

# Normalización

Lucio Salgado  
Diciembre 2008

# Introducción

- ◆ La normalización permite afrontar el problema del diseño de base de datos de una manera rigurosa y objetiva.
- ◆ El concepto de dependencia funcional es la base para definir las primeras formas normales.

# Sobre el diseño relacional

- ◆ El diseño relacional, trae consigo ciertos inconvenientes:
  - ◆ Redundancia
  - ◆ Ambigüedades
  - ◆ Pérdida de información
  - ◆ Pérdida de dependencias funcionales
  - ◆ Aparición de estados no válidos en el mundo real
  - ◆ Incapacidad de almacenar ciertos hechos

# Dependencia Funcional-Definición

- ◆ Sea la relación R definido con el conjunto de atributos A y sean X e Y subconjuntos de A. Se dice que Y depende funcionalmente de X, ó X determina o implica a Y, denotado por:

$$\mathbf{X \rightarrow Y}$$

donde :      X      determinante o implicante  
                 Y      implicado

**si, y sólo si, cada valor de X tiene asociado en todo momento un único valor de Y.**

# Dependencia Funcional-1

Ej. En la relación

ESTUDIANTE=(CI,NOMBRE,DIRECCION,S  
EM\_ING)

$CI \rightarrow NOMBRE$

$CI \rightarrow DIRECCION$

# Dependencia Funcional-2

◆ Ahora un esquema de relación se considera:

$$R=(A,DEP)$$

Donde :      A      Conjunto de atributos  
              DEP    Conjunto de dependencias

Ej.

ESTUDIANTE=({CI,NOMBRE,DIRECCION,  
SEM\_ING},{**CI→NOMBRE,**  
**CI→DIRECCION,CI→SEM\_ING,**  
**NOMBRE→DIRECCION**})

# Diagrama de Dependencia Funcional



Figura 1. Dependencia funcional

# Dependencia Funcional Completa-2

- ◆ Si el descriptor  $X$  es compuesto :  
 $X(X_1, X_2)$  se dice que  $Y$  tiene dependencia funcional completa, si depende funcionalmente de  $X$ , pero de ningún subconjunto del mismo.

$$X \rightarrow Y$$

$$X_1 \not\rightarrow Y$$

$$X_2 \not\rightarrow Y$$



# Dependencia Funcional Completa-2

Ejemplo. En la relación

REGISTRA=(CI,SIGLA,NOTA,SEM)

$CI, SIGLA, SEM \rightarrow NOTA$

$CI \twoheadrightarrow NOTA$

$SIGLA \twoheadrightarrow NOTA$

$SEM \twoheadrightarrow NOTA$

# Dependencia Funcional Trivial

- ◆ Una dependencia funcional  $X \rightarrow Y$  es trivial si  $Y$  es un subconjunto de  $X$  ( $Y \subseteq X$ )

Ejemplo:

CI, NOMBRE  $\rightarrow$  NOMBRE

# Descriptores equivalentes

◆ Se dice que  $X$  e  $Y$  son equivalentes si se cumple :  $X \rightarrow Y$

$$Y \rightarrow X$$

representado por :  $X \leftrightarrow Y$

Ej. En la relación

EMPLEADO=(CI,NSS,NOMBRE,DIREC)

$$CI \rightarrow NSS$$

$$NSS \rightarrow CI$$

$$NSS \leftrightarrow CI$$

# Dependencia Funcional Transitiva

- ◆ Sea la relación  $R(X,Y,Z)$  con las siguientes dependencias :

$$X \rightarrow Y$$

$$Y \rightarrow Z$$

$$Y \not\rightarrow X$$

- ◆ Se dice que  $Z$  tiene una dependencia transitiva respecto a  $X$  a través de  $Y$ .

