

FUNDAMENTOS DE BASES DE DATOS

Solución Examen Febrero 2005

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. (25 pts)

La organización "Músicos sin Fronteras" (MSF) organiza recitales a beneficio en todo el mundo. Dicha organización desea construir un Sistema de Información para llevar el control de los eventos realizados y facilitar la organización de los nuevos.

Para la organización de un evento, MSF cuenta con un registro de intérpretes y un registro de escenarios posibles y un registro de partituras posibles para ejecutar en un programa.

Dado el carácter internacional de MSF, se cuenta con un registro de países y ciudades. De cada país se conoce el nombre y una identificación. De cada ciudad, se conoce el nombre, en qué país se encuentra y su huso horario (ej: el de Montevideo actualmente, es +2 con respecto a Greenwich). Además, para cada pareja de ciudades se conoce un conjunto de medios de transporte mediante los que se puede viajar de una a otra. De cada medio de transporte entre dos ciudades se registra el tipo (avión, omnibus, auto, tren, etc.), el tiempo de viaje, y el costo total (incluyendo tasas, impuestos, etc.) en dólares. Hay que tener en cuenta que esta información no es simétrica, es decir, viajar por avión de San Pablo a Montevideo podría ser más caro que viajar por avión de Montevideo a San Pablo debido a las tasa de los aeropuertos.

De los intérpretes se conoce el nombre, su número de pasaporte y el conjunto de estilos que domina (clásico, popular, pop, etc.). Además se conoce la ciudad en la que viven habitualmente. Los intérpretes pueden ser músicos, directores o cantantes. De los músicos, se conoce el conjunto de instrumentos que tocan, y de los cantantes se conoce el registro de su voz.

De los instrumentos se conoce un nombre que los identifica, una descripción y un tipo (vientos, cuerda, electrónico, etc.).

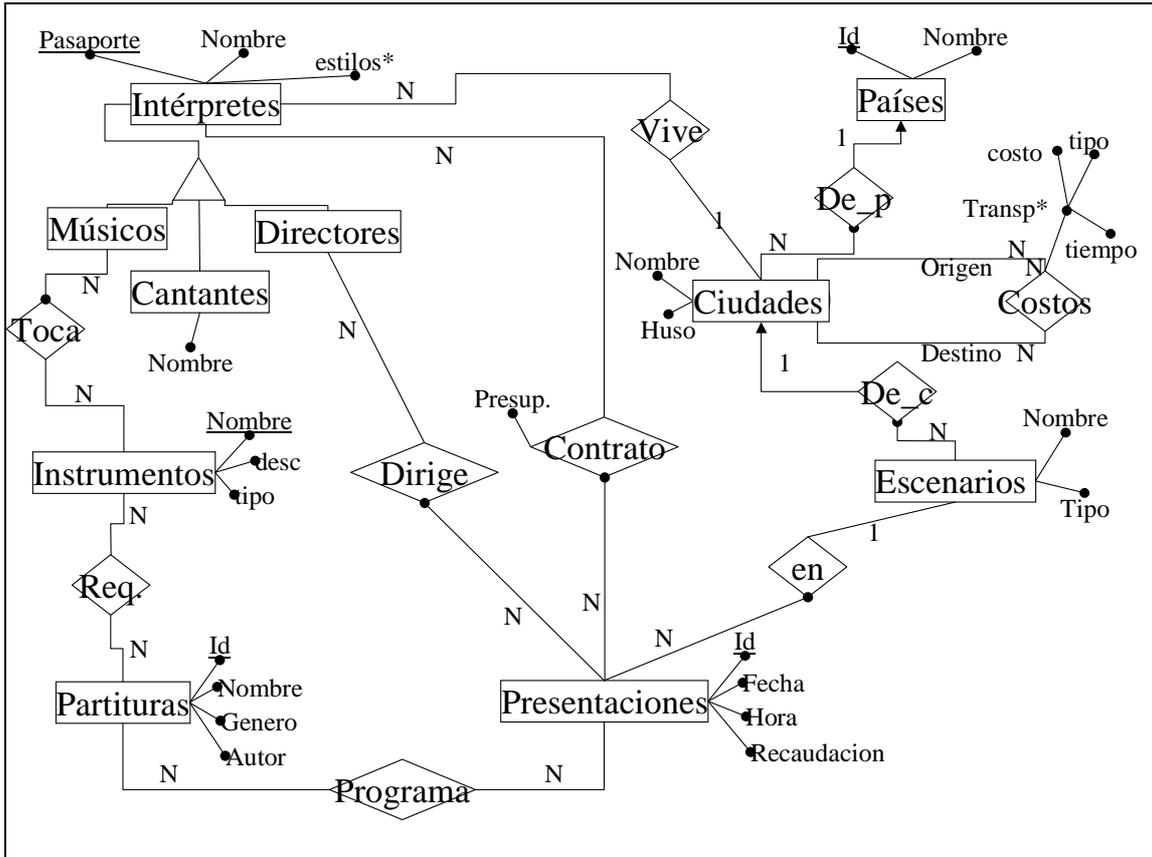
De las partituras se conoce una identificación, un nombre, un género, el nombre del autor y los instrumentos que requiere en su ejecución.

