

FUNDAMENTOS DE BASES DE DATOS

Examen Julio 2005

Solución

Parte 1. Modelo Entidad-Relación (25 puntos)

Ejercicio 1. (25 pts)

La empresa "Coordinación de Reparaciones" presta el servicio de reparaciones domésticas en general. La empresa no tiene personal propio sino que funciona como intermediaria entre los clientes y un conjunto de personas a las que les deriva los trabajos. La empresa administra la reparación de bienes inmuebles (Por ej. sanitaria) y también de artefactos diversos (desde licuadoras a DVD's, pasando por televisores y hornos de gas, etc.). Se debe diseñar la base de datos de su sistema central de acuerdo al siguiente relevamiento:

De cada cliente se conoce su teléfono de contacto, su nombre, su cédula de identidad, su dirección y un nro. de celular.

De los técnicos se conoce su nombre, su cédula de identidad, dirección, teléfono y celular.

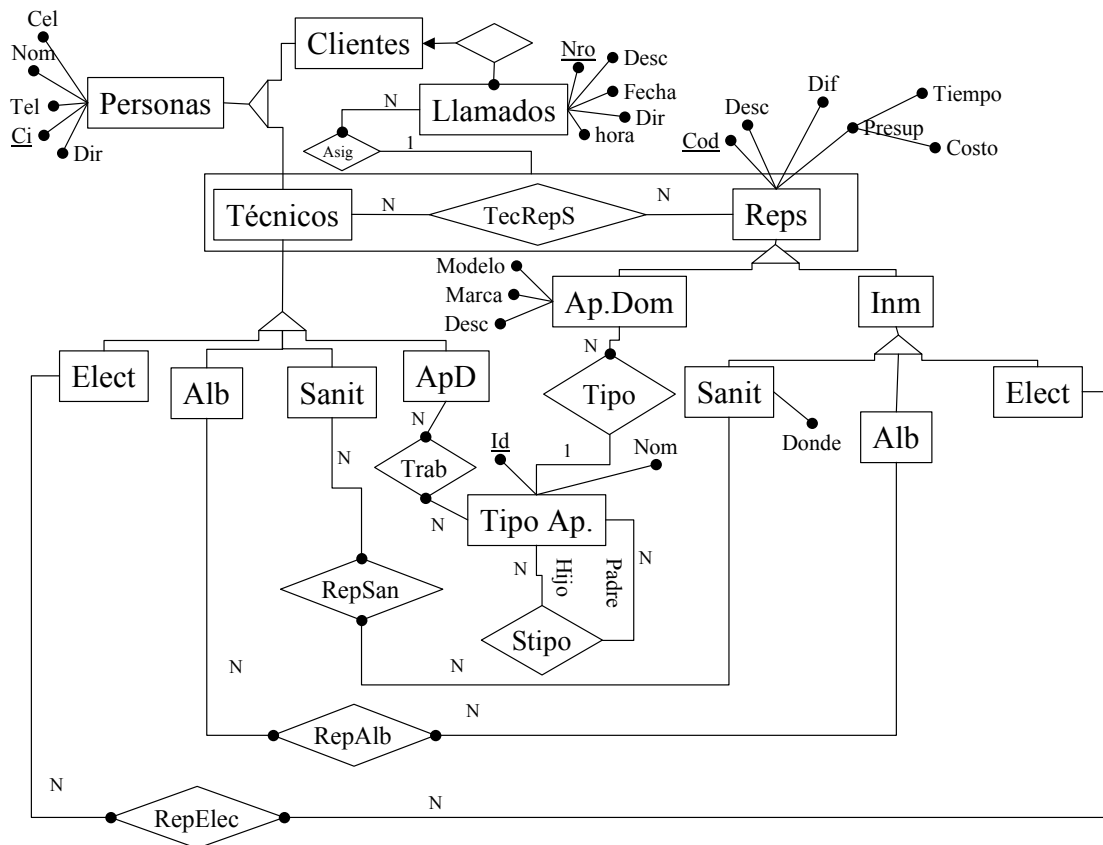
La empresa maneja un catálogo de las reparaciones para las cuales hay al menos un técnico capaz de realizarlas. De las reparaciones, se conoce un código que las identifica, una descripción, un índice de dificultad, un tiempo y un costo de presupuestación. Las reparaciones posibles se clasifican en de aparatos domésticos o de inmuebles. Las reparaciones de inmuebles pueden ser de sanitaria, de albañilería o de electricidad. Para las reparaciones de sanitaria, se conoce en qué parte del inmueble se puede realizar (cocina, baño, comedor, etc.). De las reparaciones de aparatos domésticos, se registra una marca, un modelo, una descripción y un único tipo de aparato asociado.

