

## **FUNDAMENTOS DE BASES DE DATOS**

**Diciembre 2005**

### **Solución**

#### **Parte 1. Modelo Entidad-Relación (25 puntos)**

##### **Ejercicio 1. (25 pts)**

Se desea diseñar una base de datos que dé soporte a un sitio WEB de venta de discos compactos de música. Dicha base deber permitir explorar la información desde diferentes puntos de vista.

Los discos son creaciones de artistas. De cada artista se conoce un código que lo identifica, su nombre, apodo, fecha y lugar de nacimiento. Los artistas pueden ser cantantes, músicos o compositores (aunque algunos artistas pueden desempeñar mas de uno de estos roles). De los cantantes se conoce el registro de su voz mientras que de cada músico se conoce la lista de instrumentos que toca. De cada instrumento se conoce su nombre (que lo identifica) y su tipo (viento, cuerdas, etc)

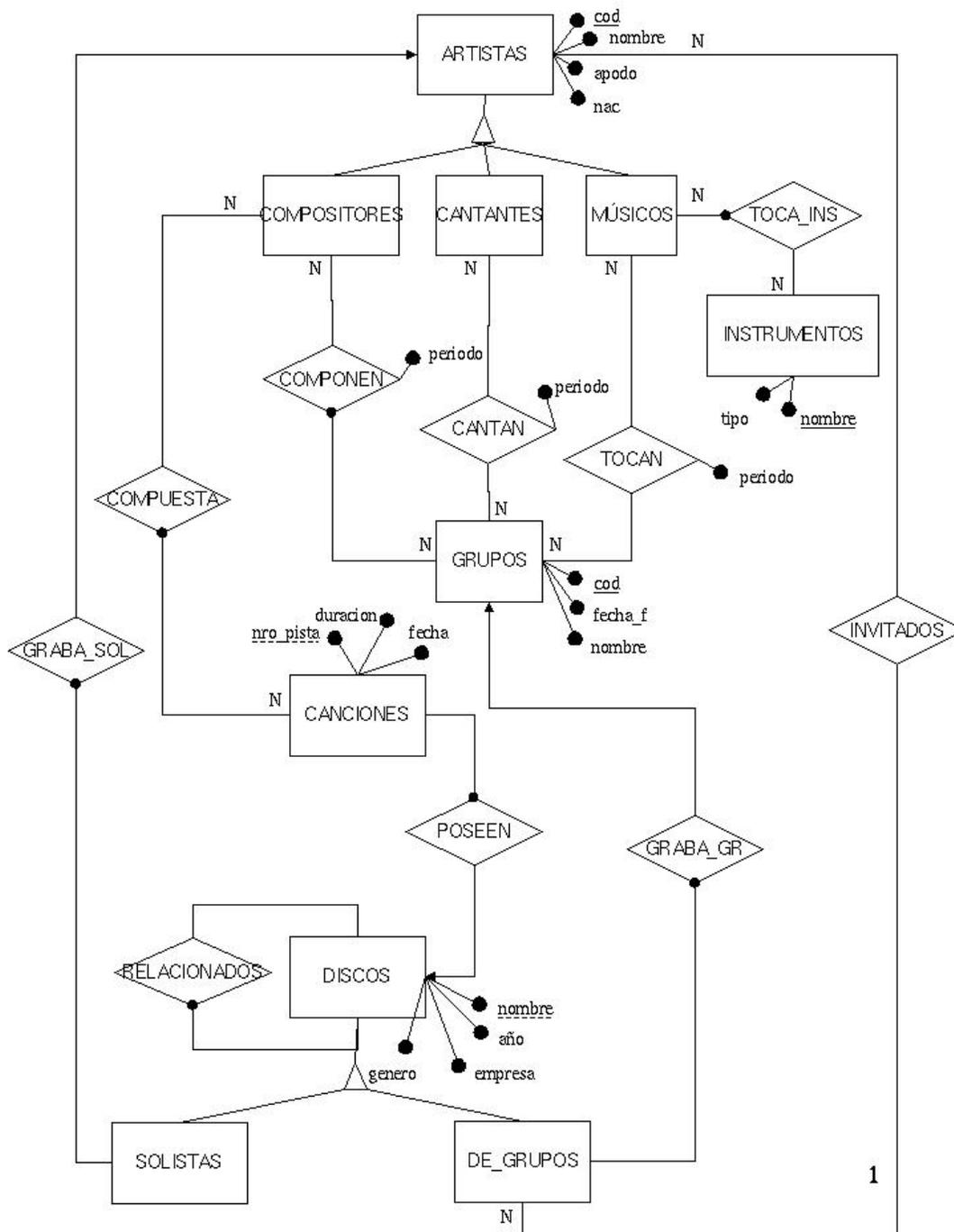
Los discos pueden ser grabados por un grupo o por una única persona (disco solista). Cada grupo se identifica mediante un código y se conoce el nombre del grupo y la fecha de su formación. De cada grupo interesa saber la lista de músicos, cantantes y compositores que lo forman. Puede haber grupos sin músicos y grupos sin cantantes pero todos los grupos deben tener al menos un compositor. En todos los casos para cada artista del grupo interesa saber el período en que formó parte del mismo. Los artistas pueden pertenecer a mas de un grupo a la vez, incluso pueden pertenecer a un mismo grupo en diferentes roles. También puede darse el caso de que un artista forme parte de un grupo y al mismo tiempo grabe discos como solista.

