

Compito di Complementi di Basi di Dati

15 settembre 2006

Esercizio 1:

Si supponga che per memorizzare i dati relativi alla gestione di un insieme di alberghi, appartenenti ad un'unica catena, sia stata usata un'unica tabella *ALBERGO* con la seguente struttura:

ALBERGO(*NomeAlbergo*, *NumeroStelle*, *Amministratore*, *MembroPersonale*,
RecapitoAmministratore, *Qualifica*, *OreSettimanali*)

Sia dato il seguente insieme di requisiti. Ogni albergo abbia un nome, che lo identifica univocamente, un amministratore e del personale. Di ogni albergo venga, inoltre, specificato il numero di stelle (1,2,3, ..). Un amministratore possa gestire più alberghi, tutti dello stesso livello (ossia con lo stesso numero di stelle). Un membro del personale (ad esempio, un addetto alla pulizia delle camere) possa lavorare per più alberghi. Ogni amministratore sia caratterizzato da un nome, che lo identifica univocamente, e da un recapito. Ogni membro del personale sia caratterizzato da un nome, che lo identifica univocamente, e da una qualifica (ad esempio, esperto contabile). Di ogni membro del personale si voglia sapere, per ogni albergo presso cui lavora, il numero di ore settimanali di lavoro presso tale albergo.

- Determinare le dipendenze funzionali della relazione *ALBERGO*, indicando, per ciascuna di esse, il requisito codificato.
- Determinare le chiavi candidate e gli attributi primi e non primi di *ALBERGO*.
- Stabilire se *ALBERGO* è o meno in 2NF.
- Stabilire se *ALBERGO* è o meno in 3NF.
- Stabilire se *ALBERGO* è o meno in BCNF.
- Nel caso in cui *ALBERGO* non sia in 3NF, fornire una scomposizione lossless join in 3NF di *ALBERGO* che conservi le dipendenze.

Esercizio 2:

Stabilire se i seguenti schedule appartengono o meno a VSR, CSR, TS, 2PL e 2PL stretto:

- $r_1(y), w_1(y), r_2(y), r_1(x), w_2(x)$
- $w_2(y), w_2(z), r_1(y), w_1(y), r_2(y), r_1(z)$
- $w_1(z), w_3(z), w_3(x), r_1(x), r_1(y), r_2(y), w_2(z)$
- $r_3(x), w_2(z), r_1(z), r_3(z), w_2(y), w_1(y), w_1(x)$

Esercizio 3:

Descrivere il calcolo dei profili dei risultati intermedi e del risultato finale (limitandosi alla cardinalità e

alla dimensione delle tabelle e al numero di valori distinti di ciascuno dei loro attributi) della seguente interrogazione¹:

$$\pi_{Cognome, Matricola, Provincia}(Studenti \bowtie_{\sigma_{Voto=27}}(Esami)),$$

relativa alle tabelle:

Studenti(*Cognome*, *Matricola*, *Provincia*);

Esami(*Materia*, *Matricola*, *Voto*),

aventi i seguenti profili:

$CARD(Studenti) = 40$, $SIZE(Studenti) = 27$, $CARD(Esami) = 1300$, $SIZE(Esami) = 22$,

$SIZE(Cognome) = 20$, $SIZE(Matricola) = 5$, $SIZE(Provincia) = 2$, $SIZE(Materia) = 15$,

$SIZE(Voto) = 2$,

$VAL(Cognome) = 38$, $VAL(Matricola) = 40$ (si assumo che tutti gli studenti abbiano superato almeno un esame e, quindi, $VAL(Matricola)$ sia uguale nelle due tabelle),

$VAL(Provincia) = 40$, $VAL(Materia) = 20$, $VAL(Voto) = 13$.

Esercizio 4:

Si fornisca una breve descrizione delle caratteristiche fondamentali delle tecniche di hashing esterno per l'organizzazione dei file di dati. Successivamente, si illustrino le differenze tra hashing (esterno) statico e dinamico. Infine, si presenti in modo sintetico la tecnica di hashing lineare.

¹Si ricordi che la formula $col(n, m, k)$ che calcola il numero di colori distinti presenti in k oggetti estratti a partire da n oggetti diversi di m colori distinti, distribuiti omogeneamente, ammette la seguente approssimazione:

(a) $col(n, m, k) = k$ se $k \leq m/2$;

(b) $col(n, m, k) = (k + m)/3$ se $m/2 \leq k \leq 2m$;

(c) $col(n, m, k) = m$ se $k \geq 2m$.