Cada escenario tiene un nombre y un tipo (estadio, teatro cerrado, teatro abierto, etc.) y la ciudad en la que se encuentra. Se asume que los nombres de los escenarios no se repiten en una misma ciudad.

Una presentación tiene una identificación, una fecha, una hora (local de acuerdo al escenario), un valor de recaudación. También se conoce el escenario en el que se va a realizar, el programa (o sea el conjunto de partituras que se van a ejecutar) y el conjunto de intérpretes contratados para la presentación. De cada intérprete contratado en una presentación se conoce el presupuesto asignado en dólares para que el intérprete participe de la presentación. Toda presentación tiene que tener al menos un director y al menos un músico o un cantante (no necesariamente los dos), no necesariamente distintos del director (o entre ellos).

MSF tiene una regla básica para la conformación de la orquesta: deben cubrirse todos los instrumentos del programa y no debe asignarse intérpretes que estén a más de 1000U\$S de "distancia", es decir, cualquiera de los pasajes (ida o vuelta) para el traslado desde donde vive el intérprete sea de más de 500U\$S.

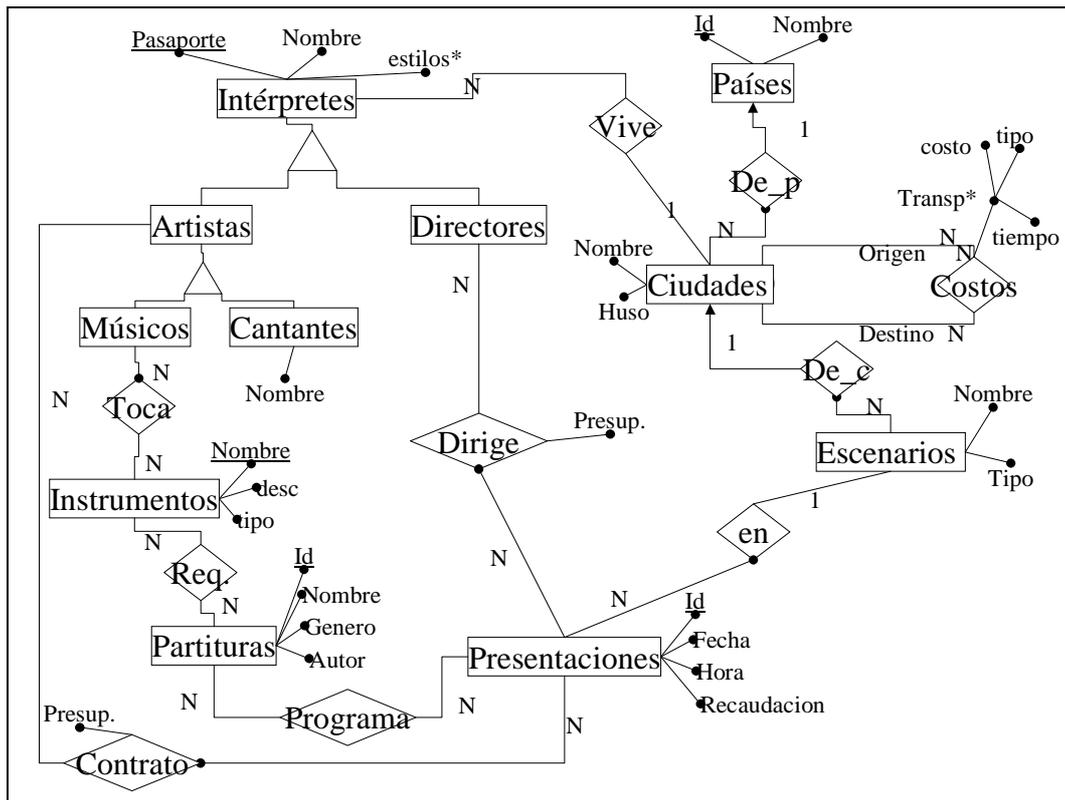
SE PIDE: Esquema Entidad Relación completo, incluyendo restricciones de integridad.



