Compito di Basi di Dati e Sistemi Informativi

23 marzo 2004

Esercizio 1:

Sia dato il seguente schema relazionale, relativo a gare internazionali di lancio del peso, comprendente le relazioni:

$Atleta(Nome, Nazione, DataNascita)$;

$Gara(\mathit{Città}, Nazione, Anno, Budget)$;

$Risultati(\mathit{Città}, Anno, Atleta, Misura)$.

La base di dati vuole memorizzare informazioni su gare di lancio del peso che si svolgono in varie città del mondo (Londra, Montreal, Berlino, Zurigo, etc.). Tali gare possono svolgersi più volte nella stessa città, ma non nel corso di uno stesso anno (non necessariamente si svolgono ogni anno in ogni città). Per ogni gara, viene memorizzata la lunghezza massima (misura) raggiunta da ogni atleta attraverso un certo numero di lanci. La classifica di ogni gara è determinata dalle misure raggiunte dai partecipanti: vince la gara l'atleta che ottiene la misura maggiore. Assumiamo che la lunghezza dei lanci venga misurata in centimetri. Assumiamo, inoltre, per semplicità, che si sia mai verificato il caso di una gara in cui due o più atleti abbiano raggiunto la stessa identica misura.

Definire preliminarmente le chiavi primarie e le eventuali chiavi esterne delle relazioni date. Successivamente, formulare opportune interrogazioni in algebra relazionale che permettano di determinare:

(a) per ogni anno, il budget complessivamente stanziato da ogni nazione (si tenga presente che in un certo anno possono svolgersi più gare in una stessa nazione, ovviamente in città diverse);

(b) gli atleti che non hanno mai vinto una gara;

(c) le gare la cui classifica vede al primo posto l'atleta più giovane, al secondo posto il secondo atleta più giovane, al terzo posto il terzo atleta più giovane, e così via (per semplicità, si assuma che non vi siano atleti nati esattamente lo stesso giorno e che le date si possano confrontare direttamente attraverso gli operatori $<, \leq, \ldots$);

(d) per ogni nazione, la media delle misure ottenute dai propri atleti nella gara svoltasi a Berlino nel 2001, limitatamente alle nazioni con almeno due atleti che hanno superato i 2000 centimetri;

(e) con riferimento alla gara svoltasi a Londra nel 2003, tutte le coppie di atleti che si sono classificati uno dopo l'altro (primo e secondo, terzo e quarto, secondo e terzo, etc.).

Esercizio 2:

Si dimostri che, dati uno schema relazionale $R(A_1, \ldots, A_n)$, un insieme di dipendenze funzionali F ad esso associato e un insieme di attributi $X \subseteq \{A_1, \ldots, A_n\}$, la dipendenza $X \rightarrow A_i$ appartiene a F^+ (chiusura di F) se e sole se A_i appartiene a X_F^+ (chiusura di X rispetto a F).

Esercizio 3:

Si supponga che per archiviare dati relativi alla gestione di un ospedale sia stata usata un'unica tabella $OSPEDALE$ con la seguente struttura:

$OSPEDALE(CFPaziente, CognomePaziente, InizioRicovero, FineRicovero, Reparto, CFPrimario, CognomePrimario)$.

Si assuma che:

- ogni paziente sia identificato univocamente dal proprio codice fiscale;
- ogni primario sia identificato univocamente dal proprio codice fiscale;
- pazienti e primari abbiano un unico cognome;
- ogni reparto sia identificato dal suo nome;
- ogni reparto abbia un unico primario e ogni primario sia primario di un unico reparto;
- ogni paziente possa essere ricoverato più volte nello stesso reparto o in reparti diversi;
- durante un ricovero, un paziente non cambi reparto.

(a) Determinare le dipendenze funzionali della relazione $OSPEDALE$.

(b) Determinare le chiavi candidate e gli attributi primi e non primi di $OSPEDALE$.

(c) Stabilire se $OSPEDALE$ è o meno in BCNF.

(d) Stabilire se $OSPEDALE$ è o meno in 3NF.

(e) Stabilire se $OSPEDALE$ è o meno in 2NF.

(f) Nel caso in cui $OSPEDALE$ non sia in 3NF, fornire una scomposizione lossless join in 3NF di $OSPEDALE$ che conservi le dipendenze.

Esercizio 4:

Si consideri un file contenente 1000000 record di dimensione prefissata pari a 100 byte, memorizzati in blocchi di dimensione pari a 1024 byte in modo unspanned. Si determini la dimensione di un indice multilivello ottenuto a partire da un indice primario costruito su un campo chiave V di 9 byte, con dimensione del puntatore a blocco pari a 6 byte. Successivamente, si determini la dimensione di un B-albero, con campo di ricerca il campo chiave V di 9 byte, 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