

La progettazione concettuale: il modello ER

Matteo Gorgone
Università degli Studi di Messina

La progettazione concettuale

La **progettazione concettuale** consiste nel riorganizzare tutti gli elementi presenti nella documentazione sulle specifiche per rappresentare la realtà di interesse in termini di una descrizione formale e completa, indipendentemente dai criteri di rappresentazione usati nei DBMS.

Il prodotto di tale attività è lo **schema concettuale**, che rappresenta il documento ufficiale di riferimento e di comunicazione per i progettisti della successiva fase di progettazione logica.

Il modello concettuale ER

Tra i possibili modelli concettuali è molto diffuso il **modello Entità/Associazioni** (in inglese **ER**, da *Entity/Relationship*).

Si tratta di un modello grafico per la descrizione dei dati e delle loro relazioni all'interno di una realtà di interesse.

È uno strumento per svolgere l'analisi delle caratteristiche di una realtà indipendentemente dagli eventi che in essa accadono.

Per la costruzione di uno schema ER si parte dal semplice concetto che la realtà da rappresentare è composta da **entità**, ognuna delle quali è caratterizzata da specifiche **proprietà** (o **attributi**).

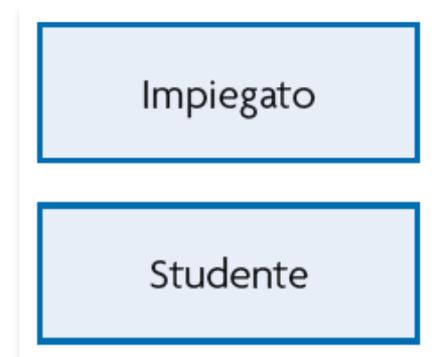
Le entità

L'**entità** permette di rappresentare una classe di oggetti (cioè fatti, persone, cose e così via) che appartengono a una realtà e che:

- hanno proprietà comuni;
- hanno esistenza autonoma;
- sono di interesse per l'applicazione.

Ogni singolo oggetto che appartiene a un'entità è detto **istanza di entità** (o anche **occorrenza di entità**).

La **rappresentazione grafica** di un'entità, in uno schema concettuale, è un rettangolo con al centro il nome che identifica univocamente l'entità in questione.



Gli attributi

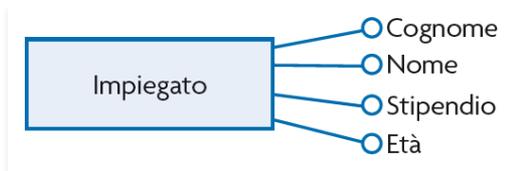
Gli **attributi** sono le proprietà ossia i fatti che si intendono rappresentare e che interessano poiché descrivono caratteristiche di un'entità.

Ogni attributo è specificato da:

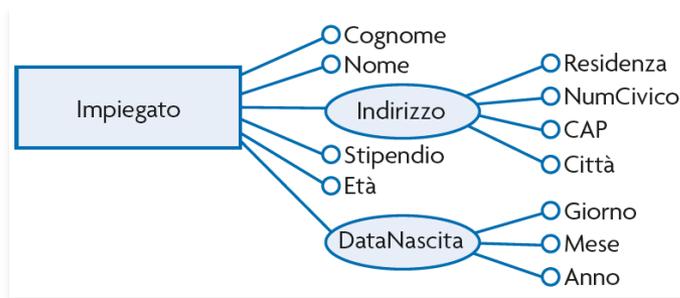
- un **nome**;
- un **formato**: il tipo di valori che può assumere (stringa, intero, reale, booleano, etc);
- una **dimensione**: la quantità massima di caratteri o cifre inseribili;
- un **valore**: i diversi valori assunti dagli attributi determinano le diverse istanze dell'entità. L'insieme dei possibili valori assunti da un attributo si chiama **dominio** dell'attributo.
- un'**opzionalità**: può non essere sempre valorizzato. Attributi obbligati e facoltativi.

Gli attributi

Un attributo si definisce **semplice** (o **elementare**) quando non è ulteriormente scomponibile, cioè rappresenta un'unità informativa di base che caratterizza un'entità e alla quale è associato un singolo valore.



Un attributo si definisce **composto** (o **aggregato**) quando è costituito da un'aggregazione di altri attributi, che a loro volta possono essere semplici o frutto di altre aggregazioni di attributi.

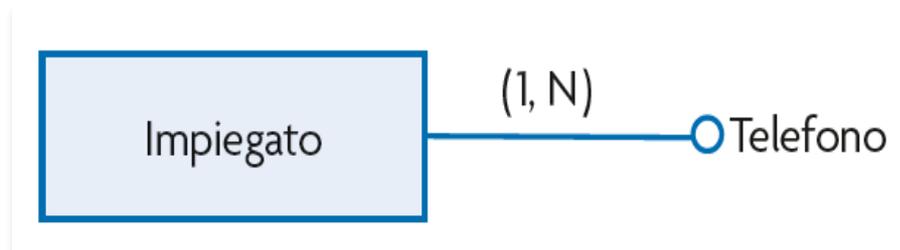


Gli attributi

Un attributo si definisce **multiplo** quando possono essergli associati contemporaneamente anche più valori dello stesso tipo.

La **cardinalità di attributo** è descritta da una coppia di valori (x, y) che specificano, rispettivamente, il numero minimo e massimo di valori dell'attributo associati a ogni istanza dell'entità.

I parametri x (cardinalità minima) e y (cardinalità massima) possono assumere i valori 0, 1 e N.



Gli attributi

Un attributo si definisce **opzionale** quando può anche non assumere un valore e quando può accettare valori nulli (altrimenti l'attributo viene detto **obbligatorio**).

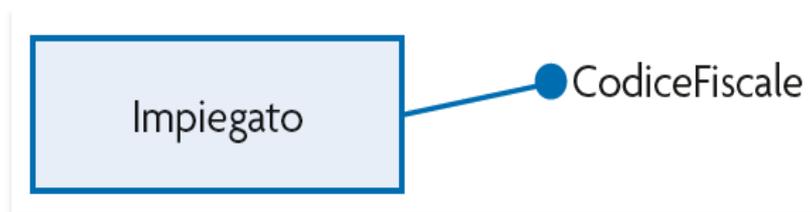
Cardinalità	Valore	Significato
Minima	0	L'attributo è opzionale (l'informazione potrebbe essere non disponibile)
Minima	1	L'attributo è obbligatorio
Massima	N	L'attributo è multivalore

Un attributo si definisce **derivato** quando il suo valore può essere ricavato implicitamente elaborando un attributo o un insieme di attributi presenti nello schema.

Gli attributi chiave

Si definisce **chiave candidata** o **superchiave** un attributo o un insieme di attributi in grado di distinguere un'istanza di entità dall'altra in modo univoco.

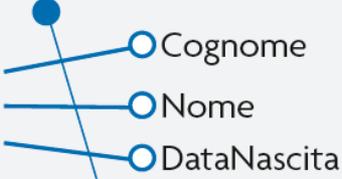
Una chiave composta da un solo attributo si rappresenta come un attributo semplice, con la sola variante che il pallino è pieno.



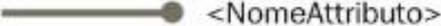
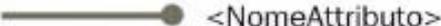
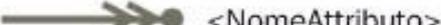
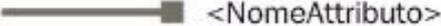
Nel caso di chiave composta da più attributi, la rappresentazione avviene con una linea terminante con un pallino pieno che unisce gli attributi che la compongono.

Tra le chiavi candidate, quella con il minor numero di attributi prende il nome di **chiave**, o meglio **chiave primaria** (*primary key*).