# Axiomas de Armstrong

- ◆ A1 : Si  $Y \subseteq X$  entonces  $X \rightarrow Y$  (reflexividad)
- ◆ A2 : Si  $Z \subseteq W$  y  $X \rightarrow Y$ , entonces,  $XW \rightarrow YZ$  (aumentatividad)
- ◆ A3 : Si  $X \rightarrow Y$  e  $Y \rightarrow Z$ , entonces  $X \rightarrow Z$  (transitividad)
- ◆ A4 : Si  $X \rightarrow YZ$ , entonces,  $X \rightarrow Y$  y  $X \rightarrow Z$  (proyectividad).
- ◆ A5 : Si  $X \rightarrow Z$  y  $X \rightarrow Y$ , entonces  $X \rightarrow YZ$  (unión o aditividad)
- ◆ A6 : Si  $X \rightarrow Y$  e  $YW \rightarrow Z$ , entonces  $XW \rightarrow Z$  (pseudotransitividad).

# Clave candidata

- ◆ Sea, **R** una relación con atributos **A1,A2,...,An**, y **X** un subconjunto de, **A1,A2,...,An**.
- ◆ Se dice que **X**, es una clave de R, si cumple:
  1. **X** → **A1,A2,...,An**, es decir todos los atributos dependen funcionalmente de **X**.
  2. No existe un **Y**  $\subset$  **X**, tal que **Y** → **A1,A2,...,An**, que representa la condición de minimalidad (necesario para que sea considerada clave candidata).

# Identificar claves candidatas

1. Partir de las dependencias funcionales, se asume que no hay más llaves que dependencias funcionales.
2. Para cada dependencia funcional, aplicando los axiomas de Armstrong se puede lograr determinar todos los atributos de la relación
3. Luego se analiza la condición de minimalidad, descrita anteriormente.

# Ejemplo

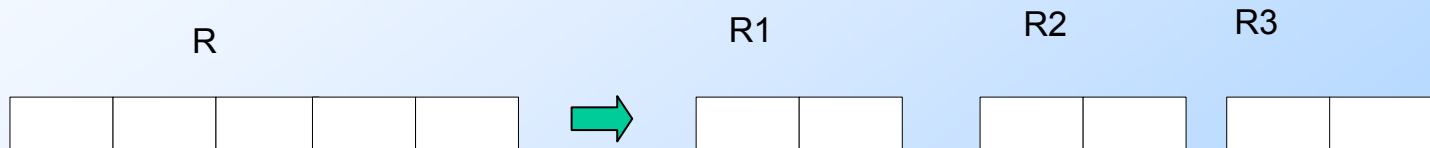
Ejemplo: Sea  $R(A,B,C,D,E)$  con  $A \rightarrow B$ ,  $BC \rightarrow D$  y  $AB \rightarrow E$

<p>1. <math>A \rightarrow B</math>  <math>A \rightarrow A</math> (<u>reflexividad</u>)  <math>A \rightarrow AB</math> (<u>aditividad</u>)  <math>A \rightarrow ABE</math> (transitividad con 3).  <math>AC \rightarrow ABEC</math> (<u>aumentatividad</u>)  <math>AC \rightarrow ABECD</math> (<u>transitiividad</u>)</p>	<p>2. <math>BC \rightarrow D</math>  <math>BC \rightarrow BC</math> (<u>reflexividad</u>)  <math>BC \rightarrow BCD</math> (<u>aditividad</u>)  <math>BCA \rightarrow BCDA</math> (<u>aumentativita</u>)  <math>BCA \rightarrow BCDAE</math> (<u>transitividad</u>)</p>	<p>3. <math>AB \rightarrow E</math>  <math>AB \rightarrow AB</math> (<u>reflexividad</u>)  <math>AB \rightarrow EAB</math> (<u>aditividad</u>)  <math>ABC \rightarrow EABC</math> (<u>aumentatividad</u>)  <math>ABC \rightarrow EABCD</math> (<u>transitividad</u>)</p>
AC es clave	ABC es clave	ABC es clave



# Normalización

## ◆ Proceso de descomposición



Se cumple:

- ◆ Conservación de la información
- ◆ Conservación de dependencias
- ◆ Mínima redundancia

# Normalización-1

- ◆ Si se cumplen esas propiedades:  
 $\{R_i\}$  es equivalente a  $R$
- ◆ Si  $R_i$  esta en formas normales más avanzadas que el esquema origen, entonces  
 $\{R_i\}$  es mejor a  $R$

# Primera Forma Normal(1FN)

◆ No debe admitir grupos repetitivos

DOCENTE			
CI	NOM	<u>SIGLA</u> M	SEM
56798	JUAN	SIS2302 SIS2301	II/97
45671	ANA	SIS1100 SIS1101	II/97

DOCENTE			
CI	NOM	<u>SIGLA</u> M	SEM
56798	JUAN	SIS2302	II/97
56798	JUAN	SIS2301	II/97
45671	ANA	SIS1100	II/97
45671	ANA	SIS1101	II/97

# Segunda Forma Normal(2FN)

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

Ej. Se tiene la siguiente relación  
REGISTRA=(CI,SIGLA,NOTA,SEM,  
NOM\_MAT)

Se tiene las siguientes dependencias :

CI,SIGLA,SEM $\rightarrow$ NOTA

SIGLA $\rightarrow$ NOM\_MAT

# Segunda Forma Normal(2FN)-1

- ◆ Para que este en 2FN, se descompone quitando las que no tienen DF completa de R.

Ej.

- ◆ REGISTRA1=(CI,SIGLA,NOTA,SEM)
- ◆ REGISTRA2=(SIGLA,NOM\_MAT)

# Tercera Forma Normal(3FN)

- ◆ Esta en 2FN
- ◆ Ningún atributo no principal depende transitivamente de ninguna clave de la relación.

Ej.

ESTUDIANTE=(CI,NOMBRE,DIR, COD\_CAR,NOM\_CAR)

Con:

$CI \rightarrow COD\_CAR$

$COD\_CAR \rightarrow NOM\_CAR$

$COD\_CAR \dashrightarrow CI$

$CI \rightarrow NOM\_CAR$

# Tercera Forma Normal(3FN)-1

◆ Para que este en 3FN se debe quitar los atributos que tienen DF transitiva

◆ Ej.

ESTUDIANTE1=(CI,NOMBRE,DIR,COD\_CAR)

ESTUDIANTE2=(COD\_CAR,NOM\_CAR)

# Forma Normal Boyce Codd

- Se dice que una relación se encuentra en FNBC si, y sólo si, todo determinante es una clave candidata.
  - ♦ Ej. Sea la relación  
DOCENTE = ( $\{CI\_D, NOM\_D, CI\_ES, NOTA\}$  ;  
 $\{CI\_D \rightarrow NOM\_D, (CI\_D, CI\_ES) \rightarrow NOTA\}$ )



# Dependencia Multivaluada

$$X \twoheadrightarrow Y$$

Se lee X multidetermina a Y,

- ◆ Donde X e Y son descriptores tales que cierto valor de X implica un conjunto bien definido de valores Y con independencia de los demás atributos de la relación

# Dependencia Multivaluada

## DM-DF

- ◆ Una Dependencia Funcional es un caso particular de una Dependencia Multivaluada.
- ◆ En la definición  $Y$  es un conjunto de valores, pero ese conjunto puede tener un solo elemento

# Cuarta Forma Normal(4FN)-1

- Se dice que se encuentra en 4FN si, y sólo si, las únicas dependencias multivaluadas no triviales son aquellas en las cuales una clave multidetermina un atributo es decir toda dependencia multivaluada viene determinada por la clave candidata.

# Cuarta Forma Normal(4FN)-2

◆ Ej.

