

Segundo Parcial de Fundamentos de Base de Datos

Noviembre 2004

Presentar la resolución del parcial:

- 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 I. DISEÑO RELACIONAL. (35 puntos)

Ejercicio 1 (9 puntos)

- 1) ¿Qué es una dependencia funcional? Defina.
- 2) Sea un esquema de relación $R(A,B,C)$. A partir de la siguiente instancia válida indicar si se cumplen o no las siguientes dependencias funcionales en el esquema R . **Justificar** la respuesta.

| | A | B | C |
|---------|-------|-------|-------|
| tupla 1 | a_1 | b_1 | c_1 |
| tupla 2 | a_2 | b_3 | c_1 |
| tupla 3 | a_1 | b_1 | c_1 |
| tupla 4 | a_2 | b_2 | c_3 |

- a) $A \rightarrow B$
- b) $B \rightarrow C$
- c) $C \rightarrow A$

3) Dependencias Multivaluadas.

Consideremos un fragmento de lo que intenta ser una descripción de la realidad de los docentes universitarios:

"...Un docente dirige a varios estudiantes. Además, este docente, participa en varios proyectos de investigación. Los proyectos de investigación no están relacionados con los estudiantes que el docente dirige..."

Supongamos que se cuenta con el siguiente esquema para modelar los aspectos de la realidad planteados.

INFO (CIDocente, CIEstudiente, NroProyecto)

Si sabemos que cada docente se identifica por su cédula (CIDocente), que cada estudiante se identifica por su cédula (CIEstudiente) y que cada proyecto se identifica por un número (NroProyecto) podemos decir que en esta realidad se cumplen las siguientes DMV:

CIDocente \twoheadrightarrow CIEstudiente
CIDocente \twoheadrightarrow NroProyecto

- a) Dar una instancia válida del esquema INFO que evidencie la existencia de las DMV planteadas.
- b) ¿Qué inconvenientes plantea el esquema INFO?
- c) Sugerir alguna forma de evitar los inconvenientes planteados en la parte b.

Ejercicio 2 (16 puntos)

Un artesano zapatero confecciona calzado para niños, a medida y por encargo y lo envía. De los niños conoce su CI, nombre, ciudad y número de calzado.

El zapatero confecciona varios modelos de calzado, para lo que utiliza diferentes materiales y herramientas. Las herramientas utilizadas son independientes de los materiales del calzado.

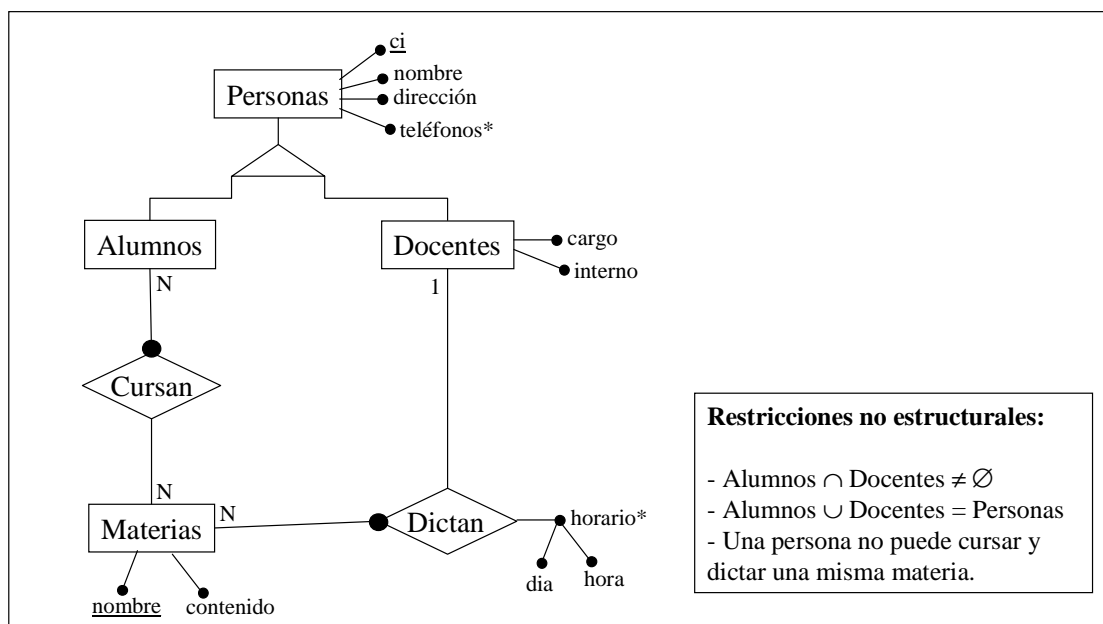
El zapatero lleva registro de sus ventas. A la venta de cada par de calzado le asigna un código único de venta, y registra además, el modelo, color y número del calzado, el niño, la fecha y el precio. El precio de la venta es una función del modelo del calzado, la ciudad donde vive el niño (por los gastos de envío) y la fecha.

