

# Unidad 4: Buenas prácticas

BBDD01, Sesión 12:

Descomposición y Normalización de BBDD

# INDICE

- Introducción
- Descomposición.
- Propiedades deseables de la descomposición.
- Primera Forma Normal
- Segunda Forma Normal
- Tercera Forma Normal
- Forma normal Boyce-Codd.
- Cuarta forma normal.
- Otras formas normales.
- Proceso general del diseño de BD

Referencias: Silberschatz 4ª Ed. pp 161-189 Elmasri 3ª Ed. Pp 439-496

# Introducción

- Las dependencias funcionales son la clave para poder realizar la normalización de la información de la BD.
- Mediante el uso de las operaciones vistas en la sesión anterior, se definirá cómo se puede segmentar una tabla.
- El uso cuidadoso de dependencias funcionales y reglas de descomposición permite la normalización

## Recordando: Dependencias Funcionales

- Ejemplo: Dado el conjunto de DF para el esquema (A,B,C)

$A \rightarrow BC$ ,  $B \rightarrow C$ ,  $A \rightarrow B$  y  $AB \rightarrow C$

- Calcular el recubrimiento canónico de F.
- Hay dos DF con el mismo conjunto de atributos a la izquierda:

$A \rightarrow BC$

$A \rightarrow B$

- Se transforman en  $A \rightarrow BC$
- A es raro en  $AB \rightarrow C$  porque F implica lógicamente a  $(F - \{AB \rightarrow C\}) \cup \{(B \rightarrow C)\}$
- C es raro en  $A \rightarrow BC$ , ya que  $A \rightarrow BC$  está implicada lógicamente en  $A \rightarrow B$  y  $B \rightarrow C$
- El recubrimiento canónico es:

$A \rightarrow B$

$B \rightarrow C$

## Descomposición

- El mal diseño anterior sugiere descomponer el esquema en varios esquemas con menos atributos cada uno.
- Ejemplo: Esquema-empréstito en

<i>nombre-sucursal</i>	<i>ciudad-sucursal</i>	<i>activo</i>	<i>nombre-cliente</i>
Centro	Arganzuela	9.000.000	Santos
Moralzarzal	La Granja	2.100.000	Gómez
Navacerrada	Aluche	1.700.000	López
Centro	Arganzuela	9.000.000	Sotoca
Becerril	Aluche	400.000	Santos
Collado Mediano	Aluche	8.000.000	Abril
Navas de la Asunción	Alcalá de Henares	300.000	Valdivieso
Segovia	Cerceda	3.700.000	López
Centro	Arganzuela	9.000.000	González
Navacerrada	Aluche	1.700.000	Rodríguez
Galapagar	Arganzuela	7.100.000	Amo

*Esquema-sucursal-cliente = (nombre-sucursal, ciudad-sucursal, activo, nombre-cliente)*

*Esquema-cliente-préstamo = (nombre-cliente, número-préstamo, importe)*

<i>nombre-cliente</i>	<i>número-préstamo</i>	<i>importe</i>
Santos	P-17	1.000
Gómez	P-23	2.000
López	P-15	1.500
Sotoca	P-14	1.500
Santos	P-93	500
Abril	P-11	900
Valdivieso	P-29	1.200
López	P-16	1.300
González	P-18	2.000
Rodríguez	P-25	2.500
Amo	P-10	2.200

## Descomposición

- Hay casos en los que hay que reconstruir la relación préstamo. Ejemplo: hallar todas las sucursales que tienen préstamos con importes inferiores a 1000€. Se podría realizar con:

*sucursal-cliente ⋈ cliente-préstamo*

<i>nombre-sucursal</i>	<i>ciudad-sucursal</i>	<i>activo</i>	<i>nombre-cliente</i>	<i>número-préstamo</i>	<i>importe</i>
Centro	Arganzuela	9.000.000	Santos	P-17	1.000
Centro	Arganzuela	9.000.000	Santos	P-93	500
Moralzarzal	La Granja	2.100.000	Gómez	P-23	2.000
Navacerrada	Aluche	1.700.000	López	P-15	1.500
Navacerrada	Aluche	1.700.000	López	P-16	1.300
Centro	Arganzuela	9.000.000	Sotoca	P-14	1.500
Becerril	Aluche	400.000	Santos	P-17	1.000
Becerril	Aluche	400.000	Santos	P-93	500
Collado Mediano	Aluche	8.000.000	Abril	P-11	900
Navas de la Asunción	Alcalá de Henares	300.000	Valdivieso	P-29	1.200
Segovia	Cerceda	3.700.000	López	P-15	1.500
Segovia	Cerceda	3.700.000	López	P-16	1.300
Centro	Arganzuela	9.000.000	González	P-18	2.000
Navacerrada	Aluche	1.700.000	Rodríguez	P-25	2.500
Galapagar	Arganzuela	7.100.000	Amo	P-10	2.200

## Descomposición

- Hay tuplas que no estaban en la original.

(Centro, Arganzuela, 9.000.000, Santos, P-93, 500)  
(Navacerrada, Aluche, 1.700.000, López, P-16, 1.300)  
(Becerril, Aluche, 400.000, Santos, P-17, 1.000)  
(Segovia, Cerceda, 3.700.000, López, P-15, 1.500)

- La consulta hallar todas las sucursales que han concedido un préstamo por importe inferior a 1500€ produce:
  - 1º Tabla:{Becerril,Collado,Mediano}
  - 2º Tabla:{Becerril,Collado, Mediano,Centro}
- Debido a que no se sabe qué préstamo pertenece a cada sucursal.
- Es una descomposición con pérdida (Mal diseño de la BD)
- Necesitamos descomposiciones sin pérdida de información.