De cada disco se conoce su nombre, año de edición, género musical al que pertenece y empresa discográfica que lo produce. Ni los grupos ni los solistas editan más de una vez un disco con igual nombre. En los discos de grupo pueden aparecer artistas invitados, o sea artistas que no pertenecen al grupo pero que colaboran en el disco. Interesa registrar para cada disco de grupo la lista de artistas invitados, en caso de que existan.

En todos los discos existen canciones. De cada canción se conoce el número de pista que le corresponde dentro del disco (EJ track1), su nombre, su duración en minutos y la fecha en que fue grabada. Interesa saber para cada canción la lista de compositores. En el caso de los discos de grupo todas las canciones son compuestas por alguno de los compositores del grupo.

De cada disco se guarda una lista de discos relacionados, lo cual permite promocionar otros discos a los compradores.

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



**RNE:**

- Los artistas invitados a un disco de grupo no pertenecen al grupo.
- Un disco no se relaciona consigo mismo.
- $\text{Discos}_{\text{solistas}} \cap \text{Discos}_{\text{de grupo}} = \emptyset$
- Los compositores de canciones incluidas en discos de grupo son compositores del grupo.

## Parte 2. Consultas (25 puntos)

### Ejercicio 2 (25 pts)

Una empresa de plaza posee un sistema para manejar los proyectos que realiza y el personal asignado a los mismos. Este sistema usa las siguientes tablas:

**PROYECTOS** (cod\_proy, cod\_cliente, fecha\_inicio, tipo)

**PERSONAL** (cod\_func, nombre, fecha\_ingreso)

**TAREAS** (cod\_tarea, descripción, tipo)

**ASIGNACION** (cod\_func, cod\_proy, cod\_tarea)

**REGISTRO\_HORAS** (cod\_func, cod\_proy, fecha, cant\_horas)

Notas:

\*No existen tablas vacías

\* $\Pi_{\text{cod\_proy}}(\text{ASIGNACION}) \subseteq \Pi_{\text{cod\_proy}}(\text{PROYECTOS})$

\* $\Pi_{\text{cod\_proy}}(\text{REGISTRO\_HORAS}) \subseteq \Pi_{\text{cod\_proy}}(\text{PROYECTOS})$

\* $\Pi_{\text{cod\_func}}(\text{ASIGNACION}) \subseteq \Pi_{\text{cod\_func}}(\text{PERSONAL})$

\* $\Pi_{\text{cod\_func}}(\text{REGISTRO\_HORAS}) \subseteq \Pi_{\text{cod\_func}}(\text{PERSONAL})$

\* $\Pi_{\text{cod\_tarea}}(\text{ASIGNACION}) \subseteq \Pi_{\text{cod\_tarea}}(\text{TAREAS})$

Se pide:

a) Resolver las siguientes consultas en **álgebra relacional**.

1) Obtener los códigos de los proyectos, cliente y tipo de los proyectos que sólo tienen asignadas personas que ingresaron en el 2005 a la empresa.

