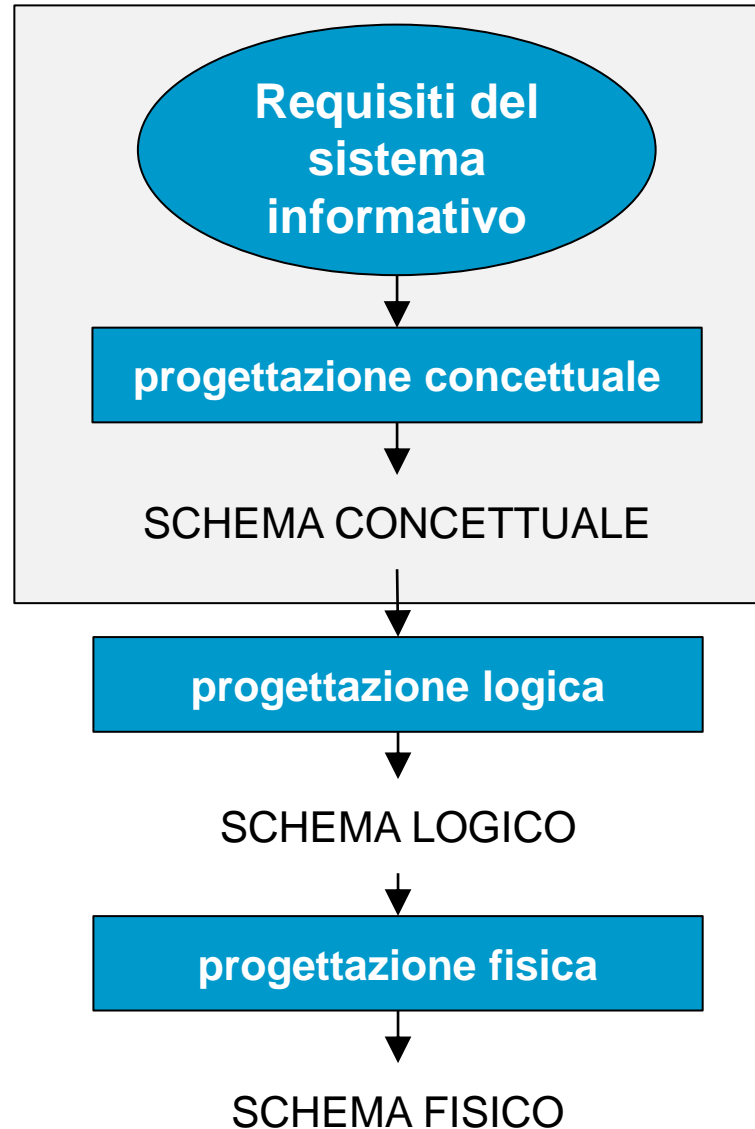




Progettazione concettuale

Il primo passo..



Obiettivi

- La fase di analisi dei requisiti e progettazione concettuale (“Analisi dei dati”) comprende attività di:
 - Acquisizione dei requisiti
 - Analisi dei requisiti
 - Costruzione dello schema concettuale (modello E/R)
 - Costruzione del glossario dei dati

Raccolta dei requisiti

- ❑ I requisiti devono innanzitutto essere acquisiti.
- ❑ Le fonti possono essere molto diversificate tra loro:
 - Utenti, attraverso:
 - Interviste
 - Documentazione apposita
 - Documentazione esistente
 - Normative (leggi, regolamenti di settore)
 - Regolamenti interni, procedure aziendali
 - Realizzazioni preesistenti
 - Modulistica
- ❑ La raccolta dei requisiti è un'attività molto difficile e non standardizzabile; in genere procede di pari passo con la fase di analisi (la prima analisi stimola nuove domande..)

Interagire con gli utenti

- ❑ È un'attività da considerare con molta attenzione, in quanto:
 - utenti diversi possono fornire informazioni diverse con riferimento allo stesso tema;
 - utenti a livello più alto hanno spesso una visione più ampia ma meno dettagliata.
- ❑ In generale, risulta utile:
 - effettuare spesso verifiche di comprensione e coerenza;
 - verificare anche per mezzo di esempi (generalisti e relativi a casi limite);
 - richiedere definizioni e classificazioni;
 - far evidenziare gli aspetti essenziali rispetto a quelli marginali

Requisiti: documentazione descrittiva

□ Regole generali:

- scegliere il corretto livello di astrazione;
- standardizzare la struttura delle frasi;
- suddividere le frasi articolate;
- separare le frasi sui dati da quelle sulle funzioni (operazioni).

□ Per meglio evidenziare i concetti che sono espressi nei requisiti, è opportuno:

- costruire un glossario dei termini;
- individuare omonimi e sinonimi e unificare i termini;
- rendere esplicito il riferimento fra termini;
- riorganizzare le frasi per concetti.

Requisiti: un esempio

- ❑ Base di dati bibliografica
- ❑ Si vogliono organizzare i dati di interesse per automatizzare la gestione dei riferimenti bibliografici.

BD bibliografica (1)

- ❑ Si vogliono organizzare i dati di interesse per automatizzare la gestione dei riferimenti bibliografici, con tutte le informazioni da riportarsi in una bibliografia.
- ❑ Per ogni **pubblicazione** deve esistere un **codice identificante** costituito da sette caratteri, indicanti le iniziali degli autori, l'anno di pubblicazione e un carattere aggiuntivo per la discriminazione delle collisioni

BD bibliografica (2)

- ❑ Si vogliono organizzare i dati di interesse per automatizzare la gestione dei riferimenti bibliografici, con tutte le informazioni da riportarsi in una bibliografia.
- ❑ Le **pubblicazioni** sono di due tipi, **monografie** (per le quali interessano editore, data e luogo di pubblicazione) e **articoli su rivista** (con nome della rivista, volume, numero, pagine e anno di pubblicazione); per entrambi i tipi si devono ovviamente riportare i nomi degli autori.
- ❑ Per ogni pubblicazione deve esistere un codice identificante

Un altro esempio più articolato

- ❑ **Società di formazione**
- ❑ Si vuole realizzare una base di dati per una società che eroga corsi di formazione; si devono rappresentare le informazioni tipiche sui **partecipanti** ai corsi e sui **docenti**.
- ❑ Per gli **studenti** (circa 5000), identificati da un codice, si vuole memorizzare il codice fiscale, il cognome, il nome, il sesso, il luogo di nascita, la data di nascita, il nome, l'indirizzo e il numero di telefono dei loro attuali datori di lavoro, i **posti** dove hanno lavorato in precedenza insieme al periodo, i **corsi** che hanno frequentato (i corsi sono in tutto circa 200) e il giudizio finale.
- ❑ Si desidera inoltre rappresentare, per ciascun **insegnamento** che gli studenti stanno attualmente frequentando, i giorni e le ore di lezione nonché i luoghi dove si svolgono le lezioni.

Società di formazione (2)

- ❑ I **corsi** hanno un codice, un titolo e possono avere varie edizioni caratterizzate ciascuna da una data di inizio, una data di fine e un numero di partecipanti.
- ❑ Se gli **studenti** sono **liberi professionisti** si vuole conoscere l'area di interesse e, se lo possiedono, il titolo professionale. Per quelli che lavorano alle **dipendenze** di altri, vogliamo conoscere invece il livello e la posizione ricoperta.
- ❑ Per gli **insegnanti** (circa 300), si vuole rappresentare il codice fiscale, il cognome, il nome, il sesso, la data di nascita, il posto dove sono nati, il nome del corso che insegnano, quelli che hanno insegnato nel passato e quelli che possono insegnare. Si vogliono memorizzare inoltre tutti i loro recapiti telefonici. I docenti possono essere dipendenti interni della società o collaboratori esterni.

Glossario dei termini, omonimi e sinonimi

- ❑ Raramente i requisiti espressi in linguaggio naturale sono privi di ambiguità. È infatti frequente il caso di:
 - **omonimi**: lo stesso termine viene usato per descrivere concetti differenti (es: libro e copia di libro, posto: di lavoro e geografico);
 - **sinonimi**: termini diversi vengono usati per descrivere lo stesso concetto (es: studente e partecipante).

- ❑ Un modo conveniente per rappresentare sinteticamente i concetti più rilevanti emersi dall'analisi è il **glossario dei termini**, il cui scopo è fornire per ogni concetto rilevante:
 - una breve descrizione del concetto;
 - eventuali sinonimi;
 - relazioni con altri concetti del glossario stesso.

Glossario dei termini: esempio

Termine	Descrizione	Sinonimi	Collegamenti
Partecipante	Persona che partecipa ai corsi. Può essere un dipendente o un libero professionista.	Studente	Corso, Datore
Docente	Docente dei corsi. Può essere un collaboratore esterno.	Insegnante	Corso
Corso	Corso organizzato dalla società. Può avere più edizioni.	Insegnamento	Docente, Partecipante
Datore	Datore di lavoro attuale o passato di un partecipante ai corsi.	Posto di lavoro	Partecipante

Ristrutturazione dei requisiti

- Oltre a costruire il glossario, per semplificare le analisi successive, è utile riformulare i requisiti:
 - eliminare le omonimie;
 - usare un termine univoco per ogni concetto;
 - riorganizzare le frasi raggruppandole in base al concetto cui si riferiscono.
 - Nell'esempio:
 - frasi di carattere generale;
 - frasi riferite ai partecipanti;
 - frasi riferite ai docenti;
 - frasi riferite ai corsi;
 - frasi riferite alle società.

Esempio: frasi relative ai partecipanti

- ❑ Per i **partecipanti** (circa 5000), identificati da un codice, rappresentiamo il codice fiscale, il cognome, il nome, la data di nascita, il sesso, la città di nascita, i nomi dei loro attuali datori di lavoro e di quelli precedenti (insieme alle date di inizio e fine rapporto), le edizioni dei corsi che stanno attualmente frequentando e quelli che hanno frequentato nel passato, con la relativa votazione finale in decimi. Per i partecipanti che sono **liberi professionisti**, rappresentiamo l'area di interesse e, se lo possiedono, il titolo professionale. Per i partecipanti che sono **dipendenti**, rappresentiamo invece il livello e la posizione ricoperta.



Frase relative ai corsi

- ❑ Per i **corsi** (circa 200), rappresentiamo il titolo e il codice, le varie edizioni con date di inizio e fine e, per ogni edizione, rappresentiamo il numero di partecipanti e i giorni e le ore di lezione nonché i luoghi dove si svolgono le lezioni.

Fraasi relative ai docenti

- Per i **docenti** (circa 300), rappresentiamo il codice fiscale, il nome, il cognome, il sesso, la data di nascita, la città di nascita, uno o più numeri di telefono, il titolo del corso che insegnano, di quelli che hanno insegnato nel passato e di quelli che possono insegnare. I docenti possono essere dipendenti interni della società di formazione o collaboratori esterni.

Dai concetti allo schema E/R

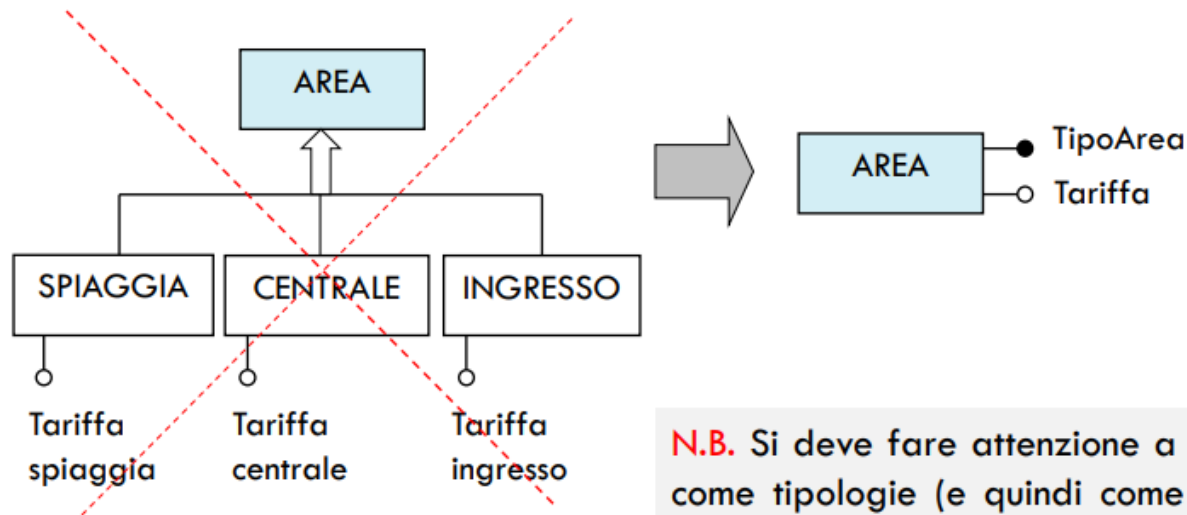
- ❑ Va sempre ricordato che un concetto non è di per sé un'entità, un'associazione, un attributo, o altro

DIPENDE DAL CONTESTO!

- ❑ Come regole guida, un concetto sarà rappresentato come
 - **Entità**: se ha proprietà significative e descrive oggetti con esistenza autonoma
 - **Attributo**: se è semplice e non ha proprietà.
 - **Associazione**: se correla due o più concetti - e non si tratta di entità.
 - **Generalizzazione / Specializzazione**: se è caso più generale/particolare di un altro.

Considerazioni di progettazione

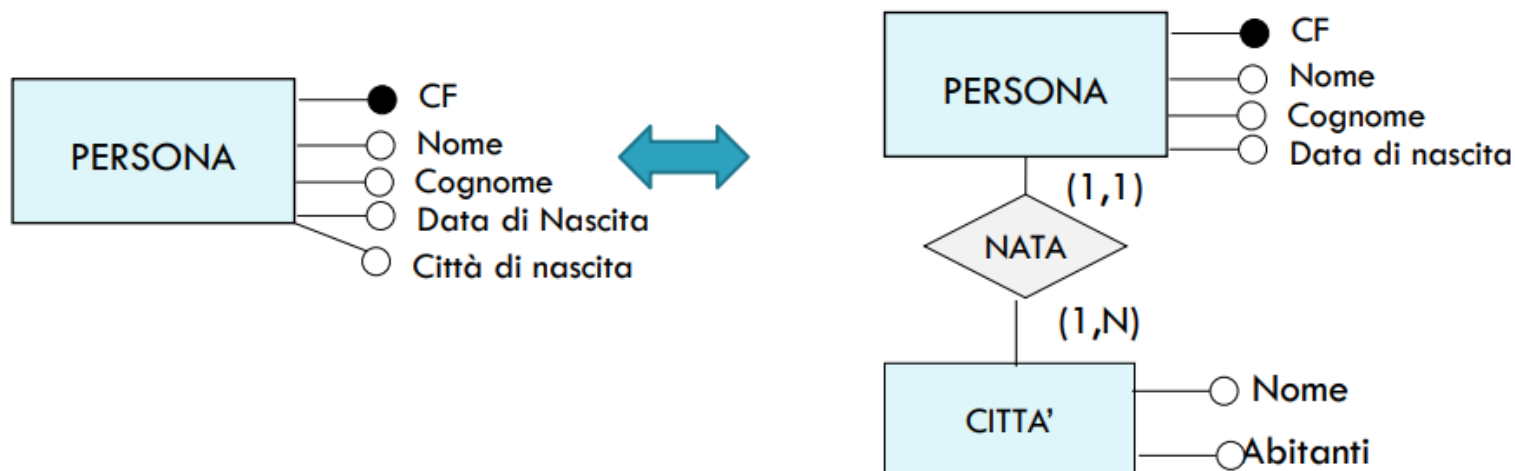
- Attenzione a non confondere entità con istanze di entità tentando di modellare attraverso gerarchie la conoscenza di specifiche istanze.
- **Esempio: ... un campeggio è diviso in tre aree (spiaggia, centrale, ingresso) ognuna delle quali è caratterizzata da una certa tariffa ...**



N.B. Si deve fare attenzione a non interpretare come tipologie (e quindi come specializzazioni di un'entità) quelle che sono solo istanze di un'entità.