Restricciones no estructurales.

1. $\forall p \in \text{Presentaciones}. \forall \langle p, p_1 \rangle \in \text{Programa}. \forall \langle i, p_1 \rangle \in \text{Req}. \exists \langle p, i \rangle \in \text{contrato}. \langle p_1, i \rangle \in \text{Toca}$
2. $\forall p \in \text{Presentaciones}. \forall \langle p, i \rangle \in \text{Contrato}. \exists \langle i, c_1 \rangle \in \text{Vive}. \exists \langle p, e \rangle \in \text{En}. \exists \langle e, c \rangle \in \text{De}_c. (\langle c, c_1 \rangle \in \text{Costos} \wedge \exists t \in \text{transp}(\langle c, c_1 \rangle). (\text{costo}(t) < 500)) \wedge \langle c_1, c \rangle \in \text{Costos} \wedge \exists t \in \text{transp}(\langle c_1, c \rangle). (\text{costo}(t) < 500))$
3. $\forall p \in \text{Presentaciones}. \forall \langle p, d \rangle \in \text{Dirige}. \langle p, d \rangle \in \text{Contrato}$
4. $\forall p \in \text{Presentaciones}. \exists \langle p, i \rangle \in \text{Contrato}. (i \in \text{Músicos} \vee i \in \text{Cantantes})$
5. $\forall c \in \text{Ciudades}. \langle c, c \rangle \notin \text{Costos}$

OTRA SOLUCION



1. $\forall p \in \text{Presentaciones}. \forall \langle p, p_1 \rangle \in \text{Programa}. \forall \langle i, p_1 \rangle \in \text{Req}. \exists \langle p, i \rangle \in \text{contrato}. \langle p_1, i \rangle \in \text{Toca}$
2. $\forall p \in \text{Presentaciones}. \forall \langle p, i \rangle \in \text{Contrato}. \exists \langle i, c_1 \rangle \in \text{Vive}. \exists \langle p, e \rangle \in \text{En}. \exists \langle e, c \rangle \in \text{De}_c. (\langle c, c_1 \rangle \in \text{Costos} \wedge \exists t \in \text{transp}(\langle c, c_1 \rangle). (\text{costo}(t) < 500)) \wedge \langle c_1, c \rangle \in \text{Costos} \wedge \exists t \in \text{transp}(\langle c_1, c \rangle). (\text{costo}(t) < 500))$
3. $\forall c \in \text{Ciudades}. \langle c, c \rangle \notin \text{Costos}$

Parte 2. Consultas (25 puntos)

Ejercicio 2 (25 pts)

El camping “Los Mosquitos” posee un sistema para manejar reservas de cabañas y predios el cual usa las siguientes tablas:

Predios (códigoP, categoría, zona)

En esta tabla se guarda la información relativa a los predios donde se puede acampar. De cada predio se conoce su código que lo identifica (códigoP), su categoría (categoría) y la zona en la que se encuentra (zona). Las categorías son ‘A’, ‘B’ o ‘C’ dependiendo de los servicios que disponga. Las zonas son ‘Sur’, ‘Este’, ‘Norte’ u ‘Oeste’.

Cabañas (códigoC, capacidad, aire_acondicionado)

En esta tabla se guarda la información relativa a las cabañas. De cada cabaña se conoce su código que la identifica (códigoC), la cantidad de personas que puede albergar (capacidad), y si dispone o no de aire acondicionado (aire_acondicionado).

Equipamiento (códigoC, númeroA, cantidad)

En esta tabla se guarda información relativa a los artículos con los que cuenta una cabaña y la cantidad que dispone de los mismos (cantidad).

Artículos (númeroA, nombre)

En esta tabla se guarda información relativa a los distintos artículos. De cada artículo se conoce su número que lo identifica (númeroA), y su nombre (nombre). Algunos ejemplos de artículos pueden ser: cama, cocina, plato, etc.

