

FUNDAMENTOS DE BASES DE DATOS

Solución Examen Julio 2011

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)

Se desea realizar el modelo conceptual para la información de una empresa de auxilio mecánico, la cual necesita registrar lo siguiente.

Existe un conjunto de vehículos de la empresa, los cuales son autos y camiones. De estos vehículos se guarda la matrícula (que identifica a cada vehículo), el año, marca y modelo. En el caso de los camiones interesa saber el peso máximo de carga que puede llevar.

Se registran los siguientes datos de los choferes de la empresa: cédula de identidad, nombre, dirección y teléfono. Algunos de los choferes manejan solamente autos, otros solamente camiones, y otros manejan tanto autos como camiones. A la vez, cada chofer que maneja autos tiene asignado un conjunto de autos que puede manejar, y análogamente los que manejan camiones. Tanto los autos como los camiones pueden ser manejados por distintos choferes.

Los vehículos de la empresa están asignados a zonas donde trabajan. De las zonas se registra un identificador, un nombre y el departamento al que pertenece. Interesa saber qué zonas son limítrofes entre sí.

Se lleva la información sobre los socios y sus vehículos asociados. De cada socio interesa su cédula de identidad, su nombre, su dirección y sus teléfonos. De cada vehículo asociado se guarda su matrícula (que lo identifica), su marca, modelo y color. Un socio puede tener varios vehículos asociados, pero para un vehículo hay un solo socio correspondiente.

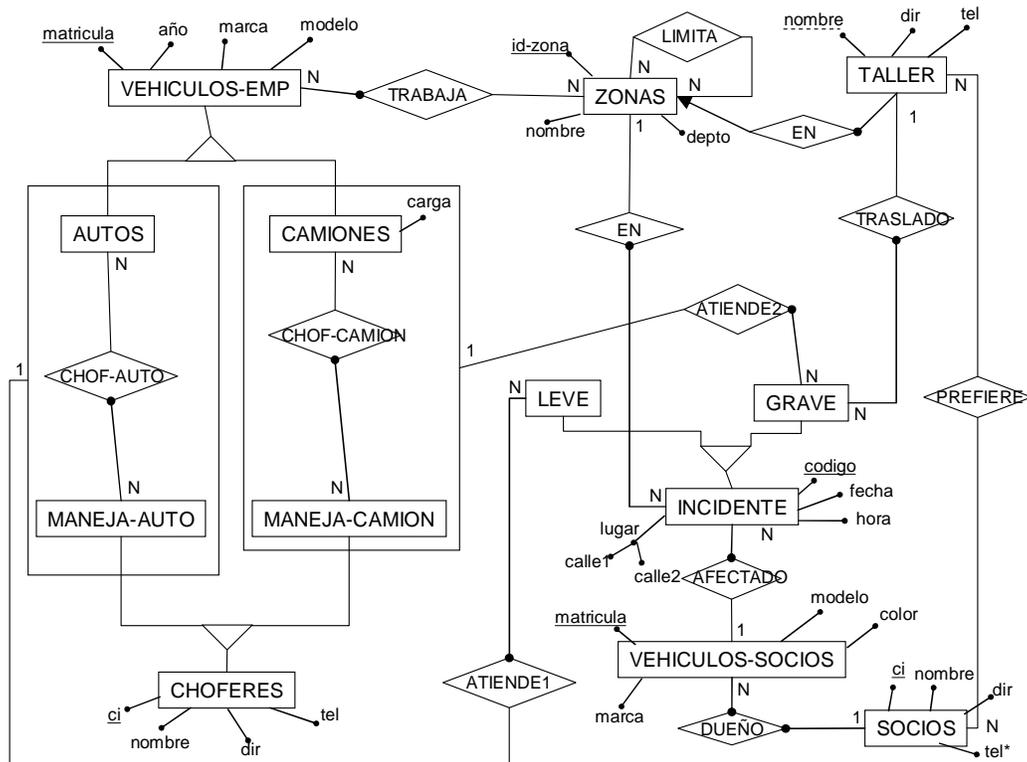
Se registran todos los incidentes que ocurren a los vehículos asociados y generan un pedido de auxilio. Los incidentes se clasifican en leves o graves. Los leves, son los que pueden ser atendidos en el lugar por un auto de la empresa, mientras que los graves son aquellos donde tiene que ir un camión para trasladar el vehículo socio a un taller.

