

UNIDAD DIDÁCTICA 5

Normalización de las relaciones

DISEÑO LÓGICO DE UNA BASE DE DATOS EN EL MODELO RELACIONAL (Teoría de la Normalización)

CONSISTE EN:

Descomposición sin pérdida de información ni de semántica de la relación universal (o de una colección de relaciones equivalentes a la misma) en una colección de relaciones en la que las anomalías de actualización (inserción, Borrado y modificación) no existan o sean mínimas.

Normalización. Objetivos.

1. Evitar redundancia de datos. De esa forma se evita uso excesivo de espacio y la posibilidad de inconsistencia.
2. Hechos distintos deben almacenarse en lugares distintos. Esto evita ciertas anomalías a la hora de actualizar los datos.

Normalización. Primera forma normal (1FN).

1.- Una tabla está en 1ª Forma Normal si todos sus dominios son atómicos, es decir, en cada campo solo hay un valor y además no se repiten grupos de campos.

Ejemplo de la tabla que no está en 1FN

Cliente

<u>ID Cliente</u>	Nombre	Apellido	Teléfono
123	Rachel	Ingram	555-861-2025
456	James	Wright	555-403-1659 555-776-4100
789	Maria	Fernandez	555-808-9633 555-456-3452

Normalización. Primera forma normal (1FN).

Solución: Los campos que se repiten, se pasan a una nueva tabla junto con la clave primaria de la tabla origen.

Cliente

<u>ID Cliente</u>	Nombre	Apellido
123	Rachel	Ingram
456	James	Wright
789	Maria	Fernandez

Teléfono del cliente

<u>ID Cliente</u>	<u>Teléfono</u>
123	555-861-2025
456	555-403-1659
456	555-776-4100
789	555-808-9633
789	555-456-3452

Nueva tabla

Normalización. Segunda forma normal (2FN).

Una tabla está en 2ª FN si está en 1ª FN y además, todos los atributos que no son clave dependen de todos los atributos de la clave primaria y no de una parte de ella

(Con tablas con una clave primaria compuesta por más de un atributo)

<u>Empleado</u>	<u>Habilidad</u>	Lugar actual de trabajo
Jones	Mecanografía	114 Main Street
Jones	Taquigrafía	114 Main Street
Jones	Tallado	114 Main Street
Bravo	Limpieza ligera	73 Industrial Way
Ellis	Alquimia	73 Industrial Way
Ellis	Malabarismo	73 Industrial Way
Harrison	Limpieza ligera	73 Industrial Way

En este ejemplo “Lugar actual de trabajo” sólo depende de “Empleado”.

Normalización. Segunda forma normal (2FN).

Solución: Se sacan los atributos no claves que dependen solo de la parte de la clave primaria a una nueva tabla con la parte de la clave primaria de la que dependen.

Empleados

<u>Empleado</u>	Lugar actual de trabajo
Jones	114 Main Street
Bravo	73 Industrial Way
Ellis	73 Industrial Way
Harrison	73 Industrial Way

Habilidades de los empleados

<u>Empleado</u>	<u>Habilidad</u>
Jones	Mecanografía
Jones	Taquigrafía
Jones	Tallado
Bravo	Limpieza ligera
Ellis	Alquimia
Ellis	Malabarismo
Harrison	Limpieza ligera

Normalización. Tercera forma normal (3FN).

Una tabla está en 3ª FN si está en 2ª FN y se comprueba que los atributos que no forman parte de la clave no dependen de otros atributos que no son clave.

Ganadores del torneo

<u>Torneo</u>	<u>Año</u>	<u>Ganador</u>	<u>Fecha de nacimiento del ganador</u>
Indiana Invitational	1998	Al Fredrickson	21 de julio de 1975
Cleveland Open	1999	Bob Albertson	28 de septiembre de 1968
Des Moines Masters	1999	Al Fredrickson	21 de julio de 1975
Indiana Invitational	1999	Chip Masterson	14 de marzo de 1977

En este ejemplo “Fecha de nacimiento del ganador” (atributo no clave) depende de Ganador (atributo no clave).

Normalización. Tercera forma normal (3FN).

Solución: Se sacarán a una tabla nueva los atributos no claves dependientes de otro atributo no clave junto a este último.

Ganadores del torneo

<u>Torneo</u>	<u>Año</u>	Ganador
Indiana Invitational	1998	Al Fredrickson
Cleveland Open	1999	Bob Albertson
Des Moines Masters	1999	Al Fredrickson
Indiana Invitational	1999	Chip Masterson

Fecha de nacimiento del jugador

<u>Jugador</u>	Fecha de nacimiento
Chip Masterson	14 de marzo de 1977
Al Fredrickson	21 de julio de 1975
Bob Albertson	28 de septiembre de 1968

Normalización. Dependencias funcionales.

Son propiedades inherentes al contenido semántico de los datos, que se han de cumplir para *cualquier extensión* del esquema de relación.

