

## **FUNDAMENTOS DE BASES DE DATOS**

### **Examen Diciembre 2003**

#### **Presentar la resolución del examen:**

- Con las hojas numeradas y escritas de un solo lado.
- Con la cantidad de hojas entregadas en la primer hoja.
- Con cédula de identidad y nombre en cada hoja.
- Escrita a lápiz y en forma prolija.

### **Parte 1. Modelo Entidad-Relación (30 puntos)**

#### **Ejercicio 1. (30 pts)**

Se quiere modelar la realidad relativa a una clínica odontológica. La clínica está compuesta por varios locales de atención, identificados por su nombre, de los cuales se conoce además su dirección dada por la ciudad donde se ubica, la calle y el número.

En cada local existen varios consultorios que se identifican por un número dentro del local y en cada consultorio existe cierto equipamiento. Dicho equipamiento se identifica globalmente mediante un número de serie, se conoce el tipo (torno, laser, etc.) e interesa mantener registro de la última fecha en que se le realizó mantenimiento.

La clínica posee dos planes diferentes de afiliación: individual y grupal. De los afiliados se conoce la CI, el nombre y uno o más teléfonos. Para los afiliados grupales interesa saber el nombre del convenio de afiliación y el porcentaje de rebaja que se debe aplicar a la cuota mensual.

En la clínica se realizan tratamientos, los cuales se identifican por su nombre y tienen un costo asociado.

Los odontólogos que trabajan en la clínica se identifican por su nombre. De ellos se conoce su especialidad principal dentro de la odontología y los diferentes tratamientos que pueden realizar.

Los odontólogos trabajan en diferentes locales y cada odontólogo puede tener distintos horarios de atención en cada local. De cada horario de atención se conoce el día de la semana, la hora de comienzo y la hora de finalización. (EJ: lunes de 16:00 a 18:30).

Los afiliados se atienden con determinados odontólogos en determinado local y además los odontólogos les realizan tratamientos. Para que un paciente pueda recibir tratamiento de un odontólogo debe ser previamente atendido por este.

Interesa mantener la historia clínica de cada afiliado, la cual consiste, por un lado, en un registro de cada consulta indicando la fecha de consulta, el odontólogo y el local y por otro en un registro de todos los tratamientos que se le han realizado. De cada tratamiento interesa saber: fecha de inicio, si el tratamiento ha sido finalizado o no y la identificación del odontólogo que lo realizó, teniendo en cuenta que como política de la clínica un odontólogo sólo puede practicar un tratamiento por vez a cada afiliado. El inicio de un tratamiento, siempre se hace en una consulta.

Para cada afiliado se mantiene una cuenta corriente donde se incluyen los costos de todos los tratamientos que han sido finalizados. Esta cuenta corriente es global a la clínica.

**SE PIDE: Modelo Entidad Relación completo, incluyendo restricciones de integridad.**

## Parte 2. Diseño Relacional (20 puntos)

### Ejercicio 2 (10 pts)

**Nota:** En todos los casos se debe justificar la respuesta.

Sea  $R(A,B,C,D,E,G,H,I)$  y  $M = \{ GH \rightarrow AB, C \rightarrow AG, D \rightarrow EB, E \rightarrow HI, B \rightarrow CD, A \rightarrow I, G \rightarrow DE \}$  el conjunto de dependencias sobre  $R$ .

- Hallar 3 claves distintas de  $R$  según  $M$ .
- ¿ $R$  se encuentra en BCNF según  $M$ ?
- Sea  $\rho = \{R_1(A,B,C,D), R_3(B,E,G,H,I)\}$  una descomposición de  $R$ 
  - ¿Es  $\rho$  una descomposición con join sin pérdida según  $M$ ?
  - ¿ $\rho$  preserva la dependencia  $GH \rightarrow A$ ?
  - ¿ $\rho$  preserva la dependencia  $C \rightarrow G$ ?
- Sea  $F = M - \{ A \rightarrow I, G \rightarrow DE \}$ . ¿Es  $F$  un conjunto minimal?
- Hallar una descomposición de  $R$  en 4NF con join sin pérdida según  $M$ .

### Ejercicio 3 (10 pts)

#### Parte A

La persona encargada del diseño de la Base de Datos de una empresa aplicó el siguiente algoritmo a un esquema  $R$  y conjunto de dependencias funcionales  $F$ :

- Para cada miembro izq  $X$  de una  $df$  que aparezca en  $G$  crear un  $er \{X, A_1, A_2, \dots, A_m\}$ , donde  $X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_m$  sean las únicas  $dfs$  en  $G$  con  $X$  como miembro izq;
- Colocar cualquier atributo restante de  $R$  en un solo  $er$ .

Esta persona afirma que la descomposición resultante de este algoritmo es en 3NF con preservación de dependencias.

Indicar si esta afirmación es cierta o no, justificando en cada caso.

#### Parte B

Dado  $R(A,B,C,D,E,G)$  y el conjunto de dependencias sobre el  $F = \{ G \rightarrow A, BC \rightarrow G, D \rightarrow B, AE \rightarrow C, B \rightarrow EA, A \rightarrow B \}$

- Hallar una descomposición de  $R$  en 3NF con join sin pérdida y preservación de dependencias, aplicando los algoritmos vistos en el curso. (justificar los pasos dados).
- Verificar si la descomposición anterior se encuentra en BCNF. Justificar la respuesta.

## Parte 3. Consultas (30 puntos)

### Ejercicio 4 (30 pts)

Al finalizar el año lectivo, los estudiantes llenan un conjunto de encuestas sobre los cursos y los docentes. Estas encuestas son procesadas y los resultados resumidos son almacenados en una base de datos. El siguiente esquema corresponde a los resúmenes de las encuestas realizadas en este año lectivo:

#### **CURSOS (cod\_curso, nom\_curso, duracion)**

Contiene los datos de los cursos considerados.

