

FUNDAMENTOS DE BASES DE DATOS

Solución Febrero 2006

Parte 1. Modelo Entidad-Relación (25 puntos)

Ejercicio 1. (25 pts)

Una empresa desea implementar un servicio de comunicación en línea y envío de noticias por Internet y le solicita a UD. el diseño conceptual de la base de datos.

Cuando se crea un nuevo usuario del servicio, se le asigna un identificador, una password, una dirección de mail y una dirección web. Estos usuarios pueden ser personales u organizacionales (de una organización). De estos usuarios además, se guarda una lista de contactos, es decir, de otros usuarios del servicio con los que éste desea comunicarse. Cuando un usuario desea comunicarse con otro por primera vez, se le pide aprobación al segundo. En caso de que el segundo usuario no esté en línea, o rechazó la solicitud, entonces se registra que la aprobación está pendiente o fue rechazada. En el momento de creación del usuario, se registra también el conjunto de idiomas que utiliza.

De los idiomas interesa mantener un código identificador, el nombre y los países que lo hablan.

Se desea llevar además, un registro de qué personas utilizan cada usuario. De las personas se conoce una identificación, el nombre completo y su profesión. Cada usuario personal puede ser usado por una sola persona pero cada usuario organizacional puede ser utilizado por un conjunto de personas distintas debidamente registradas. Para cualquier usuario, siempre se conoce la o las personas que lo utilizan. De las personas, se conoce también su ciudad de nacimiento y una de residencia actual (que no tienen por qué ser las mismas).

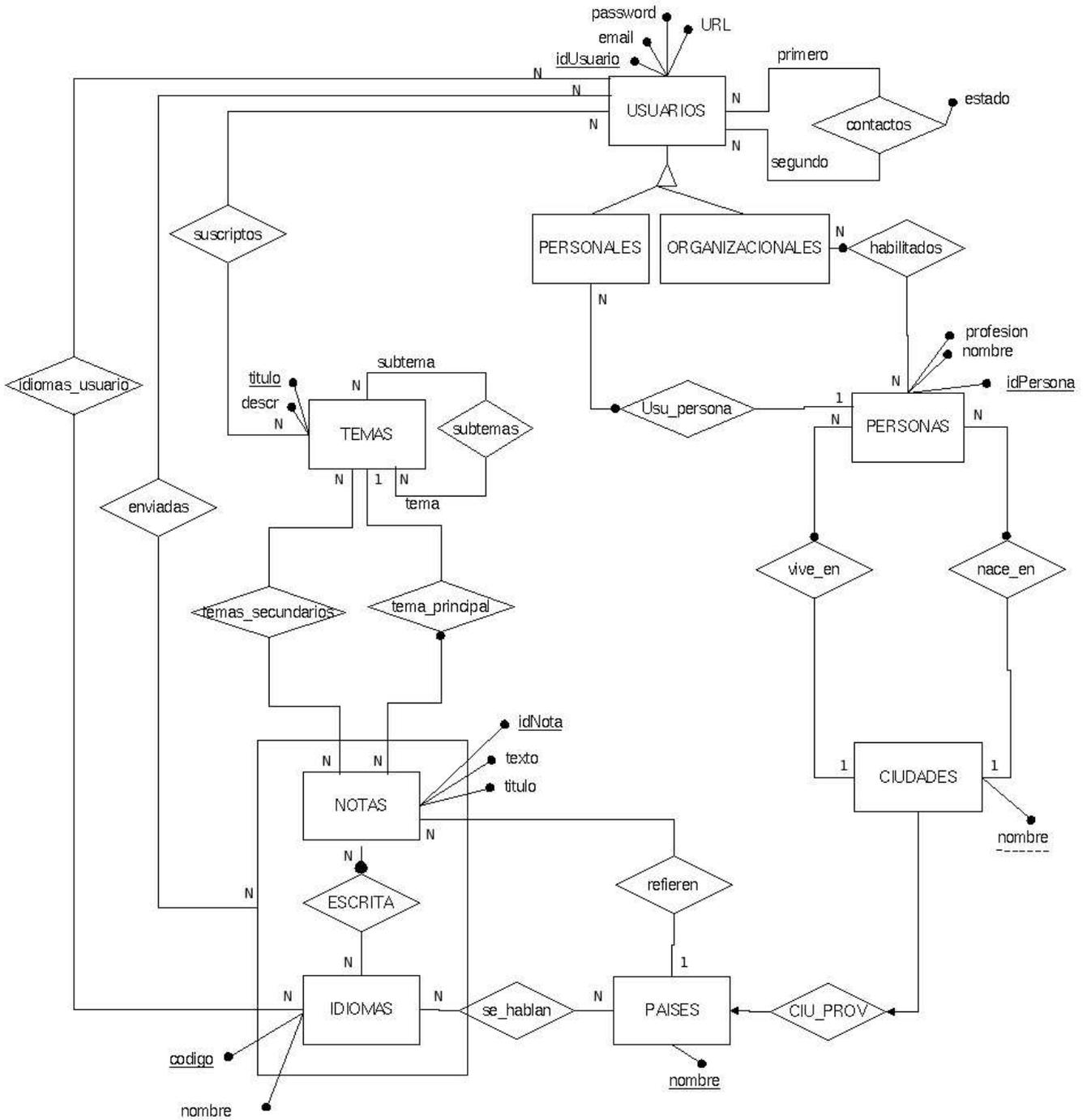
Por otra parte, se recolectan notas de prensa de todo el mundo. De las notas se conoce un identificador, un título y los idiomas en los que está escrita. Todas las notas tienen un tema principal y algunas veces, un conjunto de temas secundarios. De cada tema se conoce un título que lo identifica y una descripción. Por otro lado, los temas pueden tener subtemas.

