

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	3
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	5
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: sì/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 diagnostiche. 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 effettuata 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 tripletta 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. Il *codice ricovero* è stato selezionato come chiave primaria dell'entità in quanto univoco.
- DIAGNOSI: dai requisiti è possibile individuare le seguenti voci significative: *patologia diagnosticata*, *codice patologia*, *gravità patologia*, *data*, *medico*. Per ognuna di esse è stato creato un attributo. In particolare *patologia* è un attributo composto che si suddivide in *codice* e *gravità*. Si

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

assume che la data comprenda anche l'ora in cui è stata effettuata la diagnosi. DIAGNOSI è un'entità debole in quanto la chiave primaria è definita attraverso l'attributo *data* e la relazione EFFETTUATA-DURANTE. La chiave è stata stabilita sfruttando l'assunzione precedente riguardo la data in modo tale che per ogni ricovero non sia possibile associare più diagnosi effettuate nello stesso istante.

- FARMACO: dai requisiti è possibile individuare le seguenti voci significative: *nome commerciale, azienda produttrice, nome principi attivi, quantità principi attivi, dose giornaliera raccomandata*. Per ognuna di esse è stato creato un attributo. In particolare *nome principi attivi e quantità principi attivi* sono attributi multivalore e la chiave primaria è rappresentata da *nome commerciale* in quanto univoco.
- TERAPIA: dai requisiti è possibile individuare le seguenti voci significative: *farmaco prescritto, dose giornaliera, modalità di somministrazione, medico prescrivente*. Sono stati definiti gli attributi *dose giornaliera* e *modalità somministrazione* mentre l'informazione relativa alle altre voci verrà considerata in altre entità e relazioni. La terapia è intesa come metodo per la somministrazione di un farmaco al fine di curare una patologia, di conseguenza, la chiave primaria è composta dagli attributi di terapia e dalla relazione SOMMINISTRATO-DURANTE che rende l'entità un'entità debole.

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 modella 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.

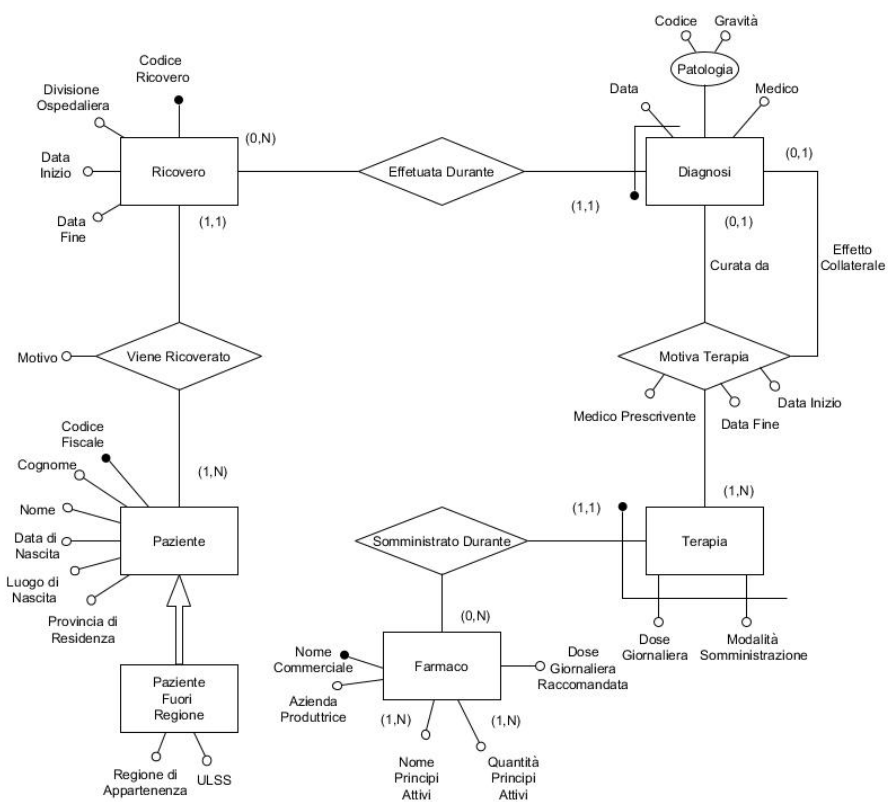


Figura 1: Lo schema Entità-Relazioni completo

2 Progettazione Logica

2.1 Ristrutturazione del Modello E-R

2.1.1 Semplificazione dei Concetti

È utile semplificare strutture come specializzazioni, attributi composti e attributi multivalore perché renderà più facile la traduzione del modello Entità-Relazioni in quello Relazionale.

Ricordando che quest'ultimo non permette di rappresentare generalizzazioni sarà necessario rimuovere tutte le specializzazioni dal nostro schema E-R. Si può notare che la specializzazione PAZIENTE-FUORI-REGIONE può essere facilmente rappresentata aggiungendo due attributi all'entità PAZIENTE. I nuovi attributi, *ULSS* e *Regione di Appartenenza*, in PAZIENTE saranno facoltativi poiché la specializzazione è parziale. La rimozione delle generalizzazioni non è sempre così immediata: se PAZIENTE-FUORI-REGIONE avesse avuto una relazione in più rispetto al genitore si sarebbero dovute valutare varie opzioni. Per esempio la creazione di una nuova relazione per simboleggiare il legame tra genitore e figlio, lasciando invariati gli attributi. Oppure la duplicazione nel figlio di tutti gli attributi e di tutte le relazioni del genitore.

