

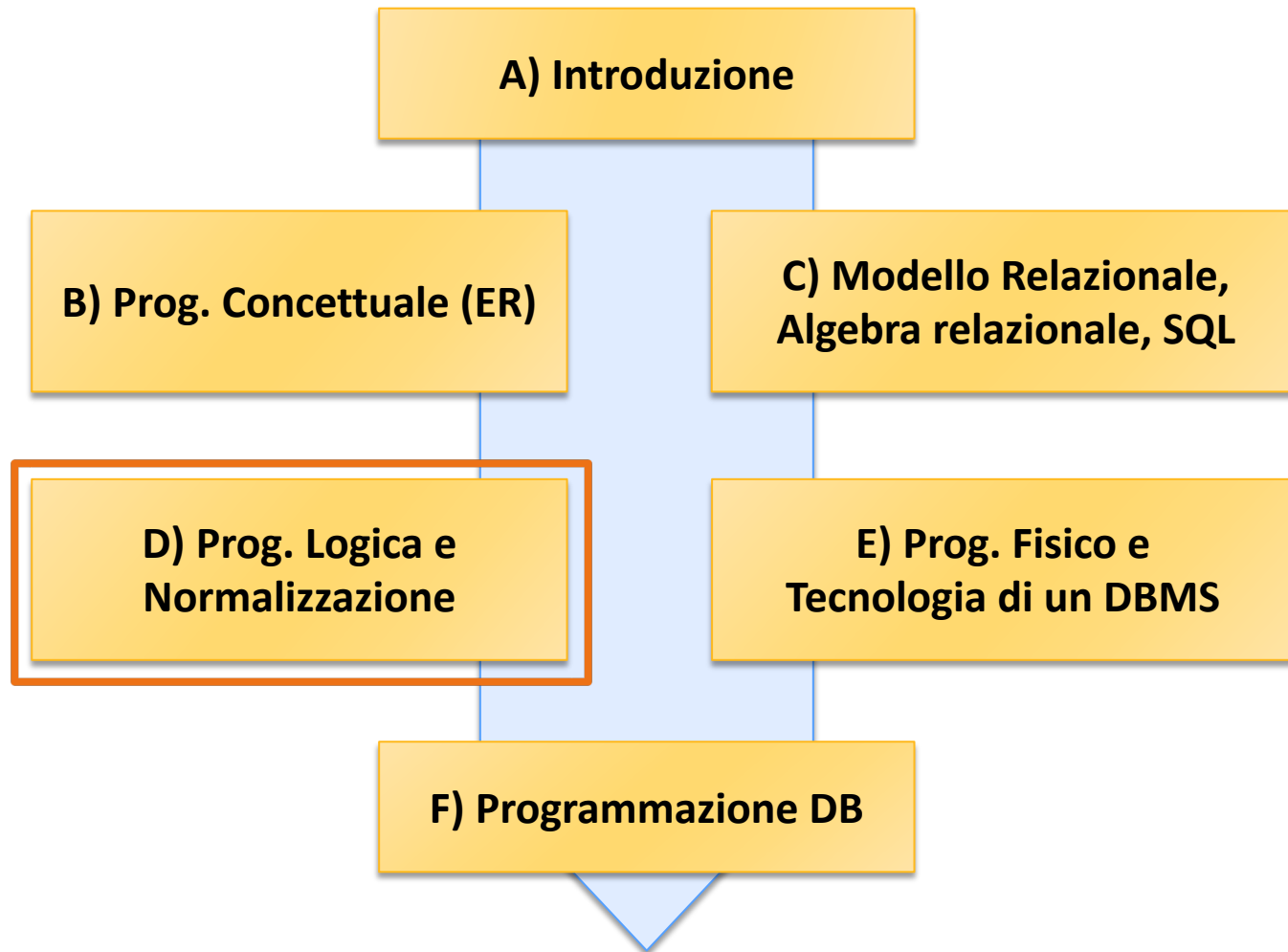


Basi di Dati



Approfondimento su Dati Derivati

Basi di Dati – Dove ci troviamo?



