

Segundo Parcial de Fundamentos de Base de Datos

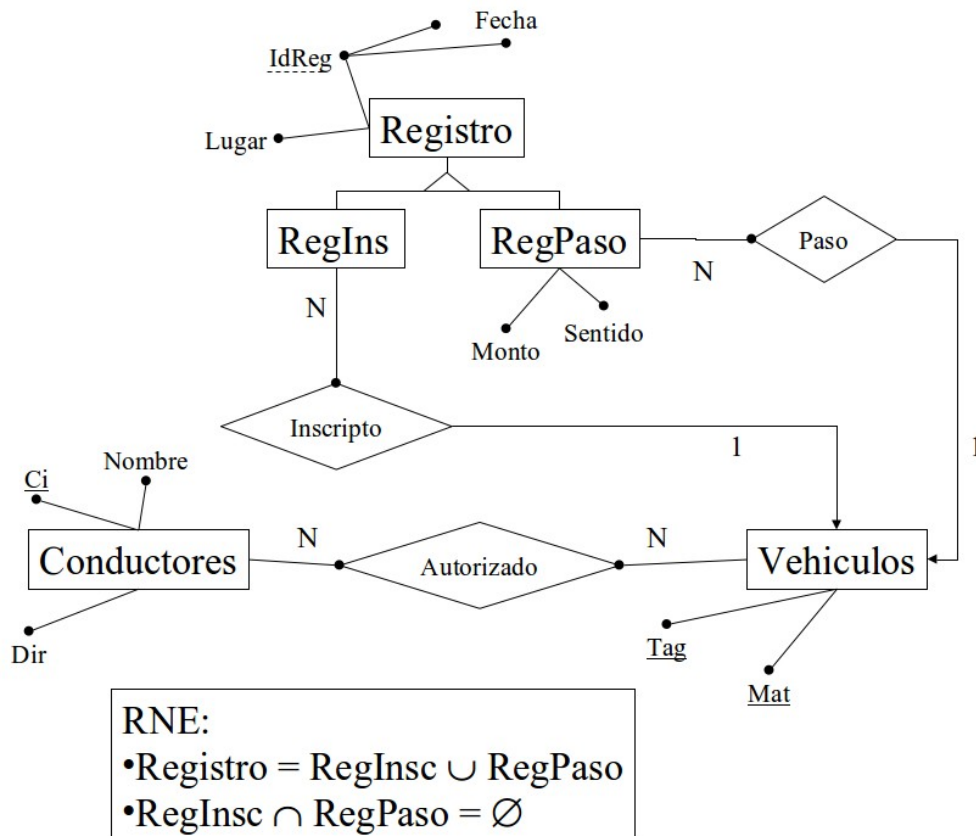
Noviembre 2011

Presentar la resolución del parcial:

- Duración: 3 horas
- Con las hojas numeradas y escritas de un solo lado.
- Con la cantidad de hojas entregadas en la primera hoja.
- Con cédula de identidad y nombre en cada hoja.
- Escrita a lápiz y en forma prolija.

Ejercicio 1 (17 puntos)

Una empresa que administra carreteras y peajes ha implementado un sistema llamado Telepeaje. En este sistema, a cada vehículo que se suscribe se le coloca un chip (llamado Tag) que es leído por antenas colocadas en cada senda del peaje y si tiene saldo, levanta la barrera correspondiente. Cada vehículo tiene un único Tag pero puede tener varios conductores autorizados. Cuando un usuario se registra en un peaje determinado, obtiene un descuento en las pasadas sólo en ese peaje. Esto hace que un usuario deba presentar el vehículo y registrarse en los diferentes peajes en los que desee descuento. Cada vez que se pasa por un peaje se registra el sentido de circulación y el saldo que queda en la cuenta. El siguiente MER representa una parte de la base de datos de ese sistema:



Se asume que este MER representa la realidad correctamente.

