

FUNDAMENTOS DE BASES DE DATOS

Examen Diciembre 2003

SOLUCION

Parte 1. Modelo Entidad-Relación (30 puntos)

Ejercicio 1. (30 pts)

Se quiere modelar la realidad relativa a una clínica odontológica. La clínica está compuesta por varios locales de atención, identificados por su nombre, de los cuales se conoce además su dirección dada por la ciudad donde se ubica, la calle y el número.

En cada local existen varios consultorios que se identifican por un número dentro del local y en cada consultorio existe cierto equipamiento. Dicho equipamiento se identifica globalmente mediante un número de serie, se conoce el tipo (torno, laser, etc.) e interesa mantener registro de la última fecha en que se le realizó mantenimiento.

La clínica posee dos planes diferentes de afiliación: individual y grupal. De los afiliados se conoce la CI, el nombre y uno o más teléfonos. Para los afiliados grupales interesa saber el nombre del convenio de afiliación y el porcentaje de rebaja que se debe aplicar a la cuota mensual.

En la clínica se realizan tratamientos, los cuales se identifican por su nombre y tienen un costo asociado.

Los odontólogos que trabajan en la clínica se identifican por su nombre. De ellos se conoce su especialidad principal dentro de la odontología y los diferentes tratamientos que pueden realizar.

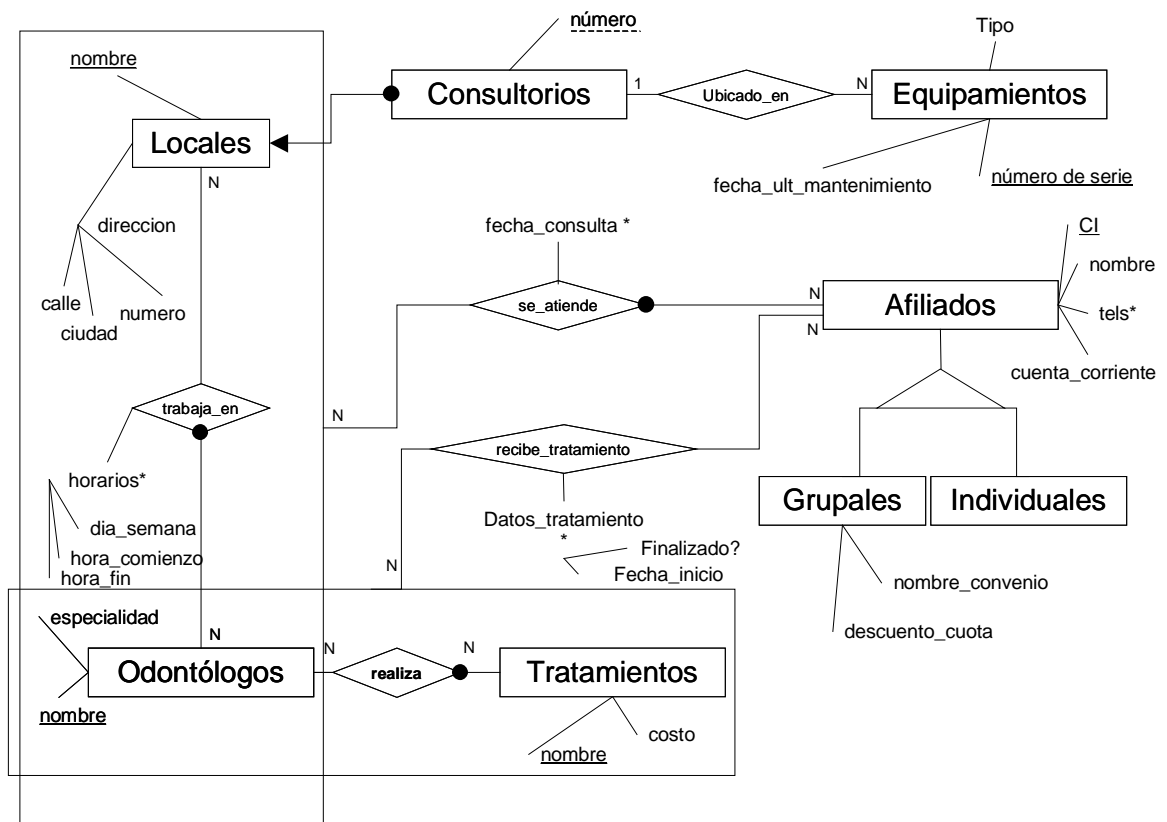
Los odontólogos trabajan en diferentes locales y cada odontólogo puede tener distintos horarios de atención en cada local. De cada horario de atención se conoce el día de la semana, la hora de comienzo y la hora de finalización. (EJ: lunes de 16:00 a 18:30).