Las notas, también hacen referencia a un país. De cada país se conoce su nombre, y de cada ciudad también. Tenga en cuenta que existen varios países con una ciudad llamada Córdoba o Montevideo.

Un usuario, puede suscribirse a un conjunto de temas de los que quiere recibir notas de prensa. Se lleva el control de las notas y el idioma en que están escritas, que son enviadas a cada usuario. Las notas enviadas al usuario, siempre deben estar escritas en alguno de los idiomas que están registrados para el usuario y además, su tema principal debe ser alguno de los temas a los que el usuario se suscribió.

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

Solución:



RNE:

- Las notas escritas en un idioma enviadas a un usuario deben estar escritas en alguno de los idiomas del usuario.
- Las notas escritas en un idioma enviadas a un usuario deben tener como tema principal alguno de los temas a los que se suscribió el usuario.
- Un tema no puede ser subtema de él mismo.
- Si un tema es tema principal de una nota no puede ser uno de los temas secundarios de la misma nota.

- Un usuario no puede estar en su lista de contactos.
- Personales \cap Organizacionales = \emptyset
- Personales \cup Organizacionales = Usuarios

Parte 2. Consultas (25 puntos)

Ejercicio 2 (25 pts)

Un servicio de comidas a domicilio ha comenzado la automatización de su sistema de menús y la base de datos cuenta con las siguientes tablas:

NUTRIENTES(cod_nut, nom_nut, desc_nut, categoria_nut)

La tabla **Nutrientes** contiene la lista de nutrientes conocidos indicando un código interno, un nombre (Calcio, Calorías, HDL, LDL, Sacarosa, etc.), una descripción y una categoría (Calorías, Hidrato de Carbono, Grasa, Proteína, Vitamina, Mineral, etc.)

INGREDIENTES (cod_ing, nom_ing, desc_ing, clase, unidad_de_medida)

La tabla **Ingredientes** contiene la lista de ingredientes que se utilizan para realizar las comidas indicando un código interno, un nombre, una descripción, una clase (Ave, Carne Roja, Verdura, Condimento, etc.) y una unidad de medida (gr, cc, etc.).