Reservas (códigoR, CI, fecha_ini, fecha_fin, cantidad_personas)

En esta tabla se guarda información relativa a las reservas. De cada reserva se conoce el código del predio o cabaña reservado (código), la cédula de identidad de la persona que realizó la reserva (CI), la fecha en que comienza la reserva (fecha_ini), la fecha en que finaliza la reserva (fecha_fin) y la cantidad de personas que van a ocupar el lugar reservado (cantidad_personas).

Notas:

*No existen tablas vacías

* $\prod_{\text{códigoC}}(\text{Equipamiento}) \subseteq \prod_{\text{códigoC}}(\text{Cabañas})$

* $\prod_{\text{códigoA}}(\text{Equipamiento}) \subseteq \prod_{\text{códigoA}}(\text{Artículos})$

* $\prod_{\text{códigoR}}(\text{Reservas}) \subseteq \prod_{\text{códigoP}}(\text{Predios}) \cup \prod_{\text{códigoC}}(\text{Cabañas})$

* $\prod_{\text{códigoP}}(\text{Predios}) \cap \prod_{\text{códigoC}}(\text{Cabañas}) = \emptyset$

Se pide:

Resolver las siguientes consultas en Álgebra Relacional:

- a) Obtener los códigos de las cabañas equipadas con el total de artículos en existencia, que nunca hayan sido reservadas.

$$A = \prod_{\text{códigoC}, \text{códigoA}}(\text{Equipamiento}) \% \prod_{\text{códigoA}}(\text{Artículos})$$
$$S = A - \prod_{\text{códigoA}}(\text{Reservas})$$

- b) Obtener los códigos de los predios o cabañas cuya reserva será ocupada más rápidamente.

$$A = \prod_{\text{códigoR}}(\text{Reservas}) \mid \text{><} \mid_{\$3>=\$7} \prod_{\text{códigoR}}(\text{Reservas})$$
$$S = A \% \prod_{\text{códigoR}}(\text{Reservas})$$

CORRECCIÓN => LO ANTERIOR ESTÁ MAL

$$A = \prod_{\$1, \$6}(\text{Reservas}) \mid \text{><} \mid_{\$3<=\$8} \text{Reservas}$$
$$S = A \% \prod_{\text{códigoR}}(\text{Reservas})$$

Resolver las siguientes consultas en Cálculo Relacional:

- c) Obtener las zonas donde sólo se encuentran predios de las categorías A y B.

$$\{t.zona / \text{Predios}(t) \wedge \neg(\exists p) (\text{Predios}(p) \wedge p.zona = t.zona \wedge p.categoría = 'C')\}$$

- d) Obtener las zonas en las que hay predios de categoría A y no quedan lugares disponibles para dicha categoría.

$$\{t.zona / \text{Predios}(t) \wedge (\forall p) (\text{Predios}(p) \wedge p.zona = t.zona \wedge p.categoría = 'A' \rightarrow (\exists r) (\text{Reservas}(r) \wedge r.códigoR = p.códigoP)) \wedge (\exists q) (\text{Predios}(q) \wedge q.zona = t.zona \wedge q.categoría = 'A')\}$$

Resolver las siguientes consultas en SQL sin utilizar vistas:

- e) Dar los códigos de las cabañas con más de 7 reservas donde al menos una de esas reservas se hizo para más personas que la capacidad de la cabaña y donde el período de reserva de la misma fue dentro del 2004.

```
select
from Cabañas C, Reservas R
where códigoC=códigoR and fecha_ini>='1/1/2004' and fecha_fin<='31/12/2004'
and exists (select R1.códigoR
            from Reservas R1
            where R1.códigoR=códigoC and R1.fecha_ini>='1/1/2004' and
              R1.fecha_fin<='31/12/2004' and R1.cantidad_personas > capacidad )
group by códigoC
having count(fecha_ini) > 7
```

Parte 3. Diseño Relacional (25 puntos)

Ejercicio 3 (15 pts)

Sea el esquema relación $R(A,B,C,D,E,G,H)$ y los siguientes conjuntos de dependencias sobre R :

$F1 = \{ BD \rightarrow AC, AB \rightarrow C, GH \rightarrow AE, BG \rightarrow E, AE \rightarrow B, A \rightarrow C, B \rightarrow A, DA \rightarrow B \}$

