



Tema IV: Teoría de la Normalización

- 4.1** Noción intuitiva de las primeras formas normales
- 4.2** Dependencias funcionales
- 4.3** Definición formal de las tres primeras formas normales
- 4.4** Forma Normal de Boyce y Codd
- 4.5** Axiomas de Armstrong y manipulación de dependencias funcionales
- 4.6** Métodos de diseño: Análisis y Síntesis



Tema 4.1: Noción intuitiva de las primeras formas normales

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.

Tema 4.1: Noción intuitiva de las primeras formas normales

OBJETIVOS DE UN DISEÑO NORMALIZADO

- Eliminar anomalías de actualización.
- Conservar la información (descomposición sin pérdida de información).
- Conservar las dependencias funcionales (descomposición sin pérdida de DF).
- No crear dependencias nuevas o interrelaciones inexistentes.
- Facilidad de uso.
- Eficiencia.

¡ A VECES NO SON COMPATIBLES !

Tema 4.1: Noción intuitiva de las primeras formas normales

EJEMPLO DE DISEÑO INADECUADO

TABLA IMPARTE

Cod_Prof	Nombre_Prof	Despacho	Cod_asignatura	Nombre_asignatura	Nº créditos
9321	J. Sánchez	2B2	ITIG0231	Diseño de BD	7
9321	J. Sánchez	2B2	ITIG0221	Ficheros y BD	7
8142	P. Martín	2B2	ITIG0231	Diseño de BD	7
8142	P. Martín	2B2	ITIG0242	Diseño Avanzado de BD	4,5
8142	P. Martín	2B2	LD0241	SGBD	6
9577	A. García	2C4	II0232	Administración de BD	6
9111	L. López	2D5	II0232	Administración de BD	6
9111	L. López	2D5	ITIG0232	BD Avanzadas	4,5
9111	L. López	2D5	ITIG0221	Ficheros y BD	7

- **Redundancias**
- **Posibles inconsistencias**
- **Imposibilidad de almacenar ciertas informaciones**
- **Necesidad de valores nulos**

Tema 4.1: Noción intuitiva de las primeras formas normales

PRIMERA FORMA NORMAL (1FN)

No puede haber grupos repetitivos.

Ejemplo:

Se soluciona repitiendo toda la tupla para cada uno de los valores del grupo repetitivo

R (DNI_P, NOMBRE_P, TELÉFONO)

DNI_P	NOMBRE_P	TELÉFONO
414618	S. HIDALGO	2343840 6544070
9528309	M. PIATTINI	7193456
.	.	.

NO ESTÁ EN 1FN

DNI_P	NOMBRE_P	TELÉFONO
414618	S. HIDALGO	2343840
414618	S. HIDALGO	6544070
9528309	M. PIATTINI	7193456
.	.	.

ESTÁ EN 1FN

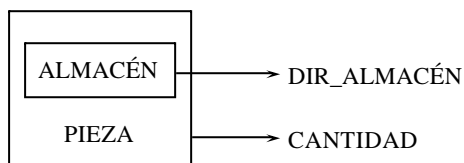
Tema 4.1: Noción intuitiva de las primeras formas normales

SEGUNDA FORMA NORMAL (2FN)

Todo atributo no clave debe ser un hecho acerca de la **totalidad** de la clave.

Ejemplo:

R (PIEZA, ALMACÉN, CANTIDAD, DIR_ALMACÉN)



Tema 4.1: Noción intuitiva de las primeras formas normales

PROBLEMAS BÁSICOS EN UN DISEÑO NO 2FN

- . Redundancia de datos
- . Necesidad de actualización en cadena
- . Posible inconsistencia de datos en las actualizaciones.
- . Imposibilidad de almacenar ciertos datos

Solución

Se descompone en:

R1 (ALMACÉN, PIEZA, CANTIDAD)
R2 (ALMACÉN, DIR_ALMACÉN)

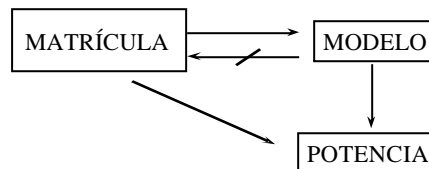
Tema 4.1: Noción intuitiva de las primeras formas normales

TERCERA FORMA NORMAL (3FN)

Todo atributo no clave debe ser un hecho **SOLO** acerca de la clave (y de la **totalidad** de la clave).

