



Dalla Progettazione Concettuale alla Progettazione Logica

Informatica ICA (LC)

23 Novembre 2017

Giacomo Boracchi

giacomo.boracchi@polimi.it



Le presenti slide sono tratte dalle slide del libro di testo
"Basi di dati", Atzeni, Ceri, Fraternali, Paraboschi, Torlone,
McGraw-Hill 4a edizione, 2014



Dallo schema ER al modello relazionale

Obiettivo:

- Passare da uno schema ER ad uno schema relazionale

Due fasi:

- La ristrutturazione dello schema ER: ottimizzazioni dello schema ER indipendenti dal modello logico scelto
- La traduzione verso il modello relazionale



Ristrutturazione dello Schema E-R

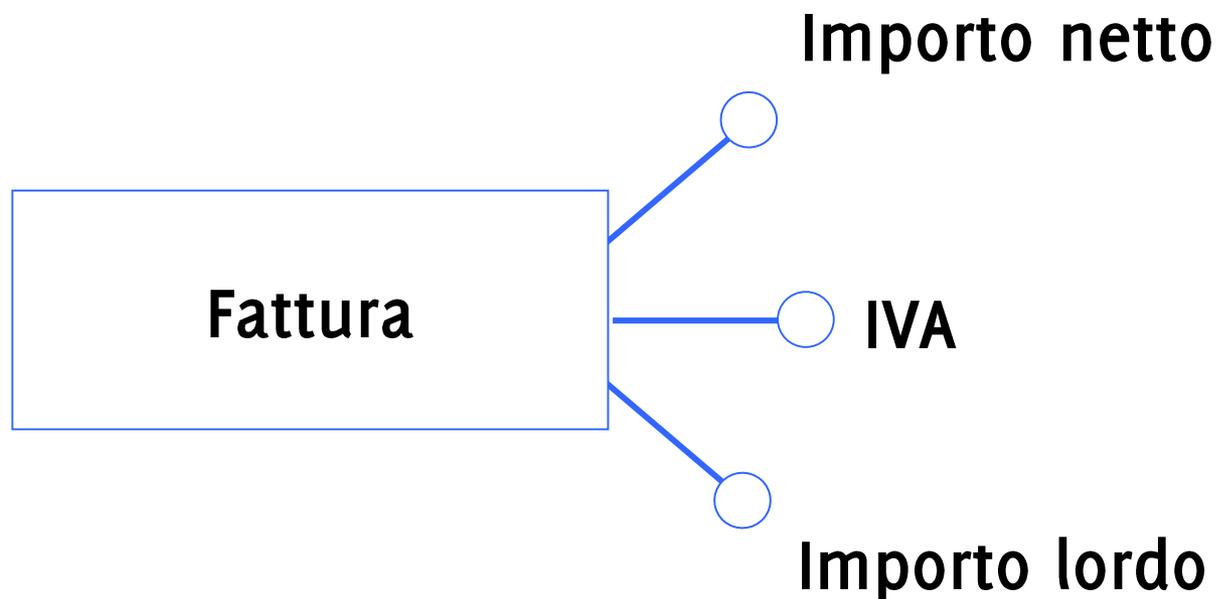
Consideriamo le seguenti fasi:

- **Analisi delle ridondanze**
- Eliminazione delle generalizzazioni
- Scelta degli identificatori primari



Analisi delle ridondanze

Una ridondanza in uno schema E-R è un'informazione significativa ma derivabile da altre.





Attributi Derivabili

Ci focalizziamo solo su questa forma di ridondanza e vediamo **attributi derivabili** da

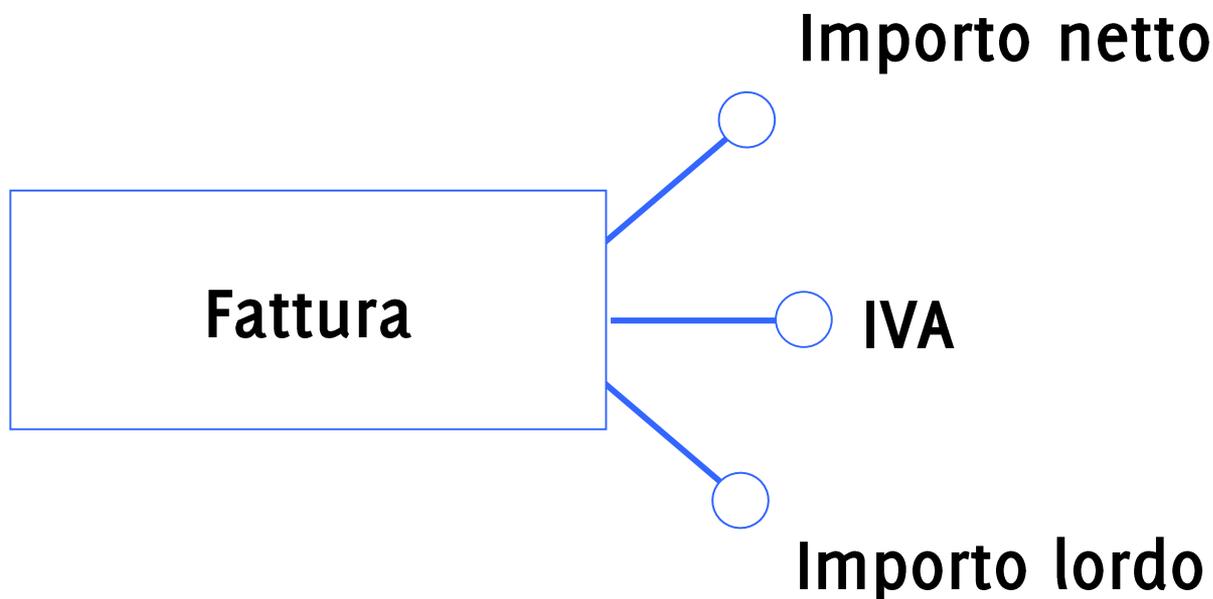
- da altri attributi della stessa entità (o associazione)
- da attributi di altre entità (o associazione)



Analisi delle ridondanze

L'importo lordo è derivabile da altri attributi della stessa entità:

$$\text{ImportoLordo} = \text{ImportoNetto} * (1 + \text{IVA})$$

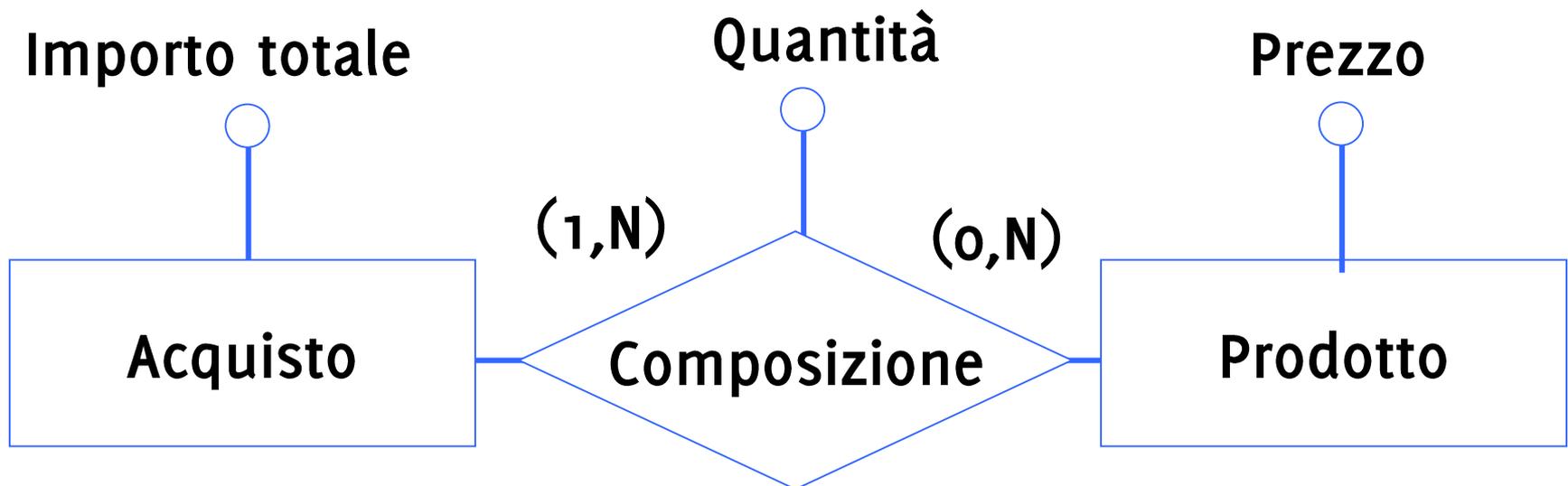




Attributo derivabile da altra entità

L'importo totale può essere derivato dall'associazione Acquisto prodotto

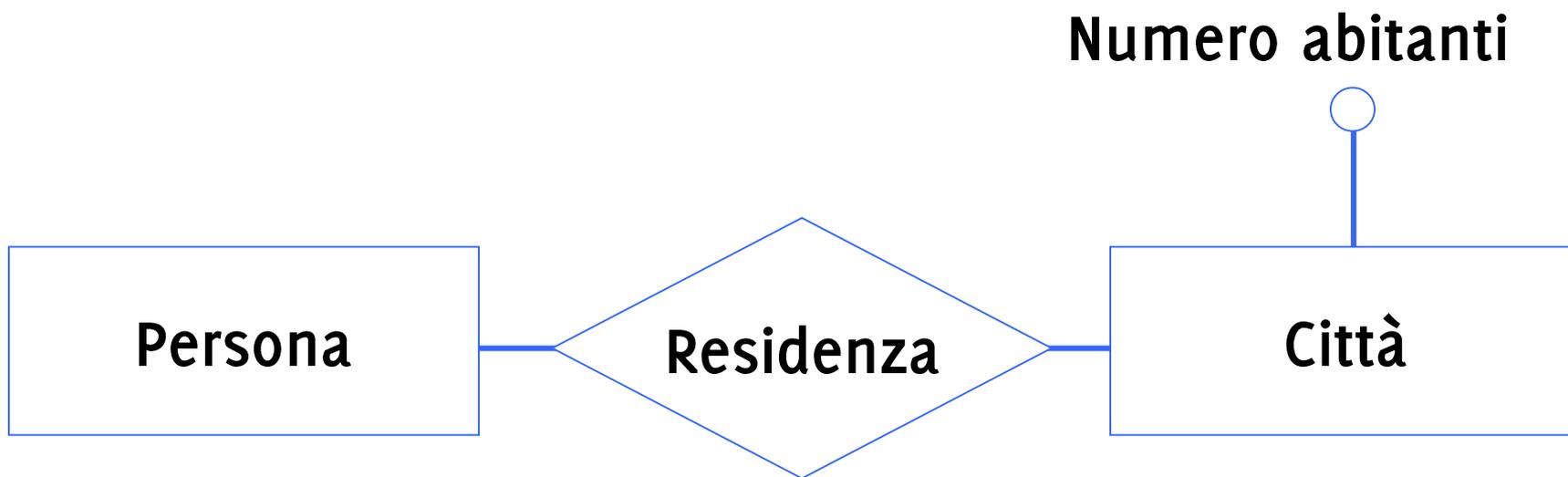
$$ImportoTotale = \sum_{\text{prodotti che compongono}} \text{Quantità} * \text{Prezzo}$$





Attributo derivabile da altra entità

Il numero di abitanti è derivabile conteggiando le occorrenze delle persone nell'associazione residenza





Vantaggi delle ridondanze

- semplificazione delle interrogazioni (il dato esiste senza dover essere derivato)

Svantaggi

- maggiore occupazione di spazio
- aggiornamenti delle informazioni più complicate

Quindi la scelta se rimuovere o mantenere una ridondanza dipende dall'utilizzo atteso del DB (quante letture e quante scritture sono attese su quelle informazioni)



Ristrutturazione dello Schema E-R

Consideriamo le seguenti fasi:

- Analisi delle ridondanze
- **Eliminazione delle generalizzazioni**
- Scelta degli identificatori primari



Eliminazione Delle Generalizzazioni

Nel modello relazionale posso rappresentare facilmente:

- entità
- associazioni

Non posso rappresentare direttamente le **generalizzazioni**

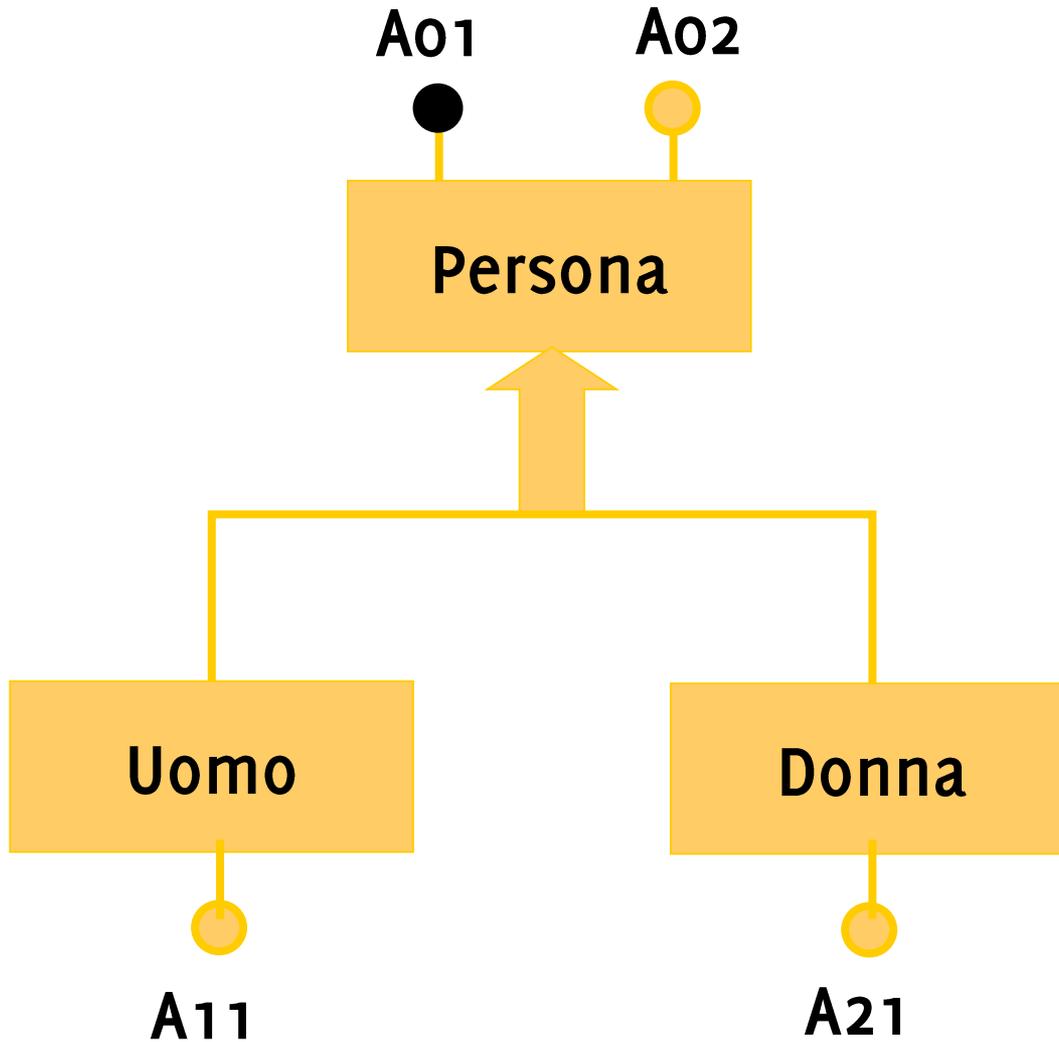
- le gerarchie vanno eliminate e sostituite con entità e associazioni

Tre possibili soluzioni:

1. accorpare le figlie della generalizzazione nel genitore
2. accorpare il genitore della generalizzazione nelle figlie
3. sostituzione della generalizzazione con associazioni