$F2 = \{ B \rightarrow A, B \rightarrow C, GH \rightarrow E, GH \rightarrow A, GH \rightarrow C, BG \rightarrow E, AE \rightarrow B, DA \rightarrow B \}$

- ¿ $F2$ es un cubrimiento minimal de $F1$?
- Calcular todas las claves de R según $F1$.
- Calcular todas las claves de R según $F2$.
- Sea $\rho = \{ (BDEGH), (ABD), (ABC) \}$ una descomposición de R .
Considerando el conjunto de dependencias $F1$: ¿ En que forma normal se encuentra ρ y cada uno de sus subesquemas?

En todos los siguientes casos SE DEBE justificar la respuesta.

Solución ejercicio 3

parte a

Para ver si $F2$ es un cubrimiento minimal de $F1$:

- $F2$ no debe tener atributos redundantes
- $F2$ no debe tener DFS redundantes
- $F1$ y $F2$ deben ser equivalentes, por lo tanto debe cumplirse $F1^+ \equiv F2^+$

Atributos redundantes:

Para cada DFS $X \rightarrow Y$ en $F2$, para cada atributo $A \in X$ calculo $(X-A)^+$.

Si $Y \subseteq (X-A)^+$ entonces A es redundante en $X \rightarrow Y$

No hay atributos redundantes.

DFS redundantes:

Para cada DFS $X \rightarrow Y$ en $F2$ calculo X^+ respecto a $(F2 - (X \rightarrow Y))$.

Si $Y \subseteq (F2 - (X \rightarrow Y))^+$ entonces $X \rightarrow Y$ es redundante.

$GH^+_{(F2 - (GH \rightarrow C))} = \{G, H, A, E, B, C\}$ $GH \rightarrow C$ es redundante, por lo tanto $F2$ no es minimal.

Parte b

DGH está en toda clave por no aparecer del lado derecho de ninguna DFS.

$DGH^+ = \{D, G, H, A, E, B, C\}$ DGH clave única.

Parte c

DGH está en toda clave por no aparecer del lado derecho de ninguna DFS.

$DGH^+ = \{D, G, H, A, E, B, C\}$ DGH clave única.

Parte d

Hay que proyectar las dependencias de F1 sobre la descomposición.

R1(BDEGH) $\Pi_{R1}=\{GH \rightarrow E, BG \rightarrow E\}$

Claves: BDGH

Por lo que está en 1NF dado que $GH \rightarrow E$ es parcial y E no es primo.

R2 (ABD) $\Pi_{R2}=\{DA \rightarrow B\}$

Claves: DA

Por lo que está en BCNF

R3(ABC) $\Pi_{R3}=\{A \rightarrow C, AB \rightarrow C, B \rightarrow A\}$

Claves: B

Por lo que está en 1NF dado que $A \rightarrow C$ es parcial y C no es primo.

El esquema está en 1NF

Ejercicio 4 (10 pts)

Sea el esquema relación R (A,B,C,D,E,G,H) y el siguiente conjunto de dependencias sobre R:

$F1 = \{ BD \rightarrow AC, AB \rightarrow C, GH \rightarrow AE, BG \rightarrow E, AE \rightarrow B, A \rightarrow C, B \rightarrow A, DA \rightarrow B\}$

- a) Llevar a BCNF con join sin pérdida sin llevar el esquema previamente a 3NF. Identifique cada paso que dá y justifique las decisiones fundamentales que toma. Discutir si hay pérdida de dependencias funcionales y en caso afirmativo indicar cuales son las dependencias que se pierden.

Solución ejercicio 4

Partiendo de la descomposición de la parte d del ejercicio 1...

$D := \{ r1(BDEGH), r2(ABD), r3(ABC)\}$

Paso 1:

$\Pi_{R1}=\{GH \rightarrow E, GH \rightarrow B, BG \rightarrow E\}$

Claves: BDGH

$\Pi_{R2}=\{DA \rightarrow B\}$

Claves: DA entonces está en BCNF

$\Pi_{R3}=\{A \rightarrow C, AB \rightarrow C, B \rightarrow A\}$

Claves: B

En R1 $GH \rightarrow E$ viola BCNF porque GH no es superclave.