Ejemplo:

R (MATRÍCULA, MODELO, POTENCIA)



Tema 4.2: Dependencias Funcionales

DEPENDENCIAS

- 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.

Tema 4.2: Dependencias Funcionales

DEFINICIÓN DE DESCRIPTOR

Dada la relación

$$R(A_1, A_2, \dots, A_n)$$

Se dice que X es un descriptor de R, si

$$X \subseteq (A_1, A_2, \dots, A_n)$$

Tema 4.2: Dependencias Funcionales

DEPENDENCIA FUNCIONAL (DF)

Definición (1):

Sea la relación

$R (A_1, A_2, \dots, A_n)$

sean:

$X \hat{=} (A_1, A_2, \dots, A_n)$

$Y \hat{=} (A_1, A_2, \dots, A_n)$

dos descriptores de R.

Se dice que Y depende funcionalmente de X si a cada valor de X le corresponde un único valor de Y.

$X \longrightarrow Y$ (X implica o determina Y)

Tema 4.2: Dependencias Funcionales

DEPENDENCIA FUNCIONAL COMPLETA (DFC)

Sea la DF: $X \longrightarrow Y$

Si el descriptor X es compuesto:

$X (X_1, X_2)$

Y tiene dependencia funcional completa respecto de X si:

$$\begin{cases} X_1 \not\rightarrow Y \\ X_2 \not\rightarrow Y \end{cases}$$

$X \longrightarrow Y$

Tema 4.2: Dependencias Funcionales

DESCRIPTORES EQUIVALENTES

Sean X e Y, son dos descriptores de R, si se cumple que:

$$\left. \begin{array}{l} X \longrightarrow Y \\ Y \longrightarrow X \end{array} \right\}$$

entonces X e Y son **EQUIVALENTES**.

$$X \longleftrightarrow Y$$

Tema 4.2: Dependencias Funcionales

EJEMPLOS DE DEPENDENCIAS

DNI	_____	NOMBRE
MATRICULA	_____	MARCA
FABRICANTE, NUM_SERIE	_____	TIPO_ARMA
CARRETERA, KM	_____	TERMINO_MUNICIPAL
JUGADOR, TEMPORADA	_____	EQUIPO
DNI_PROF	_____	TELÉFONOS
DISCO	_____	CANTANTE
AÑO, PRUEBA	_____	ATLETA

* ¿ Hay alguna que no sea dependencia funcional?

* ¿ Son todas elementales?

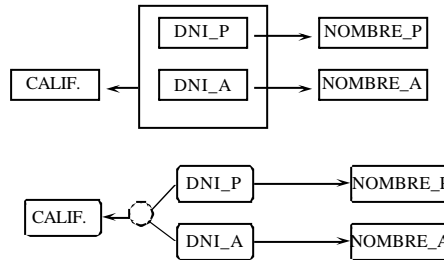
* ¿ Cuáles son completas?

Tema 4.2: Dependencias Funcionales

GRAFO DE DEPENDENCIAS FUNCIONALES

Las dependencias funcionales elementales entre un conjunto de atributos se pueden representar mediante un grafo.

EJEMPLO:



Tema 4.2: Dependencias Funcionales

DEPENDENCIA TRANSITIVA

Dada la relación

$R(X, Y, Z)$

en la que existen las siguientes dependencias funcionales:

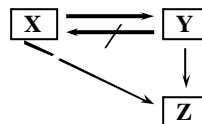
$X \longrightarrow Y$

$Y \longrightarrow Z$

$Y \not\longrightarrow X$

se dice que Z tiene una dependencia TRANSITIVA respecto a X a través de Y

$X \dashrightarrow Z$



X, Y, Z no tienen por qué ser disjuntos

Tema 4.2: Dependencias Funcionales

SUPERCLAVE Y CLAVE DE UNA RELACIÓN

Denominamos **Superclave** k de una relación $R(T)$ a un subconjunto no vacío de T , tal que T tiene dependencia funcional respecto de k ,

$$k \subseteq T \wedge k \neq \emptyset \wedge k \longrightarrow T$$

Si no existe ningún subconjunto k' de k , tal que:

$$k' \longrightarrow T$$

k es una **clave** de R .

$$k \subseteq T \wedge k \neq \emptyset \wedge k \longrightarrow T \wedge (\nexists k' \subset k : k' \longrightarrow T)$$

En la relación $R(T)$ se denominan atributos **principales** a aquellos que son elementos de alguna clave de $R(T)$.

