

Compito di Basi di dati

8 settembre 2017

Esercizio 1:

Sia dato il seguente schema relazionale relativo agli insegnamenti offerti da un corso di laurea a ciclo unico di durata quinquennale (questo è il caso, ad esempio, del corso di laurea in giurisprudenza) di una data università:

STUDENTI(*Matricola*, *Nome_studente*, *Cognome_studente*, *Anno_immatricolazione*);

INSEGNAMENTI(*Cod_insegnamento*, *Nome_insegnamento*, *Cod_docente*, *Area*, *Anno_corso*);

ESAMI(*Studente*, *Insegnamento*, *Voto*, *Lode*).

Si assuma che ogni studente sia identificato univocamente dalla matricola e sia caratterizzato dal nome, dal cognome e dall'anno di immatricolazione (ad esempio, l'anno di immatricolazione degli studenti che si iscrivono al primo anno del corso di laurea nell'anno accademico 2017/18 è il 2017). Si assuma che ogni insegnamento sia identificato univocamente da un codice e sia caratterizzato dal codice del docente che lo tiene, dall'area disciplinare in cui si colloca (ad esempio, l'insegnamento di Complementi di basi di dati si colloca nell'area Basi di dati) e dall'anno di corso in cui si tiene (l'attributo Anno assume il valore 1 per gli insegnamenti del primo anno, il valore 2 per gli insegnamenti del secondo anno e così via—assumiamo che ogni insegnamento sia associato ad uno ed un solo anno). Si assuma che vi possano essere più insegnamenti con lo stesso nome (ad esempio, due diversi insegnamenti di algoritmi e strutture dati distinti mediante il loro codice). Si assuma, inoltre, che uno stesso docente possa tenere più insegnamenti. e che possano essere offerti più insegnamenti relativi alla stessa area disciplinare. Si assuma, infine, che la relazione esami registri i voti degli studenti relativi agli insegnamenti dei quali hanno superato l'esame (voti compresi tra 18 e 30). L'attributo Lode può assumere i valori vero e falso. Il valore vero può essere associato solo al voto 30. Si noti come la base di dati possa contenere studenti che non hanno registrato alcun esame ed insegnamenti il cui esame non è stato superato da alcuno studente.

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 necessario le funzioni aggregate):

- (a) gli studenti che hanno registrato solo esami relativi ad insegnamenti del biennio iniziale del corso di laurea;
- (b) per ogni area disciplinare, l'insegnamento (gli insegnamenti se più di uno) col maggior numero di esami registrati;
- (c) gli studenti che hanno superato l'esame di un sottoinsieme proprio degli insegnamenti dei quali lo studente con matricola 152473 ha superato l'esame.

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 vuole realizzare una base di dati per la registrazione delle opere d'arte presenti nei principali musei italiani sulla base del seguente insieme di requisiti.

- Di ogni città in cui si trova almeno un museo registrato nella base di dati o in cui è nato almeno uno degli artisti presenti nella base di dati, vogliamo memorizzare il nome, che la identifica univocamente, la collocazione geografica (nord, centro, sud, isole), la dimensione (numero di residenti) e la presenza o meno di un aeroporto nelle immediate vicinanze (ad esempio, la città di Venezia ha l'aeroporto Marco Polo).

- Ogni museo è identificato univocamente dal suo nome all'interno della città ove si trova (non ci possono essere due musei a Napoli di nome Museo Nazionale, ma ci possono essere un Museo Nazionale a Napoli e un Museo Nazionale a Torino). Di ogni museo vogliamo memorizzare l'indirizzo, il sito web, un indirizzo email per il pubblico, un recapito telefonico, i giorni e gli orari di apertura, il direttore e il numero di dipendenti.
- Ogni opera d'arte ha un codice, che la identifica univocamente. Di ogni opera d'arte interessano il titolo, l'anno in cui è stata realizzata, l'artista che l'ha fatta, il valore commerciale e il museo in cui si trova. Le opere d'arte sono classificate sulla base del secolo in cui sono state realizzate (opere del sedicesimo secolo, opere del diciassettesimo secolo e così via). Esse sono, inoltre, classificate sulla base del loro soggetto (nature morte, ritratti, paesaggi, soggetti religiosi, ..).
- Di ogni artista, identificato univocamente dal suo nome d'arte (ad esempio, Pinturicchio), che può coincidere o meno col suo nome proprio, vogliamo memorizzare le opere presenti nei diversi musei inseriti nella base di dati, le date di nascita e di morte, la città in cui è nato e quella in cui è morto e, più in generale, le città ove è vissuto.

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 gestione (regole di derivazione e vincoli di integrità) necessarie per codificare alcuni dei requisiti attesi del sistema.

Esercizio 4:

Anna e Biagio eseguono concorrentemente due transazioni sullo schema $R(\underline{A} : int)$ di una base di dati. Anna esegue la transazione seguente:

```
start transaction;
set transaction isolation level repeatable read;
select sum(A) from R;
select sum(A) from R;
commit;
```

Biagio esegue la transazione seguente:

```
start transaction;
set transaction isolation level repeatable read;
insert into R values (10);
commit;
```

Si assuma che, prima dell'inizio di tali transazioni, la somma dei valori in R sia 100 e che R non contenga il valore 10. Si assuma, inoltre, che il sistema non faccia uso di lock di predicato nel livello **repeatable read**.
 QUESITO 1: si specifichi, giustificando la risposta, quali sono i possibili valori calcolati dalle due **select** di Anna.
 QUESITO 2: Se invece di un inserimento, Biagio eseguisse una cancellazione, che cosa cambierebbe (allo stesso livello d'isolamento)?

Esercizio 5:

Si consideri un file contenente 10000000 record di dimensione prefissata pari a 400 byte, memorizzati in blocchi di dimensione pari a 4096 byte in modo unspanned. La dimensione del campo chiave primaria V sia 14 byte; la dimensione del puntatore a blocco P sia 6 byte. Si chiede di confrontare fra loro le seguenti soluzioni, in termini di numero medio di accessi a blocco e di dimensione dell'indice.

- Ricerca basata su un indice multilivello statico ottenuto a partire da un indice primario costruito sul campo chiave primaria V .
- Ricerca basata su un B^+ -albero, con campo di ricerca il campo chiave primaria V , puntatore ai dati di dimensione pari a 7 byte e puntatore ausiliario di dimensione pari a 6 byte, assumendo che ciascun nodo del B^+ -albero sia pieno al 70%.
- Ricerca basata su un B -albero, con campo di ricerca il campo chiave primaria V , puntatore ai dati di dimensione pari a 7 byte e puntatore ausiliario di dimensione pari a 6 byte, assumendo che ciascun nodo del B -albero sia pieno al 70%.