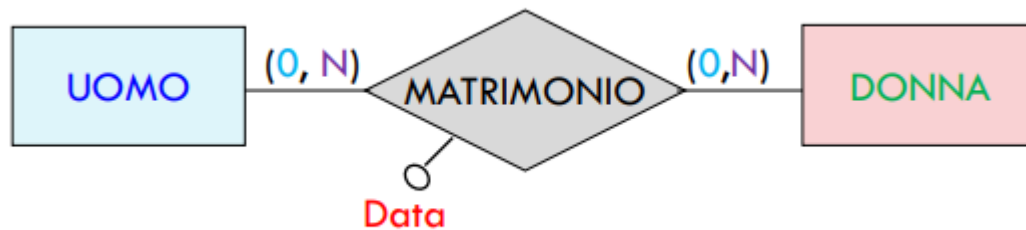
Considerazioni di progettazione

- Un concetto autonomo è un'entità. Tuttavia, se il concetto è molto semplice ed è costituito da un singolo attributo, è spesso opportuno trasformarlo in un attributo di entità.
- **Esempio: ... un campeggio è diviso in tre aree (spiaggia, centrale, ingresso) ognuna delle quali è caratterizzata da una certa tariffa ...**

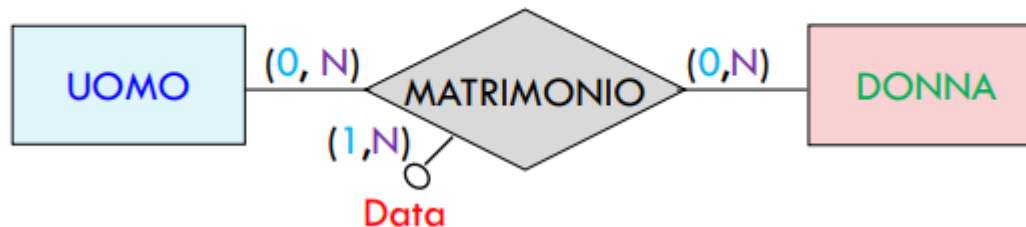


Nota. Se il concetto di "città" è solo una stringa nei requisiti del progetto, è conveniente rappresentarlo come un semplice attributo. In un database geografico, invece, la città è sicuramente un'entità perché è composta da più proprietà.

Considerazioni di progettazione



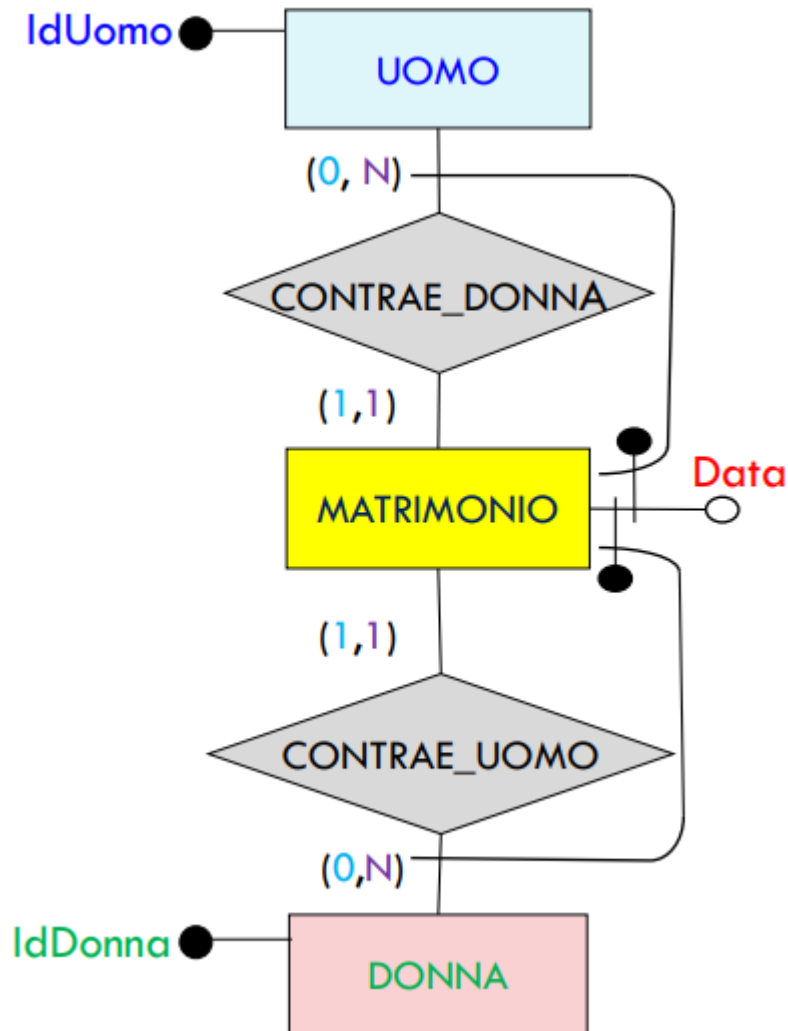
- Non possono essere presenti nell'associazione MATRIMONIO due istanze $(U_i, D_j, d1)$ e $(U_i, D_j, d2)$ dunque uno stesso uomo non può risposare la stessa donna (una stessa donna non può risposare lo stesso uomo)



- Un uomo può risposare la stessa donna più volte.

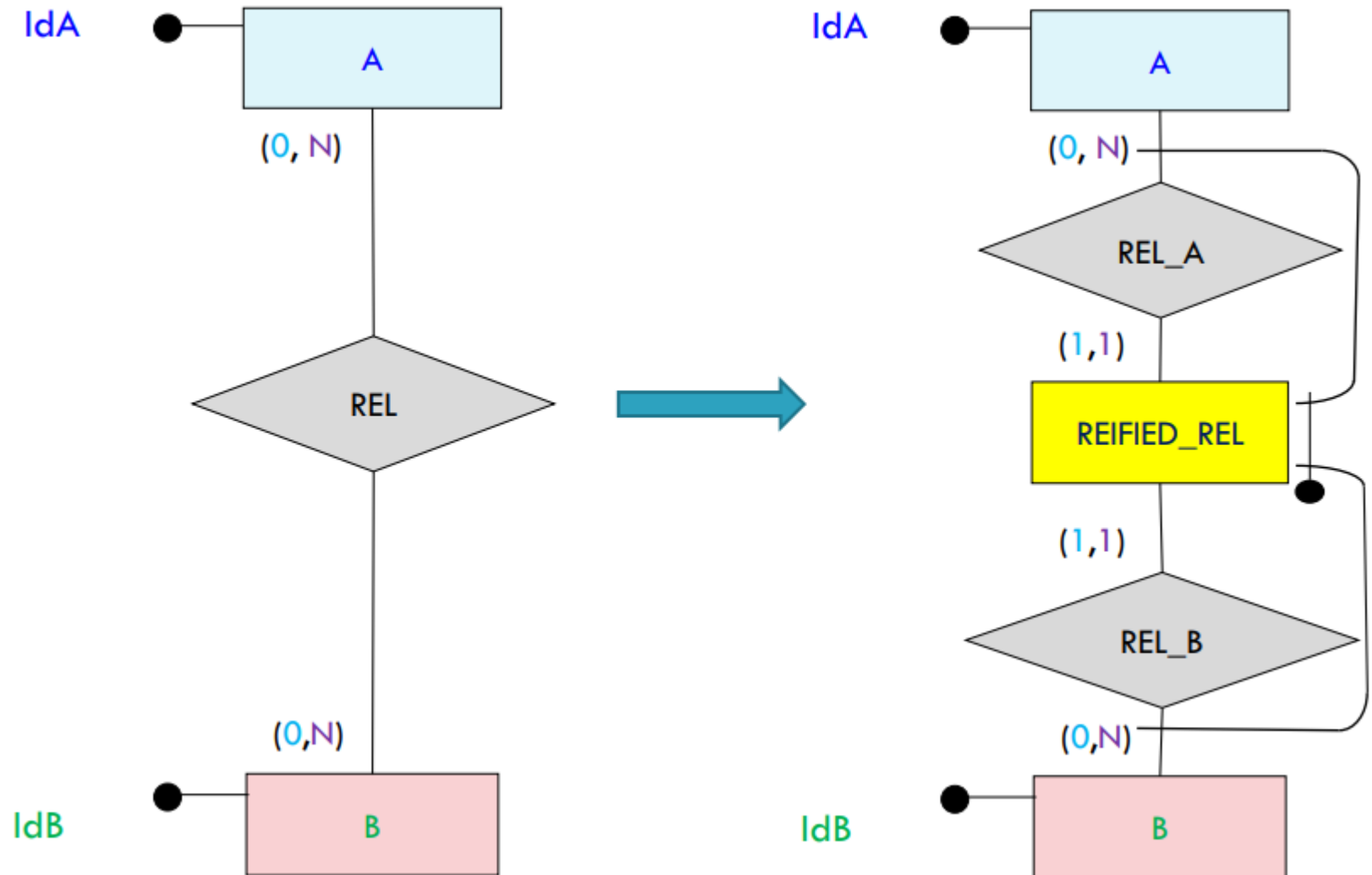
N.B. Entrambi gli schemi consentono che uno stesso uomo (o una stessa donna) possano sposarsi più volte nella stessa data!

Considerazioni di progettazione

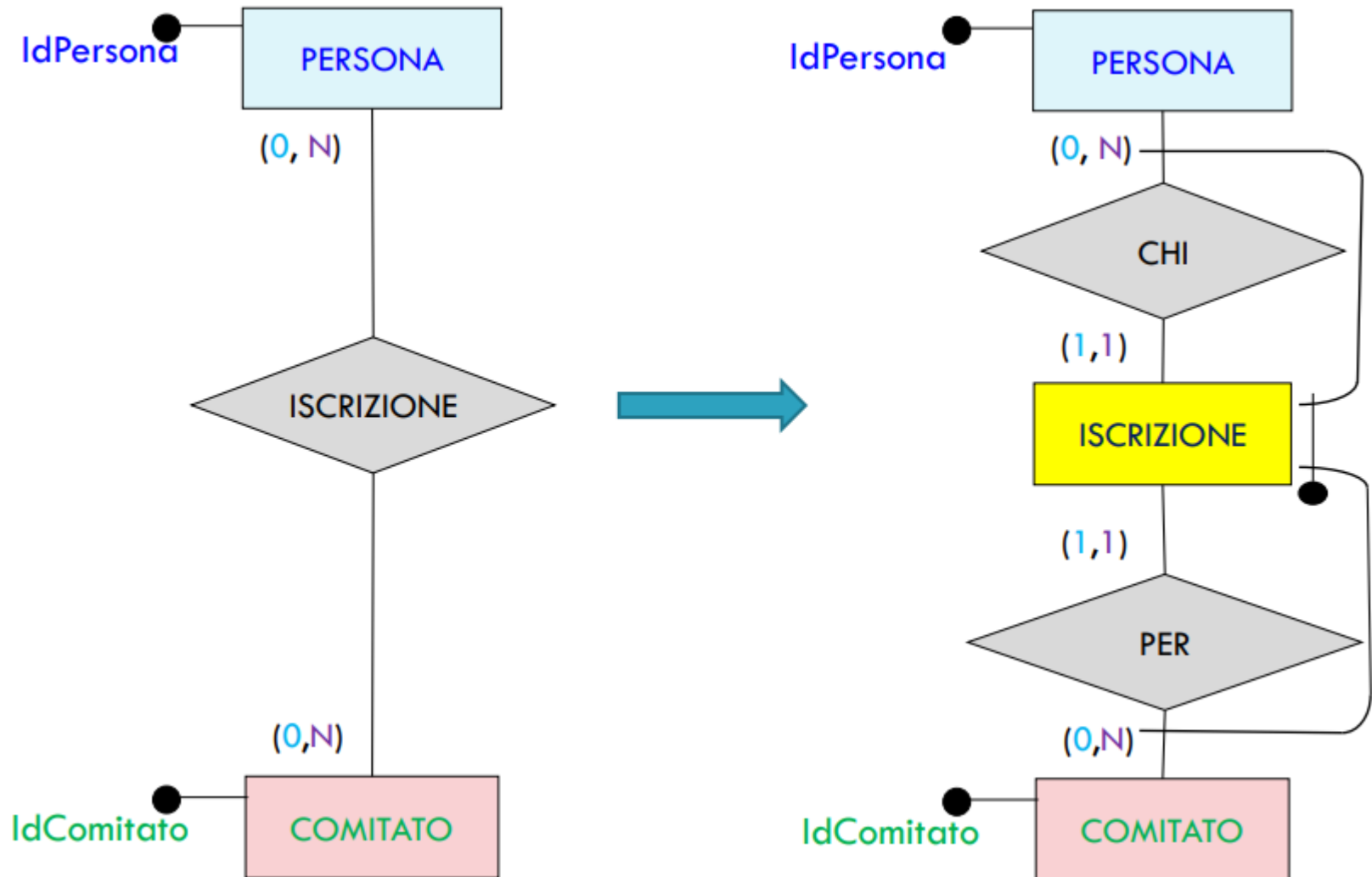


- Soluzione? **Reificazione**: trasformazione di un'associazione in un'entità
- Uno stesso uomo e una stessa donna possono anche risposarsi tra loro.
- Un uomo (donna) non può contrarre più di un matrimonio nello stesso giorno.
- N.B. Entrambi i vincoli devono essere rispettati, pertanto a livello di schema relazionale si sceglierà come chiave primaria di **MATRIMONIO** uno dei due identificatori {**IdDonna**, **Data**} o {**IdUomo**, **Data**}, e si imporrà il rispetto dell'altro vincolo.

Reificazione

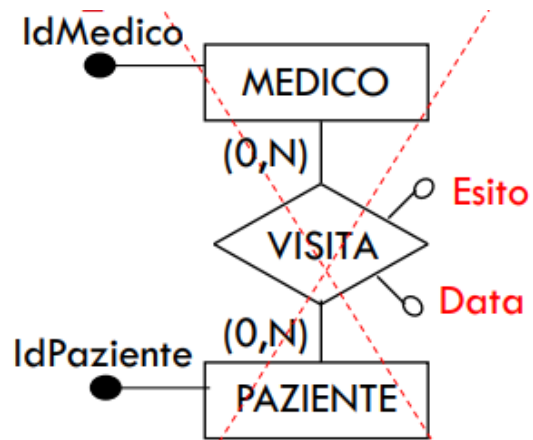


Reificazione: esempio

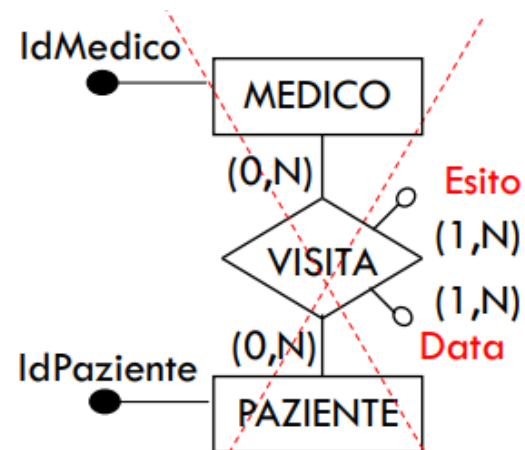


Reificazione: esempio

Dominio: medico-paziente, compresa la registrazione di tutte le date e gli esiti della visita



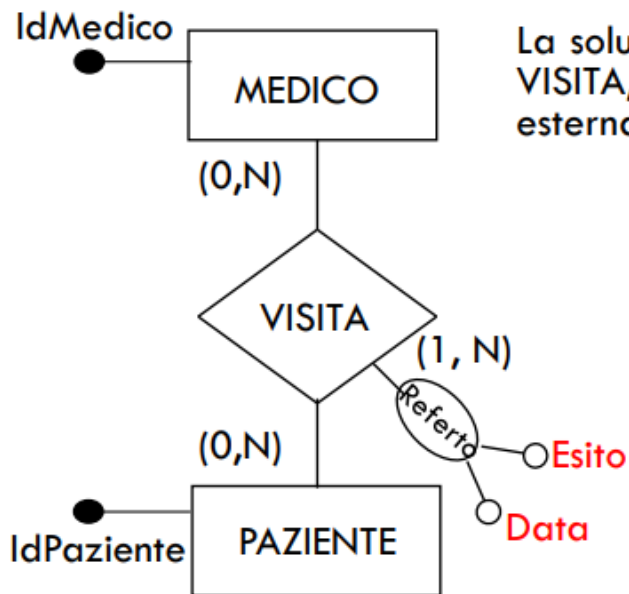
Un medico non può visitare un paziente più di una volta



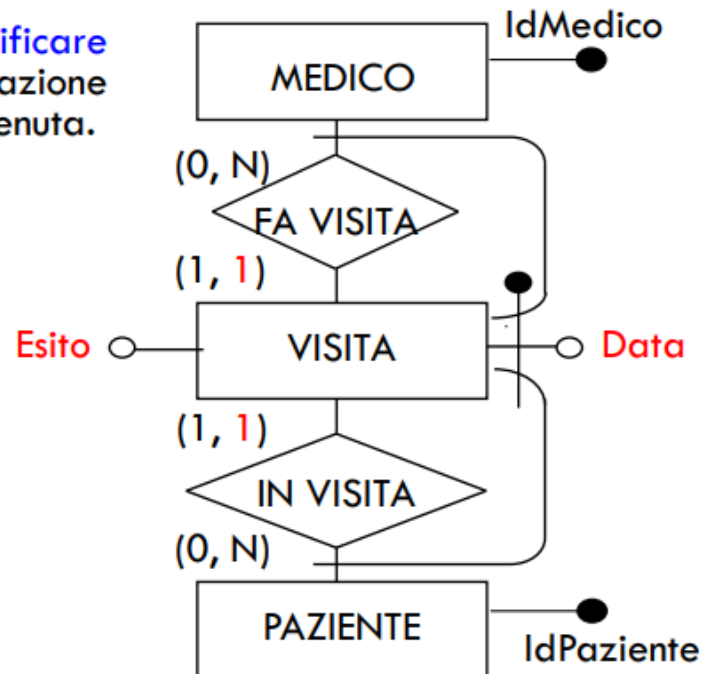
Non è possibile associare l'esito della visita alla data in cui è stata effettuata