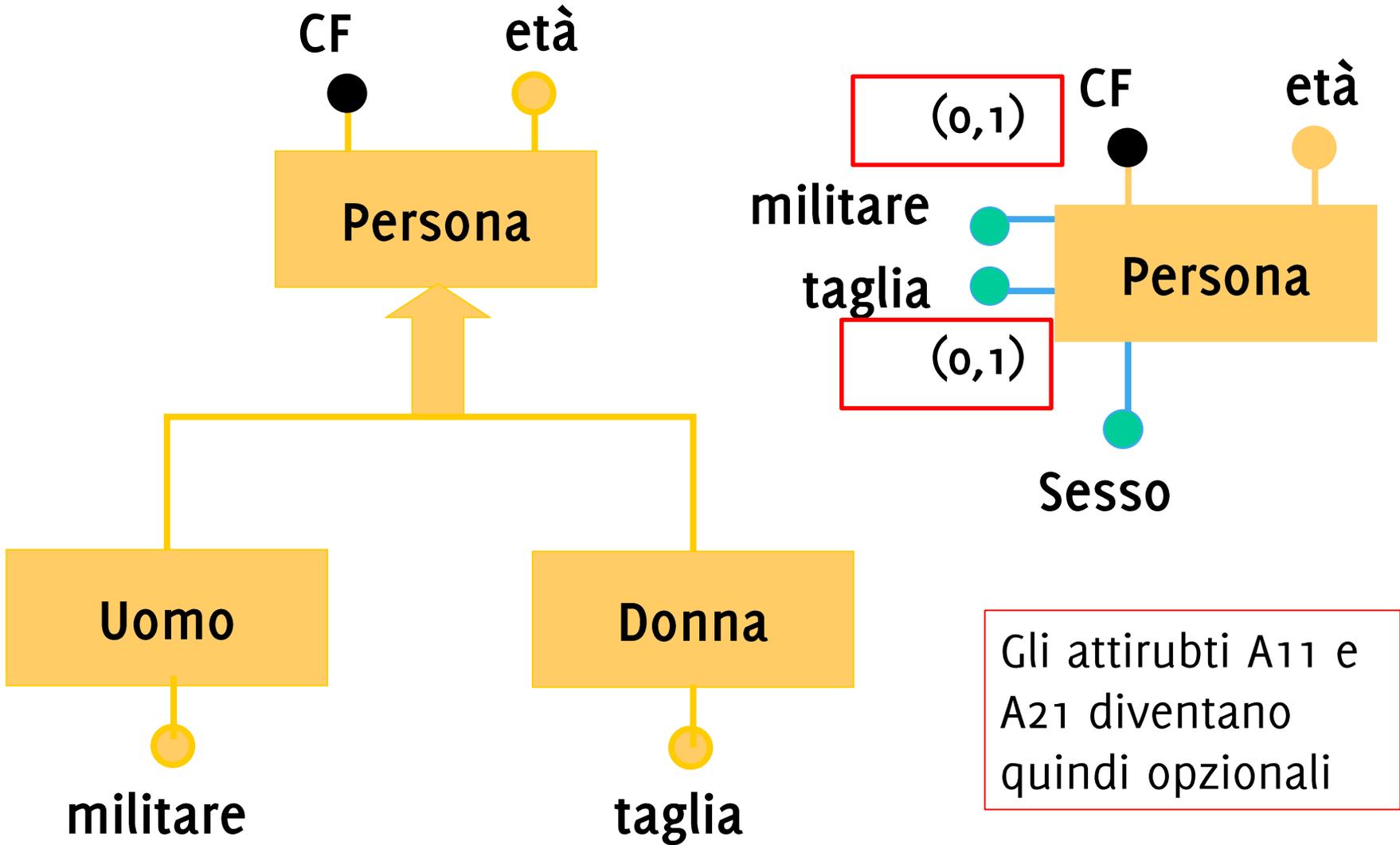


Esempio: accorpare le figlie nel genitore



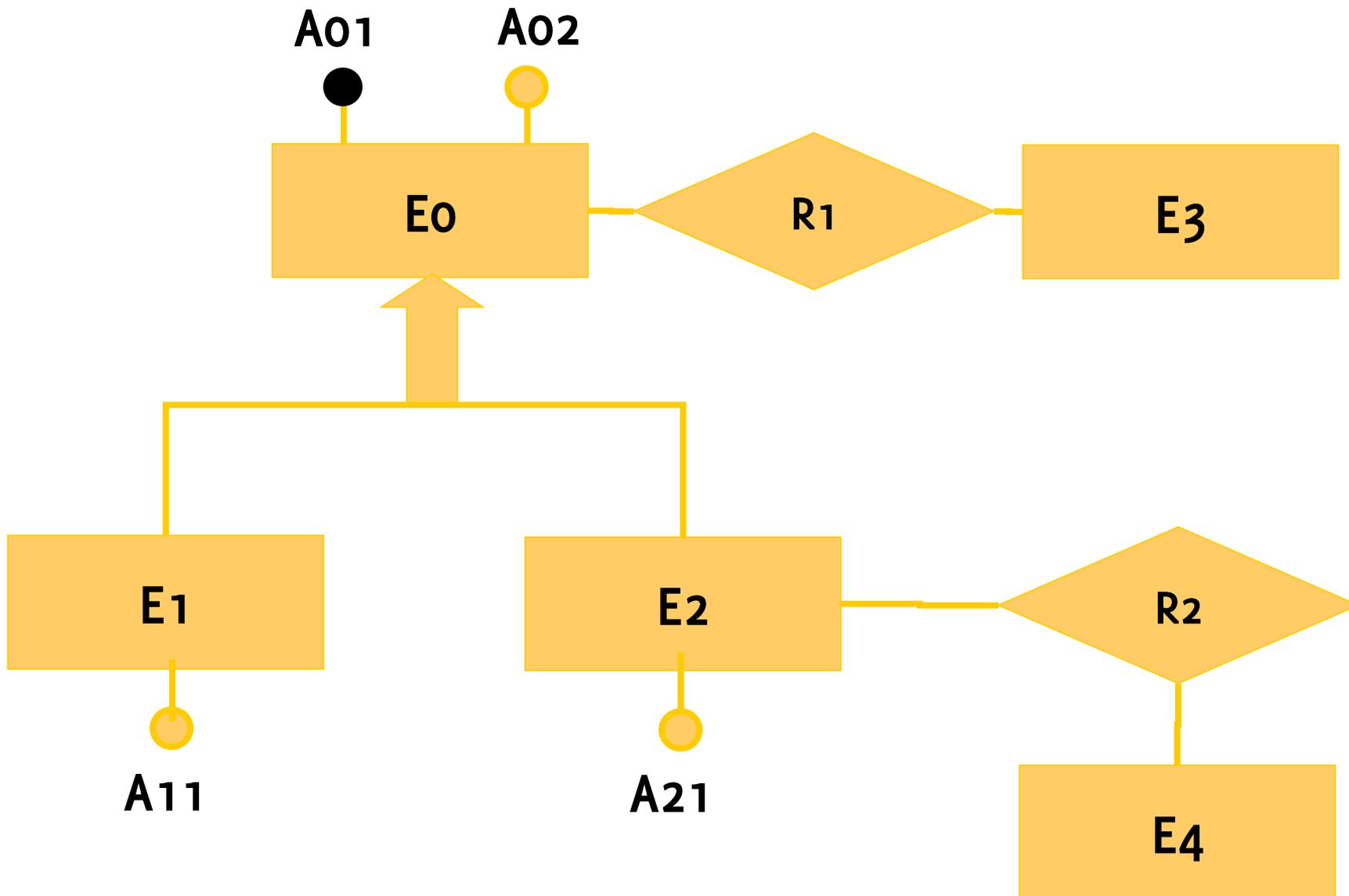


Esempio: accorpare le figlie nel genitore



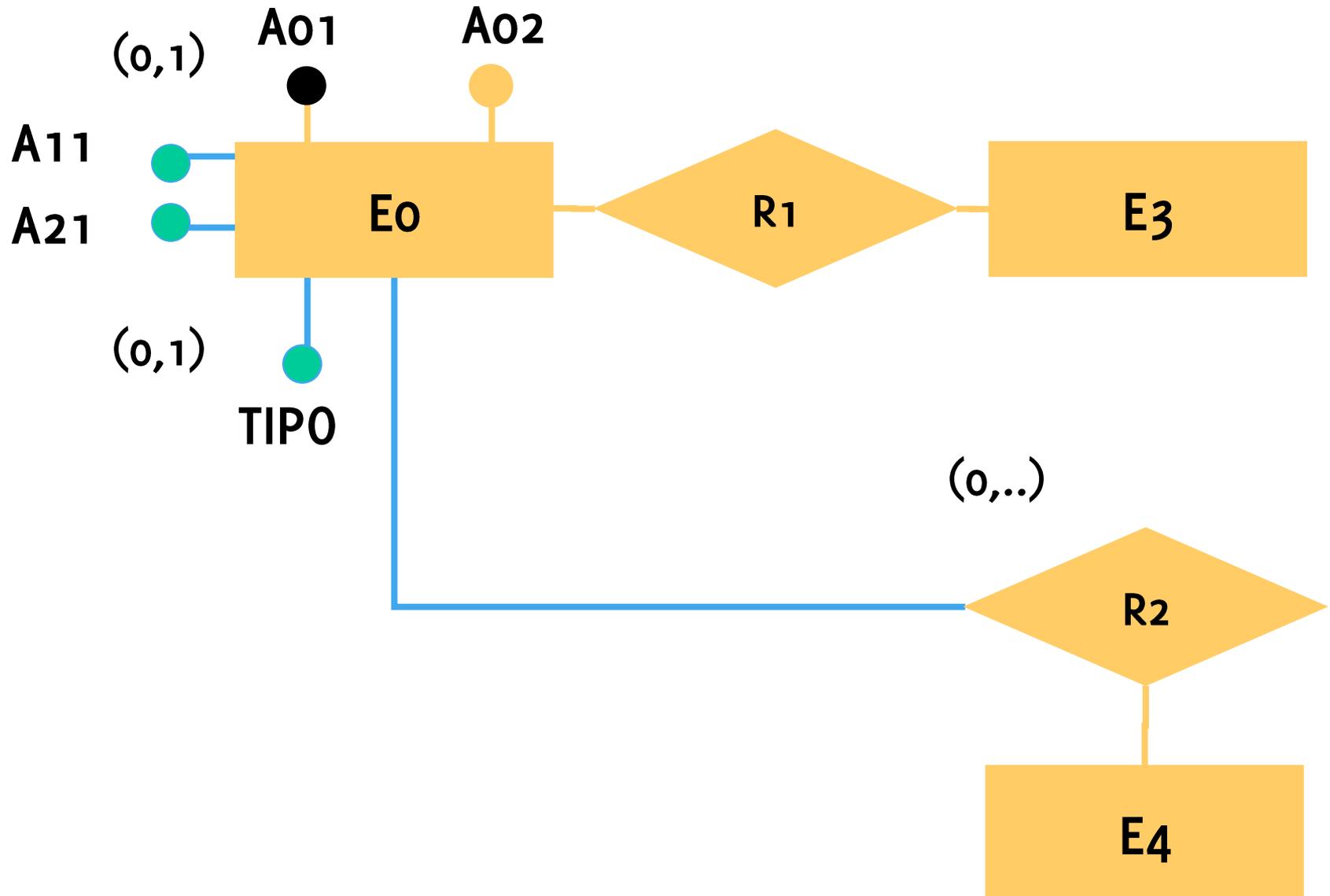


Esempio di Schema da semplificare



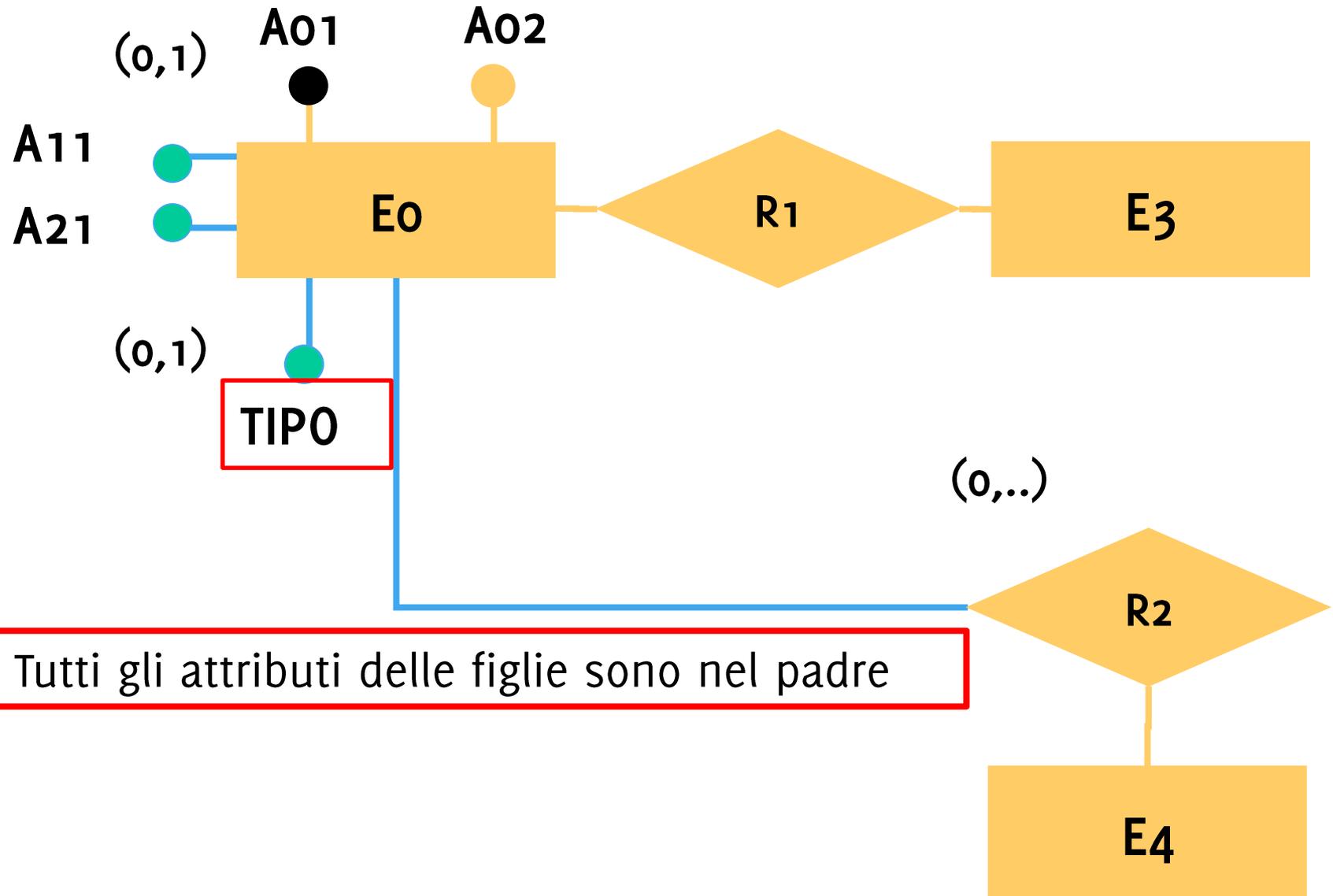


Accorpare le figlie nel genitore





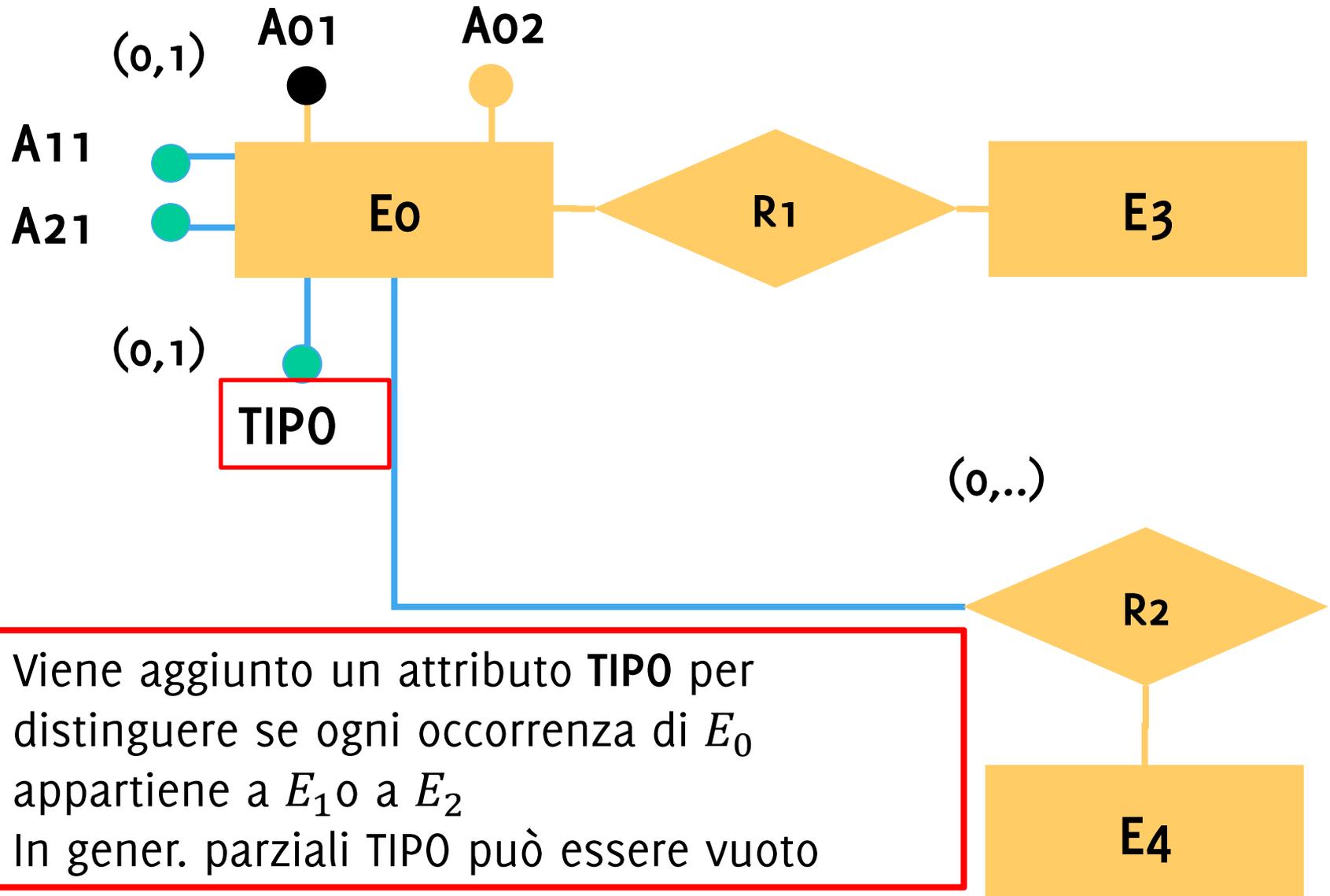
Accorpare le figlie nel genitore



- Tutti gli attributi delle figlie sono nel padre

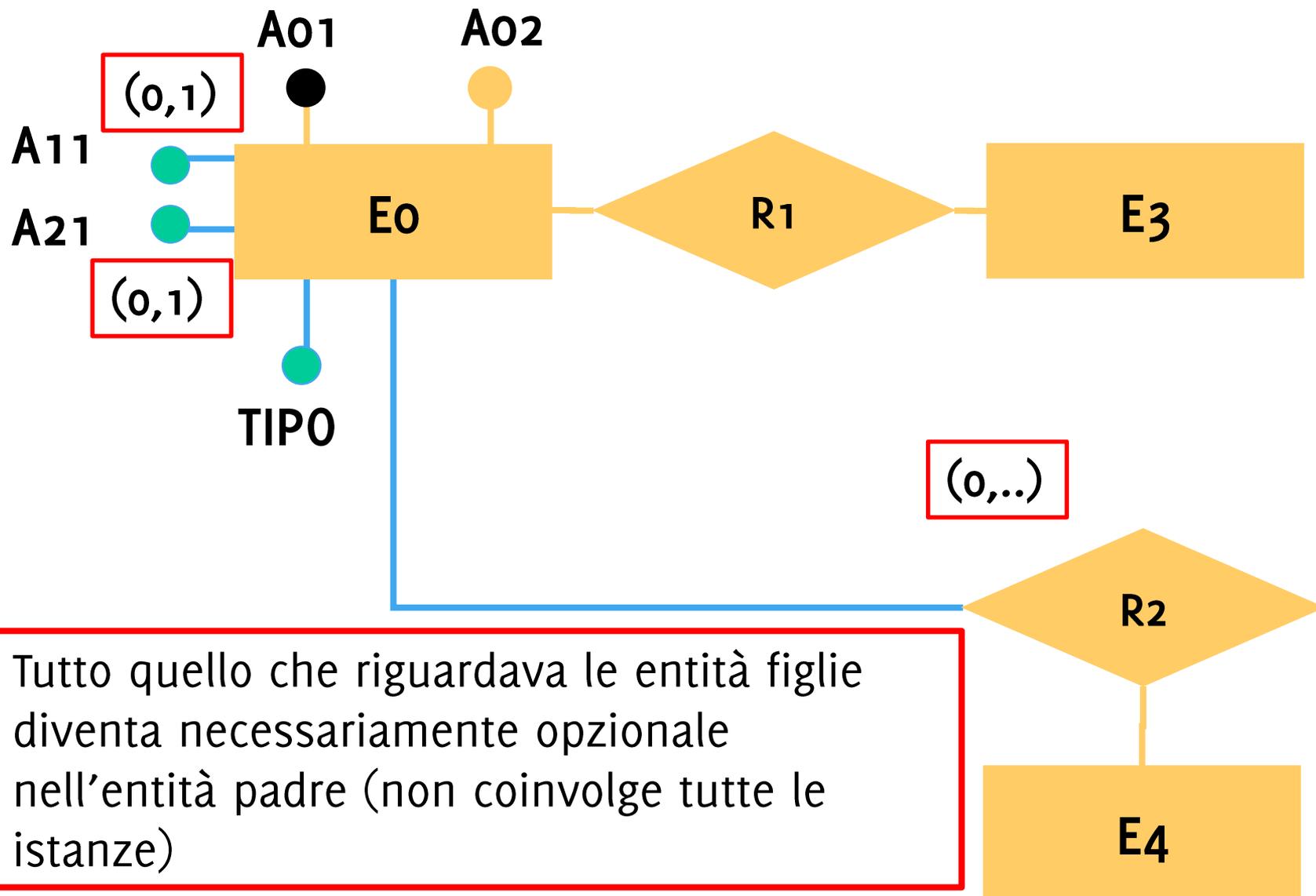


Accorpare le figlie nel genitore





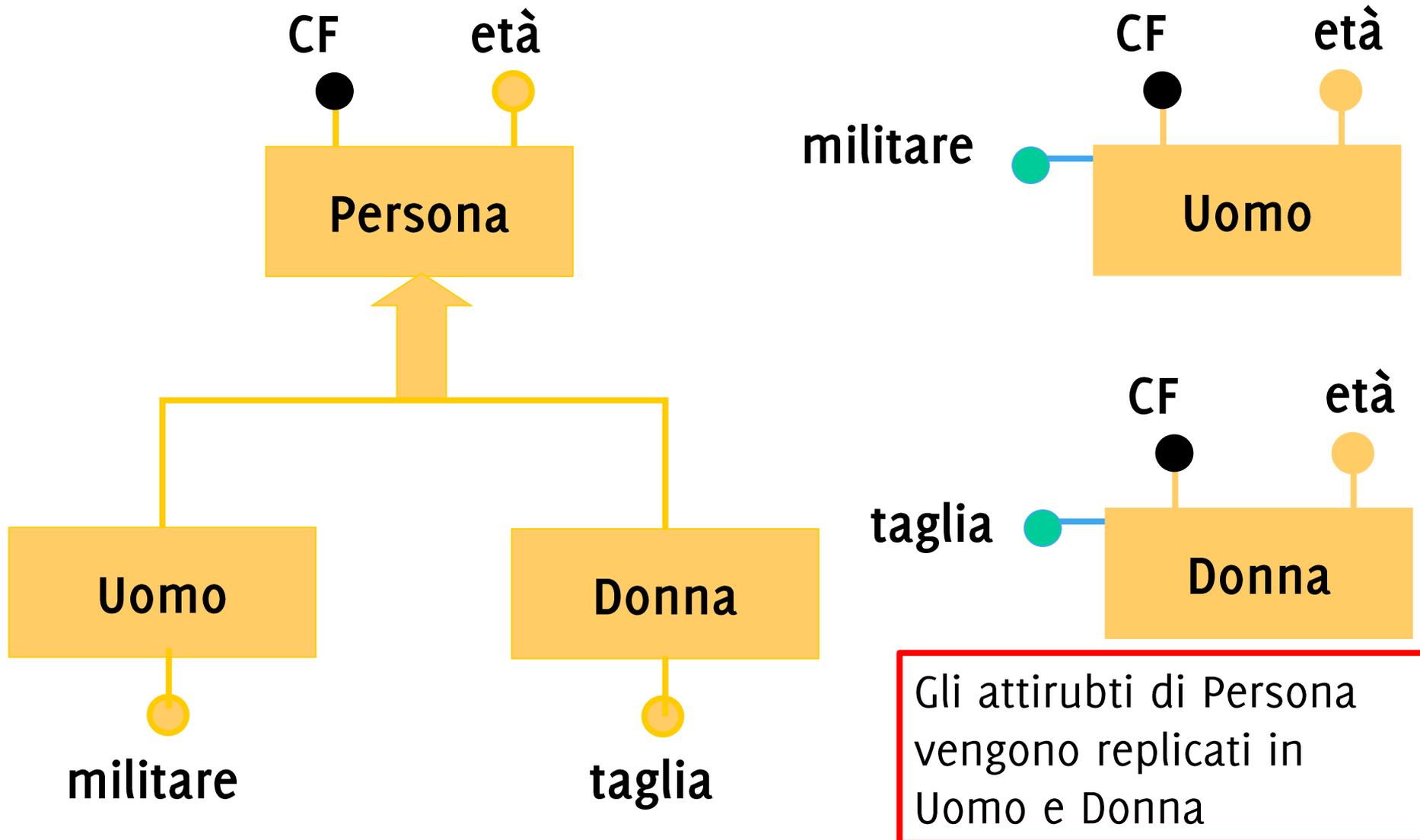
Accorpare le figlie nel genitore



- Tutto quello che riguardava le entità figlie diventa necessariamente opzionale nell'entità padre (non coinvolge tutte le istanze)