A los demás atributos se les denomina **no principales**.

Tema 4.3: Definición formal de las tres primeras FN

PRIMERA FORMA NORMAL (1FN)

$R(T)$ está en 1FN si todos sus atributos, para cada tupla, toman un único valor del dominio simple subyacente, es decir, no hay grupos repetitivos.

SEGUNDA FORMA NORMAL (2FN)

$R(T)$ está en 2FN si:

- Está en 1FN
- Cada atributo no principal tiene dependencia funcional completa respecto de cada una de las claves.

TERCERA FORMA NORMAL (3FN)

$R(T)$ está en 3FN si:

- Está en 2FN
- Ningún atributo no principal depende transitivamente de las claves (no existen dependencias entre atributos no principales).

Tema 4.4: Forma Normal de Boyce-Codd

- Es más fuerte que la 3FN.
- Ocurre cuando existen claves candidatas que se solapan.
- Fué definida por Boyce y Codd en 1974.
- Las relaciones en 3FN pueden tener todavía anomalías de actualización.

Definición:

“Para que una relación se encuentre en FNBC, todo determinante HA DE SER una clave candidata”

(NO: formar parte)

Tema 4.4: Forma Normal de Boyce-Codd

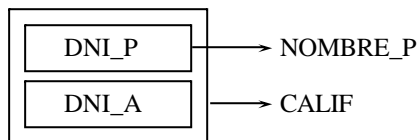
Ejemplo I:

NOTA (DNI_P, NOMBRE_P, DNI_A, CALIF)

¿Está en 3FN?
¿Está en FNBC?

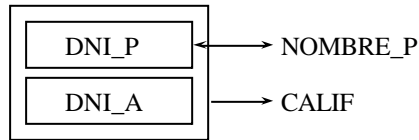
Supongamos las siguientes dependencias funcionales:

Supuesto A)



Tema 4.4: Forma Normal de Boyce-Codd

Supuesto B)



Está en 2FN porque NOMBRE_P es atributo principal, y en 3FN porque no hay dependencias transitivas.

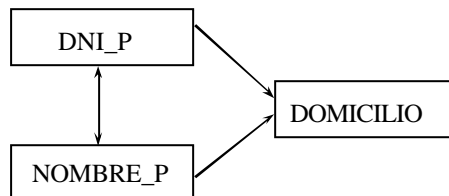
No está en FNBC (DNI_P y NOMBRE_P son determinantes y no son claves candidatas).

Tema 4.4: Forma Normal de Boyce-Codd

Ejemplo II:

PROFESOR (DNI_P, NOMBRE_P, DOMICILIO, ...)

Dependencias funcionales



¿Está en 3FN?

¿Está en FNBC?

Tema 4.5: Axiomas de Armstrong y manipulación de DF

A-1: Reflexividad:

Si $Y \subseteq X$, entonces $X \longrightarrow Y$

A-2: Transitividad:

Si $X \longrightarrow Y$ e $Y \longrightarrow Z$
entonces $X \longrightarrow Z$

A-3: Aumentatividad:

Si $X \longrightarrow Y$ y $Z \subseteq W$
entonces $XW \longrightarrow YZ$

Tema 4.5: Axiomas de Armstrong y manipulación de DF

Axiomas derivados:

A-4: Proyectividad:

Si $X \longrightarrow Y$ e $Y' \subseteq Y$
entonces $X \longrightarrow Y'$

A-5: Unión o aditividad:

Si $X \longrightarrow Y$ y $X \longrightarrow Z$
entonces $X \longrightarrow YZ$

A-6: Pseudotransitividad:

Si $X \longrightarrow Y$ e $YW \longrightarrow Z$
entonces $XW \longrightarrow Z$

