

FUNDAMENTOS DE BASES DE DATOS

Examen Febrero 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 (25 puntos)

Ejercicio 1.

La “Corporación Turística Latinoamericana Inc.” es una asociación de agencias de viajes que provee diferentes servicios a sus socios. Entre otros servicios, la asociación registra los paquetes que ofrecen los mayoristas mundiales considerando su composición (tantos días en tal ciudad de tal país, en tal hotel con tales paseos, etc.).

De los mayoristas se conoce una identificación, su nombre, sus teléfonos, la dirección, al menos un número de fax y los idiomas que hablan. De las agencias de viajes asociadas, se conoce una identificación, un nombre, sus teléfonos y su dirección (sin incluir la ciudad o país). Los mayoristas ofrecen paquetes que tienen un nombre que los identifica (Ej. Europa Clásica). Los diferentes mayoristas pueden ofrecer paquetes con el mismo nombre.

Las agencias de viajes comercializan determinados paquetes. Toda agencia comercializa al menos un paquete aunque podrían existir paquetes que no fueran comercializados por nadie. Para cada paquete que es comercializado la agencia organiza salidas desde su base registrando para cada salida, la fecha, la hora, la cantidad de lugares totales, la cantidad de lugares ocupados y el precio por persona base single y doble. Una agencia puede vender lugares a otra agencia en una salida. Interesa registrar en estos casos cuantos lugares fueron comprados por cada agencia en cada salida, teniendo en cuenta que no está permitida la sobreventa, o sea no se pueden vender mas lugares de los existentes.

Cada paquete está conformado por un conjunto de hospedajes. De cada hospedaje se conoce un identificador, el nombre, la dirección y los teléfonos. Los hospedajes puede ser hoteles, hosterías o albergues. Si son hoteles, se conoce la cantidad de estrellas, los horarios de entrada y de salida. Si son hosterías, se conoce la hora de servicio de desayuno, almuerzo y cena. Un paquete no sólo incluye la información referente a que hospedajes se visitarán, sino también cuantos días se reservan en cada uno y en que orden serán visitados. Hay que tener en cuenta que los pasajeros que toman cierto paquete pueden pasar más de una vez por el mismo hospedaje.

Los hospedajes se encuentran en ciudades de las que se conoce su nombre, la dirección del ayuntamiento (intendencia), la lista de países que tienen consulado en esa ciudad, la característica telefónica de la ciudad y una descripción de los principales atractivos. Cada ciudad se encuentra en una determinada región. De cada región, se conocen su nombre, el tipo (departamento, provincia, región, etc) y una descripción de sus atractivos turísticos. Cada región está en un país determinado del que se conoce el nombre (que lo identifica), su característica telefónica y su ciudad capital. Hay que tener en cuenta que en diferentes países pueden existir regiones con el mismo nombre y en diferentes regiones, ciudades con el mismo nombre. Interesa registrar en qué ciudad se encuentra cada agencia de viajes o mayorista.

Se pide: Modelo Entidad-Relación completo.

Parte 2 – Diseño Relacional (25 puntos)

Ejercicio 2.

Una emergencia médico móvil administra en forma manual el despacho de ambulancias: Cada vez que suena el teléfono de emergencias, una recepcionista toma la llamada, elige una ambulancia a la que le pasa el llamado, y llena una tarjeta. En esta tarjeta se indica el día, la recepcionista y la ambulancia a la que se le comunicó el llamado. Además existen otros registros (también manuales) que indican la conformación de la ambulancia en cuanto al equipamiento y personal para cada día.

La emergencia móvil cuenta con un conjunto de ambulancias. De cada ambulancia se posee una ficha con la matrícula que la identifica y su capacidad. En la ficha de la ambulancia, también aparece la lista de los aparatos con que está equipada incluyendo, para cada aparato, un nombre, marca modelo y antigüedad. El nombre del aparato es identificatorio dentro de una ambulancia, pudiendo haber dos aparatos con el mismo nombre en dos ambulancias distintas

Actualmente, existen 3 registros independientes de médicos, recepcionistas y choferes. Es una política de la empresa que ningún funcionario tenga más de una función. Dicho de otra forma, nunca puede aparecer una misma persona en dos de esos registros. De cualquier miembro del personal se conoce su nombre, cédula de identidad que lo identifica, domicilio y teléfono. De los recepcionistas se conoce su horario de trabajo (hora de inicio y hora de fin). De los choferes se conoce su edad y la fecha en la cual sacaron la libreta de conducir. De los médicos se conoce su especialidad y una lista de las mutualistas en las que trabajan.

