

Progettazione logica

Progettazione Logica



Il prodotto della progettazione logica è uno schema logico che rappresenta le informazioni contenute nello schema E-R in modo corretto ed efficiente.

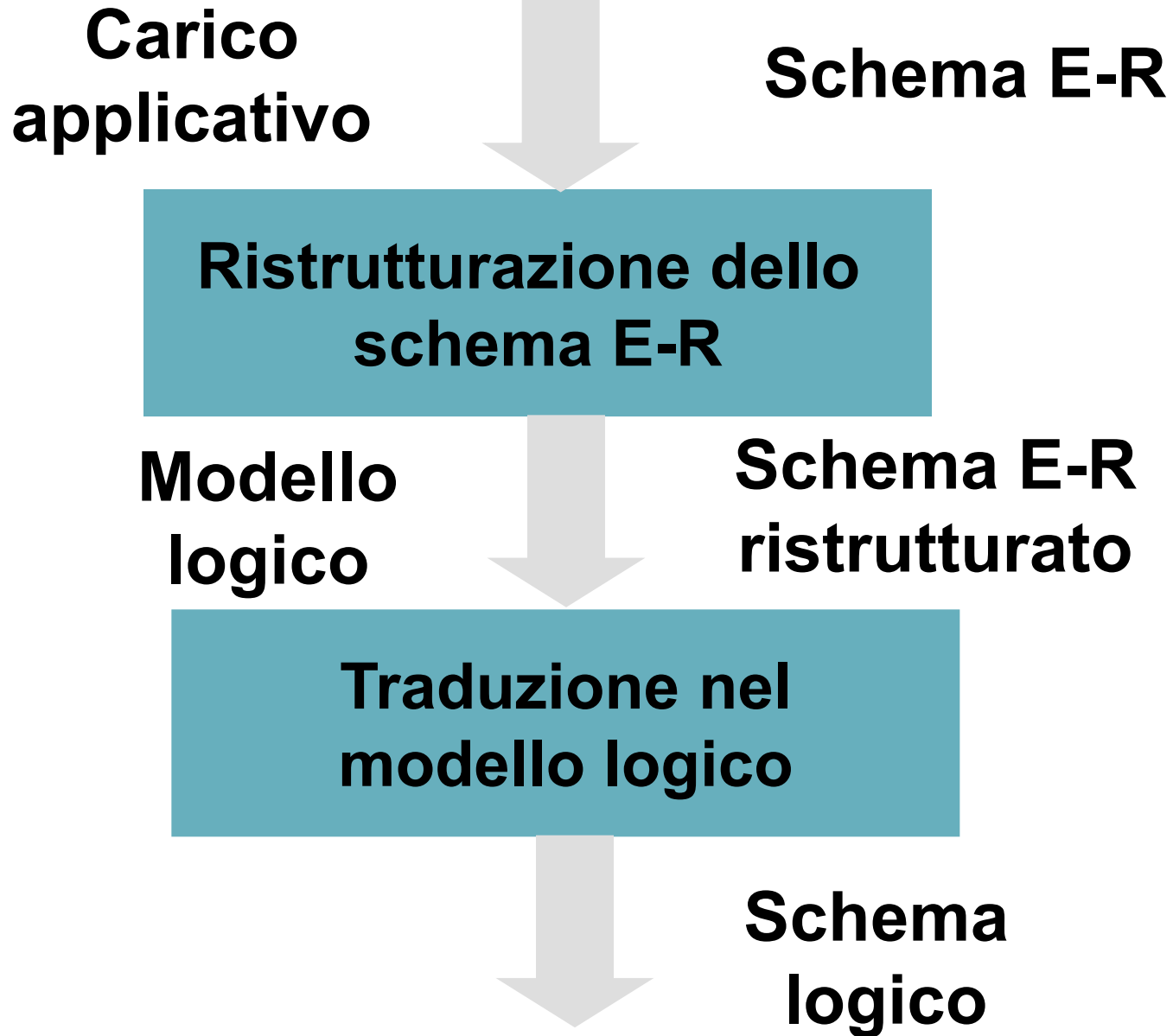
Richiede non solo una *traduzione* da uno schema all'altro ma anche una *ristrutturazione* dello schema.

