

Progetto Basi di Dati

Paolo Addis
Samuele Poz
Tristano Munini

RENDERE CARINO

Indice

1	Progettazione Concettuale	1
1.1	Raccolta ed Analisi dei Requisiti	1
1.2	Progettazione del Modello E-R	2
1.2.1	Le Entità	2
1.2.2	Le Relazioni	4
2	Progettazione Logica	5
2.1	Ristrutturazione del Modello E-R	5
2.1.1	Semplificazione dei Concetti	5
2.2	Analisi delle Ridondanze	5
2.3	Traduzione da ER a Relazionale	7
3	Progettazione Fisica	8
3.1	Nuovi Indici	8
4	Definizione della Base di Dati in SQL	9
4.1	Definizione delle Tabelle	9
4.2	Definizione di Query Significative	9
5	Analisi Dati	10
5.1	Popolamento della Base di Dati	10
5.2	Analisi Dati in R	10
6	Portale Web	11
6.1	Interfaccia Grafica	11
6.2	Comunicare con pgAdmin	11

1 Progettazione Concettuale

1.1 Raccolta ed Analisi dei Requisiti

Si vuole modellare il seguente insieme di informazioni riguardanti un sistema per la gestione delle diagnosi e delle terapie dei pazienti ricoverati in un dato ospedale.

- Di ogni ricovero, il sistema deve memorizzare il codice univoco, il nome della divisione ospedaliera (Cardiologia, Reumatologia, Ortopedia, ...), il paziente ricoverato, le date di inizio e fine del ricovero e il motivo principale del ricovero.
- Di ogni paziente, il sistema deve memorizzare il codice sanitario (univoco), il cognome, il nome, la data di nascita, il luogo di nascita e la provincia di residenza. Per i pazienti residenti fuori regione, vengono memorizzati anche il nome della ULSS e la regione di appartenenza.
- Di ogni diagnosi effettuata durante il ricovero del paziente, sono memorizzati la patologia diagnosticata, col suo codice ICD10 (classificazione internazionale delle patologie) e l'indicazione della sua gravità (grave: si/no), la data e il nome e cognome del medico che ha effettuato la diagnosi.
- Nella base di dati si tiene traccia delle terapie prescritte ai pazienti durante il ricovero. Di ogni terapia, si memorizzano il farmaco prescritto, la dose giornaliera, le date di inizio e di fine della prescrizione, la modalità di somministrazione ed il medico che ha prescritto la terapia.
- Di ogni farmaco sono memorizzati il nome commerciale (univoco), l'azienda produttrice, il nome e la quantità dei principi attivi contenuti e la dose giornaliera raccomandata.
- Si tiene, infine, traccia delle diagnosi che hanno motivato le terapie. In particolare, ogni terapia è prescritta al fine di curare una o più patologie diagnosticate. Può capitare anche che una nuova patologia (registrata come nuova diagnosi) sia causata, come effetto collaterale, da una terapia precedentemente prescritta. Tale legame causa-effetto va registrato nella base di dati.

Tramite un portale web dovrà essere possibile accedere a tale base di dati ed aggiungere o rimuovere pazienti, ricoveri, diagnosi con relative terapie e farmaci.

1.2 Progettazione del Modello E-R

Inizialmente l'attenzione è stata concentrata sulle entità PAZIENTE e RICOVERO in quanto ritenute fondamentali: si può dire che in assenza di un'entità che rappresenti i pazienti, l'intera base di dati perderebbe significato. Infatti per ogni ricovero o diagnosi deve essere possibile risalire al soggetto che è stato ricoverato o a cui è stata effettuato una diagnosi, altrimenti non si potrebbero utilizzare correttamente le informazioni raccolte in passato. RICOVERO, invece, segna i momenti in cui il paziente raggiunge e lascia la struttura ospedaliera, inoltre permette di indicare in modo univoco le visite e le cure alle quali è stato sottoposto durante questo periodo. Le due entità sono in relazione tra di loro attraverso una VIENE-RICOVERATO che intuitivamente collega una entità PAZIENTE a molte entità RICOVERO. È stata poi costruita la tripla FARMACO, SOMMINISTRATO-DURANTE e TERAPIA composta rispettivamente da: un'entità che rappresenta il farmaco con i relativi attributi; una relazione che permette di indicare che farmaco viene usato durante una terapia; infine un'entità che raccoglie i dettagli della terapia, ad esempio la dose giornaliera di farmaco da assumere. L'entità DIAGNOSI e le rimanenti relazioni, in particolare CAUSA-EFFETTO, verranno analizzate successivamente.