Se pide:

1. Deducir dependencias funcionales y multivaluadas. Utilizar los siguientes nombres de atributos: modelo, color, numero, material, herramienta, CI, nombre, ciudad, fecha, cod-venta y precio.
2. Calcular todas las claves. Justificar.
3. Indicar las dependencias que se proyectan en los siguientes sub-esquemas:
 - R1 (CI, nombre, ciudad, numero)
 - R2 (modelo, material, herramienta)
 - R3 (cod-venta, CI, modelo, color, numero, fecha, precio)
4. Indicar en qué forma normal están dichos sub-esquemas. Justificar.
5. Indicar qué dependencias funcionales se preservan y cuáles no.
6. Llevar los sub-esquemas anteriores a 4NF aplicando los algoritmos de normalización.

Ejercicio 3 (10 puntos)

Pasar el siguiente Esquema Entidad-Relación a un Esquema Relacional especificando: relaciones, claves y dependencias de inclusión.



PARTE II. OPTIMIZACIÓN DE CONSULTAS. (15 puntos)

Ejercicio 4 (6 puntos)

Las siguientes formulas de cálculo de costo son para distintas implementaciones de la operación de selección σ_{cond} de R , cuyo resultado es la relación S .

- a) $x + 1$
- b) $x + b_S$
- c) $x + |S|$

En las fórmulas anteriores x representa la cantidad de niveles del índice utilizado, b_S representa la cantidad de bloques del resultado de la selección, y $|S|$ representa la cantidad de tuplas del resultado de la selección.

Explicar a qué implementación corresponde cada una y por qué son sus diferencias.

Ejercicio 5 (9 puntos)

Las siguientes tablas contienen información sobre los estudiantes de facultad.

Estudiantes(CiEst, Nom, Fnac)

Contiene información de los estudiantes: Cédula de identidad (CiEst), nombre completo (Nom) y fecha de nacimiento (Fnac).

Curso_Insc(CiEst, IdCur, Ai)

Contiene información de los cursos en los cuales se inscribió un estudiante: Cedula del estudiante (CiEst), identificación del curso (IdCur), Año de su última reglamentación (Ai)

Exámenes(CiEst, IdCur, Fecha, Nota)