ING_NUT(cod_ing, cod_nut, cant_nut)

La tabla **Ing_nut** contiene la cantidad de nutriente que el ingrediente tiene cada 100 unidades. En caso que un ingrediente no tenga un nutriente no aparece en esta tabla.

RECETAS(cod_rec, nom_rec, preparacion_rec, cant_porciones, tecnica_basica_rec, complejidad, tiempo_total, categoria_rec)

La tabla **Recetas** es el registro de las recetas que maneja la empresa. De cada receta se conoce un código interno, un nombre, las instrucciones para la preparación, la cantidad de porciones que salen con las medidas indicadas, la técnica básica que se utiliza (horneado, parrilla, frito, guisado, etc.), un índice de la complejidad del 1 al 10, el tiempo total incluyendo la cocción y la preparación de los ingredientes y una categoría para indicar el uso (postre, plato principal, guarnición, entrada o merienda).

LISTA_ING (cod_rec, cod_ing, cant_ing, tecnica_basica_ing, preparacion_ing).

La tabla **Lista_Ing** contiene la lista de ingredientes de cada receta. Para cada ingrediente en cada receta se conoce la cantidad de unidades de ese ingrediente expresada en la unidad de ese ingrediente, la técnica básica con que se trata ese ingrediente en la receta y una descripción de la preparación.

Notas:

- * No existen tablas vacías
- * $\Pi_{\text{cod_nut}}(\text{ING_NUT}) \subseteq \Pi_{\text{cod_nut}}(\text{NUTRIENTES})$
- * $\Pi_{\text{cod_ing}}(\text{ING_NUT}) \subseteq \Pi_{\text{cod_ing}}(\text{INGREDIENTES})$
- * $\Pi_{\text{cod_ing}}(\text{LISTA_ING}) \subseteq \Pi_{\text{cod_ing}}(\text{INGREDIENTES})$
- * $\Pi_{\text{cod_rec}}(\text{LISTA_ING}) \subseteq \Pi_{\text{cod_rec}}(\text{RECETAS})$

Se pide:

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

- 1) Nombre de los platos principales que pueden comer los celíacos, es decir, que ninguno de sus ingredientes tiene el nutriente "gluten".

$$\begin{aligned}
 A &= \prod_{\text{Id_Rec}} (\sigma_{\text{categoría}='plato principal'} (\text{RECETAS})) \\
 B &= \prod_{\text{Cod_Ing}} (\text{ING_NUT} \mid > < \mid \$2=\$4 \wedge \$5='gluten' \text{ NUTRIENTES}) \\
 C &= \prod_{\text{Id_Rec}} (\text{LISTA_ING} * B) \\
 \text{SOL} &= \prod_{\text{nombre}} ((A - C) * \text{RECETAS})
 \end{aligned}$$

- 2) Identificación y nombre de los postres que contienen a todos los nutrientes de categoría "vitamina".

$$\begin{aligned}
 A &= \prod_{\text{cod_nut}} (\sigma_{\text{categoría}='vitamina'} (\text{NUTRIENTES})) \\
 B &= \prod_{\text{Id_Rec}, \text{Cod_Nut}} (\text{LISTA_ING} * \text{ING_NUT}) \\
 \text{SOL} &= \prod_{\text{Id_Rec}, \text{nombre}} (\sigma_{\text{categoría}='Postre'} ((B \% A) * \text{RECETAS}))
 \end{aligned}$$

b) Resolver las siguientes consultas en **cálculo relacional**.

- 3) De las recetas que tienen complejidad 3 o menos, devolver identificación y nombre de las que tienen la cantidad máxima del ingrediente "aceite de oliva".