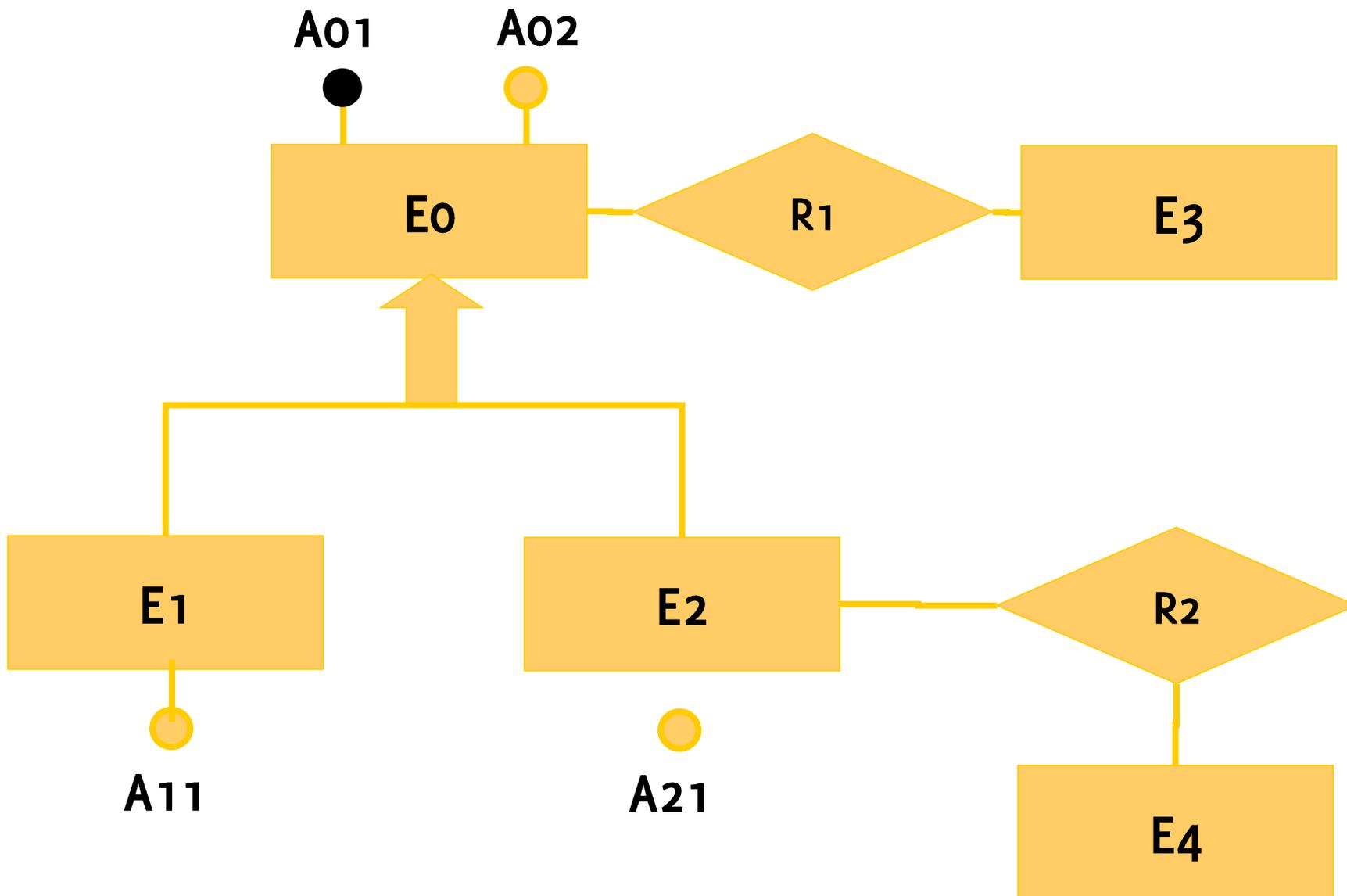


Esempio: accorpare il genitore nelle figlie



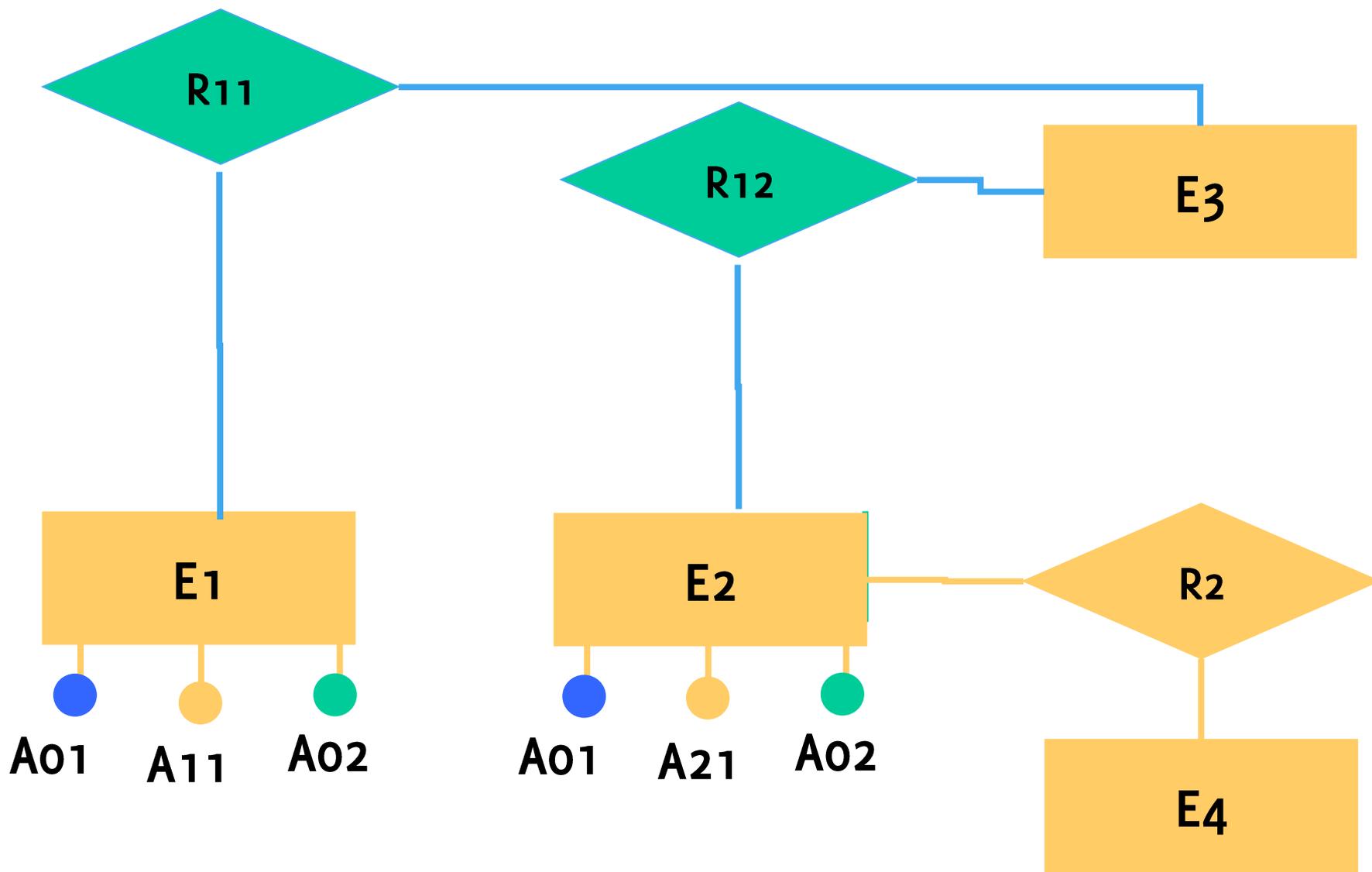


Esempio di Schema da semplificare



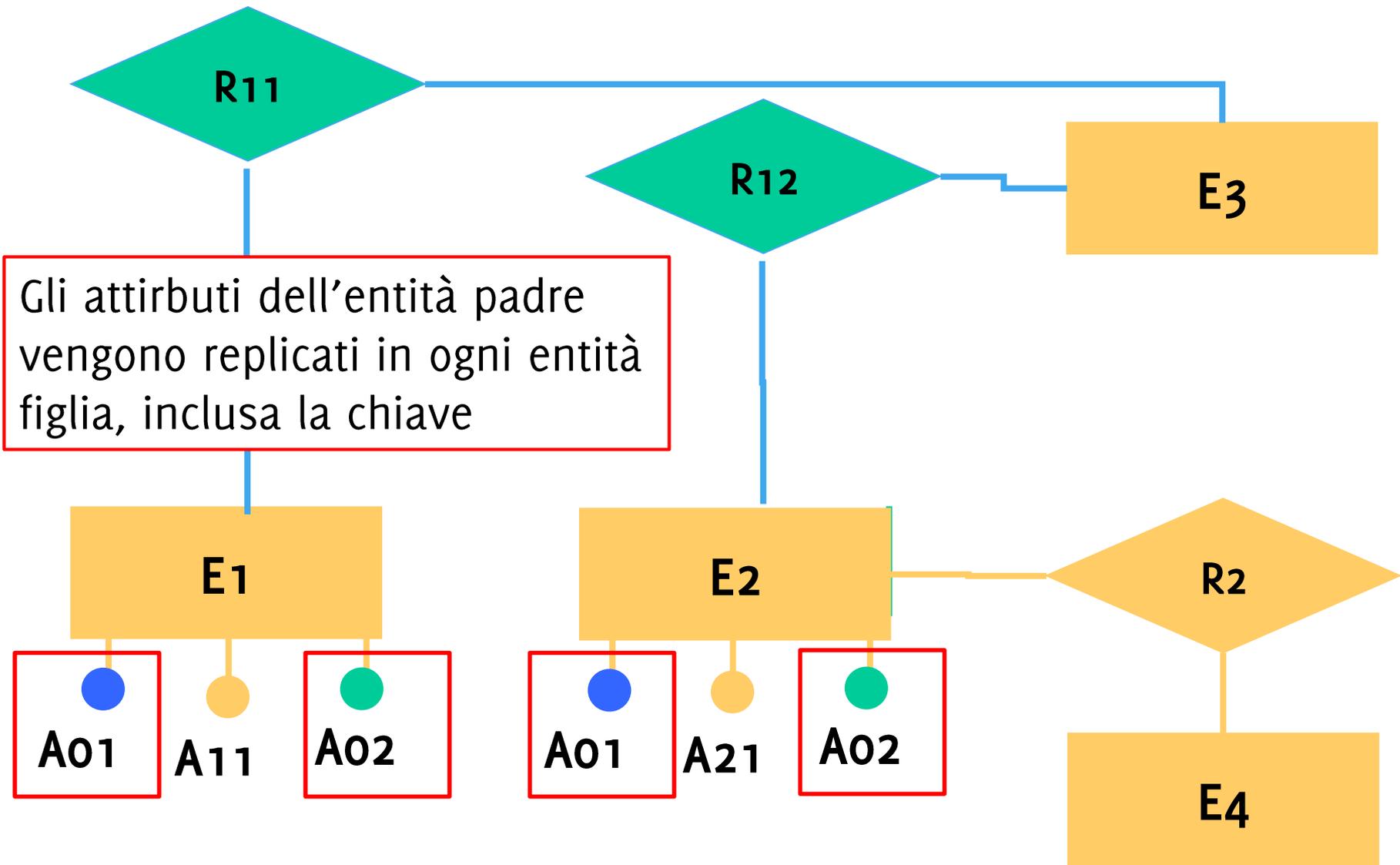


Accorpare il genitore nelle figlie



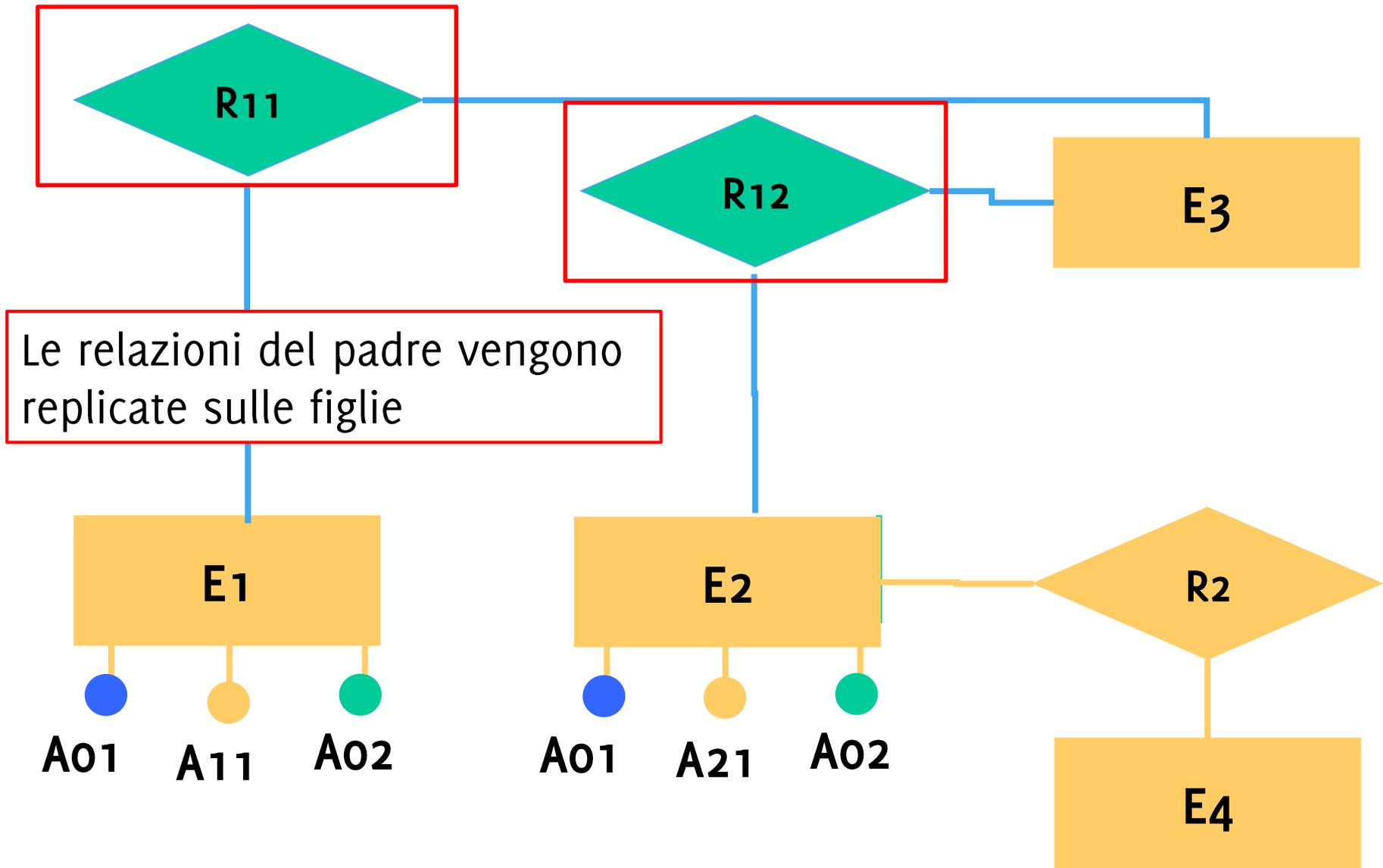


Accorpare il genitore nelle figlie