**A =  $\Pi_{\text{cod\_func}}(\sigma_{\text{fecha\_ingreso} \geq 1/1/05}(\text{PERSONAL}))$**

**B =  $\Pi_{\text{cod\_func}}(\text{PERSONAL}) - A$**

**C =  $\Pi_{\text{cod\_proy}}(\text{ASIGNADOS} * A)$**

**D =  $\Pi_{\text{cod\_proy}}(\text{ASIGNADOS} * B)$**

**SOL =  $\Pi_{\text{cod\_proy}, \text{cod\_cliente}, \text{tipo}}((C - D) * \text{PROYECTOS})$**

2) Obtener los nombres de los funcionarios que se encuentran asignados a todos los proyectos de tipo consultoría o que tienen asignadas todas las tareas de tipo *Gestión*.

**A =  $\Pi_{\text{cod\_proy}}(\sigma_{\text{tipo}=\text{consultoría}}(\text{PROYECTOS}))$**

**B =  $\Pi_{\text{cod\_tarea}}(\sigma_{\text{tipo}=\text{Gestión}}(\text{TAREAS}))$**

**S1 =  $(\Pi_{\text{cod\_func}, \text{cod\_proy}}(\text{ASIGNACION}) \% A)$**

**S2 =  $(\Pi_{\text{cod\_func}, \text{cod\_tarea}}(\text{ASIGNACION}) \% B)$**

**SOL =  $S1 \cup S2$**



### Parte 3. Diseño Relacional (26 puntos)

#### Ejercicio 3. (13 pts.)

Una empresa de transporte desea implementar una base de datos con la información de su negocio.

La empresa cuenta con ómnibus de los cuales se conoce su matrícula que los identifica, su capacidad y fecha de la última inspección.

Entre el personal de la empresa se cuenta con choferes, de los cuales se conoce su número de funcionario (que los identifica), su nombre y fecha ingreso a la empresa.

Cada ómnibus tiene un único chofer titular asignado pero un chofer puede ser titular en más de un ómnibus.

La empresa realiza un conjunto de recorridos. Cada recorrido tiene un código que lo identifica y una duración estimada.

En cada fecha/horario a cada recorrido se le asigna un ómnibus para realizarlo.

- a) Determinar las dependencias funcionales que se cumplen en esta realidad. Utilice los siguientes nombres de atributos:

MAT\_OM → matrícula de los ómnibus  
CAP\_OM → capacidad de los ómnibus  
FEC\_OM → fecha de inspección de los ómnibus  
NRO\_FUN → número de funcionario de los choferes  
NOM\_CH → nombre de los choferes  
FEC\_CH → fecha de ingreso de los choferes  
COD\_REC → código de los recorridos  
DUR\_REC → duración estimada de los recorridos  
FECHA → fecha/horario

**SOL:**

**mat\_om → cap\_om, fec\_om**

**nro\_fun → nom\_ch, fec\_ch**

**mat\_om → nro\_fun**

**cod\_rec → dur\_rec**

**fecha, cod\_rec → mat\_om**

- b) Sabiendo que la política de trabajo de la empresa incluye:
- No todos los ómnibus realizan todos los recorridos, o sea que cada ómnibus tiene un conjunto de recorridos que realiza.
  - Cada ómnibus tiene un conjunto de choferes suplentes asignados. (atributo: NRO\_FUN\_S)

Suponga que en el diseño de la base de datos se obtiene un esquema con la siguiente información:

$E(\text{MAT\_OM}, \text{CAP\_OM}, \text{COD\_REC}, \text{NRO\_FUN\_S})$

Donde se tiene para cada ómnibus su capacidad, código de recorrido que realiza y número de funcionario de los choferes suplentes.

- i) Indique las dependencias funcionales que se cumplen en este esquema.

**mat\_om → cap\_om**

- ii) Indique si existen y cuáles son las dependencias multivaluadas que se cumplen en este esquema.

$\text{mat\_om, cap\_om} \twoheadrightarrow \text{cod\_rec}$

$\text{mat\_om, cap\_om} \twoheadrightarrow \text{nro\_fun\_s}$

#### Ejercicio 4. (13 pts).

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

$F = \{ADG \rightarrow BE, C \rightarrow B, AB \rightarrow H, G \rightarrow A, DH \rightarrow CG, CG \rightarrow H\}$

**Se pide** (en todos los casos **JUSTIFICAR SU RESPUESTA**)

- a) Calcular todas las claves de  $R$  según  $F$ .

