

Compito di Basi di dati - Informatica 9 CFU

16 luglio 2012

Esercizio 1:

Sia dato il seguente schema relazionale di una base di dati contenente informazioni relative ad un turno elettorale:

partito(*nomePartito*, *annoFondazione*, *cfSegretario*);

comizi(*candidato*, *città*);

candidati(*cfCandidato*, *nomeCandidato*, *cognomeCandidato*, *annoDiNascita*, *partitoCandidato*, *votiRicevuti*).

Si assuma che ogni partito sia identificato univocamente dal suo nome e ogni candidato dal suo codice fiscale. Non possano essere esclusi casi di omonimia. Si assuma che alcuni candidati possano aver tenuto più comizi in città diverse (ma non nella stessa città) e che altri possano non aver tenuto alcun comizio. Si assuma che ogni elettore non possa votare per un partito senza indicare il nome di un candidato (di quel partito) e che possa votare uno e un solo candidato. Per i candidati che non hanno ricevuto alcun voto, il valore dell'attributo *votiRicevuti* sia 0. Il totale dei voti ricevuti da un partito sia, quindi, pari alla somma dei voti ricevuti dai suoi candidati.

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 le funzioni aggregate solo se necessario):

- i partiti che hanno uno o due candidati nati nell'anno 1980;
- per ogni partito, il totale dei voti ricevuti, i voti ricevuti dal candidato (dai candidati, se più d'uno) che ha ricevuto il numero massimo di voti e la media dei voti ricevuti dai candidati;
- i partiti che hanno tenuto un comizio in almeno una città in cui nessun altro partito ha tenuto un comizio.

Esercizio 2:

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

Esercizio 3:

Si voglia modellare il seguente insieme di informazioni riguardanti gli incidenti stradali, che non hanno coinvolto pedoni, avvenuti nel territorio di un determinato comune.

- **Veicolo.** Di ogni veicolo coinvolto in almeno un incidente vengano memorizzati il numero di matricola, che lo identifica univocamente, l'anno di costruzione, la marca, il modello e la cilindrata. Si assuma che un veicolo possa essere coinvolto in più incidenti.
- **Passeggero.** Di ogni persona coinvolta in almeno un incidente vengano memorizzati il codice fiscale, o documento equivalente nel caso di stranieri, che lo identifica univocamente, il nome, il cognome e la data di nascita. Si assuma che una persona possa essere coinvolta in più incidenti.

- **Incidente.** Ogni anno, ogni incidente sia identificato univocamente da un numero progressivo (primo incidente del 2011, secondo incidente del 2011, etc.). Di ogni incidente, si vuole tener traccia del luogo (identificato dalle coordinate geografiche e dal nome della via, assumendo che anche gli incidenti avvenuti all'incrocio di più vie vengano comunque assegnati ad un'unica via) e dell'ora in cui è avvenuto, e dei veicoli e delle persone coinvolti. Di ogni persona coinvolta in un incidente vanno indicati il veicolo in cui si trovava (non si può escludere che una stessa persona possa essere coinvolta in più incidenti come passeggero di veicoli diversi), la posizione all'interno del veicolo (alla guida, a fianco del guidatore, in uno dei sedili posteriori) e le conseguenze riportate (in riferimento a 4 categorie: illeso, ferito lieve, ferito grave, deceduto). Sulla base delle conseguenze sui passeggeri coinvolti, gli incidenti vengano classificati in tre categorie principali: (i) incidenti mortali, quando almeno uno dei passeggeri è deceduto; (ii) incidenti gravi (quando almeno uno dei passeggeri è stato ferito in modo grave, ma nessun passeggero è deceduto); (iii) incidenti non gravi (quando nessun passeggero è deceduto o è rimasto gravemente ferito).
- **Via.** Ogni via sia identificata univocamente dal suo nome. Per ogni via, si tenga traccia del numero di incidenti mortali, gravi e non gravi in essa avvenuti suddivisi per anno.
- **Poliziotto.** Per ogni incidente, si vuole tener traccia dei poliziotti (uno o più) che hanno effettuato la rilevazione del sinistro. Ogni poliziotto sia identificato univocamente dal suo codice fiscale e sia caratterizzato da un nome, un cognome e un grado.

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 vincoli di integrità necessari per codificare alcuni dei requisiti attesi del sistema.

Esercizio 4:

Si consideri un file contenente 40000000 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, evidenziandone vantaggi e svantaggi.

- Ricerca basata su un indice primario costruito sul campo chiave primaria V .
- Ricerca basata su un indice secondario costruito sul campo chiave primaria V .
- Ricerca basata su un indice multilivello statico ottenuto a partire dall'indice primario di cui al punto (a).
- Ricerca basata su un indice multilivello statico ottenuto a partire dall'indice secondario di cui al punto (b).
- 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%.

Esercizio 5:

Stabilire se i seguenti schedule sono o meno serializzabili rispetto alle viste, ai conflitti, al metodo del locking a due fasi e al metodo del locking a due fasi stretto (nel caso di schedule appartenenti a VSR e in quello di schedule appartenenti a CSR, si fornisca almeno uno schedule seriale equivalente):

- $s_1 : r_3(y), r_1(x), r_3(z), w_5(x), w_1(t), r_2(z), w_4(x), w_5(z), r_5(t), w_2(z), r_4(t), w_1(y);$
- $s_2 : r_2(x), r_4(z), r_1(x), w_2(x), w_1(z), r_3(x), r_3(y), w_1(x), r_5(z), r_3(x), w_1(y), r_5(y), w_5(x).$