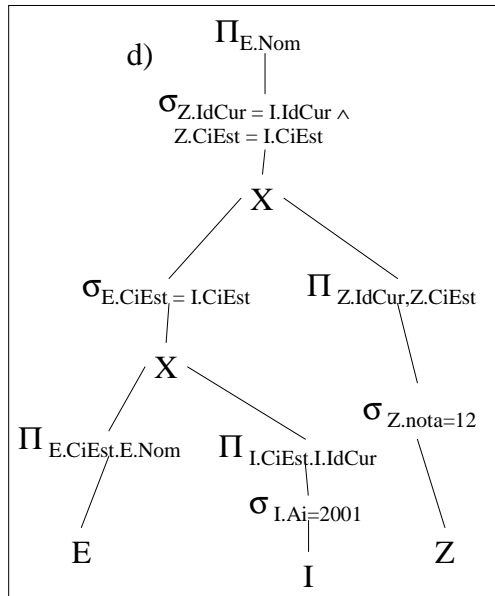
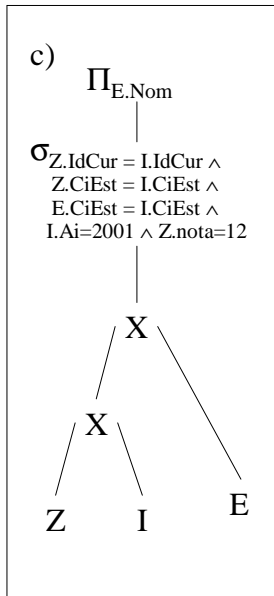
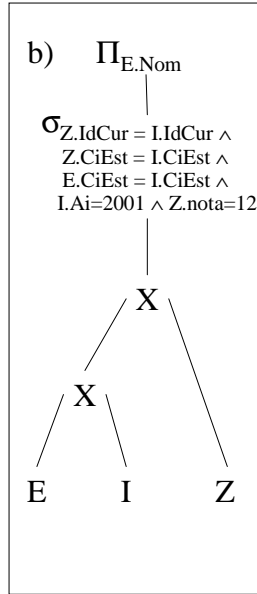
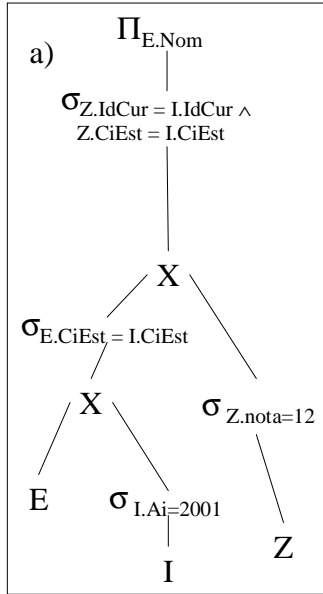
Contiene información de los exámenes rendidos por cada estudiante: Cédula del estudiante (CiEst), identificación del Curso (IdCur), fecha (Fecha) en que el estudiante rindió el examen de ese curso, y la nota (Nota) que obtuvo. En esta tabla sólo hay información de los estudiantes que efectivamente se presentaron a cada examen.

NOTA: En esta tabla sólo hay datos de exámenes rendidos por estudiantes que se inscribieron al curso correspondiente a dicho examen en el año 2001.

Considere la siguiente consulta en SQL:

```
Select E.Nom
From Exámenes Z, Curso_Insc I, Estudiantes E
Where Z.IdCur = I.IdCur and
      Z.CiEst = I.CiEst and
      E.CiEst = I.CiEst and
      I.Ai = 2001 and
      Z.nota = 12
```

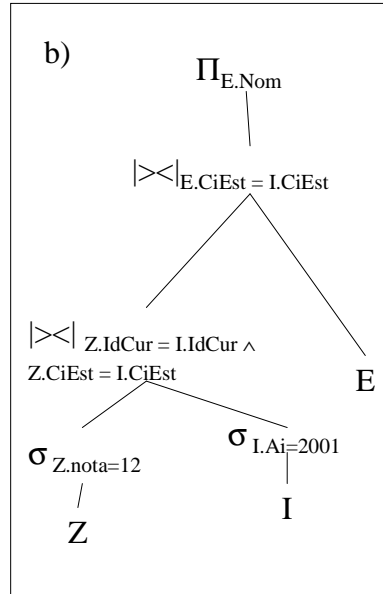
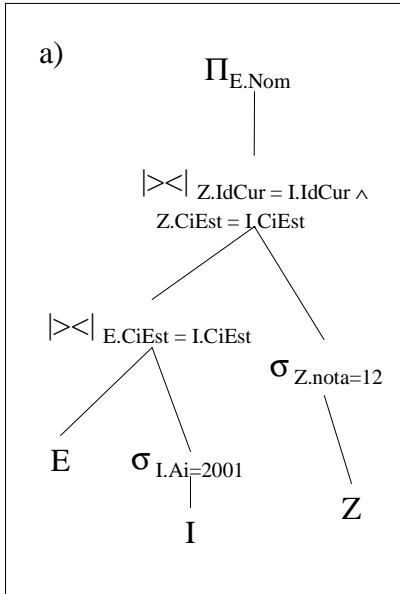
a) Indique cual de los siguientes árboles canónicos es el correcto.