$D$  pertenece a todas las claves por no figurar en el lado derecho de las dependencias funcionales.

$E$  no pertenece a las claves por sólo figurar en el lado derecho de las dependencias funcionales.

Verificó si  $D$  es clave

$D^+ = \{D\}$  por lo tanto  $D$  no es clave

Busco claves de 2 atributos:

$(DA)^+ = \{D,A\}$

$(DB)^+ = \{D,B\}$

$(DC)^+ = \{D,C,B\}$

$(DG)^+ = \{D,G,A,B,E,H,C\}$

Por lo tanto  $DG$  es clave

$(DH)^+ = \{D,H,C,G,A,B,E\}$

Por lo tanto  $DH$  es clave

Determino si existen más claves:

$(ABCDE)^+ = \{A,B,C,D,E,H,G\}$

Por lo tanto hay más claves

Busco claves de 3 atributos:

$(DAB)^+ = \{D,A,B,H,C,G,E\}$

Por lo tanto  $DAB$  es clave

$(DAC)^+ = \{D,A,C,B,H,G,E\}$

Por lo tanto  $DAC$  es clave

Determino si existen más claves:

$(DAE)^+ = \{D,A,E\}$

$$(DBCE)^+ = \{D,B,C,E\}$$

Por lo tanto NO HAY MÁS CLAVES

Claves : DG,DH,DAB,DAC

b) Determinar si F es un conjunto de dependencias minimal.

No es minimal ya que en los lados derechos de las dependencias hay más de un atributo.

c) Hallar una descomposición  $\rho_1$  de R en 3NF con join sin pérdida y preservación de dependencias utilizando el algoritmo visto en el curso.

1. Calculamos un cubrimiento minimal de F.

$$F_1 = \{ADG \rightarrow B \\ ADG \rightarrow E, \\ C \rightarrow B, \\ AB \rightarrow H, \\ G \rightarrow A, \\ DH \rightarrow C \\ DH \rightarrow G, \\ CG \rightarrow H \}$$

¿Atributos redundantes ?

Analizamos  $ADG \rightarrow B$  y  $ADG \rightarrow E$

$$(A)^+_{F_1} = \{A\}$$

$$(D)^+_{F_1} = \{D\}$$

$$(G)^+_{F_1} = \{G,A\}$$

$$(DG)^+_{F_1} = \{D,G,A,B,E,\dots\dots\}$$

Por lo tanto A es redundante en estas 2 dependencias.

$$F_2 = \{DG \rightarrow B \text{ (1)} \\ DG \rightarrow E, \text{ (2)} \\ C \rightarrow B, \text{ (3)} \\ AB \rightarrow H, \text{ (4)} \\ G \rightarrow A, \text{ (5)} \\ DH \rightarrow C \text{ (6)} \\ DH \rightarrow G \text{ (7)}\}$$

$$CG \rightarrow H \text{ (8)}$$

Analizamos la dependencia  $AB \rightarrow H$

$$(A)^+_{F_2} = \{A\}$$

$$(B)^+_{F_2} = \{B\}$$

Por lo tanto no hay atributos redundantes

Analizamos las dependencias  $DH \rightarrow C$  y  $DH \rightarrow G$

$$(D)^+_{F_2} = \{D\}$$

$$(H)^+_{F_2} = \{H\}$$

Por lo tanto no hay atributos redundantes

Dependencias redundantes?

Las posibles dependencias redundantes son: 1) – 3) y 4)-8).

El resto de las dependencias son las únicas formas de determinar a los atributos: E, A, C y G respectivamente por lo tanto no pueden ser redundantes.

$DG \rightarrow B$  ?

$$F_3 = F_2 - \{DG \rightarrow B\}$$

$$(DG)^+_{F_3} = \{D,G,E,A\}$$

Por lo tanto  $DG \rightarrow B$  no es redundante.

$$F_3 = F_2 - \{C \rightarrow B\}$$

$$(C)^+_{F_3} = \{C\}$$

Por lo tanto  $C \rightarrow B$  no es redundante.

$$F_3 = F_2 - \{AB \rightarrow H\}$$

$$(AB)^+_{F_3} = \{A,B\}$$

Por lo tanto  $AB \rightarrow H$  no es redundante.

$$F_3 = F_2 - \{CG \rightarrow H\}$$

$$(CG)^+_{F_3} = \{C, G, B, A, H, \dots\}$$

Por lo tanto  $CG \rightarrow H$  es redundante.