De los incidentes se registra un código, la fecha, la hora y el lugar donde ocurrió. El lugar está dado por el cruce de dos calles, de las cuales se guarda su nombre. Interesa registrar qué auto o camión atendió el incidente y quién fue el chofer actuante. También interesa registrar qué vehículo socio fue el afectado, y si el incidente fue grave, a qué taller fue hecho el traslado. Del taller se guarda el nombre, el cual junto con la zona a la cual pertenece, lo identifica, dirección y teléfono. Se registra también en qué zona fue el incidente. La zona del incidente debería ser una zona asignada al vehículo de la empresa que lo atendió.

Los socios declaran un conjunto de talleres como de su preferencia. Cuando se traslada un vehículo socio se intenta hacerlo a un taller de su preferencia, aunque no siempre es posible.

Se pide: Modelo Entidad Relación completo.

Solución:



RNE:

- $AUTOS \cup CAMIONES = VEHICULOS-EMP$
- $AUTOS \cap CAMIONES = \emptyset$
- $MANEJA-AUTO \cup MANEJA-CAMION = CHOFERES$
- $LEVE \cup GRAVE = INCIDENTE$
- $LEVE \cap GRAVE = \emptyset$
- LIMITA es irreflexiva.
- $\langle l, \langle a, ma \rangle \rangle \in ATIENDE1, \langle a, z \rangle \in ZONAS \rightarrow \langle l, z \rangle \in EN$
- $\langle g, \langle c, mc \rangle \rangle \in ATIENDE2, \langle c, z \rangle \in ZONAS \rightarrow \langle g, z \rangle \in EN$

Parte 2. Consultas (25 puntos)

Ejercicio 2 (25 pts)

Una federación de ajedrez mantiene la información de los clubes afiliados a la misma, de los jugadores y de los torneos que se organizan en una base de datos con el siguiente esquema:

CLUB (nroClub, nomClub, dirClub, dirigente)

Contiene información de los clubes afiliados a la federación. De cada club se conoce un número que permite identificarlo en la federación, su nombre, dirección de la sede y el nombre del dirigente responsable frente a la federación.

JUGADORES (nroJug, nomJug, nroClub, categoria, elo)

Contiene información de los jugadores registrados en la federación. De cada jugador se conoce su nombre, su categoría, su elo y el número del club al que pertenece y por el cuál esta registrado. A cada uno de estos jugadores la federación le asigna un número que lo identifica.

TORNEOS (nroTorneo, nroClub, tipo, fecha)

Contiene información de los torneos oficiales de la federación.

De estos torneos se conoce un número que lo identifica, el número club encargado de la organización, tipo del torneo y fecha de realización.

PARTICIPACIÓN (nroTorneo, nroJug, puntos)

Contiene información sobre la participación de los jugadores en los torneos registrando en cada caso el total de puntos obtenidos por el jugador.

En este esquema no existen tablas vacías y se cumplen las siguientes dependencias de inclusión.

- $\prod_{\text{nroClub}} (\text{JUGADORES}) \subseteq \prod_{\text{nroClub}} (\text{CLUB})$
- $\prod_{\text{nroClub}} (\text{TORNEOS}) \subseteq \prod_{\text{nroClub}} (\text{CLUB})$
- $\prod_{\text{nroTorneo}} (\text{PARTICIPACION}) \subseteq \prod_{\text{nroTorneo}} (\text{TORNEOS})$
- $\prod_{\text{nroJug}} (\text{PARTICIPACION}) \subseteq \prod_{\text{nroJug}} (\text{JUGADORES})$

Resolver en álgebra relacional las siguientes consultas:

- a) Obtener el nombre de los jugadores que han obtenido más de 0 puntos sólo en los torneos organizados por el club al cual pertenecen.

$$A = \prod_{\text{nroJug}} (\sigma_{\text{puntos} > 0} (\text{JUGADORES} * \text{PARTICIPACIÓN} * \text{TORNEOS}))$$