Carico di lavoro

- Il **carico di lavoro** sul DB è rappresentato
 - dalla **dimensione dei dati**
 - dalle **operazioni** più significative che si stima saranno eseguite sul DB
- **Regola 20-80**: il 20% delle operazioni produce l'80% del carico

Carico di lavoro

- **Volume dei dati**

- Numero medio di istanze di ogni entità e associazione
- Cardinalità e dimensioni di ciascun attributo
- Percentuali di copertura di gerarchie

Tabella dei volumi:

Concetto	Tipo	Volume dei dati
----------	------	-----------------

Nome

E (entità) R (associazione)
A (attributo)

Carico di lavoro

- **Descrizione delle operazioni**

- **Tipo di operazione:** Interattiva o batch
- **Frequenza:** numero medio di esecuzioni in un certo periodo di tempo
- **Schema di operazione:** frammento di E-R interessato dall'operazione con “cammino logico” da percorrere per accedere alle informazioni di interesse

Tabella delle operazioni:

Operazione	Tipo (I o B)	Frequenza
------------	--------------	-----------

Carico di lavoro

- Con le informazioni viste è possibile fare una stima del **costo** di un'operazione contando il **numero di accessi** alle istanze e associazioni necessario per eseguire l'operazione

Tabella degli accessi:

Concetto	Accessi	Tipo
----------	---------	------

- Le operazioni di **scrittura (S)** sono generalmente più onerose di quelle in **lettura (L)**: il peso degli accessi in scrittura è doppio di quello in lettura.

Dati derivati

- **DEFINIZIONE:** un **dato derivato** è un dato che può essere ottenuto attraverso una serie di operazioni da altri dati
 - Sulla base delle operazioni e delle loro frequenze è possibile valutare se è conveniente o meno mantenere nello schema attributi derivati
- **Vantaggi:** a tempo di accesso non è richiesta alcuna operazione per ricavare il valore dell'attributo
- **Svantaggi:** occorre eseguire operazioni di aggiornamento per mantenere la consistenza dei dati; si spreca memoria