Los afiliados se atienden con determinados odontólogos en determinado local y además los odontólogos les realizan tratamientos. Para que un paciente pueda recibir tratamiento de un odontólogo debe ser previamente atendido por este.

Interesa mantener la historia clínica de cada afiliado, la cual consiste, por un lado, en un registro de cada consulta indicando la fecha de consulta, el odontólogo y el local y por otro en un registro de todos los tratamientos que se le han realizado. De cada tratamiento interesa saber: fecha de inicio, si el tratamiento ha sido finalizado o no y la identificación del odontólogo que lo realizó, teniendo en cuenta que como política de la clínica un odontólogo sólo puede practicar un tratamiento por vez a cada afiliado. El inicio de un tratamiento, siempre se hace en una consulta.

Para cada afiliado se mantiene una cuenta corriente donde se incluyen los costos de todos los tratamientos que han sido finalizados. Esta cuenta corriente es global a la clínica.

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



RNE:

Un afiliado puede recibir un tratamiento de un odontólogo sólo si se_atiende con dicho odontólogo en alguna fecha_consulta menor o igual que la fecha_inicio del tratamiento.

Si un afiliado recibe tratamiento de un odontólogo y este tratamiento no ha finalizado, no puede recibir otro tratamiento del mismo odontólogo con fecha_inicio mayor o igual que la del tratamiento no finalizado.

(FALTAN RNE GRUPALES Ç INDIVIDUALES = Æ)

Parte 2. Diseño Relacional (20 puntos)

Ejercicio 2 (10 pts)

Nota: En todos los casos se debe justificar la respuesta.

Sea $R(A,B,C,D,E,G,H,I)$ y $M = \{ GH \rightarrow AB, C \rightarrow AG, D \rightarrow EB, E \rightarrow HI, B \rightarrow CD, A \rightarrow I, G \rightarrow DE \}$ el conjunto de dependencias sobre R .

a) Hallar 3 claves distintas de R según M .

No hay atributos que deban pertenecer a todas las claves.

A no pertenece a ninguna clave (no está del lado izquierdo de ninguna df)

Considero posibles claves de un elemento.

$B^+ = \{ B, C, D, A, E, G, H, I \}$ por lo tanto B es clave

$C^+ = \{ C, A, G, \}$ - no es clave

$D^+ = \{ D, E, B, H, I, C, A, G \}$ por lo tanto D es clave

$E^+ = \{ E, H, I, \}$ por lo tanto E no es clave

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

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

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

Considero posibles claves de 2 elementos.

$GH^+ = \{ G, H, A, B, C, D, E, I \}$ por lo tanto es clave.

Tres claves de R según M son B, D, GH .

b) ¿ R se encuentra en BCNF según M ?

No. Por ejemplo en la dependencia funcional $C \rightarrow A$ que pertenece a M , C no es superclave, por lo tanto R no está en BCNF.

c) Sea $r = \{ R1(A,B,C,D), R3(B,E,G,H,I) \}$ una descomposición de R

1. ¿Es r una descomposición con join sin pérdida según M ?

Aplicaremos el teorema para determinar jsp en descomposiciones en dos esquemas.

$$R1 \cap R3 = B$$

$$R1 - R3 = ACD$$

$$R3 - R1 = EGH$$

En este caso como B es clave de R según M pertenecen a M^+ tanto $B \rightarrow ACD$ como $B \rightarrow EGH$. Por lo tanto la descomposición es con JSP.

2. ¿ r preserva la dependencia $GH \rightarrow A$?

$GH \rightarrow B$ se proyecta en $R3$, $B \rightarrow A$ se proyecta en $R1$, por lo tanto $GH \rightarrow A$ se preserva.

3. ¿ r preserva la dependencia $C \rightarrow G$?

Calculo C^+ con respecto a $M1 \cup M3$, con $M1 = \{ B \rightarrow ACD, D \rightarrow B, C \rightarrow A \}$

$M3 = \{ B \rightarrow EGH$

$EG \rightarrow B$

$GH \rightarrow B$

$E \rightarrow HI \}$

$C^+ = \{ C, A \}$ Por lo tanto la dependencia $C \rightarrow G$ se pierde en la descomposición-

d) Sea $F = M - \{ A \twoheadrightarrow I, G \twoheadrightarrow DE \}$. ¿Es F un conjunto minimal?