Reificazione: esempio

Dominio: medico-paziente, compresa la registrazione di tutte le date e gli esiti della visita

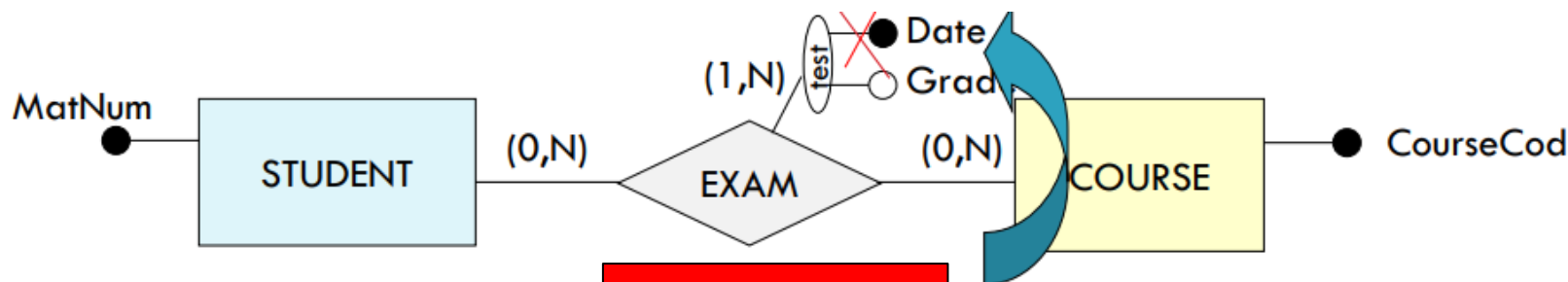


Questa soluzione accoppia data e esito... ma consente più visite a un paziente da parte dello stesso medico nello stesso giorno

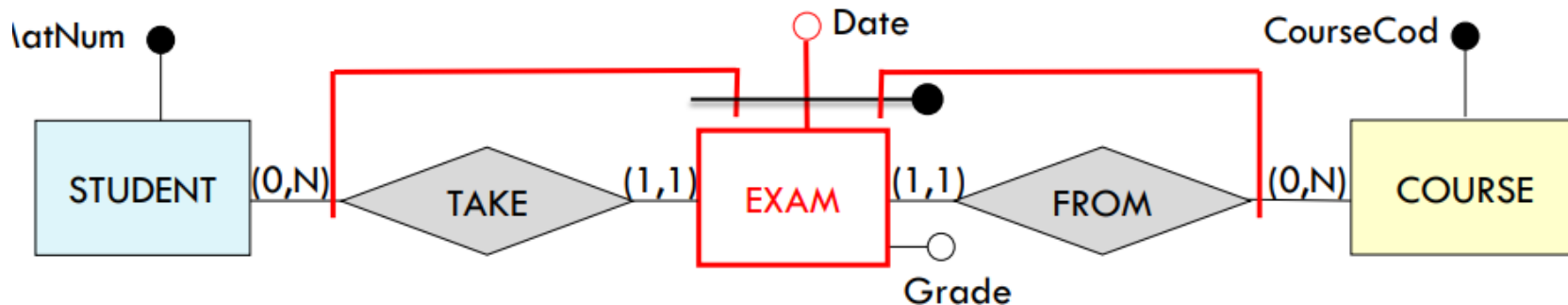


Reificazione: riassunto

Se una relazione ha un attributo composto e ripetuto, e uno dei componenti attributi è necessario per identificare la relazione, dovresti reificare la relazione e creare un identificatore misto.

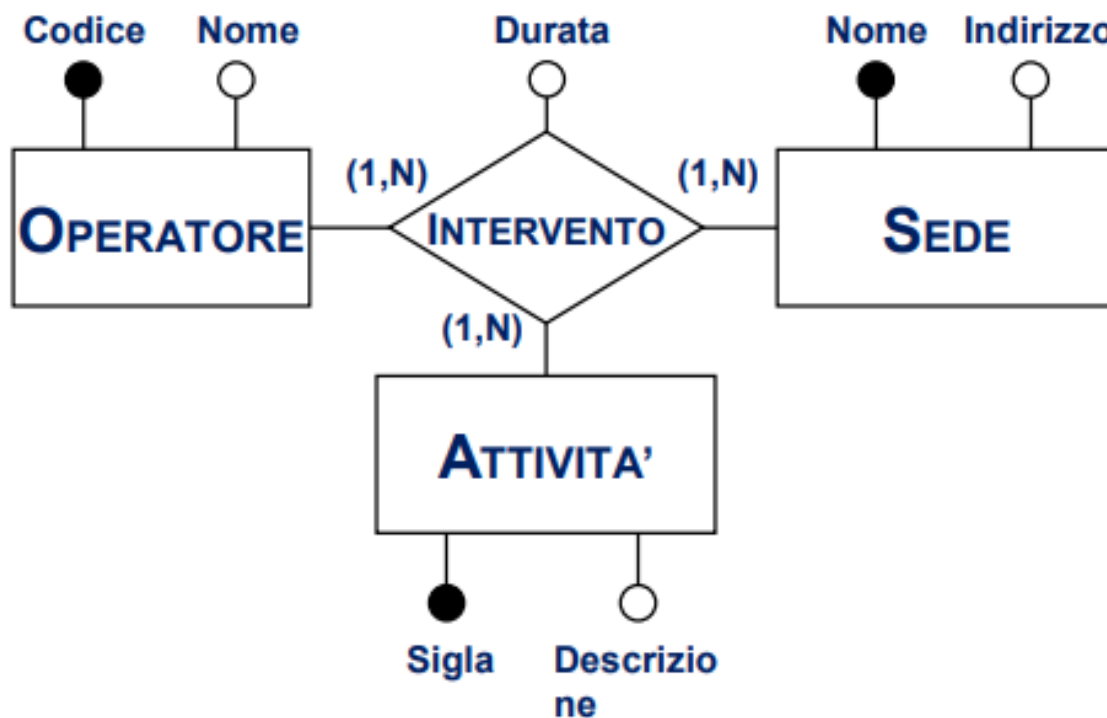


ERRATO

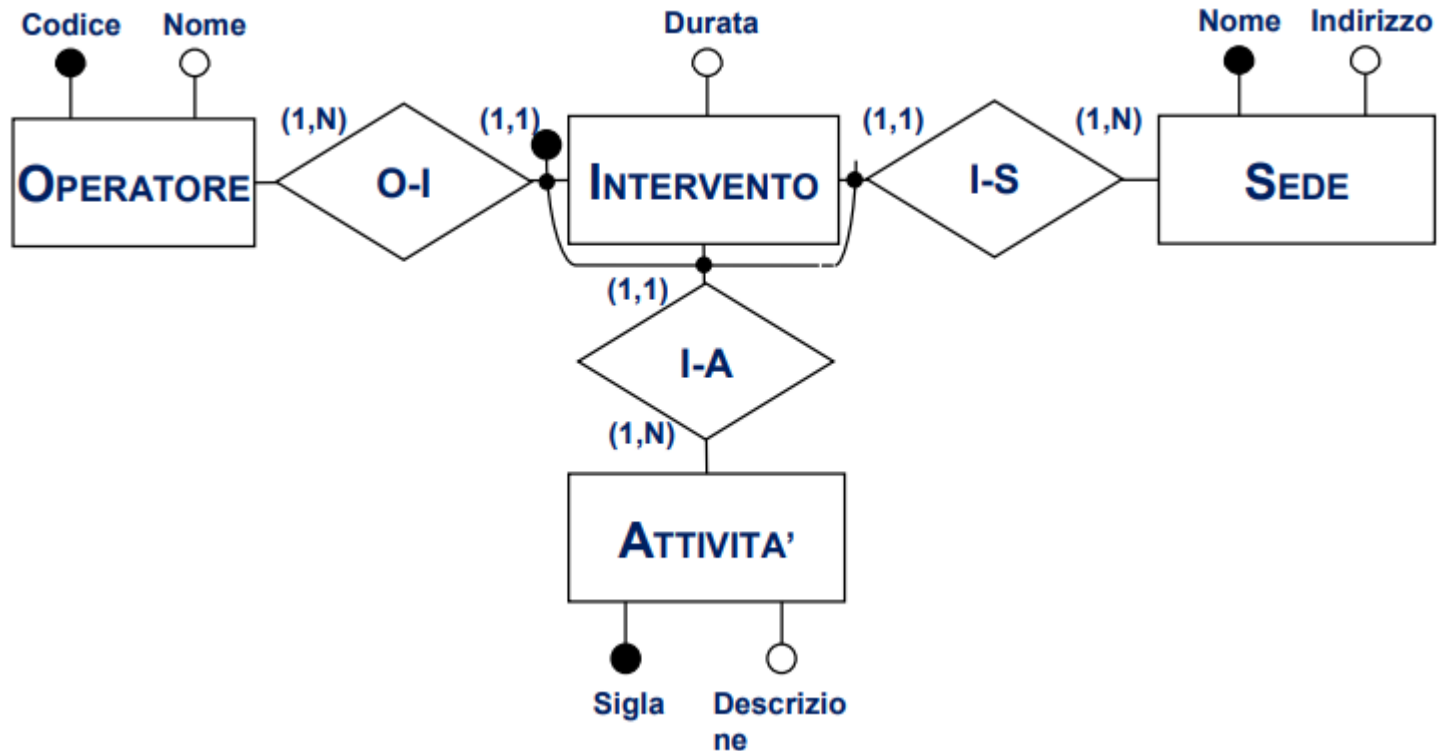


Reificazione di un'associazione ternaria

- Lo schema riportato modella correttamente il caso in cui un operatore può effettuare interventi che consistono in diverse attività in sedi diverse.



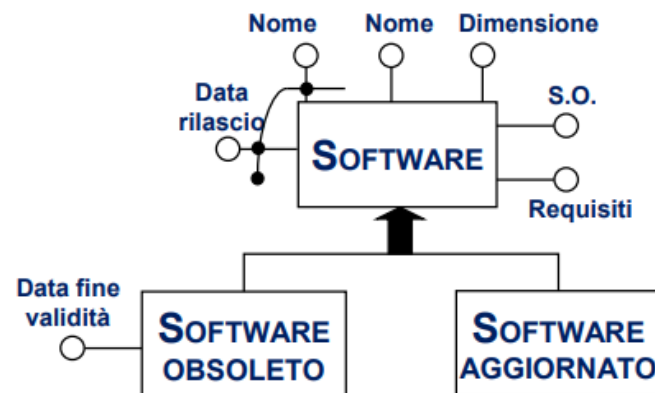
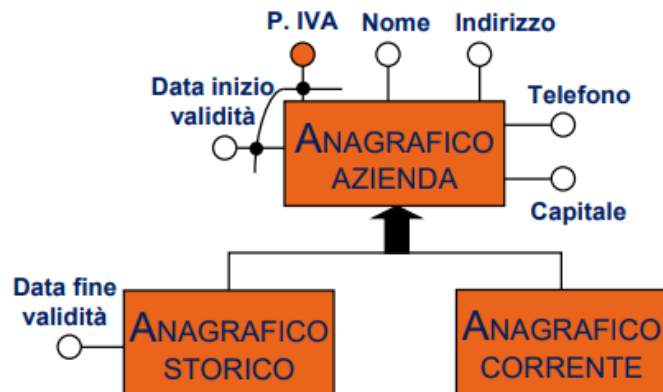
Reificazione di un'associazione ternaria



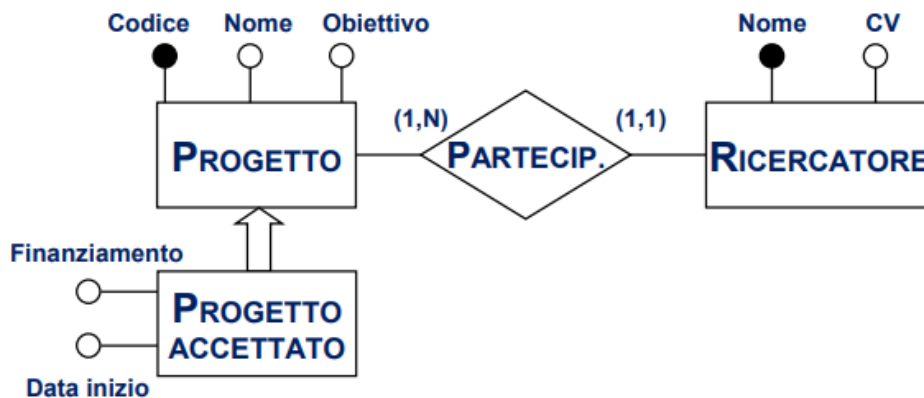
- Modificando l'identificatore di intervento si possono modellare specifiche diverse:
 - Sede + operatore: in ogni sede ogni operatore svolge la stessa attività
 - Sede + attività: in ogni sede ogni attività è svolta da un solo operatore
 - Operatore + attività: ogni operatore svolge ogni attività in una sola sede

