

FUNDAMENTOS DE BASES DE DATOS

Examen Febrero 2019

La duración del examen es de 3 horas y media.

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.
- Comenzando cada ejercicio en una nueva hoja

Ejercicio 1 (25 puntos).

Se desea modelar la realidad correspondiente a ferias de libros que se realizan en cierto país. Cada feria tiene un número identificador, un nombre, un costo para la entrada de ingreso, una fecha de inicio, una fecha de fin y la dirección donde se realiza el evento. De esta última se conoce la ciudad, la calle y el número de puerta.

En cada feria se realizan presentaciones. Cada una de las presentaciones se identifica por el número de feria al cual pertenece, conjuntamente con la fecha y la hora en la que se realiza. Además, se conoce la duración de la misma. Hay dos tipos de presentaciones, de libros y tertulias literarias. De las primeras interesa saber qué libro se presentó (en cada presentación se presenta uno sólo) y de las últimas interesa saber los temas tratados.

Los libros que se presentan en las ferias se identifican por su código (ISBN). Tienen un resumen, cantidad de páginas, nombre del editor, género e idioma. Un libro puede ser escrito por varios autores y un autor puede escribir muchos libros. Se sabe que todos estos autores escribieron al menos un libro que es presentado en alguna feria. Cada libro es presentado por uno solo de los autores y una sola vez en una única feria. Sólo se modelan libros presentados en alguna feria.

Hay dos tipos de personas que están relacionadas con las ferias y/o los libros: autores e invitados especiales. Tanto de los autores como de los invitados especiales se conoce la CI, el número de pasaporte, el nombre, la profesión y todos sus e-mails. Tanto la CI como el pasaporte sirven para identificar a las personas. Para cada feria al que asiste un autor interesa registrar la institución que lo registró (tenga en cuenta que un autor puede ir a un evento por una institución y a otro por otra). De los invitados especiales, interesa registrar la fecha en la que llega al evento. Tenga en cuenta que alguien que es autor en un evento puede ser invitado especial en otro (o incluso en el mismo evento).

Cada presentación tiene un único coordinador, que es uno de los invitados especiales. Todos los invitados especiales coordinan al menos una presentación. También se sabe que cuando un invitado especial se encuentra con un autor discuten al menos uno de los libros de dicho autor. Interesa registrar las notas que surgen de las discusiones de cada libro, para generar posteriormente material para la prensa.

Se pide: Modelo Entidad-Relación completo del problema.

Ejercicio 2 (20 puntos).

Considere la tabla $R(A,B,C,D,E)$ con las siguientes dependencias:

$$F = \{ EC \rightarrow D, D \rightarrow BC, B \rightarrow E, AD \rightarrow E \}$$

- Determine todos los atributos primos. Justifique su respuesta.
- Determine la forma normal en que se encuentra R . Justifique su respuesta.
- Encuentre una descomposición en dos tablas que NO tenga join sin pérdida pero tenga al menos un atributo en común entre las dos tablas. Justifique su respuesta.
- Determine si se perdieron dependencias de F en la descomposición de la parte c). Justifique su respuesta.
- Determine la forma normal en que se encuentra esa descomposición. Justifique su respuesta.
- Encuentre un cubrimiento minimal de F .

Ejercicio 3 (20 puntos)

Consideremos un manejador de bases de datos, que registra el uso del espacio en disco por parte de los usuarios así como las operaciones que realizan éstos sobre las tablas. Se dispone asimismo de cierta información relativa a la seguridad del sistema. Un proceso agrega información a las siguientes tablas:

Sesiones (idSesion, usuario, pwd, fecha, hora, duracion, IP)

Contiene información de las sesiones establecidas en el manejador. A cada sesión se le asigna un idSesion y se conoce el nombre de usuario, la contraseña utilizada (pwd), la fecha y hora de comienzo de la sesión, la duración de la sesión medida en segundos y la IP de la máquina desde la cual se realizó la conexión.

Tablas (idTabla, nombre, largoTuplaProm, cantTuplaProm)

Esta describe las tablas del sistema. Cada tabla se identifica por un idTabla, y se conoce su nombre. Además, largoTuplaProm es el largo promedio en bytes de una tupla de la tabla y cantTuplaProm es la cantidad promedio de tuplas en la tabla.