Por lo tanto un cubrimiento minimal de F es:

$$F_{\min} = \{DG \rightarrow B, \\ DG \rightarrow E, \\ C \rightarrow B, \\ AB \rightarrow H, \\ G \rightarrow A, \\ DH \rightarrow C, \\ DH \rightarrow G\}$$

Aplicamos el algoritmo visto en el curso que permite obtener una descomposición en 3NF con join sin pérdida y preservación de dependencias.

R1(DGBE)

R2(CB)

R3(ABH)

R4(GA)

R5(DHCG)

R1 contiene una clave (DG)

Por lo tanto  $\rho_1 = \{R1, R2, R3, R4, R5\}$  es la descomposición buscada.

- d) Determinar si la descomposición hallada en la parte c) se encuentra en BCNF. En caso negativo obtener una, aplicando el algoritmo de BCNF visto en el curso partiendo de  $\rho_1$ . Llamamos  $\rho_2$  a la descomposición de R que se encuentra en BCNF.

Proyecto las dependencias de  $F_{\min}$  en cada uno de los esquemas de  $\rho_1$ .

R1(DGBE)

$DG \rightarrow BE$

Clave: DG

Por lo tanto esta en BCNF

R2(CB)

$C \rightarrow B$

Clave: C

Por lo tanto esta en BCNF

R3(ABH)

$AB \rightarrow H$   
Clave: AB  
Por lo tanto esta en BCNF

R4(GA)  
 $G \rightarrow A$   
Clave: G  
Por lo tanto esta en BCNF

R5(DHCG)  
 $DH \rightarrow CG$   
 $DG \rightarrow HC$   
 $CG \rightarrow H$   
Claves : DH,CG  
Por lo tanto  $CG \rightarrow H$  viola BCNF

Por lo tanto descomponemos R5 utilizando el algoritmo del curso que permite obtener una descomposición en BCNF con join sin pérdida.

R51(CGH)  
 $CG \rightarrow H$   
Claves: CG

R52(D,C,G)  
 $DG \rightarrow C$   
Claves: DG

Por lo tanto  $\rho_2 = \{R1, R2, R3, R4, R51, R52\}$

e) Sea  $J = F \cup \{D \twoheadrightarrow B \mid GE, A \twoheadrightarrow B \mid H\}$ , asociado a R

Determinar si  $\rho_2$  está en 4NF. En caso negativo encontrar una descomposición  $\rho_3$  que se encuentre en 4NF con join sin pérdida partiendo de  $\rho_2$ .

R1(DGBE)  
 $DG \rightarrow BE$   
 $D \twoheadrightarrow B$   
Clave: DG  
Por lo tanto D no es superclave, R1 no esta en 4NF

R2(CB)  
 $C \rightarrow B$   
Clave: C  
No se proyectan dep. multivaluadas no triviales. Por lo tanto esta en 4NF

R3(ABH)

$AB \rightarrow H$

$A \twoheadrightarrow B$

Clave: AB

Por lo tanto A no es superclave, R3 no esta en 4NF

R4(GA)

$G \rightarrow A$

Clave: G

No se proyectan dep. multivaluadas no triviales. Por lo tanto esta en 4NF

R51(CGH)

$CG \rightarrow H$

Claves: CG

No se proyectan dep. multivaluadas no triviales. Por lo tanto esta en 4NF

R52(D,C,G)

$DG \rightarrow C$

Claves: DG

No se proyectan dep. multivaluadas no triviales. Por lo tanto esta en 4NF

R11(DB)

R12(DGE)

R31(AB)

R32(AH)

Por lo tanto  $\rho_3 = \{R11, R12, R2, R31, R32, R4, R51, R52\}$  es la descomposición buscada