# Propiedades deseables de la Descomposición

- Ejemplo sobre el esquema:

*Esquema-empréstito = (nombre-sucursal,  
ciudad-sucursal, activo, nombre-cliente,  
número-préstamo, importe)*

- Que se exige que cumpla las DF:

*nombre-sucursal → ciudad-sucursal activo  
número-préstamo → importe nombre-sucursal*

- Descomposición de reunión sin pérdida
- Sea R y sea F conjunto de DF de R. R se puede descomponer sin pérdida en  $R_1$  y  $R_2$  si al menos una de las siguientes dependencias se halla en  $F^+$

- $R_1 \cap R_2 \rightarrow R_1$
- $R_1 \cap R_2 \rightarrow R_2$



## Propiedades deseables de la Descomposición

- Ejemplo: se descompone Esquema-empréstito en dos esquemas

*Esquema-sucursal = (nombre-sucursal,  
ciudad-sucursal, activo)*  
*Esquema-info-préstamo = (nombre-sucursal,  
nombre-cliente, número-préstamo, importe)*

- Dado que nombre-sucursal  $\rightarrow$  ciudad-sucursal activo
- Utilizando aumentatividad:
- Nombre-sucursal  $\rightarrow$  nombre-sucursal ciudad-sucursal activo
- Como  $\text{Esquema-sucursal} \cap \text{Esquema-info-prestamo} = \{\text{nombre-sucursal}\}$ , la reunión es sin pérdida.

# Propiedades deseables de la Descomposición

- Se descompone Esquema-info-préstamo en

*Esquema-préstamo = (número-préstamo,  
nombre-sucursal, importe)*  
*Esquema-prestatario = (nombre-cliente,  
número-préstamo)*

- Es una descomposición de reunión sin pérdida ya que numero-préstamo es un atributo común y numero-préstamo → importe nombre-sucursal.

# Propiedades deseables de la Descomposición

## Conservación de las dependencias

- Hay que conservar las dependencias.
- Cuando se haga una actualización sobre la BD hay que cumplir todas las DF dadas.
- Comprobar de manera eficiente las actualizaciones sin calcular las reuniones.
- Sea  $F$  un conjunto de DF del esquema  $R$  y  $R_1, R_2, \dots, R_n$  una descomposición de  $R$ .
- La restricción de  $F$  a  $R_i$  es el conjunto de todas las DF de  $F^+$  que sólo incluyen atributos de  $R_i$ .
- Puesto que todas las DF de una restricción únicamente implican atributos de un esquema de relación, es posible comprobar el cumplimiento de una dependencia verificando sólo una relación.

## Propiedades deseables de la Descomposición

- El conjunto de restricciones  $F_1, F_2, \dots, F_n$  es el conjunto de dependencias que pueden comprobarse de manera eficiente.
- Sea  $F' = F_1 \cup F_2 \cup \dots \cup F_n$  un conjunto de DF del esquema de R.
- Las descomposiciones que tienen la propiedad de  $F'^+ = F^+$  son descomposiciones que conservan las dependencias.

```
calcular  $F^+$ ;  
for each esquema  $R_i$  de  $E$  do  
  begin  
     $F_i :=$  la restricción de  $F^+$  a  $R_i$ ;  
  end  
 $F' := \emptyset$   
for each restricción  $F_i$  do  
  begin  
     $F' = F' \cup F_i$   
  end  
calcular  $F'^+$ ;  
if ( $F'^+ = F^+$ ) then return (true)  
else return (false);
```

## Propiedades deseables de la Descomposición

- Se puede demostrar que la descomposición de Esquema-empréstito conserva las dependencias.
  - Esquema-sucursal(nombre-sucursal, ciudad-sucursal, activo) conserva la DF nombre-sucursal → ciudad-sucursal activo
  - Esquema-préstamo(nombre-sucursal, número-prestamo, importe) conserva la DF numero-prestamo → importe nombre-sucursal
- Si puede comprobarse cada miembro de F en una relación de la descomposición  $\Rightarrow$  conserva las dependencias

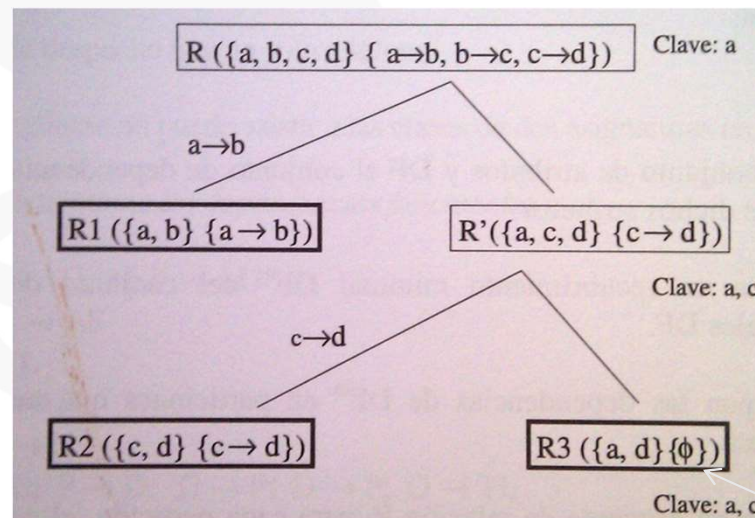
# Propiedades deseables de la Descomposición

## Repetición de la información

- No es deseable la repetición de la información en los diseños de la BD.
- La descomposición de Esquema-empréstito no lo sufre.
- El grado hasta el que se puede conseguir la falta de redundancia viene representado por varias formas normales.
- Teoría de Normalización.