Se trata de restricciones de integridad que permiten conocer qué interrelaciones existen entre los atributos del mundo real.

Existen otras restricciones de integridad que no se pueden expresar por medio de dependencias.

Son invariantes en el tiempo.

Normalización. Dependencias funcionales.

Dependencia funcional.

Se dice que un conjunto de atributos $Y(Y_1, Y_2, \dots, Y_n)$ depende funcionalmente de otro conjunto de atributos $X(X_1, X_2, \dots, X_n)$ si para cada valor de X hay un único valor posible para Y . Simbólicamente se denota por $X \rightarrow Y$.

Por ejemplo el nombre de una persona depende funcionalmente del DNI; es decir para un DNI concreto sólo hay un nombre posible.

Al conjunto X del que depende funcionalmente el conjunto Y se le llama **determinante**. Al conjunto Y se le llama **implicado**.

DNI \rightarrow Nombre

Normalización. Dependencias funcionales.

Dependencia funcional completa.

Un conjunto de atributos $Y(Y_1, Y_2, \dots, Y_n)$ tiene una dependencia funcional completa de otro conjunto de atributos $X(X_1, X_2, \dots, X_n)$ si Y tiene dependencia funcional de X y además no se puede obtener de X un subconjunto de atributos que consiga una dependencia funcional sobre Y (es decir, no hay en X un **determinante** formado por atributos más pequeños).

Por ejemplo en una tabla de clientes, el conjunto de atributos formado por el nombre y el DNI producen una dependencia funcional sobre el atributo apellidos. Pero no es completa ya que el DNI individualmente también produce una dependencia funcional sobre apellidos. El DNI sí produce una dependencia funcional completa sobre el campo apellidos.

Una dependencia funcional completa se denota como $X \Rightarrow Y$

{Nombre, DNI} → Apellidos

DNI ⇒ Apellidos

Normalización. Dependencias funcionales.

Dependencia funcional elemental.

Se produce cuando X e Y forman una dependencia funcional completa y además Y es un único atributo.

Normalización. Dependencias funcionales.

Dependencia funcional transitiva.

Se produce cuando tenemos tres conjuntos de atributos X, Y y Z. Y depende funcionalmente de X ($X \rightarrow Y$), Z depende funcionalmente de Y ($Y \rightarrow Z$). Además X no depende funcionalmente de Y ($Y \not\rightarrow X$).

Entonces ocurre que X produce una dependencia funcional transitiva sobre Z.

Esto se denota como: ($X \twoheadrightarrow Z$).

Normalización. Definición formal de 2FN.

Segunda forma normal.

1. Cuando una tabla está en 1FN y,
2. Cada atributo que no sea clave depende de forma **funcional completa** respecto de la clave. Toda la clave principal debe hacer dependientes al resto de atributos, si hay atributos que depende sólo de parte de la clave, entonces esa parte de la clave y esos atributos formarán otra tabla.

<u>Empleado</u>	<u>Habilidad</u>	Lugar actual de trabajo
Jones	Mecanografía	114 Main Street
Jones	Taquigrafía	114 Main Street
Jones	Tallado	114 Main Street
Bravo	Limpieza ligera	73 Industrial Way
Ellis	Alquimia	73 Industrial Way
Ellis	Malabarismo	73 Industrial Way
Harrison	Limpieza ligera	73 Industrial Way

**{Empleado, Habilidad} \Rightarrow Lugar actual de trabajo, porque
Empleado \Rightarrow Lugar actual de trabajo**

Normalización. Definición formal de 3FN.

Tercera forma normal.

1. Cuando una tabla está en 2FN y,
 2. Ningún atributo que no sea clave depende **transitivamente** de las claves de la tabla.
- Es decir, una tabla no está en tercera forma normal cuando algún atributo depende funcionalmente de atributos que no son clave.

Ganadores del torneo

<u>Torneo</u>	<u>Año</u>	Ganador	Fecha de nacimiento del ganador
Indiana Invitational	1998	Al Fredrickson	21 de julio de 1975
Cleveland Open	1999	Bob Albertson	28 de septiembre de 1968
Des Moines Masters	1999	Al Fredrickson	21 de julio de 1975
Indiana Invitational	1999	Chip Masterson	14 de marzo de 1977

X

Y

Z

$\{\text{Torneo, Año}\} \rightarrow \text{Ganador} \text{ (X} \rightarrow \text{Y)}$

$\text{Ganador} \rightarrow \text{Fecha de nacimiento del ganador} \text{ (Y} \rightarrow \text{Z)}$

$\text{Ganador} \not\rightarrow \{\text{Torneo, Año}\} \text{ (Y} \not\rightarrow \text{X)}$

$\{\text{Torneo, Año}\} \not\rightarrow \text{Fecha de nacimiento del ganador} \text{ (X} \not\rightarrow \text{Z)}$

Normalización. Forma normal de Boyce-Codd (FNBC).

Ocurre si una tabla está en 3FN y además todo determinante de la tabla sea una clave candidata.