1.2.1 Le Entità

- PAZIENTE: dai requisiti è possibile individuare le seguenti voci significative: *codice sanitario, cognome, nome, data di nascita, luogo di nascita e provincia di residenza* del paziente. Per ognuna di esse è stato creato un attributo. Il codice fiscale (d'ora in poi detto anche CF) è stato selezionato come chiave primaria in quanto, come definito nei requisiti, univoco¹. Per salvare gli attributi esclusivi per i pazienti fuori regione riguardanti il nome della ULSS e la regione di appartenenza, è stata creata l'entità PAZIENTE-FUORI-REGIONE, specializzazione dell'entità precedente. Poiché solo una frazione dell'insieme dei pazienti sarà di questo tipo la specializzazione è parziale.
- RICOVERO: dai requisiti è possibile individuare le seguenti voci significative: *codice ricovero, divisione ospedaliera, paziente ricoverato, data inizio, data fine* del ricovero. Per ognuna di esse è stato creato un attributo, fatta eccezione per *paziente ricoverato*. Quest'ultima informazione verrà memorizzata attraverso la relazione VIENE-RICOVERATO, descritta in seguito.
- TERAPIA:
- FARMACO
- DIAGNOSI

¹Proprietà fondamentale di un attributo per la definizione di esso come chiave primaria.

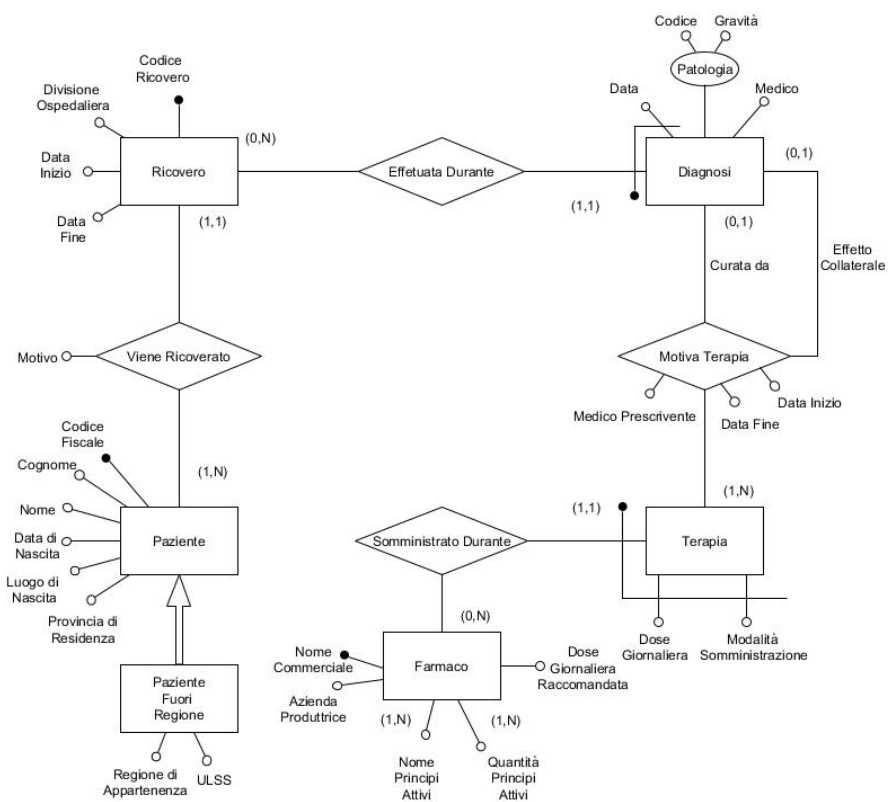


Figura 1: Lo schema Entità-Relazioni completo

1.2.2 Le Relazioni

- VIENE-RICOVERATO
- SOMMINISTRATO-DURANTE
- EFFETTUATA-DURANTE
- CARTELLA-CLINICA

Abbiamo prestato particolare attenzione a CAUSA-EFFETTO. Questa relazione potrebbe sembrare eccessiva essendo sia ternaria che ricorsiva, in realtà modello molto bene un requisito fondamentale: *“Può capitare anche che una nuova patologia (registrata come nuova diagnosi) sia causata, come effetto collaterale, da una terapia precedentemente prescritta.”* Da un altro requisito sappiamo che una DIAGNOSI può essere in relazione con una TERAPIA in quanto la terapia viene prescritta solo se esiste una diagnosi che la motivi.