Modellazione delle gerarchie

➤ Storicizzazione

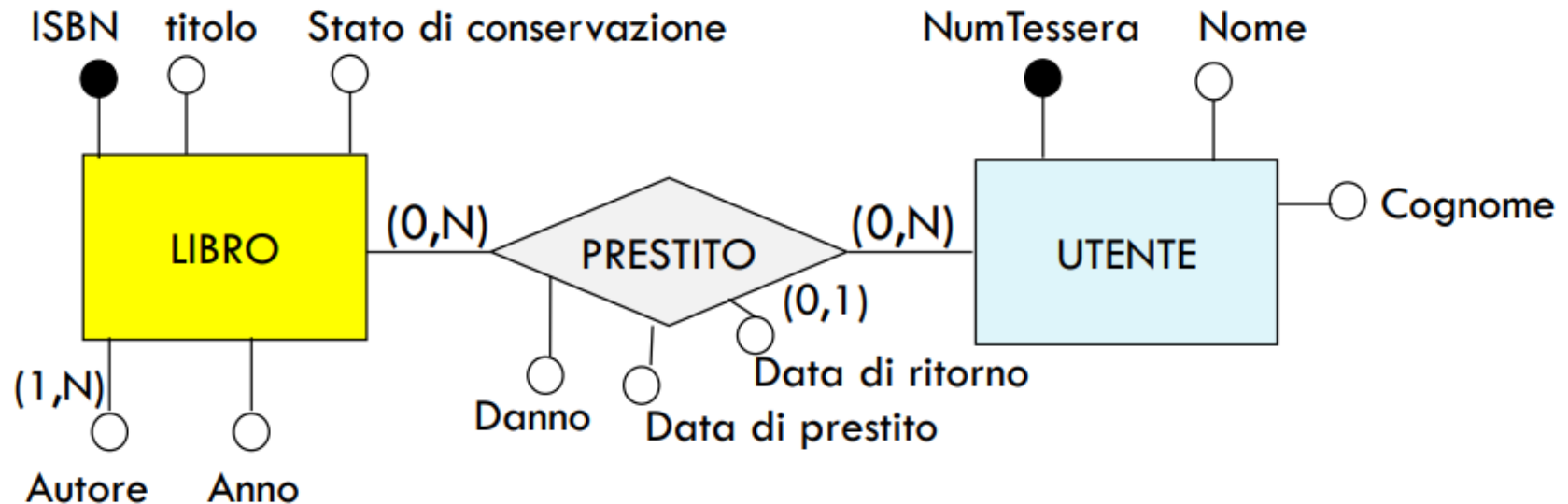


➤ Evoluzione di concetto



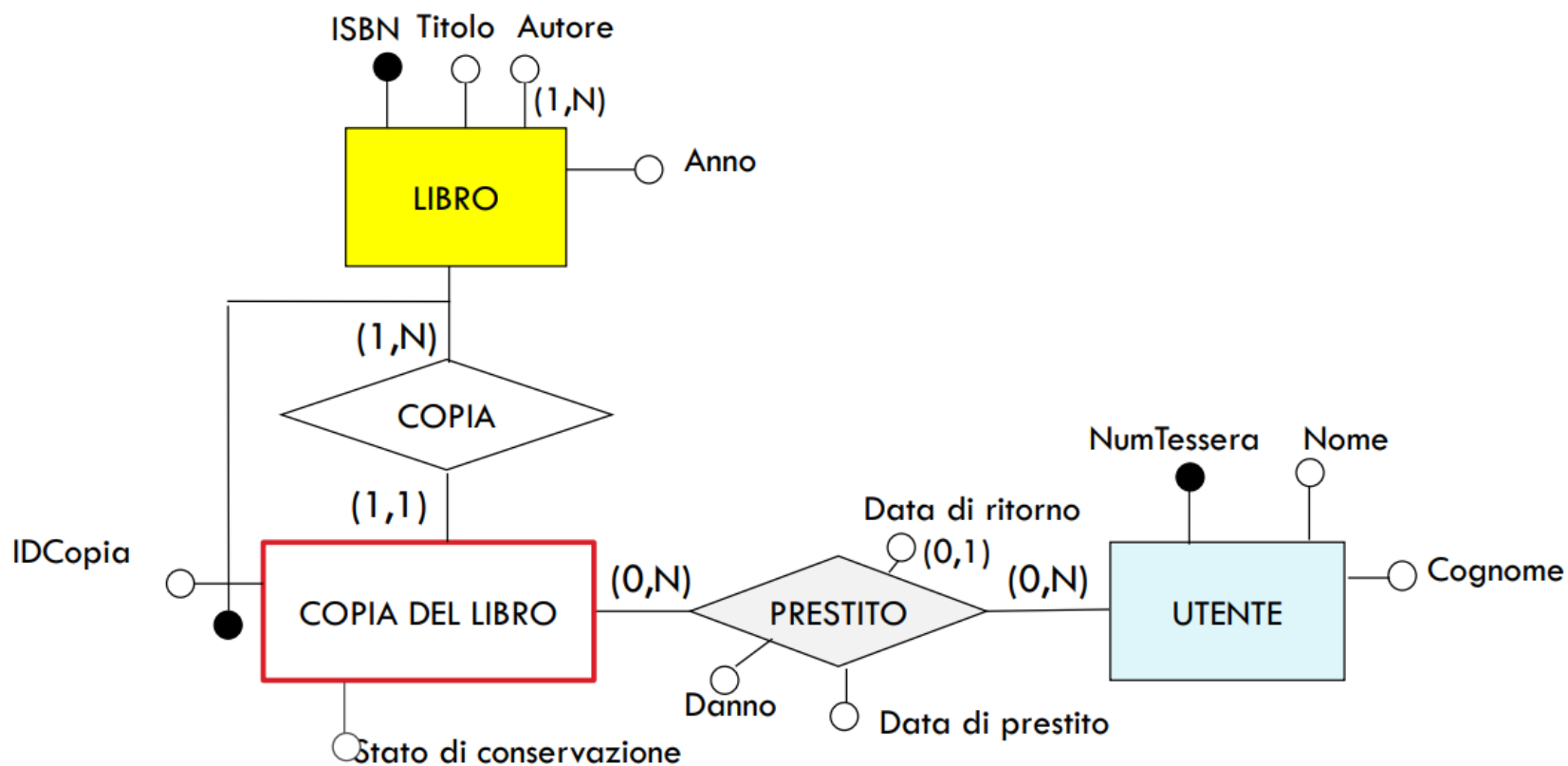
Pattern Instance-of

Esempio: In una biblioteca si vogliono conservare le informazioni sui **libri** (titolo, autore, anno, codice ISBN, stato di conservazione), sugli **utenti** (numTessera, nome, cognome) e sul **prestito** (data di prestito, eventuale data di restituzione, utente), **segnalando eventuali danni arrecati al libro...**



Pattern Instance-of

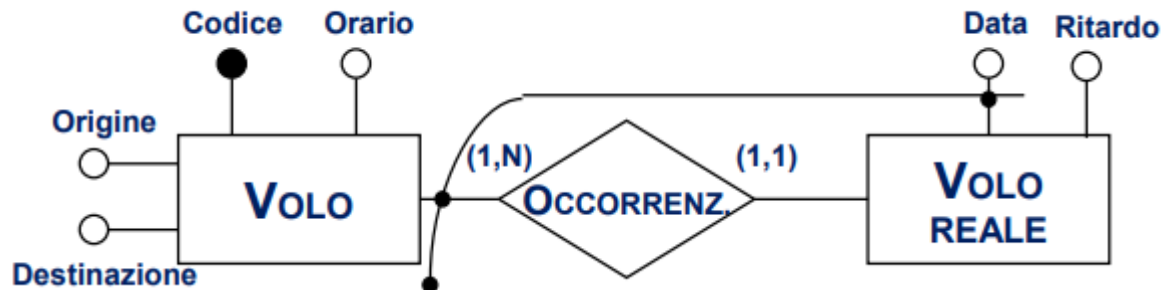
LIBRO \neq COPIA DEL LIBRO



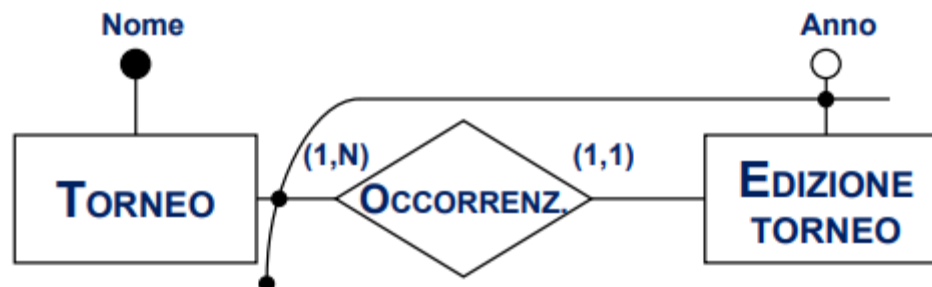
Pattern Instance-of

Le occorrenze di un'entità sono istanze delle occorrenze di un'altra entità.

- Il concetto astratto di volo vs. il volo reale: l'istanza ha gli stessi valori per gli attributi dell'occorrenza astratta, ma ha anche altri attributi specifici.

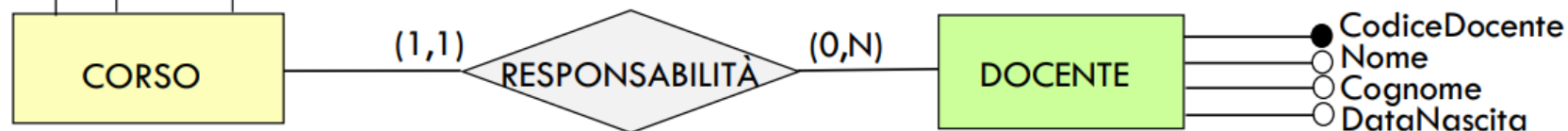


- L'identificazione dell'istanza in genere è mista e comprende l'identificatore della classe astratta e un attributo temporale

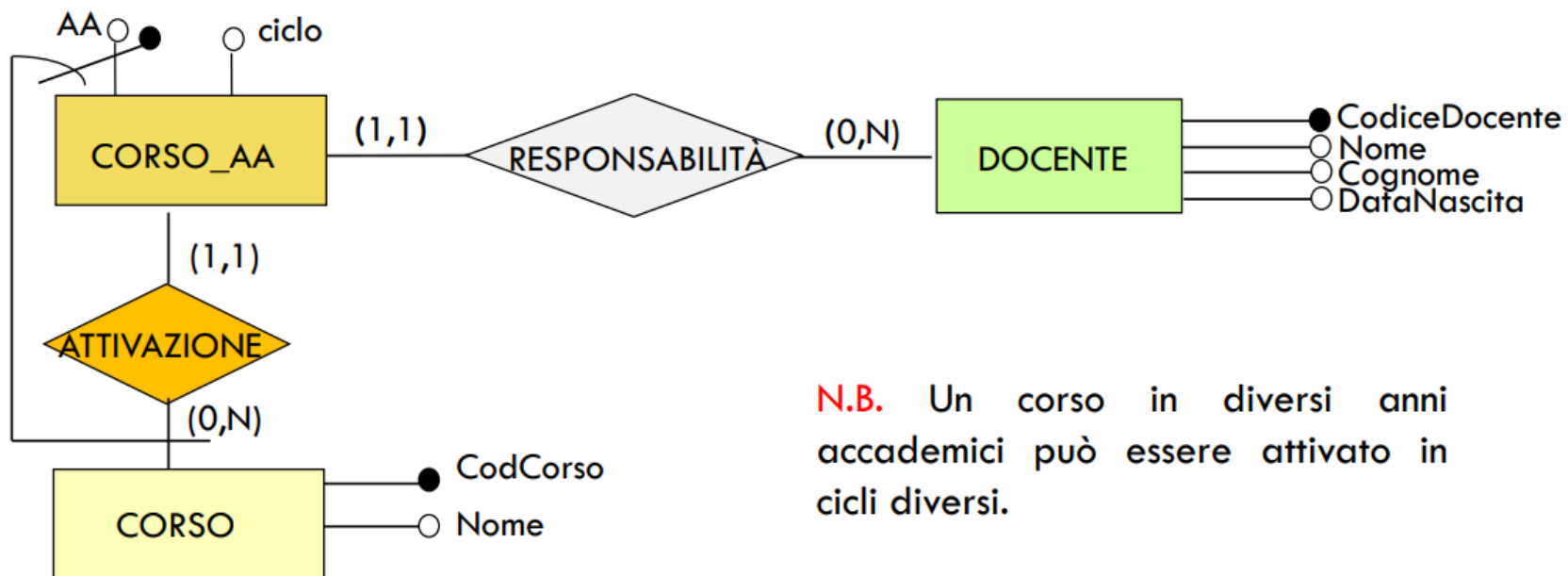


Storicizzazione

CodCorso Nome Ciclo ■ Fotografia dei corsi attivati in un anno accademico:



■ La storia considerando più anni accademici:

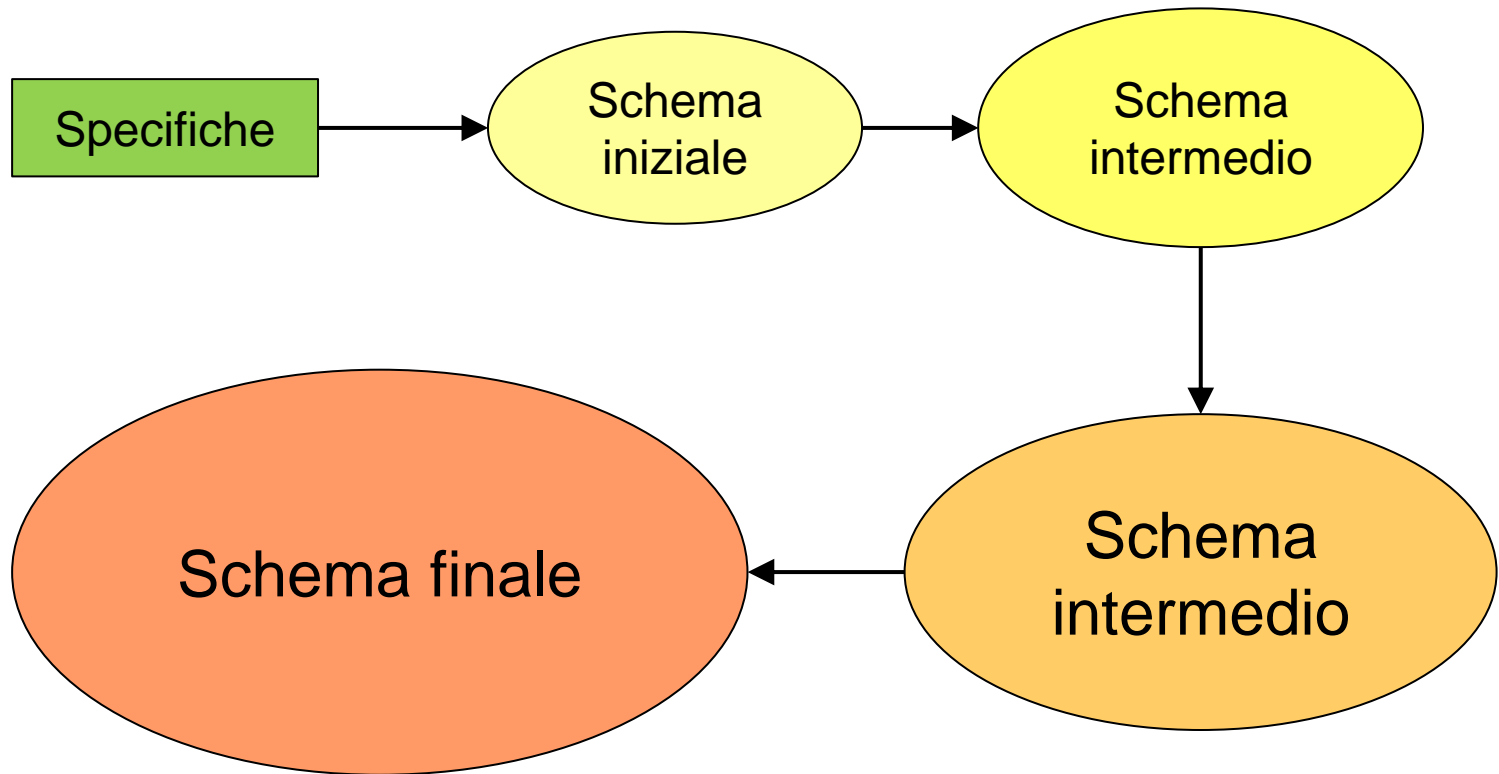


N.B. Un corso in diversi anni accademici può essere attivato in cicli diversi.

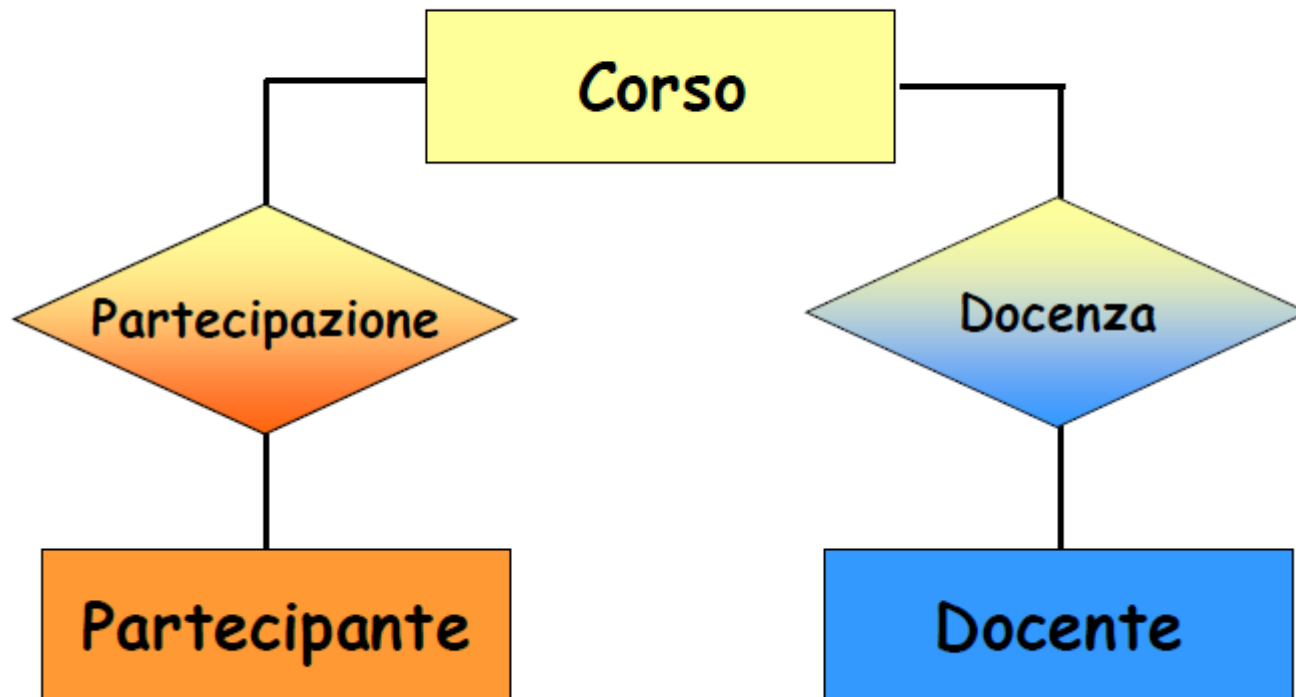
Strategie di progettazione

- ❑ Per affrontare progetti complessi è opportuno adottare uno specifico modo di procedere, ovvero una strategia di progettazione.
- ❑ Strategia **top-down**
 - Si parte da uno schema iniziale molto astratto ma completo, che viene successivamente raffinato fino ad arrivare allo schema finale.
- ❑ Strategia **bottom-up**
 - Si suddividono le specifiche in modo da sviluppare semplici schemi parziali ma dettagliati, che poi vengono integrati tra loro.
- ❑ Strategia **inside-out**
 - Lo schema si sviluppa “a macchia d’olio”, partendo dai concetti più importanti, aggiungendo quelli a essi correlati, e così via.
 - È un caso particolare della strategia bottom-up, in cui il raffinamento inizia dai concetti principali per poi estendersi a quelli più lontani attraverso una “navigazione” delle specifiche