## Ejemplo de descomposición

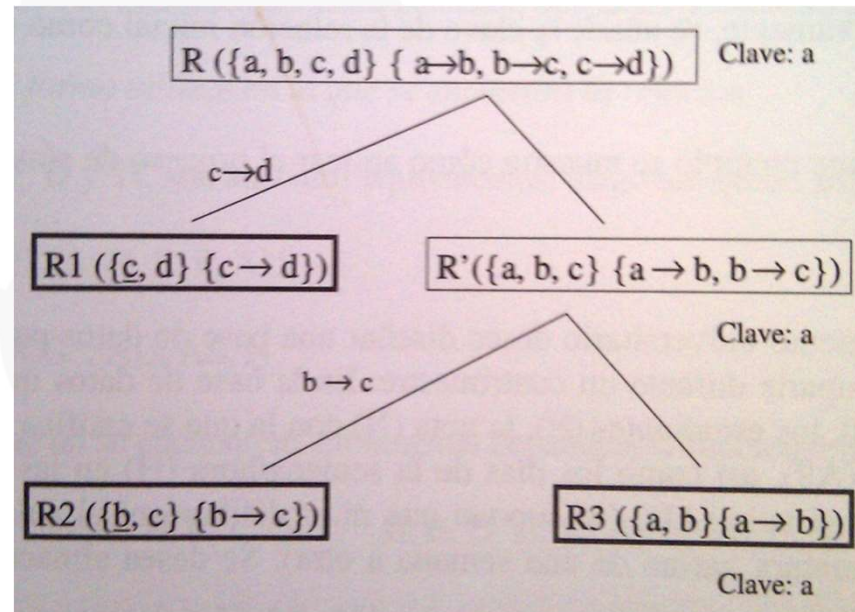
Ejemplo ERRONEO:



Se pierde una DF

## Ejemplo de descomposición

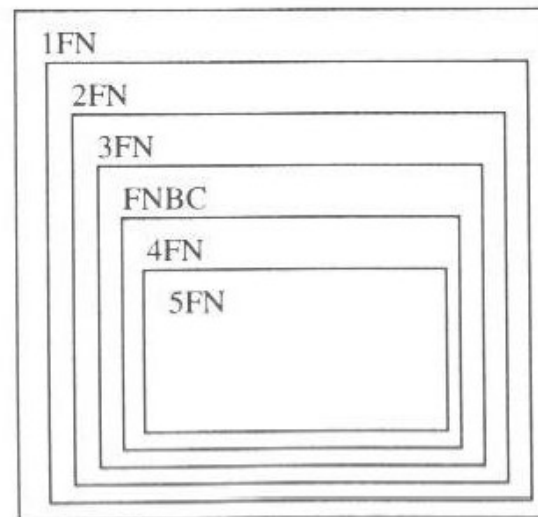
Ejemplo CORRECTO:





# Teoría de Normalización

- Dependencias funcionales dadas
- Cada relación una clave primaria
- Realizan condiciones para satisfacer formas normales  $\Rightarrow$  Proceso de normalización.
- Formas normales:



# Teoría de Normalización

- Serie de pruebas para ver si satisface una forma normal determinada (Codd 1972)
- Codd propuso tres formas normales: 1FN, 2FN, 3FN
- Boyce-Codd  $\Rightarrow$  FNBC
- Estas 4 se basan en dependencias funcionales.
- Más adelante  $\Rightarrow$  4FN (dependencias multivaluadas) y 5FN (dependencias de reunión)
- Normalización  $\Rightarrow$  análisis de los esquemas de relación en base a sus DF y claves primarias
  - Minimizar redundancia.
  - Minimizar anomalías de inserción, actualización y borrado.
- Los esquemas que no superan las pruebas  $\Rightarrow$  se descomponen en esquemas más pequeños que si lo cumplen.

# Teoría de Normalización

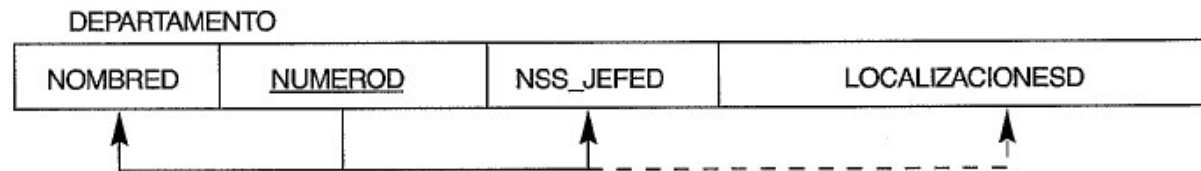
- No es necesario sólo comprobar por separado los esquemas de relación
- También se debe de cumplir:
  - Propiedad de descomposición sin pérdida (no tuplas espúreas)
  - Conservación de las dependencias.
- Se suele normalizar hasta FNBC ó 4FN
- No es necesario normalizar hasta la forma normal más alta por problemas de rendimiento  $\Rightarrow$  desnormalización.
- Conceptos: clave, superclave, clave candidata y clave primaria
- Atributo primo  $\Rightarrow$  miembro de alguna clave.