Tema 4.5: Axiomas de Armstrong y manipulación de DF

Dado el esquema de relación:

$$R(A,B,C,D,E; A \longrightarrow B, C \longrightarrow D, D \longrightarrow E)$$

Demostrar que : $AC \longrightarrow ABCDE$

- 1) $A \longrightarrow B$ (dada)
- 2) $AC \longrightarrow ABC$ (aumentatividad de 1 por AC)
- 3) $C \longrightarrow D$ (dada)
- 4) $D \longrightarrow E$ (dada)
- 5) $C \longrightarrow E$ (transitividad de 3 y 4)
- 6) $C \longrightarrow DE$ (unión de 3 y 5)
- 7) $ABC \longrightarrow ABCDE$ (aumentatividad de 6 por ABC)
- 8) $AC \longrightarrow ABCDE$ (transitividad de 2 y 7)

Luego AC implica todos los atributos

Tema 4.5: Axiomas de Armstrong y manipulación de DF

CIERRE TRANSITIVO DE UN DESCRIPTOR

Dada la relación

$$R \langle A, DF \rangle$$

El cierre transitivo de X - denotado por X_{DF}^+ - respecto al conjunto de dependencias funcionales DF , es un descriptor Y de R , tal que cualquier dependencia:

$$X \longrightarrow Y \hat{I} DF^+$$

siendo Y máximo en el sentido de que la adición de cualquier atributo vulneraría la condición anterior.

Tema 4.5: Axiomas de Armstrong y manipulación de DF

ALGORITMO T1 (Cierre de un descriptor)

Entrada : Un conjunto de dependencias DF

Un descriptor X

Salida: X^+ , cierre de X respecto a DF

Proceso:

1) $X^+ = X$

2) Repetir hasta que no se añadan más atributos a X^+

- Por cada dependencia $Y \longrightarrow A$ en DF.

Si Y pertenece a X y A no está en X^+ entonces $X^+ = X^+ \cup A$.

Tema 4.5: Axiomas de Armstrong y manipulación de DF

Ejemplo:

Sean las dependencias:

A	→	B
C,A	→	D
D	→	F,E
F	→	G
H	→	K

El cierre transitivo de {A,D } respecto a éste conjunto de dependencias es:

$X^{(0)} = AD$

$X^{(1)} = ADB$

$X^{(2)} = ADBFE$

$X^{(3)} = ADBFEG = X^+$

Por tanto: $AD^+ = ADBFEG$

Tema 4.5: Axiomas de Armstrong y manipulación de DF

El cálculo del cierre nos va a permitir determinar si una dependencia

$X \twoheadrightarrow Y$

- ✓ está implicada por un conjunto de dependencias
- ✓ obtener las claves de un esquema
- ✓ saber si un descriptor es clave.
- ✓ la cobertura minimal de un conjunto de dependencias



es la base para dar solución a los problemas que se plantean en el proceso de normalización.

Tema 4.5: Axiomas de Armstrong y manipulación de DF

B) Recubrimiento minimal

Podemos definir un **recubrimiento minimal** o **irredundante** de un conjunto de dependencias funcionales asociadas a un conjunto de atributos A, como un subconjunto de las dependencias elementales del conjunto inicial DF, tal que cumpla:

- Ninguna de las dependencias funcionales elementales en DF es redundante, es decir, si se elimina cualquiera de las dependencias de DF, el nuevo conjunto de dependencias DF' no es equivalente a DF (no tiene el mismo cierre).
- Todas las dependencias funcionales entre los atributos A están en DF⁺

La definición de recubrimiento minimal se basa en los conceptos de **dependencia redundante** y de **atributo extraño**.

Tema 4.5: Axiomas de Armstrong y manipulación de DF

Dependencia redundante.

Una dependencia funcional d de DF se dice que es redundante si puede derivarse de $\{DF - d\}$ mediante la aplicación de los axiomas de Armstrong. Es decir, $X \rightarrow Y$ es redundante si $Y \in X^+_{\{DF-d\}}$

Atributo extraño

Dada la dependencia $X \rightarrow Y$ de DF, un atributo A perteneciente a X se dice que es un atributo extraño en la dependencia, si la dependencia $(X - A) \rightarrow Y$ se encuentra en DF^+ , es decir, $Y \in (X-A)^+_{\{DF\}}$

La dependencia $X \rightarrow Y$ no es plena; el atributo que "sobra" es el atributo ajeno.

