

Esercitazione Ristrutturazione

“Basi di dati – Modelli e linguaggi di interrogazione”

Paolo Atzeni, Stefano Ceri, Stefano Paraboschi, Riccardo Torlone

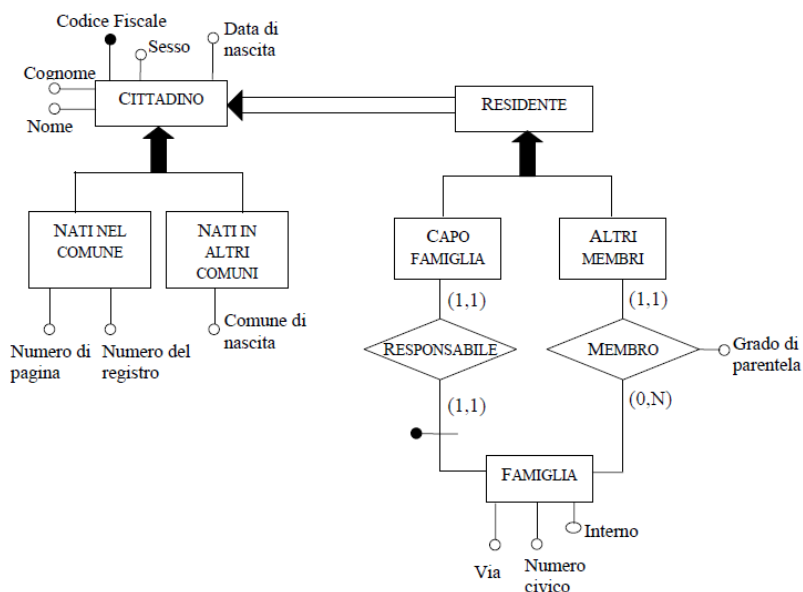
Questa esercitazione è tratta dal libro:

Esercizio 8.1

Si consideri lo schema Entità-Relazione ottenuto come soluzione dell'esercizio 7.4. Fare delle ipotesi sul volume dei dati e sulle operazioni possibili su questi dati e, sulla base di queste ipotesi, effettuare le necessarie ristrutturazioni dello schema. Effettuare poi la traduzione verso il modello relazionale.

Soluzione:

Questo è lo schema prodotto nell'esercizio 7.4



Le seguenti tavole contengono ipotesi sui volumi e sulle operazioni:

Volumi:

Concetto	Tipo	Volume
Cittadino	E	1.100.000
Nati nel comune	E	1.000.000
Nati in altri comuni	E	100.000
Residente	E	1.000.000
Capo famiglia	E	250.000
Altri membri	E	750.000
Famiglia	E	250.000
Responsabile	R	250.000
Membro	R	750.000

Operazioni:

Operazione	Descrizione	Frequenza	Tipo
1	Aggiungere un nuovo cittadino nato nel comune	100 al giorno	I
2	Aggiungere un nuovo cittadino residente nel comune ma nato in un altro comune	20 al giorno	I
3	Aggiungere una nuova famiglia	20 al giorno	I
4	Cancellare un cittadino	100 al giorno	I
5	Cancellare una famiglia	5 al giorno	I
6	Visualizzare il numero di cittadini residenti nel comune	1 al giorno	B
7	Visualizzare un numero di residenti uomini e donne	1 al giorno	B

Potrebbe essere utile per aggiungere un attributo ridondante “Numero di Componenti” all’entità FAMIGLIA. Senza questo attributo, l’operazione 6 ha bisogno di 1.000.000 di accessi in lettura all’entità RESIDENTE ogni giorno. Con questo attributo ridondante, l’operazione 6 ha bisogno di soli 250.000 accessi in lettura all’entità FAMIGLIA.

Comunque, la presenza di questo attributo cambia il costo delle operazioni 1, 2 e 4; infatti queste 3 operazioni hanno ora bisogno, oltre agli accessi che già avevano, anche di un accesso in lettura a CAPO FAMIGLIA (o ad ALTRI MEMBRI), un accesso a RESPONSABILE (o a MEMBRO), un accesso in lettura ed uno in scrittura all’entità FAMIGLIA (per aggiornare l’attributo “Numero di componenti”).