Strategia top-down

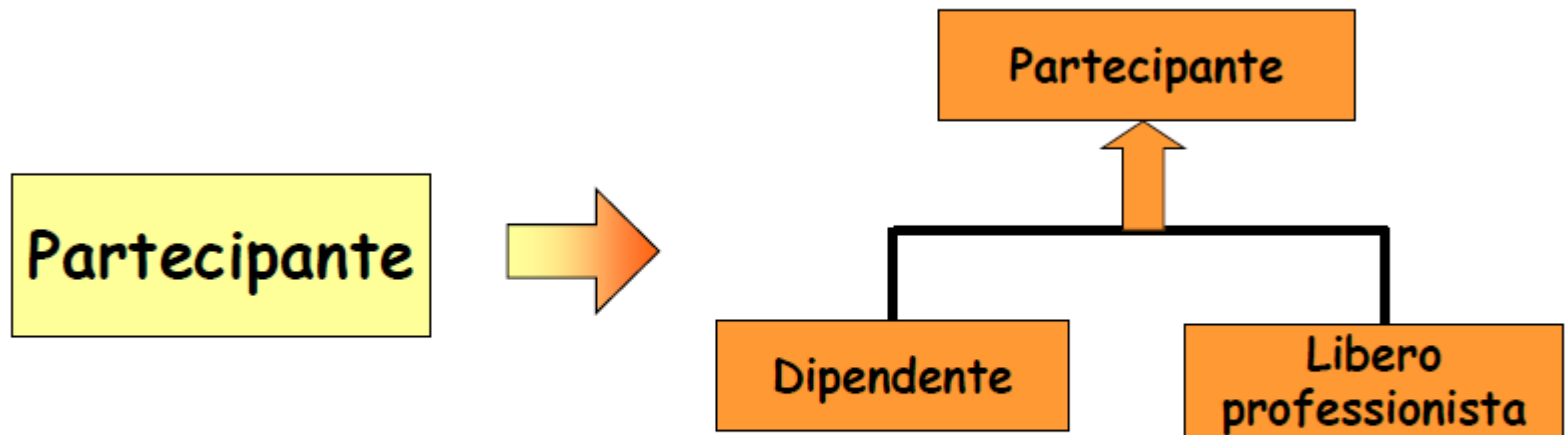


Schema scheletro



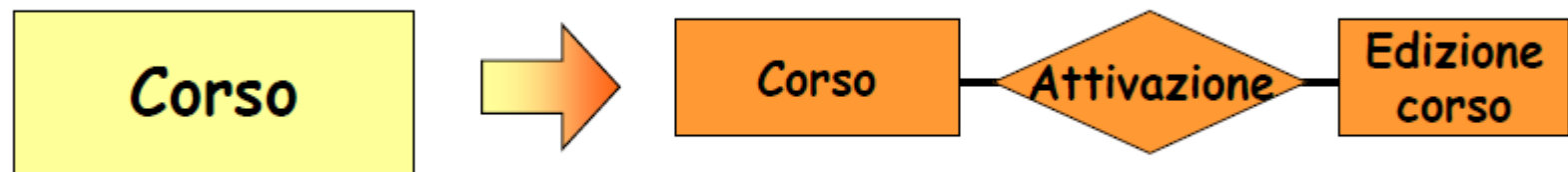
Primitive di raffinamento top-down (1)

- Entità → Gerarchia di entità

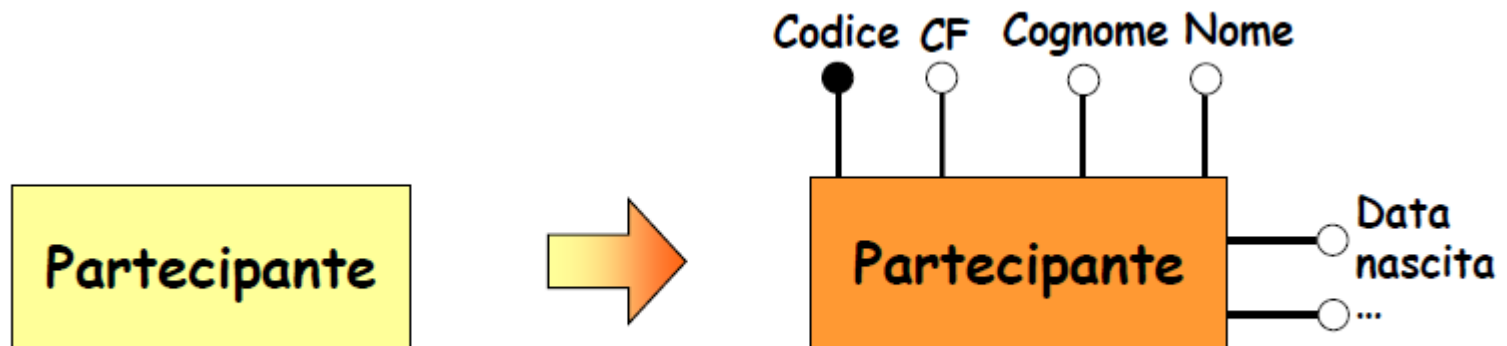


Primitive di raffinamento top-down (2)

- Entità → Entità e associazioni

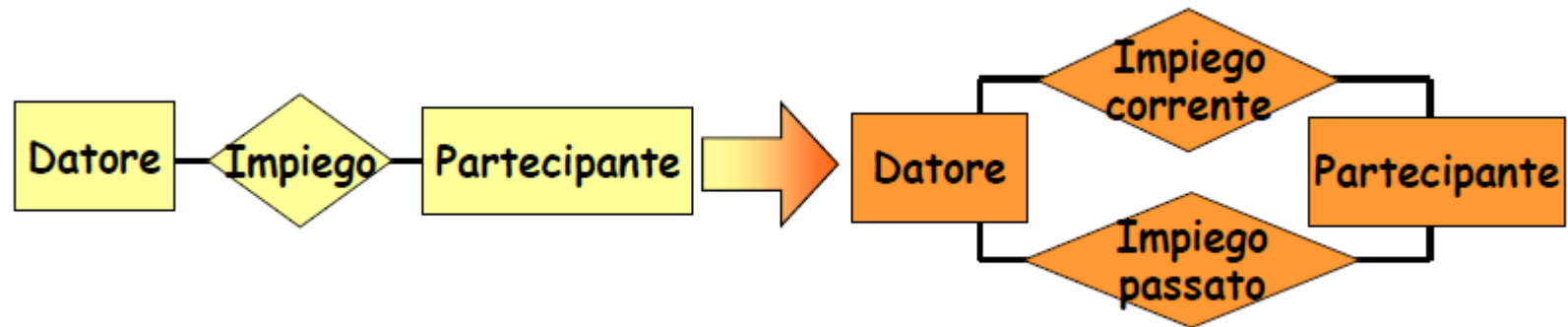


- Entità → Entità e attributi

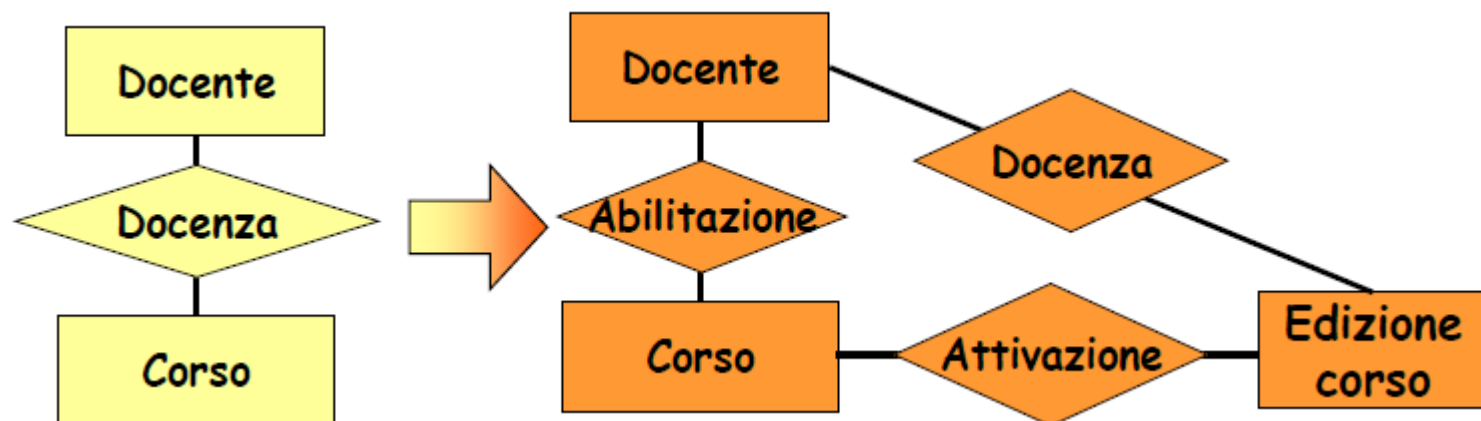


Primitive di raffinamento top-down (3)

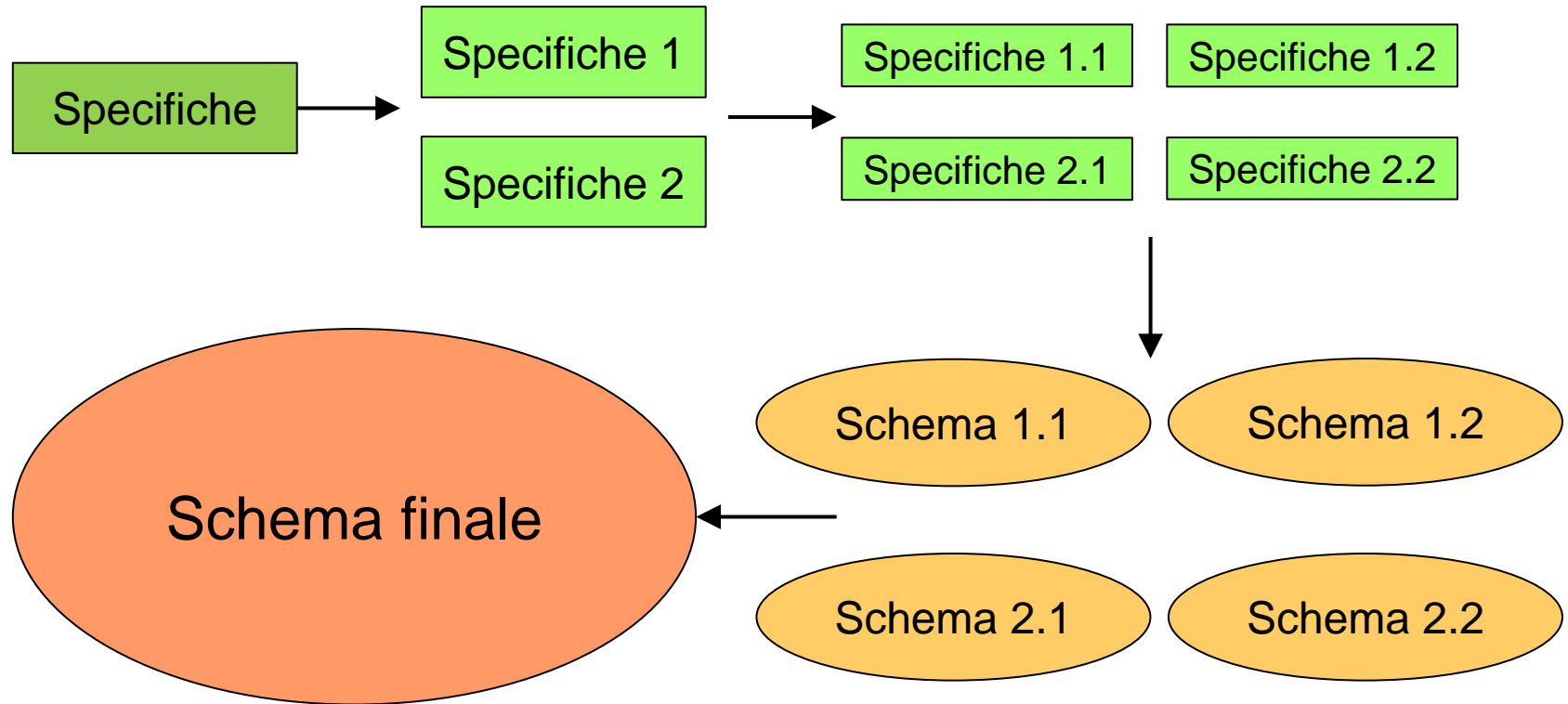
- Associazioni → Più associazioni



- Associazioni → Entità e associazioni



Strategia bottom-up



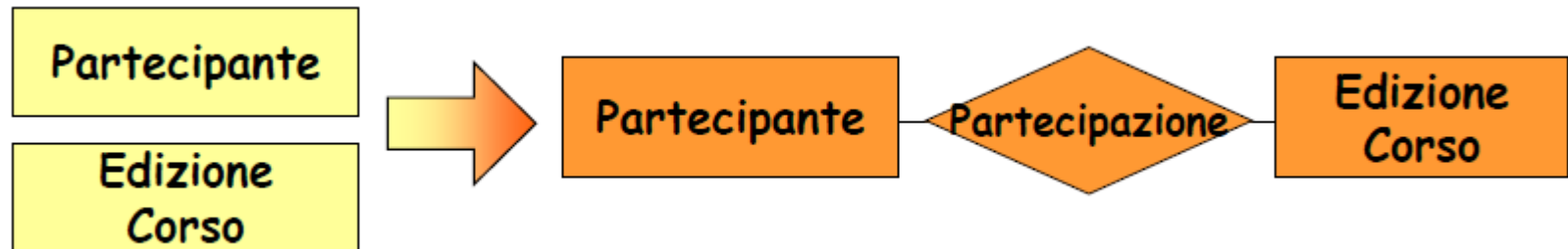
Primitive di raffinamento bottom-up (1)

- ❑ Specifiche su concetto → Nuova entità

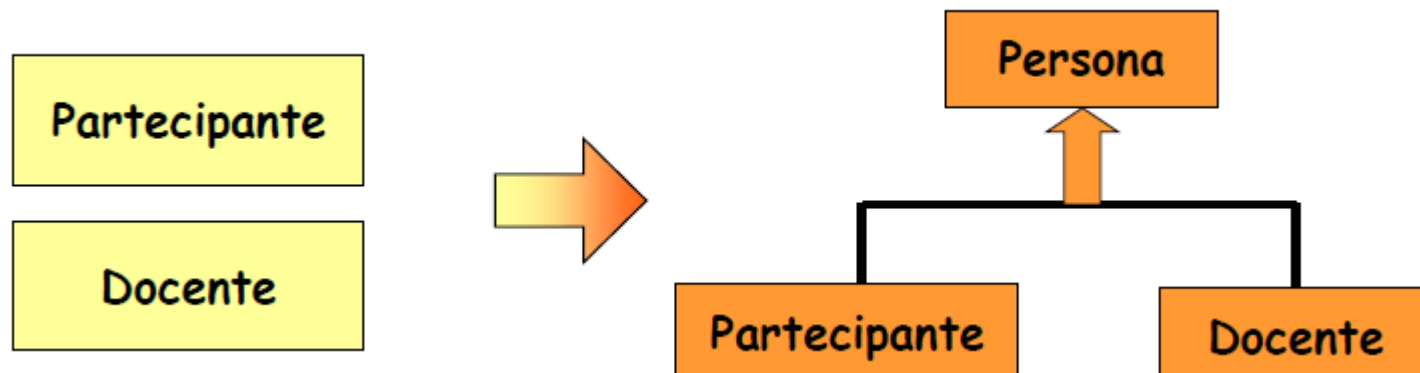


Primitive di raffinamento bottom-up (2)

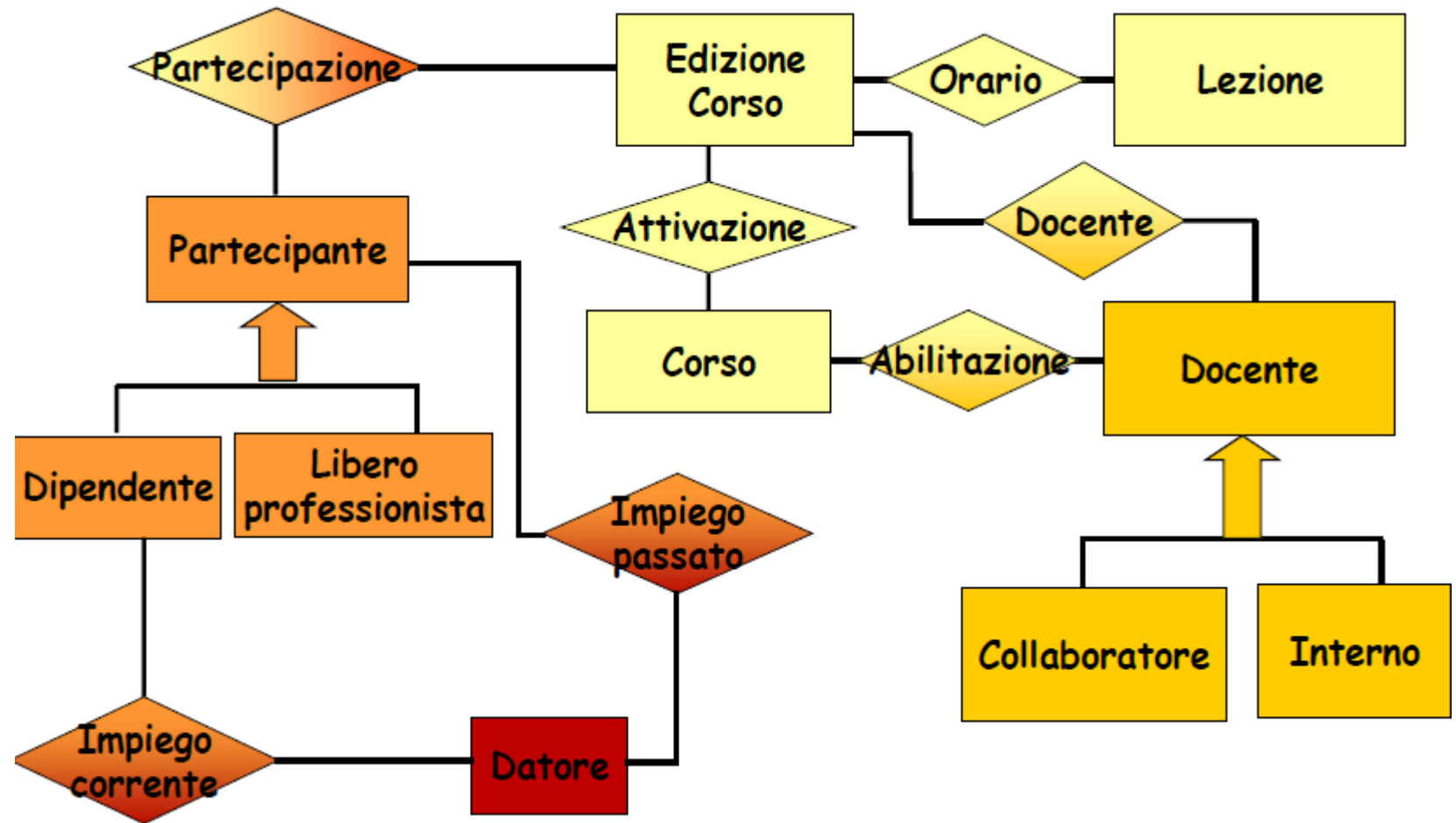
- Entità isolate → Associazioni tra le entità



- Entità isolate → Gerarchie di generalizzazione



Strategia inside-out



Strategie: pro e contro

❑ Top-down

- + non è inizialmente necessario specificare i dettagli;
- richiede sin dall'inizio una visione globale del problema, non sempre ottenibile in casi complessi.