Tema 4.5: Axiomas de Armstrong y manipulación de DF

Ejemplo:

LIBRO ({*cod_libro, isbn, editorial, país*},
{*cod_libro* \rightarrow *isbn*, *isbn* \rightarrow *cod_libro*, *isbn* \rightarrow (*editorial, país*),
cod_libro \rightarrow *editorial*, *editorial* \rightarrow *país* })

los conjuntos de dependencias

$DF_1 = \{ \textit{cod_libro} \rightarrow \textit{isbn}, \textit{isbn} \rightarrow \textit{cod_libro}, \textit{isbn} \rightarrow \textit{editorial},$
 $\textit{editorial} \rightarrow \textit{país} \}$

$DF_2 = \{ \textit{cod_libro} \rightarrow \textit{isbn}, \textit{isbn} \rightarrow \textit{cod_libro}, \textit{cod_libro} \rightarrow \textit{editorial},$
 $\textit{editorial} \rightarrow \textit{país} \}$

son recubrimientos minimales, mientras que el conjunto original no lo era

¿ POR QUÉ ?

Tema 4.5: Axiomas de Armstrong y manipulación de DF

No es irredundante el siguiente conjunto:

$$DF_3 = \{ \text{cod_libro} \rightarrow \text{isbn}, \text{isbn} \rightarrow \text{cod_libro}, \text{isbn} \rightarrow \text{país}, \\ \text{isbn} \rightarrow \text{editorial}, \text{editorial} \rightarrow \text{país} \}$$

ya que $\text{isbn} \rightarrow \text{país}$ es una dependencia redundante que puede ser deducida a partir de :

$$\text{isbn} \rightarrow \text{editorial} \text{ y } \text{editorial} \rightarrow \text{país}.$$

Tampoco es un recubrimiento minimal (del anterior conjunto de dependencias):

$$DF_4 = \{ \text{cod_libro} \rightarrow \text{isbn}, \text{isbn} \rightarrow \text{editorial}, \text{editorial} \rightarrow \text{país} \}$$

ya que la dependencia $\text{isbn} \rightarrow \text{cod_libro}$ falta, y no puede ser deducida de las demás.

Tema 4.5: Axiomas de Armstrong y manipulación de DF

C) Determinación de si un descriptor es clave de una relación

Otro de los problemas que se plantea en relación con la manipulación de dependencias es cómo determinar si un descriptor es o no clave de una relación.

Dado el esquema de la relación $R(A, DF)$, se denomina **superclave** SK de la relación R a un subconjunto no vacío de A , tal que $SK \rightarrow A$ es una consecuencia lógica de DF , siendo, por tanto, un elemento de su cierre, es decir:

$$SK \neq \Phi \wedge SK \rightarrow A \in DF^+$$

Para el mismo esquema anterior decimos que K es una **clave candidata** de R si, además de ser una superclave, no existe ningún subconjunto estricto K' de K tal que K' implique también a A (condición de **minimidad**); expresado formalmente:

$$K \neq \Phi \wedge K \rightarrow A \in DF^+ \wedge (\forall K' \subset K : K' \not\rightarrow A)$$

La clave es, por tanto, un caso especial de superclave.

Tema 4.5: Axiomas de Armstrong y manipulación de DF

Dado el esquema de relación:

$R(A,B,C,D,E; A \longrightarrow B, C \longrightarrow D, D \longrightarrow E)$

determinar si AC es una clave

$AC^+ = ABCDE$

$A^+ = AB$

$C^+ = DE$

Luego AC es la clave

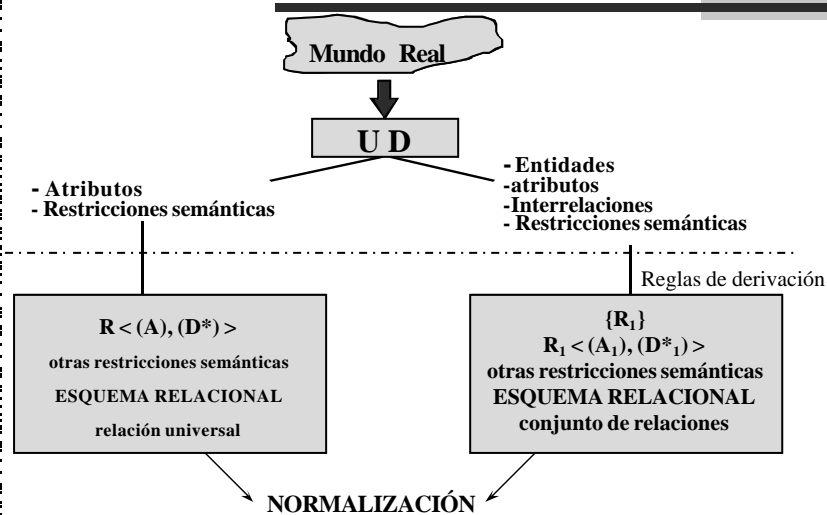
Tema 4.5: Axiomas de Armstrong y manipulación de DF

D) Determinación de las claves de una relación

Algoritmo:

1. Se calcula el cierre transitivo de todos los atributos de la relación $R(A,DF)$.
2. Separar los atributos independientes (formarán después parte de todas las claves).
3. Calcular los descriptores equivalentes, eligiendo un representante y eliminando las dependencias correspondientes.
4. Comprobar si existe algún descriptor X tal que $X^+ \rightarrow A$. Los atributos que aparecen sólo en la parte izquierda de las DF formarán parte de todas las claves.
 - Si es así, X es clave. Ir al paso 5.
 - Si no, ir al paso 5.
5. Calcular el cierre transitivos de las combinaciones binarias, ternarias, etc. de descriptores (observando cardinalidades de los cierres y aspecto de las DF) e ir al paso 3.
 - Si ya no hay más combinaciones posibles FIN

Tema 4.6: Métodos de Diseño



Tema 4.6: Métodos de Diseño

1. **ANÁLISIS**: Analiza una estructura relacional existente –puede ser la relación universal que contiene todos los atributos-, determinando la FN en la que se encuentra y descomponiéndola según un conjunto de dependencias funcionales en dos nuevas relacionales más regulares que cumplen unas determinadas propiedades y así sucesivamente (árbol de análisis).
2. **SÍNTESIS**: Utiliza el conjunto de dependencia funcionales directamente para obtener una descomposición con determinadas propiedades (generalmente una relación por cada dependencia funcional).

Tema 4.6: Métodos de Diseño

PROPIEDADES DESEABLES DE UNA DESCOMPOSICIÓN (1)

La descomposición (o diseño utilizando teoría de la normalización) es un proceso de refinamientos sucesivos que debiera conducir a aislar los tipos de entidades y tipos de interrelación del mundo real.

La descomposición se basa en dos operadores del álgebra relacional.

PROYECCIÓN/ COMBINACIÓN NATURAL

Tema 4.6: Métodos de Diseño

PROPIEDADES DESEABLES DE UNA DESCOMPOSICIÓN (4)

Descomposición en Proyecciones Independientes (Condición de Rissanen)

Para que se produzca la descomposición de una relación sin pérdida de información y sin pérdida de dependencias funcionales se debe cumplir que:

Sea R una relación y R_1 y R_2 dos de sus proyecciones, se dice que dichas proyecciones son independientes si, y sólo si,

- a) sus atributos comunes son la clave primaria de, al menos, una relación.
- b) cada dependencia funcional en R puede deducirse de las de R_1 y R_2 .

Tema 4.6: Métodos de Diseño

PROPIEDADES DESEABLES DE UNA DESCOMPOSICIÓN (5)

- Sin pérdida de información.
- Sin pérdida de dependencias funcionales.