1. Considere la siguiente tabla que tiene todos los atributos que aparecen en el MER:
R(CI,Nombre,Dir,Tag,Mat,Fecha,Hora,Lugar,Monto,Sentido)
 - a. Indique cuáles son las dependencias funcionales que se cumplen sobre R correspondiente a ese MER. Justifique por qué agrega cada dependencia.
 - b. Indique en qué forma normal se encuentra esa tabla. Justifique.

2. Alguien construyó las siguientes tablas en base a esa realidad.

Telepague(CI,Lugar,Fecha,Hora,Tag)

Conductores(CI,Nombre,Dir)

Pasadas(Mat,Tag,Lugar,Fecha,Hora,Sentido,Monto)

- a. En qué forma normal está este esquema relacional con respecto al conjunto de dependencias funcionales de la parte 1.a.? Justifique.
 - b. ¿Se cumple sobre la tabla Telepague la dependencia $Tag \twoheadrightarrow Ci$? Justifique.
3. Construya un esquema relacional a partir del Mer. Sobre cada tabla indique todas las claves, las dependencias funcionales y de inclusión. Indique en qué forma normal se encuentra este esquema relacional.

Ejercicio 2 (15 puntos)

Dado el esquema relación R (A, B, C, D, E, G, H), y el conjunto de dependencias funcionales

$F = \{ AB \rightarrow G, B \rightarrow GH, E \rightarrow CDH, BD \rightarrow E, G \rightarrow H \}$

- a) Hallar todas las claves de R según F.
- b) Hallar un cubrimiento minimal de F, mostrando los pasos seguidos.
- c) Llevar R a 3NF con JSP y preservación de dependencias, justificando cada una de estas propiedades.
- d) Decir cuáles de los sub-esquemas obtenidos se encuentran en BCNF y cuáles no, justificando.
- e) Dar una descomposición de R distinta de la obtenida en la parte c), que no se encuentre en BCNF, justificando.
- f) Considere que se cumple la dependencia multivaluada embebida $D \twoheadrightarrow E \mid B$. Llevar la descomposición obtenida en la parte c) a 4NF, utilizando el algoritmo dado en el curso.

Ejercicio 3 (13 puntos)

La siguiente realidad corresponde a una peluquería femenina en la cual se registra la atención de sus clientas desde el momento en que abrió por primera vez. Las tablas relevantes de la base son 3:

Cientas (ciCli, nombre, nacionalidad, teléfono) : Contiene la información de las clientas de la peluquería.

Peluqueras (ciPel, nombre, teléfono, cargo): Contiene la información de todas las peluqueras que trabajaron en algún momento en la peluquería.

Atencion (ciPel, ciCli, fechaAtención): Contiene la información de cada vez que una peluquera atendió a una clienta.

Sea la siguiente consulta SQL:

```
SELECT      A.fechaAtencion
FROM        Clientas C, Peluqueras P, Atencion A
WHERE       A.ciPel = P.ciPel and A.ciCli = C.ciCli and
           C.nacionalidad = "uruguaya" and P.cargo = "peinadora";
```

Y se conoce la siguiente Información:

Relación R	Tamaño tupla (bytes)	bf _R	n _R	Atributos	Índices
Peluqueras P	200	100	1000	V(nombre, P) = 25 V(cargo, P) = 8 (distribución uniforme)	Índice Primario de un nivel sobre ciPel Índice secundario árbol b+ de 3 niveles sobre telefono, cargo.
Clientas C	80	250	2000	V(nacionalidad, C) = 25 (distribución uniforme)	Índice Primario de un nivel sobre ciCli Índice secundario árbol b+ de 3 niveles sobre nacionalidad, nombre.
Atención A	20	1000	50000	El 60% de las tuplas corresponden a la atención por parte de peluqueras de cargo 'peinadora', y el 30% de dichas atenciones fueron hechas a clientas uruguayas.	Índice Primario de un nivel sobre ciPel, ciCli, fechaAtención.

