studenti (ma i dettagli non ci interessano)



Documentazione associata agli schemi concettuali

- dizionario dei dati
 - entità
 - relationship
- vincoli non esprimibili



Dizionario dei dati (entità)

Entità	Descrizione	Attributi	Identificatore
Impiegato	Dipendente dell'azienda	Codice, Cognome, Stipendio	Codice
Progetto	Progetti aziendali	Nome, Budget	Nome
Dipartimento	Struttura aziendale	Nome, Telefono	Nome, Sede
Sede	Sede dell'azienda	Città, Indirizzo	Città

Dizionario dei dati (relationship)

Relazioni	Descrizione	Componenti	Attributi
Direzione	Direzione di un dipartimento	Impiegato, Dipartimento	
Afferenza	Afferenza a un dipartimento	Impiegato, Dipartimento	Data
Partecipazione	Partecipazione a un progetto	Impiegato, Progetto	
Composizione	Composizione dell'azienda	Dipartimento, Sede	

Vincoli non esprimibili

Vincoli di integrità sui dati

- (1) Il direttore di un dipartimento deve afferire a tale dipartimento
- (2) Un impiegato non deve avere uno stipendio maggiore del direttore del dipartimento al quale afferisce
- (3) Un dipartimento con sede a Roma deve essere diretto da un impiegato con più di dieci anni di anzianità
- (4) Un impiegato che non afferisce a nessun dipartimento non può partecipare a nessun progetto

Modellazione dei dati in UML

- In alternativa al modello ER per la rappresentazione concettuale dei dati viene talvolta utilizzato il linguaggio/modello UML
- Si fa uso dei diagrammi delle classi
- Cambia la rappresentazione diagrammatica ma non l'approccio alla progettazione
- Non approfondiamo, vediamo solo una figura
- UML viene presentato in dettaglio nel corso di Analisi e Progettazione del Software

Uno schema concettuale in UML