Anche nel caso degli attributi composti la semplificazione è immediata, infatti nello schema l'unico attributo composto è *Patologia* in DIAGNOSI. Può essere trasformato in due nuovi attributi *Codice Patologia* e *Gravità Patologia* che contengono esattamente la stessa informazione.

Bisogna prestare più attenzione agli attributi multivalore: in FARMACO sono presenti *Nome Principi Attivi* e *Quantità Principi Attivi* entrambi obbligatori e multivalore. L'unica semplificazione attuabile è creare una nuova entità chiamata PRINCIPIO-ATTIVO e metterla in relazione Molti-A-Molti con FARMACO attraverso la nuova relazione CONTIENE. PRINCIPIO-ATTIVO avrà un solo attributo, scelto anche come chiave primaria, chiamato *Nome*, mentre alla relazione CONTIENE verrà aggiunto l'attributo *Quantità*. In questo modo saranno modellate esattamente le stesse informazioni precedentemente modellate dagli attributi multivalore. Rimozione e Scelta delle chiavi primarie

2.2 Analisi delle Ridondanze

2.3 Traduzione da ER a Relazionale

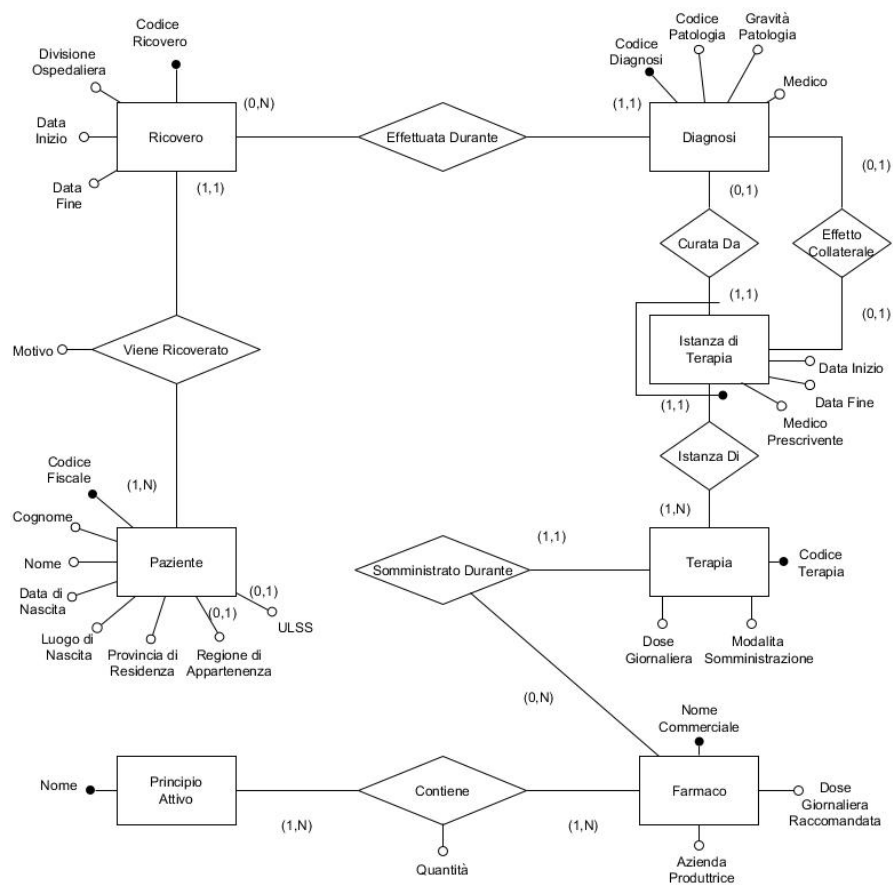


Figura 2: Lo schema Entità-Relazioni completo

Tabella 1: Tabella dei volumi

Concept	Type	Volume
Paziente	E	100000
Ricovero	E	500000
Diagnosi	E	2500000
Terapia	E	30000
Istanza di Terapia	E	2500000
Farmaco	E	10000
Viene Ricoverato	R	500000
Effettuata Durante	R	2500000
Diagnosi Effettuate	R	2500000
Curata Da	R	2500000
Istanza Di	R	2500000
Effetto Collaterale	R	100000
Somministrato Durante	R	10000

Tabella 2: Tabella delle operazioni

Operation	Type	Frequency
Inserimento Diagnosi	I	1400/giorno
Stampa Storico Diagnosi	I	280/giorno

Tabella 3: Tabella dei costi caso con ridondanza

Concept	Type	Access	Type
Inserimento Diagnosi			
Diagnosi	E	1	W
Effettutata Durante	R	1	W
Diagnosi Effettuate	R	1	W
Stampa Storico Diagnosi			
Paziente	E	1	R
Diagnosi Effettuate	R	25	R
Diagnosi	E	25	R

Tabella 4: Tabella dei costi caso senza ridondanza

Concept	Type	Access	Type
Inserimento Diagnosi			
Diagnosi	E	1	W
Effettutata Durante	R	1	W
Stampa Storico Diagnosi			
Paziente	E	1	R
Viene Ricoverato	R	5	R
Effettutata Durante	R	25	W
Diagnosi	E	25	R

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