Ad es. le generalizzazioni non hanno una traduzione diretta fra modello E-R e modello relazionale. Inoltre bisogna tenere conto delle prestazioni e quindi la riorganizzazione deve avere come obiettivo l'efficienza.

Come passo preliminare bisognerà eseguire un'analisi delle prestazioni attese dal sistema, per ottimizzarne la realizzazione.

**Carico
applicativo**

Schema E-R



```
graph TD; A[Carico applicativo] --> B[Ristrutturazione dello schema E-R]; B --> C[Modello logico]; C --> D[Traduzione nel modello logico]; D --> E[Schema logico];
```

**Ristrutturazione dello
schema E-R**

**Modello
logico**

**Schema E-R
ristrutturato**

**Traduzione nel
modello logico**

**Schema
logico**

Analisi delle prestazioni su schemi E-R

In tutti i sistemi software i due principali indici di prestazione che si considerano sono:

- Costo delle operazioni
- Occupazione di memoria

Nel caso di una base di dati, il costo di un'operazione è misurato in funzione del numero di occorrenze delle entità e delle associazioni fra entità coinvolte.

Per fare questa analisi è necessario conoscere:

- Volume dei dati (occorrenze, dimensioni degli attributi)
- Caratteristiche delle operazioni (tipo, frequenza, dati coinvolti)

Analisi delle prestazioni su schemi E-R

In genere basta limitarsi ad analizzare il costo delle operazioni più frequenti (regola dell'80/20 "*l'80% del tempo è occupato dal 20% delle operazioni*")

Per descrivere volume dei dati e caratteristiche delle operazioni si possono costruire la *tavola dei volumi*, in cui vengono riportati tutti i concetti dello schema con i volumi previsti a regime, e la *tavola delle operazioni*, in cui si riportano il tipo e la frequenza delle operazioni.

Analisi delle prestazioni su schemi E-R

Per descrivere le operazioni si può usare il frammento dello schema E-R che comprende i concetti coinvolti nell'operazione, sul quale vengono riportati i cammini logici attraverso i quali si accede ai dati.

Nella tavola degli accessi si riportano i concetti coinvolti, il numero medio degli accessi e il tipo (lettura, scrittura).

Operazione 1: assegna un impiegato a un progetto.

Operazione 2: trova i dati di un impiegato, del dipartimento nel quale lavora e dei progetti ai quali partecipa.

Operazione 3: trova i dati di tutti gli impiegati di un certo dipartimento.

Operazione 4: per ogni sede, trova i suoi dipartimenti con il cognome del direttore e l'elenco degli impiegati del dipartimento.

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

Tavola delle operazioni

Operazione	Tipo	Frequenza
Op. 1	I	50 al giorno
Op. 2	I	100 al giorno
Op. 3	I	10 al giorno
Op. 4	B	2 a settimana

Ristrutturazione di schemi E-R

Può essere divisa in più passi successivi:

- **Analisi delle ridondanze:** si decide se mantenere o eliminare le eventuali ridondanze
- **Eliminazione delle generalizzazioni:** si sostituiscono le generalizzazioni con costrutti del modello logico utilizzato
- **Partizionamento/accorpamento di entità e associazioni:** si decide se raggruppare più concetti o partizionare in più concetti i concetti dello schema E-R
- **Scelta degli identificatori primari:** si selezionano gli identificatori per le entità che ne hanno più di uno

Ridondanze

- **Vantaggi**

- semplificazione delle interrogazioni: se lo stesso dato compare in più tabelle, è minore la probabilità di dovere fare un join per recuperarlo.