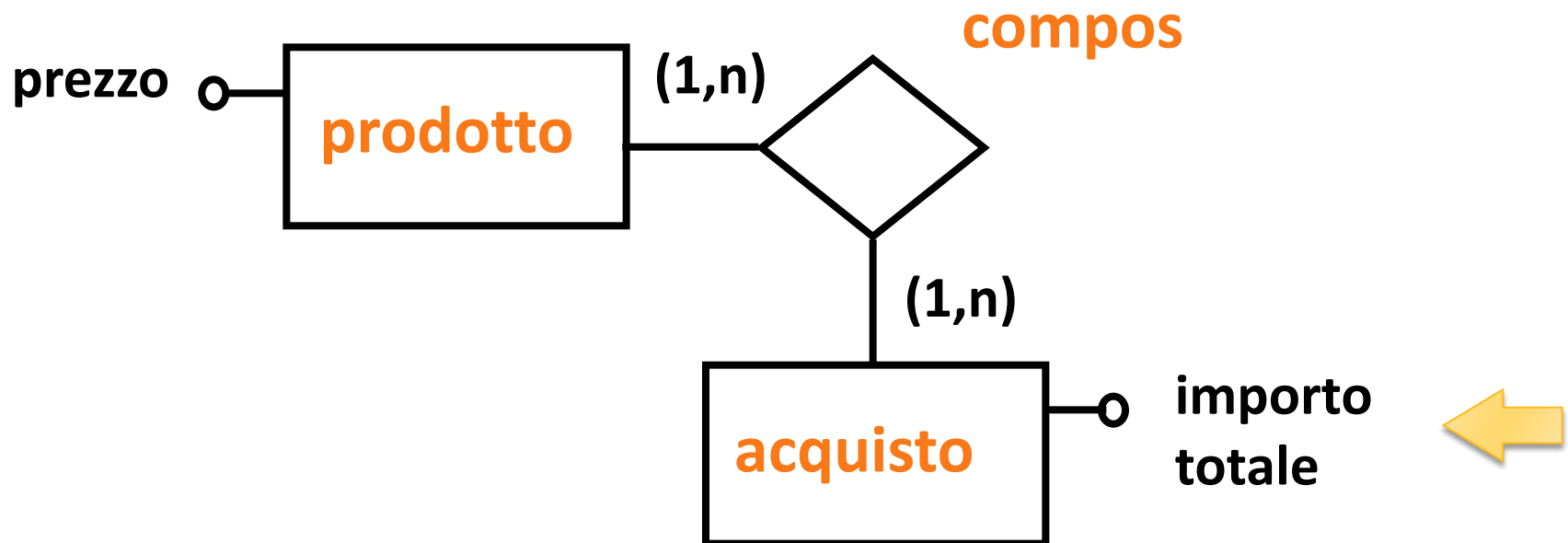
Dati derivati

- **Attributi derivabili da altri attributi della stessa entità o associazione**



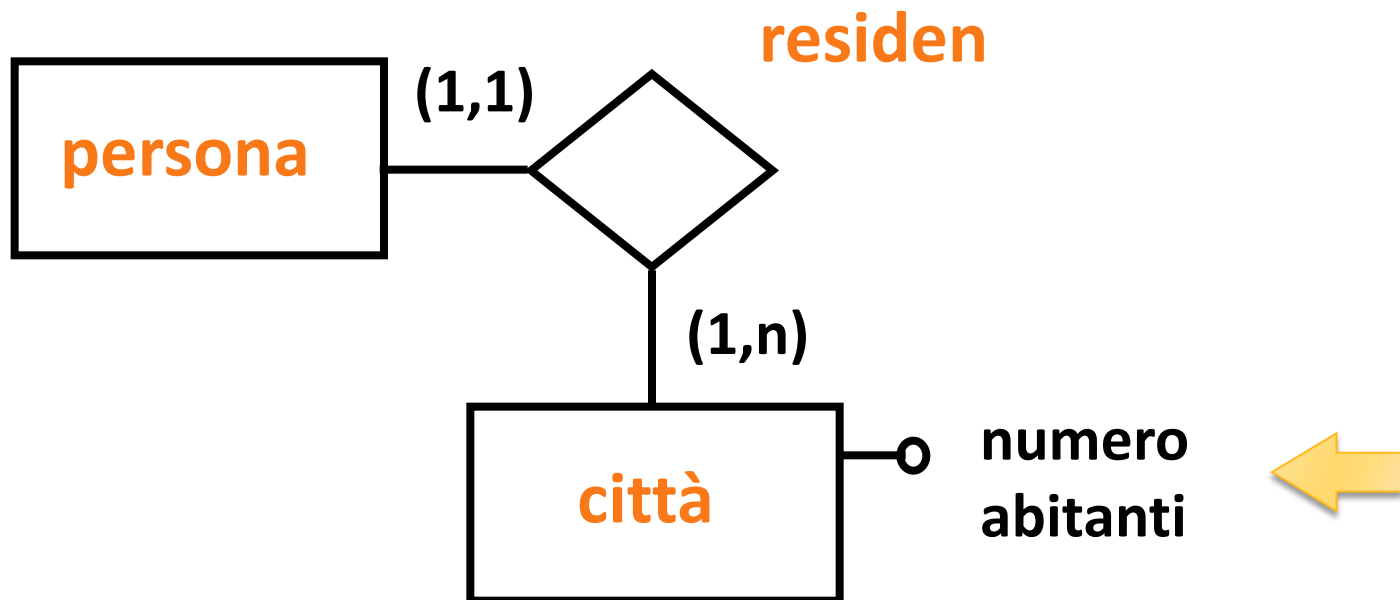
Dati derivati

- **Attributi derivabili da attributi di altre entità o associazioni**



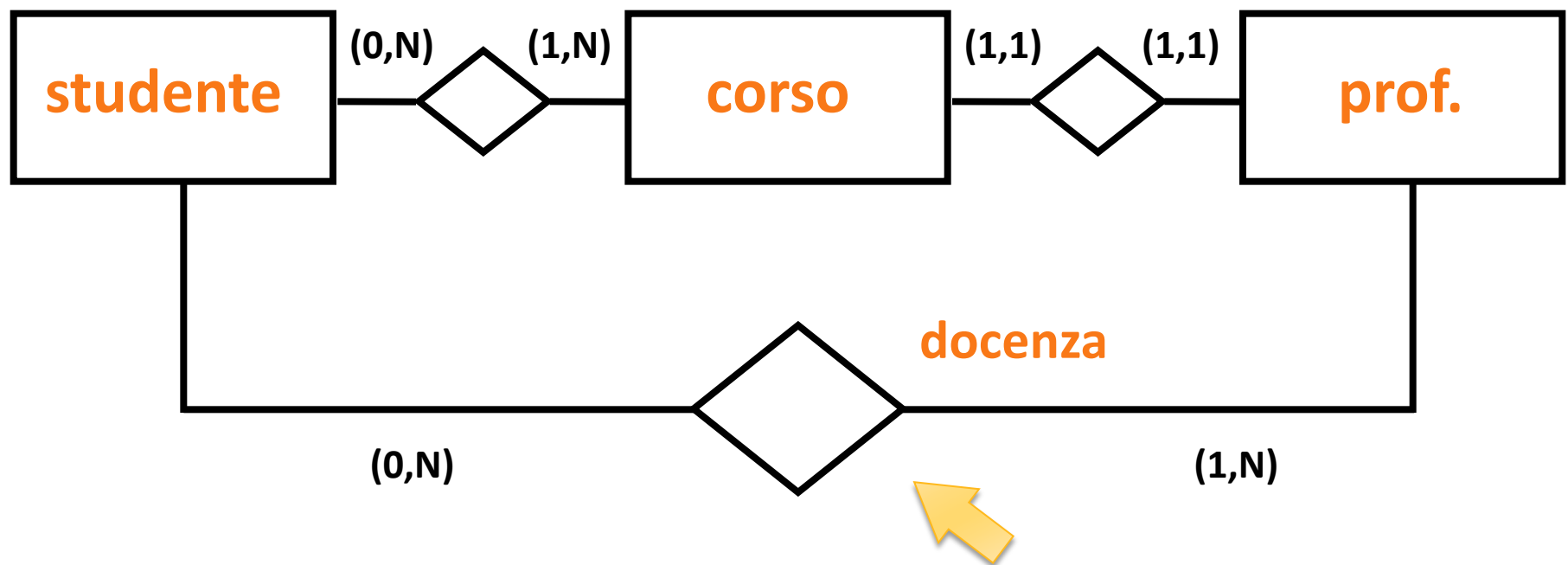
Dati derivati

