

Clase Auxiliar 10¹

Normalización

La normalización se usa para ver si una tabla está bien o mal diseñada. Una tabla está bien diseñada si no tiene redundancia (datos repetidos) y está mal en caso contrario.

Dependencia Funcional: En una tabla la columna Y depende funcionalmente de la columna X si cada valor de la columna Y está determinado por el valor de la columna X en la misma fila.

Formalmente : $X \rightarrow Y$

Es decir, para todo par de filas en la tabla, tales que tienen igual valor en la columna X, entonces también deben tener igual valor en la columna Y.

Tanto el *determinante* (a la izquierda de la flecha, X) como el *dependiente* (a la derecha de la flecha, Y) pueden ser compuestos (más de una columna).

Las dependencias funcionales son una propiedad del mundo real representado mediante la tabla. No son una propiedad del contenido de la tablas en un instante.

Llave Primaria y Dependencia Funcional: Para toda tabla se cumple que toda columna que no es parte de la llave depende funcionalmente de la llave.

Llave Candidata:

Es un conjunto de atributos que cumple dos propiedades:

1. Unicidad: En cualquier instante dado, no existen dos instancias iguales del conjunto de atributos.
2. Minimalidad: Si se elimina algún atributo del conjunto, se deja de cumplir la condición 1.

Primera Forma Normal: Una tabla está en 1FN si todos los dominios de columnas contienen sólo valores atómicos o escalares.

Segunda Forma Normal: Una tabla está en 2FN si está en 1FN y además se cumple que toda columna que no es parte de la llave de la tabla depende funcionalmente de toda la llave.

Tercera Forma Normal: Una tabla está en 3FN si está en 2FN y además se cumple que toda columna que no es parte de la llave de la tabla depende sólo de la llave, o sea si eliminamos las *dependencias transitivas*.

Forma Normal de Boyce Codd (BCNF): Una relación está en BCNF, si y sólo si sólo los determinantes son llaves candidatas.

¹ Basada en clase auxiliar por Felipe Aguilera V.

Problema #1

El centro de computación de la Facultad posee un grupo de Consultoría dedicado a la atención de usuarios. Dicho grupo está formado por alumnos regulares de la Facultad, los que tienen dominio de algún software específico. Para cada consultor existe un horario de atención preestablecido y el propio jefe del grupo está a cargo de controlar el funcionamiento del servicio.

Se pide normalizar el modelo relacional hasta la 3FN. Para ello determine claramente todas las dependencias funcionales existentes. Es importante que justifique todos sus pasos.

horario					asistencia		
matricula	num_dia	cod_mod	hora_inicio	hora_fin	fecha	matricula	cod_mod
2039	1	1	08,50	10,00	24/11/98	2039	1
	2	2	10,25	11,75	25/11/98	2029	2
2020	3	1	08,50	10,00	25/11/98	2020	1

consultor					
cod_dpto	Nom_dpto	nombre	matricula	cod_soft	descripcion
10	Computación	Alfredo	2039	0001	Excel 5,0
				0002	Word 6,0
				0003	Pascal
				0004	Turing
20	Matemáticas	Claudia	2020	0005	C
				0006	PL/SQL
				0003	Pascal

Problema #2 (basado en P2, C2, 2004/01)

a.) Transforme, PASO A PASO, el siguiente esquema relacional en otros hasta BCNF. Especifique claramente el conjunto de dependencias funcionales de los esquemas resultantes, así como las dependencias que se pierden (en el caso que se pierda alguna).

R (a, b, c, d, e, f, g) { $abc \rightarrow d$; $abc \rightarrow e$; $d \rightarrow b$; $e \rightarrow f$; $bc \rightarrow g$ }

b.) Cual de las siguientes afirmaciones es correcta. Justifique la respuesta. Respuesta sin justificación no tendrá ninguna validez.

- i.) Una relación cuya clave principal es un único atributo está siempre en 2FN.
- ii.) Una relación cuya clave principal es un único atributo está siempre en 2FN y 3FN.
- ii.) Una relación cuya clave es un único atributo está siempre en BCNF.

Solución

Problema #1

horario:

No se encuentra en ninguna forma normal, ya que para estar en primera forma normal, todos los dominios de las columnas deben contener valores atómicos. Por lo tanto, separamos en dos tablas:

1FN:

horario

<u>matrícula</u>	<u>num_día</u>	<u>cod_mod</u>
2039	1	1
2039	2	2
2020	3	1

modulo

<u>cod_mod</u>	hora_inicio	hora_fin
1	8:50	10:00
2	10:25	11:75

2FN: Ahora debemos ver que toda columna que no es parte de la llave dependa funcionalmente de toda la llave. Como horario no tiene columnas que no sean parte de la llave, está en segunda forma normal. Y en módulo, sólo hay una columna conformando la llave, por ende las otras dos columnas dependen de toda la llave.