## Parte 4 Optimización y Concurrencia (24 puntos)

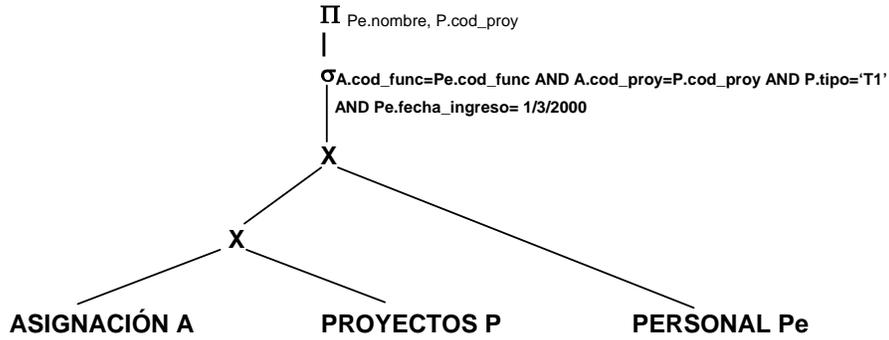
### Ejercicio 5. (12 pts).

Dado el esquema del ejercicio 2 y la siguiente consulta:

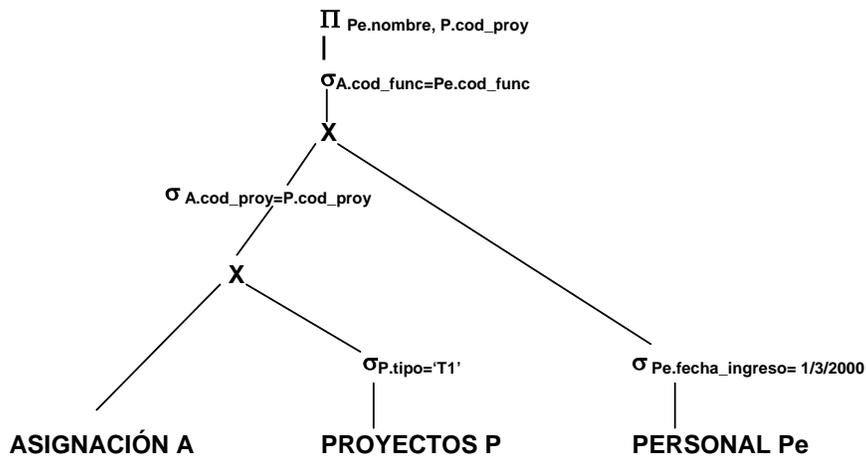
```
SELECT Pe.nombre, P.cod_proy
FROM Asignacion A, Proyectos P, Personal Pe
WHERE A.cod_func = Pe.cod_func AND A.cod_proy = P.cod_proy
      AND P.tipo = 'T1' AND Pe.fecha_ingreso = 1/3/2000
```

- a) Dar un plan lógico de la consulta optimizado según las heurísticas. (Describir los pasos aplicados).