Gli attributi: riepilogo

Tipo di attributo	Rappresentazione
Semplice	
Composto	
Multiplo (obbligatorio)	
Opzionale	
Multiplo opzionale	
Chiave (composta da un solo attributo)	
Chiave (composta da più attributi)	

Convenzioni

Oggetto	Descrizione
Entità	
Attributo non chiave	
Attributo chiave	
Attributo multiplo	
Attributo composto	

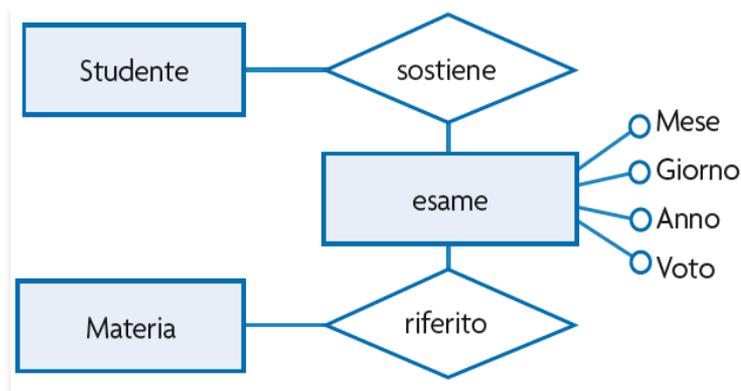
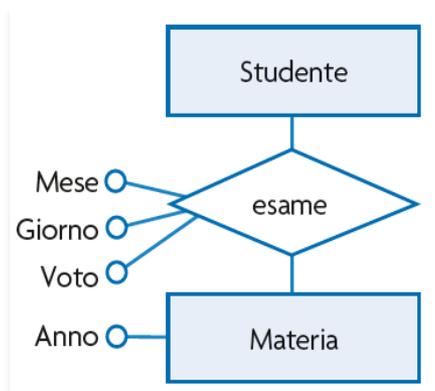
- Nomi al singolare per le entità;
- Iniziali maiuscole per nomi delle entità e degli attributi;
- Iniziali maiuscole per i nomi di attributi composti da più parole.



Modellare con attributi o entità

Quando modelliamo una realtà di interesse, dobbiamo decidere se modellare un particolare concetto come entità o come attributo.

Esempio



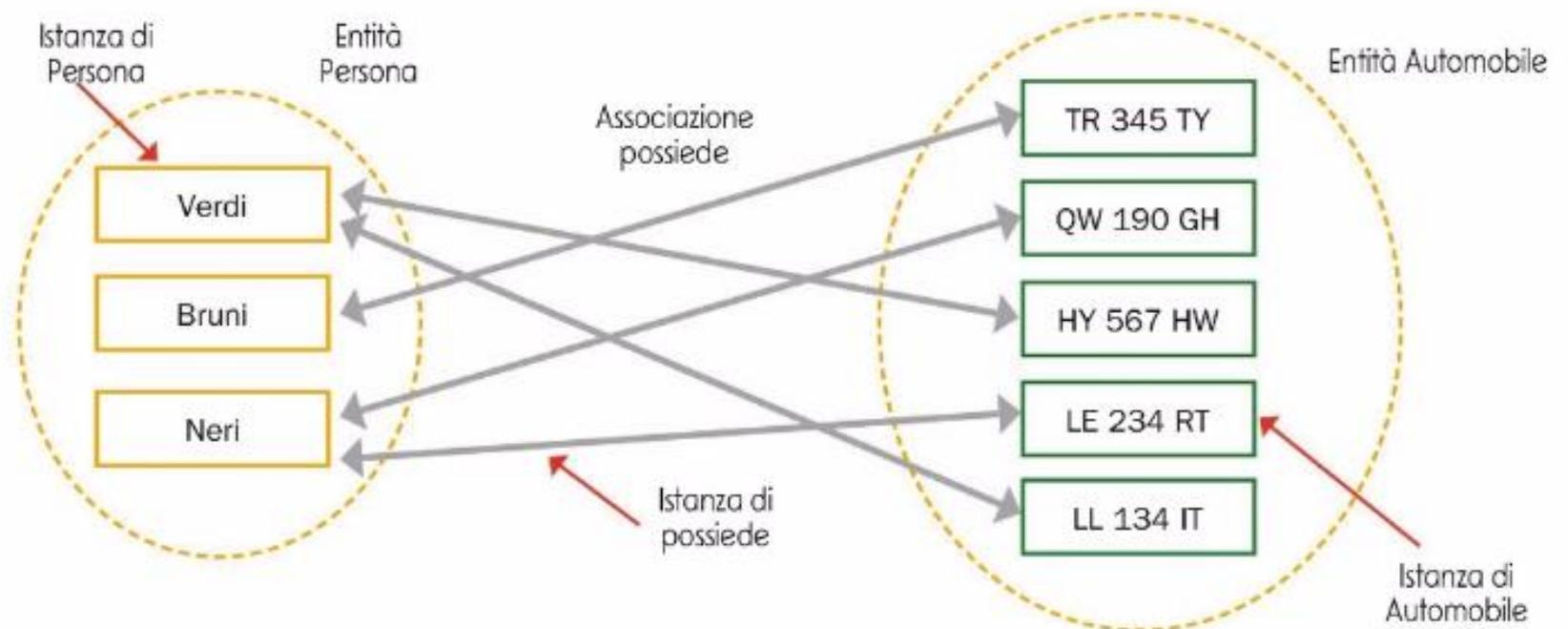
Le associazioni

L'**associazione** (in inglese *relationship*) è un legame logico tra due o più entità rilevanti nella realtà che si sta considerando.



A livello estensionale, un'associazione A tra le entità E e F è costituita da un insieme di coppie (k, h) , tali che k è un'istanza di E , e h è un'istanza di F . Ogni coppia è detta **istanza dell'associazione A** .