- **Attributi derivabili da operazioni di conteggio di istanze**



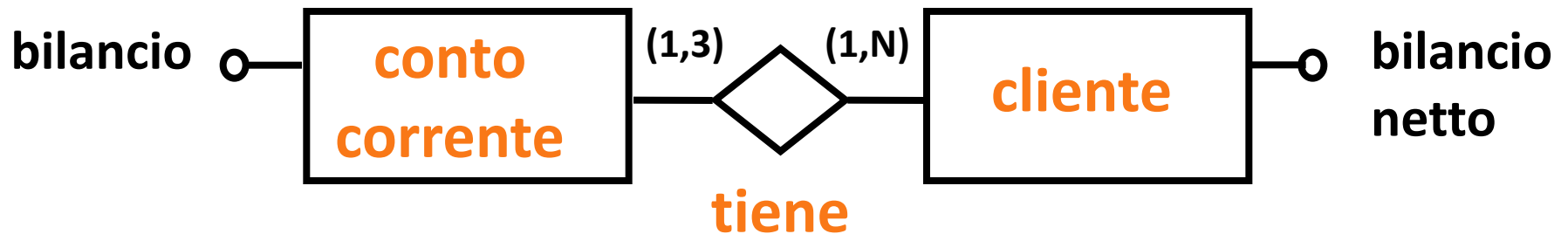
Dati derivati

- **Attributi derivabili dalla composizione di altre associazioni**



Esempio

- Dato derivato: **Bilancio Netto**



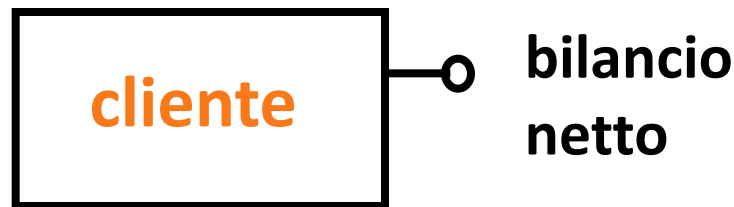
- Operazione 1:
lettura del bilancio netto di un cliente
- Operazione 2:
deposito su un conto corrente