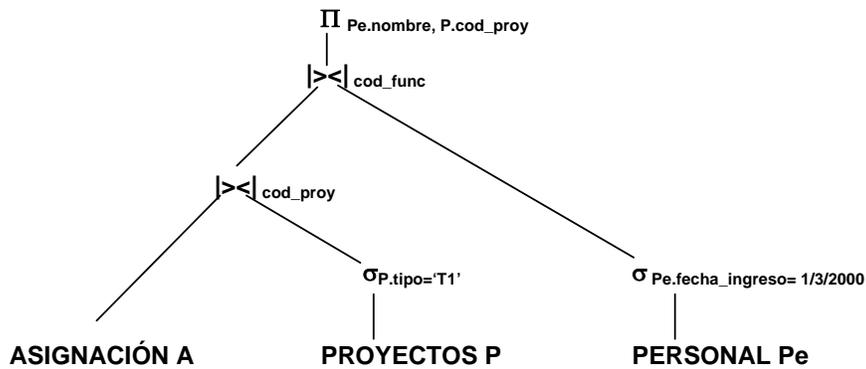
**Arbol canónico:**



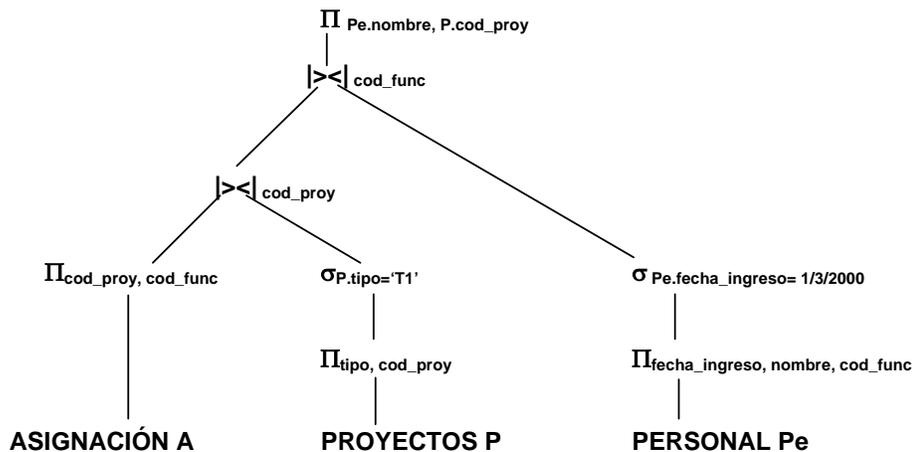
**Bajamos selecciones:**



**Fusionamos productos cartesianos con selecciones:**

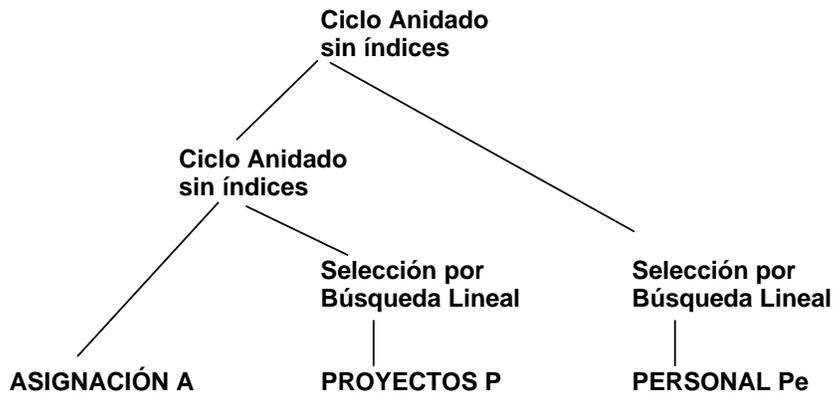


Bajamos proyecciones:



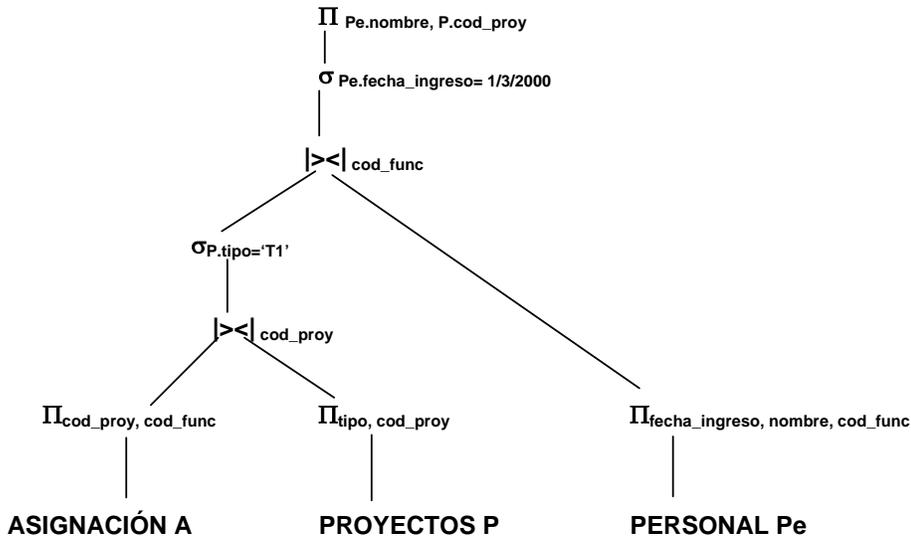
b) Considerando que existe un índice primario Ind\_func por el atributo cod\_func en la tabla Personal y otro índice primario Ind\_proy por el atributo cod\_proy en la tabla Proyectos:

i) Dar un plan físico asociado al plan lógico anterior.