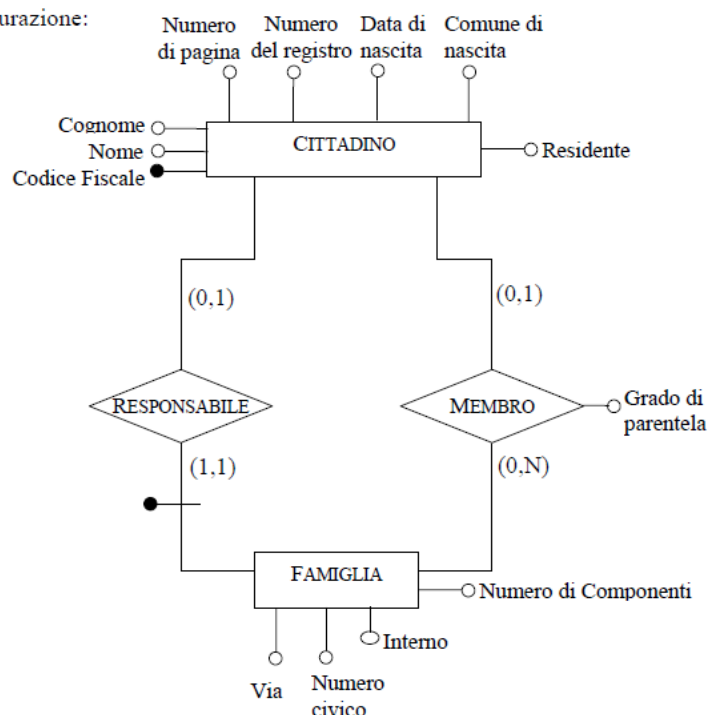
Supponendo che un accesso in scrittura abbia il costo di 2 accessi in lettura, il costo totale è $(1+1+1+2)*90 + (1+1+1+2)*20 + (1+1+1+2)*100 = 1.050$

La frequenza dell’operazione 1 è 90 perché non tutti i cittadini nati nel comune sono residenti, ma solo il 90%.

Così, il vantaggio dell’attributo ridondante è $750.000 - 1.050 = 748.950$ accessi al giorno.

Alla fine l’attributo ridondante viene mantenuto. La ristrutturazione ed il relativo mapping proposti sono leggermente differenti da quella vista durante l’esercitazione.

Ristrutturazione:



Traduzioni:

CITTADINO(Codice Fiscale, Cognome, Nome, Numero di pagina, Numero del registro, Data di nascita, Comune di nascita, Residente)

FAMIGLIA(Capo Famiglia, Via, Numero civico, Interno, Numero di Componenti) con vincolo di integrità referenziale tra Capo Famiglia e la relazione CITTADINO.

MEMBRO(Cittadino, Famiglia, Grado di parentela) con vincolo di integrità referenziale tra Cittadino e la relazione CITTADINO e tra Famiglia e la relazione FAMIGLIA.

Esercizio 8.7

Si consideri lo schema concettuale di Figura 8.38, che descrive i dati di conti correnti bancari. Si osservi che un cliente può essere titolare di più conti correnti e che uno stesso conto corrente può essere intestato a diversi clienti. Si supponga che su questi dati, sono definite le seguenti operazioni principali:

Operazione 1: Apri un conto corrente ad un cliente.

Operazione 2: Leggi il saldo totale di un cliente.

Operazione 3: Leggi il saldo di un conto.

Operazione 4: Ritira i soldi da un conto con una transazione allo sportello.

Operazione 5: Deposita i soldi in un conto con una transazione allo sportello.

Operazione 6: Mostra le ultime 10 transazioni di un conto.

Operazione 7: Registra transazione esterna per un conto.

Operazione 8: Prepara rapporto mensile dei conti.

Operazione 9: Trova il numero dei conti posseduti da un cliente.

Operazione 10: Mostra le transazioni degli ultimi 3 mesi dei conti delle società con saldo negativo.

Si supponga infine che, in fase operativa, i dati di carico per questa applicazione bancaria siano quelli riportati in figura 8.39.

Effettuare la fase di progettazione logica sullo schema E-R tenendo conto dei dati forniti. Nella fase di ristrutturazione si tenga conto del fatto che sullo schema esistono due ridondanze: Gli attributi **Saldo Totale** e **Numero di Conti** dell'entità **CLIENTE**. Essi possono infatti essere derivati dall'associazione **TITOLARITÀ** e dall'entità **CONTO**.

Anche l'attributo **Saldo** di conto può essere derivato dall'entità **Transazione**.