$$\begin{aligned}
 &\{ t.\text{Id_Rec}, t.\text{nombre} / \text{RECETAS}(t) \wedge t.\text{complejidad} \leq 3 \wedge \\
 &\quad (\exists i)(\text{INGREDIENTES}(i) \wedge i.\text{nombre} = \text{'Aceite de Oliva'} \wedge \\
 &\quad (\exists s)(\text{LISTA_ING}(s) \wedge t.\text{Id_Rec} = s.\text{Id_Rec} \wedge \\
 &\quad \quad i.\text{cod_ing} = s.\text{cod_ing} \wedge \\
 &\quad \quad \neg (\exists x)(\text{RECETA}(x) \wedge x.\text{Id_Rec} \neq t.\text{Id_Rec} \wedge \\
 &\quad \quad \quad x.\text{complejidad} \leq 3 \wedge \\
 &\quad \quad (\exists n)(\text{LISTA_ING}(n) \wedge n.\text{Id_Rec} = x.\text{Id_Rec} \wedge \\
 &\quad \quad \quad n.\text{cod_ing} = i.\text{cod_ing} \wedge \\
 &\quad \quad \quad n.\text{cantidad} > s.\text{cantidad}) \\
 &\quad \quad) \\
 &\quad) \\
 &\}
 \end{aligned}$$

- 4) Identificación y Nombre de las recetas de postres que generan al menos 20 porciones y no contienen grasas.

$$\begin{aligned}
 &\{ t.\text{Id_Rec}, t.\text{nombre} / \text{RECETAS}(t) \wedge t.\text{categoría} = \text{'Postre'} \wedge \\
 &\quad t.\text{cant_porciones} \geq 20 \wedge \\
 &\quad (\forall s)(\text{LISTA_ING}(s) \wedge s.\text{Id_Rec} = t.\text{Id_Rec} \\
 &\quad \rightarrow (\forall x)(\text{ING_NUT}(x) \wedge x.\text{cod_ing} = s.\text{cod_ing} \\
 &\quad \rightarrow \neg (\exists n)(\text{NUTRIENTES}(n) \wedge \\
 &\quad \quad n.\text{cod_nut} = x.\text{cod_nut} \wedge \\
 &\quad \quad n.\text{categoría} = \text{'Grasa'}) \\
 &\quad) \\
 &\quad) \\
 &\}
 \end{aligned}$$

Otra posible solución.

```

{ t.Id_Rec, t.nombre / RECETAS(t) ∧ t.categoria = 'Postre' ∧
  t.cant_porciones >= 20 ∧
  ¬(∃s)(LISTA_ING(s) ∧ s.Id_Rec = t.Id_Rec
    ∧ (∃x)(ING_NUT(x) ∧ x.cod_ing = s.cod_ing
      ∧ (∃n)(NUTRIENTES(n) ∧
        n.cod_nut = x.cod_nut ∧
        n.categoria = 'Grasa'
      )
    )
  )
}

```

c) Resolver la siguientes consulta en **SQL** (sin utilizar vistas)

- 5) Identificación y nombre de las recetas “horneadas”, con el número total de nutrientes de cada categoría de nutriente utilizada en la receta.

```

SELECT R.Id_Rec, R.nombre, N.categoria, count(distinct N.cod_nut)
FROM RECETAS R, LISTA_ING LI, ING_NUT IN, NUTRIENTES N
WHERE R.tecnica_basica = 'horneado' AND
      R.Id_Rec = LI.Id_Rec AND
      LI.cod_ing = IN.cod_ing AND
      IN.cod_nut = N.cod_nut
GROUP BY R.Id_Rec, R.nombre, N.categoria

```

Parte 3. Diseño Relacional (25 puntos)

Ejercicio 3. (8 pts.)

Sea R (A,B,C,D,E,G) un esquema relación, $F = \{A \rightarrow CG, BD \rightarrow E, C \rightarrow DB, E \rightarrow G\}$ el conjunto de dependencias sobre R.

Sea $\rho = (R1, R2, R3, R4)$ una descomposición de R donde:

R1 (A,C,B) R2 (C,D) R3 (B,D,E) R4(E,G)

- a) Determinar si la descomposición ρ es con join sin pérdida respecto a F. Justificar la respuesta.