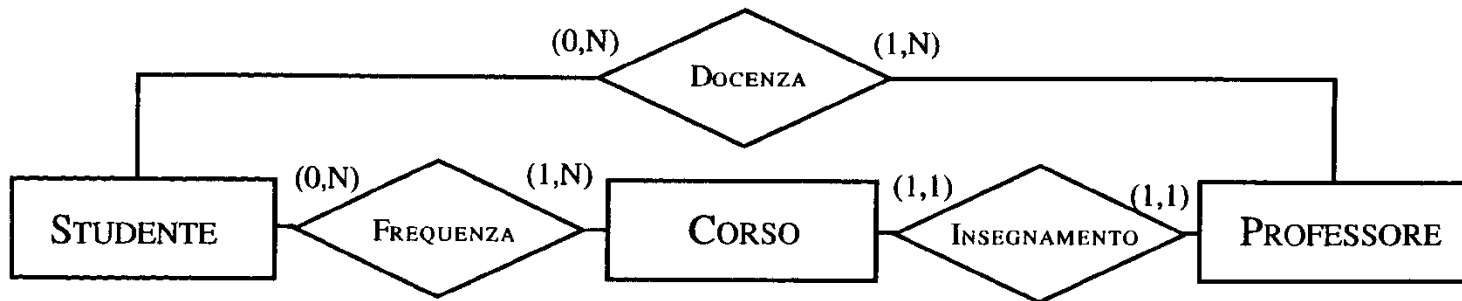
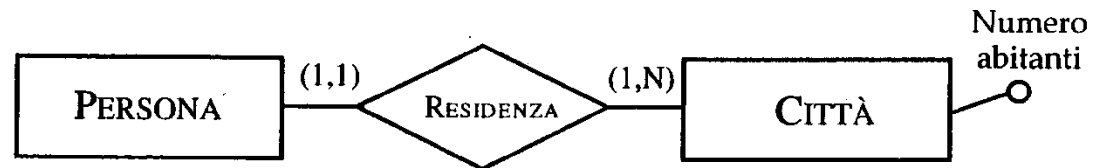
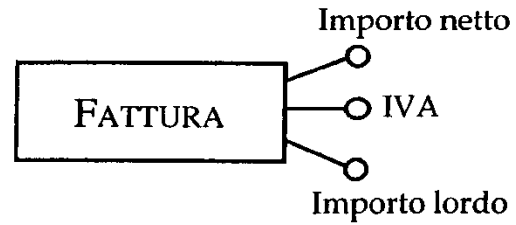
- **Svantaggi**

- appesantimento degli aggiornamenti: se modifico un'occorrenza del dato devo modificare anche tutte le 'copie'
- maggiore occupazione di memoria

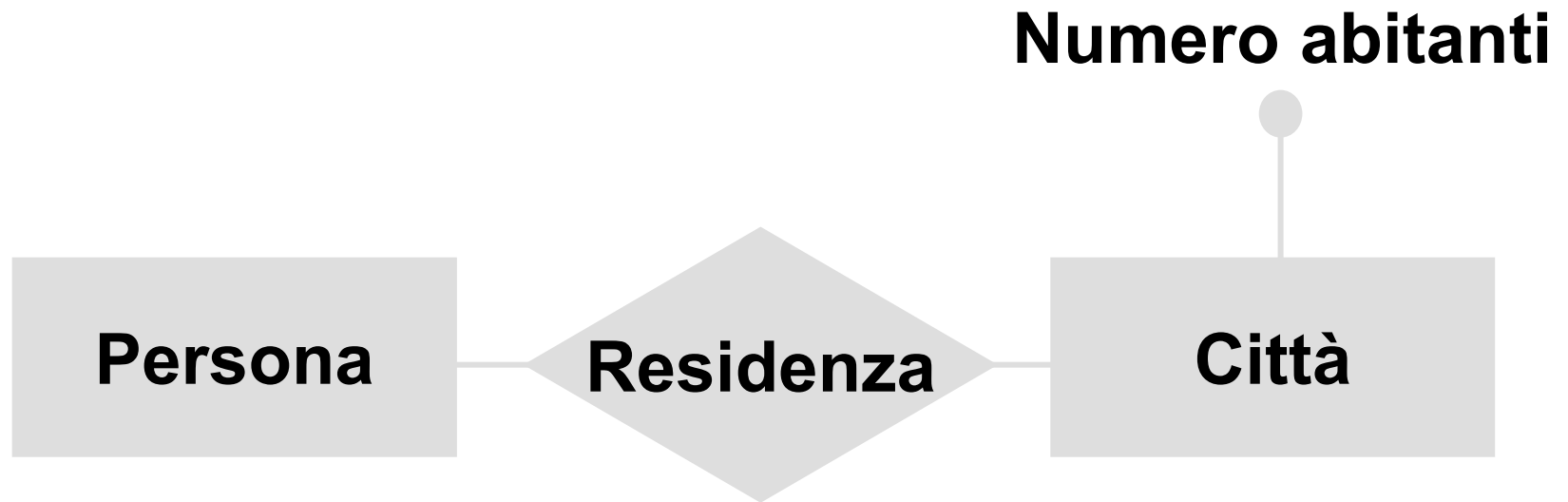
Analisi delle ridondanze

Le ridondanze di uno schema E-R nascono quando un concetto può essere derivato da altri.