❑ Bottom-up

- + permette una ripartizione delle attività;
- richiede una fase di integrazione.

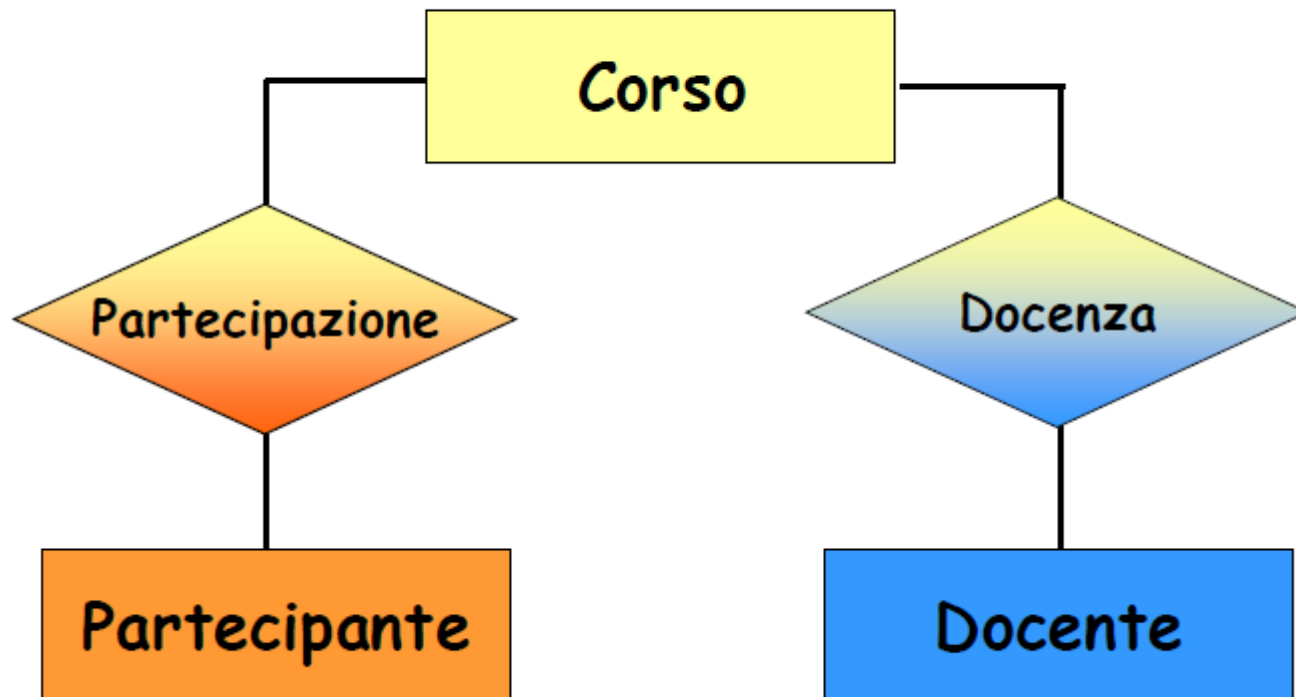
❑ Inside-out

- + non richiede passi di integrazione;
- richiede a ogni passo di esaminare tutte le specifiche per trovare i concetti non ancora rappresentati.

Un approccio “misto”

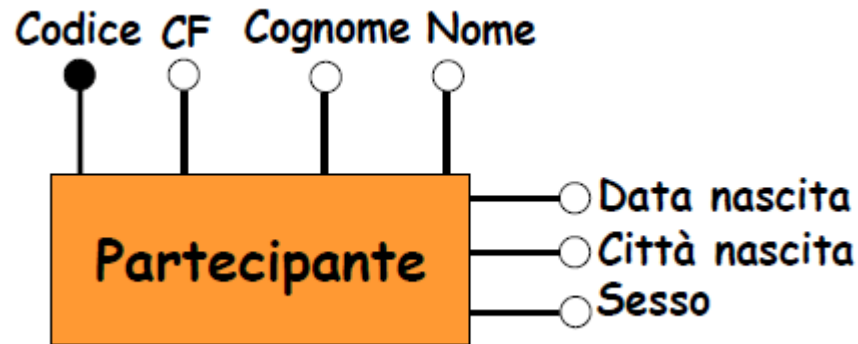
- Nella pratica si fa spesso uso di una strategia ibrida:
 1. Si individuano i concetti principali e si realizza uno **schema scheletro**, che rappresenta solamente i concetti più importanti;
 2. sulla base dello schema scheletro si può decomporre;
 3. successivamente si raffina, si espande, si integra
- Vediamo che cosa succede nel caso della società di formazione...

Schema scheletro



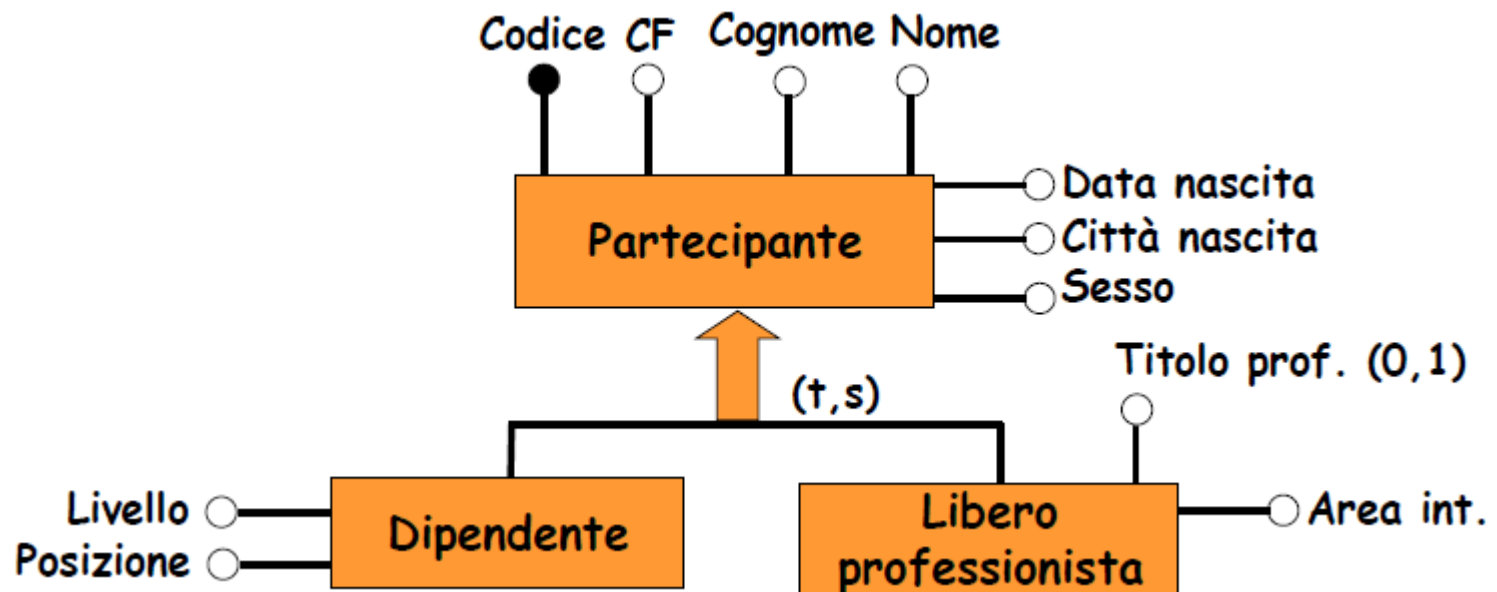
I partecipanti...

- Per i **partecipanti** (circa 5000), identificati da un codice, rappresentiamo il codice fiscale, il nome, il cognome, la data di nascita, il sesso, la città di nascita...



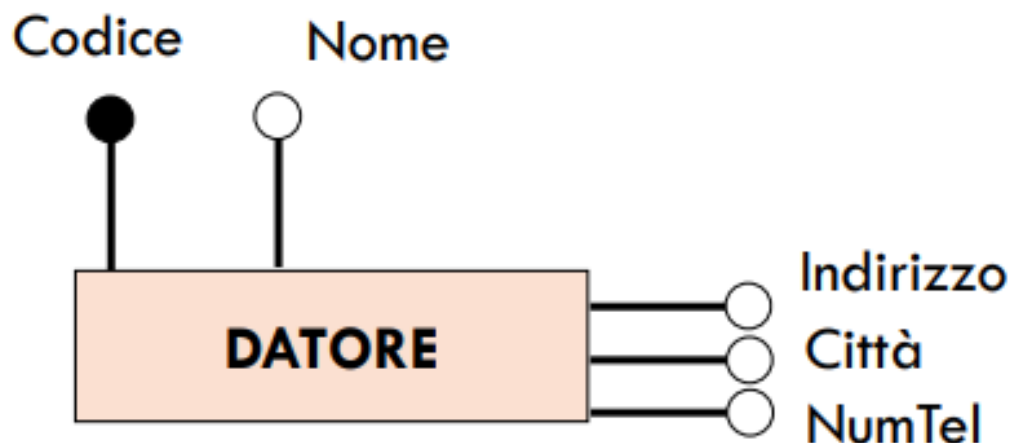
Raffinamento di partecipanti

- Per i partecipanti che sono **liberi professionisti**, rappresentiamo l'area di interesse e, se lo possiedono, il titolo professionale. Per i partecipanti che sono **dipendenti**, rappresentiamo invece il livello e la posizione ricoperta.

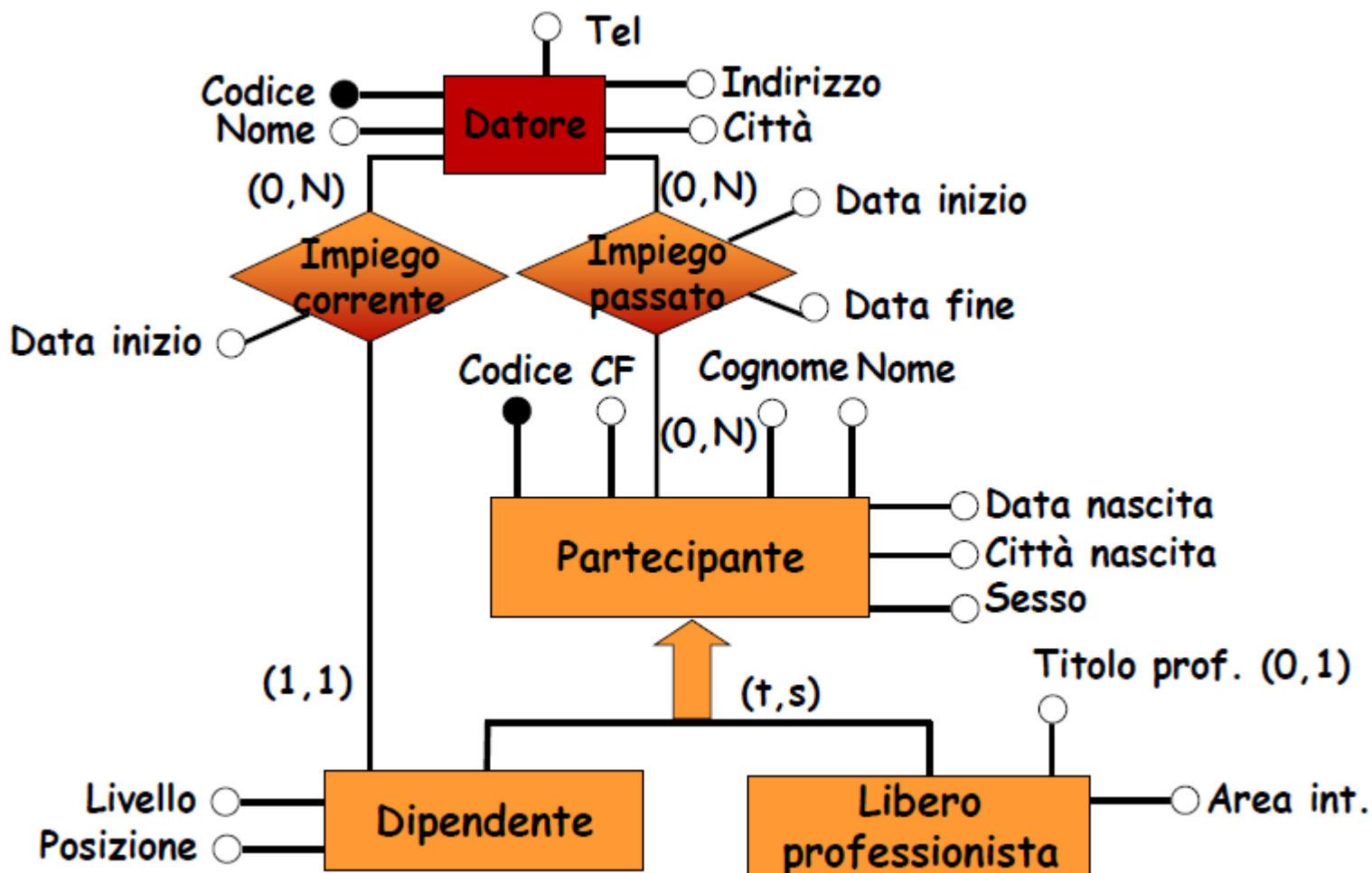


I datori di lavoro...

- Relativamente ai **datori di lavoro** presenti e passati dei partecipanti rappresentiamo il nome, l'indirizzo, la città e il numero di telefono. Per ciascun partecipante memorizziamo il nome dell'attuale datore di lavoro e di quelli precedenti (insieme alle date di inizio e fine rapporto).