Todos los años se hacen asignaciones tanto de médicos como de choferes a las ambulancias. Para cada fecha del año y para cada ambulancia se asigna un chofer y un médico. Tanto un chofer como un médico puede estar asignado a lo sumo a una ambulancia en una fecha dada, y por otro lado, puede estar asignado a una misma ambulancia en varias fechas distintas.

Se Pide:

- Diseño relacional hasta 3NF, justificando por que esta en esa forma normal. Incluir todas las dependencias funcionales y multivaluadas que se encuentren.
- A partir del diseño resultante de la parte A, obtener un diseño en 4NF con join sin pérdida.

Ejercicio 3.

Dado el esquema R (A, B, C, D, E, F, G, H)

y las dependencias

A → F
C → B D E H
H → D

C →→ G
C →→ AF

responder las siguientes preguntas.

- ¿Cuáles son claves candidatas de la tabla?
- ¿En que forma normal está el esquema? Justificar.
- Pasar el esquema a 4NF con join sin perdida.
- Indique que dependencias se perdieron en el proceso.

Parte 3 – Consultas (25 puntos)

Ejercicio 4.

Un canal de televisión desea obtener estadísticas en base a la información que dispone respecto a la programación del canal, los empleados que allí trabajan y las preferencias de algunos televidentes encuestados. Todos los programas emitidos por el canal tienen nombres diferentes, pero un mismo programa puede ser emitido un solo día o varios días a la semana y eventualmente en diferentes horarios. Existen varios géneros, por ejemplo “película”, “infantil”, etc que son utilizados para clasificar los programas.

Se tiene una base de datos con las siguientes relaciones:

- **PROGRAMAS (nomProg, genero)**

Representa los programas emitidos por el canal.

- **PROGRAMACIÓN (nomProg, dia, horaComienzo, horaFin)**

Representa la programación del canal. Para cada programa se indican los días de la semana en que se emite y las horas de comienzo y fin para cada día.

- **EMPLEADOS (nomEmp, ciEmp, dirEmp, telEmp, genero, cargo, horaEntrada, horaSalida)**

Representa a los empleados del canal. De ellos se conoce el nombre, la cédula, la dirección, el teléfono, el género de programas televisivos en los que trabaja, y su cargo (por ejemplo “encargado”, “asistente”, etc). Además se conoce el horario que cumple, el cual es fijo. El horario de un empleado no coincide necesariamente con el horario de los programas en los que trabaja.

- **TELEVIDENTES (nomTele, ciTele, dirTele, telTele, barrio, edad, sexo, ocupacion)**

Representa a los televidentes encuestados. De ellos se conocen el nombre, la cédula, la dirección, el teléfono, el barrio donde vive, la edad, el sexo y la ocupación.

- **PREFERENCIAS (nomProg, dia, HoraComienzo, ciTele)**

Representa las preferencias de los televidentes encuestados. Se tiene registro, para cada persona encuestada, de aquellos programas que habitualmente mira y los días de la semana y horarios en que los mira. Se registran sólo aquellos televidentes que miran todo el programa.

SE PIDE :

1) Escribir en álgebra relacional las siguientes consultas:

- a) Devolver las CI de los televidentes que miran todas las emisiones de “Gran Hermano”.
- b) Devolver el nombre de los programas que son emitidos Lunes y Viernes, y no son vistos por nadie el día Lunes o no son vistos por nadie el día Viernes

2) Escribir en cálculo relacional las siguientes consultas:

- c) Devolver la CI y el nombre de los empleados que trabajan con un género de programa y que durante su horario de trabajo NO se emiten programas de dicho género.
- d) Devolver los nombres de los televidentes que los jueves sólo miran programas de género “película”.

3) Escribir en SQL las siguientes consultas, teniendo en cuenta que si es necesario se pueden utilizar vistas o tablas temporales.

- e) Devolver cantidad de televidentes, por ocupación, que miran “TeleInforma” los Jueves.
- f) Devolver los nombres de los programas más vistos el día Lunes.

Parte 4 – Optimización, Concurrencia y Recuperación (25 puntos)

Ejercicio 5.

