

Basi di Dati

Soluzione della prova scritta del 6 Luglio 2004

Esercizio 1 (2 punti)

La società ACME ha commissionato lo sviluppo e del proprio sistema informativo alla società SOFTSYS, che ha deciso di utilizzare un DBMS prodotto dalla società DBSNC. La gestione operativa è affidata a personale della società ACME. A quali organizzazioni appartengono persone delle seguenti categorie?

1. Progettisti della base di dati
2. Progettisti del DBMS
3. Amministratori della base di dati
4. Utenti finali
5. Programmatori delle applicazioni

Soluzione

1 (SOFTSYS), 2 (DBSNC), 3 (ACME), 4 (ACME), 5 (SOFTSYS)

Schema Relazionale per gli Esercizi 2 e 3

CD(Codice,Autore,CasaDiscografica)

CLIENTE(NTessera,Nome,Indirizzo)

ACQUISTO(CodiceCD,NTesseraCliente,Data,Quantità)

Foreign key CodiceCD *References* CD(Codice)

Foreign key NTesseraCliente *References* CLIENTE(NTessera)

Il *Cliente* identificato da *Ntessera* ha acquistato in una certa *Data* un numero di copie del *CD* identificato dal *Codice* pari a *Quantità*

Esercizio 2 (8 punti)

Formulare in algebra relazionale (4 punti) e in SQL (4 punti) la seguente interrogazione: *Selezionare i dati dei clienti che dopo il 1/1/2000 non hanno acquistato nessun CD prodotto dalla casa discografica DDD.*

Soluzione

Algebra Relazionale

$CLIENTE - \pi_{NTessera, Nome, Indirizzo} (CLIENTE \Join_{NTessera = NTesseraCliente} ((\sigma_{CasaDiscografica = 'DDD'} (CD) \Join_{Codice = CodiceCD} (\sigma_{Data > '1/1/2000'} (ACQUISTO))))$

SQL

```
SELECT *
FROM CLIENTE
WHERE NTessera NOT IN (SELECT NTesseraCliente
                        FROM ACQUISTO, CD
                        WHERE CodiceCD = Codice AND
                              CasaDiscografica = 'DDD' AND
                              Data > '1/1/2000')
```

Esercizio 3 (5 punti)

Formulare in SQL la seguente interrogazione: *selezionare, per ogni autore, il numero di tessera del cliente che ha acquistato il maggior numero di copie di CD di quell'autore.*

```
SELECT Autore, NTessera
FROM ACQUISTO AX, CD
WHERE AX.CodiceCD = CD.Codice
GROUP BY Autore, NTessera
HAVING SUM(Quantità) >= ALL
      (SELECT SUM(Quantità)
       FROM ACQUISTO AY, CD
       WHERE AY.CodiceCD = CD.Codice
             AND AY.Autore = AX.Autore
       GROUP BY NTessera)
```

Esercizio 4 (7 punti)

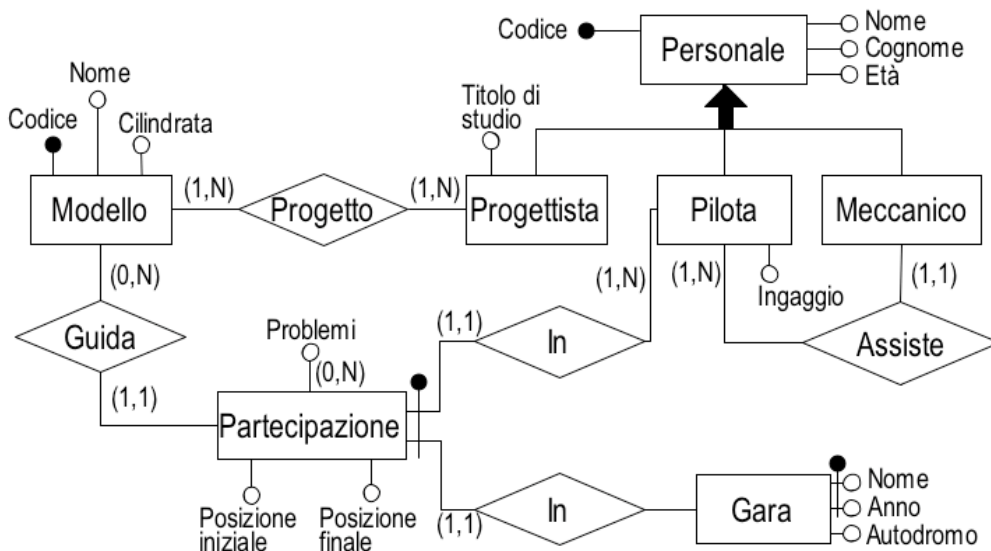
Mostrare uno schema E-R che rappresenti la seguente realtà.