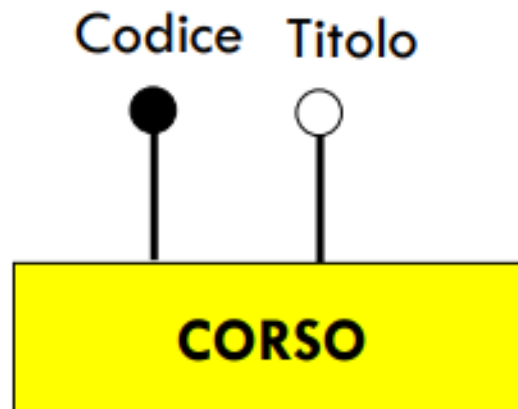


Raffinamento di datori di lavoro

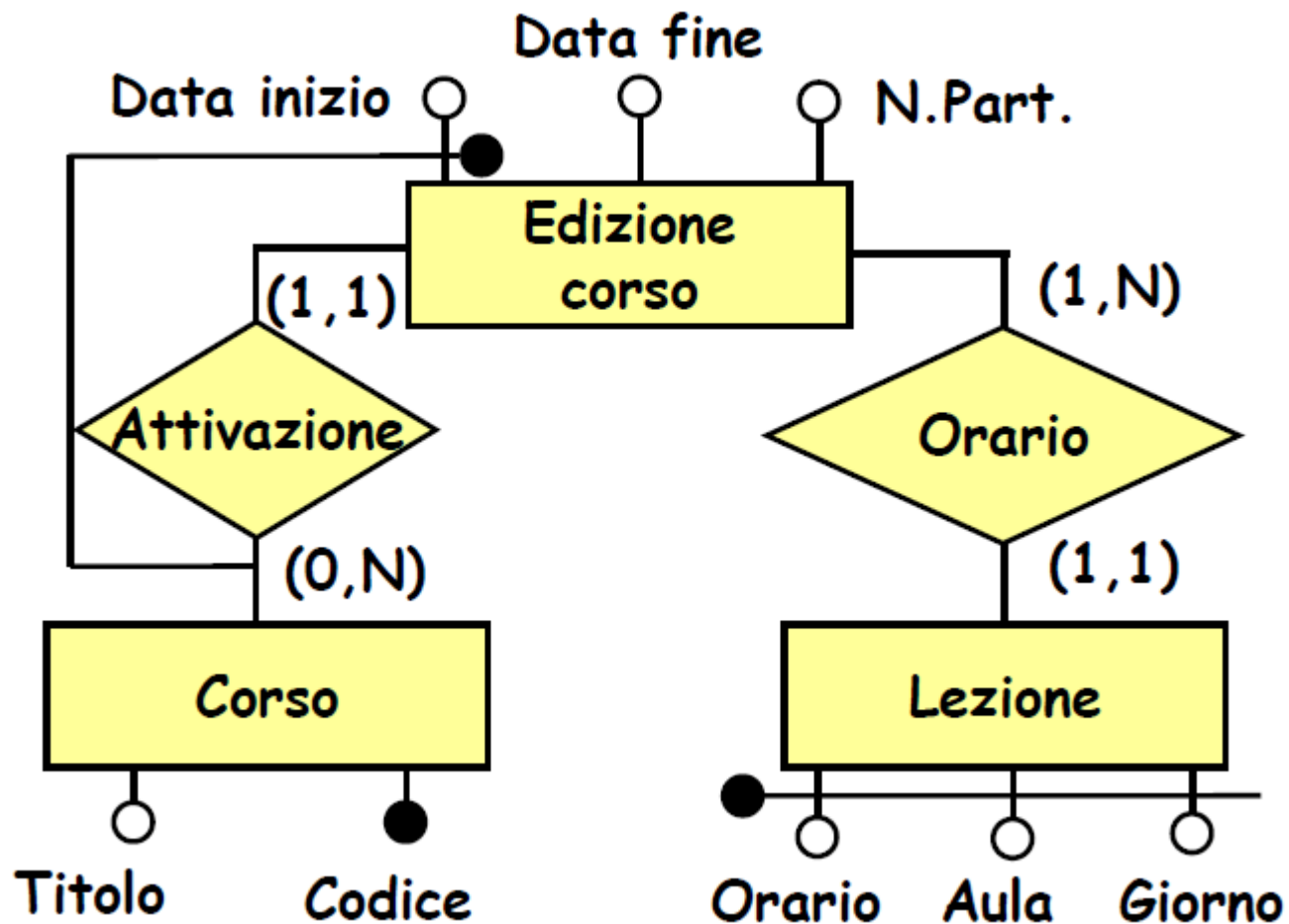


I corsi...

- ❑ Per i **corsi** (circa 200), rappresentiamo il titolo e il codice, le varie edizioni con date di inizio e fine e, per ogni edizione, rappresentiamo il numero di partecipanti e i giorni e le ore di lezione nonché i luoghi dove si svolgono le lezioni.



Raffinamento di corsi

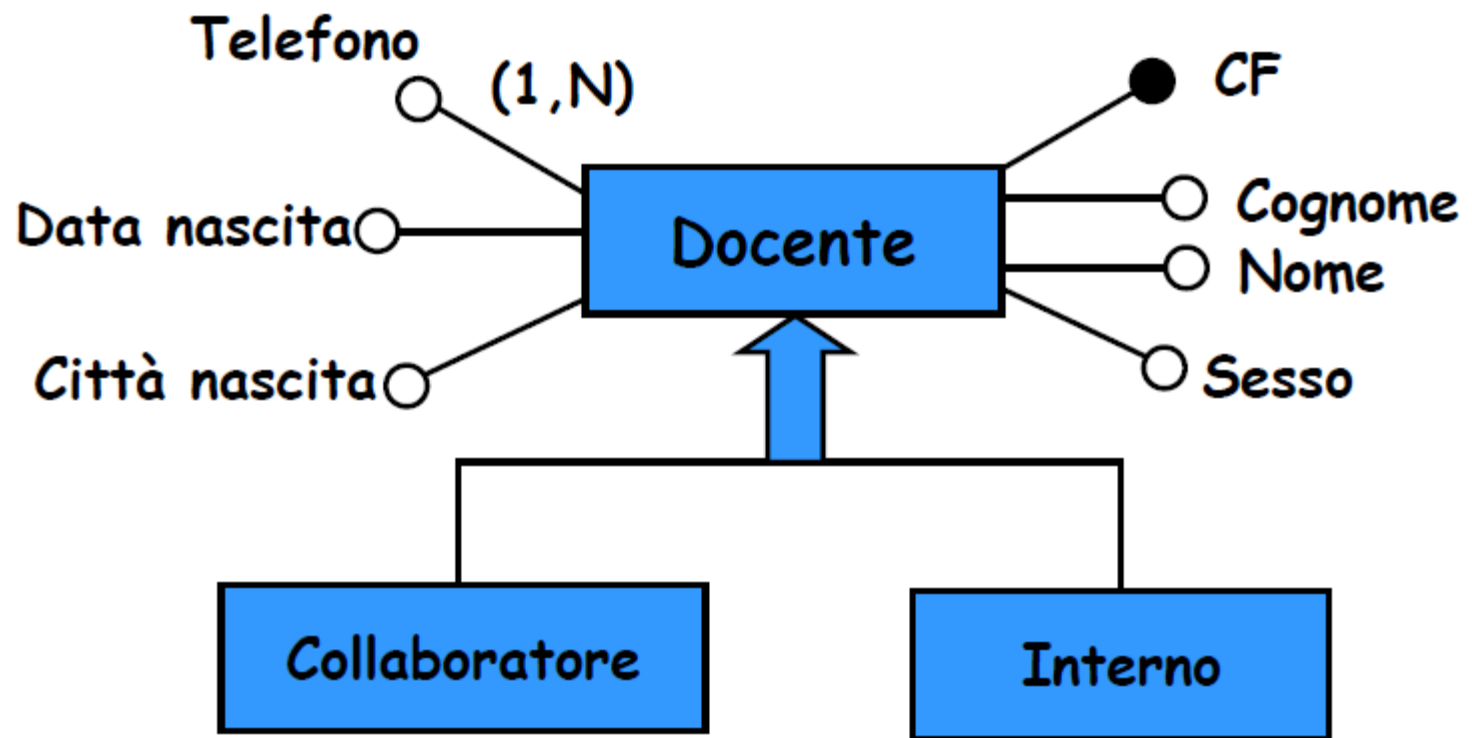


I docenti...

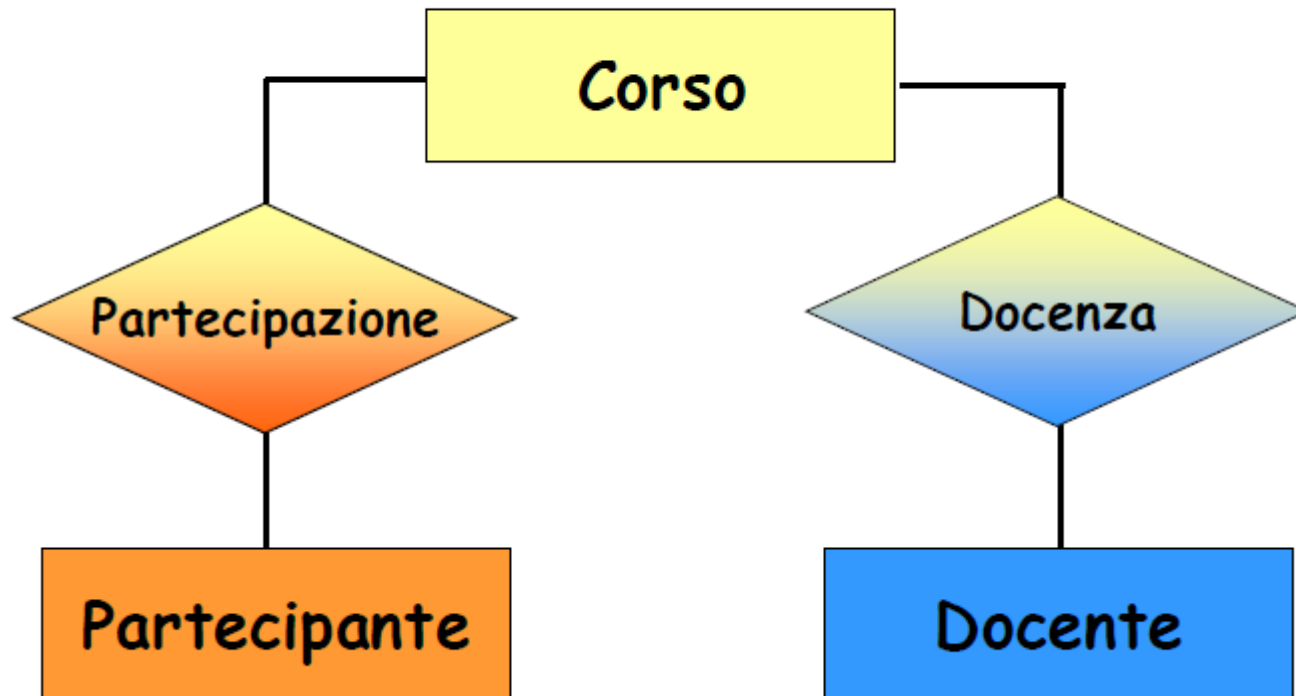
- Per i **docenti** (circa 300), rappresentiamo il codice fiscale, il nome, il cognome, il sesso, la data di nascita, la città di nascita e uno o più numeri di telefono. I docenti possono essere dipendenti interni della società di formazione o collaboratori esterni.



Raffinamento di docente

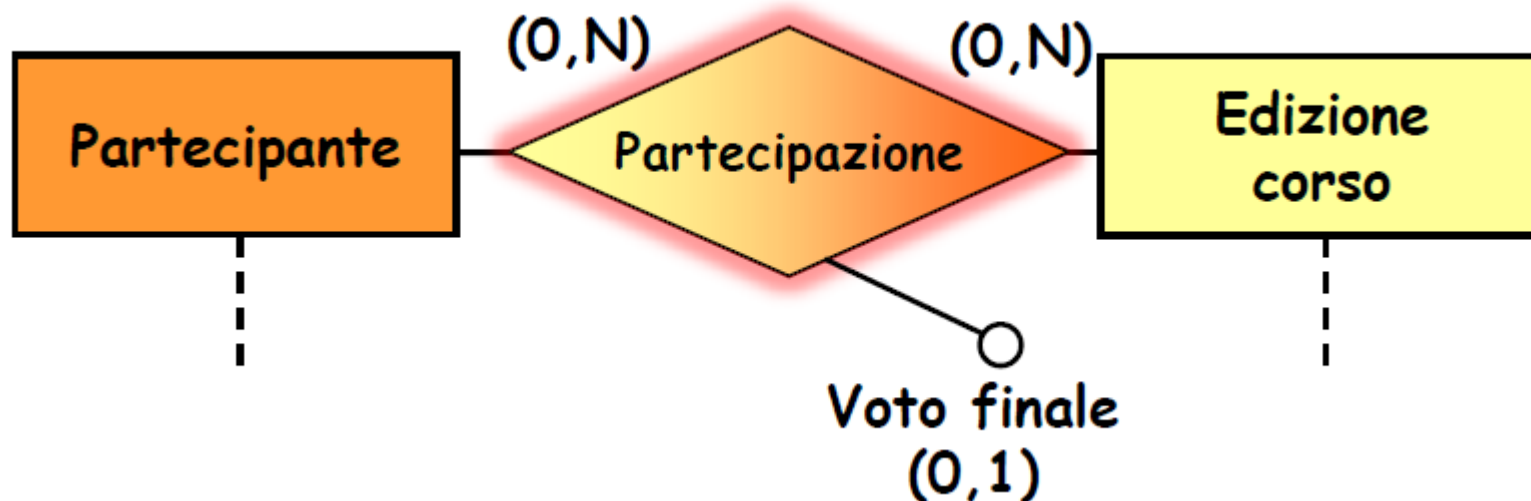


Integrazione: schema di riferimento



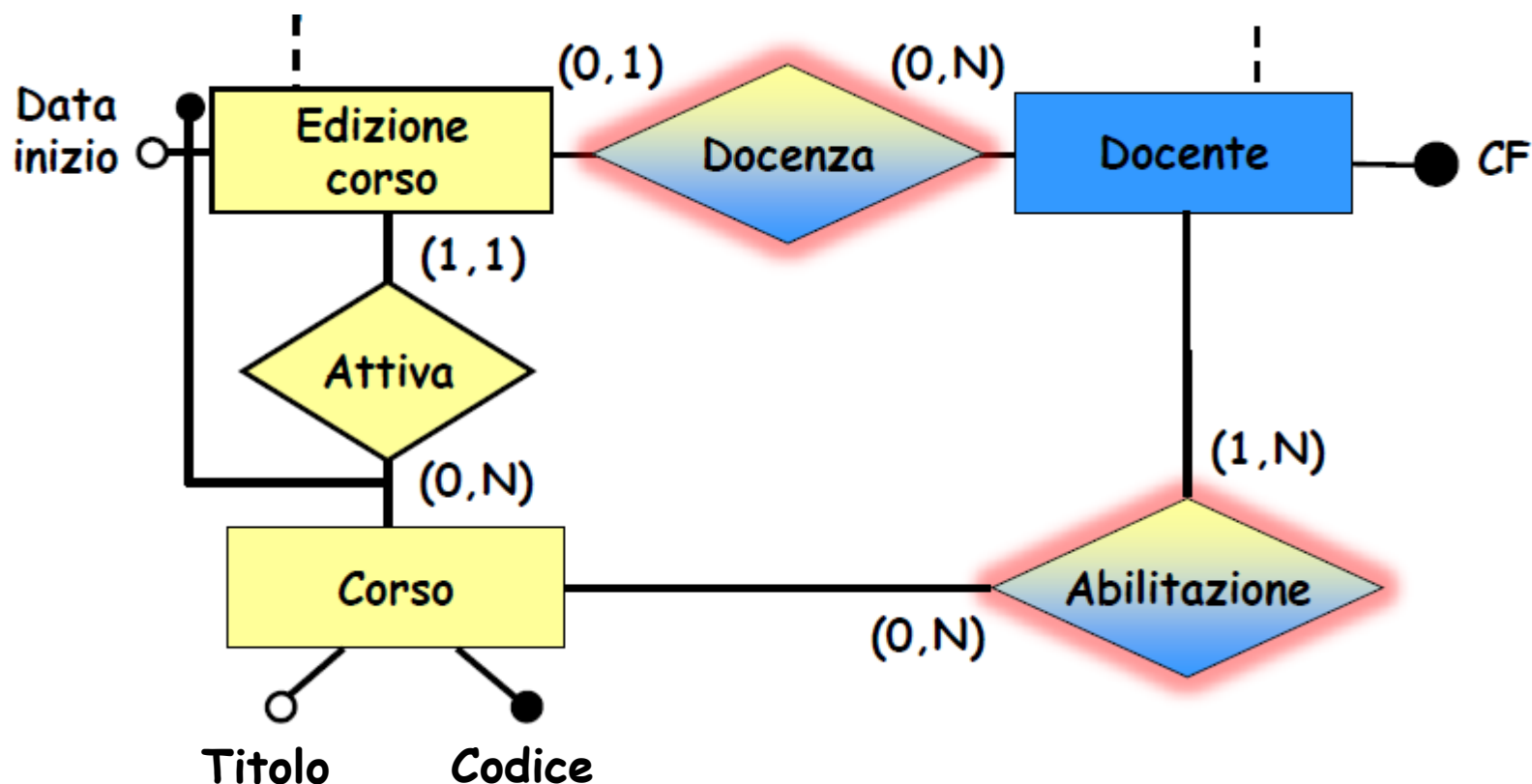
Integrazione: Partecipante e Corso

- Per i partecipanti si memorizzano le edizioni dei corsi che stanno attualmente frequentando e quelli che hanno frequentato nel passato, con la relativa votazione finale in decimi.