De cada tipo de aparato se conoce un identificador y un nombre. Los tipos de aparatos pueden tener diversos subtipos que también son tipos de aparatos. Hay que tener en cuenta que cada tipo de aparato puede tener uno o más supertipos. Por ejemplo, como tipos de aparato podemos tener "digital", "analógico", "video", "audio", "dvd", etc. El tipo "dvd" debería tener como supertipos "video", "audio" y "digital".

Los técnicos, se clasifican en sanitarios, albañiles, electricistas y de aparatos domésticos. Cada técnico trabaja en su especialidad, pudiendo tener más de una especialidad. De cada técnico de aparatos domésticos, se conoce el conjunto de tipos de aparatos con el que trabaja y se asume que es capaz de reparar cualquier aparato de esos tipos.

A cada llamado de un cliente se le asigna el número de llamado con respecto a los realizados por ese cliente, la dirección donde se debe realizar el trabajo o pasar a levantar el aparato, la fecha y hora y una descripción de los problemas reportados. A cada llamado también se le asigna una reparación del catálogo de reparaciones y un técnico adecuado para la atención.

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



Restricciones:

$TecRepS \subseteq$
 $RepAlb \cup RepElec \cup \{ \langle x,y \rangle / \exists b \in TipoAp (\langle x,b \rangle \in Trab \wedge \langle b,y \rangle \in Tipo) \} \cup RepSan$

$Elect \cup Alb \cup Sanit \cup ApD = Técnicos$

Parte 2. Consultas (25 puntos)

Ejercicio 2 (25 pts)

Una empresa que se dedica a la comercialización de prendas de vestir posee un sitio de comercio electrónico para la venta y promoción de sus productos. Los clientes que ingresan al sitio pueden recorrer las distintas secciones agregando al carrito los productos de su preferencia, los cuales se registran en una orden de compra. La información del sitio se encuentra almacenada en una base de datos con las siguientes tablas:

Productos (Id_Producto, Nombre_Producto, Descripcion_Producto, Precio_Lista, Imagen_Producto)

Esta tabla contiene todos los productos que se ofrecen en el sitio. De cada uno de ellos se conoce un identificador (Id_Producto), su nombre (Nombre_Producto), una descripción (Descripcion_Producto), su precio de lista (Precio_Lista) y la URL de la imagen del producto (Imagen_Producto)

Ofertas (Id_Producto, Inicio_Oferta, Fin_Oferta, Porcentaje_Descuento)

Esta tabla contiene la información de las ofertas. De cada oferta se conoce el producto sobre el que se aplica la oferta (Id_Producto), el período sobre el que tiene vigencia la oferta, determinado por una fecha de inicio (Inicio_Oferta) y una fecha de fin (Fin_Oferta) y el porcentaje de descuento que se aplica sobre el precio de lista (Porcentaje_Descuento)

Secciones (Nom_Seccion, Descripcion_Seccion, Imagen_Seccion)

Esta tabla contiene la información de las secciones en las que se divide el sitio (por ejemplo: damas, niños, etc). De cada sección se conoce su nombre que la identifica (Nom_Seccion), una descripción de la sección (Descripcion_Seccion) y la URL de la imagen de la sección (Imagen_Seccion)

Secciones_Productos (Nom_Seccion, Id_Producto)

Esta tabla contiene los productos que se ofrecen a la venta en cada sección.