# Primera Forma Normal (1FN)

- Se definió para prohibir atributos multivaluados, compuestos y combinaciones.
- Establece:
  - Dominio de un atributo debe de ser atómico (simple, indivisible)
  - Valor de cualquier atributo debe ser un valor individual de ese dominio
- Ejemplo:
  - Esquema relación departamento con PK numerod
  - Cada departamento puede tener varios lugares

DEPARTAMENTO			
NOMBRED	<u>NUMEROD</u>	NSS_JEFED	LOCALIZACIONESD

## Primera Forma Normal (1FN)



DEPARTAMENTO

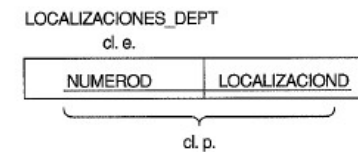
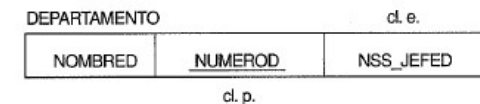
NOMBRED	<u>NUMEROD</u>	NSS_JEFED	LOCALIZACIONESD
Investigación	5	333445555	{Bellaire, Sugarland, Houston}
Administración	4	987654321	{Stafford}
Dirección	1	888665555	{Houston}

- No está en 1FN  $\Rightarrow$  dominio es atómico, pero hay conjunto de valores.
- Cómo pasar a 1FN?

# Primera Forma Normal (1FN)

- Eliminar el atributo que viola 1FN y colocarlo en otra relación.

- Ampliar la PK



DEPARTAMENTO			
NOMBRED	<u>NUMEROD</u>	NSS_JEFED	<u>LOCALIZACIONESD</u>
Investigación	5	333445555	Bellaire
Investigación	5	333445555	Sugarland
Investigación	5	333445555	Houston
Administración	4	987654321	Stafford
Dirección	1	888665555	Houston

- Desventaja  $\Rightarrow$  redundancia de tuplas
- Si se sabe máximo número de localizaciones  $\Rightarrow$  3 campos: localizaciónd1, localizaciónd2 y localizaciónd3

# Primera Forma Normal (1FN)

- Se prohíben atributos compuestos (son multivaluados)
- Se prohíben relaciones anidadas

EMP_PROY			
NSS	NOMBREE	PROY	
		NUMEROP	HORAS

EMP_PROY			
NSS	NOMBREE	NUMEROP	HORAS
123456789	Smith, John B.	1	32,5
		2	7,5
666884444	Narayan, Ramesh K.	3	40,0
453453453	English, Joyce A.	1	20,0
		2	20,0
333445555	Wong, Franklin T.	2	10,0
		3	10,0
		10	10,0
		20	10,0
999887777	Zelaya, Alicia J.	30	30,0
		10	10,0
987987987	Jabbar, Ahmad V.	10	35,0
		30	5,0
987654321	Wallace, Jennifer S.	30	20,0
		20	15,0
888665555	Borg, James E.	20	nulo

- Normalizar a 1FN

EMP_PROY1	
<u>NSS</u>	NOMBREE

EMP_PROY2		
<u>NSS</u>	NUMEROP	HORAS

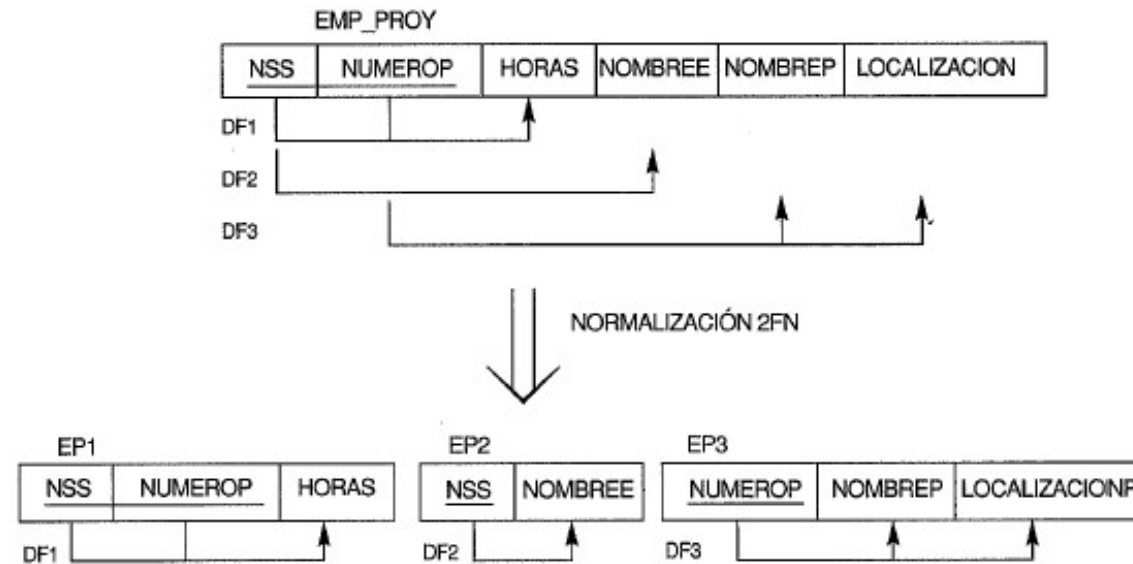
## Segunda Forma Normal (2FN)

- Se basa en el concepto de dependencia funcional total.
- DF total  $X \rightarrow Y$ , si la eliminación de cualquier atributo A de X hace que la dependencia deje de ser válida.
- DF parcial, si al eliminar A,  $(X-A) \rightarrow Y$  es válida.
- $NSS, NUMEROP \rightarrow HORAS$ , es total
- $NSS, NUMEROP \rightarrow NOMBREE$ , es parcial
- 2FN  $\rightarrow$  Si está en 1FN y la verificación de DF cuyos atributos del lado izquierdo son parte de la PK
- Si la PK contiene un solo atributo  $\Rightarrow$  está en 2FN.
- R está en 2FN si todo atributo no primo A de R depende funcionalmente de manera total de la PK de R.