Esempi

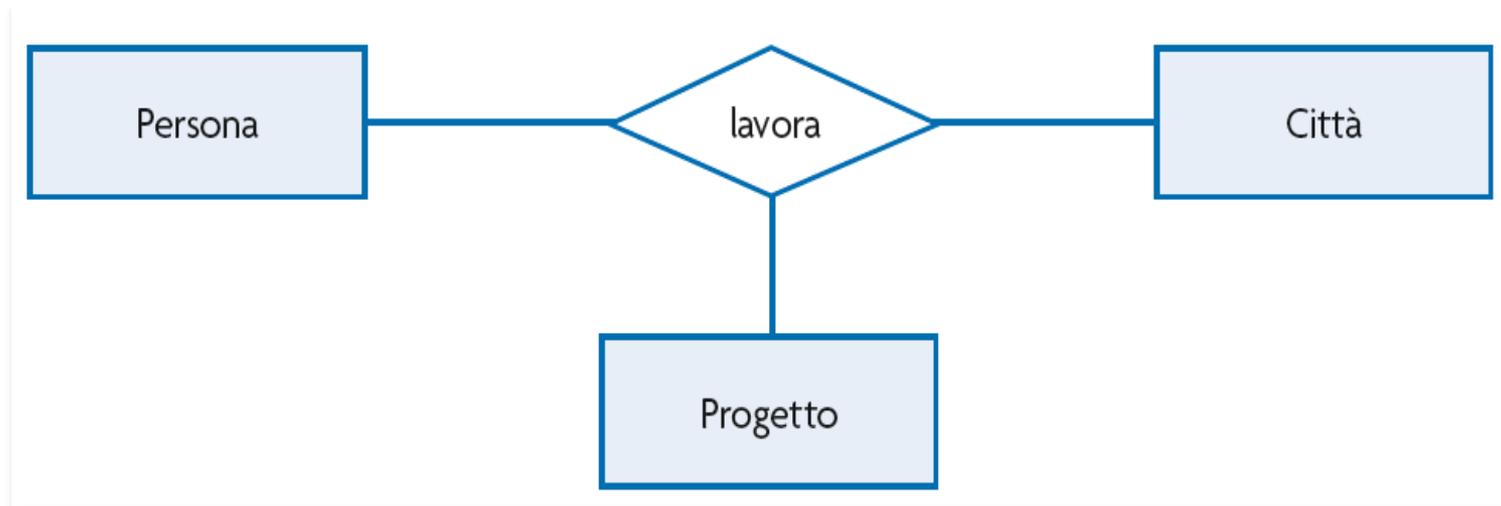


Il grado di associazione

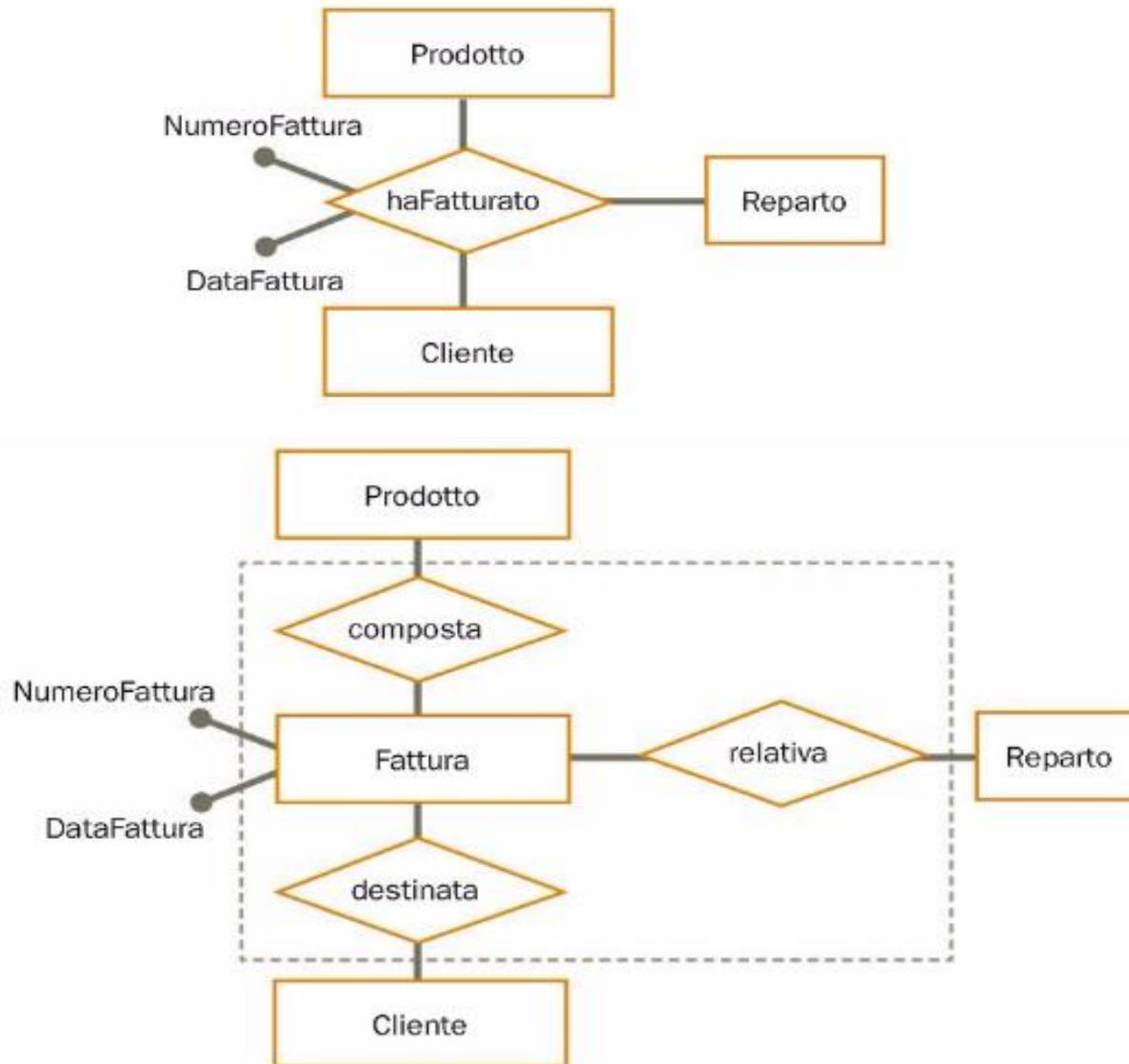
Si definisce **grado** il numero di entità che partecipano all'associazione.

Esempio

La seguente, invece, è un'associazione di grado 3, definita **ternaria** (o, genericamente, **n-aria**):



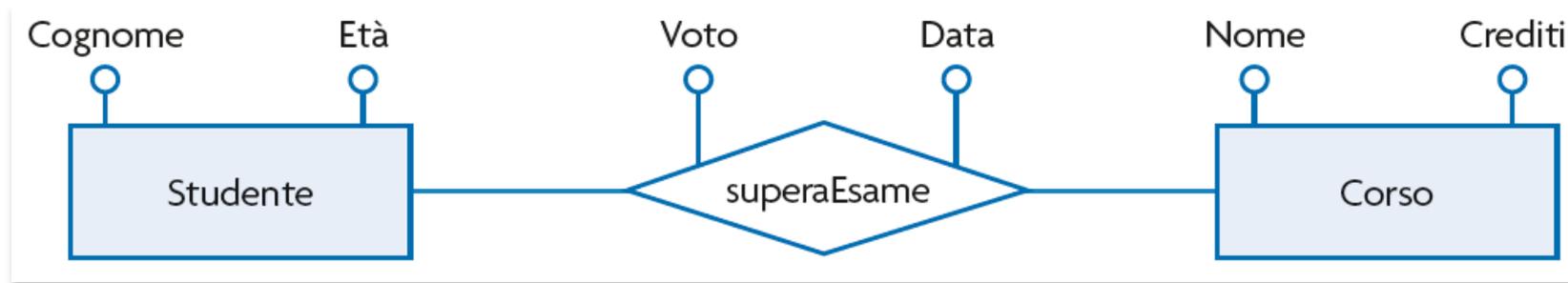
Ogni associazione multipla può essere trasformata in associazione binaria.



Gli attributi dell'associazione

Un **attributo di associazione** è una proprietà locale di un'associazione, di interesse ai fini dell'applicazione.

Esempio



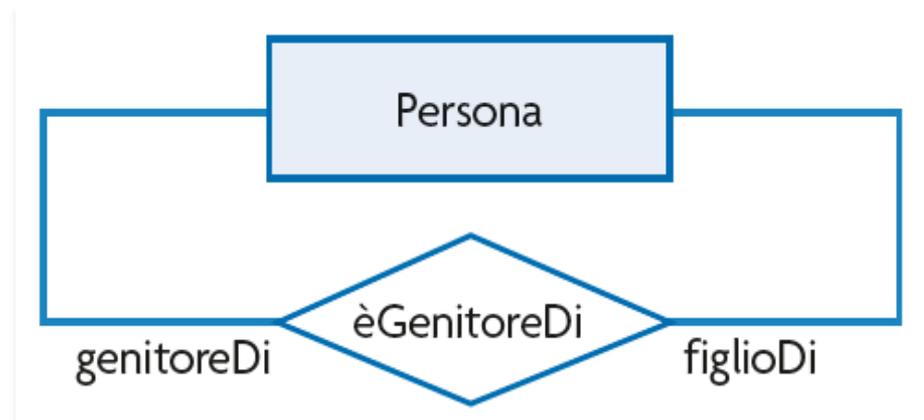
Gli attributi *Voto* e *Data* non sono proprietà né di uno studente né di un corso, **ma del legame Studente-Corso** che si crea in occasione dello svolgimento di un esame.

Le associazioni ricorsive e i ruoli

Quando un'associazione crea un legame tra una stessa entità, si ha il caso particolare di un'**associazione ricorsiva** (o **unaria**).

Esempio

Abbiamo distinto quale ramo rappresenta il genitore e quale ramo rappresenta il figlio: *genitoreDi* e *figlioDi* sono anche chiamati **ruoli** dell'entità nell'ambito dell'associazione.