Clientes (Id_Cliente, Nom_Cliente, Direccion, Telefono, E-mail)

Esta tabla contiene la información de las personas que alguna vez realizaron alguna compra del sitio. De cada cliente se conoce un identificador (Id_Cliente), su nombre (Nom_Cliente), su dirección postal (Direccion) , su teléfono (Telefono) y su dirección de correo electrónico (E-mail).

Ordenes (Id_Orden, Id_Clientes, Monto, Fecha)

Esta tabla contiene la información de las órdenes de compra de los clientes. De cada orden se conoce su identificador (Id_Orden), el cliente que la realiza (Id_Cliente), el monto total de la compra (Monto) y la fecha de compra (Fecha).

Items (Id_Orden, Nro_Item, Id_Producto, Cantidad, Precio_Unitario)

Cada orden de compra se compone de ítems. De cada ítem se conoce la orden a la que corresponde (Id_Orden), un número de ítem que es único dentro de cada orden (Nro_Item), el producto al que refiere el ítem (Id_Producto), la cantidad de unidades compradas de dicho producto (Cantidad) y el precio de cada unidad de producto (Precio_Unitario). Si el ítem corresponde a un producto en oferta el precio unitario ya tiene aplicado el descuento sobre el precio de lista.

Observación: Asumir que ninguna tabla está vacía.

i) **Resolver en álgebra las siguientes consultas:**

- 1) Devolver los identificadores de los productos comprados por clientes que nunca compraron productos que alguna vez estuvieron en oferta.

$$B = \prod_{Id_Producto} (OFERTAS)$$

$$C = \text{clientes compradores de productos } B$$

$$C = \prod_{Id_Clientes} (ORDENES \mid \>\< \mid \prod_{Id_orden} (ITEMS \mid \>\< \mid B))$$

$$D = A - C = \text{clientes que nunca compraron productos alguna vez en oferta}$$

$$D = \prod_{Id_Clientes} (CLIENTES) - C$$

$$\text{Resultado} = \prod_{Id_Producto} (ORDENES \mid \>\< \mid \prod_{Id_orden} (ITEMS \mid \>\< \mid D))$$

- 2) Devolver los nombres de las secciones en las que están a la venta los productos más caros considerando sólo el precio de lista.

Productos con mayor precio de lista

$$A = \prod_{Id_Producto} (PRODUCTOS |><| PRODUCTOS) \% \prod_{Id_Producto} (PRODUCTOS) \\ \$4>=\$9$$

Secciones en las que se ofrecen productos más caros

$$Resultado = \prod_{Nom_Seccion} (SECCIONES_PRODUCTOS |><| A)$$

ii) Resolver en cálculo las siguientes consultas:

- 3) Devolver los identificadores de los clientes que hayan comprado todos los productos.

$$\{o.Id_Cliente / ORDENES(o) \wedge (\forall p) (PRODUCTOS(p) \rightarrow (\exists i) (ITEMS(i) \wedge (\exists u) (ORDENES(u) \wedge i.Id_Orden = u.Id_Orden \wedge i.id_producto = p.id_producto \wedge u.Id_Cliente = o.Id_Cliente)))\}$$

- 4) Devolver los nombres de los productos comprados en alguna de las órdenes de mayor valor total.

$$\{p.Nom_Producto / PRODUCTOS(p) \wedge (\exists o) (ORDENES(o) \wedge (\exists i) (ITEMS(i) \wedge i.Id_Orden = o.Id_Orden \wedge i.id_producto = p.id_producto \wedge \neg(\exists o1)(ORDENES(o1)/o1.Monto > o.Monto))\}$$

iii) Resolver en SQL las siguientes consultas, sin utilizar vistas:

- 5) Devolver los nombres y teléfonos de clientes que hicieron alguna compra de más de 3 ítems.

```
SELECT C.NOM_CLIENTE, C.TELEFONO
FROM CLIENTES C, ORDENES O
WHERE C.ID_CLIENTE=O.ID_CLIENTE AND ID_ORDEN IN (
    SELECT O.ID_ORDEN
    FROM ORDENES O, ITEMS I
    WHERE O.ID_ORDEN = I.ID_ORDEN
    GROUP BY ID_ORDEN
    HAVING COUNT(*) >3)
```

- 6) Devolver la identificación de los productos y las fechas en las que se vendieron esos productos durante Julio de 2004, tales que la cantidad vendida en al menos una orden de esa fecha es mayor que la cantidad vendida total en el mes de junio de 2004 para ese mismo producto.