No.

Por lo calculado en la parte a) sabemos que en $GH \rightarrow A$ no hay atributos redundantes.

Calculamos GH^+ con respecto a $F - \{GH \rightarrow A\}$. $GH^+ = \{G, H, B, C, D, A, E, I\}$. Por lo tanto la dependencia $GH \rightarrow A$ es redundante en F , F no es minimal.

e) Hallar una descomposición de R en 4NF con join sin pérdida según M .

$R(A, B, C, D, E, G, H, I)$ y

$M = \{$
 $GH \rightarrow A,$
 $GH \rightarrow B,$
 $C \rightarrow A,$
 $C \rightarrow G,$
 $D \rightarrow E,$
 $D \rightarrow B,$
 $E \rightarrow H,$
 $E \rightarrow I,$
 $B \rightarrow C,$
 $B \rightarrow D,$
 $A \twoheadrightarrow I,$

$G \twoheadrightarrow DE \}$

La dependencia $A \twoheadrightarrow I$ viola 4NF ya que A no es superclave.

$R_1 (A I)$

$R_2 (ABCDEGH)$

En R_2 se proyecta $G \twoheadrightarrow DE$ que viola 4NF

$R_{21} (GDE)$

$R_{22} (ABCGH) \quad GH \rightarrow AB, C \rightarrow AG, B \rightarrow CAGH$

La dependencia $C \twoheadrightarrow AG$ viola 4NF.

$R_{221} (CAG) \quad C \rightarrow AG$

$R_{222} (CBH) \quad B \rightarrow CH$

La descomposición resultado es : $R_1, R_{21}, R_{221}, R_{222}$.
 $(AI)(GDE) (CAB) (CAG) (CBH)$.

Ejercicio 3 (10 pts)

Parte A

La persona encargada del diseño de la Base de Datos de una empresa aplicó el siguiente algoritmo a un esquema R y conjunto de dependencias funcionales F:

1. Para cada miembro izq X de una df que aparezca en G crear un er $\{X, A_1, A_2, \dots, A_m\}$, donde $X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_m$ sean las únicas dfs en G con X como miembro izq;
2. Colocar cualquier atributo restante de R en un solo er.

Esta persona afirma que la descomposición resultante de este algoritmo es en 3NF con preservación de dependencias.

Indicar si esta afirmación es cierta o no, justificando en cada caso.

Con lo visto en el curso, la única forma de poder afirmar que se obtiene una descomposición en 3NF con preservación de dependencia es si corresponde al algoritmo visto en el curso.

La diferencia con el algoritmo planteado es que en ese caso no se asegura que el conjunto de dependencias sea minimal.

Por lo tanto la afirmación no es correcta.

Parte B

Dado $R(A,B,C,D,E,G)$ y el conjunto de dependencias sobre el $F = \{ G \rightarrow A, BC \rightarrow G, D \rightarrow B, AE \rightarrow C, B \rightarrow EA, A \rightarrow B \}$

- a) Hallar una descomposición de R en 3NF con join sin pérdida y preservación de dependencias, aplicando los algoritmos vistos en el curso.(justificar los pasos dados).

A) **Primer Paso** - Obtener un cubrimiento minimal de F.

$$F_1 = \{ G \rightarrow A, BC \rightarrow G, D \rightarrow B, AE \rightarrow C, B \rightarrow E, B \rightarrow A, A \rightarrow B \}$$

2) Eliminación de atributos redundantes

2.1. Considero $BC \rightarrow G$

$B^+_{F_1} = \{ B, E, A, C, G \}$, $G \in B^+_{F_1}$, por lo tanto C es redundante en la dependencia considerada.

$$\text{Sea } F_2 = \{ G \rightarrow A, B \rightarrow G, D \rightarrow B, AE \rightarrow C, B \rightarrow E, B \rightarrow A, A \rightarrow B \}$$

2.2. Considero $AE \rightarrow C$

$A^+_{F_2} = \{ A, B, E, C, G \}$, $C \in A^+_{F_2}$, por lo tanto E es redundante en la dependencia considerada.

$$\text{Sea } F_3 = \{ G \rightarrow A, B \rightarrow G, D \rightarrow B, A \rightarrow C, B \rightarrow E, B \rightarrow A, A \rightarrow B \}$$

3) Eliminación de dependencias redundantes

Las dependencias $B \rightarrow G, A \rightarrow C, B \rightarrow E$ no son redundantes ya que son las únicas formas de determinar los atributos de la derecha.