3FN: Ahora debemos eliminar (si es que existen) las dependencias transitivas en las columnas que no conforman la llave. Como horario no tiene columnas que no sean parte de la llave, está en tercera forma normal. Veamos ahora las dependencias de modulo:

$\text{cod_mod} \rightarrow \text{hora_inicio}$

$\text{cod_mod} \rightarrow \text{hora_fin}$

por lo tanto, está en tercera forma normal.

asistencia:

(*) aquí está la tabla corregida (con la llave definida correctamente):

asistencia

<u>fecha</u>	<u>matricula</u>	cod_mod
24/11/98	2039	1
25/11/98	2039	2
25/11/98	2020	1

1FN: ✓

2FN: Como cada día un alumno tiene un horario determinado de atención, cod_mod depende funcionalmente de la fecha y de la matrícula.

(fecha,matricula) → cod_mod

3FN: Como sólo hay una columna que no es parte de la llave, no pueden haber dependencias transitivas (entre las columnas que no conforman la llave). Por lo tanto, está en tercera forma normal.

consultor:

1FN:

consultor

cod_dpto	nom_dpto	nombre	<u>matricula</u>
10	Computación	Alfredo	2039
20	Matemáticas	Claudia	2020

especialidad

<u>matricula</u>	<u>cod_soft</u>	descripcion
2039	0001	Excel 5.0
2039	0002	Word 6.0
2039	0003	Pascal
2039	0004	Turing
2020	0005	C
2020	0006	PL/SQL
2020	0003	Pascal

2FN: consultor ✓

especialidad:

matricula → cod_soft

cod_soft → descripcion

descripcion no depende de toda la llave, así que debemos separar en dos tablas:

especialidad

<u>matricula</u>	<u>cod_soft</u>
2039	0001
2039	0002
2039	0003
2039	0004
2020	0005
2020	0006
2020	0003

software

<u>cod_soft</u>	descripcion
0001	Excel 5.0
0002	Word 6.0
0003	Pascal
0004	Turing
0005	C
0006	PL/SQL

3FN:

consultor:

matricula → nombre

matricula → cod_dpto

matricula → nom_dpto

cod_dpto → nom_dpto

como nom_dpto depende también de cod_dpto, separamos en dos tablas:

consultor

cod_dpto	nombre	<u>matricula</u>
10	Alfredo	2039
20	Claudia	2020

departamento

<u>cod_dpto</u>	nom_dpto
10	Computación
20	Matemáticas

especialidad ✓

software ✓

Problema #2

a) **1FN:** no podemos decir nada sobre su primera forma normal. Por lo tanto asumiremos que ya está en primera forma normal.

2FN: primero se debe elegir una llave. Como d,e dependen funcionalmente de abc, g de bc y f de e, nos conviene tomar como llave (a,b,c)

R

<u>a</u>	<u>b</u>	<u>c</u>	d	f	e	g
----------	----------	----------	---	---	---	---

Como $(b,c) \rightarrow g$, debemos separar en dos tablas:

R1

<u>a</u>	<u>b</u>	<u>c</u>	d	f	e
----------	----------	----------	---	---	---

R2

<u>b</u>	<u>c</u>	g
----------	----------	---

R1: $\{abc \rightarrow d; abc \rightarrow e; d \rightarrow b; e \rightarrow f\}$

R2: $\{bc \rightarrow f\}$

3FN:

R1: como $e \rightarrow f$, debemos separar R1 en dos tablas:

R11

<u>a</u>	<u>b</u>	<u>c</u>	d	e
----------	----------	----------	---	---

R12

<u>e</u>	f
----------	---

R11: $\{abc \rightarrow d; abc \rightarrow e; d \rightarrow b\}$

R12: $\{e \rightarrow f\}$

(Nótese que $d \rightarrow b$ no se vió afectado por la normalización).

R2 ✓

BCNF:

R11: Como $d \rightarrow b$ y d no es llave candidata, entonces no está en BCNF. Para normalizarlo, debemos separar en dos tablas:

R111

<u>a</u>	<u>b</u>	<u>c</u>	e
----------	----------	----------	---

R112

<u>d</u>	b
----------	---

R111: $\{abc \rightarrow e\}$

R112: $\{d \rightarrow b\}$

Aquí se pierde la df $abc \rightarrow d$, por lo tanto hay una pérdida de información.

R12 ✓

R2 ✓

- b) i) Correcta, puesto que si hay un sólo atributo como llave, siempre dependerán las demás columnas de toda la llave (recordemos que siempre en una tabla los atributos que no conforman la llave, dependen funcionalmente de ella).
- ii) No es correcta, puesto que podría fácilmente haber dependencias transitivas (ver en el problema 1, la tabla consultor en su primera [y segunda] forma normal).
- iii) No es correcta, puesto que en el caso que no esté en BCNF, tampoco estaría en 3FN ($3FN \Rightarrow BCNF$)