Esempio

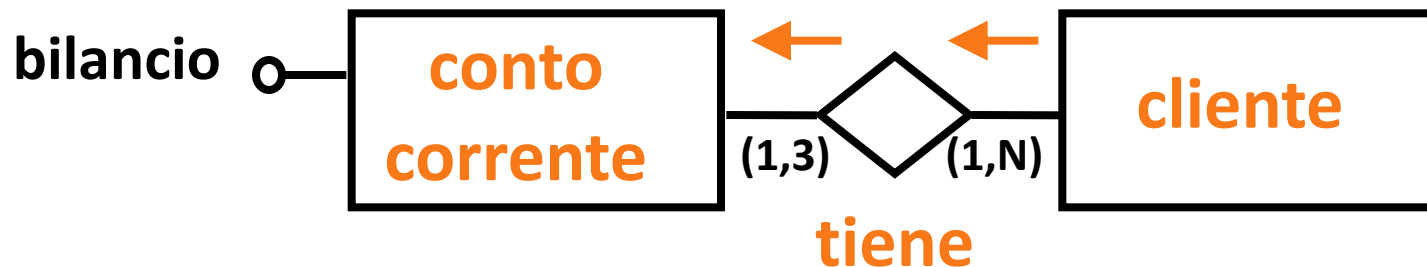
– Operazione 1:

lettura del bilancio netto di un cliente

Con il dato derivato:



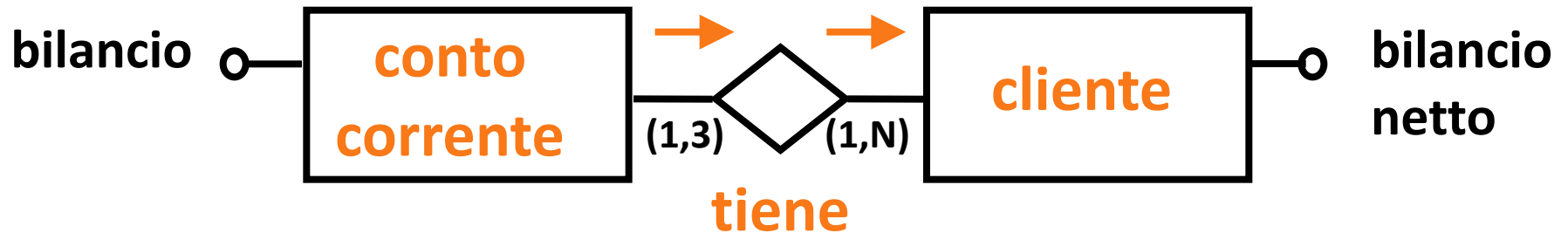
Senza il dato derivato:



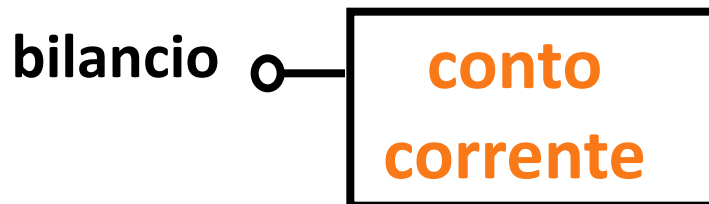
Esempio

- Operazione 2:
deposito su un conto corrente

Con il dato derivato:



Senza il dato derivato:



Esempio

– Tabella dei volumi:

Concetto	Tipo	Volume dati
Cliente	E	15000
ContoCorrente	E	20000
Tiene	R	30000

– Tabella delle operazioni:

Operazione	Tipo	Frequenza
Oper. 1	I	3000/giorno
Oper. 2	I	1000/giorno

Esempio

Con il dato derivato:

- **Occupazione di memoria:**
Se ogni valore di “Bilancio Netto” richiede
6 byte di memoria