Accesos (idTabla, idSesion, cantUpd, cantDel, cantIns, duracion)

Se almacenan aquí los accesos de una determinada sesión idSesion a una tabla idTabla y se guarda la cantidad de tuplas modificadas, eliminadas e insertadas, cantUpd, cantDel, cantIns respectivamente.

Se pide:

En Cálculo Relacional:

- Devolver los usuarios que han borrado tuplas en todas sus sesiones.

En Algebra Relacional:

- Devolver las tablas con máxima cantidad promedio de tuplas que han sido accedidas alguna vez, pero nunca para borrar sus tuplas.

En SQL (sin utilizar sub-consultas en el FROM) ni en el SELECT:

- Encontrar las sesiones que han realizado la menor cantidad total de inserciones. Devolver el id de sesión, los usuarios involucrados, así como una estimación de la cantidad de bytes agregados.

Ejercicio 4 (20 puntos).

Dadas las siguientes transacciones:

T1: r1(X), r1(Y), w1(Y), c1

T2: w2(X), r2(Y), r2(Z), w2(Z), c2

1. Dar dos historias entrelazadas con T1 y T2 tales que son equivalentes, pero una evita abortos en cascada (EAC) y la otra es recuperable, pero no EAC. Justifique.
2. Escribir T1 y T2 siguiendo el protocolo **2PL básico** (utilizando locks de lectura y de escritura). Justifique.
3. Dar una historia entrelazada de las T1 y T2 de la parte anterior. Decir si la historia es serializable, justificando.

Ejercicio 5 (15 puntos)

Considere el siguiente esquema:

ACTIVIDADES (IdAct, Nombre, Duracion)

ASIGNACION_ACT (IdDoc, IdAct, DiaSemana, Hora)

A continuación se presenta una consulta sobre dicho esquema:

```
SELECT S.IdDoc
FROM ACTIVIDADES A, ASIGNACION_ACT S
WHERE S.IdAct = A.IdAct
AND A.Nombre = 'GIMNASIA'
```

Con el siguiente catálogo:

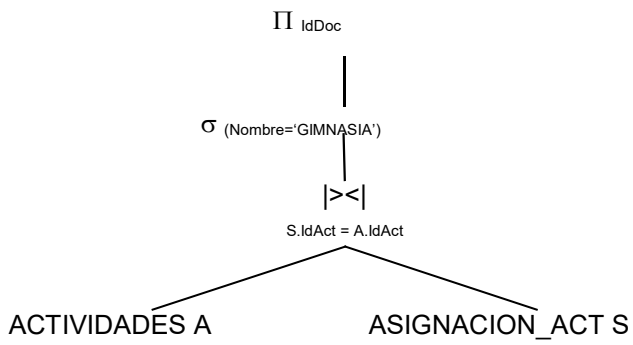
Tabla	Columna	Valores distintos (distrib. uniforme)
ACTIVIDADES	IdAct	50
ACTIVIDADES	Nombre	50
ASIGNACION_ACT	IdAct	50
ASIGNACION_ACT	IdDoc	100

Tabla	Cant. tuplas
ACTIVIDADES	50
ASIGNACION_ACT	400

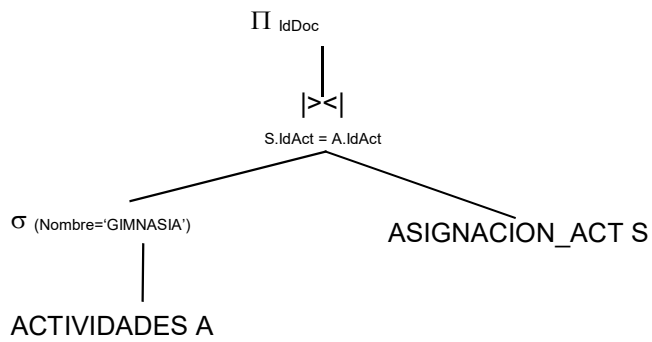
Indice	Tabla/Atributo	Tipo
Act_Nombre	ACTIVIDADES/Nombre	Secundario

Dados los siguientes planes lógicos:

Plan 1)



Plan 2)



Se pide:

- Calcular los tamaños intermedios y finales para cada plan.
- Indicar cuál debería tener mejor performance, justificando su respuesta.