b) Disponiendo de la siguiente información:

- Hay 10000 Estudiantes distintos.
- En el año 2001 estuvieron activos solo 1000 estudiantes, los cuales se inscribieron a 5 cursos cada uno.
- Solo un 10% de los exámenes se aprueba con la máxima calificación (12).
- La tabla de exámenes contiene 10000 tuplas.

(i) Estime los tamaños de los resultados (en cantidad de tuplas) de todas las operaciones en cada uno de los siguientes planes lógicos.



(ii) Indique cual de los anteriores planes lógicos es el mejor, justificando.

PARTE III. CONCURRENCIA Y RECUPERACION. (10 puntos)

Ejercicio 6 (10 puntos).

Es necesario migrar un programa sobre una base relacional con un scheduler capaz de generar los locks automáticamente a una en la que los locks se deben poner a mano en el programa. En el sistema, sólo se ejecuta este programa todas las veces que se requiera en forma concurrente. El pseudocódigo del programa es el siguiente:

```

Transfer(C_origen: int, C_destino: int, Monto: real)
SaldoO, SaldoD: real;
1. Begin
2. SaldoO=0; /* Inicializamos Variables */
3. SaldoD=0;
4. Begin Transaction /* Iniciamos el modo transaccional */
5.   read(C_origen, SaldoO); /* Leemos el saldo de la c. Origen .. */
6.   if (read_ok<>0) /* si la c.origen no existe, abortamos */
7.   then
8.     Abort;
9.   endif
10.  read(C_destino, SaldoD); /* Leemos la c. destino.... */
11.  if (read_ok<>0) /* si la c. origen no existe, abortamos */
12.  then
13.    Abort;
14.  endif
15.  if (SaldoO < Monto) /* Si no hay suficiente c. origen, abortamos */
16.  then
17.    Abort;
18.  endif
19. /* Como todo está bien, actualizamos las cuentas ... */
20.   SaldoO=SaldoO-Monto;
21.   write(C_origen,SaldoO);
22.   SaldoD=SaldoD+Monto;
23.   write(C_destino,SaldoD);
24.   Commit;
25. End Transaction;
26. End

```

Se Pide:

Dada la siguiente tabla

| La Instrucción debe ubicarse entre las líneas: | No sigue 2pl básico | 2pl básico | 2pl conservador | 2pl estricto | 2pl riguroso |
|--|---------------------|------------|-----------------|--------------|--------------|
| r_lock(C_origen) | 4-5 | | | | |
| r_lock(C_destino) | 9-10 | | | | |
| W_lock(C_origen) | 20-21 | | | | |
| W_lock(C_destino) | 22-23 | | | | |
| unlock(C_origen) | 21-22 | | | | |
| unlock(C_destino) | 23-24 | | | | |

Completarla de forma que en cada celda queden los números de línea (según el listado anterior) entre los que debe ubicarse la instrucción de la fila indicada, para que las transacciones resultantes de la ejecución del programa anterior sigan **exclusivamente** los protocolos indicados por las columnas (o sea no siga ninguno de los protocolos siguientes). A modo de ejemplo se presenta completa la primer columna.