## Segunda Forma Normal (2FN)

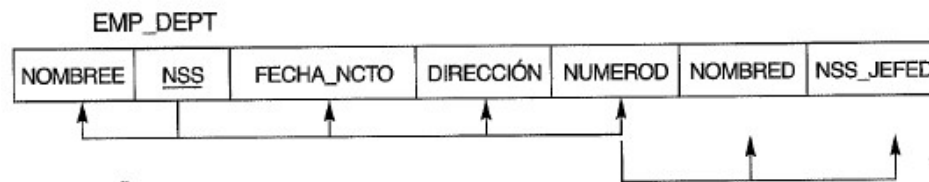
○ Ejemplo:



○ Normalización a 2FN  $\Rightarrow$  nuevos esquemas con DF totales para los atributos no primos.

## Tercera Forma Normal (3FN)

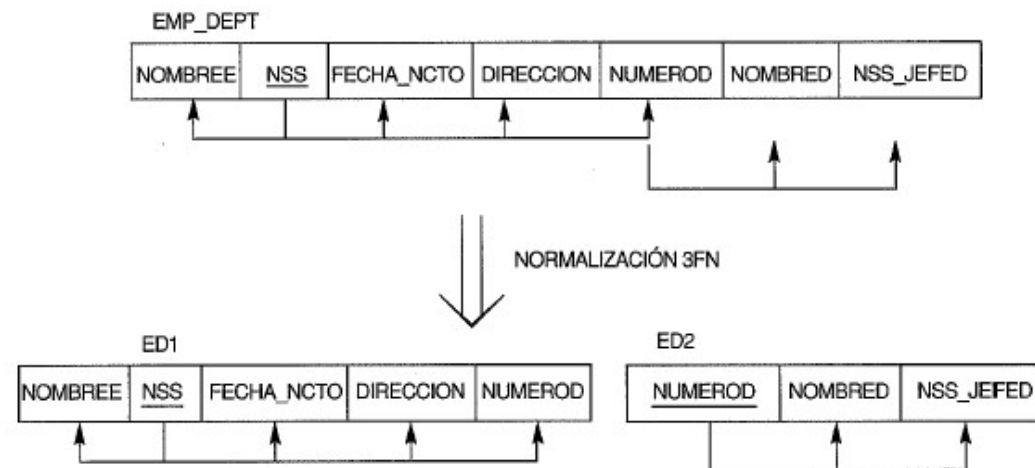
- Se basa en el concepto de DF transitiva
- Una dependencia  $X \rightarrow Y$  de R es transitiva si existe un subconjunto de atributos Z que no sea un subconjunto de cualquier clave de R y se cumple que  $X \rightarrow Z$  y  $Z \rightarrow Y$



- $NSS \rightarrow NUMEROD$
- $NUMEROD \rightarrow NSS\_JEFED$
- Por lo que  $NSS \rightarrow NSS\_JEFED$  y además NUMEROD no es clave ni forma parte de alguna clave

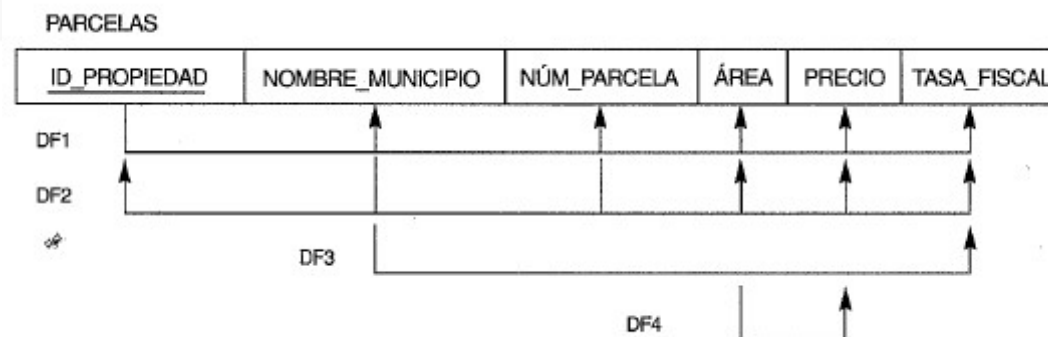
## Tercera Forma Normal (3FN)

- Un esquema R está en 3FN si lo está en 2FN y ningún atributo no primo de R depende transitivamente de la PK.



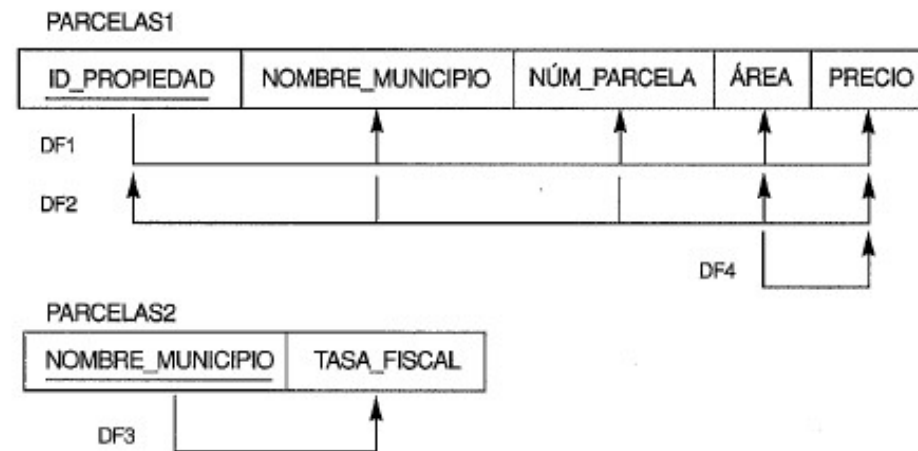
## Definición General 2FN y 3FN

- Los esquemas no deben de tener dependencias parciales ni transitivas ya que provocan anomalías de actualización.
- Hasta ahora se han visto sobre claves primarias
- Las definiciones generales de 2FN y 3Fn se hacen sobre cualquier campo que sea clave: primaria y candidata
- 2FN  $\Rightarrow$  ningún atributo no primo de R depende parcialmente de alguna clave de R.