Select I.id_prod,O.Fecha

from Items I, Ordenes O

where I.id_orden=O.id_orden and O.fecha >= '1/7/2004' and O.fecha <='31/7/2004' and

I.cantidad > (select sum(I1.cantidad)

from Ordenes O1, Items I1

where O1.Id_Orden=I1.id_Orden and O1.Fecha >='1/6/2004' and O1.fecha <= '30/6/2004' and

I1.id_Producto=I.id_producto)

Parte 3. Diseño Relacional (25 puntos)

Ejercicio 3

Considere la siguiente relación con el siguiente conjunto de dependencias funcionales:

$R(A,B,C,D,E,G,H)$ $F=\{AB \rightarrow D, A \rightarrow G, BG \rightarrow H, CE \rightarrow B, B \rightarrow C\}$

- 1) Indicar cuáles de las siguientes dependencias están en F^+ y cuáles no. Justifique.
 - a. $DA \rightarrow C$
 - b. $ECA \rightarrow H$
- 2) Considerando ahora $F'=F \cup \{EC \rightarrow \rightarrow GH\}$ Indicar si la siguiente dependencia está en la clausura de dependencias. Justifique.
 $EC \rightarrow G$
- 3) Hallar todas las claves
- 4) Llevar a 3NF con join sin pérdida aplicando los algoritmos vistos en el curso.
- 5) Llevar a 4NF con join sin pérdida aplicando los algoritmos vistos en el curso, a partir del resultado en 3NF que obtuvo en la parte 4. Discuta si se pierden dependencias funcionales o no.

Solución:

- 1)
 - a. Calculo la clausura de DA , $DA^+=\{D,A,G\}$. C no está en la clausura, por lo tanto la dependencia no se cumple.
 - b. Calculo la clausura de ECA . $ECA^+=\{E,C,A,B,G,H\}$ Como H está en la clausura, entonces la dependencia se cumple.
- 2) La dependencia se cumple por la regla de coalescencia, que dice:
 $X \rightarrow \rightarrow Y, W \cap Y = \{\}, Z \subseteq Y, W \rightarrow Z \models X \rightarrow Z$.
En este caso, $W=\{A\}$ $Z=\{G\}$ $X=\{E,C\}$, $Y=\{G,H\}$.

3)

A, E no aparecen del lado derecho en ninguna df => pertenecen a toda clave
D, H no aparecen del lado izquierdo en ninguna df, pero sí del lado derecho
=> no pertenecen a ninguna clave

$$(AE)^+ = \{ A E G \}$$

$$(AEB)^+ = \{ A E B D G H C \} \Rightarrow (AEB) \text{ es clave}$$

$$(AEC)^+ = \{ A E C G B D H \} \Rightarrow (AEC) \text{ es clave}$$

$$(AEG)^+ = \{ A E G \}$$

Cualquier otra clave debe contener a AE, no puede contener a B ni a C (porque sería superclave), no puede contener a D ni a H. Solo puede contener a G, o sea que quedaría AEG, que ya se probó. Por lo tanto no hay más claves.

4)

(i) Cubrimiento minimal de F.

1er. Paso – Las dfs ya tienen un solo atributo a la derecha.

2º. Paso – Buscamos atributos redundantes a la izquierda.

$$(A)^+ = \{ A G \}$$

$$(B)^+ = \{ B C \} \quad AB \rightarrow D \text{ no tiene atributos redundantes.}$$

$$(G)^+ = \{ G \} \quad BG \rightarrow H \text{ no tiene atributos redundantes.}$$

$$(C)^+ = \{ C \}$$

$$(E)^+ = \{ E \} \quad CE \rightarrow B \text{ no tiene atributos redundantes.}$$

3er. Paso – Buscamos dfs redundantes.