ESTUDIA(CI,ESPECIALIDAD,ACTI\_DEP)

◆ Con:

- ◆  $CI \twoheadrightarrow ESPECIALIDAD$

- ◆  $CI \twoheadrightarrow ACTI\_DEP$

◆ No está en 4FN, pero ahora si:

- ◆ ESTUDIA1(CI,ESPECIALIDAD)

- ◆ ESTUDIA2(CI, ACTI\_DEP)

# Dependencia de Combinación

- ◆ Una dependencia de combinación es cuando una relación se puede descomponer en otras relaciones (proyecciones) sin pérdida de información, de tal manera que la relación original se obtiene realizando la operación de combinación de las relaciones descompuestas.

$$R = R_1 \mid \mathbf{X} \mid R_2 \mid \mathbf{X} \mid \dots \mid \mathbf{X} \mid R_j$$

# Quinta Forma Normal(5FN)

- ◆ Se dice que una relación está en 5FN si, y sólo si está en todas las formas anteriores formas normales y toda dependencia de combinación está implicada por una clave candidata

# Forma Normal Dominio-Clave(FNDC)

- ◆ Una relación  $R$  se encuentra en FNDC si, y sólo si, toda restricción de  $R$  es una consecuencia lógica de las restricciones de dominio y las restricciones de clave que se cumplen en  $R$ .
- ◆ Toda relación que se encuentra en FNDC se encuentra necesariamente en 5FN.

# Algoritmo de Normalización

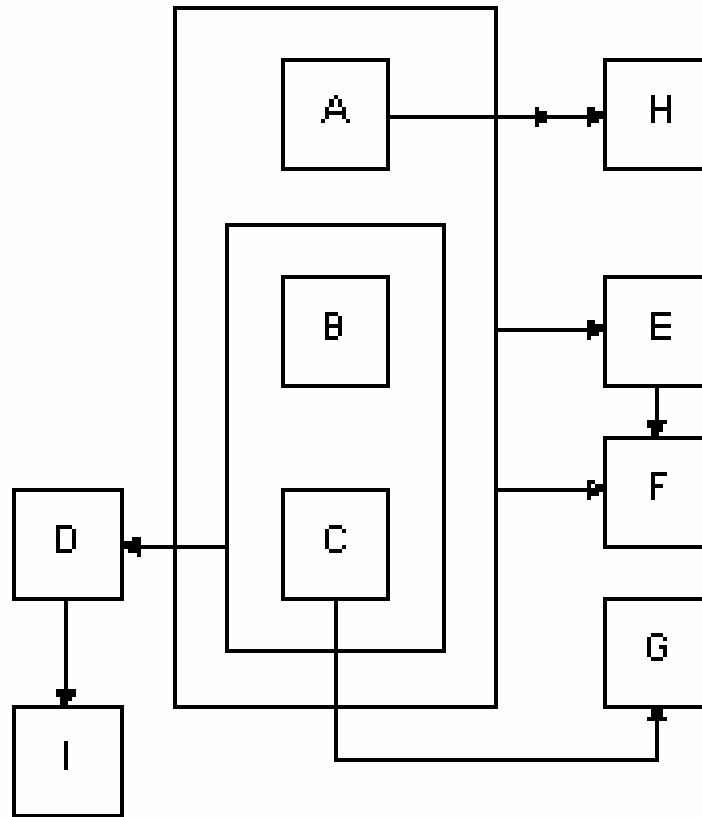
- ◆ Paso 1: Realizar el diagrama de dependencias funcionales
- ◆ Paso 2: Descomponer las dependencias multivaluadas
- ◆ Paso 3: Descomponer las dependencias funcionales transitivas
- ◆ Paso 4: Descomponer las dependencias funcionales completas