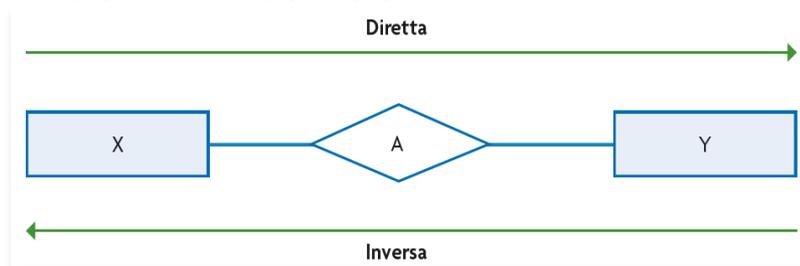


Le proprietà delle associazioni

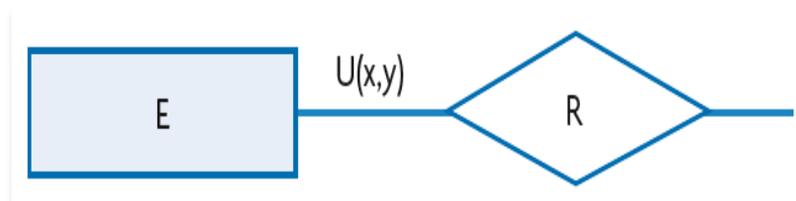
Data un'associazione A tra le entità X e Y , si attribuisce un verso da X a Y che definisce l'**associazione diretta** da X a Y .

Definiamo **inversa** l'associazione da Y a X , che indichiamo con A^{-1} .

Possiamo allora dire che la coppia (y, x) con $y \in Y$ e $x \in X$ appartiene ad A^{-1} se, e solo se, la coppia (x, y) appartiene ad A .



Un **vincolo di cardinalità** si associa a un ruolo U in una relazione R e impone un limite minimo x e un limite massimo y di istanze della relazione a cui può partecipare ogni istanza dell'entità E nel ruolo U .



Le proprietà delle associazioni

Un vincolo di cardinalità si esprime mediante una coppia (x,y) associata al ruolo U della relazione R , dove:

- x è la cardinalità minima, cioè un intero ≥ 0 ;
- y è la cardinalità massima, che è n , oppure un intero positivo $\geq x$.

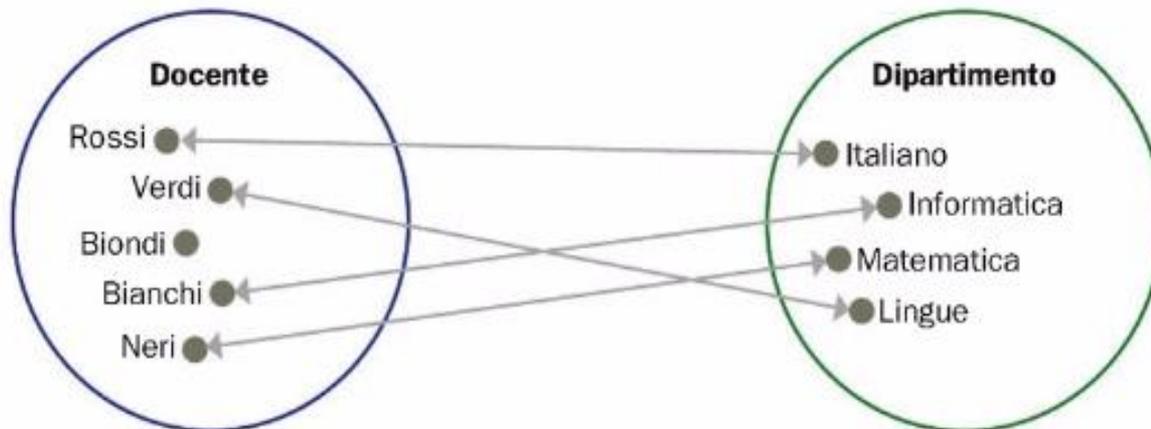
Nella maggiore parte dei casi, per esprimere la cardinalità è sufficiente utilizzare solo tre valori: 0, 1 e N .

Se la **cardinalità minima** è:

- 0, la partecipazione dell'entità E in R è **opzionale** (o **parziale**);
- 1, la partecipazione dell'entità E in R è **obbligatoria** (o **totale**).

Se la **cardinalità massima** è 1 o N , le istanze dell'entità partecipano al più 1 volta o un numero N di volte alle istanze dell'associazione rispettivamente.

Esempi



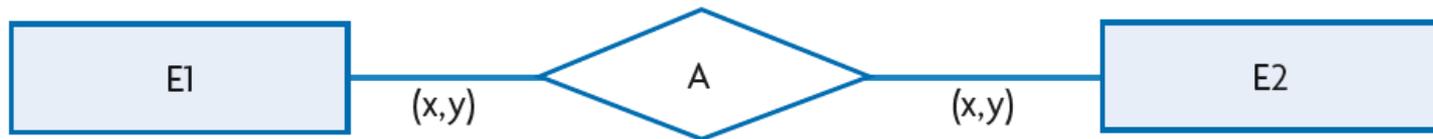
Esempi



Un'associazione si definisce **univoca** da X a Y se ogni istanza di X si può trovare in relazione con al più una istanza di Y. Se tale vincolo non esiste si parla di associazione **multipla o multi-valore**.

Tipi di associazioni

In base al valore assunto dal parametro della cardinalità massima, un'associazione binaria A tra due entità $E1$ e $E2$:

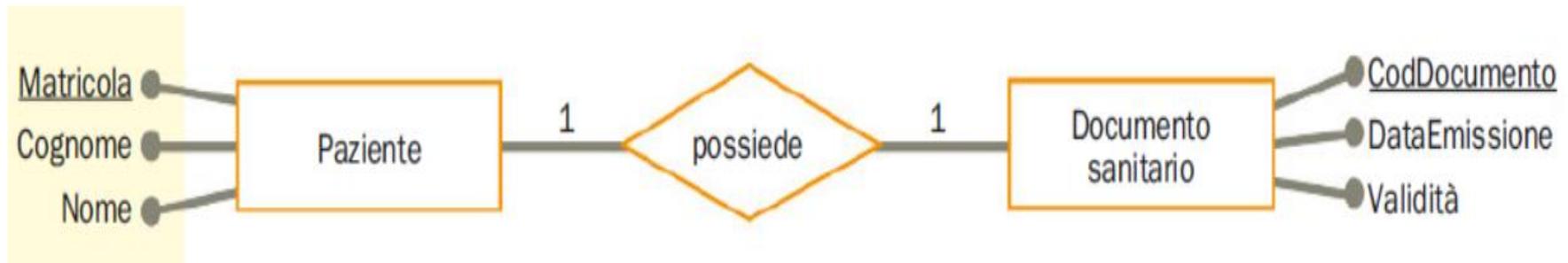


è definita come segue:

- **Uno a Uno (1, 1)** o **biunivoca**, se le cardinalità massime di entrambe le entità rispetto ad A sono 1;
- **Uno a Molti (1, N)** o **semplice**, se la cardinalità massima di $E1$ rispetto ad A è 1 e la cardinalità massima di $E2$ rispetto ad A è N ;
- **Molti a Molti (N, N)** o **complessa**, se le cardinalità massime di entrambe le entità rispetto ad A sono N .

Associazione 1:1 (uno a uno)

Il seguente schema ER descrive l'associazione *possiede* che esiste tra i pazienti e il loro documento sanitario. L'associazione è di tipo 1:1, la diretta e l'inversa sono totali.