Nro de los jugadores que han obtenido puntos en torneos organizados por su club.

$$B = \prod_{\$2} (\sigma_{\text{puntos} > 0} ((\text{PARTICIPACIÓN} * \text{TORNEOS}) \triangleright \triangleleft \text{JUGADORES}))$$

\$(2=\$7 ^ \$4≠\$8

Nro de los jugadores que han obtenido puntos en torneos organizados por otros clubes.

$$\text{SOL} = \prod_{\text{nomJug}} ((A - B) * \text{JUGADORES})$$

- b) Obtener el nombre de los clubes que uno de sus jugadores participó en todos los torneos de tipo ABIERTO.

$$A = \prod_{\text{nroTorneo}} (\sigma_{\text{tipo} = \text{ABIERTO}} (\text{TORNEOS}))$$

Nro de los torneos de tipo abierto.

$$B = \prod_{nroJug, nroTorneo} (PARTICIPACION) \% A$$

Jugadores que participaron en todos los torneos abiertos.

$$SOL = \prod_{nomClub} (B * JUGADORES * CLUB)$$

Resolver en cálculo relacional las siguientes consultas:

- c) Obtener el nombre de los jugadores con elo 0 que han obtenido el mayor puntaje en algún torneo en el cual también participaron jugadores con elo mayor a 1200.

```
{ j.nomJug / JUGADORES(j) ^ j.elo = 0 ^
  (∃p)(PARTICIPACION(p) ^ p.nroJug = j.nroJug ^
    (¬(∃p2)(PARTICIPACION(p2) ^ p2.nroTorneo = p.nroTorneo ^
      p2.puntos > p.puntos
    )
  ) ^
  (∃p3) (PARTICIPACION(p3) ^ p3.nroTorneo = p.nroTorneo ^
    (∃j2)(JUGADORES(j2) ^ p3.nroJug = j2.nroJug ^ j2.elo > 1200)
  )
}
```

- d) Obtener el nombre de los dirigentes de los clubes tales que todos sus jugadores participaron en todos los torneos realizados antes del 1/1/2010.

```
{ c.dirigente / CLUB(c) ^ (∃j)(JUGADORES(j) ^ j.nroClub = c.nroClub ^
  (∀j2)(JUGADORES(j2) ^ j2.nroClub = j.nroClub →
    (¬(∃t)(TORNEOS(t) ^ t.fecha < 1/1/2010 ^
      (¬(∃p)(PARTICIPACION(p) ^ p.nroJug = j2.nroJug ^
        p.nroTorneo = t.nroTorneo)
    )
  )
}
```

Resolver en SQL, sin utilizar vistas ni sub-consultas en el FROM, las siguientes consultas:

- e) Obtener el número del jugador, la cantidad de torneos en los que participó y los puntos totales obtenidos para aquellos jugadores que pertenecen a categorías con por lo menos 20 jugadores registrados

```
SELECT P.nroJug, count(*), sum(J.puntos)
FROM PARTICIPACION P NATURAL JOIN JUGADORES J
WHERE J.categoria IN
  (SELECT J1.categoria
   FROM JUGADORES J1
   GROUP BY J1.categoria
   HAVING count(*) > 20
  )
GROUP BY P.nroJug
```

- f) Obtener las parejas (nro de torneo, nombre del club) tales que para ese torneo el club indicado tuvo la mayor cantidad de jugadores participantes en el mismo.

```
SELECT P.nroTorneo, C.nomClub
FROM PARTICIPACION P NATURAL JOIN JUGADORES J NATURAL JOIN CLUB C
GROUP BY P.nroTorneo, J.nroClub, C.nomClub
HAVING count(*) >= ALL
      (SELECT count(*)
       FROM PARTICIPACION P2 NATURAL JOIN JUGADORES J2
       WHERE P2.nroTorneo = P.nroTorneo
       GROUP BY J2.nroClub
      )
```

Parte 3. Diseño Relacional (25 puntos)

Ejercicio 3. (25 pts.)

