

FUNDAMENTOS DE BASES DE DATOS

Examen Febrero 2006

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)

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 porqué 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.

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

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".

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

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:

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

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

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

Justificar la respuesta.

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 :
- AE
 - GAB
 - GAH
 - GBE
 - GC
- b) Calcular todas las claves de R según F .
- c) Hallar una descomposición ρ_1 de R en 3NF con join sin pérdida y preservación de dependencias utilizando el algoritmo visto en el curso.
- 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.
- e) Sea $J = F \cup \{E \twoheadrightarrow A \mid H, G \twoheadrightarrow B \mid A, B \twoheadrightarrow C \mid ED\}$, asociado a R
Determinar si ρ_2 está en 4NF. En caso negativo encontrar una descomposición ρ_3 que se encuentre en 4NF con join sin pérdida partiendo de ρ_2 .

Parte 4 Optimización (25 puntos)

Ejercicio 5. (25 pts).

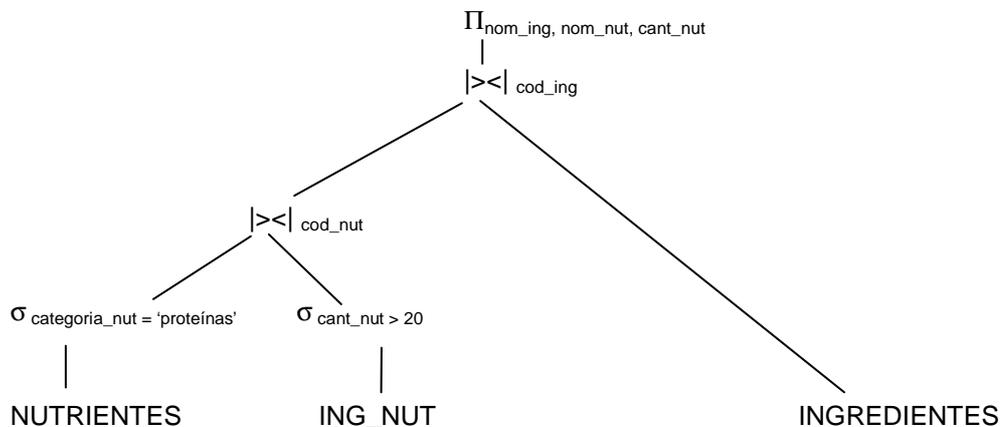
Dadas las siguientes relaciones del esquema del ejercicio 2.

NUTRIENTES(cod_nut, nom_nut, desc_nut, categoria_nut)

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

ING_NUT(cod_ing, cod_nut, cant_nut)

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 categoria_nut.		- Índice primario por cod_ing

Además:

- Factor de bloqueo de $NUTRIENTES \bowtie ING_NUT = 5$
- En todos los índices nombrados la cantidad de niveles es 3.

Implementaciones de los Operadores.

Oper.	Algoritmo	Costo	Condición	Organización
$\sigma_c(R)$	Búsqueda Lineal	b_R peor caso, $b_R/2$ promedio	Cualquier Caso	Cualquiera
	Búsqueda Binaria	$\log_2 b_R + \lceil s/bf_R \rceil - 1$	Cualquier caso	Registros ordenados
	Índice Primario	$x + 1$	Por igualdad a un valor	Registros Ordenados
	Hash	1 o 2 según el tipo	Por igualdad a un valor	Cualquiera
	Índices Primario	$x + (b/2)$ (promedio)	Por relación de orden.	Índice ordenado
	Índice Cluster	$x + \lceil s/bf_R \rceil$	Cualquier Caso	Registros Ordenados

Oper.	Algoritmo	Costo	Cond.	Organización
$\sigma_c(R)$	Índice secundario B+	$x + s$ peor caso	Cualquier Caso	Cualquiera
	Grabación Intermedia	s/bf_R	Cualquier caso	Cualquiera
$R X _cS$	Loop Anidado (registros)	$b_R + (n_R * b_s)$	Cualquier caso	Cualquiera
	Loop Anidado (bloque)	$b_R + \lceil b_R/(M-2) \rceil * b_s$	Cualquier caso	Cualquiera
	Sort Merge	$b_R + b_s + \text{costo ords.}$	Cualquier caso	índice en disco
	Index join	$b_R + n_R * Z$	Cualquier caso	índice en disco

donde Z depende del tipo de índice:

secundario: $Z = x + ss$, cluster: $Z = x + \lceil ss/bfs \rceil$, primario: $Z = x+1$, hash = h