#### **CARRERA\_CURSOS (cod\_curso, cod\_carrera, semestre, tipo)**

Contiene los datos de los cursos por carrera. El atributo *tipo* puede tomar el valor "obligatorio" o el valor "opcional".

#### **ASIGNACION (cod\_curso, nro\_docente, tipo\_tarea)**

Contiene los datos de la asignación de los docentes a los cursos. El atributo *tipo\_tarea* puede ser "teórico", "práctico", "laboratorio", "otros". Observar que un mismo docente puede estar asignado a varias tareas dentro de un mismo curso.

#### **ENC\_CURSO (cod\_curso, prom\_material, prom\_evaluacion, prom\_global, cant\_enc)**

Contiene los datos de las encuestas relativos a los cursos. El atributo *prom\_material* es el promedio obtenido en los puntos de la encuesta correspondientes al material. El atributo *prom\_evaluación* es análogo al anterior con respecto a la evaluación. El atributo *cant\_enc* contiene, para cada curso, la cantidad de estudiantes que llenaron encuestas.

#### **ENC\_DOCENTE (cod\_curso, nro\_docente, prom\_actitud, prom\_metodologia, cant\_enc)**

Contiene los datos de las encuestas relativos a los docentes en los cursos. El atributo *prom\_actitud* es el promedio obtenido en los puntos de la encuesta correspondientes a la actitud del docente. El atributo *prom\_metodología* es análogo al anterior con respecto a la metodología del docente en el curso. El atributo *cant\_enc* contiene, para cada docente en cada curso, la cantidad de estudiantes que llenaron encuestas.

**NOTA:** Suponga que en todos los cursos al menos un estudiante llenó una encuesta para el curso y para cada docente.

Resolver las siguientes consultas en Álgebra Relacional:

- Obtener los códigos de los cursos tales que todos los docentes asignados a ellos tuvieron un promedio\_actitud superior a 3, en ese curso.
- Obtener el número de los docentes asignados a todos los cursos opcionales del semestre 2 de la carrera 72.

Resolver las siguientes consultas en Cálculo Relacional:

- Obtener las parejas *nom\_curso*, *nro\_docente* tales que el valor del atributo *prom\_metodología* es el más alto de las encuestas.
- Obtener los códigos de curso tales que todos sus docentes están asignados a él con un solo tipo de tarea, es decir, que hay docentes que están asignados a práctico y otros docentes a teórico, pero no hay docentes en el curso que a la vez estén asignados a práctico y teórico.

Resolver las siguientes consultas en SQL:

- e) Para cada carrera, por semestre obtener el promedio obtenido en sus cursos en `prom_global`, considerando solamente las encuestas de más de 100 estudiantes.
- f) Obtener el número de los docentes tales que obtuvieron el valor máximo de la encuesta en `prom_actitud` para algún curso y el valor mínimo registrado en la encuesta en `prom_actitud` para otro curso.

## Parte 4 Optimización y Concurrencia (20 puntos)

### Ejercicio 5. (20 pts.)

Dados el esquema relacional del **Ejercicio 4** y la siguiente consulta

```
SELECT C.cod_curso, C.duración
FROM Cursos C, Enc_Docente E, Asignación A
WHERE A.cod_curso = E.cod_curso AND A.nro_docente = E.nro_docente, AND
C.cod_curso = A.cod_curso AND A.tipo = "teorico" AND E.prom_actitud > 3
```

- 1) Dar un plan lógico de la consulta, aplicando las heurísticas y calculando los tamaños intermedios.
- 2) Elegir una implementación para la primer selección y el primer join que se realice según el plan lógico, y calcular su costo estimado. (En el costo se deben incluir los costos de grabar los resultados.)

#### DATOS:

	CURSOS	ASIGNACION	ENC_DOCENTE
Cantidad tuplas	300	2 500	100 000
Indices primarios	cod-curso (niveles: 2)		(cod-curso, nro- docente) (niveles: 1)
Indices secundarios (B+)		tipo-tarea (niveles: 3)	
Cantidad de tuplas por bloque	50	60	30
Observaciones		- En promedio, se asignan a teorico 2 docentes por curso.	- En prom-actitud los valores van de 1 a 5, y su distribución es uniforme.

	CURSOS  ><  ASIGNACION	ASIGNACION  ><  ENC_DOCENTE
Cantidad de tuplas por bloque	30	20

Fórmulas para cálculo de costo y de tamaños:

*Notación:* b – cantidad de bloques

fbl – factor de bloqueo

x – cantidad de niveles del índice

|R| - cantidad de tuplas de R

s – cardinalidad de la selección (tamaño obtenido)

js – selectividad del join

V(A,R) - cantidad de valores distintos del atributo A en R

<b>Selección (R)</b>	Búsqueda lineal	B
	Búsqueda binaria	$\log_2 b + \lceil s / fb_{lR} \rceil - 1$
	Índice primario	$x + 1$
	Índice secundario (B+)	$x + s$
<b>Join (R,S)</b>	Nested Loop (ciclo anidado) sin utilizar índices	$b_R + b_R * b_S$
	Nested Loop (ciclo anidado) utilizando índice secundario para recuperar tuplas que matchean	$b_R +  R  * (x + s)$
	Nested Loop (ciclo anidado) utilizando índice primario para recuperar tuplas que matchean	$b_R +  R  * (x + 1)$
<b>Escribir resultados en disco</b>	Join (R,S)	$(js *  R  *  S ) / fb_{lRS}$
<b>Selectividad</b>	Selección (atrib. A)	$1 / V(A,R)$
	Join (atrib. A)	$1 / \text{Max}(V(A,R), V(A,S))$