Il sistema informativo della sezione "corse" di una casa automobilistica memorizza dati su: produzione di veicoli da competizione, disputa di gare internazionali e personale impiegato in tali attività in base alle seguenti specifiche

- Ogni modello da competizione ha un codice, un nome e una cilindrata.
- Ogni membro del personale ha: un nome, un cognome, età e un codice univoco. Il personale è diviso in piloti, meccanici e progettisti: nessuno può far parte di due categorie. Ogni meccanico è assegnato ad un pilota; un pilota può avere assegnati al massimo tre meccanici. I piloti hanno un ingaggio. I progettisti hanno un titolo di studio. Un progettista può aver progettato uno o più modelli di veicoli da competizione. Un modello è stato progettato da uno o più progettisti.
- Ogni gara ha un nome (esempio: G.P. d'Italia, di Germania, ecc.), un anno e un autodromo in cui si è disputata. Una gara non viene disputata sempre nello stesso autodromo, e può essere disputata in più anni, ma viene disputata al massimo una volta sola nello stesso anno ed in un unico autodromo.
- Un pilota partecipa ad una gara guidando un preciso modello. Per ognuna di queste partecipazioni si deve memorizzare la posizione iniziale e finale del pilota, e zero o più problemi riscontrati durante la gara. Un modello può essere stato impiegato in una o più gare.

(Punteggio pieno per schemi E-R corretti che contengono solo associazioni binarie. Se lo schema è corretto ma si è fatto uso di associazioni n-arie, 5 punti.)

Soluzione



Esercizio 5 (4 punti)

Tradurre lo schema E-R ottenuto nell'esercizio precedente in uno schema logico del modello relazionale.

Soluzione

La generalizzazione relativa al personale può essere eliminata accorpendo l'entità genitore nelle figlie, in assenza di altre informazioni relative alle operazioni.

PROGETTISTA(Codice, Nome, Cognome, Età, TitoloDiStudio)

PILOTA(Codice, Nome, Cognome, Età, Ingaggio)

MECCANICO(Codice, Nome, Cognome, Età, Pilota)

con vincolo di integrità referenziale fra l'attributo Pilota e l'attributo Codice della relazione PILOTA

MODELLO(Codice, Nome, Cilindrata)

PROGETTAZIONE(Modello, Progettista)

con vincoli di integrità referenziale fra l'attributo Modello e l'attributo Codice della relazione

MODELLO e fra l'attributo Progettista e l'attributo Codice della relazione PROGETTISTA

GARA(Nome, Anno, Autodromo)

PARTECIPAZIONE(Pilota, Gara, Anno, Modello, PosizioneIniziale, PosizioneFinale)

con vincoli di integrità referenziale fra l'attributo Pilota e l'attributo Codice della relazione PILOTA, fra

gli attributi Gara e Anno e gli attributi Nome e Anno della relazione GARA e fra l'attributo Modello e

l'attributo Codice della relazione MODELLO

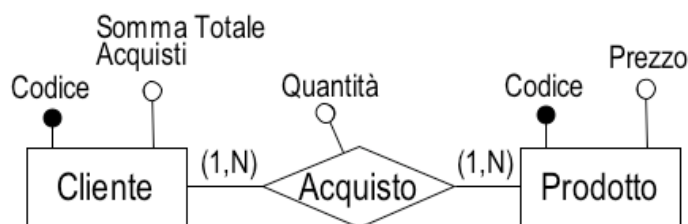
PROBLEMI(Pilota, Gara, Anno, Problema)

con vincoli di integrità referenziale fra l'attributo Pilota e l'attributo Codice della relazione PILOTA e

fra gli attributi Gara e Anno e gli attributi Nome e Anno della relazione GARA

Esercizio 6 (4 punti)

Dato il seguente schema E-R, con il seguente volume dei dati e le seguenti operazioni, decidere se è conveniente conservare nello schema l'attributo derivato SommaTotaleAcquisti, trascurando l'occupazione di memoria di tale dato.



OP 1: Dati i codici di un cliente e di un prodotto già esistenti, inserire l'acquisto del prodotto da parte del cliente

OP 2: Visualizzare i dati di un cliente

Tavola dei volumi

Concetto	Tipo	Volume
Prodotto	E	600
Cliente	E	300
Acquisti	R	12000

Tavola delle operazioni

Operazione	Tipo	Frequenza
OP 1	I	400/Giorno
OP 2	I	10/Giorno

Soluzione

Nota: consideriamo il costo di un accesso in lettura pari a 1 e il costo di un accesso in scrittura pari a 2

Con l'attributo ridondante

Costo di OP 1: 1 accesso in scrittura su Acquisti, 1 accesso in lettura su Prodotto (lettura Prezzo del prodotto acquistato), 1 accesso in lettura su Cliente (lettura del valore corrente di SommaTotaleAcquisti), 1 accesso in scrittura su Cliente (aggiornamento SommaTotaleAcquisti) = 6

Costo di OP 2: 1 accesso in lettura su Cliente (per codice e SommaTotaleAcquisti)

Costo medio totale al giorno: $2400 + 10 = 2410$

Senza l'attributo ridondante

Costo di OP 1: 1 accesso in scrittura su Acquisti

Costo di OP 2: 1 accesso in lettura su Cliente, 40 ($12000/300$) accessi in lettura su Acquisti, 40 accessi in lettura su Prodotto (lettura del prezzo dei prodotti acquistati) = 81

Costo medio totale al giorno: $800 + 810 = 1610$

Conviene eliminare l'attributo ridondante

Esercizio 7 (3 punti)

Illustrare sinteticamente quali tipi di conoscenza possono essere estratti dalle basi di dati per mezzo di tecniche di Data Mining