Descompongo R1 en R11(BDGH) R12(GHE)

$\Pi_{R11}=\{GH \rightarrow B\}$

$\Pi_{R12}=\{GH \rightarrow E\}$

Ambos subesquemas están en BCNF.

Paso 2:

En R3 $A \rightarrow C$ viola BCNF porque A no es superclave.

Descompongo R3 en R31(AB) y R32(AC)

$\Pi_{R31}=\{B \rightarrow A\}$

$\Pi_{R32}=\{A \rightarrow C\}$

Ambos subesquemas están en BCNF.

La descomposición resultante es: $\rho = \{ (BDGH), (GHE), (ABD), (AB), (AC) \}$

Para que se preserven las DFS la clausura de la unión de los conjuntos resultantes de proyectar las DFS de F1 sobre cada uno de los elementos de la descomposición debe ser igual a la clausura de F1.

Sean :

$F1 = \{BD \rightarrow AC, AB \rightarrow C, GH \rightarrow AE, BG \rightarrow E, AE \rightarrow B, A \rightarrow C, B \rightarrow A, DA \rightarrow B\}$

$J = \{GH \rightarrow B, GH \rightarrow E, DA \rightarrow B, B \rightarrow A, A \rightarrow C\}$

¿ $F1^+ = J^+$?

$B^+_J = \{B, A, C\}$

Se conservan $BD \rightarrow AC$ y $B \rightarrow A$

$A^+_J = \{A, C\}$

Se conservan $AB \rightarrow C$ y $A \rightarrow C$

$GH^+_J = \{G, H, E, B, A, C\}$

Se conservan $GH \rightarrow AE$

$BG^+_J = \{B, G, A, C\}$

Se pierde $BG \rightarrow E$

$AE^+_J = \{A, E, C\}$

Se pierde $AE \rightarrow B$

Parte 4 Optimización y Concurrency (25 puntos)

Ejercicio 5 (12 pts).

- Explicar brevemente que es el protocolo 2PL básico y cual es su utilidad.
- Dadas las transacciones:

T1: r1(x) w1(x) r1(y) w1(y) c1

T2: r2(x) r2(y) w2(x) c2

- Escribir una historia con bloqueos y desbloqueos (de lectura y escritura) donde T1 y T2 cumplan con el protocolo 2PL básico. (La historia debe ser entrelazada).
- Escribir una historia con bloqueos y desbloqueos (de lectura y escritura) donde T1 y T2 cumplan con el protocolo 2PL riguroso.

Solución ejercicio 5

- Una transacción cumple el protocolo 2PL-básico si todos sus bloqueos preceden a los desbloqueos, y cuando se realiza una degradación ya no se puede aumentar los bloqueos nuevamente (solo se pueden degradar). Las historias donde todas sus transacciones cumplen 2PL-básico, son serializables. (Ver texto del curso, Elmasri-Navathe).

b)

- r1(x) r1(x) w1(x) w1(x) r1(y) u1(x) r2(x) r2(x) r1(y) w1(y) u1(y) c1 r2(y) r2(y) w2(x) w2(x) u2(y) u2(x) c2
- 2PL-Riguroso: En cada transacción los unlock (de lect. y escrit.) se hacen después del commit.
w1(x) r1(x) w1(x) w1(y) r1(y) w1(y) c1 u1(x) u1(y) r2(x) r2(x) r2(y) r2(y) w2(x) w2(x) c2 u2(x) u2(y)

Ejercicio 6. (13 pts.)

En la base de datos de un taller mecánico se tienen las siguientes tablas relacionales:

Abonados(mat, vendedor, tipo-abono, v-cuota, fecha-susc, u_mes_p, marca, modelo)

Esta tabla representa los vehículos abonados al taller.

En esta tabla, para cada matrícula (mat) se tiene el vendedor que realizó la suscripción, el valor actual de la cuota, la fecha en que se realizó la suscripción, el último mes que se pagó, la marca y el modelo.

Servicios(mat, Cod-servicio, fecha, importe, status)

Esta tabla representa los servicios realizados a cada vehículo. Para cada matrícula, dado un código de servicio y una fecha, la tabla contiene el importe de ese servicio y el status. El atributo status indica si el importe fue saldado al contado, si es a crédito o si está con mora.

Cientes(CI_Cli, mat)

Esta tabla contiene los vehículos abonados junto con los clientes dueños de los mismos.