2 Progettazione Logica

2.1 Ristrutturazione del Modello E-R

2.1.1 Semplificazione dei Concetti

È necessario semplificare strutture come specializzazioni o generalizzazioni, attributi composti o multivalore in questo modo renderemo più facile la traduzione del kk

2.2 Analisi delle Ridondanze

Consideriamo il ciclo PAZIENTE, RICOVERO, DIAGNOSI dato dalle relazioni VIENE RICOVERATO, EFFETTUATA DURANTE, EFFETTUATA A.

Procediamo a valutare se è conveniente mantenere la relazione EFFETTUTA A tra PAZIENTE e DIAGNOSI in funzione della frequenza delle operazioni di inserimento nuove diagnosi e stampa dello storico diagnosi di un paziente.

Riportiamo i volumi stimati durante la fase dei requisiti in Tabella 2.2 e la frequenza delle operazioni interessate in Tabella 2.2.

Concept	Type	Volume
Paziente	E	100000
Ricovero	E	300000
Diagnosi	E	1200000
Terapia	E	30000
Istanza di Terapia	E	1200000
Farmaco	E	10000
Viene Ricoverato	R	300000
Effettuata Durante	R	1200000
Effettuta A	R	1200000
Curata Da	R	1200000
Istanza Di	R	1200000
Effetto Collaterale	R	50000
Somministrato Durante	R	10000

Tabella 1: Tabella dei volumi

Concept	Type	Volume
Paziente	E	100000
Ricovero	E	300000
Diagnosi	E	1200000
Viene Ricoverato	R	300000
Effettuata Durante	R	1200000
Effettuata A	R	1200000

Tabella 2: Tabella dei volumi ridotta

Operation	Type	Frequency
Inserimento Diagnosi	I	700/giorno
Stampa Storico Diagnosi	I	1400/giorno

Tabella 3: Tabella delle operazioni ridotta

Nel caso con ridondanza per l'operazione di inserimento di una nuova diagnosi sono necessari 1 accesso in scrittura per l'entità DIAGNOSI e 1 accesso in scrittura per entrambe le relazioni EFFETTUATA DURANTE e EFFETTUA-TA A. Per l'operazione di stampa delle diagnosi di un paziente sono necessari 12 accessi in lettura per la relazione EFFETTUATA A e 12 accessi in lettura per l'entità DIAGNOSI.

Considerando il costo di 1 accesso in scrittura come 2 accessi in lettura, calcoliamo il numero di accessi complessivo giornaliero per le 2 operazioni nel caso con ridondanza:

$$3W \times 700 + 24R \times 1400 \approx 37800R \quad (1)$$

Concept	Type	Access	Type
Inserimento Diagnosi			
Diagnosi	E	1	W
Effettutata Durante	R	1	W
Effettuata A	R	1	W
Stampa Storico Diagnosi			
Effettuata A	R	12	R
Diagnosi	E	12	R

Tabella 4: Tabella dei costi, caso con ridondanza

Nel caso senza ridondanza per l'operazione di inserimento di una nuova diagnosi sono necessari 1 accesso in scrittura per l'entità DIAGNOSI e 1 accesso in scrittura per la relazione EFFETTUATA DURANTE. Per l'operazione di stampa delle diagnosi di un paziente sono necessari 3 accessi in lettura per la relazione

VIENE RICOVERATO, 12 accessi in lettura per la relazione EFFETTUATA DURANTE e 12 accessi in lettura per l'entità DIAGNOSI.

Considerando il costo di 1 accesso in scrittura come 2 accessi in lettura, calcoliamo il numero di accessi complessivo giornaliero per le 2 operazioni nel caso senza ridondanza:

$$2W \times 700 + 27R \times 1400 \approx 40600R \quad (2)$$

Concept	Type	Access	Type
Inserimento Diagnosi			
Diagnosi	E	1	W
Effettutata Durante	R	1	W
Stampa Storico Diagnosi			
Viene Ricoverato	R	3	R
Effettutata Durante	R	12	R
Diagnosi	E	12	R

Tabella 5: Tabella dei costi, caso senza ridondanza

Concludiamo che è conveniente mantenere la ridondanza visto il costo minore delle operazioni giornaliere.

2.3 Traduzione da ER a Relazionale

3 Progettazione Fisica

3.1 Nuovi Indici

4 Definizione della Base di Dati in SQL

4.1 Definizione delle Tabelle

4.2 Definizione di Query Significative

5 Analisi Dati

5.1 Popolamento della Base di Dati

5.2 Analisi Dati in R

6 Portale Web

6.1 Interfaccia Grafica

6.2 Comunicare con pgAdmin