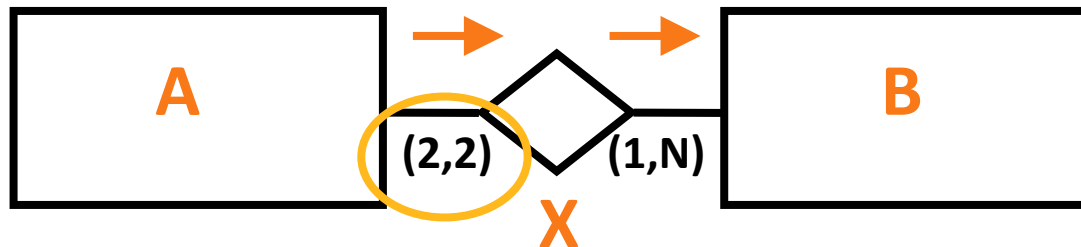
 **Memoria richiesta: 90 KB**

Tipi di operazioni: lettura e scrittura

- ▶ In caso di **lettura** di una istanza
 - ▶ una operazione di lettura;
- ▶ In caso di **aggiornamento** di una istanza esistente
 - ▶ necessaria una lettura (per leggere la tupla passata) e una scrittura (per aggiornarla);
- ▶ In caso di **inserimento** di una nuova istanza o **cancellazione** di una istanza esistente:
 - ▶ necessaria una scrittura.
- ▶ Nota: attenzione, in caso di scritture su entità, alla necessità di effettuare scritture anche sulle relative associazioni!

Calcolo di cardinalità

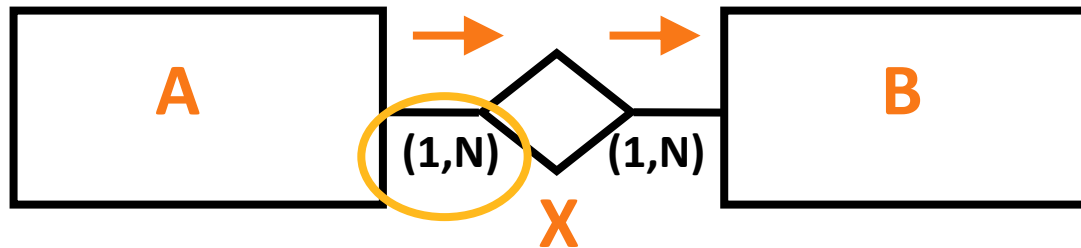
- ▶ In generale, per spostarsi da un'entità A ad una B, è necessario calcolare la cardinalità di accessi all'associazione X (e, di conseguenza, all'entità B)



- ▶ Per farlo, dobbiamo prima di tutto guardare la cardinalità dell'associazione dal lato dell'entità A
 - ▶ Se la **cardinalità minima è uguale alla massima**, questa coincide con la cardinalità di accessi a X e a B
 - ▶ Nell'esempio qui sopra avremo, per un'istanza di A, 2 istanze di X e 2 istanze di B collegate

Calcolo di cardinalità

- ▶ In tutti gli altri casi, lo schema E/R non ci aiuta a stimare la cardinalità e **dovremo fare il calcolo** sfruttando la tabella dei volumi.



Concetto	Tipo	Volume
A	E	10000
X	R	30000

$$\begin{aligned}\text{Card}(A \rightarrow X) &= \text{Vol}(X) / \text{Vol}(A) = \\ &= 30000 / 10000 = \\ &= 3\end{aligned}$$

- ▶ Nell'esempio qui sopra avremo, per un'istanza di A, mediamente 3 istanze di X (e, di conseguenza, 3 istanze di B collegate)

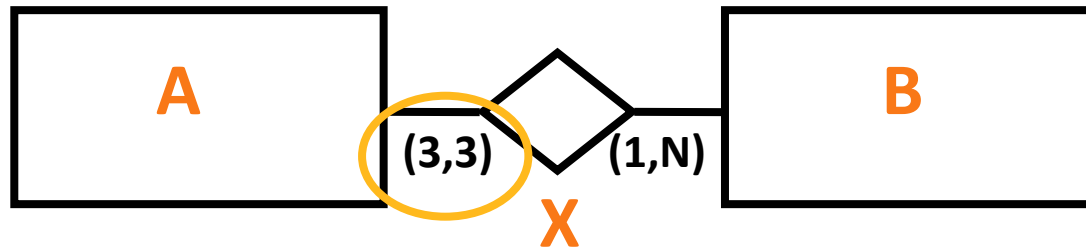
Calcolo di cardinalità

► Attenzione!