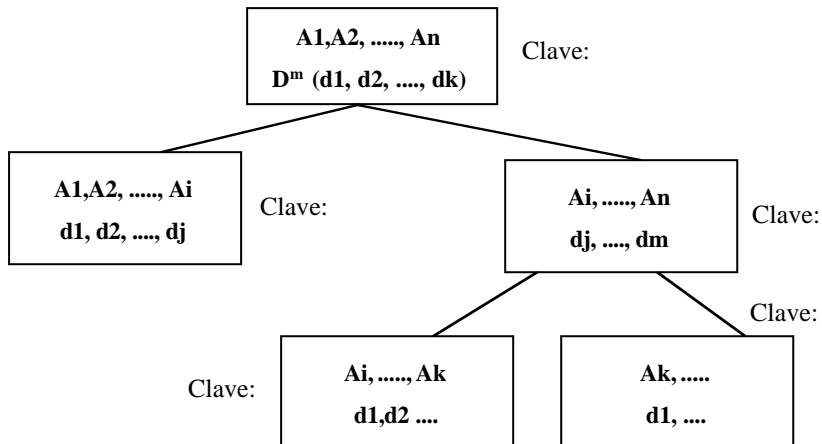
Ejemplo:

R(DNI_P, CIUDAD, PROVINCIA)

R11(<u>DNI_P</u> , CIUDAD)	}	Bien
R12(<u>CIUDAD</u> , PROVINCIA)		
R21(<u>DNI_P</u> , CIUDAD)	}	Pérdida de dependencia funcional CIUDAD → PROVINCIA
R22(<u>DNI_P</u> , PROVINCIA)		
R31(<u>DNI_P</u> , PROVINCIA)	}	Pérdida de información y de DNI_P CIUDAD
R32(<u>CIUDAD</u> , PROVINCIA)		

Tema 4.6: Métodos de Diseño

Método de Análisis (1)



Tema 4.6: Métodos de Diseño

Método de Análisis (2)

Supongamos la relación con esquema: R(A, DF) y que sufre anomalías por lo que se quiere descomponer. Los pasos a seguir son:

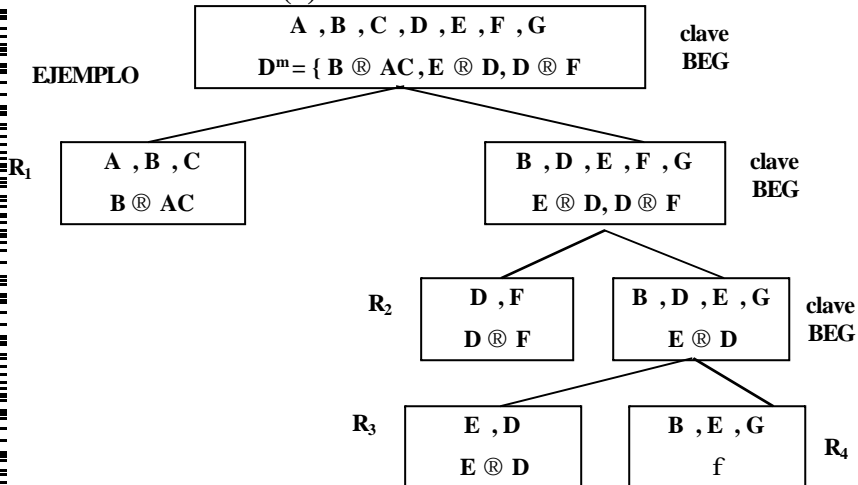
1. Hallar un recubrimiento minimal de DF
2. Determinar la(s) clave(s) así como los atributos principales y no principales
3. Identificar la FN en que se encuentra la relación.

Si se desea llegar a una FN más avanzada:

4. Agrupar las DF que tengan el mismo implicante.
5. Obtener proyecciones independientes sobre cada una de las dependencias funcionales (o de los grupos), de forma que los atributos que aparecen en la correspondiente dependencia constituyen una nueva relación y el implicado de la dependencia, así como esta, desaparezcan de la relación origen
6. Proseguir la descomposición repitiendo el paso 5 hasta que todas las dependencias estén implicadas por una clave

Tema 4.6: Métodos de Diseño

Método de Análisis (3)





Bibliografía

- ♦ Dolores Cuadra, Elena Castro, Ana Iglesias, Paloma Martínez, Javier Calle, César de Pablo, Harith Al-Jumaily y Lourdes Moreno (2007): Desarrollo de Bases de Datos: casos prácticos desde el análisis a la implementación. RA-MA
- ♦ Silberschatz, A.; Korth, H.; Sudarshan, S., 2006: Fundamentos de bases de datos. (5ª edición). McGraw-Hill.
- ♦ Levene, M. y Loizou, G., 1999: A Guided Tour of Relational Databases and Beyond. Springer Verlag.