- Un attributo può essere derivato da attributi della stessa entità o relazione
- Un attributo può essere derivato da attributi di altre entità e relazioni
- Un attributo può essere derivato mediante conteggio di occorrenze
- Associazioni possono essere derivate da altre associazioni in presenza di cicli



Analisi di una ridondanza



Numero abitanti è pari al numero di istanze della relazione Residenza

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

- *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)

Presenza di ridondanza

Operazione 1

Concetto	Costrutto	Accessi	Tipo
Persona	Entità	1	S
Residenza	Relazione	1	S
Città	Entità	1	L
Città	Entità	1	S

Operazione 2

Concetto	Costrutto	Accessi	Tipo
Città	Entità	1	L

Assenza di ridondanza

Operazione 1

Concetto	Costrutto	Accessi	Tipo
Persona	Entità	1	S
Residenza	Relazione	1	S

Operazione 2

Concetto	Costrutto	Accessi	Tipo
Città	Entità	1	L
Residenza	Relazione	5000	L

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

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

Eliminazione delle gerarchie

Per eliminare le generalizzazioni (non presenti nel modello relazionale) si possono seguire 3 strategie alternative:

1. **Accorpamento delle figlie nel padre :**

si usa una sola entità avente anche tutti gli attributi delle figlie più un attributo che indica il “tipo”.

Richiede più memoria (saranno presenti valori nulli in corrispondenza degli attributi specifici per una delle figlie) ma meno accessi (per ogni accesso accedo sempre e solo alla entità padre).