- Il numero di accessi all'entità “destinazione” B **coincide** con il numero di accessi all'associazione X attraversata
- Il numero di accessi all'entità “destinazione” B **non dipende** dal volume dei dati presenti in B! (questi saranno invece utili per spostarsi da B verso A, ad esempio)

Completamento tabella dei volumi

- ▶ A volte, la tabella dei volumi non contiene in esplicito le informazioni per tutte le entità e associazioni, perché queste possono essere desunte dallo schema. Per poter effettuare il calcolo visto, è quindi utile prima di tutto **completarla**:



Concetto	Tipo	Volume
A	E	10000
B	E	20000



Concetto	Tipo	Volume
A	E	10000
B	E	20000
X	R	30000

Tornando all'esempio...

Con il dato derivato:

Operazione 1

1 accesso in lettura

1* 3000 = **3000** / giorno

Operazione 2

4 accessi in lettura

2.5 accessi in scrittura

9*1000 = **9000** / giorno

Concetto	Acc.	Tipo
Cliente	1	L
ContoCorrente	1	L
ContoCorrente	1	S
Tiene	1.5	L
Cliente	1.5	L
Cliente	1.5	S

Esempio

Senza il dato derivato:

Operazione 1

5 accessi in lettura

$5 * 3000 = \mathbf{15000}$ / giorno

Operazione 2

1 accesso in lettura

1 accesso in scrittura

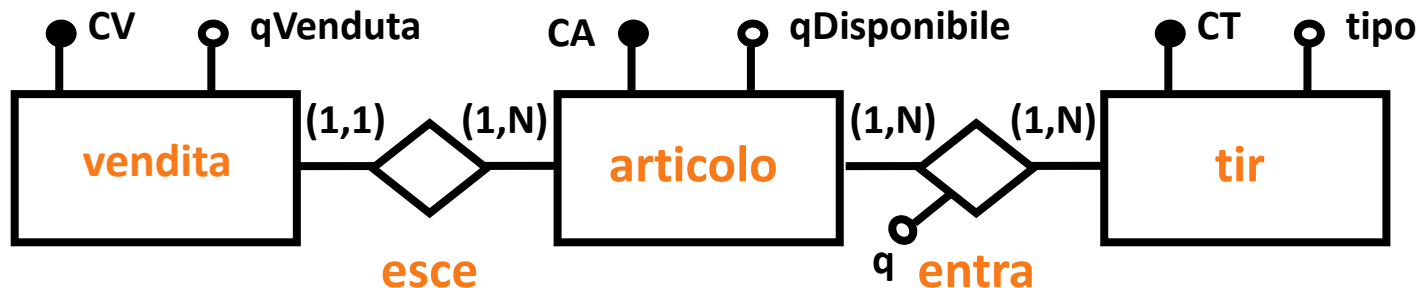
$3 * 1000 = \mathbf{3000}$ / giorno

Concetto	Acc.	Tipo
Cliente	1	L
Tiene	2	L
ContoCorrente	2	L
ContoCorrente	1	L
ContoCorrente	1	S

CONCLUSIONE: conviene tenere il dato derivato

Esercizio 1

- **Dato derivato: quantità disponibile (qDisponibile)**, calcolata come la differenza tra la somma delle quantità che entrano tramite tir e la somma delle quantità che escono tramite vendita



- **Operazione 1:**
inserimento di una nuova vendita (si suppone noto e valido il codice dell'articolo venduto)
- **Operazione 2:**
visualizzazione di tutti i dati di un articolo, inclusa la qDisponibile

Esercizio 1

– Tabella dei volumi:

Concetto	Tipo	Volume dati
Articolo	E	200
Vendita	E	1000000
Entra	R	800

– Tabella delle operazioni:

Operazione	Tipo	Frequenza
Oper. 1	I	1000/giorno
Oper. 2	I	3/giorno

Esercizio 1

Con il dato derivato:

TOT: 7003 / giorno

	Concetto	Acc.	Tipo
Operazione 1 1 accesso in lettura 3 accesso in scrittura 7*1000 = 7000 / giorno	Vendita	1	S
	Esce	1	S
	Articolo	1	L
	Articolo	1	S
Operazione 2 1 accesso in lettura 1*3 = 3 / giorno	Articolo	1	L

Esercizio 1

Senza il dato derivato:

TOT: 34015 / giorno

Operazione 1 2 accessi in scrittura 4*1000 = 4000 / giorno	Concetto	Acc.	Tipo
	Vendita	1	S
Operazione 2 10005 accessi in lettura 10005*3 = 30015 / giorno	Esce	1	S
	Articolo	1	L
	Esce	5000	L
	Vendita	5000	L
	Entra	4	L

Esercizio 1

Con il dato derivato:

TOT: 7003 / giorno

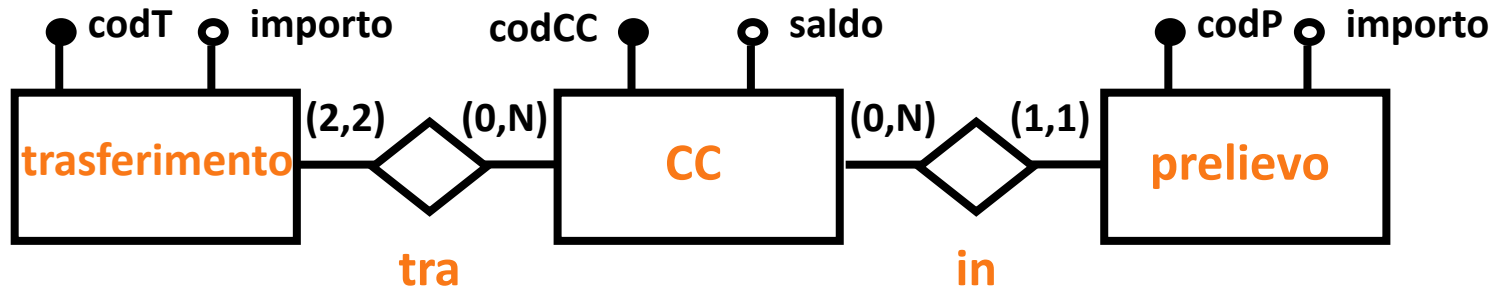
Senza il dato derivato:

TOT: 34015 / giorno

Conclusione: conviene mantenere il dato derivato

Esercizio 2

- Dato derivato: **saldo**



- Operazione 1:
inserimento di un nuovo prelievo su un dato **codCC** (si suppone che il CC sia valido e già presente)
- Operazione 2:
inserimento di un nuovo trasferimento da un **codCC** a un altro (si suppone che i due CC siano validi e già presenti)
- Operazione 3:
dato il **codCC**, visualizzare tutti i dati di un CC, incluso il **saldo**

Esercizio 2

– Tabella dei volumi:

Concetto	Tipo	Volume dati
Trasferimento	E	10000
CC	E	1000
Prelievo	E	30000

– Tabella delle operazioni:

Operazione	Tipo	Frequenza
Oper. 1	I	200/giorno
Oper. 2	I	10/giorno
Oper. 3	I	4/giorno

Esercizio 2

Con il dato derivato:

TOT: 1524 / giorno

	Concetto	Acc.	Tipo
Operazione 1 1 accesso in lettura 3 accesso in scrittura 7*200 = 1400 / giorno	Prelievo	1	S
	In	1	S
	CC	1	L
	CC	1	S
Operazione 2 2 accessi in lettura 5 accessi in scrittura 12*10 = 120 / giorno	Trasferimento	1	S
	Tra	2	S
	CC	2	L
	CC	2	S
Operazione 3 1 accesso in lettura 1*4 = 4 / giorno	CC	1	L

Esercizio 2

Senza il dato derivato:

TOT: 1264 / giorno

Operazione 1 2 accessi in scrittura $4 \cdot 200 = \mathbf{800}$ / giorno	Concetto	Acc.	Tipo
	Prelievo	1	S
Operazione 2 3 accessi in scrittura $6 \cdot 10 = \mathbf{60}$ / giorno	In	1	S
	Trasferimento	1	S
	Tra	2	S
Operazione 3 101 accessi in lettura $101 \cdot 4 = \mathbf{404}$ / giorno	CC	1	L
	In	30	L
	Prelievo	30	L
	Tra	20	L
	Trasferimento	20	L

Esercizio 2

Con il dato derivato:

TOT: 1524 / giorno

Senza il dato derivato:

TOT: 1264 / giorno

Conclusione: non conviene mantenere il dato derivato