ORGANIZACIÓN		
Trabajador	Departamento	Responsable
Alex	Producción	Felipa
Arturo	Producción	Martín
Carlos	Ventas	Julio
Carlos	Producción	Felipa
Gabriela	Producción	Higinio
Luisa	Ventas	Eva
Luisa	Producción	Martín
Manuela	Ventas	Julio
Pedro	Ventas	Eva

La cuestión es que un trabajador o trabajadora puede trabajar en varios departamentos. En dicho departamento hay varios responsables, pero cada trabajador sólo tiene asignado uno. El detalle importante que no se ha tenido en cuenta, es que el o la responsable sólo puede ser responsable en un departamento.

Este detalle último produce una dependencia funcional ya que:

Responsable → Departamento



Es determinante y no es clave candidata

Por lo tanto hemos encontrado un determinante que no es clave candidata. No está en FNBC. En este caso la redundancia ocurre por mala selección de clave. La redundancia del departamento es completamente evitable.

Normalización. Forma normal de Boyce-Codd (FNBC).

ORGANIZACIÓN		
<u>Trabajador</u>	<u>Departamento</u>	Responsable
Alex	Producción	Felipa
Arturo	Producción	Martín
Carlos	Ventas	Julio
Carlos	Producción	Felipa
Gabriela	Producción	Higinio
Luisa	Ventas	Eva
Luisa	Producción	Martín
Manuela	Ventas	Julio
Pedro	Ventas	Eva

La solución sería:

PERSONAL	
<u>Trabajador</u>	Responsable
Alex	Felipa
Arturo	Martín
Carlos	Julio
Carlos	Felipa
Gabriela	Higinio
Luisa	Eva
Luisa	Martín
Manuela	Julio
Pedro	Eva

RESPONSABLES	
<u>Responsables</u>	Departamento
Felipa	Producción
Martín	Producción
Julio	Ventas
Higinio	Producción
Eva	Ventas

Normalización. Vídeos Youtube.

Vídeos explicativos Youtube:

Dependencias funcionales: <http://youtu.be/pxAIPtrHKM4>

Normalización: <http://youtu.be/-5QMUXKPd40>

Normalización. Ejercicio resuelto.

- **Considere la siguiente relación:**
- **VENTA_COCHES (Vendedor#,Coche#, Fecha_Venta, Comisión, cant_descuento)**
- Suponga que un coche puede ser vendido por múltiples vendedores y **por lo tanto**, {Vendedor#, Coche#} es la clave primaria. Otras dependencias adicionales son:
- Fecha_Venta \rightarrow cant_descuento y
- Vendedor#, \rightarrow %Comisión.

Normalización. Ejercicio resuelto.

Venta_Coches

<u>Vendedor#</u>	<u>Coche#</u>	Fecha_Venta	Comisión	Cant_descuento
------------------	---------------	-------------	----------	----------------

¿Cumple la 1FN?

Todos los atributos son de cardinalidad 0 y 1. No existen grupos repetidos, por lo tanto cumple con la 1FN

¿Cumple con la 2FN?

De acuerdo con una de las dependencias funcionales definidas:

Vendedor#, \rightarrow %Comisión.

Existe dependencia parcial, en el atributo comisión (ya que depende de un subconjunto de la clave), por lo tanto no cumple con la 2FN

¿Qué se debe hacer?

Normalización. Ejercicio resuelto.

Crear una segunda tabla con la columna que no cumple la 2FN y la PK de la cual depende.

Venta_Coches

<u>Vendedor#</u>	<u>Coche#</u>	Fecha_Venta	Cant_descuento
------------------	---------------	-------------	----------------

Vendedor

<u>Vendedor#</u>	comisión
------------------	----------

Normalización. Ejercicio resuelto.

¿La relación Venta_Coches cumple la 3FN?

Venta_Coches

<u>Vendedor#</u>	<u>Coche#</u>	Fecha_Venta	Cant_descuento
------------------	---------------	-------------	----------------

Se tiene definida otra dependencia funcional:

Fecha_Venta \rightarrow cant_descuento

Recordar que la 3FN, se refiere a dependencias transitivas, es decir un atributo **no** clave depende de otro **no** clave

Por lo tanto esta relación no cumple con la 3FN

Debemos descomponer la relación, de tal manera que:

Venta_Coches

<u>Vendedor#</u>	<u>Coche#</u>	Fecha_Venta
------------------	---------------	-------------

Fecharventa_descuento

<u>Fecha_Venta</u>	Cant_desc
--------------------	-----------

Normalización. Ejercicio resuelto.

Las relaciones resultantes son:

Venta_Coches

<u>Vendedor#</u>	<u>Coche#</u>	Fecha_Venta
------------------	---------------	-------------

Vendedor

<u>Vendedor#</u>	comisión
------------------	----------

Fechaventa_descuento

<u>Fecha_Venta</u>	Cant_desc
--------------------	-----------