	A	B	C	D	E	G
R1	a ₁	a ₂	a ₃	b ₁₄	b ₁₅	b ₁₆
R2	b ₂₁	b ₂₂	a ₃	a ₄	b ₂₅	b ₂₆
R3	b ₃₁	a ₂	b ₃₃	a ₄	a ₅	b ₃₆
R4	b ₄₁	b ₄₂	b ₄₃	b ₄₄	a ₅	a ₆

$C \rightarrow DB$

	A	B	C	D	E	G
R1	a ₁	a ₂	a ₃	a ₄	b ₁₅	b ₁₆
R2	b ₂₁	a ₂	a ₃	a ₄	b ₂₅	b ₂₆
R3	b ₃₁	a ₂	b ₃₃	a ₄	a ₅	b ₃₆
R4	b ₄₁	b ₄₂	b ₄₃	b ₄₄	a ₅	a ₆

$E \rightarrow G$

	A	B	C	D	E	G
R1	a_1	a_2	a_3	a_4	b_{15}	b_{16}
R2	b_{21}	a_2	a_3	a_4	b_{25}	b_{26}
R3	b_{31}	a_2	b_{33}	a_4	a_5	a_6
R4	b_{41}	b_{42}	b_{43}	b_{44}	a_5	a_6

$BD \rightarrow E$

	A	B	C	D	E	G
R1	a_1	a_2	a_3	a_4	a_5	b_{16}
R2	b_{21}	a_2	a_3	a_4	a_5	b_{26}
R3	b_{31}	a_2	b_{33}	a_4	a_5	a_6
R4	b_{41}	b_{42}	b_{43}	b_{44}	a_5	a_6

$E \rightarrow G$

	A	B	C	D	E	G
R1	a_1	a_2	a_3	a_4	a_5	a_6
R2	b_{21}	a_2	a_3	a_4	a_5	a_6
R3	b_{31}	a_2	b_{33}	a_4	a_5	a_6
R4	b_{41}	b_{42}	b_{43}	b_{44}	a_5	a_6

b) Determinar si la siguiente afirmación es correcta.

La dependencia $A \rightarrow G$ se pierde en la descomposición ρ .

Justificar la respuesta.

$F = \{A \rightarrow CG, BD \rightarrow E, C \rightarrow DB, E \rightarrow G\}$

R1 (A,C,B)

Entre las df que se proyectan en este esquema se encuentran:

$A \rightarrow C, C \rightarrow B$

R2 (C,D)

Entre las df que se proyectan en este esquema se encuentran:

$C \rightarrow D$

R3 (B,D,E)

Entre las df que se proyectan en este esquema se encuentran:

$BD \rightarrow E$

R4(E,G)

Entre las df que se proyectan en este esquema se encuentran:

$E \rightarrow G$

Sea $J = \{A \rightarrow C, C \rightarrow B, C \rightarrow D, BD \rightarrow E, E \rightarrow G\}$, $J \subseteq \cup_{i=1..4} \Pi_{R_i}(F)$

$A^+_J = \{A, C, B, D, E, G\}$, por lo tanto $A \rightarrow G \in J^+$, $A \rightarrow G \in (\cup_{i=1..4} \Pi_{R_i}(F))^+$

Por lo tanto la dependencia $A \rightarrow G$ se preserva en la descomposición.

Ejercicio 4. (17 pts).

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

$$F = \{H \rightarrow AE, BE \rightarrow CD, AG \rightarrow B, C \rightarrow H, GAC \rightarrow D\}$$

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

a) Para cada uno de los siguientes conjuntos de atributos determinar si son claves de R según F:

i. AE

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

Por lo tanto AE no es clave

ii. GAB

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

Por lo tanto GAB no es clave

iii. GAH

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

Por lo tanto (GAH) es superclave, se debe verificar que es minimal

$$(GA)^+ = \{G,A,B\}$$

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

Por lo tanto (GAH) no es mínima, no es clave

iv. GBE

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

Por lo tanto (GBE) es superclave, se debe verificar que es minimal

$$(GB)^+ = \{G,B\}$$

$$(GE)^+ = \{G,E\}$$

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

No es posible eliminar atributos y que el conjunto siga siendo superclave, por lo tanto es mínima.