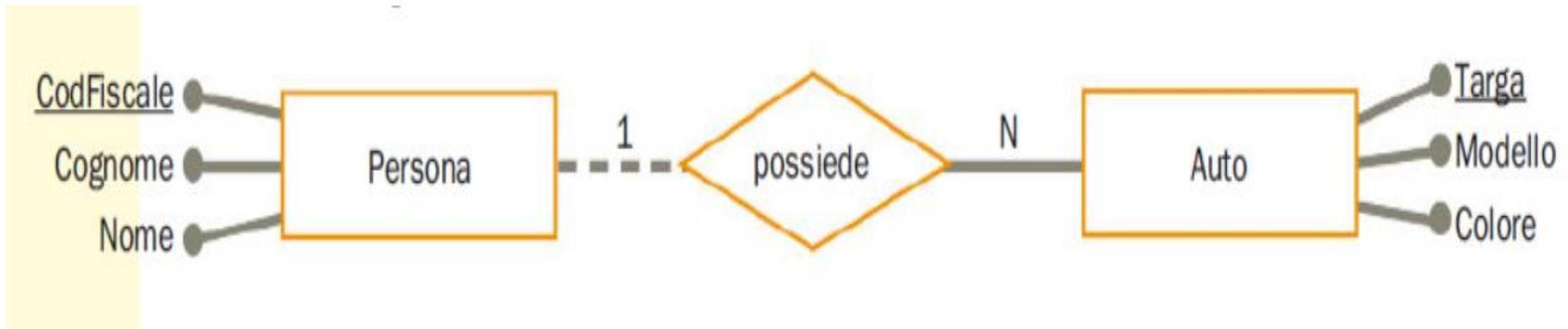
Il seguente schema ER descrive l'associazione *coordina* che esiste tra il docente e la loro classe. L'associazione è di tipo 1:1, la diretta è parziale, l'inversa è totale.



Associazione 1:N (uno a molti)

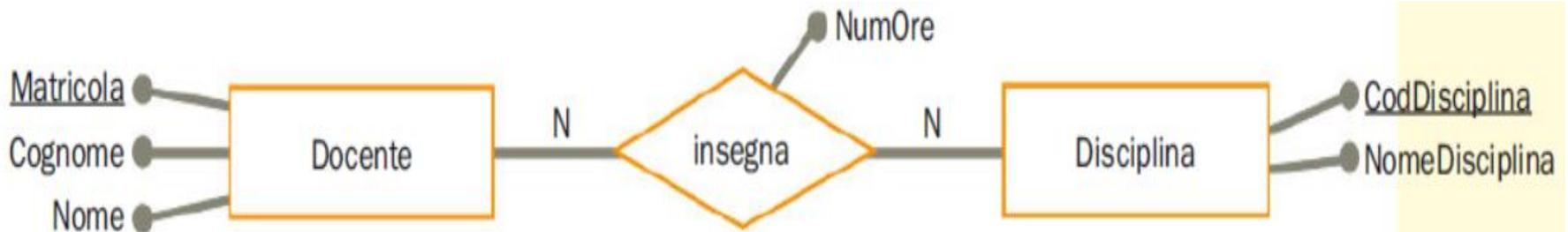
L'associazione è di tipo 1:N, poiché l'associazione diretta è multipla e l'inversa è univoca. Infatti, una persona può possedere più automobili, ma un'automobile ha un solo proprietario.

L'associazione diretta è parziale poiché non tutte le persone posseggono un'automobile, mentre l'inversa è totale poiché ogni automobile deve avere un proprietario.



Associazione N:N (molti a molti)

L'associazione è di tipo N:N poiché entrambe le associazioni sono multiple. Infatti, un docente può insegnare più discipline e una disciplina può essere insegnata da più docenti. Entrambe le associazioni sono inoltre totali, poiché ogni docente deve insegnare almeno una disciplina e ogni disciplina deve essere insegnata da almeno un docente.



Tipi di associazioni: riepilogo

Cardinalità	Partecipazione
(1,1)	Obbligatoria, una sola volta
(1,N)	Obbligatoria, almeno una volta
(0,1)	Opzionale, una sola volta
(0,N)	Opzionale, N volte

I vincoli di integrità

Possiamo definire **vincolo di integrità** una proprietà che deve essere soddisfatta dalle istanze che rappresentano informazioni corrette per l'applicazione che utilizza la base di dati. Il concetto di integrità si traduce nel tenere sotto controllo la consistenza e la precisione dei dati.

I **vincoli impliciti** sono quelli imposti dalla semantica e struttura dei dati e si dividono in:

- **vincoli di chiave primaria:** impongono che le istanze di una categoria debbano essere tutte diverse tra loro;
- **vincoli referenziali:** date due entità A e B e un'associazione tra loro, questi vincoli impongono che non si possa inserire un elemento in A (o non esista un elemento in A) non associato a un elemento in B.

I **vincoli espliciti** sono quelli che impongono delle restrizioni sul modo in cui i dati possono cambiare.

Collezione di entità e gerarchie

Quando alcune entità possono avere delle istanze in comune, oppure quando nasce l'esigenza di voler particolareggiare o generalizzare qualche entità, è utile procedere con un'organizzazione per gerarchie.

Una **gerarchia** ha una struttura ad albero: la radice rappresenta l'entità che contiene le caratteristiche comuni (**entità padre**), mentre i vari nodi (**entità figlie**) rappresentano le entità che contengono caratteristiche più specifiche.

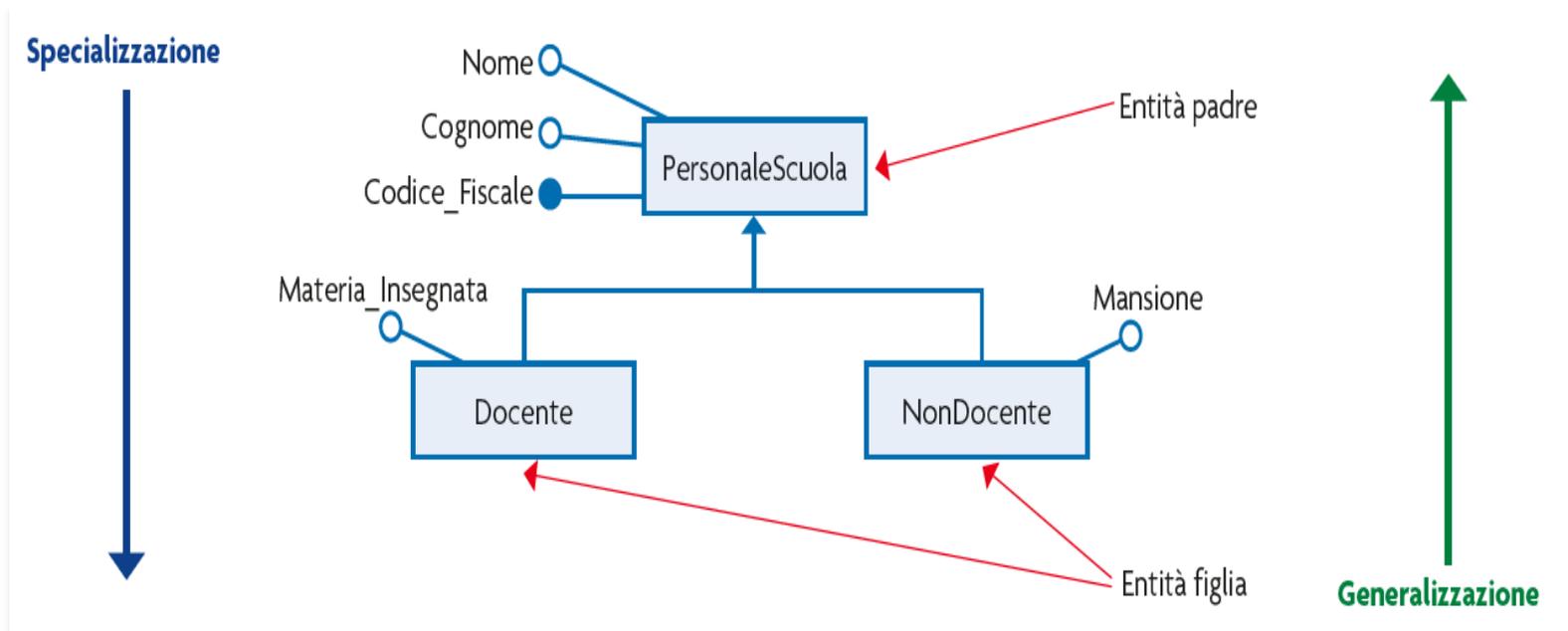
Per creare una gerarchia si attua un processo di **specializzazione**, ossia si "specializza" l'entità padre creando entità figlie sempre più specifiche.

