

Compito di Basi di dati - 9 CFU

15 settembre 2014

Esercizio 1:

Sia dato il seguente schema di una base di dati relazionale contenente alcune informazioni relative alle visite eseguite dai pazienti di una data unità sanitaria locale:

paziente(*codiceTesseraSanitaria*, *nome*, *cognome*, *numeroTelefono*, *città*);

visita(*Paziente*, *medico*, *dataInizio*, *oraInizio*, *durata*);

medico(*CodFisc*, *cognome*, *nome*, *specialità*).

Si assuma che ogni paziente sia identificato univocamente dal codice della sua tessera sanitaria e sia caratterizzato da nome, cognome, numero di telefono e città). Si assuma che pazienti diversi possano avere lo stesso recapito telefonico. Si assuma, inoltre, che ogni medico sia identificato dal suo codice fiscale e sia caratterizzato da nome, cognome e specialità. Si assuma anche che per ogni specialità siano disponibili uno o più medici. Infine, si assuma che ogni visita sostenuta da un paziente con un dato medico sia caratterizzata dal giorno e dall'ora in cui inizia e dalla durata. Si assuma che ogni visita sia effettuata da un unico medico e che un paziente possa sostenere più visite con lo stesso medico lo stesso giorno (ovviamente in ore diverse).

Definire preliminarmente le chiavi primarie, le eventuali altre chiavi candidate e, se ve ne sono, le chiavi esterne delle relazioni date. Successivamente, formulare opportune interrogazioni in algebra relazionale che permettano di determinare (senza usare l'operatore di divisione e usando solo se e quando necessario le funzioni aggregate):

- per ogni paziente visitato almeno una volta da un medico specializzato in Neurologia, il numero e la durata media delle visite fatte nel primo semestre 2012;
- il cognome e il nome dei medici specializzati in Ortopedia che nel mese di maggio 2013 non hanno fatto visite a pazienti di Verona;
- i medici che hanno visitato un sottoinsieme dei pazienti visitati dal medico Vincenzo Rossi (si assuma che vi sia un solo medico con tale nome e cognome).

Esercizio 2:

Con riferimento all'Esercizio 1, formulare opportune interrogazioni in SQL che permettano di determinare quanto richiesto (senza usare l'operatore CONTAINS e usando solo se e quando necessario le funzioni aggregate).

Esercizio 3:

Si voglia 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 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.

Si definisca uno schema Entità-Relazioni che descriva il contenuto informativo del sistema, illustrando con chiarezza le eventuali assunzioni fatte. Lo schema dovrà essere completato con attributi ragionevoli per ciascuna entità (identificando le possibili chiavi) e relazione. Vanno specificati accuratamente i vincoli di cardinalità e partecipazione di ciascuna relazione. Si definiscano anche eventuali regole di derivazione e/o vincoli di integrità necessari per codificare alcuni dei requisiti attesi del sistema.

Esercizio 4:

Si consideri il seguente schema di base di dati per la gestione dell'informazione relativa a un campionato mondiale di calcio:

PARTITA(squadra1, squadra2, goal1, goal2, data, orario, sede)

Chiave esterna: *squadra1* → SQUADRA(*codice*)

Chiave esterna: *squadra2* → SQUADRA(*codice*)

Vincolo Not Null: {*sede*, *orario*}

squadra1 ≠ *squadra2*

SQUADRA(codice, nome, gruppo : {A, ..., H})

Vincolo Not Null: {*nome*, *gruppo*}

Vincolo di unicità: {*nome*}

Ciascuna squadra è identificata da un codice di tre lettere (ad esempio, ESP per Spagna, BRA per il Brasile, e così via) ed è assegnata a un gruppo di squadre, denotato da una lettera da A a H. Ciascuna tupla di PARTITA memorizza i dati di una partita "squadra1-squadra2" (in quest'ordine). Gli attributi "goal1" e "goal2" rappresentano i goal segnati dalla prima e dalla seconda squadra, rispettivamente.

Si traduca lo schema relazionale dato in SQL usando opportuni domini e includendo tutti i vincoli d'integrità sopra specificati.

Esercizio 5:

Dato il seguente insieme di chiavi:

$A, B, S, T, Z, O, P, F, G, M, N, Q, C, V, D, E, R,$

mostrare il B^+ -albero, con ordine dei nodi interni $p = 5$ e ordine dei nodi foglia $p_{leaf} = 4$, ottenuto inserendo un elemento dopo l'altro nell'ordine dato (riportando la sequenza di alberi generata dal processo di inserimento).

Successivamente, si identifichino i nodi del B^+ -albero visitati nella ricerca di rispettivamente: (i) il record contraddistinto dal valore H (point query), (ii) il record contraddistinto dal valore G (point query) e (iii) i record con valori compresi nell'intervallo $F-M$ (range query).