3.1 Considero $G \rightarrow A$

Considero $F4 = \{ B \rightarrow G, D \rightarrow B, A \rightarrow C, B \rightarrow E, B \rightarrow A, A \rightarrow B \}$
calculo $G_{F4}^+ = \{G\}$, $A \notin G_{F4}^+$, por lo tanto la dependencia considerada no es redundante en $F3$.

3.2 Considero $D \rightarrow B$

Sea $F5 = \{ G \rightarrow A, B \rightarrow G, A \rightarrow C, B \rightarrow E, B \rightarrow A, A \rightarrow B \}$
calculo $D_{F5}^+ = \{D\}$, $B \notin D_{F5}^+$, por lo tanto la dependencia considerada no es redundante en $F3$.

3.3 Considero $B \rightarrow A$

Sea $F6 = \{ G \rightarrow A, B \rightarrow G, D \rightarrow B, A \rightarrow C, B \rightarrow E, A \rightarrow B \}$
calculo $B_{F6}^+ = \{B, G, A, E, C\}$, $A \in B_{F6}^+$, por lo tanto la dependencia considerada es redundante en $F3$, continuo con $F6$.

3.4. Considero $A \rightarrow B$

Sea $F7 = \{ G \rightarrow A, B \rightarrow G, D \rightarrow B, A \rightarrow C, B \rightarrow E \}$
calculo $A_{F7}^+ = \{ A, C \}$, $B \notin A_{F7}^+$, por lo tanto la dependencia considerada no es redundante en $F6$.

Conjunto minimal de dependencias: $F = \{ G \rightarrow A, B \rightarrow G, D \rightarrow B, A \rightarrow C, B \rightarrow E, A \rightarrow B \}$

Segundo Paso : Calculo las claves de R según F .

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

Calculo $D^+ = \{ D, B, E, G, A, C \}$. Por lo tanto D es clave y única. (tiene que estar en todas y cualquier otro conjunto que la incluya sería superclave).

Tercer Paso Aplico el algoritmo visto en el curso y obtengo la siguiente descomposición.

$\rho = \{R1 (GA), R2(BGE), R3(DB), R4(ABC)\}$

No es necesario agregar nuevos esquemas ya que todos los atributos de R están en por lo menos un R_i y $R3$ contiene la clave de R . Por lo tanto ρ es la descomposición buscada.

b) Verificar si la descomposición anterior se encuentra en BCNF. Justificar la respuesta.

Proyecto las df en cada uno de los R_i de ρ .

$R1 (GA) G \rightarrow A$
 $R2(BGE), B \rightarrow GE$
 $R3(DB) D \rightarrow B$
 $R4(ABC) A \rightarrow CB$

En cada uno de los esquemas son de la forma tal que todos los lados izquierdos son superclaves, por lo tanto cada R_i esta en BCNF, por lo tanto la descomposición ρ esta en BCNF.

Parte 3. Consultas (30 puntos)

Ejercicio 4 (30 pts)

Al finalizar el año lectivo, los estudiantes llenan un conjunto de encuestas sobre los cursos y los docentes. Estas encuestas son procesadas y los resultados resumidos son almacenados en una base de datos. El siguiente esquema corresponde a los resúmenes de las encuestas realizadas en este año lectivo:

CURSOS (cod_curso, nom_curso, duracion)

Contiene los datos de los cursos considerados.

CARRERA_CURSOS (cod_curso, cod_carrera, semestre, tipo)

Contiene los datos de los cursos por carrera. El atributo *tipo* puede tomar el valor "obligatorio" o el valor "opcional".

ASIGNACION (cod_curso, nro_docente, tipo_tarea)

Contiene los datos de la asignación de los docentes a los cursos. El atributo *tipo_tarea* puede ser "teórico", "práctico", "laboratorio", "otros". Observar que un mismo docente puede estar asignado a varias tareas dentro de un mismo curso.

ENC_CURSO (cod_curso, prom_material, prom_evaluacion, prom_global, cant_enc)

Contiene los datos de las encuestas relativos a los cursos. El atributo *prom_material* es el promedio obtenido en los puntos de la encuesta correspondientes al material. El atributo *prom_evaluación* es análogo al anterior con respecto a la evaluación. El atributo *cant_enc* contiene, para cada curso, la cantidad de estudiantes que llenaron encuestas.