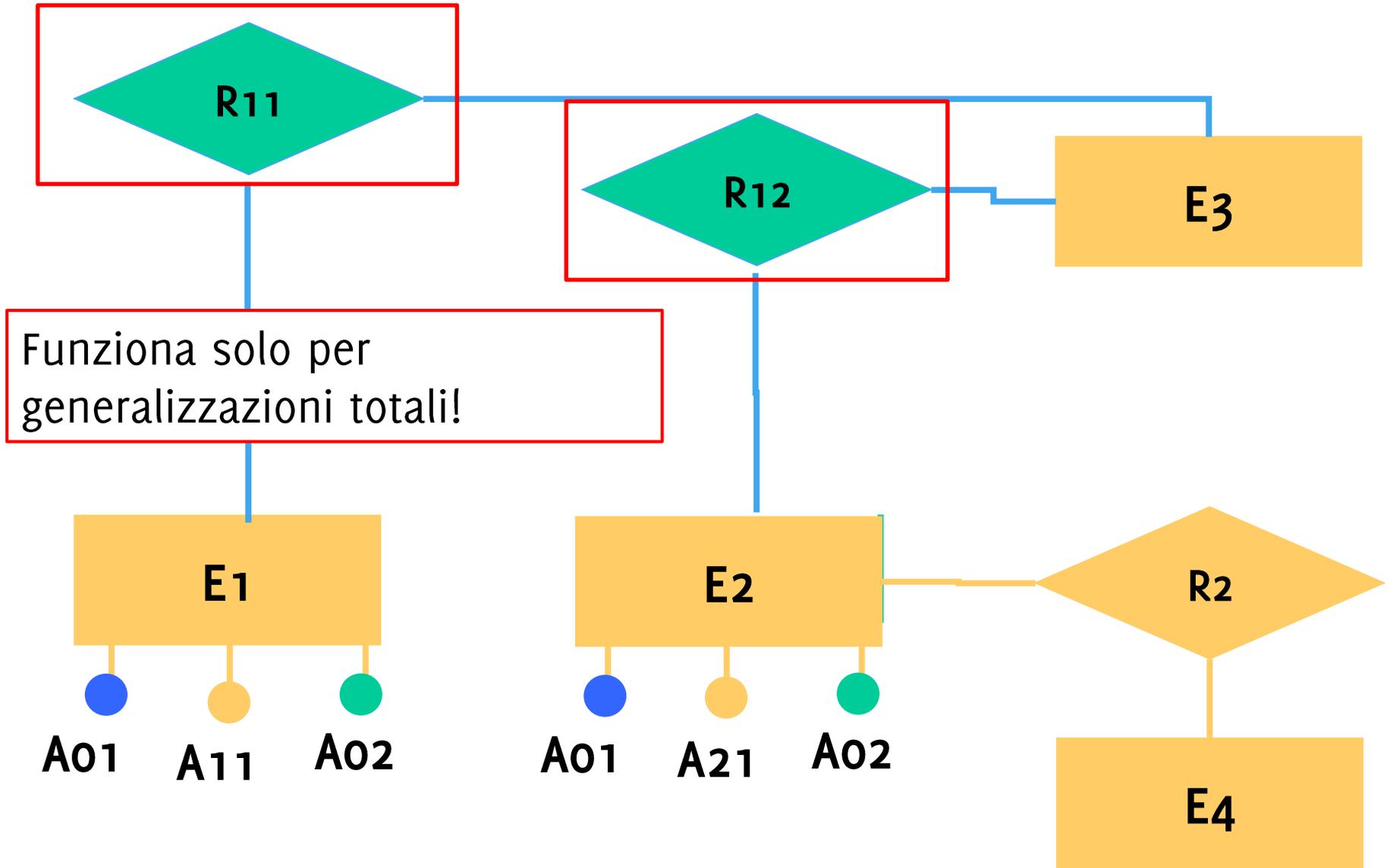


Accorpare il genitore nelle figlie



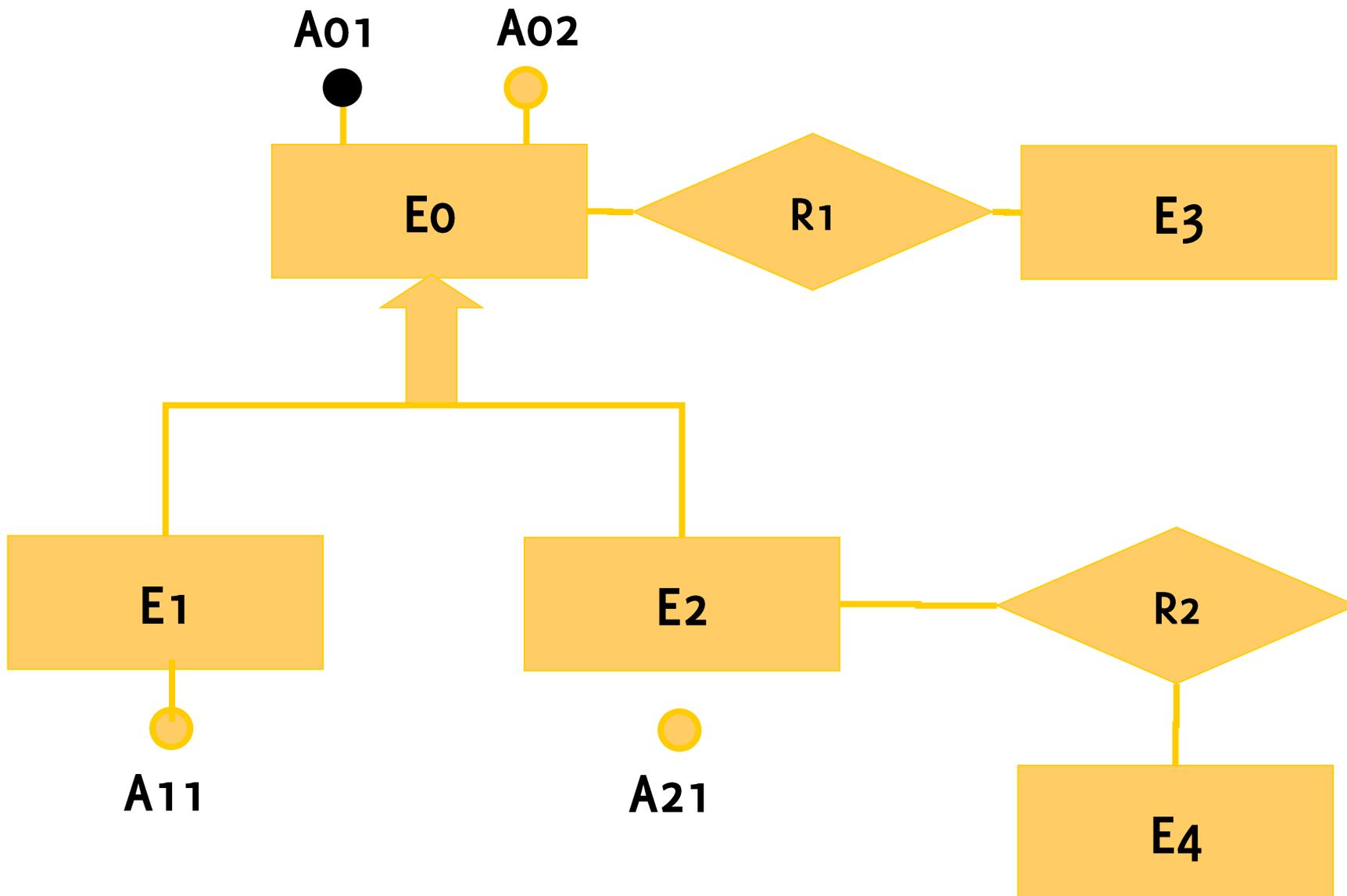


Accorpare il genitore nelle figlie



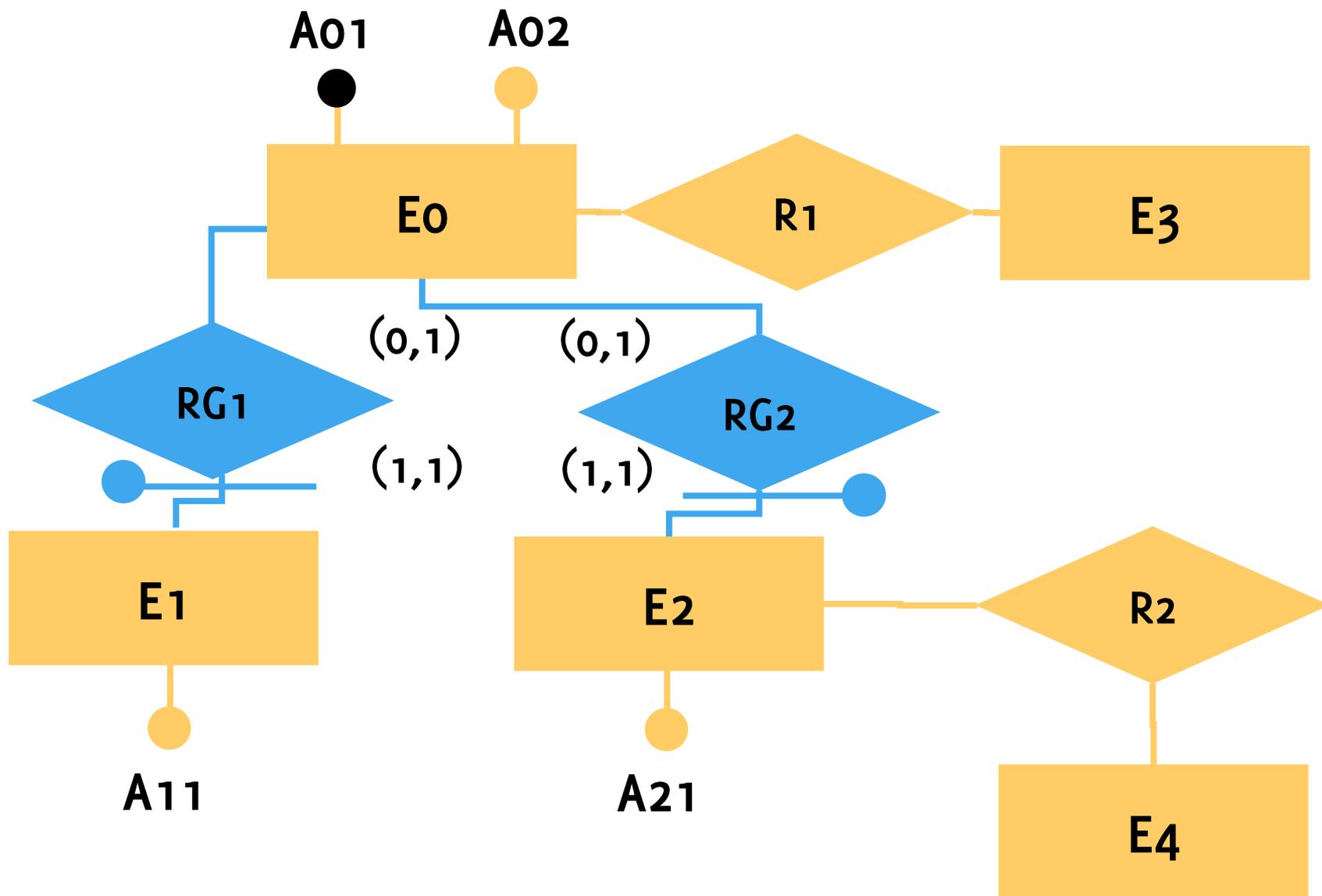


Esempio di Schema da semplificare



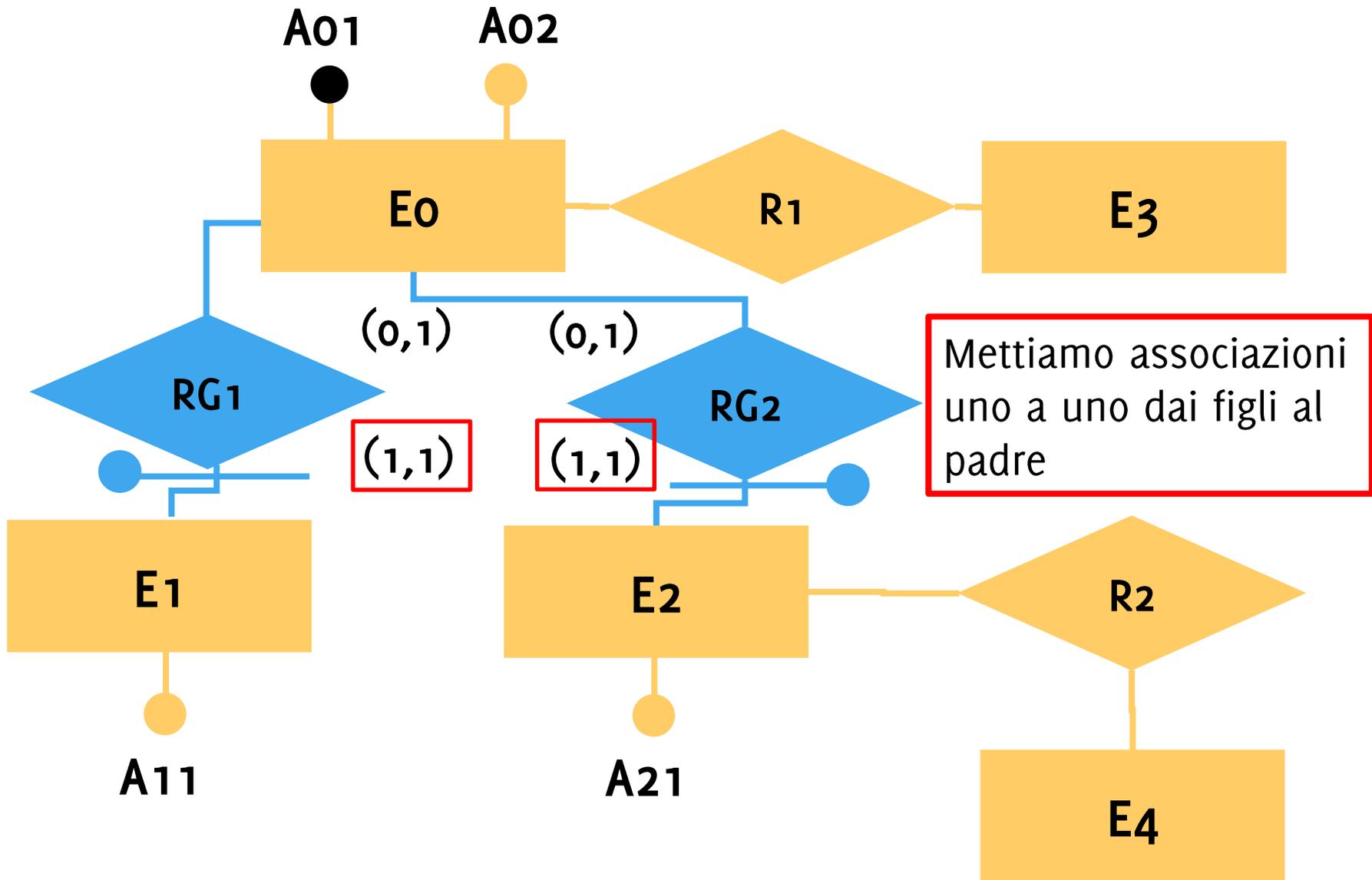


Sostituire la generalizzazione con associazioni