Percorrendo la gerarchia dal basso verso l'alto si attua automaticamente un processo di **generalizzazione**.

Collezione di entità e gerarchie

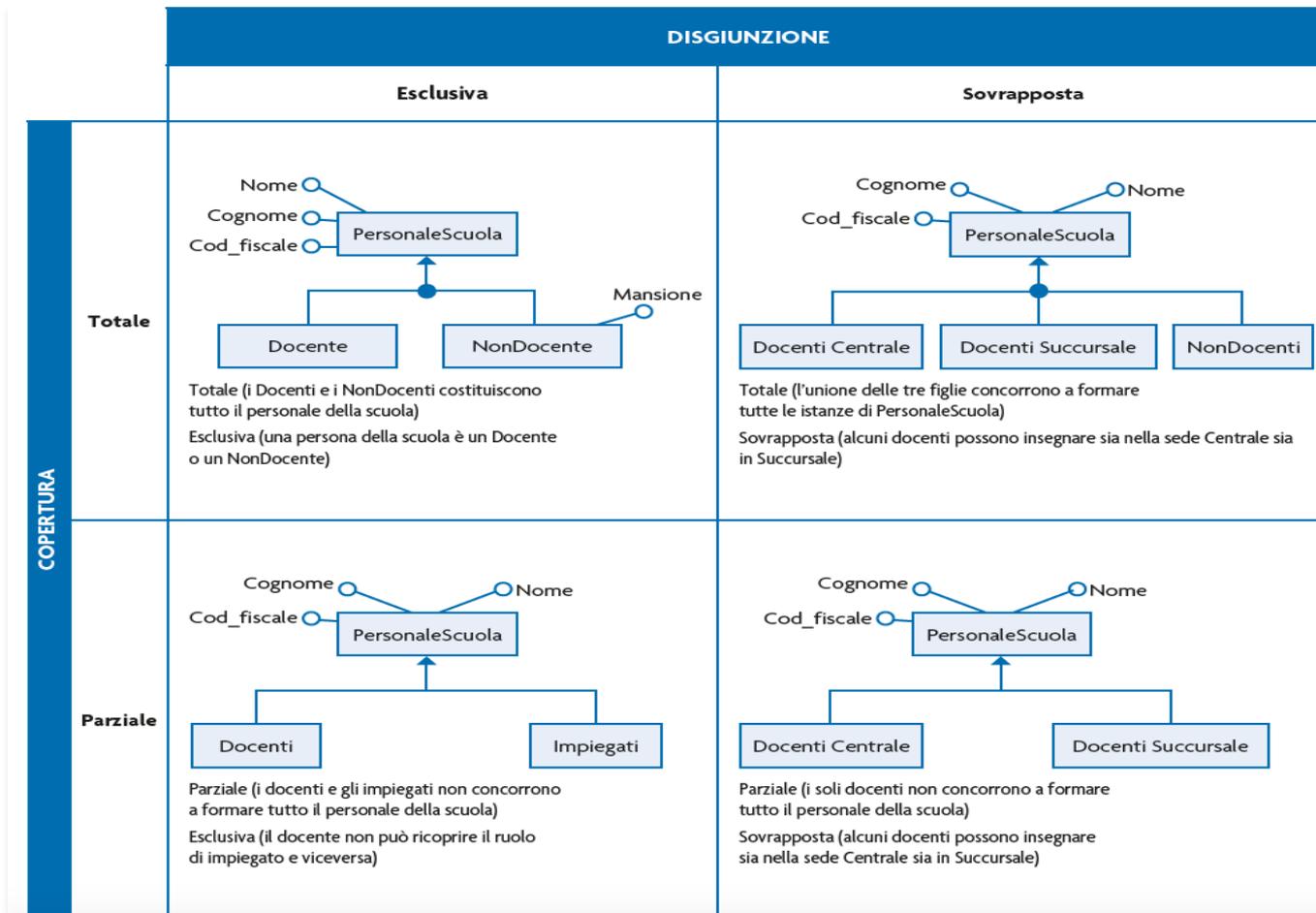
Esempio

L'entità *PersonaleScuola* può essere pensata come una generalizzazione delle entità *Docente* e *NonDocente*.



Collezione di entità e gerarchie

Le generalizzazioni possono essere classificate in base a due proprietà, **copertura** e **disgiunzione**, che sono ortogonali tra loro.



Esercizio: azienda e dipendenti

Un'azienda ha bisogno di gestire una serie di dati sui propri dipendenti, i propri reparti e i figli dei propri dipendenti (assumendo che solo uno dei genitori di ciascun figlio lavori nell'azienda).

L'azienda tiene conto dei dati dei figli solo se il genitore è dipendente dell'azienda.

Riguardo ai propri dipendenti i dati a disposizione sono: il codice fiscale, lo stipendio e il numero di telefono.

Riguardo ai reparti i dati disponibili sono il numero e il nome.

Riguardo ai figli dei dipendenti si conoscono il nome e l'età.

I dipendenti lavorano nei vari reparti. Ciascun reparto è gestito da un solo dipendente.

Analisi dei dati (entità)

Le entità del problema sono:

- **Dipendente:** serve a rappresentare i dipendenti dell'azienda;
- **Reparto:** serve a rappresentare i reparti dell'azienda;
- **Figlio:** serve a rappresentare i figli dei dipendenti dell'azienda.

Analisi dei dati (entità)

Attributi delle entità:

- **Dipendente:** Codice Fiscale, Stipendio, Telefono: tutti di tipo Stringa;
- **Reparto:** Numero Reparto, Nome Reparto: entrambi di tipo Stringa;
- **Figlio:** Nome, Età: il primo di tipo Stringa, il secondo di tipo Numerico.

Analisi delle associazioni

- L'associazione «È genitore di» lega l'entità Figlio con l'entità Dipendente ed è una associazione 1:N, poiché un ragazzo può avere un unico genitore all'interno dell'azienda e un dipendente può avere più figli. L'associazione diretta è totale (partecipazione dell'entità Figlio in «È genitore di»), poiché un figlio deve avere un genitore nell'azienda e uno solo; l'associazione inversa è parziale (partecipazione dell'entità «Dipendente» in «È genitore di»), poiché un dipendente può non avere figli all'interno dell'azienda.