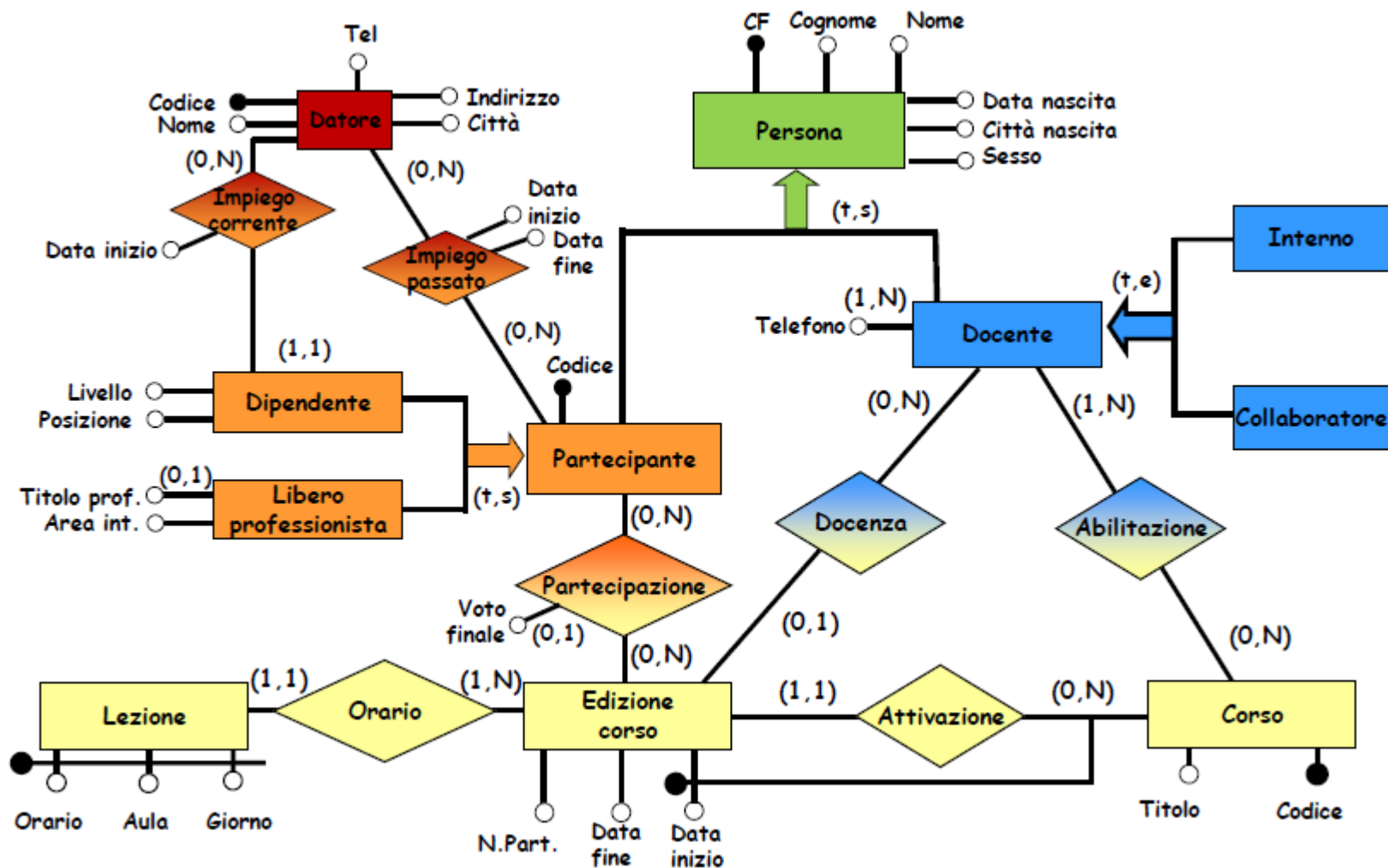


Integrazione: Docente e Corso

- Per i docenti si memorizzano il titolo del corso che insegnano, di quelli che hanno insegnato nel passato e di quelli che possono insegnare.



Schema finale



Qualità di uno schema concettuale

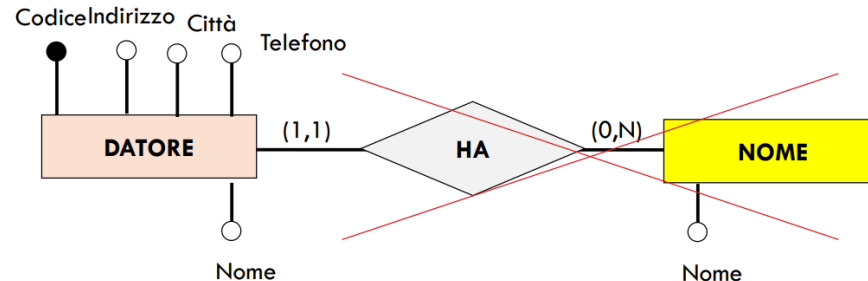
- ❑ Lo schema E/R deve essere verificato accuratamente per verificare che risponda a requisiti di:
 - **Correttezza**: non devono essere presenti errori (sintattici o semantici).
 - **Completezza**: tutti i dati di interesse devono essere specificati.
 - **Leggibilità**: riguarda anche aspetti prettamente estetici dello schema.
 - **Minimalità**: è importante capire se esistono elementi ridondanti nello schema; in alcuni casi ciò non è un problema, ma può essere viceversa una scelta di progettazione volta a favorire l'esecuzione di certe operazioni.

Correttezza

Definizione

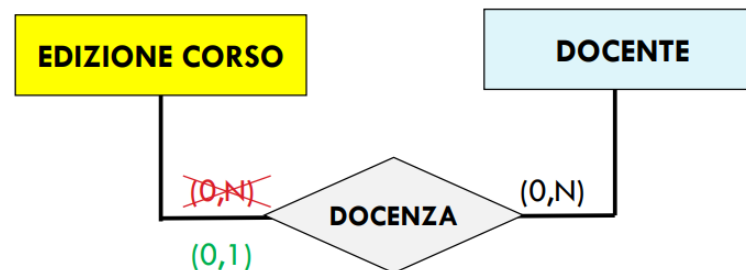
Rispetto al modello

La correttezza sintattica significa che i costrutti di modellazione nello schema vengono utilizzati in base alla loro definizione e proprietà.



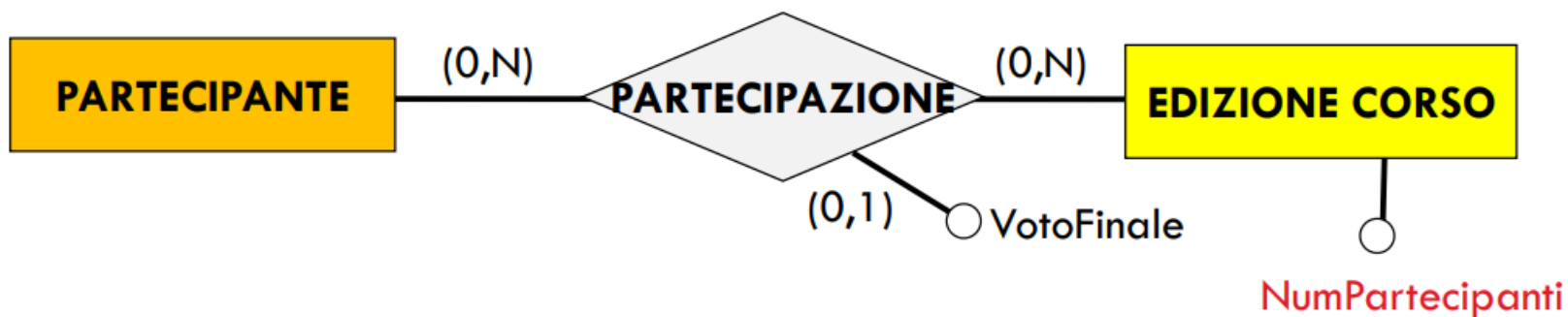
Rispetto ai requisiti

La correttezza rispetto ai requisiti riguarda la corretta rappresentazione dei requisiti in termini di costrutti del modello.



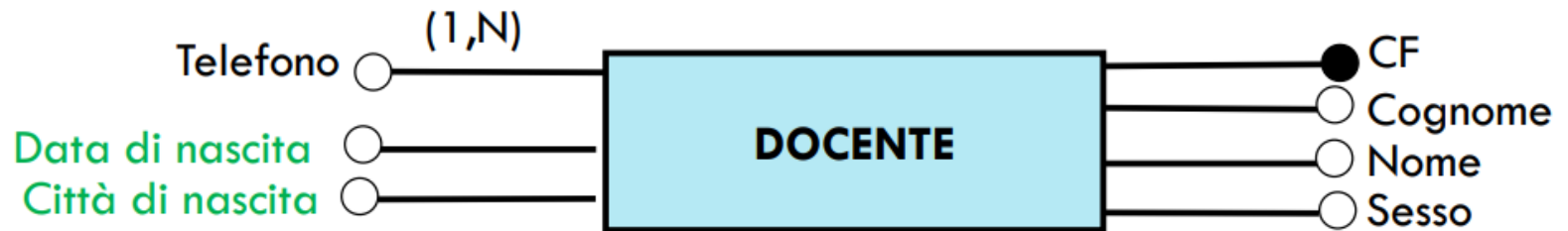
Minimalità

- **Definizione** - La minimalità afferma che i concetti ridondanti, vale a dire i concetti che possono essere espressi in termini di altri concetti nello schema, sono assenti (o evidenziati nella documentazione).



Completezza

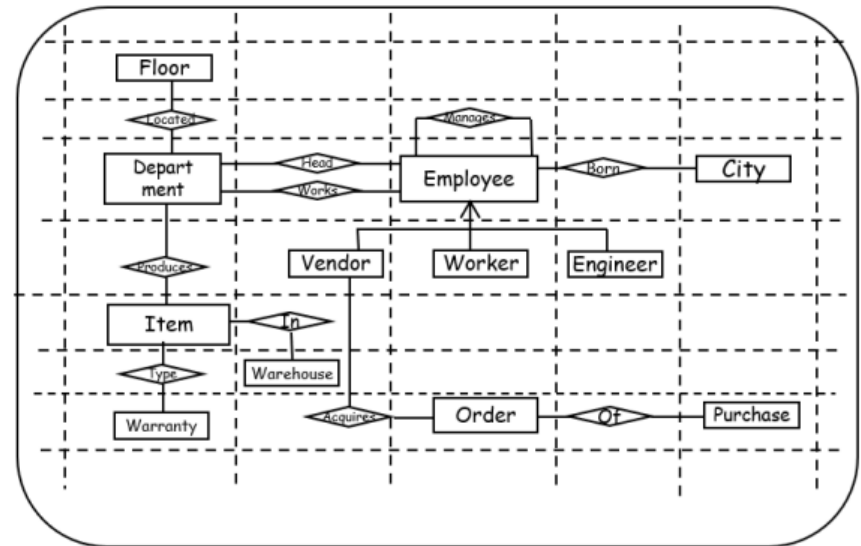
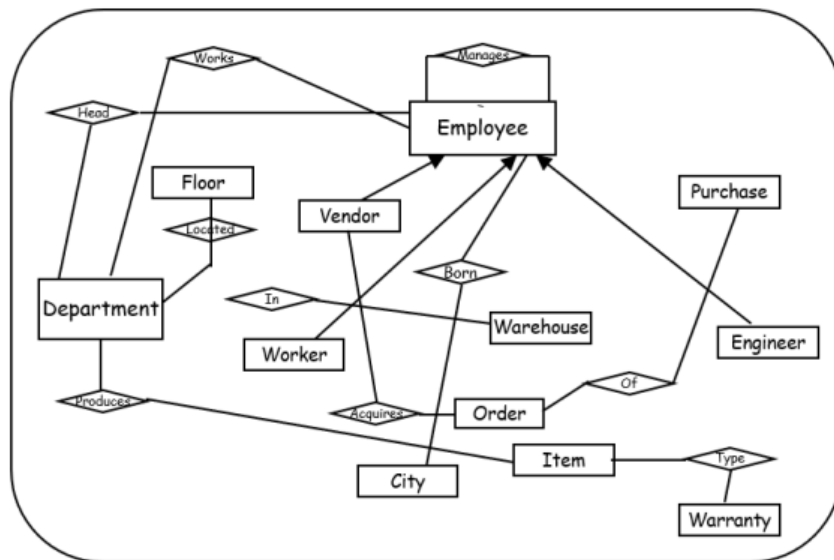
- ❑ **Definizione** - La completezza misura la misura in cui uno schema concettuale include tutti gli elementi concettuali necessari per soddisfare i requisiti.



- ❑ Non erano specificati nello schema precedente!

Leggibilità

- ❑ **Definizione** - la leggibilità corrisponde alla percezione della chiarezza con cui viene disegnato il diagramma e alla relativa facilità di comprensione



Pertinenza

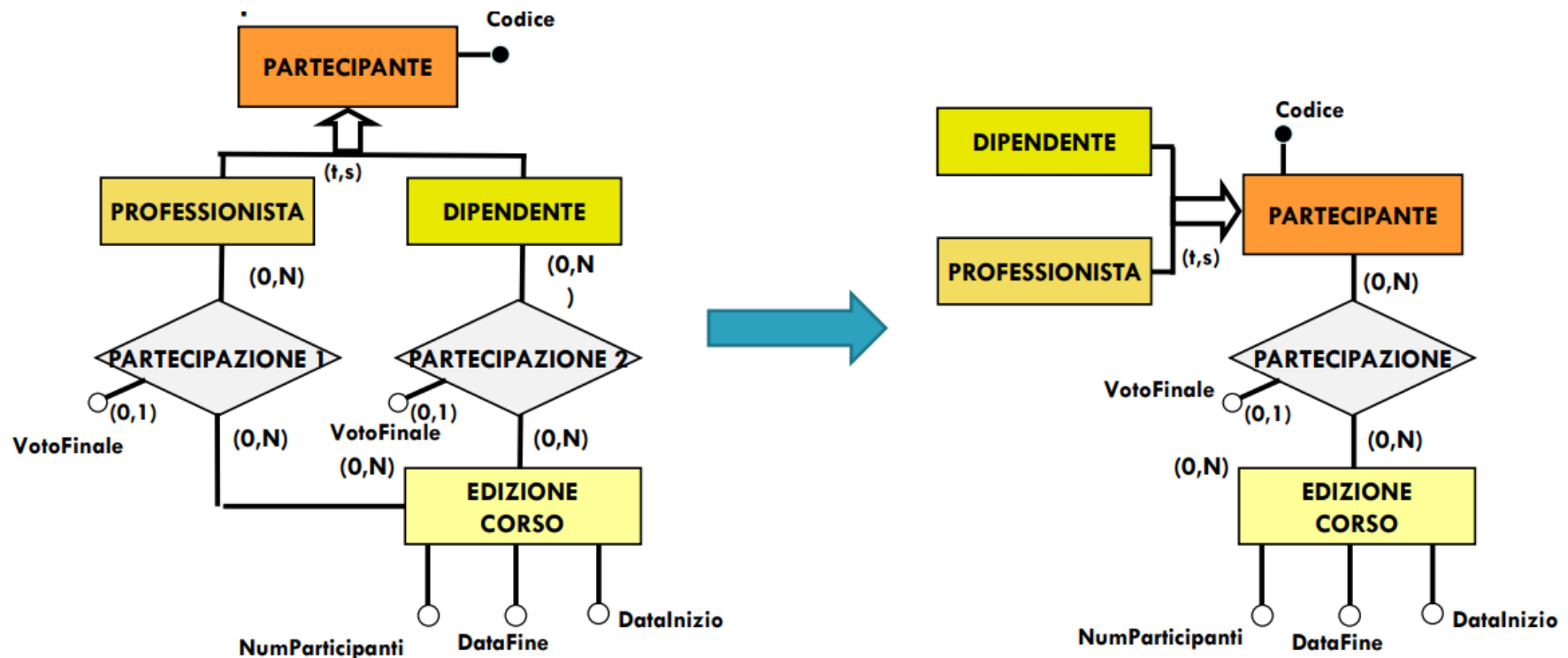
- **Definizione** - La pertinenza è la qualità complementare della completezza e misura quanti elementi concettuali non necessari sono inclusi nello schema concettuale



L'indirizzo non è richiesto

Compattezza

- ❑ **Definizione** - La compattezza della rappresentazione dello schema porta a scegliere, tra i diversi schemi concettuali che rappresentano in modo equivalente una certa realtà, quello o quelli più compatti, cioè utilizzare pochi concetti.



Metodologia basata sulla strategia mista

- ❑ Analisi dei requisiti
 - analizzare i requisiti ed eliminare le ambiguità;
 - costruire un glossario dei termini, raggruppare i requisiti.
- ❑ Passo base
 - definire uno schema scheletro con i concetti più rilevanti.
- ❑ Passo di decomposizione (se necessario o appropriato)
 - decomporre i requisiti con riferimento ai concetti nello schema scheletro.
- ❑ Passo iterativo (da ripetere finché non si è soddisfatti)
 - raffinare i concetti presenti sulla base delle loro specifiche;
 - aggiungere concetti per descrivere specifiche non descritte.
- ❑ Passo di integrazione (se si è decomposto)
 - integrare i vari sottoschemi in uno schema complessivo, facendo riferimento allo schema scheletro.
- ❑ Analisi di qualità (ripetuta e distribuita)
 - verificare le qualità dello schema e modificarlo.

Sommario

- ❑ La fase di analisi dei requisiti è fondamentale per poter progettare una base di dati che rispetti i requisiti.
- ❑ Non è possibile standardizzare tale fase; ci si avvale dunque di regole di buon senso e di una serie di strumenti che riducono il rischio di commettere errori grossolani e permettono di costituire una valida documentazione.
- ❑ Per la progettazione dello schema E/R sono possibili diverse strategie, di cui quella mista è senz'altro la più diffusa e adeguata anche nel caso di progetti estremamente complessi.