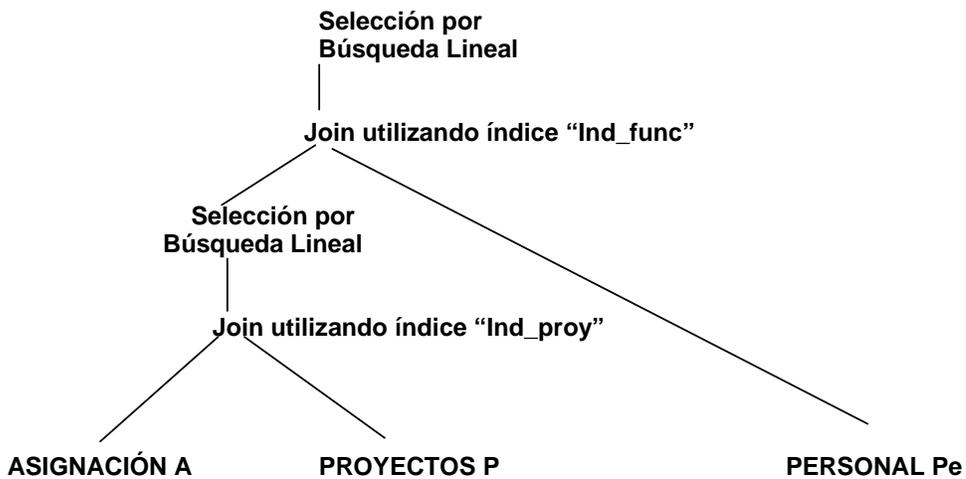


- ii) Proponer otro plan lógico y su correspondiente plan físico, tal que usted considere que puede ser más eficiente que los anteriores, justificando.

**Plan lógico:**



**Plan físico:**



Este plan físico puede ser más eficiente que el de la parte b) porque utiliza los índices para resolver los joins. En la parte b) no se podía utilizar los índices, ya que había una selección antes de cada join, que hacía que el join se ejecutara sobre resultados intermedios y no sobre las tablas originales (por lo tanto los índices dados no servían).

### Ejercicio 6. (12 pts.)

Considerando las siguientes transacciones:

T1: sigue el protocolo 2PL estricto.  
T2: sigue el protocolo 2PL básico.  
T3:  $r_3(X)$ ;  $r_3(Y)$ ;  $w_3(Y)$ ;  $c_3$ .

- a) En caso de que sea posible, agregar los bloqueos y desbloqueos a T3 de manera que cualquier historia que contenga a T1, T2 y T3 sea serializable. Justifique.

**T3:  $rl_3(X)$   $r_3(X)$   $wl_3(Y)$   $u_3(X)$   $r_3(Y)$   $w_3(Y)$   $u_3(Y)$   $c_3$**

**Cualquier historia de T1, T2 y T3 es serializable ya que cada una de ellas cumple 2PL, y esto me asegura la seriabilidad.**

- b) Sabiendo que:

T1:  $r_1(Y)$ ;  $w_1(Z)$ ;  $c_1$ .  
T2:  $r_2(X)$ ;  $w_2(Z)$ ;  $c_2$ .

Notar que T1 y T2 cumplen las condiciones antes mencionadas.

- i) Muestre, para T1 y T2, la secuencia completa de operaciones mostrando bloqueos y desbloqueos (exclusivos y compartidos).

**T1:  $rl_1(Y)$   $r_1(Y)$   $wl_1(Z)$   $w_1(Z)$   $u_1(Y)$   $c_1$   $u_1(Z)$**

**T2:  $rl_2(X)$   $r_2(X)$   $wl_2(Z)$   $u_2(X)$   $w_2(Z)$   $u_2(Z)$   $c_2$**

- ii) Dar una historia entrelazada de T1, T2 y T3.

**H:  $rl_1(Y)$   $r_1(Y)$   $rl_2(X)$   $r_2(X)$   $wl_2(Z)$   $u_2(X)$   $rl_3(X)$   $r_3(X)$   $w_2(Z)$   $u_2(Z)$   $c_2$   $wl_1(Z)$   $w_1(Z)$   $u_1(Y)$   $c_1$   $u_1(Z)$   $wl_3(Y)$   $u_3(X)$   $r_3(Y)$   $w_3(Y)$   $u_3(Y)$   $c_3$**