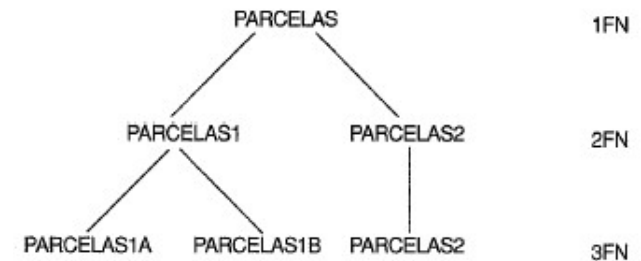
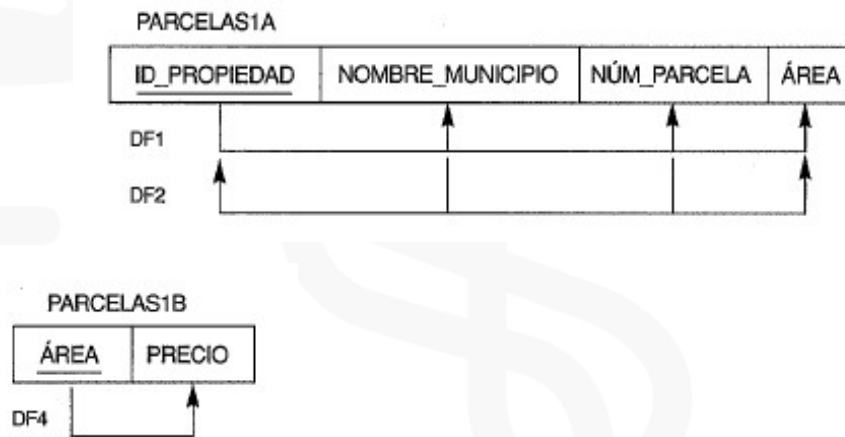
# Definición General 2FN y 3FN

- Normalización a 2FN



## Definición General 2FN y 3FN

- 3FN  $\Rightarrow$  siempre que en una DF no trivial  $X \rightarrow A$  se cumple en R, o bien (A) X es una superclave de R o (B) A es un atributo primo de R.
- Si viola ambas condiciones  $\Rightarrow$  no está en 3FN
- Ejemplo: parcelas2 está en 3FN y Parcelas1 no.



## Definición General 2FN y 3FN

- Definición alternativa 3FN:
- Un esquema de relación R está en 3FN si todo atributo no primo de R es:
  - Dependiente funcionalmente de manera total de toda clave de R.
  - Dependiente de manera no transitiva de toda clave de R.

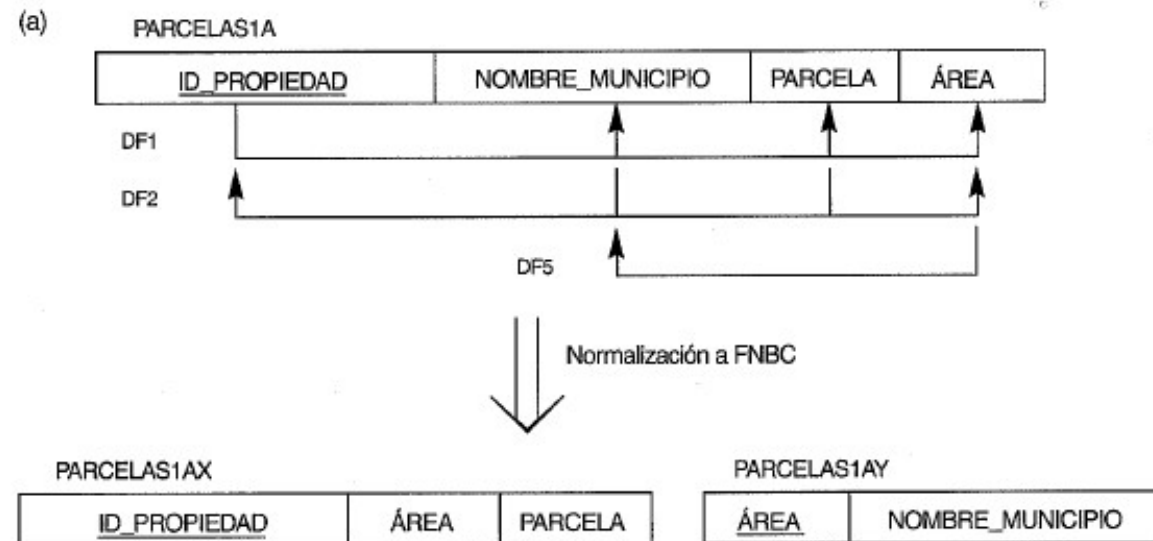
## Definición Forma Normal Boyce-Codd (FNBC)

- Más estricta que la 3FN.
- FNBC debe de estar en 3FN.
- FNBC  $\Rightarrow$  siempre que una DF no trivial  $X \rightarrow A$  es válida en R, entonces X es una superclave de R.
- Diferencia con 3FN es que la condición (B) no está presente (atributo A primo)



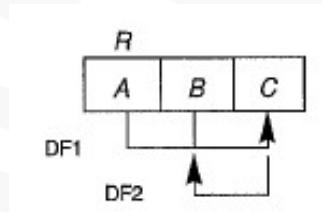
# Definición Forma Normal Boyce-Codd (FNBC)

- Ejemplo: Normalización a FNBC



## Definición Forma Normal Boyce-Codd (FNBC)

- En la práctica casi todos los esquemas de relación que están en 3FN estarán en FNBC salvo :
  - Exista una dependencia  $X \rightarrow A$  tal que  $X$  no es una superclave y  $A$  es un atributo primo.