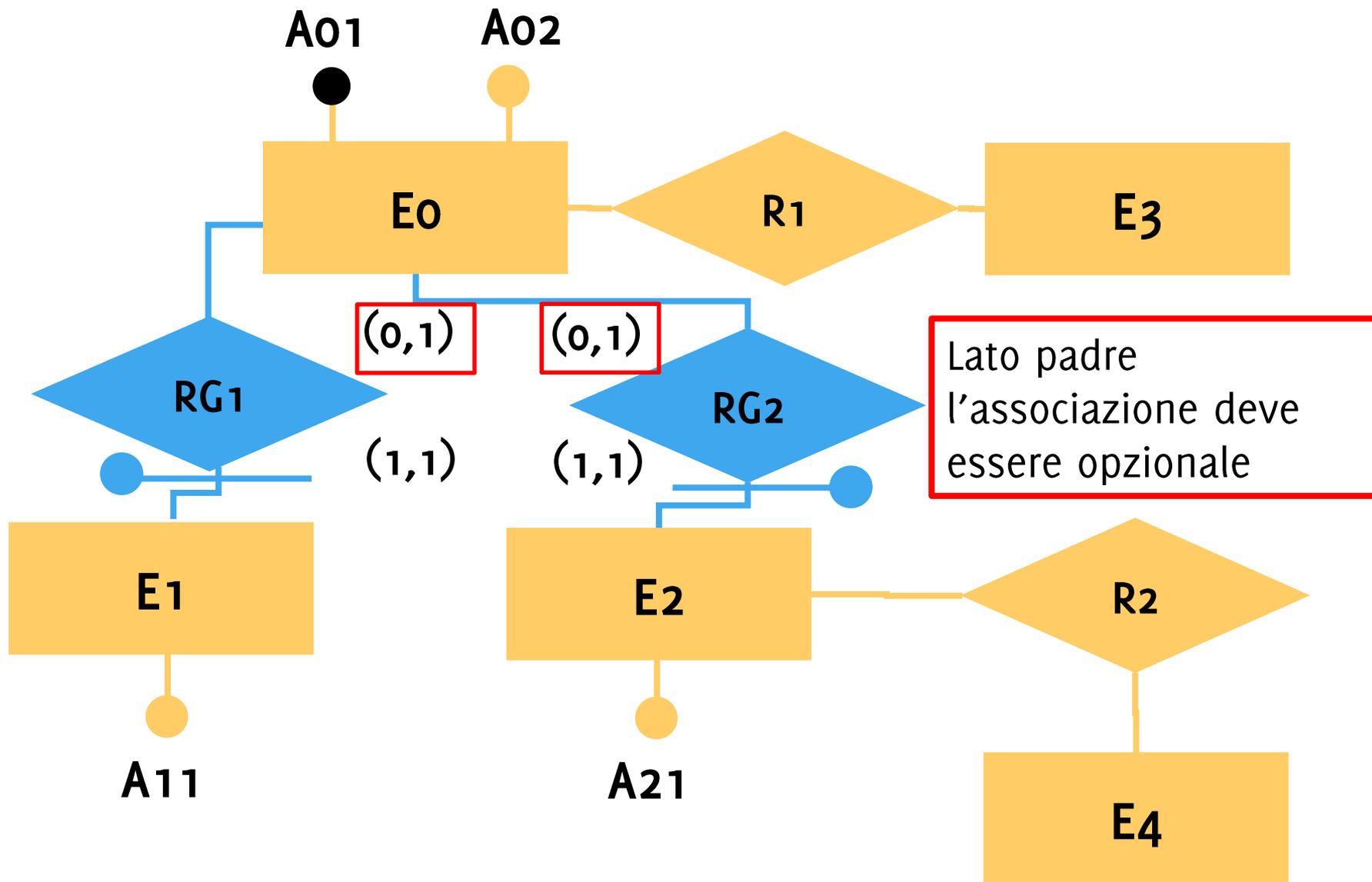


Sostituire la generalizzazione con associazioni



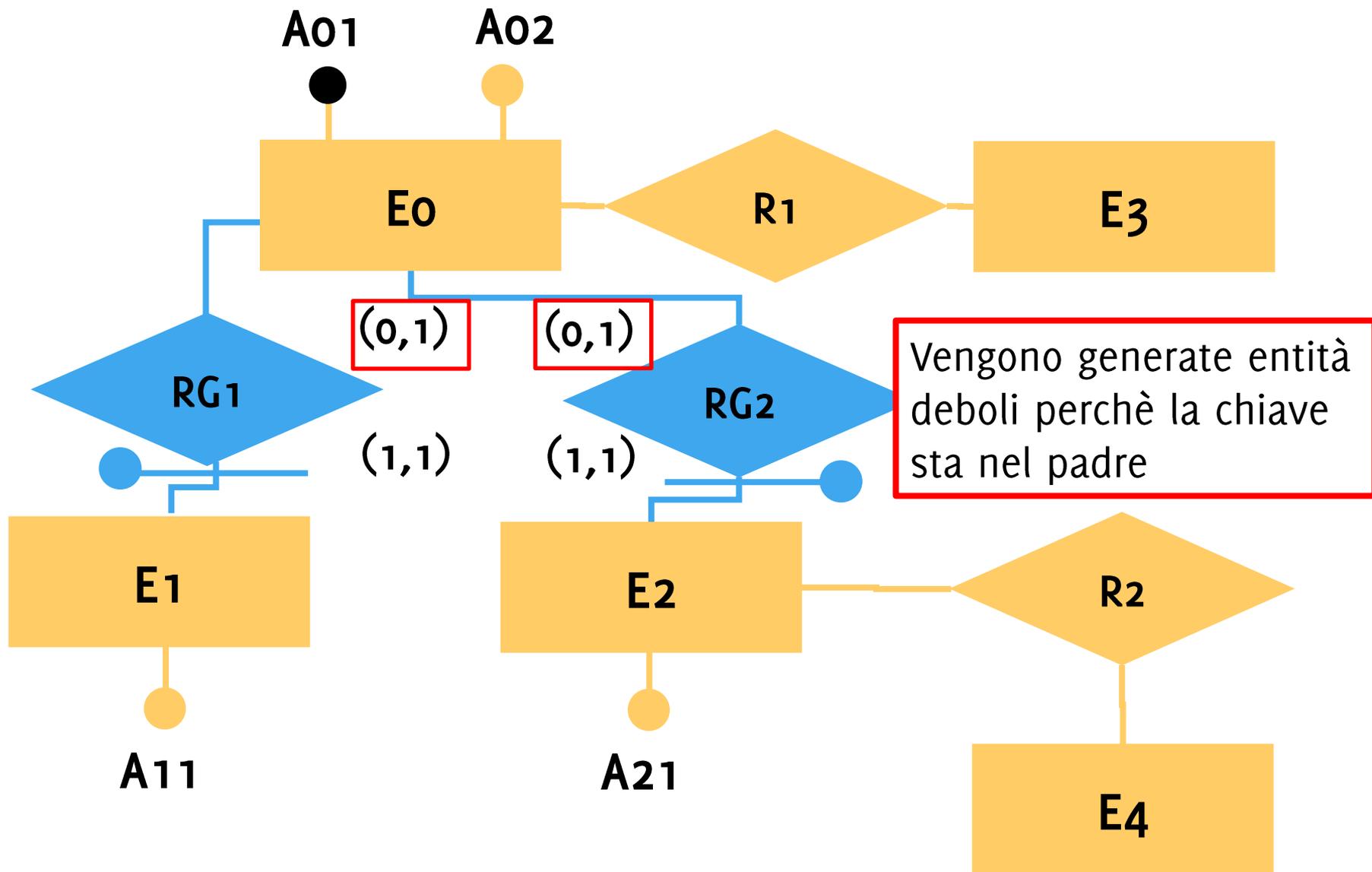


Sostituire la generalizzazione con associazioni



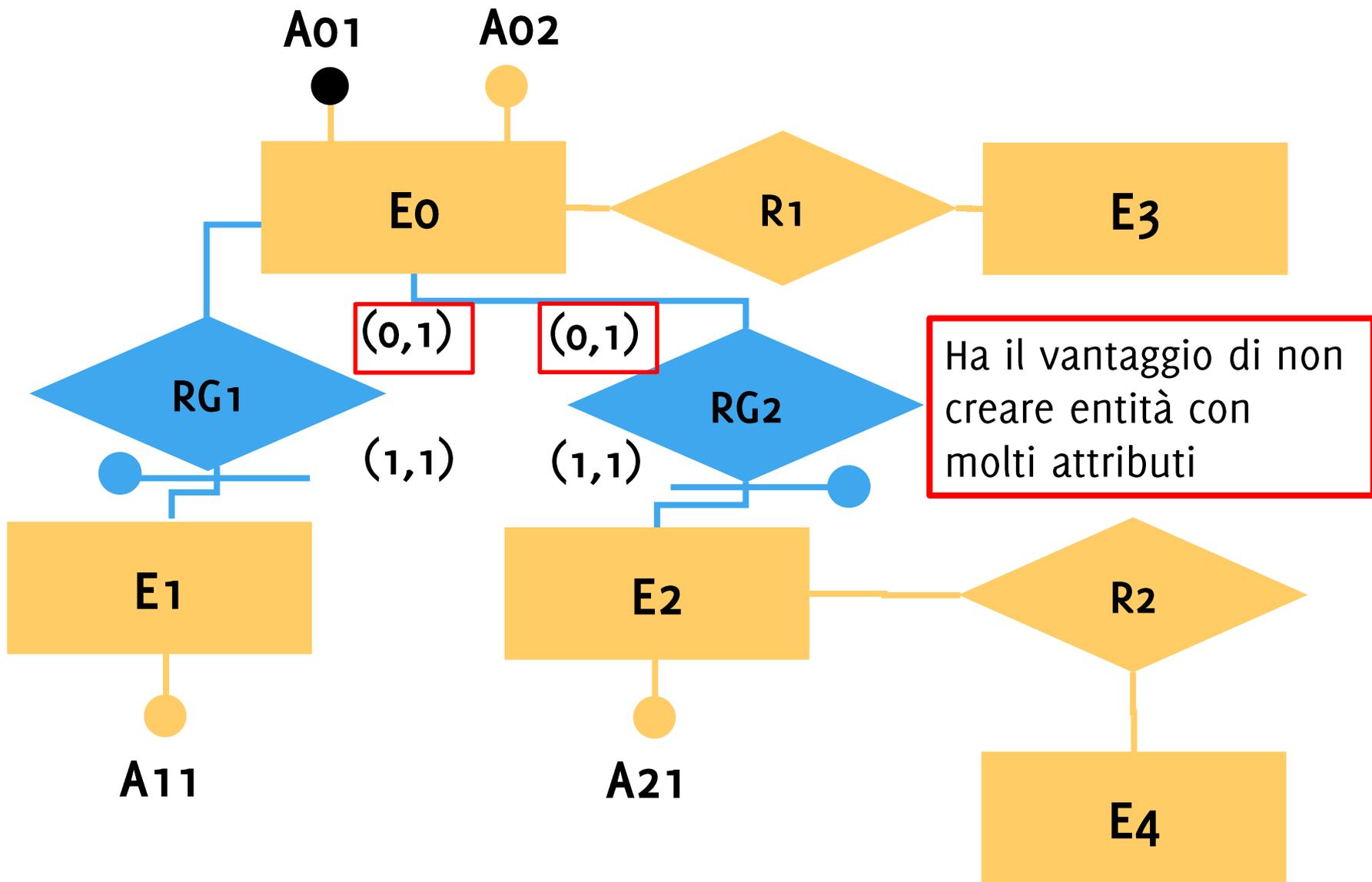


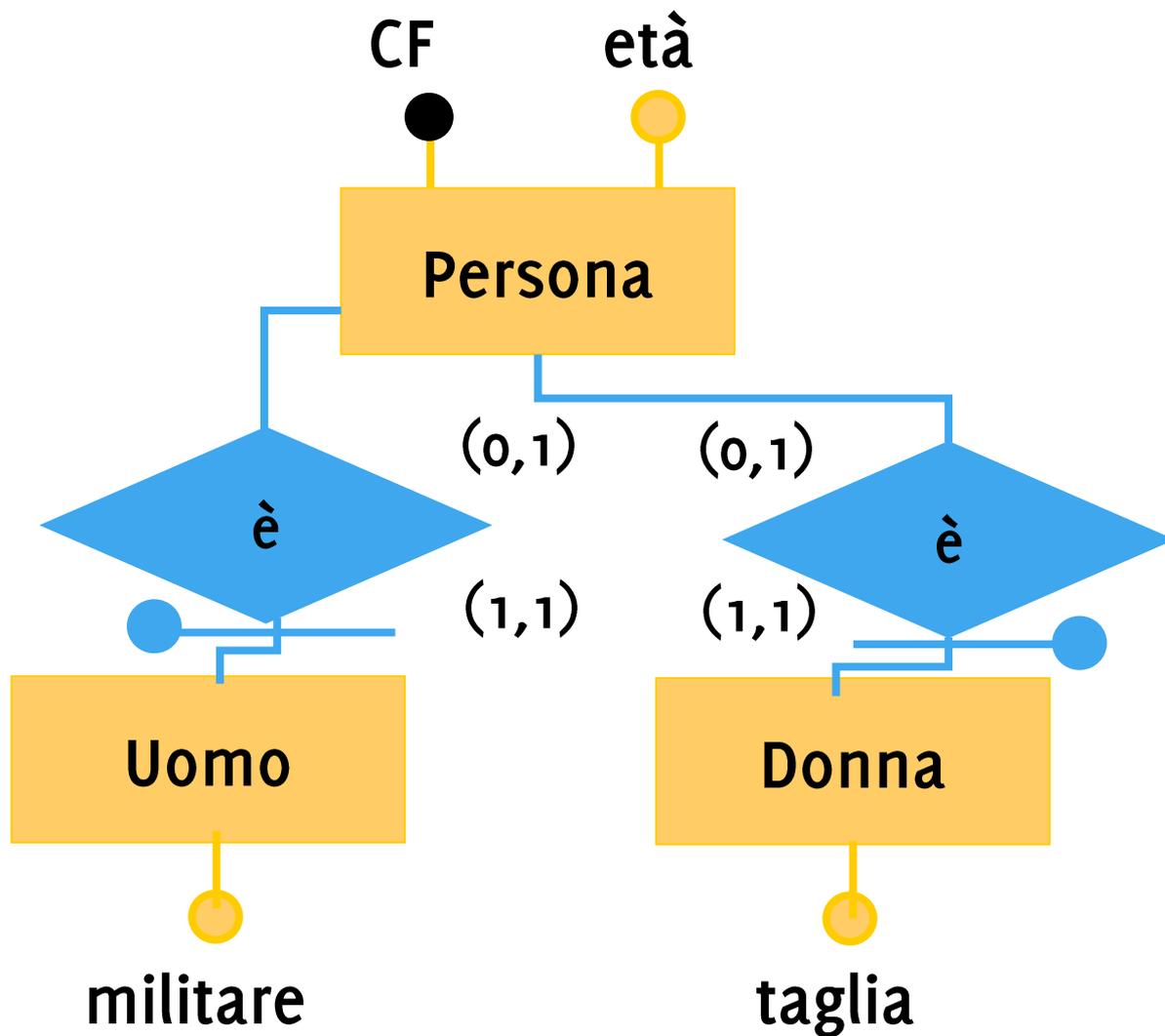
Sostituire la generalizzazione con associazioni