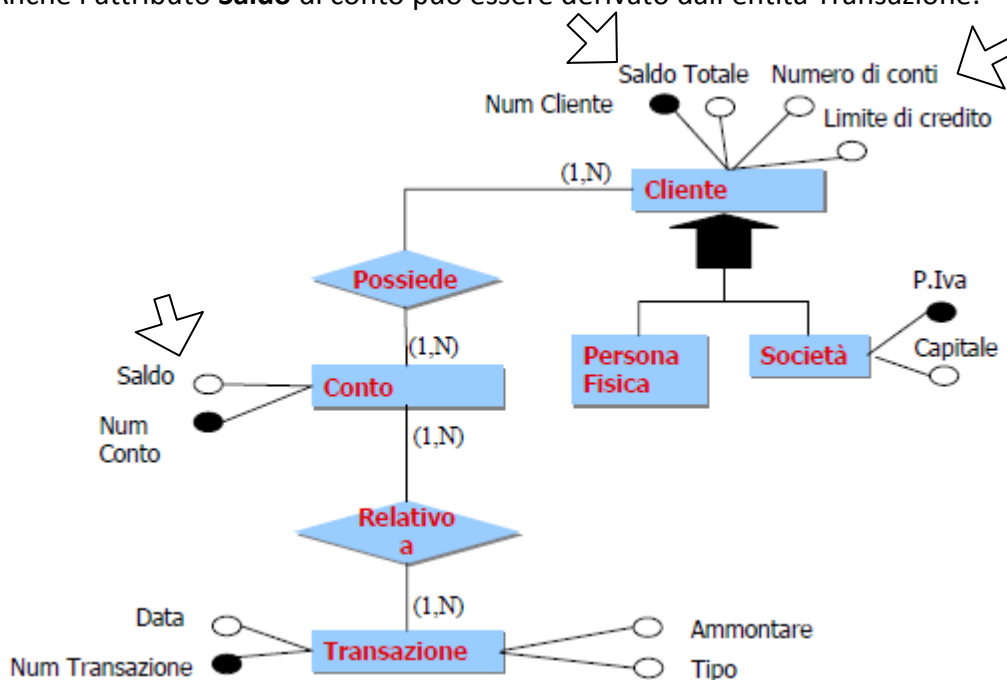


Figura 8.38

Tavola dei volumi		
Concetto	Tipo	Volume
Cliente	E	15000
Conto	E	20000
Transazione	E	600000
Persona Fisica	E	14000
Società	E	1000
Titolarità	R	30000
Operazione	R	800000

Tavola delle operazioni		
Operazione	Tipo	Frequenza
Op. 1	I	100/giorno
Op. 2	I	2000/giorno
Op. 3	I	1000/giorno
Op. 4	I	2000/giorno
Op. 5	I	1000/giorno
Op. 6	I	200/giorno
Op. 7	B	1 /giorno
Op. 8	B	1/mese
Op. 9	B	75/giorno
Op. 10	I	20/giorno

Tavola 8.39 Tavole dei volumi e delle operazioni per lo schema in figura 8.38

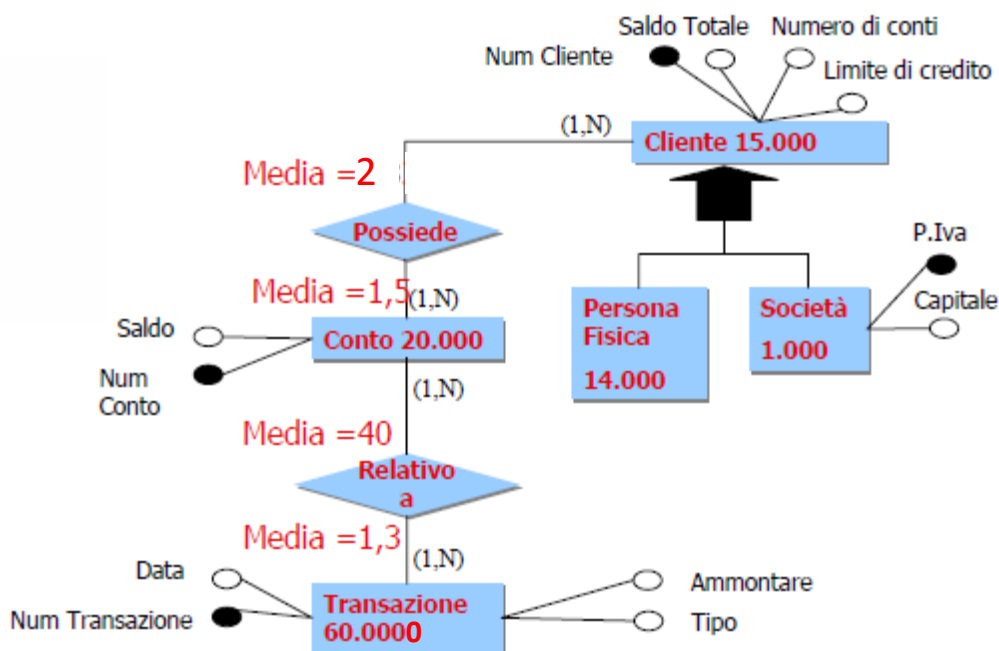


Figura 8.40 Volumi delle relazioni

La tavola degli accessi per tutte le operazioni: (L'operazione 8 è rara).