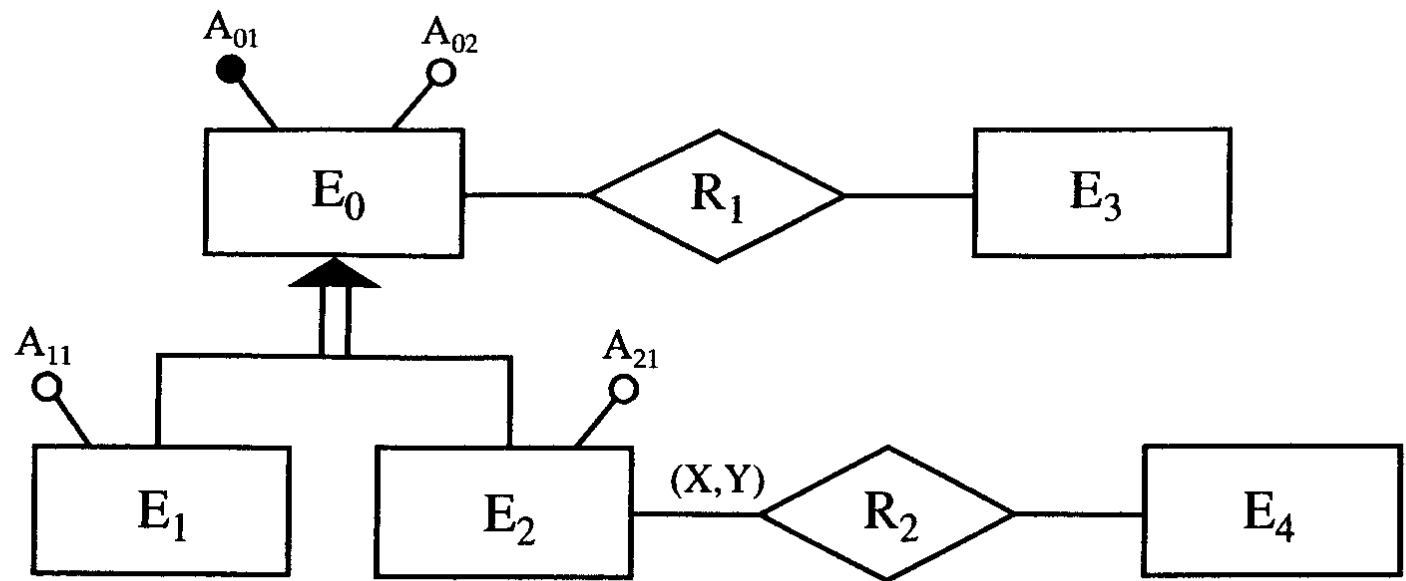
Eliminazione delle gerarchie

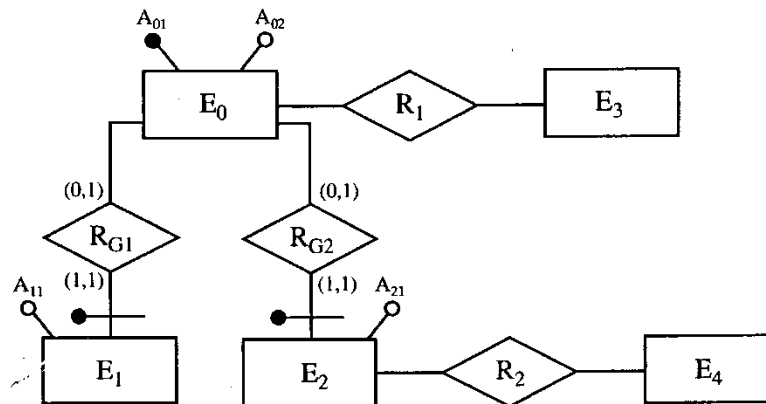
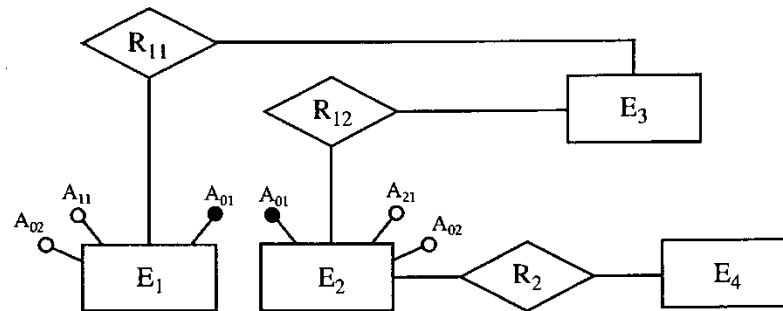
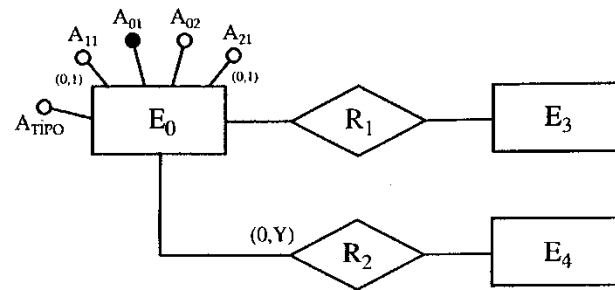
2. **Accorpamento del padre nelle figlie:**

si elimina il padre ed i suoi attributi vengono inseriti in ciascuna delle figlie (possibile solo se la generalizzazione è totale. Richiede meno memoria rispetto al caso 1, meno accessi rispetto al 3).

3. **Sostituzione della generalizzazione con associazioni:**

la generalizzazione si trasforma in una serie di associazioni, con il vincolo che, se è totale, ogni occorrenza della entità padre deve partecipare ad almeno una di tali associazioni (meno memoria rispetto a 1, più accessi a causa del vincolo aggiuntivo).