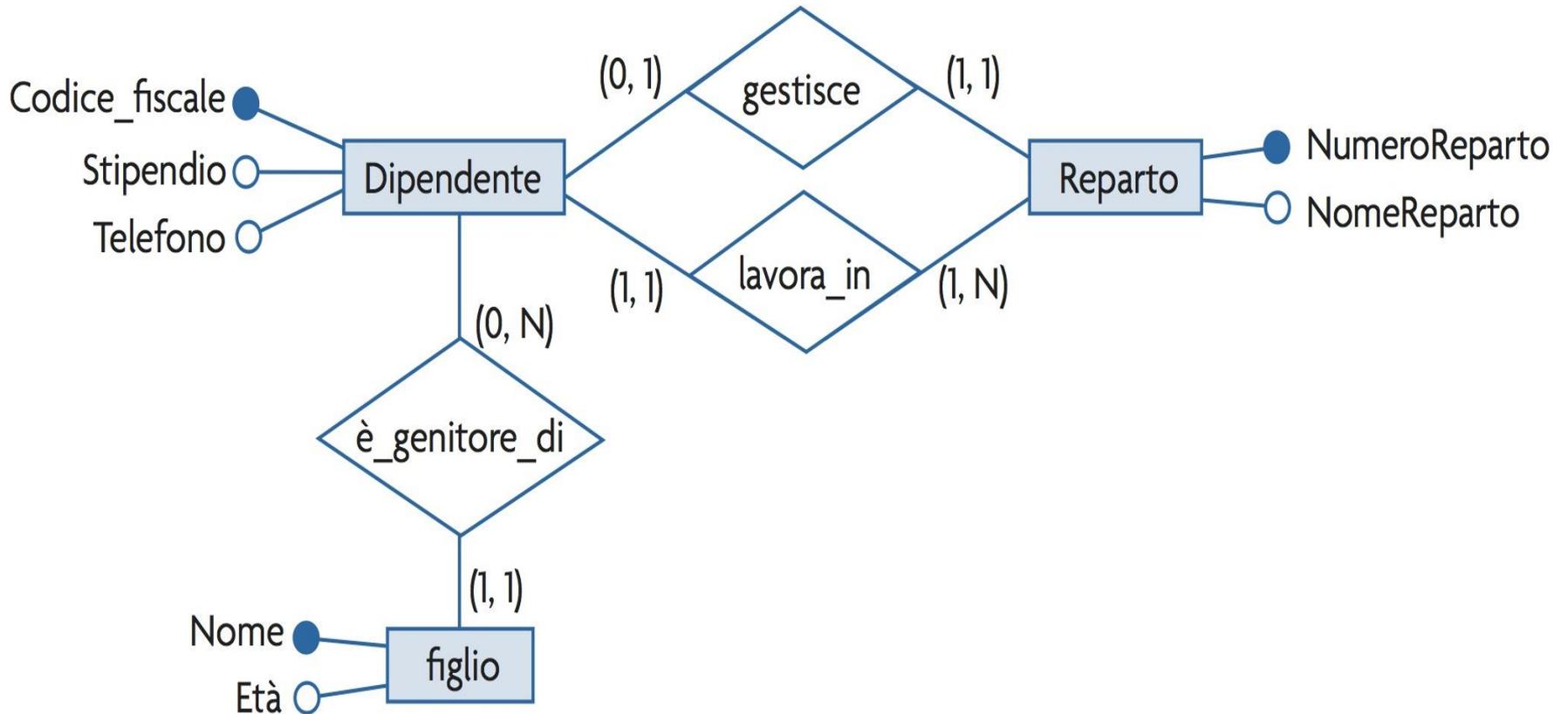
Analisi delle associazioni

- L'**associazione «Gestisce»** lega l'entità Reparto con l'entità Dipendente ed è una associazione 1:1, poiché ogni reparto deve essere gestito al più da un dipendente e un dipendente può gestire un solo reparto. L'associazione diretta è totale (partecipazione dell'entità Reparto in «Gestisce»), poiché un reparto deve essere gestito da un dipendente e da uno solo; l'associazione inversa è parziale (partecipazione dell'entità Dipendente in «Gestisce»), poiché ogni dipendente non gestisce necessariamente un reparto.

Analisi delle associazioni

- L'**associazione «Lavora in»** lega l'entità Reparto con l'entità Dipendente ed è una associazione 1:N, poiché in un reparto possono lavorare più dipendenti e ogni dipendente può lavorare in un solo reparto. L'associazione diretta è totale (partecipazione dell'entità Reparto in «Lavora in»), poiché in un reparto devono lavorare uno o più dipendenti; l'associazione inversa è totale (partecipazione dell'entità Dipendente in «Lavora in»), poiché ogni dipendente deve lavorare in un reparto e uno solo.

Schema concettuale (schema ER)



Esercizio: prestito di DVD in un negozio

La videoteca gestisce il noleggio di DVD; ogni DVD è caratterizzato da alcune informazioni generali e dai dati relativi agli attori (con i rispettivi ruoli) e ai registi, inoltre appartiene a una o più categorie (film d'azione, romantico, etc.). I DVD vengono noleggiati dai clienti, i quali devono aver richiesto e ottenuto una tessera valida fino al termine dell'anno solare di emissione e sono identificati dal numero di tale tessera. Quando un cliente richiede la tessera, non è obbligato a noleggiare subito un film, ma nel momento in cui lo fa si apre un contratto di noleggio al quale sono associati il cliente e il DVD (almeno uno) noleggiato. Vogliamo avere disponibile sia la situazione dello stato di noleggio dei DVD, sia la descrizione completa di tutti i noleggi effettuati. Nota che un cliente può restituire in date diverse i DVD noleggiati con un unico contratto.

Analisi dei dati (entità)

Le entità del problema sono:

- **Cliente:** serve a rappresentare i clienti della videoteca;
- **DVD:** serve a rappresentare le caratteristiche del DVD;
- **Attore:** serve a rappresentare gli attori che interpretano il film registrato su DVD;
- **Regista:** serve a rappresentare i registi che dirigono il film;
- **Categoria:** serve a rappresentare i tipi di film (azione, avventura, horror, etc.).

Analisi dei dati (entità)

Attributi delle entità:

- **Cliente:** Numero Tessera, Data Emissione Tessera, Nome, Cognome, Indirizzo, Città di residenza, Telefono: Numero Tessera è di tipo Numerico, gli altri di tipo Stringa;
- **DVD:** Codice DVD, Titolo, Durata: tutti di tipo Stringa;
- **Attore:** Codice Attore, Nome, Cognome: tutti di tipo Stringa;
- **Regista:** Codice Regista, Nome, Cognome: tutti di tipo Stringa;
- **Categoria:** Codice Categoria, Nome: entrambi di tipo Stringa.

Analisi delle associazioni

- L'**associazione «Noleggia»** lega il cliente ai DVD noleggiati ed è una associazione N:N, in quanto un cliente può noleggiare più DVD e un DVD può essere noleggiato da più clienti. L'associazione diretta è parziale, poiché un cliente può non aver noleggiato alcun DVD; l'associazione inversa è totale, poiché il noleggio di un DVD deve essere sempre riferito a un cliente.
- L'**associazione «Noleggia»** ha 3 attributi: Data Noleggio, Data Scadenza e Tariffa Giornaliera, cioè le date di noleggio e restituzione del DVD e la tariffa applicata per il servizio.

Analisi delle associazioni

- L'**associazione «Interpretato»** lega l'entità DVD e l'entità Attore. È una associazione N:N, in quanto un film registrato su DVD può essere interpretato da più attori e più attori possono interpretare lo stesso film. L'associazione diretta è totale, poiché un film registrato su DVD deve essere interpretato da almeno un attore; l'associazione inversa è parziale, poiché un attore può non recitare in alcun film.
- L'**associazione «Interpretato»** possiede un attributo: Ruolo, che identifica il ruolo di un attore nell'ambito di un film.