(GBE) es clave

v. GC

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

Por lo tanto (GC) es superclave, se debe verificar que es minimal

$$G^+ = \{G\}$$

$$C^+ = \{C,H,A,E\}$$

No es posible eliminar atributos y que el conjunto siga siendo superclave, por lo tanto es mínima.

(GC) es clave

b) Calcular todas las claves de R según F.

Por lo calculado en la parte anterior es posible afirmar que:

GBE,GC son claves

Y que GH es por lo menos superclave

$$G^+ = \{G\}$$

$$H^+ = \{H,A,E\}$$

Por lo tanto GH es clave

G no pertenece al lado derecho de ninguna df, por lo tanto pertenece a todas las claves.

Si existen otras claves deben estar contenidas en : GABD o GADE, cualquier otro conjunto de atributos es superclave.

$$(GABD)^+ = \{G,A,B,D\} \text{ no es superclave}$$

$$(GADE)^+ = \{G,A,D,E,B,C,H\} \text{ por lo tanto es superclave, contiene una clave}$$

$(GA)^+ = \{G,A,B\}$
 $(GD)^+ = \{G,D\}$
 $(GE)^+ = \{G,E\}$
 $(GAD)^+ = \{G,A,D,B\}$
 $(GAE)^+ = \{G,A,E,B,C,D,H\}$
 $(GDE)^+ = \{G,D,E\}$

Por lo tanto GAE es clave y no hay mas claves.

En resumen todas las claves son: GBE,GC,GH,GAE

c) Hallar una descomposición **p1** de R en 3NF con join sin pérdida y preservación de dependencias utilizando el algoritmo visto en el curso.

R no esta en 3NF ya que $BE \rightarrow D$ viola la def. de 3NF

1. Hallar un cubrimiento minimal de F

Sea $F1 = \{$

	$H \rightarrow A,$
	$H \rightarrow E,$
$BE \rightarrow C,$	
$BE \rightarrow D,$	
$AG \rightarrow B,$	
$C \rightarrow H,$	
$GAC \rightarrow D \}$	

1.1. Eliminación de atributos redundantes

Considero $BE \rightarrow C$

$B^+ = \{B\}$

$E^+ = \{E\}$

Por lo tanto no hay atributos redundantes

Considero $BE \rightarrow D$

$B^+ = \{B\}$

$E^+ = \{E\}$

Por lo tanto no hay atributos redundantes

Considero $AG \rightarrow B$

$A^+ = \{A\}$

$G^+ = \{G\}$

Por lo tanto no hay atributos redundantes

Considero $GAC \rightarrow D$

$A^+ = \{A\}$

$G^+ = \{G\}$

$C^+ = \{C,H,A,E\}$

$(GC)^+ = \{G,C,H,A,E,B,D\}$, por lo tanto A es redundante en esta dependencia.

Sea $F2 = \{$

$H \rightarrow A,$	(1)
$H \rightarrow E,$	(2)
$BE \rightarrow C,$	(3)
$BE \rightarrow D,$	(4)
$AG \rightarrow B,$	(5)
$C \rightarrow H,$	(6)
$GC \rightarrow D \}$	(7)

1.2. Eliminación de dependencias redundantes

Las únicas posibles dependencias redundantes son la (4) o (7)

Las restantes dependencias son la única forma de determinar a los atributos que figuran a la derecha de la misma por lo tanto no pueden ser redundantes.

Sea $F_3 = F_2 - \{ BE \rightarrow D \}$

$(BE)^+_{F_3} = \{B,E,C,H,A,E\}$ no incluye a D por lo tanto $BE \rightarrow D$ no es redundante en F_2

Sea $F_4 = F_2 - \{ GC \rightarrow D \}$

$(GC)^+_{F_3} = \{G,C,H,A,E,B,D\}$ incluye a D por lo tanto $GC \rightarrow D$ es redundante en F_2

Por lo tanto un cubrimiento minimal de F es

Sea $F_M = \{$
 $H \rightarrow A,$
 $H \rightarrow E,$
 $BE \rightarrow C,$
 $BE \rightarrow D,$
 $AG \rightarrow B,$
 $C \rightarrow H,$
 $\}$

2. Genero esquemas por cada grupo de df del cubrimiento minimal

R1(H,A,E)
R2(B,E,C,D)
R3(A,G,B)
R4(C,H)

3. Agrego un esquema conteniendo una clave.

R5(G,C)

Por lo tanto $\rho_1 = \{R1,R2,R3,R4,R5\}$

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 ρ_1 .