ENC_DOCENTE (cod_curso, nro_docente, prom_actitud, prom_metodologia, cant_enc)

Contiene los datos de las encuestas relativos a los docentes en los cursos. El atributo *prom_actitud* es el promedio obtenido en los puntos de la encuesta correspondientes a la actitud del docente. El atributo *prom_metodología* es análogo al anterior con respecto a la metodología del docente en el curso. El atributo *cant_enc* contiene, para cada docente en cada curso, la cantidad de estudiantes que llenaron encuestas.

NOTA: Suponga que en todos los cursos al menos un estudiante llenó una encuesta para el curso y para cada docente.

Resolver las siguientes consultas en Álgebra Relacional:

- Obtener los códigos de los cursos tales que todos los docentes asignados a ellos tuvieron un promedio_actitud superior a 3, en ese curso.
- Obtener el número de los docentes asignados a todos los cursos opcionales del semestre 2 de la carrera 72.

Resolver las siguientes consultas en Cálculo Relacional:

- Obtener las parejas *nom_curso*, *nro_docente* tales que el valor del atributo *prom_metodología* es el más alto de las encuestas.
- Obtener los códigos de curso tales que todos sus docentes están asignados a él con un solo tipo de tarea, es decir, que hay docentes que están asignados a práctico y otros docentes a teórico, pero no hay docentes en el curso que a la vez estén asignados a práctico y teórico.

Resolver las siguientes consultas en SQL:

- e) Para cada carrera, por semestre obtener el promedio obtenido en sus cursos en `prom_global`, considerando solamente las encuestas de más de 100 estudiantes.
- f) Obtener el número de los docentes tales que obtuvieron el valor máximo de la encuesta en `prom_actitud` para algún curso y el valor mínimo registrado en la encuesta en `prom_actitud` para otro curso.

Parte 4 Optimización y Concurrencia (20 puntos)

Ejercicio 5. (20 pts.)

Dados el esquema relacional del **Ejercicio 4** y la siguiente consulta

```
SELECT C.cod_curso, C.duración
FROM Cursos C, Enc_Docente E, Asignación A
WHERE A.cod_curso = E.cod_curso AND A.nro_docente = E.nro_docente, AND
C.cod_curso = A.cod_curso AND A.tipo = "teorico" AND E.prom_actitud > 3
```

- 1) Dar un plan lógico de la consulta, aplicando las heurísticas y calculando los tamaños intermedios.
- 2) Elegir una implementación para la primer selección y el primer join que se realice según el plan lógico, y calcular su costo estimado. (En el costo se deben incluir los costos de grabar los resultados.)

DATOS:

	CURSOS	ASIGNACION	ENC_DOCENTE
Cantidad tuplas	300	2 500	100 000
Indices primarios	cod-curso (niveles: 2)		(cod-curso, nro- docente) (niveles: 1)
Indices secundarios (B+)		tipo-tarea (niveles: 3)	
Cantidad de tuplas por bloque	50	60	30
Observaciones		- En promedio, se asignan a teorico 2 docentes por curso.	- En prom-actitud los valores van de 1 a 5, y su distribución es uniforme.

	CURSOS >< ASIGNACION	ASIGNACION >< ENC_DOCENTE
Cantidad de tuplas por bloque	30	20

Fórmulas para cálculo de costo y de tamaños:

Notación: b – cantidad de bloques

fbl – factor de bloqueo

x – cantidad de niveles del índice

|R| - cantidad de tuplas de R

s – cardinalidad de la selección (tamaño obtenido)

js – selectividad del join

V(A,R) - cantidad de valores distintos del atributo A en R

Selección (R)	Búsqueda lineal	B
	Búsqueda binaria	$\log_2 b + \lceil s/fbl_R \rceil - 1$
	Índice primario	x + 1
	Índice secundario (B+)	x + s
Join (R,S)	Nested Loop (ciclo anidado) sin utilizar índices	$b_R + b_R * b_S$
	Nested Loop (ciclo anidado) utilizando	$b_R + R * (x + s)$

	índice secundario para recuperar tuplas que matchean	
	Nested Loop (ciclo anidado) utilizando índice primario para recuperar tuplas que matchean	$b_R + R * (x + 1)$
Escribir resultados en disco	Join (R,S)	$(j_s * R * S) / f_{bl_{RS}}$
Selectividad	Selección (atrib. A)	$1 / V(A,R)$
	Join (atrib. A)	$1 / \text{Max}(V(A,R), V(A,S))$