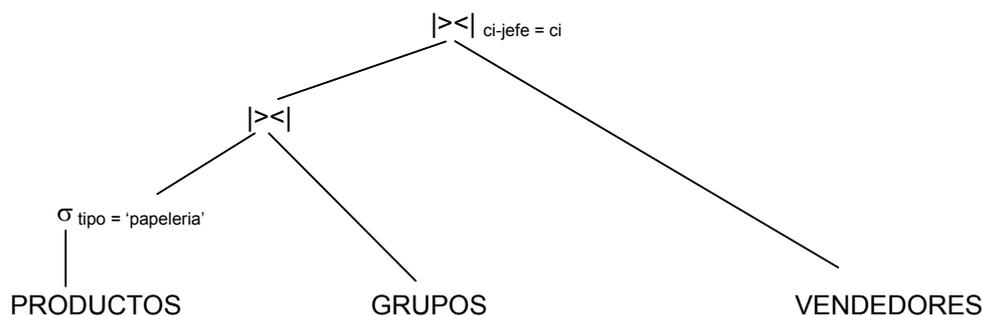
Dado el siguiente esquema relacional:

Productos (tipo, nroprod, nrogrupo)

Grupos (nrogrupo, ci-jefe)

Vendedores (ci, nrogrupo, salario)

y el siguiente plan lógico correspondiente a una consulta:



Y considerando además que:

El resultado de la selección se debe guardar en una tabla temporal Temp1 y el resultado del primer join se debe guardar en una tabla temporal Temp2.

- 1) Dar el plan físico que le parezca mejor para ese plan lógico.
- 2) Calcular los tamaños (cantidad de tuplas) de Temp1 y Temp2.
- 3) Calcular el costo total del plan. (Nota: No olvidar sumar los costos de grabar Temp1 y Temp2)

DATOS:

	PRODUCTOS	GRUPOS	VENEDORES	TEMP2
Cantidad tuplas	2000	50	10000	
Indices primarios			ci (niveles: 2)	
Indices secundarios (B+)	tipo (niveles: 2)		salario (niveles: 3)	
Cantidad tuplas por bloque	20	10	5	5
Observaciones	- Tipos distintos: 200			

Fórmulas para cálculo de costo y de tamaños:

Selección	Búsqueda lineal.	B
	Con índice primario.	$x + 1$
	Con índice secundario.	$x + s$
Join (R,S)	Nested Loop (ciclo anidado) sin utilizar índices	$b_R + b_R * b_S$
	Nested Loop (ciclo anidado) utilizando índice primario para recuperar tuplas que matchean	$b_R + R * (x + 1)$
Selectividad	Selección (atrib. A)	$1 / V(A,R)$
	Join (atrib. A)	$1 / \text{Max}(V(A,R), V(A,S))$

Notación: **b** – cantidad de bloques
x – cantidad de niveles del índice
|R| – cantidad de tuplas de R

s – cardinalidad de la selección
V(A,R) – cantidad de valores distintos del atributo A en R

Ejercicio 6.

1. Enumere y explique las 4 propiedades más importantes de las transacciones en los sistemas de bases de datos.
2. Indique qué problemas tienen los siguientes planes.

P1)

	T1	T2
t1	read(a)	
t2		read(a)
t3	write(a)	
t4		write(a)

P2)
Valores Iniciales: C1 = 40, C2 = 50, C3 = 30.

	T1	T2
t1	read(C1)	
t2	suma := suma + C1	
t3	read(C2)	
t4	suma := suma + C2	
t5		read(C3)
t6		C3 := C3 - 10
t7		write(C3)
t8		read(C1)
t9		C1 := C1 + 10
t10		write(C1)
t11	read(C3)	
t12	suma := suma + C3	

Explique el protocolo que considere mejor para evitar estos problemas.
Muestre planes nuevos para P1 y P2 usando el protocolo propuesto.
Indique si aparecen nuevos problemas.

3. Sean el siguiente plan de las transacciones T1, T2 y T3:

	T1	T2	T3
t1	V1 := Read(a)		
t2	V1 := V1 + 1		
t3		V2 := Read(a)	
t4		V2 := V2 x 2	
t5			V3 := Read(a)
t6			V3 := V1 + 1
t7			Write(a, 1)
t8		Write(a, V2)	
t9	Write(a, V1)		

Notación: Write(i,v) significa escribir en el ítem i el valor v

Indique si este plan producirá en cualquier caso (independiente de los valores de a) un valor correcto.
¿Cómo sabe la respuesta sin ejecutar el plan?

4. Indique qué consecuencias tienen en la recuperación los siguientes actos:
 - a) Nunca escribe físicamente los buffers en la base de datos antes del commit.
 - b) Fuerza los buffers hacia la base de datos en el momento del commit.