# Definición Forma Normal Boyce-Codd (FNBC)

- Ejemplo:
- ESTUDIANTE, CURSO → PROFESOR
- PROFESOR → CURSO

IMPARTE

ESTUDIANTE	CURSO	PROFESOR
Narayan	Base de datos	Mark
Smith	Base de datos	Navathe
Smith	Sistemas operativos	Ammar
Smith	Teoría	Schulman
Wallace	Base de datos	Mark
Wallace	Sistemas operativos	Ahamad
Wong	Base de datos	Omicinski
Zelaya	Base de datos	Navathe

1. {ESTUDIANTE, PROFESOR} y {ESTUDIANTE, CURSO}.
2. {CURSO, PROFESOR} y {CURSO, ESTUDIANTE}
3. {PROFESOR, CURSO} y {PROFESOR, ESTUDIANTE}.

## Cuarta Forma Normal (4FN)

- Dependencia funcional multivaluada (DMV)
- Son una consecuencia de la 1FN
- Si tenemos 2 o más atributos multivaluados independientes en el mismo esquema de relación  $\Rightarrow$  repetir valores de cada valor de un atributo con cada valor del otro atributo. Dependencia multivaluada.

- Ejemplo:

EMP		
NOMBREE	NOMBREP	NOMBRED
Smith	X	John
Smith	Y	Anna
Smith	X	Anna
Smith	Y	John

- Siempre que en una relación se mezclan 2 relaciones 1:N independientes A:B y A:C puede surgir una DMV

## Cuarta Forma Normal (4FN)

- DMV:  $X \twoheadrightarrow Y$  sobre R, donde X e Y son subconjuntos en R:
- Existen 2 tuplas  $t_1$  y  $t_2$  en r tales que  $t_1[X]=t_2[X]$ , entonces deberían de existir también dos tuplas  $t_3$  y  $t_4$  de r con las siguientes propiedades:

$$t_3[X]=t_4[X]=t_1[X]=t_2[X]$$

$$t_3[Y]=t_1[Y] \text{ y } t_4[Y]=t_2[Y]$$

$$t_3[Z]=t_2[Z] \text{ y } t_4[Z]=t_1[Z]$$

Donde  $Z=R-(XUY)$

Entonces:  $X \twoheadrightarrow Y$  y también  $X \twoheadrightarrow Z$

$NOMBREE \twoheadrightarrow NOMBREP$

$NOMBREE \twoheadrightarrow NOMBRED$

EMP

NOMBREE	NOMBREP	NOMBRED
Smith	X	John
Smith	Y	Anna
Smith	X	Anna
Smith	Y	John

## Cuarta Forma Normal (4FN)

- DMV Trivial:  $X \twoheadrightarrow Y$  si  $Y \subseteq X$  ó  $X \cup Y = R$ . (Se cumple siempre)
- Existen reglas de inferencia para DMV.

(R11) (Regla reflexiva para DF): Si  $X \subset Y$ , entonces  $X \rightarrow Y$ .

(R12) (Regla de aumento para DF):  $\{X \rightarrow Y\} \cup \{XZ \rightarrow YZ\}$ .

(R13) (Regla transitiva para DF):  $\{X \rightarrow Y, Y \rightarrow Z\} \cup \{X \rightarrow Z\}$ .

(R14) (Regla de complemento para DMV):  $\{X \twoheadrightarrow Y\} \cup \{X \twoheadrightarrow (R - (X \cup Y))\}$ .

(R15) (Regla de aumento para DMV): Si  $X \twoheadrightarrow Y$  y  $W \subset Z$  entonces  $WX \twoheadrightarrow YZ$ .

(R16) (Regla transitiva para DMV):  $\{X \twoheadrightarrow Y, Y \twoheadrightarrow Z\} \cup \{X \twoheadrightarrow (Z - Y)\}$ .

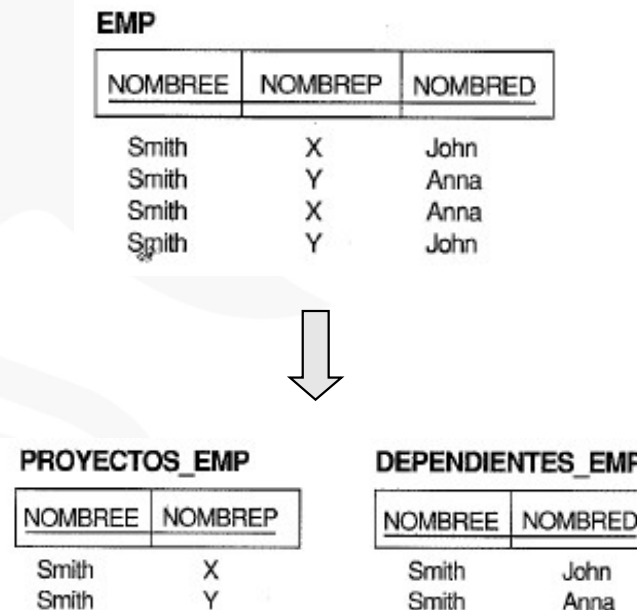
(R17) (Regla de réplica (DF a DMV)):  $\{X \rightarrow Y\} \cup \{X \twoheadrightarrow Y\}$ .

(R18) (Regla de combinación para DF y DMV): Si  $X \twoheadrightarrow Y$  y existe  $W$  con las propiedades de que (a)  $W \cap Y$  está vacío, (b)  $W \rightarrow Z$  y (c)  $Y \subset Z$ , entonces  $X \rightarrow Z$ .