Se observa que no se puede eliminar ninguna df, porque en cada df que consideramos, los atributos de la derecha no aparecen a la derecha en ninguna otra df, por lo tanto no podrían ser determinados de ninguna otra forma.

(ii) Formamos las relaciones según lo indica el algoritmo.

R1 (ABD)

R2 (AG)

R3 (BGH)

R4 (CEB)

R5 (BC) - Se elimina por estar incluida en R4

(iii) Agregamos una relación con una clave, para asegurar JSP.

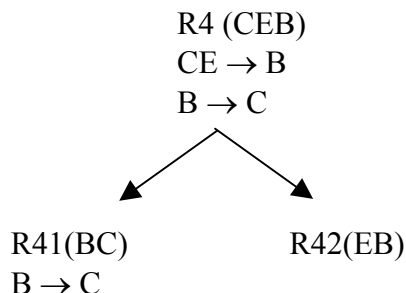
R5 (AEB)

Descomposición en 3NF con preservación de dfs y JSP:

R1(ABD), R2(AG), R3(BGH), R4(CEB), R5(AEB)

5)

R1, R2, R3 y R5 ya están en 4NF.



Se pierde la df $CE \rightarrow B$, ya que la misma no puede ser derivada de la unión de las dfs de R41 y R42.

Descomposición en 4NF con JSP:

R1(ABD), R2(AG), R3(BGH), R41(BC), R42(EB), R5(AEB)

Parte 4 Optimización y Concurrencia (25 puntos)

Ejercicio 4 (13 pts).

Dado el siguiente esquema:

VENTAS (NroVta, Fecha, Monto, Producto, NroCliente)

CLIENTES (NroCliente, Direccion, Telefono, FNac)

$\Pi_{NroCliente} (VENTAS) \sqcap \Pi_{NroCliente} (CLIENTES)$

Comparar el costo de 2 planes físicos (correspondientes a 2 planes lógicos diferentes, no necesariamente optimizados) para la siguiente consulta:

```

SELECT *
FROM VENTAS V, CLIENTES C
WHERE V.NroCliente = C.NroCliente AND V.Monto = 40000
    
```

Especificar plan lógico y plan físico en cada caso. No olvidar considerar los costos de grabar para los resultados intermedios. Considere que se tienen 3 buffers disponibles.

Datos:

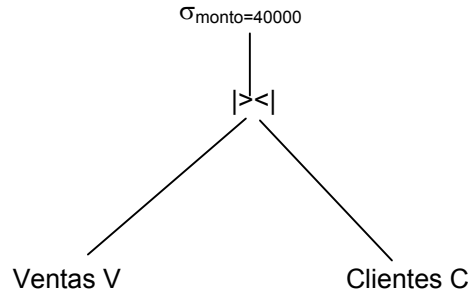
Tabla	Columna	Valores distintos
Clientes	NroCliente	50
Ventas	NroCliente	50
Ventas	Monto	500 (distr. uniforme)

Tabla	Cant. tuplas	Factor de bloqueo
Clientes	50	10
Ventas	10000	5
Clientes >< Ventas		3

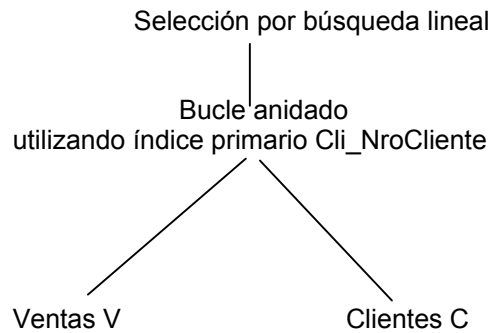
Indice	Tabla/Atributo	Tipo	Cant. niveles
Cli_NroCliente	Clientes/NroCliente	Primario	1
Vtas_Monto	Ventas/Monto	Secundario	1

Solución:

Plan lógico 1:



Plan físico 1:



Costo:

$$C(\text{join}) = b_V + n_V * (x+1) = 2000 + 10000 * 2 = 22000$$

$$C(\text{grabar}) = b_{\text{join}} = |\text{join}| / \text{factor de bloqueo} = 10000 / 3 = 3334$$

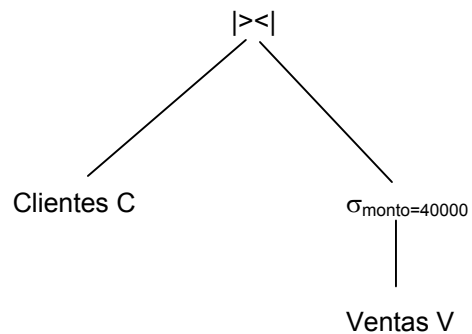
$$C(\text{selección}) = b_{\text{join}} = 3334$$

Aclaración:

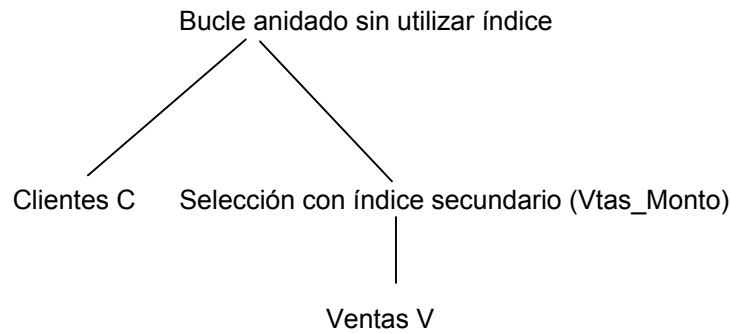
$|\text{join}| = |V| = 10000$, porque NroCliente es clave en Clientes y clave foránea en Ventas

Ctotal = 18668

Plan lógico 2:



Plan físico 2:



Costo:

$$C(\text{seleccion}) = x + |\text{selección}| = 1 + 20 = 21$$

$$C(\text{grabar}) = b_{\text{seleccion}} = |\text{seleccion}| / \text{factor de bloqueo} = 20 / 5 = 4$$

$$C(\text{join}) = b_C + b_C * b_{\text{seleccion}} = 5 + 5 * 4 = 25$$

Aclaración:

$$|\text{seleccion}| = |\text{Ventas}| / V(\text{monto}, \text{Ventas}) = 10000 / 500 = 20$$

Ctotal = 50

Vemos que el plan 2 es mucho menos costoso que el plan 1.

Ejercicio 5. (12 pts.)

(i) Dar las siguientes definiciones:

- i. Historia Serializable
- ii. Historia Recuperable
- iii. Historia Estricta
- iv. Historia que Evita Abortos en Cascada

(ii) Sea un sistema de ventas de pasajes, existiendo distintos puntos de venta.

Procedure Venta_pasajes

cant := input(c);

if pasajes_disp > 0 then

pasajes_disp := read(X);

write (X, pasajes_disp - c);

recaudacion := read(Y);

write(Y, recaudacion + valor_pasaje * c);

end;

end;

Escribir historias H1 donde se ejecuten 2 transacciones de Venta_pasajes, que cumplan:

- I. H1 serializable y no recuperable.

- II. H1 no serializable y estricta.
- III. H1 serial.
- IV. H1 no serializable y que evita abortos en cascada.

Solución.

(ii)

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

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

- I. H1: r1(X) w1(X) r2(X) w2(X) r1(Y) w1(Y) r2(Y) w2(Y) c2 c1
- II. H1: r2(X) r1(X) w1(X) r1(Y) w1(Y) c1 w2(X) r2(Y) w2(Y) c2
- III. H1: r1(X) w1(X) r1(Y) w1(Y) c1 r2(X) w2(X) r2(Y) w2(Y) c2
- IV. H1: r1(X) r2(X) w2(X) w1(X) r1(Y) w1(Y) c1 r2(Y) w2(Y) c2