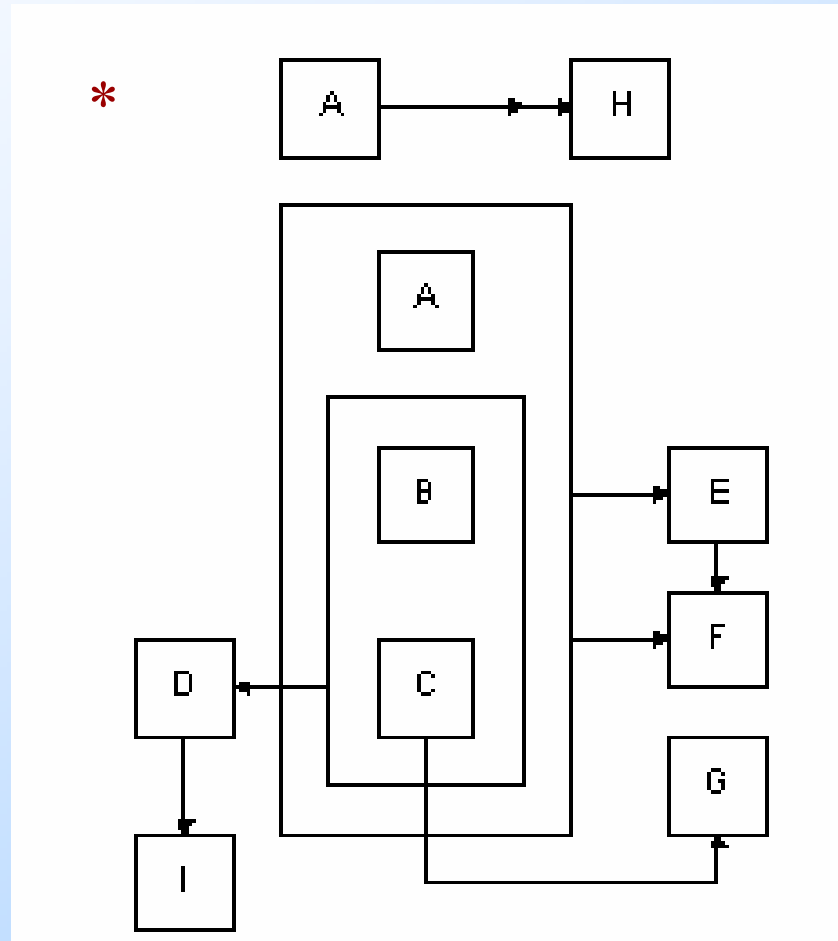
# Ejemplo

- ◆ Dada la relación
  - ◆  $R(A,B,C,D,E,F,G,H)$
- ◆ Con:

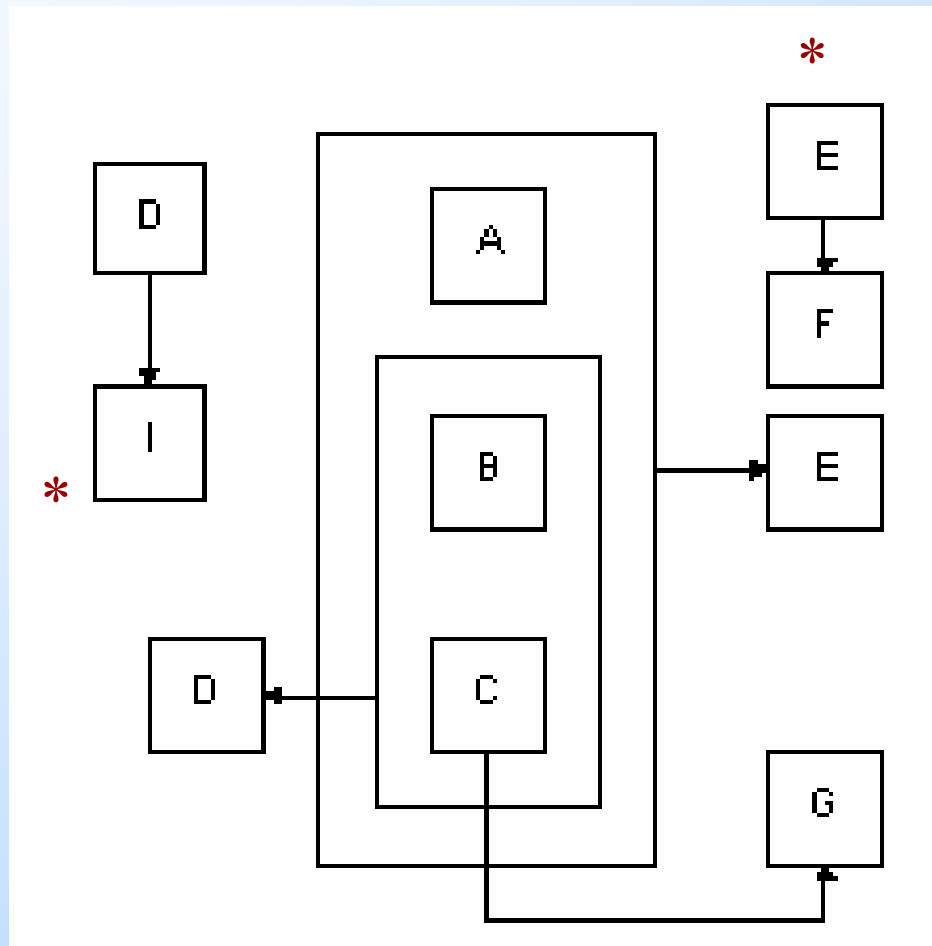
# Paso 1



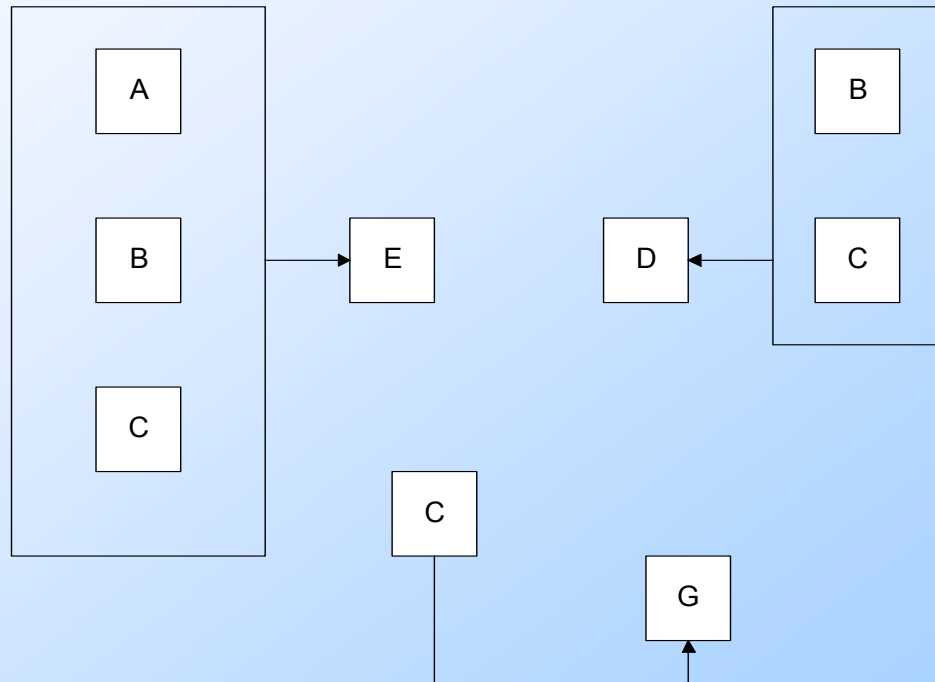
# Paso 2



# Paso 3



# Paso 4



PASO 4

# Ejercicio

- ◆ Dada la siguiente relación normalizar:  
ESTUDIANTE=(ci,nombre,direccion,sem\_ing,sigla,sem,nota,nom\_mat,creditos\_mat)
- ◆ Las dependencias funcionales:  
ci→nombre                      ci→direccion  
ci→sem\_ing                      ci.sigla,sem→nota  
sigla→nom\_mat                  sigla→creditos\_mat

¿Preguntas ?