Operazione	Concetto	Tipo	L-S	Accessi
Op1	Conto	E	S	100
	Cliente	E	S	150
	Possiede	R	S	150
Op2	Cliente	E	L	2.000
Op3	Conto	E	L	1.000
Op4	Conto	E	L	2.000
	Conto	E	S	2.000
	Transazione	E	S	2.000
	Relativa a	R	S	2.000
	Possiede	R	L	3.000
	Cliente	E	S	3.000

Operazione	Concetto	Tipo	L-S	Accessi
Op5	Conto	E	L	1.000
	Conto	E	S	1.000
	Transazione	E	S	1.000
	Relativa a	R	S	1.000
	Possiede	R	L	1.500
	Cliente	E	S	1.500
Op6	Conto	E	L	200
	Relativa a	R	L	8.000
	Transazione	E	L	2.000
Op10	Società	E	L	400
	Possiede	R	L	400
	Conto	E	L	400
	Relativa	R	L	1.000
	Transazione	E	L	1.000

Operazione	Concetto	Tipo	L-S	Accessi
Op7	Conto	E	L	1
	Conto	E	S	1
	Transazione	E	S	1
	Relativa a	R	S	1
	Possiede	R	L	1,5
	Cliente	E	S	1,5
Op9	Conto	E	L	75

Analisi ridondanze:

Saldo in conto riguarda Op2, Op3, Op4, Op5, Op7, Op8 ed Op10.

Mantenendo la ridondanza:

- I benefici sono legati ad Op3: il suo costo sarebbe di 1.000 accessi in L al giorno.
- I costi riguardano Op4 (2.000/gg accessi in S) ed Op5 (1.000/gg accessi in S); l'occupazione di memoria di un intero (6B) per un totale di 120.000 B.

Togliendo la ridondanza:

- L'esecuzione di Op3 richiederebbe ancora l'accesso a Conto + quello a Transazione per un totale di $1.000 \times (40+1)$ /gg.

Quindi la transazione fa risparmiare $(41.000 - 3.000 =) 38.000$ transazioni /gg contro 120.000 B di memoria.

Saldo totale in Cliente riguarda Op2, Op4, Op5, ed Op10.

Mantenendo la ridondanza:

- I costi sono:
 - $6B \times 15.000$ (istanze di cliente) = 90.000 memoria in più;
 - L'Op4 costerebbe $2.000 \times 1,5 = 3000$ e l'Op5 costerebbe $1.000 \times 1,5 = 1.500$, più la riscrittura di Cliente, per un totale di 4.500 accessi in L e S /gg

Togliendo la ridondanza

- L'Op2 richiederebbe la visita della relazione Possiede e l'entità conto per una media di $2.000 \times 2 = 4.000$ accessi in più /gg.
- L'Op10 dovrebbe eseguire $20 \times 1.000 \times 2 = 40.000$ accessi in L /gg, contro i 400 presenti in caso di ridondanza, con un costo aggiuntivo di 39.600.
- Il costo totale sarebbe di 43.600

43.600 contro 4.500 (S) + 4.500 (L) + 90KB: manteniamo la ridondanza.

Consideriamo la ridondanza Numero di Conti in Cliente. Rigarda le Op1 e le Op9.

Mantenendo la ridondanza:

- Op9 farebbe risparmiare 150 letture
- La memorizzazione di un intero 2 B per ogni istanza di Cliente, pari a 30.000 B.
- Op1 necessita di $100 \times 1,5 = 150$ accessi a Cliente $\times 2 = 300$ accessi a Possiede e a Conto, per un totale di 150 S di Cliente.

150 L risparmiati contro 450 L, 150 S e 30 KB, per cui la ridondanza è eliminata.

Eliminazione delle gerarchie

La gerarchia è totale ed esaustiva.