- Se dispone de 5 buffers de memoria.

Se pide:

- 1) Construir el árbol canónico de la consulta SQL dada.
- 2) Construir un plan lógico optimizado utilizando las heurísticas y teniendo en cuenta los tamaños si es necesario.
- 3) Calcular los tamaños de los resultados de cada operación del plan lógico construido.
- 4) Dar un plan físico correspondiente al plan lógico dado en la parte 2.

Ejercicio 4 (15 puntos)

Dadas las siguientes transacciones:

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

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

Parte 1)

- a) Dar una historia de T1 y T2 entrelazada, que sea recuperable, justificando.
- b) ¿Qué ventaja tiene el hecho de que una historia sea recuperable?
- c) Dar una historia de T1 y T2 entrelazada, que Evite Abortos en Cascada, justificando.

Parte 2) Ahora considere a T1 y T2 con los siguientes bloqueos y desbloques:

T1: r1(X) r1(X) w1(X) w1(X) r1(Y) r1(Y) w1(X) w1(Y) u1(X) w1(Y) u1(Y) c1

T2: r2(Y) r2(Y) w2(Y) w2(Y) r2(X) u2(Y) r2(X) u2(X) c2

- a) Para cada una de las siguientes historias decir si puede suceder en el sistema o no, justificando.

H1: r1(X) r1(X) w1(X) r2(Y) r2(Y) w2(Y) w2(Y) r2(X) u2(Y) r2(X) u2(X) c2 w1(X) r1(Y) r1(Y) w1(X) w1(Y) u1(X) w1(Y) u1(Y) c1

H2: r2(Y) r2(Y) w2(Y) w2(Y) r2(X) r1(X) r1(X) u2(Y) r2(X) u2(X) c2 w1(X) w1(X) r1(Y) r1(Y) w1(X) w1(Y) u1(X) w1(Y) u1(Y) c1

- b) Decir para cada una de las transacciones T1 y T2, si siguen el protocolo 2PL básico.
- c) Para cada una de las historias posibles de la parte a) decir si es serializable, justificando.

Implementaciones de los Operadores.

Oper.	Algoritmo	Costo	Condición	Organización
$\sigma_c(R)$	Búsqueda Lineal	b_R peor caso, $b_R/2$ promedio	Cualquier Caso	Cualquiera
	Búsqueda Binaria	$\log_2 b_R + \lceil s/bf_R \rceil - 1$	Cualquier caso	Registros ordenados
	Índice Primario	$x + 1$	Por igualdad a un valor	Registros Ordenados
	Hash	1 o 2 según el tipo	Por igualdad a un valor	Cualquiera
	Índices Primario	$x + (b/2)$ (promedio)	Por relación de orden.	Índice ordenado
	Índice Cluster	$x + \lceil s/bf_R \rceil$	Cualquier Caso	Registros Ordenados

CSI-INCO *Fundamentos de Bases de Datos* 1 --

Implementaciones de los Operadores.

Oper.	Algoritmo	Costo	Cond.	Organización
$\sigma_c(R)$	Índice secundario B+	$x + s$ peor caso	Cualquier Caso	Cualquiera
	Grabación Intermedia	s/bf_R	Cualquier caso	Cualquiera
$R \bowtie X \bowtie S$	Loop Anidado (registros)	$b_R + (n_R * b_s)$	Cualquier caso	Cualquiera
	Loop Anidado (bloque)	$b_R + \lceil b_R / (M-2) \rceil * b_s$	Cualquier caso	Cualquiera
	Sort Merge	$b_R + b_s + \text{costo ords.}$	Cualquier caso	índice en disco
	Index join	$b_R + n_R * Z$	Cualquier caso	índice en disco

CSI-INCO *Fundamentos de Bases de Datos* 1

Donde Z depende del tipo de índice:

- primario: $Z = x + 1$, hash = h

- secundario: $Z = x + ss$, cluster: $Z = x + \lceil ss/bfs \rceil$