Llamamos ρ_2 a la descomposición de R que se encuentra en BCNF.

R1(H,A,E)
 $H \rightarrow AE$
Clave: H
BCNF

R2(B,E,C,D)
 $BE \rightarrow CD$
 $C \rightarrow E$ viola BCNF
 $BC \rightarrow ED$
Claves: BE,BC

R1(H,A,E)
Clave: H
E ->> A

viola 4NF

R11(E,A)
Clave: EA
4NF

R12(H,E)
H → E
Clave H
4NF

R3(A,G,B)
Clave: AG
G ->> B

viola 4NF

R31(G,B)
Clave: GB
4NF

R32(G,A)
Clave: GA
4NF

Por lo tanto $\rho_3 = \{R11, R12, R21, R22, R31, R32, R4, R5\}$

Parte 4 Optimización (25 puntos)

Ejercicio 5. (25 pts).

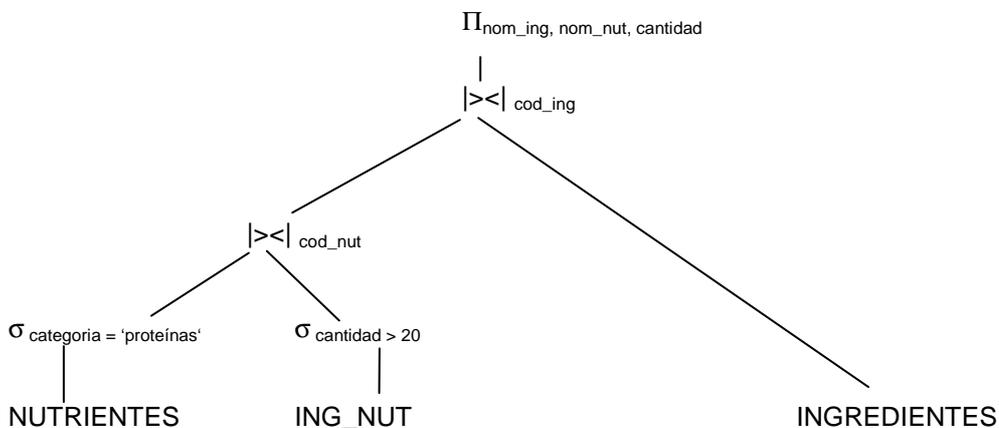
Dadas las siguientes relaciones del esquema del ejercicio 2.

NUTRIENTES(cod_nut, nom_ing, descripción, categoría)

INGREDIENTES (cod_ing, nom_nut, descripción, clase, unidad_de_medida)

ING_NUT(Cod Ing, Cod Nut, cantidad)

y el siguiente plan lógico correspondiente a una consulta sql:



- Calcular los tamaños en cada uno de los resultados intermedios y en el final.
- Dar un plan físico para este plan lógico.
- Calcular el costo total del plan. Considerar costos de grabaciones intermedias, excluyendo la del resultado final. La cantidad de buffers disponibles es 3.