Sostituire la generalizzazione con associazioni







Ristrutturazione dello Schema E-R

Consideriamo le seguenti fasi:

- Analisi delle ridondanze
- Eliminazione delle generalizzazioni
- **Scelta degli identificatori primari**



Scelta degli identificatori primari

È un'operazione indispensabile per la traduzione nel modello relazionale

Criteri per le chiavi

- Assenza di opzionalità (non devono ammettere valori nulli)
- Semplicità (è preferibile considerare uno o pochi identificatori per ogni entità)

Se nessun candidato identificatore soddisfa i requisiti:

- Si introducono nuovi attributi – **codici** -- contenenti valori speciali generati appositamente per questo scopo



Traduzione da Schema E-R a Modello Relazionale

Tradurre lo Schema E-R semplificato in un Modello Relazionale Equivalente



Nella traduzione di uno **schema E-R** in uno **schema relazionale** valgono i seguenti principi:

- le **entità diventano relazioni** aventi stessi attributi
- le **associazioni diventano relazioni** che coinvolgono gli identificatori delle entità coinvolte (oltre agli attributi dell'associazione)

La traduzione delle entità è immediata

La traduzione delle associazioni dipende dalla molteplicità dell'associazione:

- Molti a molti
- Uno a molti
- Uno a uno
- Entità deboli

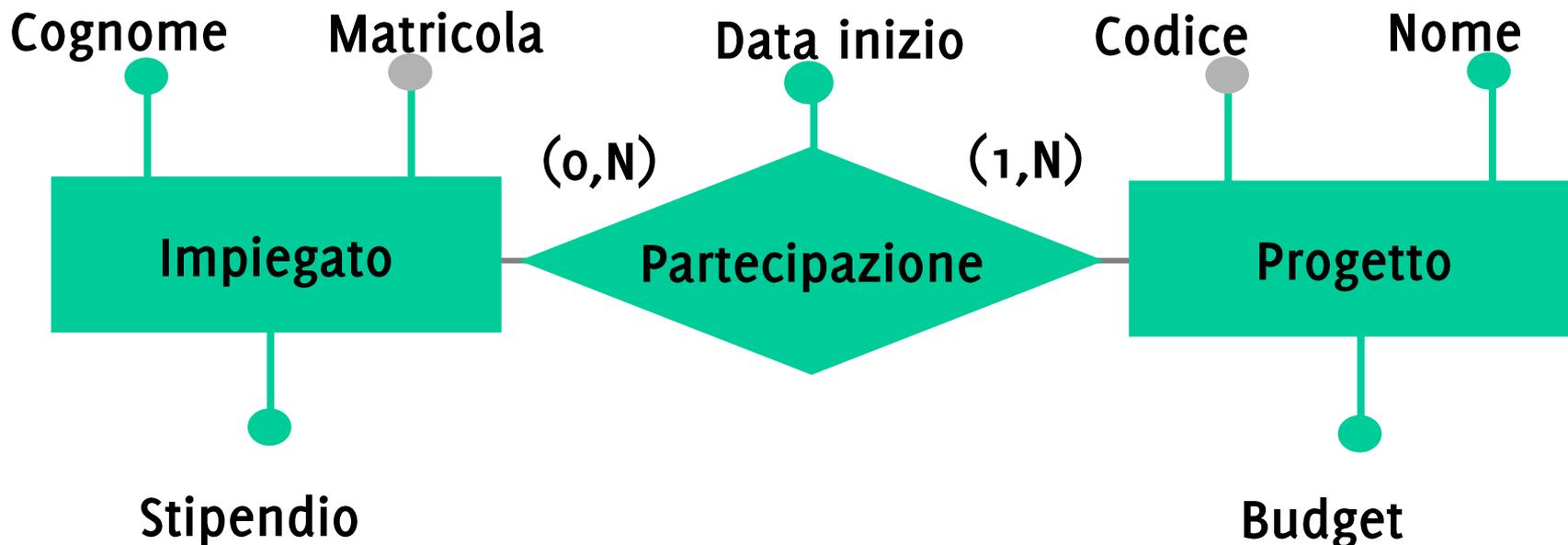


Associazione Molti a Molti

Ogni **entità** coinvolta nell'associazione **diventa una relazione** con gli stessi attributi

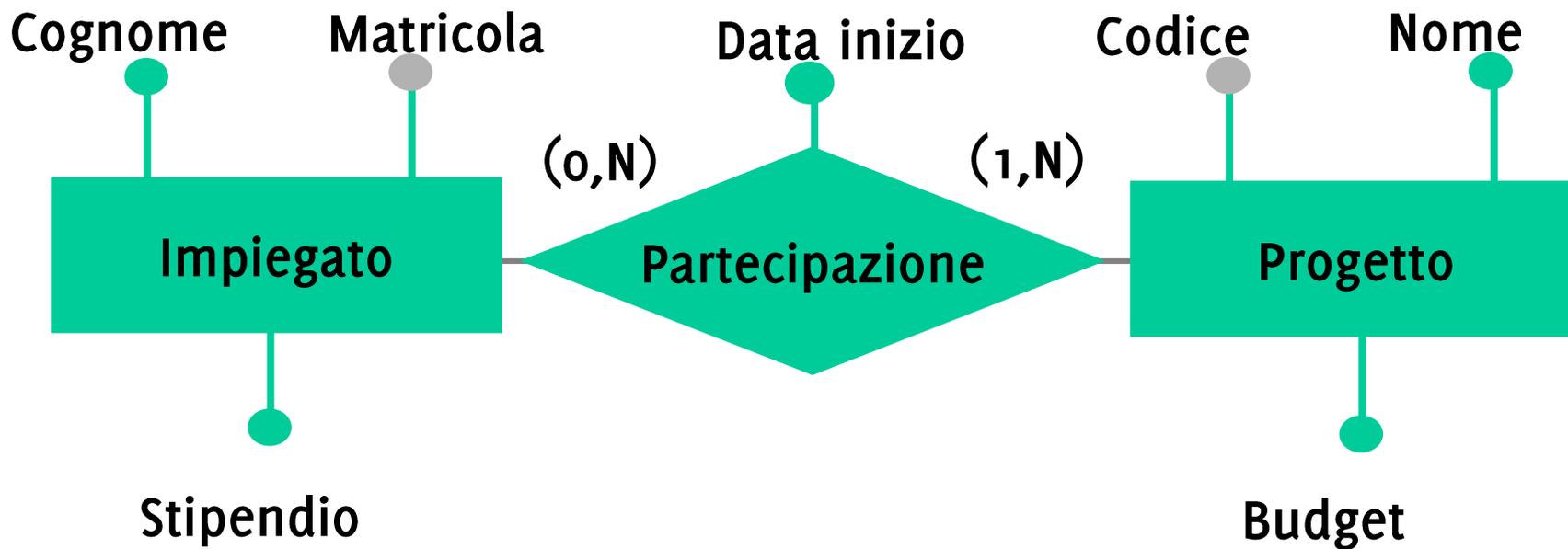
L'associazione diventa una **relazione** con attributi:

- I suoi **attributi originali**
- Le **chiavi delle entità** coinvolte (con **vincoli integrità referenziale**)





Associazione Molti a Molti



Impiegato(Matricola, Cognome, Stipendio)

Progetto(Codice, Nome, Budget)

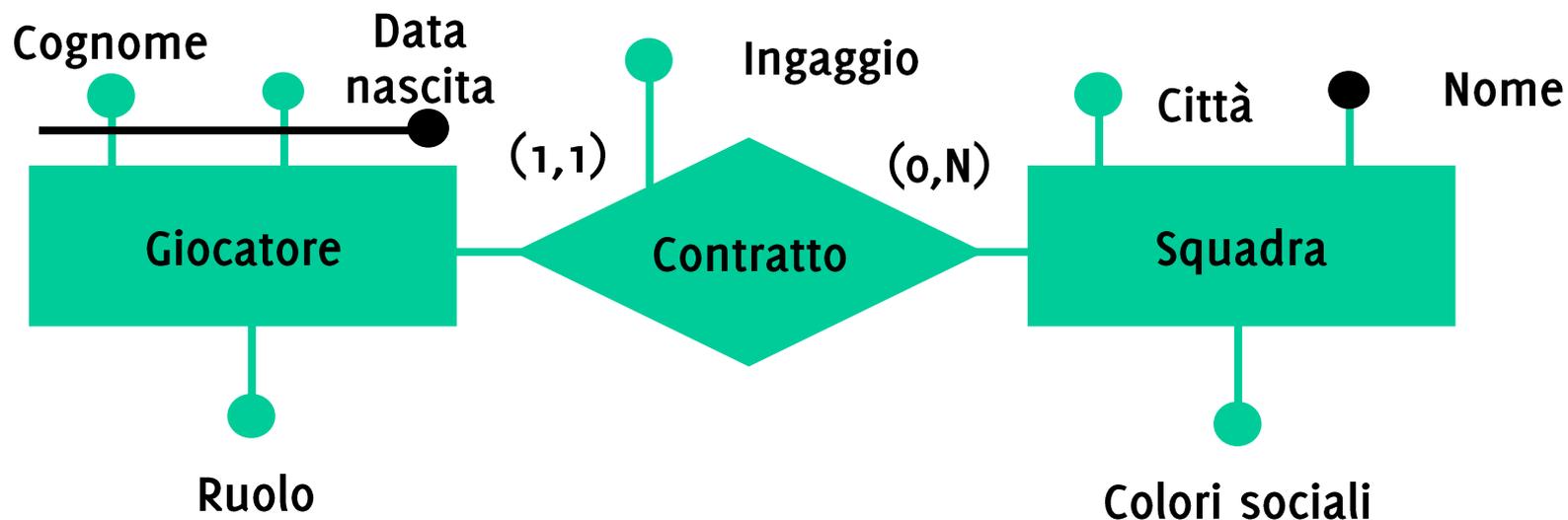
Partecipazione(Matricola, Codice, DataInizio)



Relazione uno a molti

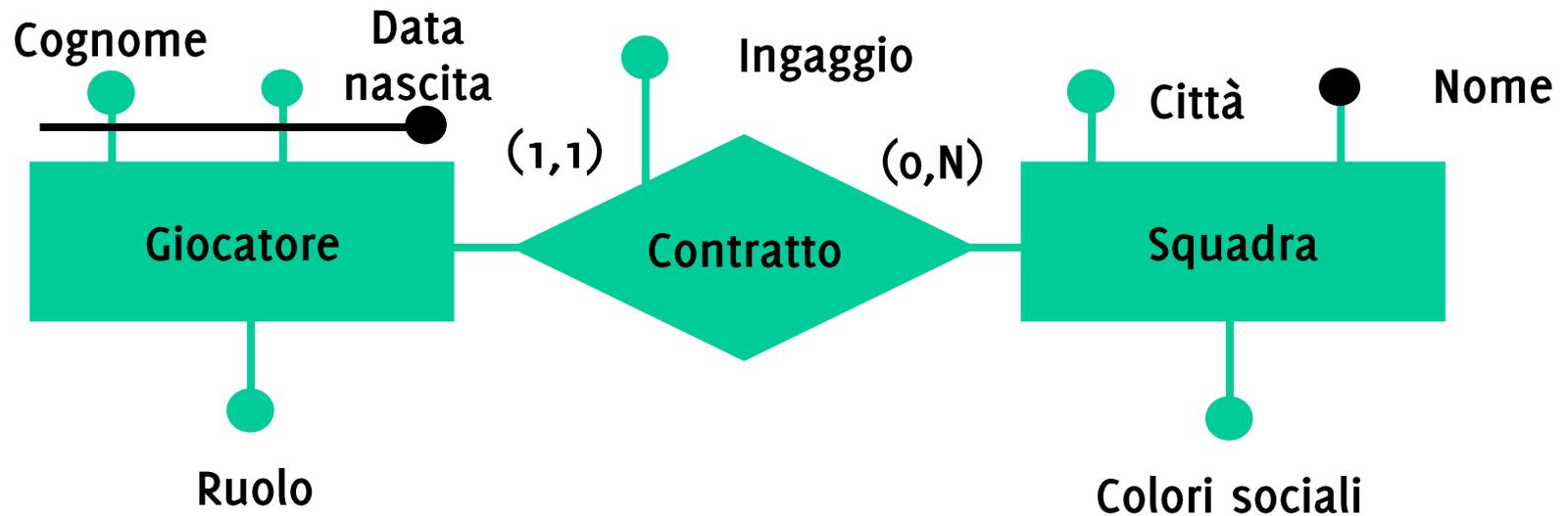
Ogni **entità** coinvolta nell'associazione **diventa una relazione** con gli stessi attributi

L'associazione non ha bisogno di un'entità a se stante, basta includere la chiave dell'entità «lato molti» negli attributi dell'entità «lato uno»





Esempio



Giocatore(Cognome, DataNasc, Ruolo, Squadra, Ingaggio)

Squadra(Nome, Città, ColoriSociali)

- con vincolo di integrità referenziale fra Squadra in Giocatore e la chiave di Squadra
- se la cardinalità minima lato Giocatore fosse 0, allora Squadra in Giocatore deve ammettere valore nullo

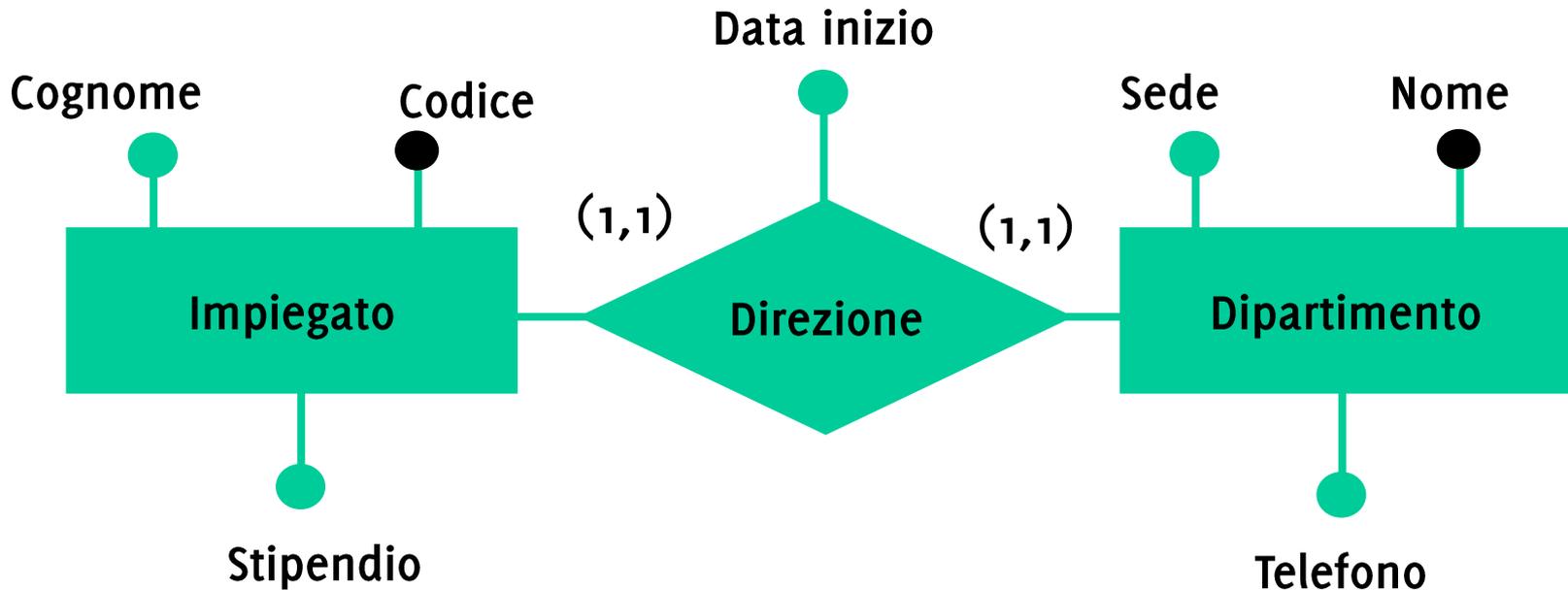


In generale:

- Creo due relazioni corrispondenti alle entità
- Non serve alcuna relazione relativa all'associazione, basta mettere una chiave di un'entità come attributo dell'altra
- Gli attributi dell'associazione vengono inseriti nell'entità che prende la chiave
- Se non ci sono opzionalità l'associazione uno a uno può essere tradotta indifferentemente in due modi



Associazioni Uno a Uno



Può diventare:

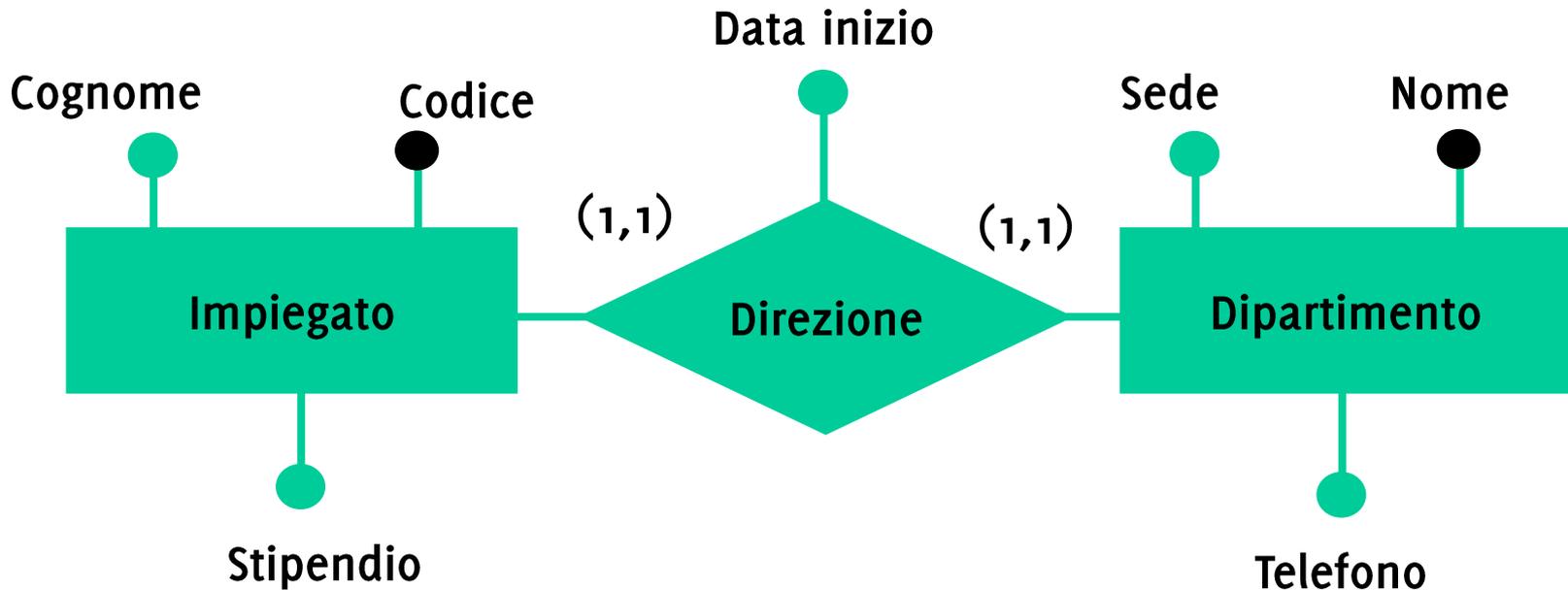
Impiegato (Codice, Cognome, Stipendio)

Dipartimento (Nome, Sede, Telefono, Direttore, InizioD)





Associazioni Uno a Uno



Oppure:

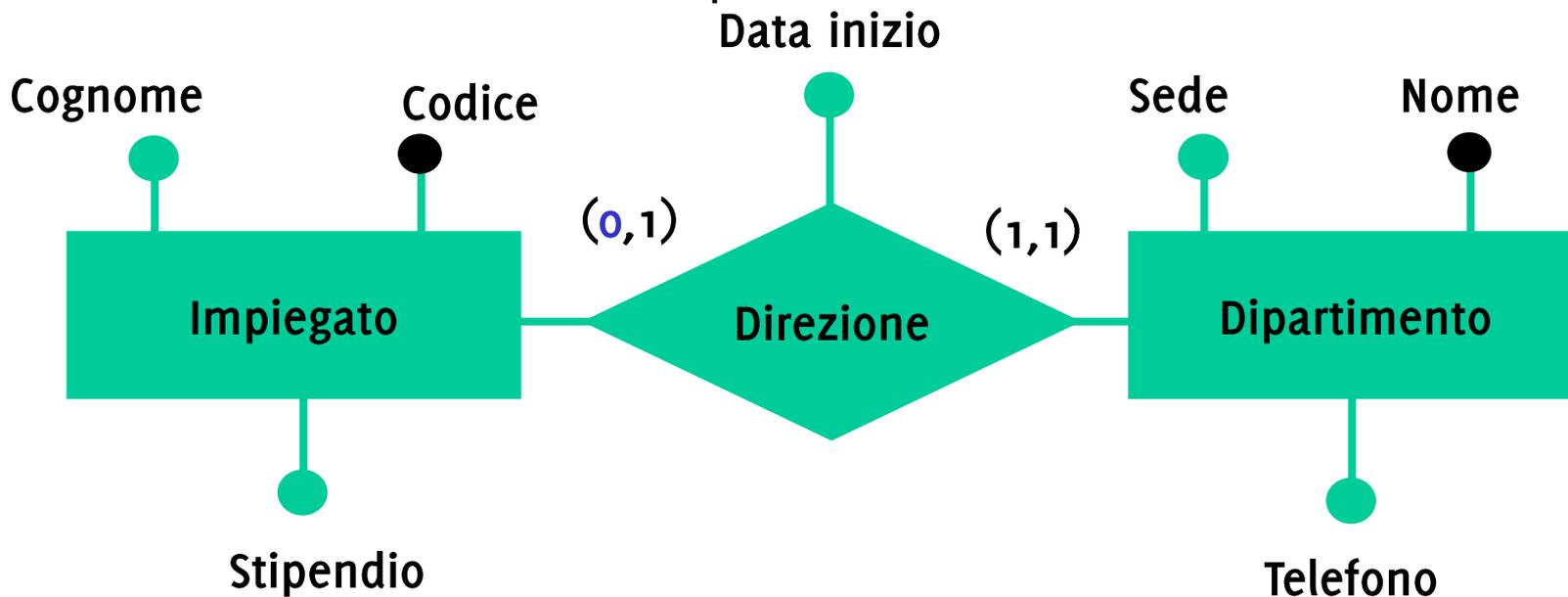
Impiegato (Codice, Cognome, Stipendio, DipartimentoDiretto, InizioD)

Dipartimento (Nome, Sede, Telefono,)



Associazioni Uno a Uno con opzionalità

Nel caso in cui ci fosse un'opzionalità nell'associazione



Allora è preferibile

Impiegato (Codice, Cognome, Stipendio)

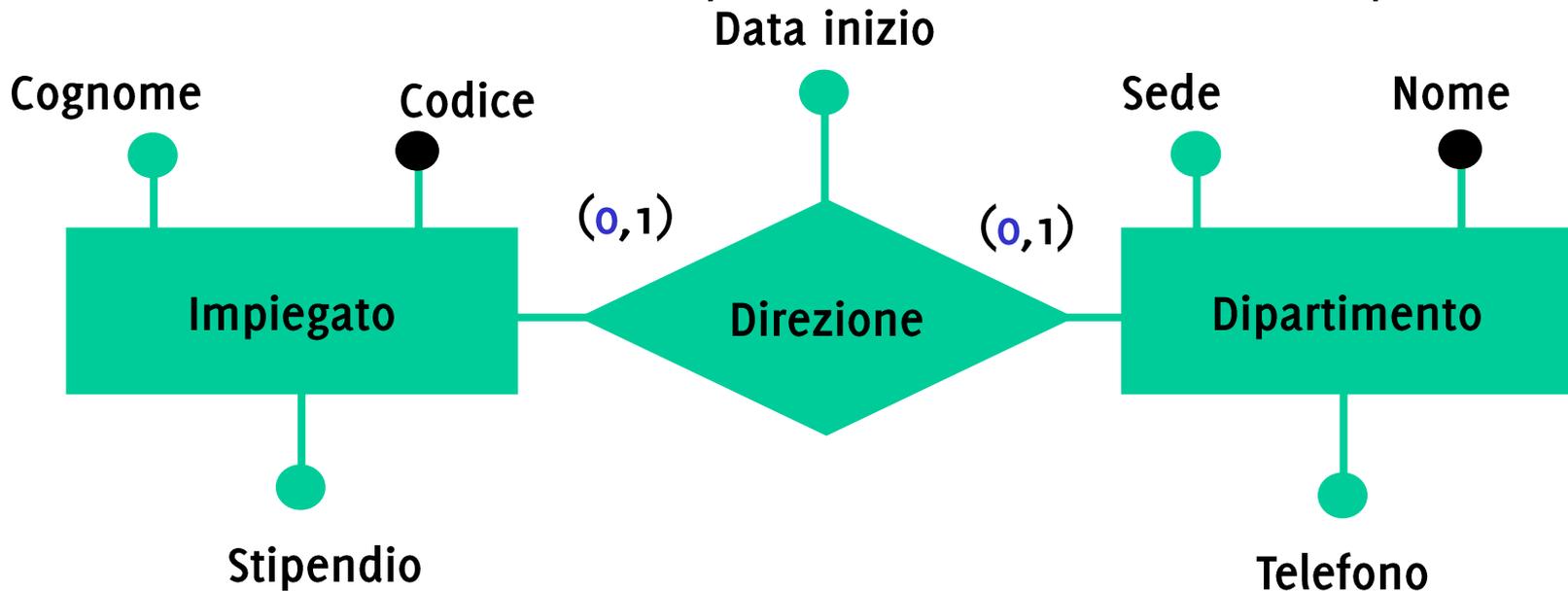
Dipartimento (Nome, Sede, Telefono, Direttore, InizioD)

Nell'altra opzione avremmo invece possibili valori nulli



Associazioni Uno a Uno con opzionalità

Nel caso in cui ci fosse un'opzionalità da entrambe le parti



Allora per evitare di avere valori nulli faccio 3 relazioni:

Impiegato (Codice, Cognome, Stipendio)

Dipartimento (Nome, Sede, Telefono)

Direzione(Direttore, Dipartimento, InizioD)



Esempio TDE 16 Febbraio 2012

Utilizzando un diagramma E-R modella la base di dati necessaria a memorizzare le informazioni per la gestione di un vivaio.

Il gestore del vivaio deve tenere traccia di tutte le piante disponibili. Ogni pianta, identificata da un codice e dalla posizione all'interno del magazzino (fila, colonna) appartiene ad una specie. Le specie sono identificate dal loro nome botanico e da una descrizione nonché da una foto esemplificativa.

Un gruppo di giardinieri si occupa di tali piante. Ogni giardiniere, identificato da nome e cognome, è specializzato su una o più specie ed è obbligato, a seguito di ogni trattamento effettuato sulla pianta, a riportare la data in cui viene effettuato il trattamento, l'ora, e il tipo di trattamento (e.g., innaffiatura, potatura, concimatura).



Partendo dal modello concettuale dell'esercizio precedente ricavarne lo schema relazionale (indicando esplicitamente i vincoli di integrità referenziale).

Sulla base dello schema relazionale ottenuto, si definiscano in algebra relazionale le seguenti interrogazioni:

- elencare i trattamenti sulle piante della fila 5
- elencare i giardinieri che sono specializzati per operare sulle piante della fila 5

Si indichi inoltre in SQL le seguenti interrogazioni

- il numero di trattamenti subiti dalle piante in fila 5 nel mese di maggio 2011
- il numero medio di trattamenti subiti nel 2011 per ogni specie di pianta presente nel vivaio.



Esempio TDE Febbraio 2016 (Recupero)

Si progetti una base dati atta a descrivere gli esami del sangue che vengono svolti in una clinica privata. L'anagrafica dei clienti della clinica viene registrata nel sistema ed i clienti sono identificati da un univoco codice cliente. La clinica svolge un numero ben definito di tipologie di esami del sangue. Ogni esame è identificato dal nome (es. Trigliceridi), e da un intervallo di normalità (valore massimo, valore minimo), da un costo e da un codice identificativo. Quando un cliente fa un esame viene registrata la data in cui questo avviene ed i rispettivi valori. Tra i clienti vi è particolare attenzione per le donne incinta e per queste viene tenuta traccia della data attesa del parto, del numero di figli che già ha. La clinica esegue, solo per le donne incinta, esami particolari (e.g. ecografia morfologica) di cui si registra, viene dispensato dal SSN, se è un'esame a rischio, e se può essere fatto a digiuno. Per ogni esame svolto da una donna incinta si tiene traccia, oltre che del valore, del mese di gestazione. Infine, tutti i referti degli esiti vengono tenuti in una cartella clinica riposta in un magazzino pieno di scaffali numerati e con riferimenti di posizione. La cartella clinica e tutti i referti di un paziente possono stare su un solo scaffale

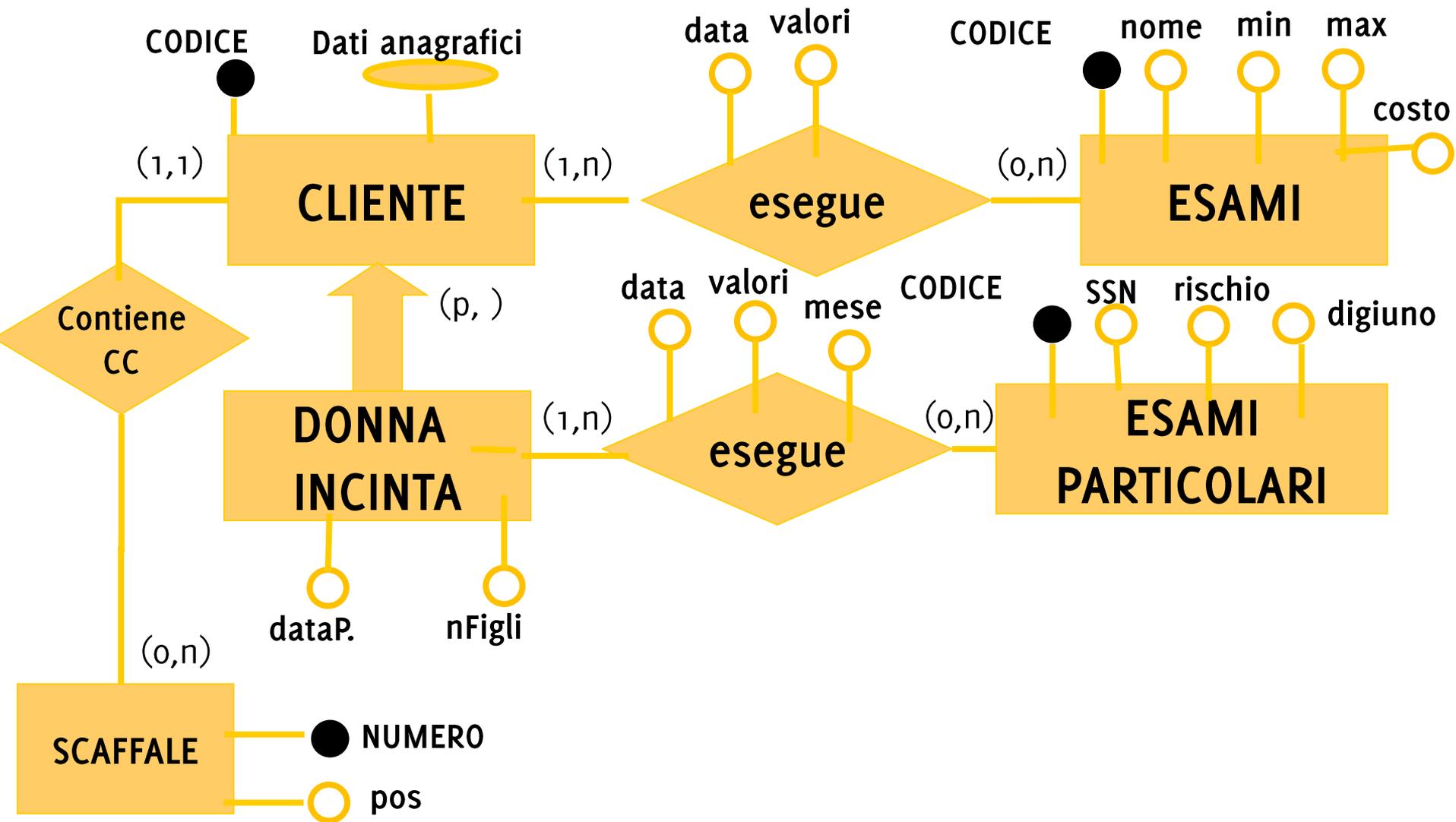


Esempio TDE Febbraio 2016 (Recupero)

- Progettare lo schema **Entità-relazione** per progettare la base dati del sistema sopra descritto (si inseriscano brevi commenti solo se necessari per giustificare alcune scelte progettuali)
- Progettare lo **schema relazionale** della base dati. (si inseriscano brevi commenti solo se necessari per giustificare alcune scelte progettuali)



Schema Entità-relazione





Schema Relazionale

Cliente (Codice, Dati Anagrafici ..., nrScaffale)

DonnaIncinta(Codice, nrFigli, dataPresunta, nrScaffale)