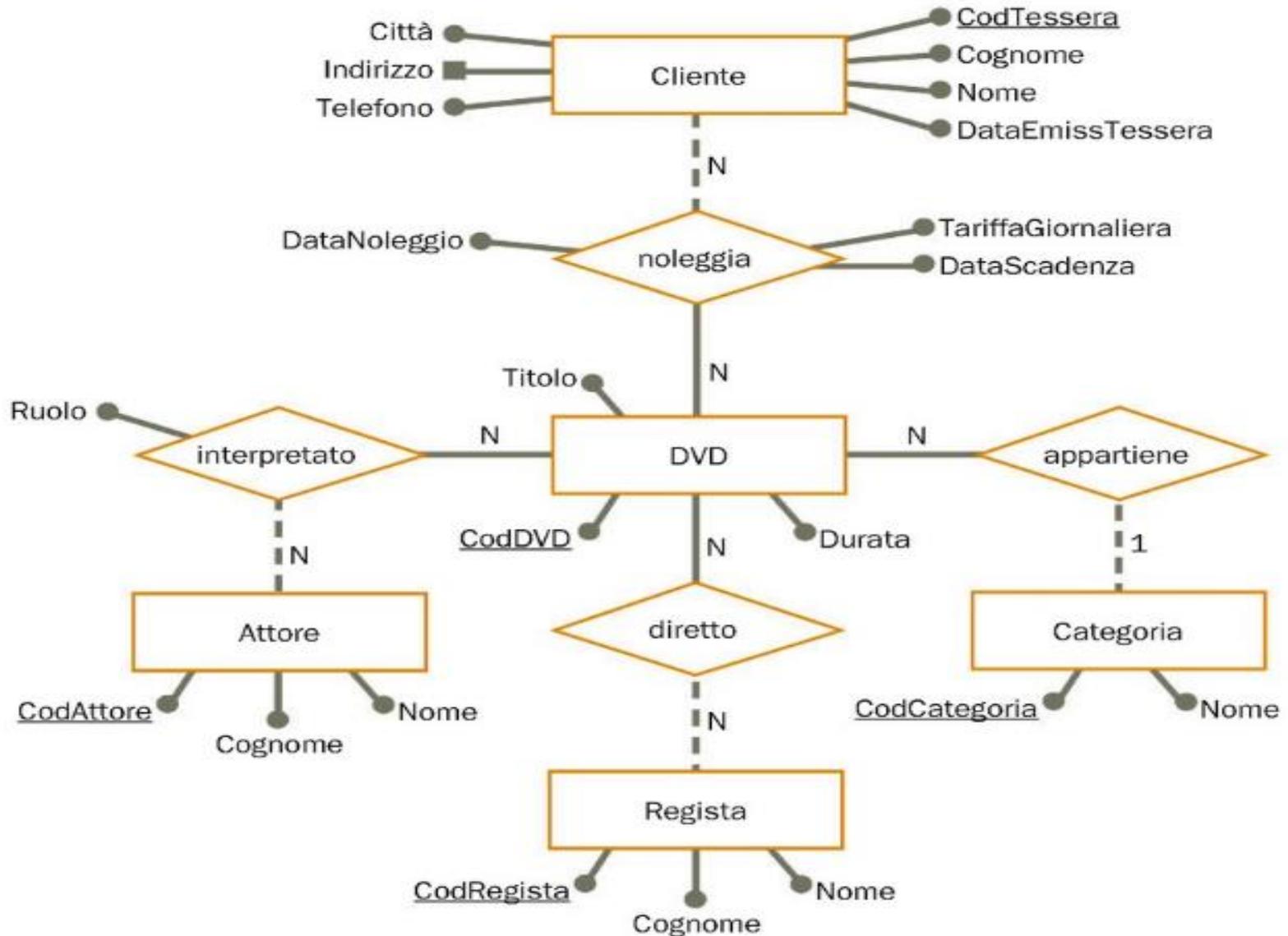
Analisi delle associazioni

- L'**associazione «Appartiene»** lega l'entità DVD e l'entità Categoria. È una associazione 1:N, in quanto un film registrato su DVD appartiene ad una categoria e ad una categoria possono appartenere più film. L'associazione diretta è totale, poiché un film registrato su DVD deve appartenere ad almeno una categoria; l'associazione inversa è parziale, poiché ad una categoria può anche non appartenere alcun DVD della videoteca.

Analisi delle associazioni

- L'**associazione «Diretto»** lega l'entità DVD e l'entità Regista. È una associazione N:N, in quanto un film registrato su DVD può essere diretto da più registi e un regista può dirigere più film. L'associazione diretta è totale, poiché un film registrato su DVD deve essere diretto da almeno un regista; l'associazione inversa è parziale, poiché un regista potrebbe anche non dirigere alcun film.

Schema concettuale (schema ER)



Esercizio: affitti ed appartamenti

Un'agenzia immobiliare gestisce affitti di appartamenti. Gli appartamenti in affitto sono collocati in condomini, ognuno di essi è gestito da un amministratore unico di condominio.

I clienti dell'agenzia sono identificati da: Codice Fiscale, nome, cognome, data di nascita e comune di residenza.

Per ogni affitto si conoscono la data di inizio e di fine del contratto, il canone di affitto nonché il proprietario dell'appartamento.

L'agenzia immobiliare deve poter conoscere quali sono gli appartamenti affittati, e a quali condomini appartengono, quali clienti hanno contratti di affitto in scadenza e quali sono i canoni di affitto pagati da ogni cliente. Inoltre l'agenzia deve poter contattare, telefonicamente/per email ogni amministratore, ogni cliente e ogni proprietario di appartamento.

Analisi dei dati (entità)

Le entità del problema sono:

- **Cliente:** serve a rappresentare i clienti dell'agenzia, cioè coloro che affittano l'appartamento o gli appartamenti;
- **Proprietario:** serve a rappresentare i proprietari degli appartamenti che vengono affittati;
- **Condominio:** serve a rappresentare i condomini di cui fanno parte gli appartamenti in affitto;
- **Appartamento:** serve a rappresentare gli appartamenti gestiti dall'agenzia per l'affitto;
- **Amministratore:** serve a rappresentare gli amministratori dei condomini in cui si trovano gli appartamenti dati in affitto.

Analisi dei dati (entità)

Attributi delle entità:

- **Cliente:** Codice Fiscale, Nome, Cognome, Email, Telefono, Indirizzo (attributo composto): tutti di tipo Stringa;
- **Proprietario:** Codice Proprietario, Nome, Cognome, Email, Telefono: tutti di tipo Stringa;
- **Condominio:** Codice Condominio, Indirizzo (attributo composto): tutti di tipo Stringa;
- **Appartamento:** Codice Appartamento, Numero Vani, Metratura, Piano: tutti di tipo Stringa;
- **Amministratore:** Codice Amministratore, Nome, Cognome, Email, Telefono: tutti di tipo Stringa.

Analisi delle associazioni

- L'**associazione «Affitta»** lega il cliente agli appartamenti affittati ed è una associazione N:N, in quanto un cliente può affittare più appartamenti e un appartamento può essere affittato da più clienti. L'associazione diretta è parziale, poiché un cliente può non aver affittato alcun appartamento; l'associazione inversa è totale, poiché l'affitto di un appartamento deve essere sempre riferito a un cliente.
- L'**associazione «Affitta»** ha 3 attributi: Data Inizio, Data Fine e Canone che stabiliscono la durata del contratto e il suo canone.

Analisi delle associazioni

- L'**associazione «Possiede»** lega il proprietario di un appartamento agli appartamenti di sua proprietà. È una associazione N:N, in quanto un proprietario può possedere più appartamenti e un appartamento può essere posseduto da più proprietari. Le associazioni diretta e inversa sono entrambi totali, poiché un proprietario deve possedere almeno un appartamento e un appartamento deve essere posseduto da almeno un proprietario.

Analisi delle associazioni

L'**associazione «Possiede»** relativa ai condomini definisce quali appartamenti appartengono ad un condominio. È di tipo 1:N, perché un condominio può avere più appartamenti, ma un appartamento può appartenere ad un solo condominio. Le associazioni diretta e inversa sono entrambe totali perché un condominio deve possedere necessariamente almeno un appartamento e ciascun appartamento deve essere posseduto da un condominio.

Analisi delle associazioni

- L'**associazione «Amministra»** lega ogni amministratore al condominio che amministra: è di tipo 1:N perché un amministratore può gestire più condomini, ma un condominio può essere gestito da un solo amministratore. L'associazione diretta è parziale perché non tutti gli amministratori devono gestire un condominio, quella inversa è totale perché ogni condominio deve possedere un amministratore.
- L'**associazione «Amministra»** ha due attributi: Data Inizio e Tariffa che stabiliscono la data di inizio del contratto dell'amministratore e la sua paga.

Schema concettuale (schema ER)