Dada la siguiente consulta SQL:

```
SELECT CI_Cli, vendedor
FROM Clientes C, Servicios S, Abonados A
WHERE A.mat = S.mat AND
      A.mat = C.mat AND
      A.marca = 'Fiat' AND
```

S.status = 'Mora'

Considerando la siguiente información:

Tabla	Columna	Valores Distintos
Servicios	Status	3 (dist. Uniforme)
Abonados	Marca	25 (dist. Uniforme)

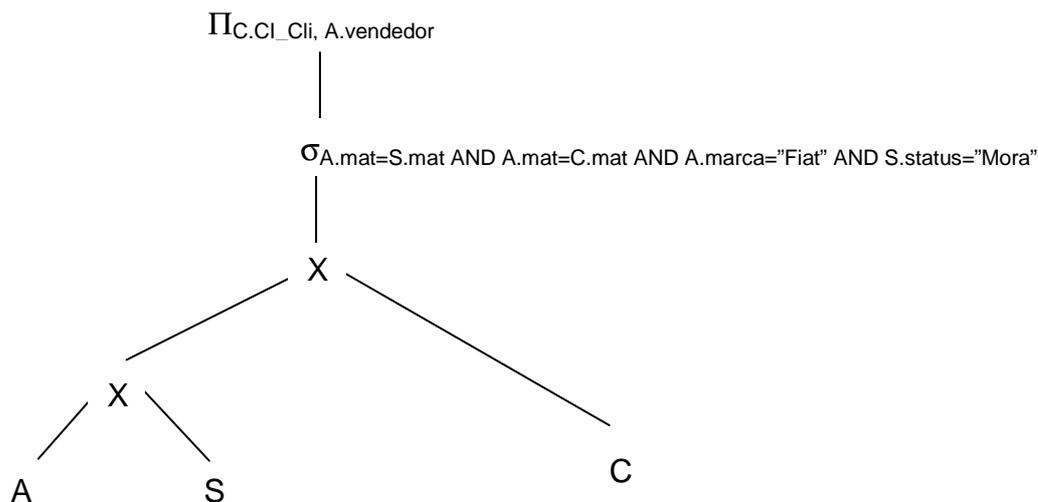
Tabla	Cant. Tuplas	Aclaraciones
Abonados	630	
Servicios	2000	- Un vehículo puede estar en estado de mora en un solo servicio. - Solo 10 vehículos de los que se encuentran en mora son marca "Fiat"
Cientes	1300	- En esta tabla se encuentran todas las matriculas abonadas. - Existen en promedio 2 clientes por cada vehículo.

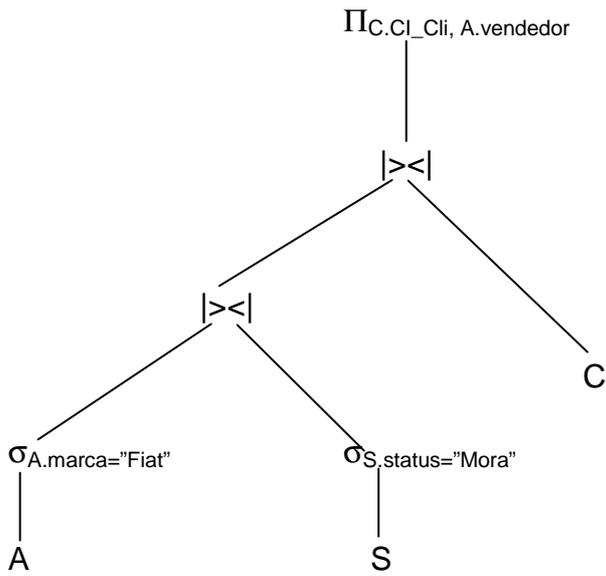
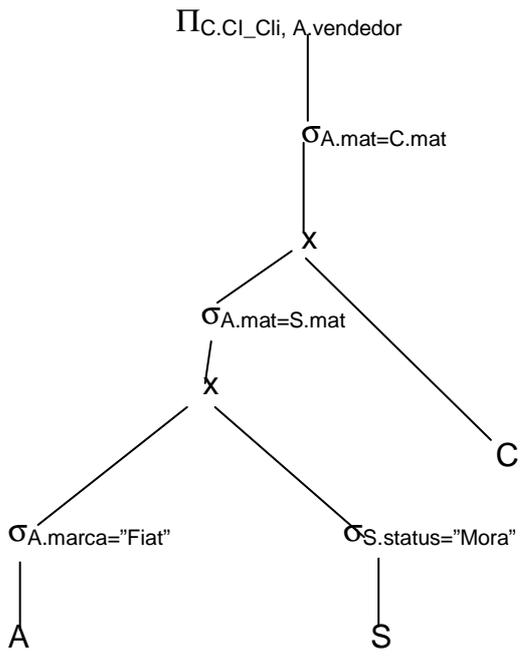
Indice	Tabla/Atributo	Tipo
Ind_Cli1	Cientes/CI_Cli	Secundario
Ind_Cli2	Cientes/mat	Secundario
Ind_Abo1	Abonados/marca	Secundario
Ind_Abo2	Abonados/mat	Primario