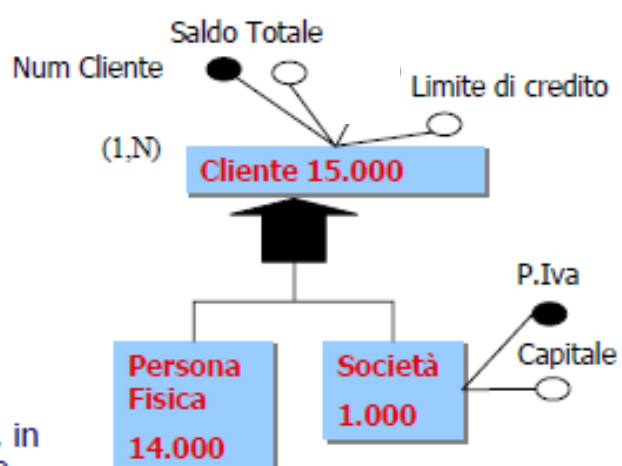
L'unica operazione che effettua la distinzione tra le due entità è Op10 (Mostrare le transazioni degli ultimi tre mesi dei conti delle società con saldo negativo)

Nessuna operazione utilizza P.Iva e Capitale.

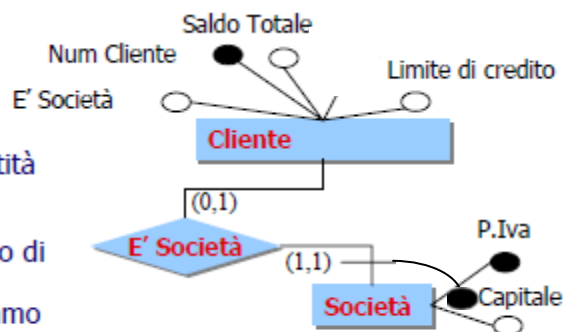
Nessuna operazione accede in modo esclusivo a Persona Fisica, che non ha attributi propri.

Non è conveniente fondere le tre entità, in quanto 14/15 sono persone fisiche, che avrebbero valori nulli per P.Iva (6 B) e Capitale (6B).

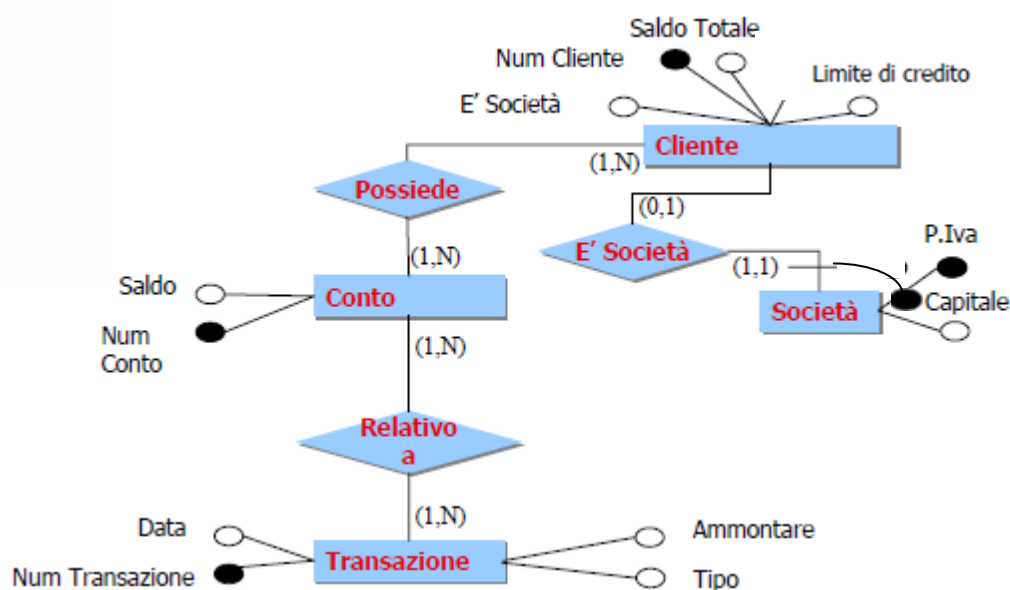
Non conviene tenere solo le entità figlie, perché non vi è un accesso privilegiato rispetto al padre.



Non ci rimane che mantenere le tre entità ed introdurre due relazioni.
 Notiamo che non avendo attributi aggiuntivi, le persone non necessitano di una entità a parte.
 Per migliorare le prestazioni, introduciamo un attributo E' Società nel cliente, per ridurre gli accessi.



Schema ristrutturato



Mapping nel modello relazionale

CLIENTE (Numero Cliente, Saldo Totale, Limite di Credito, E' Società)

SOCIETÀ (Numero Cliente, Partita Iva, Capitale)

CONTO (Numero Conto, Saldo)

TRANSAZIONE (Numero Transazione, Data, Tipo, Ammontare)

POSSIEDE (Numero Cliente, Numero Conto)

RELATIVA A (Numero Transazione, Numero Conto)