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

$F = \{B \rightarrow D, CD \rightarrow E, GHD \rightarrow E, E \rightarrow GH, GH \rightarrow C, B \rightarrow A\}$

- Indicar TODAS las claves. Justificar.
- Hallar un cubrimiento minimal de F . Justificar.
- Dada la siguiente descomposición de R :
 $R_1(A, B, C, E)$
 $R_2(B, D, E, G, H)$
 - Verificar si cumple con la propiedad de JSP
 - ¿Cuál es la máxima forma normal en que se encuentra R_1 ?
- Hallar una descomposición de R que esté en BCNF y tenga JSP, aplicando el algoritmo visto en el curso. Determine si dicha descomposición preserva las dependencias funcionales.

SOLUCIÓN:

- (a) Como B no es determinado por otros atributos, debe aparecer en toda clave. Calculamos su clausura:

$B^+ = \{A, B, D\}$, por lo que B no es clave.

A no aparece en ningún lado derecho, por lo que no puede estar en ninguna clave.

Calculamos clausuras de pares de atributos posibles que cumplen estas condiciones:

$BA^+ = \{A, B, D\}$, como $B \rightarrow A$, agregar A no aporta nada

$BC^+ = \{A, B, C, D, E, G, H\}$, BC es clave

$BD^+ = \{A, B, D\}$, como $B \rightarrow D$, agregar D no aporta nada

$BE^+ = \{A, B, C, D, E, G, H\}$, BE es clave

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

$BH^+ = \{A, B, D, H\}$

Clausuras de atributos de a tres (sin C o E , porque ya son parte de otras claves, ni A o D , porque no aportan si ya se tiene B):

$BGH^+ = \{A, B, C, D, E, G, H\}$, BGH es clave

- (b) Aplicamos el algoritmo.

Paso 1.

$F' = F$

Paso 2.

$F' = \{B \rightarrow D, CD \rightarrow E, GHD \rightarrow E, E \rightarrow G, E \rightarrow H, GH \rightarrow C, B \rightarrow A\}$

Paso 3 (eliminar atributos redundantes).

Se chequean atributos redundantes en las dependencias con más de un atributo del lado izquierdo: $CD \rightarrow E$, $GHD \rightarrow E$ y $GH \rightarrow C$

E no está en C^+ ni en D^+ , por lo tanto no hay atributos redundantes en $CD \rightarrow E$.

E no está en HD^+ ni en GD^+ ni en GH^+ , por lo tanto no hay atributos redundantes en $GHD \rightarrow E$.

C no está en G^+ ni en H^+ , por lo tanto tampoco hay atributos redundantes en $GH \rightarrow C$.

No hay atributos redundantes en F' .

Paso 4 (eliminar dependencias redundantes).

GHD^+ respecto a $F' - (GHD \rightarrow E) = \{G, H, D, C, E\}$, por lo que $GHD \rightarrow E$ es redundante y se elimina.

Resultado: $F' = \{B \rightarrow D, CD \rightarrow E, E \rightarrow G, E \rightarrow H, GH \rightarrow C, B \rightarrow A\}$

(c)

1.

Aplicamos la siguiente propiedad vista en teórico:

$D = (R_1, R_2)$ de R tiene JSP respecto a F sobre R si y solo si:

- la $df (R_1 \cap R_2) \rightarrow (R_1 - R_2)$ está en F^+

ó

- la $df (R_1 \cap R_2) \rightarrow (R_2 - R_1)$ está en F^+

En este caso

$(R_1 \cap R_2) = \{B, E\}$

BE es clave así que $BE \rightarrow R_1 - R_2$ y $BE \rightarrow R_2 - R_1$ están ambas en F^+ .

Por lo tanto la descomposición tiene JSP.

2.

$B \rightarrow A$ se proyecta en R_1 , y como A es un atributo no primo y B es parte de una clave, esta dependencia viola 2NF, por lo que R_1 está en 1NF.

(d) Aplicamos el algoritmo visto en el teórico.

$D := \{R\}$,

R no está en BCNF, $B \rightarrow D$ viola BCNF porque B no es superclave.

$D := \{R_1(B, D), R_2(A, B, C, E, G, H)\}$,

R_1 está en BCNF por tener dos atributos. R_2 no está en BCNF porque $E \rightarrow G$ viola BCNF.