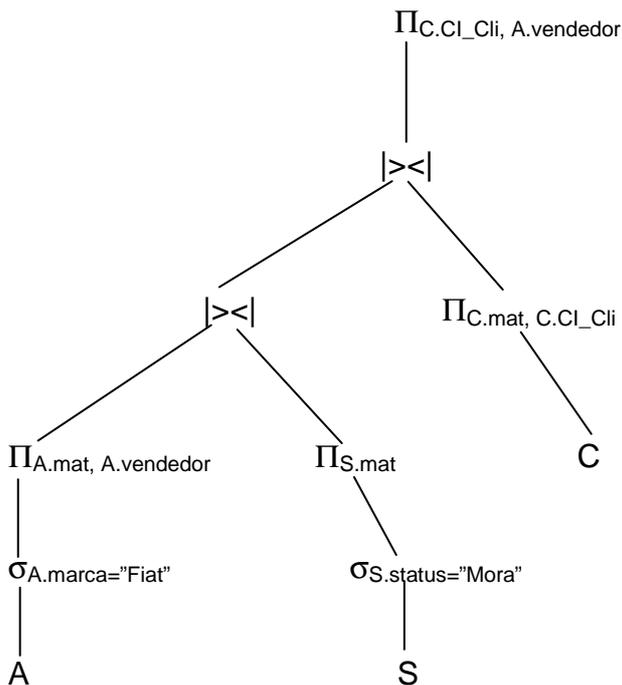
- Dar un plan lógico para la consulta, optimizado mediante las heurísticas vistas en el curso. Mostrar los pasos aplicados.
- Calcular los tamaños de los resultados de cada operación aplicada (considerando solamente las operaciones de selección y join).
- Dar un plan físico adecuado.

Solución ejercicio 6

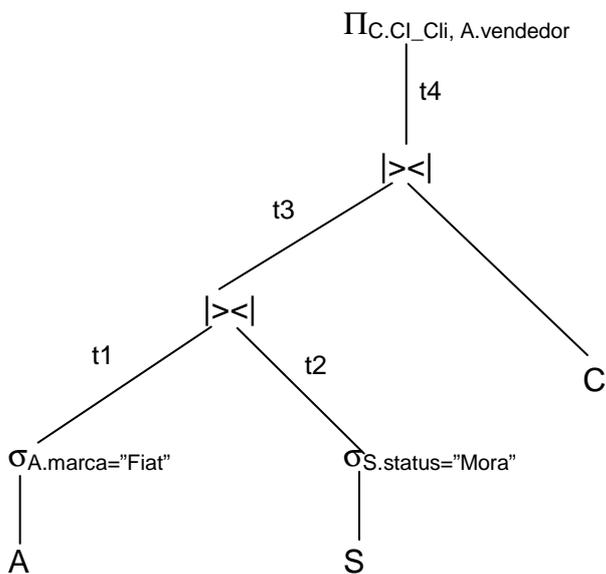
Parte a)







b)



$|A| = 630$
 $|S| = 2000$
 $|C| = 1300$

$$|t1| = \lceil |A| / V(A, \text{marca}) \rceil = \lceil 630 / 25 \rceil = 26$$

$$|t2| = \lceil |S| / V(S, \text{status}) \rceil = \lceil 2000 / 3 \rceil = 667$$

$|t3| = 10$ (solo 10 de los de t2 están en t1 y en t2 no hay repetidos)

$|t4| = 20$ (prom. 2 vehiculos por cliente)

c)