- Donde  $X, Y, Z$  son un subconjunto de  $R(A_1, A_2, \dots, A_n)$

## Cuarta Forma Normal (4FN)

- 4FN  $\Rightarrow$  no cumple cuando hay DMV no deseables y que se utilizan para descomponer el esquema de relación.
- Está en 4FN respecto a un conjunto F de DF y DMV, si para cada DMV no trivial  $X \rightarrow \rightarrow Y$  en  $F^+$ , X es una superclave de R.



## Cuarta Forma Normal (4FN)

(a) EMP

NOMBREE	NOMBREPR	NOMBRED
Silva	X	Juan
Silva	Y	Ana
Silva	X	Ana
Silva	Y	Juan
Bravo	W	Jaime
Bravo	X	Jaime
Bravo	Y	Jaime
Bravo	Z	Jaime
Bravo	W	Juana
Bravo	X	Juana
Bravo	Y	Juana
Bravo	Z	Juana
Bravo	W	Beto
Bravo	X	Beto
Bravo	Y	Beto
Bravo	Z	Beto

(b) PROYECTOS\_EMP

NOMBREE	NOMBREPR
Silva	X
Silva	Y
Silva	W
Bravo	X
Bravo	Y
Bravo	Z

DEPENDIENTES\_EMP

NOMBREE	NOMBRED
Silva	Ana
Silva	Juan
Bravo	Jaime
Bravo	Juana
Bravo	Beto

- Necesario que la descomposición sea sin pérdida:
- $R_1 = (X \cup Y)$  y  $R_2 = (R - Y)$  para una DMV  $X \twoheadrightarrow Y$



## Quinta Forma Normal (5FN)

- Basada en la Dependencia de Reunión DR,  $DR(R_1, R_2, \dots, R_n)$
- Especifica una restricción sobre los estados  $r$  de  $R$
- Todo estado permitido de  $r$  de  $R$  debe de tener una descomposición de reunión sin pérdida para dar  $R_1, R_2, \dots, R_n$

$$* (\pi_{\langle R_1 \rangle}(r), \pi_{\langle R_2 \rangle}(r), \dots, \pi_{\langle R_n \rangle}(r)) = r$$

- Una DMV es un caso especial de una DR donde  $n=2$
- $DR(R_1, R_2)$  implica una DMV :
  - $(R_1 \cap R_2) \rightarrow \rightarrow (R_1 - R_2)$
  - $(R_1 \cap R_2) \rightarrow \rightarrow (R_2 - R_1)$
- Una DR es trivial si un esquema  $R_i$  en  $DR(R_1, R_2, \dots, R_n)$  es igual a  $R$

## Quinta Forma Normal (5FN)

- 5FN  $\Rightarrow$  Un esquema R está en 5FN respecto a un conjunto F de DF, DMV y DR si para cada dependencia de reunión no trivial  $DR(R_1, R_2, \dots, R_n)$  en  $F^+$  toda  $R_i$  es una superclave de R

- Ejemplo:

**SUMINISTRAR**

NOMPROV	NOMBRECOMP	NOMPROY
Smith	Perno	ProyX
Smith	Tuerca	ProyY
Adamsky	Perno	ProyY
Walton	Tuerca	ProyZ
Adamsky	Clavo	ProyX
Adamsky	Perno	ProyX
Smith	Perno	ProyY

- Sólo tiene claves

## Quinta Forma Normal (5FN)

- Restricción adicional:
  - Proveedor v suministra un componente c
  - Proyecto p utiliza el componente c
  - Proveedor v suministra al menos un componente al proyecto p
  - Proveedor v suministra el componente c al proyecto p
- Esta restricción especifica una  $DR(R_1, R_2, R_3)$  , entre tres proyecciones de SUMINISTRAR
  - $R_1(NOMPROV, NOMBRECOMP)$
  - $R_2(NOMPROV, NOMPROY)$
  - $R_3(NOMBRECOMP, NOMPROY)$

## Quinta Forma Normal (5FN)

**SUMINISTRAR**

NOMPROV	NOMBRECOMP	NOMPROY
Smith	Perno	ProyX
Smith	Tuerca	ProyY
Adamsky	Perno	ProyY
Walton	Tuerca	ProyZ
Adamsky	Clavo	ProyX
Adamsky	Perno	ProyX
Smith	Perno	ProyY

**R1**

NOMPROV	NOMBRECOMP
Smith	Perno
Smith	Tuerca
Adamsky	Perno
Walton	Tuerca
Adamsky	Clavo

**R2**

NOMPROV	NOMPROY
Smith	ProyX
Smith	ProyY
Adamsky	ProyY
Walton	ProyZ
Adamsky	ProyX

**R3**

NOMBRECOMP	NOMPROY
Perno	ProyX
Tuerca	ProyY
Perno	ProyY
Tuerca	ProyZ
Clavo	ProyX

- Difícil deducir DR. En la práctica se utiliza poco.

# Proceso General Diseño Bases de Datos

- Varios modos de obtener un esquema R:
  1.  $R$  puede haberse generado al convertir un diagrama E-R en un conjunto de tablas.
  2.  $R$  puede haber sido una sola relación que contuviera *todos* los atributos que resultan de interés. El proceso de normalización divide a  $R$  en relaciones más pequeñas.
  3.  $R$  puede haber sido el resultado de algún diseño ad hoc de relaciones, que hay que comprobar para verificar que satisface la forma normal deseada.
- Proceso desnormalización  $\Rightarrow$  esquema normalizado y hacerlo no normalizado para mejorar rendimiento.
- Duplicación de información + más cuidado en actualizaciones
- Ejemplo: nombre titular + importe préstamo