$D := \{R_1(B, D), R_{21}(E, G), R_{22}(A, B, C, E, H)\}$,

R_{21} está en BCNF por tener dos atributos. En R_{22} $E \rightarrow H$ viola BCNF.

$D := \{R_1(B, D), R_{21}(E, G), R_{221}(E, H), R_{222}(A, B, C, E)\}$,

R_{221} está en BCNF por tener dos atributos. En R_{222} $B \rightarrow A$ viola BCNF.

$D := \{R_1(B, D), R_{21}(E, G), R_{221}(E, H), R_{2221}(A, B), R_{2222}(B, C, E)\}$,

R_{2221} está en BCNF. En R_{2222} $E \rightarrow C$ viola BCNF.

$D := \{R_1(B, D), R_{21}(E, G), R_{221}(E, H), R_{2221}(A, B), R_{22221}(C, E), R_{22222}(B, E)\}$,

R_{22221} y R_{22222} están ambas en BCNF por tener dos atributos.

Resultado $\{R_1(B, D), R_2(E, G), R_3(E, H), R_4(A, B), R_5(C, E), R_6(B, E)\}$

$CD \rightarrow E$, $GH \rightarrow C$ y $GHD \rightarrow E$ no están en $\{B \rightarrow D, E \rightarrow G, E \rightarrow H, B \rightarrow A, E \rightarrow C\}^+$ por lo que no se preservan las dependencias.

Parte 4 Optimización (25 puntos)

Ejercicio 4. (15 pts).

Dado el siguiente esquema:

BOLETOS (NroBoleto, IdParadaSube, Fecha, Compañía, Línea)

PARADAS (IdParada, CoordX, CoordY, Calle)

$\Pi_{IdParada}$ (BOLETOS) \square $\Pi_{IdParada}$ (PARADAS)

y la siguiente consulta:

```
SELECT *
FROM BOLETOS B, PARADAS P
WHERE P.IdParada = B.IdParadaSube AND
      B.Línea = '117'
```

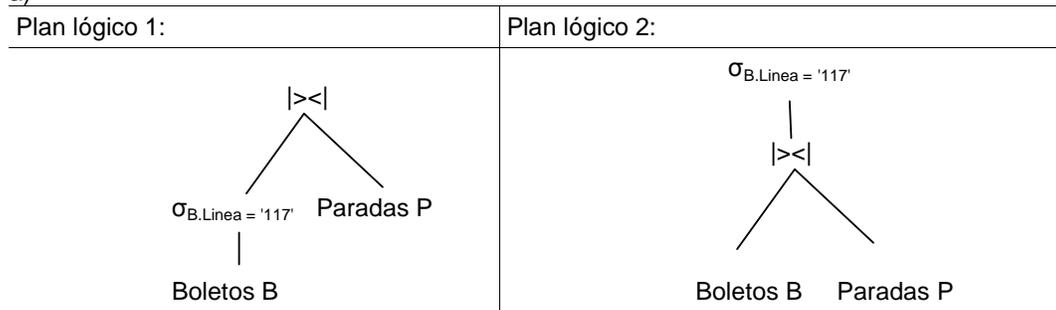
- Dar 2 planes lógicos posibles para la consulta.
- Dar un plan físico asociado a cada plan lógico dado.
- Calcular los costos de los planes físicos dados y compararlos. No olvidar considerar los costos de grabar para los resultados intermedios. Considere que se dispone de 3 buffers.

Tabla	Columna	Valores distintos
Boletos	IdParadaSube	1000
Boletos	Línea	20 (distr. uniforme)
Paradas	IdParada	1000

Tabla	Cant. tuplas	Factor de bloqueo
Paradas	1000	4
Boletos	10000	5
Paradas \bowtie Boletos		3

Indice	Tabla/Atributo	Tipo	Cant. niveles
Parada_IdParada	Paradas/IdParada	Primario	1
Boleto_Línea	Boletos/Línea	Secundario B+	1

a)



b) Dar un historia entrelazada de T1 y T2 con bloqueos y desbloqueos, tal que las transacciones sigan 2PL. Utilizar bloqueos de escritura y de lectura. Decir si la historia es serializable, justificando.

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

La historia es serializable, ya que toda historia formada por transacciones que siguen 2PL es serializable.