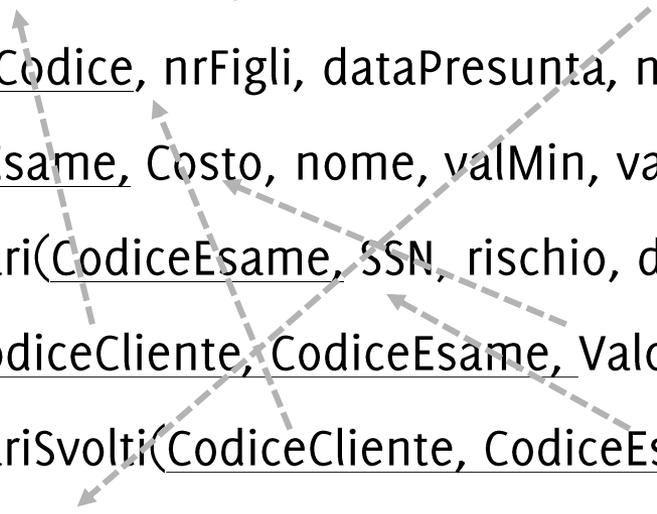
Esami(CodiceEsame, Costo, nome, valMin, valMax)

EsamiParticolari(CodiceEsame, SSN, rischio, digiuno)

EsamiSvolti(CodiceCliente, CodiceEsame, Valori, Data)

EsamiParticolariSvolti(CodiceCliente, CodiceEsame, Mese, Valori, Data)

Scaffale(Numero, Posizione)



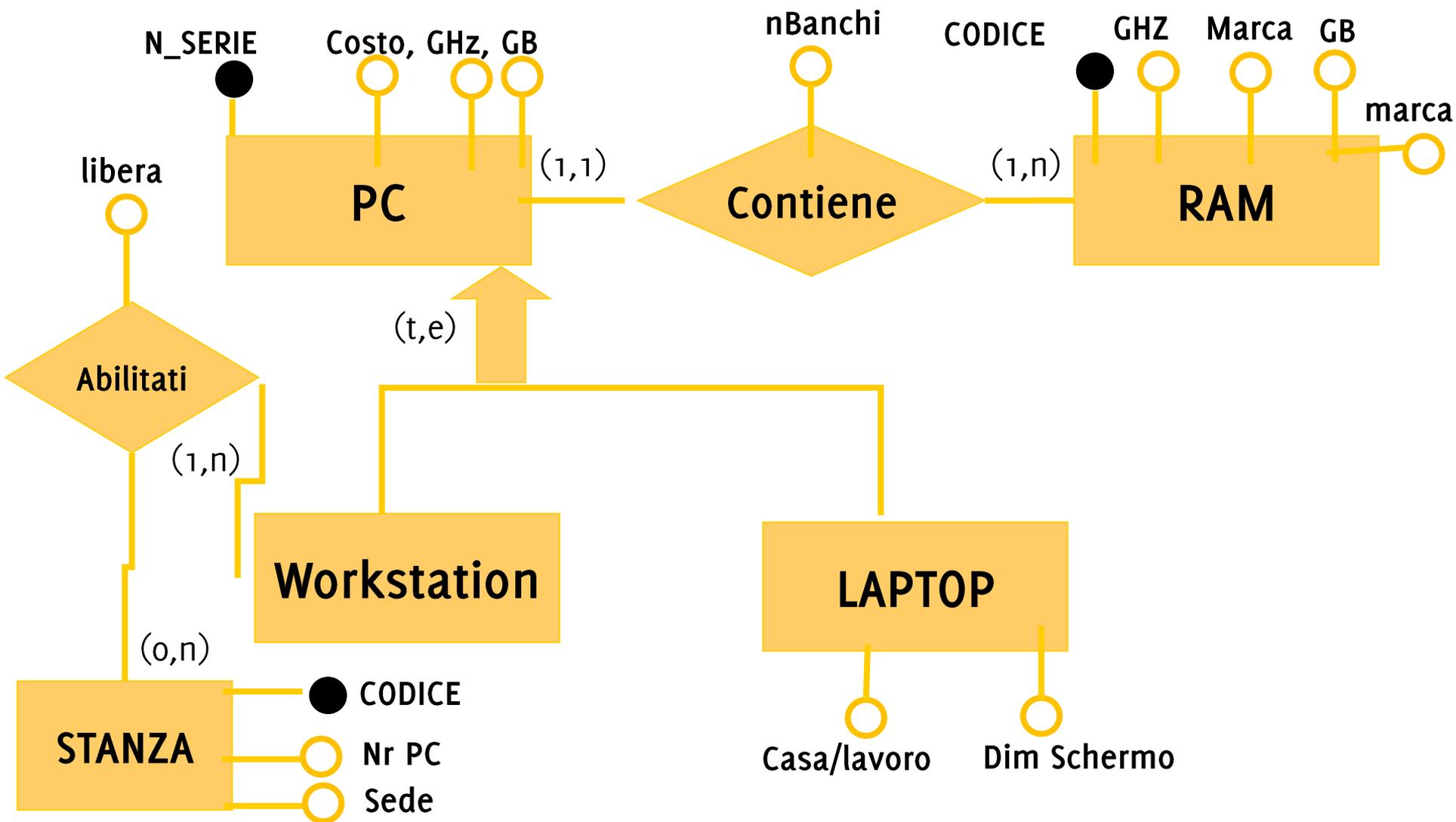


Si progetti una base dati per gestire i PC dei dipendenti del Politecnico di Milano. Al momento dell'acquisto, viene assegnato ad ogni PC un numero di serie e viene registrato il costo, le dimensioni del disco (in GB) e la velocità del processore (in GHz). C'è invece una politica particolare per le informazioni della RAM. La RAM è venduta in banchi e per ogni PC è necessario registrare i seguenti dati relativi alla RAM: il codice identificativo, la velocità (in GHz), la marca, il numero di banchi, la dimensione di un banco (in GB). Si assuma che i banchi di RAM in un PC siano identici in tutte le caratteristiche (codice, dimensioni, marca, velocità) e che quindi il codice identifichi la tipologia di banco (es 4GB Kingston 3GHz) . I PC possono essere sia portatili che workstation. Per i portatili occorre registrare anche le dimensioni dello schermo e se il dipendente è autorizzato a portare il laptop fuori dall'ufficio.

Per ogni workstation occorre registrare le stanze (possono essere più di una) dove è abilitata alla navigazione e se può navigare liberamente o solo all'interno del sito www.polimi.it. Le stanze sono identificate da un codice univoco, dal numero di PC che possono collegarsi in rete simultaneamente e dalla sede in cui si trovano (Como, Milano, Lecco, Piacenza).

1. Progettare lo schema Entità-Relazione della base dati sopra descritta
2. Progettare lo schema relazionale della base dati.

Si inseriscano brevi commenti solo se necessari per giustificare alcune scelte progettuali





Schema Relazionale

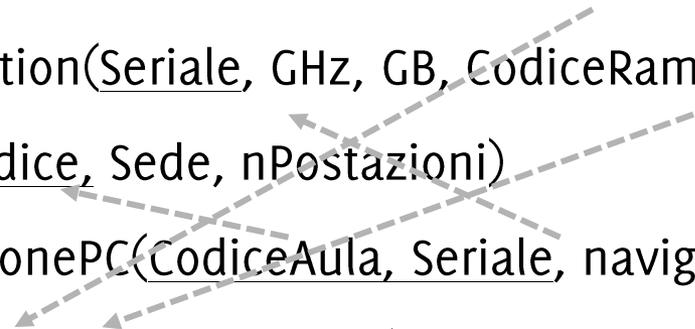
Laptop (Seriale, GHz, GB, CodiceRam)

Workstation(Seriale, GHz, GB, CodiceRam)

Aula(Codice, Sede, nPostazioni)

AbilitazionePC(CodiceAula, Seriale, navigazioneLibera)

RAM(Codice, GHz, GB, tipo)





Febbraio 2016 (Recupero)

Si consideri un sistema informativo in grado di immagazzinare i dati di diverse corse automobilistiche. Si consideri il seguente schema logico

Pilota (Numero, Nome, Cognome, DataNascita, Nazione)

Costruttore (Nome, Nazione, NomePresidente, CognomePresidente)

Auto (Pilota, CostruttoreTelaio, CostruttoreMotore, SponsorPrincipale)

Circuito(Nome, DataGara, Km)

ClassificaCircuito(Posizione, NomeCircuito, Pilota)

Sottolineare le chiavi primarie ed indicare i vincoli di integrità referenziale (con frecce tratteggiate).



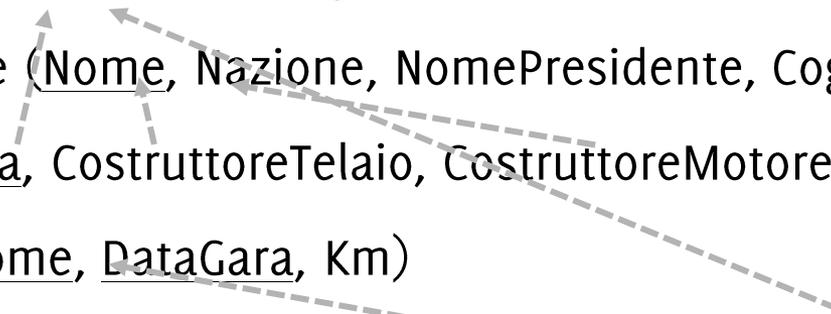
Pilota (Numero, Nome, Cognome, DataNascita, Nazione)

Costruttore (Nome, Nazione, NomePresidente, CognomePresidente)

Auto (Pilota, CostruttoreTelaio, CostruttoreMotore, SponsorPrincipale)

Circuito(Nome, DataGara, Km)

ClassificaCircuito(Posizione, NomeCircuito, Pilota)





Definire le seguenti interrogazione 2 in algebra relazionale e 2 in SQL (a vostra scelta):

- Elencare le auto in cui sia il telaio che il motore sono forniti dallo stesso costruttore
- Fornire nome, cognome dei primi tre piloti classificati del gran premio di Monaco,
- Fornire i nomi dei costruttori le cui auto non hanno mai vinto una gara (per auto si intende auto che adotta un motore o un telaio di questo costruttore)
- Fornire i nomi degli sponsor principali che hanno percorso almeno 3000 Km nel 2015



Soluzione Query

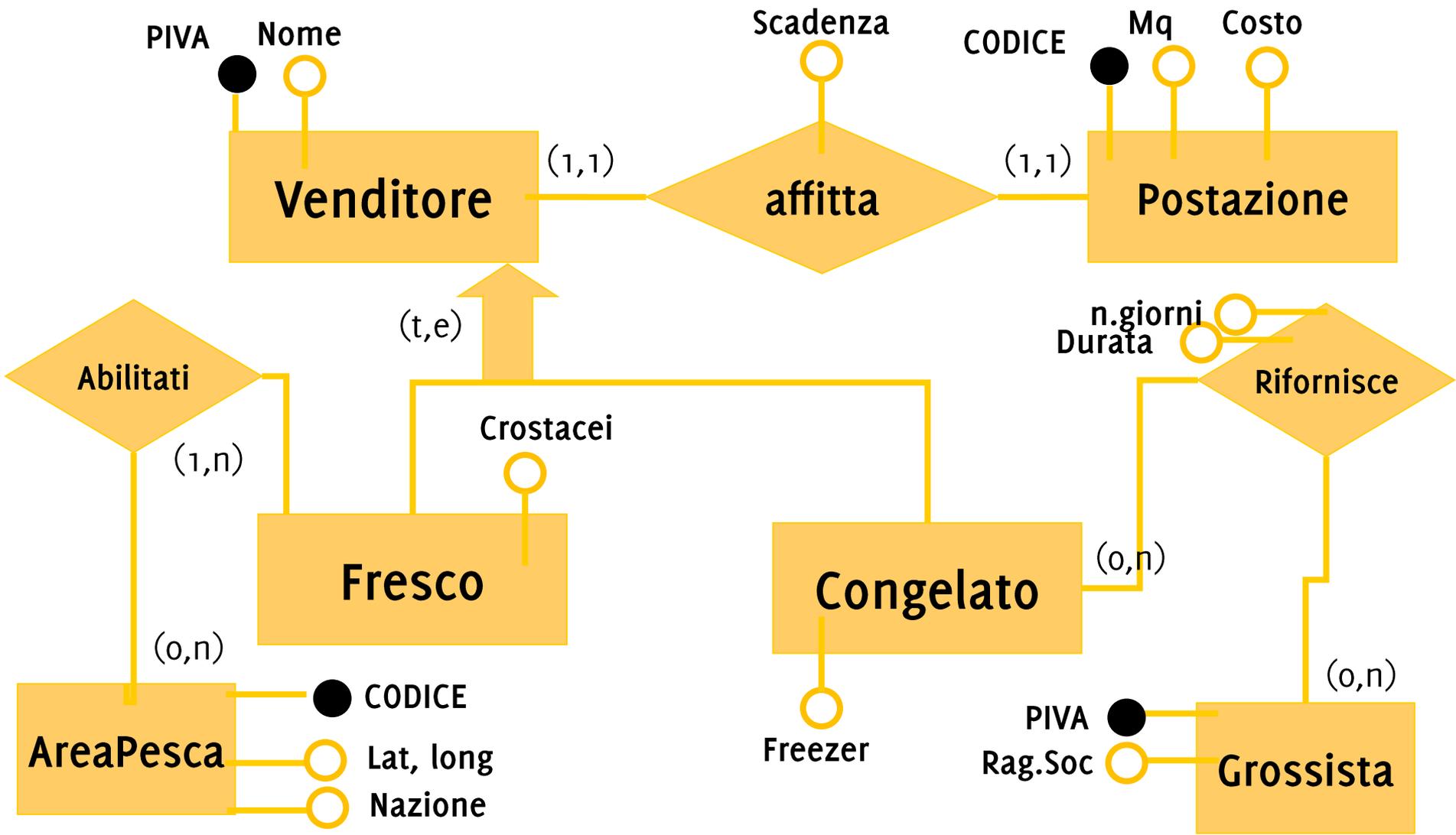
- PROJ_Pilota (SEL_CostruttoreTelaio=CostruttoreMotore (Auto))
- SELECT Nome, Cognome
FROM ClassificaCircuito, Pilota
WHERE Pilota=Numero and Posizione>=1 and Posizione<=3 and
NomeCircuito='Monaco'
- PROJ_nome (Costruttore) \ PROJ_nome((Costruttore JOIN_{Costruttore.Nome =
auto.CostruttoreTelaio} Auto) U (Costruttore JOIN_{Costruttore.Nome =
auto.CostruttoreMotore} Auto))
- SELECT Auto.SponsorPrincipale,
FROM ClassificaCircuito, Circuito, Auto
WHERE ClassificaCircuito.Pilota = Auto.Pilota AND Circuito.Nome =
ClassificaCircuito.NomeCircuito AND YEAR(Circuito.DataGara) = 2015 GROUP BY
Auto.SponsorPrincipale
HAVING SUM(Circuito.Km) > 3000



Si progetti la base dati per gestire il mercato ittico di Lecco. Nel mercato esistono diversi venditori, ciascuno identificato dalla P.ta IVA e dal suo nome. Ogni venditore occupa una precisa postazione con la sua bancarella e ha un'autorizzazione per la vendita del pesce fresco o, in alternativa, di pesce congelato. Per tutti i venditori abilitati alla vendita del pesce fresco occorre registrare se può vendere anche crostacei e le aree di pesca dove da dove proviene il pesce in vendita. Per i venditori di pesce congelato occorre registrare i contratti con i grossisti da cui possono rifornirsi e la temperatura minima raggiunta dai freezer. Ogni contratto è caratterizzato dalla durata e da un numero di giorni massimo cui il venditore può tenere in carico la merce.

Per ogni grossista è necessario registrare la P.ta IVA e la ragione sociale, il nome e l'anno di inizio dell'attività. Le aree di pesca sono identificate da un codice, dalle coordinate (lat, long) del suo centro e dalla nazione cui fa' riferimento (eventualmente INTERNAZIONALE). Infine, per le postazioni di vendita, occorre registrare un codice, la scadenza della concessione, i metri quadri ed il costo dell'affitto mensile.

- Progettare lo schema **Entità-relazione** per la base dati del sistema sopra descritto (si inseriscano brevi commenti solo se necessari per giustificare alcune scelte progettuali)
- Progettare lo **schema relazionale** della base dati. (si inseriscano brevi commenti solo se necessari per giustificare alcune scelte progettuali)





VenditoreFresco(PIVA, Crostacei, CodicePostazione, ScadenzaAffitto)

VenditoreCongelato(PIVA, TempFrezer, CodicePostazione, ScadenzaAffitto)

Postazione(Codice, Mq, CostoMensile)

Grossista(PIVA, Rag.Sociale)

RifornimentiGrossista(PIVA_Venditore, PIVA_Grossista, nGiorni, DurataContratto)

AreaPesca(Codice, Lat, Long, Nazione)

AbilitazioneAree(CodiceArea, PIVA_Venditore)



VenditoreFresco(PIVA, Crostacei, CodicePostazione, ScadenzaAffitto)

VenditoreCongelato(PIVA, TempFrezer, CodicePostazione, ScadenzaAffitto)

Postazione(Codice, Mq, CostoMensile)

Grossista(PIVA, Rag.Sociale)

RifornimentiGrossista(PIVA_Venditore, PIVA_Grossista, nGiorni, DurataContratto)

AreaPesca(Codice, Lat, Long, Nazione)

AbilitazioneAree(CodiceArea, PIVA_Venditore)