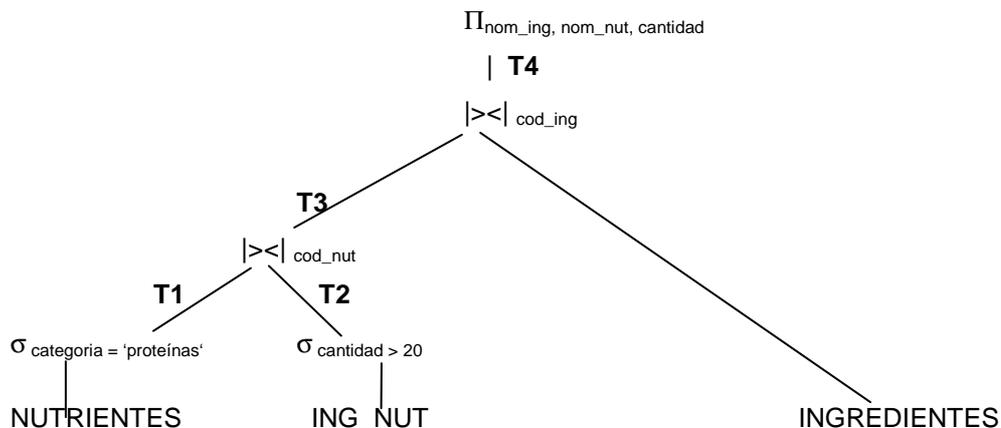
Datos a tener en cuenta:

	NUTRIENTES	ING_NUT	INGREDIENTES
TAMAÑO	100	1500	400
FACTOR DE BLOQUEO	8	10	5
VALORES ATRIBUTOS	10 categorías diferentes, distribuidas uniformemente	El 60% de las tuplas de Ing_Nut tiene la cantidad > 20, y de esas tuplas, la mitad son de nutrientes tipo proteína.	
INDICES	- Índice primario por cod_nut. - Índice secundario por categoría.		- Índice primario por cod_ing

Además:

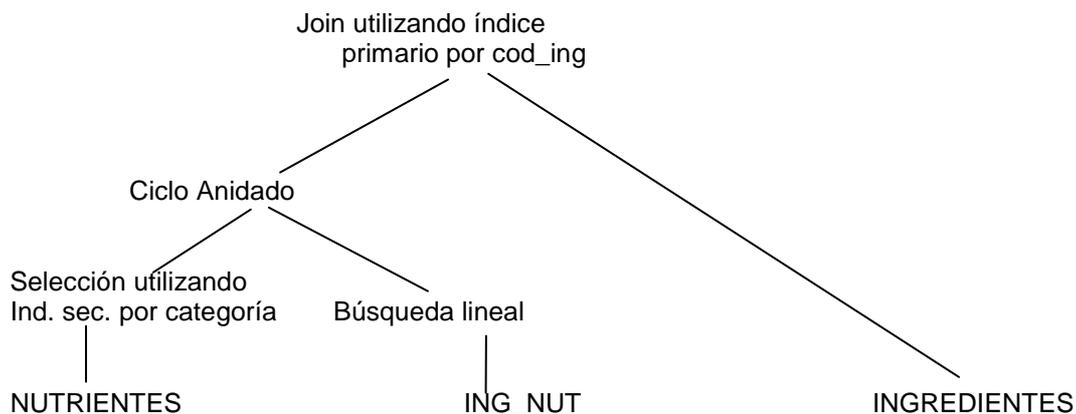
- Factor de bloqueo de **NUTRIENTES|><|ING_NUT = 5**
- En todos los índices nombrados la cantidad de niveles es 3.

Solución:



- a) $T(T1) = 100/10 = 10$
 $T(T2) = 1500 * 0.6 = 900$
 $T(T3) = 450$
 $T(T4) = 450$

b) Plan físico :



- c) $\text{Costo } (T1) = x + s = 3 + 10 = 13$
 $\text{Costo (grabar } T1) = 10 / 8 = 2$
 $\text{Costo } (T2) = |\text{Ing_Nut}| = 1500$
 $\text{Costo (grabar } T2) = 900 / 10 = 90$
 $\text{Costo } (T3) = b_{T1} + b_{T1}/(M-2) * b_{T2} = 2 + 2 * 90 = 182$
 $\text{Costo (grabar } T3) = 450 / 5 = 90$
 $\text{Costo } (T4) = b_{T3} + |T3| (x + 1) = 90 + 450 * 4 = 1890$

 $\text{Costo total} = 13 + 2 + 1500 + 90 + 182 + 90 + 1890 = 3767$