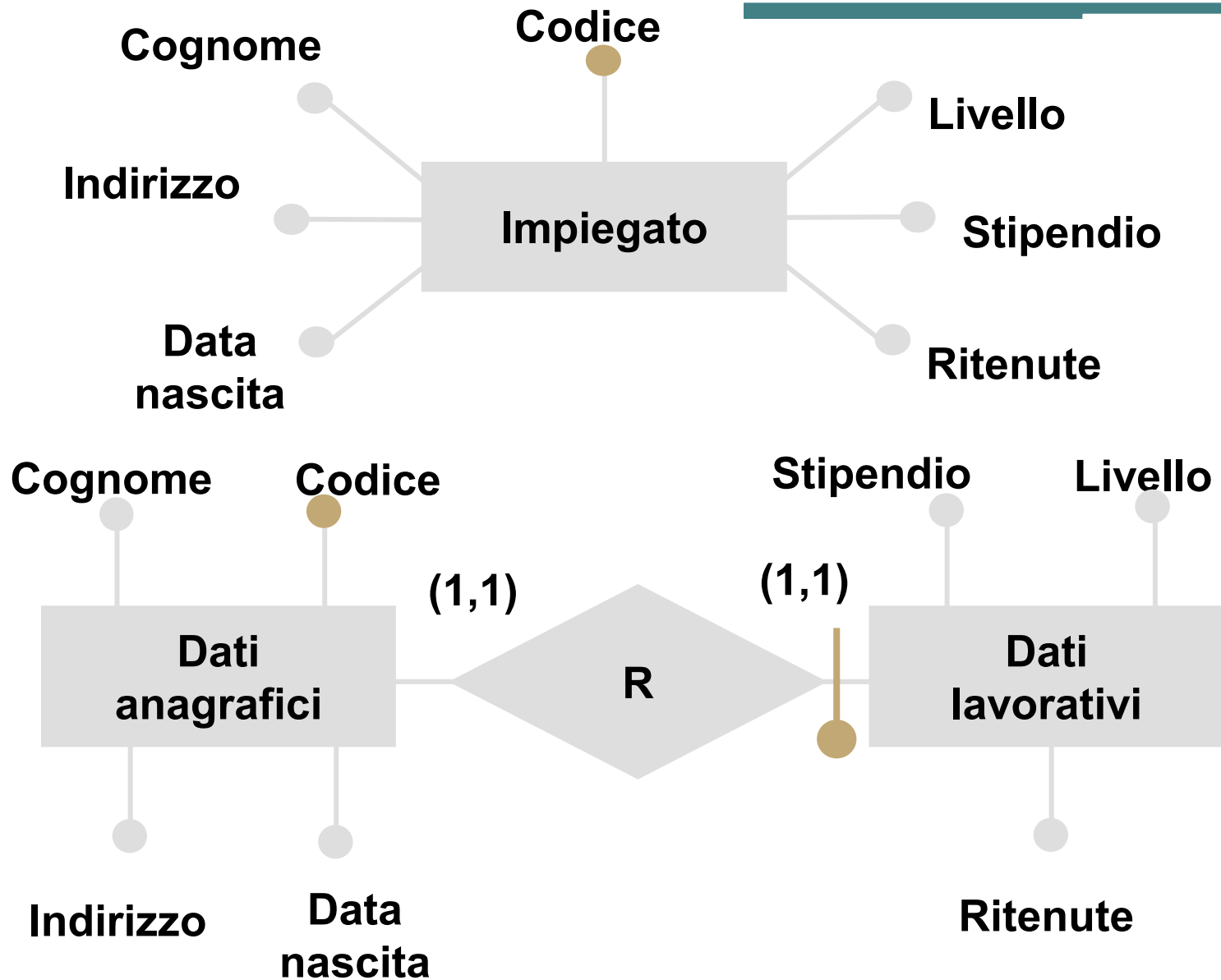
Partizionamento/accorpamento di concetti

Gli accessi possono essere ridotti separando gli attributi di uno stesso concetto cui fanno accesso operazioni diverse e raggruppando attributi di concetti diversi cui fanno accesso le stesse operazioni.

Partizionamenti di entità :

Si possono fare *in verticale*, decomponendo l'entità in più entità sulla base degli attributi, o in modo *orizzontale*, decomponendo l'entità sulla base delle sue occorrenze.

Una decomposizione orizzontale corrisponde ad una generalizzazione. Costringe a duplicare tutte le associazioni cui l'entità originaria partecipa.



Partizionamento/accorpamento di concetti

Eliminazione di attributi multivalore :

il modello relazionale non prevede questo tipo di attributo e richiede che venga sostituito da una relazione o nella decomposizione diretta dell'attributo multivalore in più attributi.

Accorpamento di entità :

spesso conviene accorpare due entità coinvolte spesso nella stessa operazione in una sola entità. Questo può causare la presenza di valori nulli.

Si accorpano di solito entità legate da relazioni di tipo uno a uno, per evitare la creazione di ridondanze.

E' possibile anche accorpare/partizionare associazioni.

