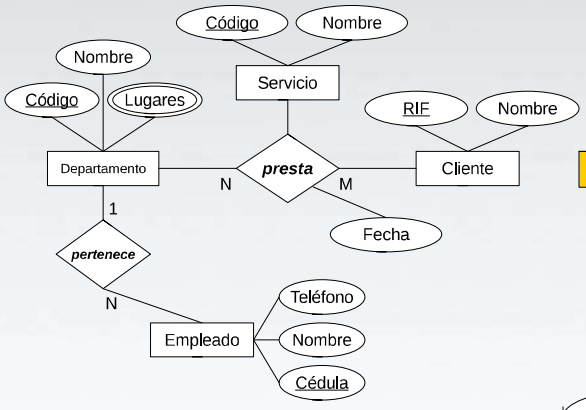
Dependencias Funcionales y Normalización

(...o el arte de no redundar...)

Universidad de los Andes
Demián Gutierrez
Enero 2010

Diseño Descendente



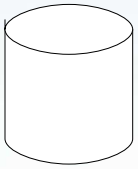
Departamento (Código, Nombre)

Empleado (Cédula, Nombre, Teléfono, CodDpto)

Cliente (RIF, Nombre)

Servicio (Código, Nombre)

Presta (CódDpto, CodServ, RIF, Fecha)



Base de Datos Relacional

El Diseño Descendente se basa en el uso de los modelos semánticos de datos.

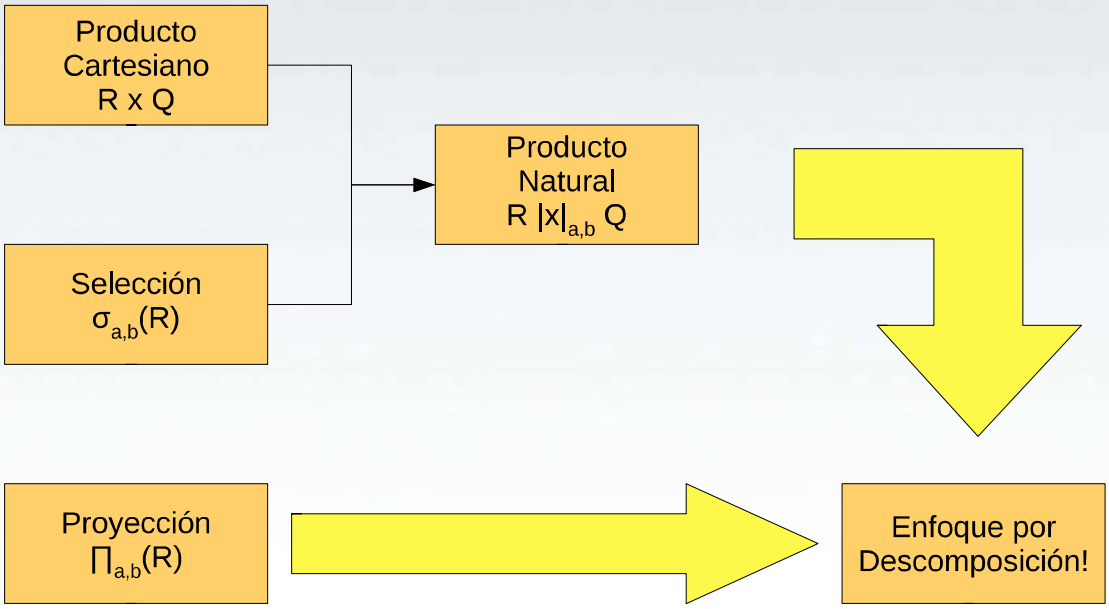
Enfoque por Descomposición

Enfoque por Descomposición o
Síntesis Relacional

Enfoque por Descomposición

- Definir ***relaciones universales*** compuestas de todos los atributos de la base de datos y luego descomponerlas, utilizando el proceso de normalización de las relaciones, en subrelaciones que no sufran anomalías
- Es un proceso de ***refinamiento paso a paso***, que lleva al ***aislamiento de las entidades y asociaciones*** del mundo real [Codd, 1979]
- Se basa las operaciones fundamentales del álgebra relacional (Proyección / Reunión Natural)

Algunas Operaciones del Álgebra Relacional



Proyección

$$\Pi_{placa,\ marca}(Carro)$$

$$\Pi_{marca,modelo,color}(Carro)$$

Carro	placa	marca	modelo	color
	MBO34L	Ford	Ka	verde
	LDA75K	Toyota	corollaXL	blanco
	ADA89A	Fiat	siena	gris
	LBF78G	Toyota	corollaXL	blanco
	XSA67D	Ford	Ka	rojo

Se eliminan todas salvo una de las tuplas repetidas (Ej, <*Toyota, corollaXL, blanco*>)

R	placa	***marca***
	MBO34L	Ford
	LDA75K	Toyota
	ADA89A	Fiat
	LBF78G	Toyota
	XSA67D	Ford

Q	***marca***	modelo	color
	Ford	Ka	verde
	Fiat	siena	gris
	Toyota	corollaXL	blanco
	Ford	Ka	rojo

Producto Cartesiano

UNIVERSIDAD DE LOS ANDES VENEZUELA

$$Carro = R \times Q$$

R	placa	***marca***
	ADA89A	Fiat
	LBF78G	Toyota
	XSA67D	Ford

Q	***marca***	modelo	color
	Fiat	siena	gris
	Toyota	corollaXL	blanco
	Ford	Ka	rojo

Carro	placa	marcaR	marcaQ	modelo	color
	ADA89A	Fiat	Fiat	siena	gris
	ADA89A	Fiat	Toyota	corollaXL	blanco
	ADA89A	Fiat	Ford	Ka	rojo
	LBF78G	Toyota	Fiat	siena	gris
	LBF78G	Toyota	Toyota	corollaXL	blanco
	LBF78G	Toyota	Ford	Ka	rojo
	XSA67D	Ford	Fiat	siena	gris
	XSA67D	Ford	Toyota	corollaXL	blanco
	XSA67D	Ford	Ford	Ka	rojo

No es relevante para el producto cartesiano, pero es notable, que en algunas tuplas marcaR = marcaQ

Selección

UNIVERSIDAD DE LOS ANDES VENEZUELA

$$R = \sigma_{marca=Ford}(Carro)$$

Carro	placa	marca	color
	MBO34L	Ford	verde
	LDA75K	Toyota	blanco
	ADA89A	Fiat	gris
	LBF78G	Toyota	blanco
	XSA67D	Ford	rojo

R	placa	marca	color
	MBO34L	Ford	verde
	XSA67D	Ford	rojo

$$R = \sigma_{marca=Ford \wedge color=rojo}(Carro)$$

Carro	placa	marca	color
	MBO34L	Ford	verde
	LDA75K	Toyota	blanco
	ADA89A	Fiat	gris
	LBF78G	Toyota	blanco
	XSA67D	Ford	rojo

R	placa	marca	color
	XSA67D	Ford	rojo

Reunión Natural

$$Carro = R |x|_{marca} Q$$

R	placa	***marca***
	MBO34L	Ford
	LDA75K	Toyota
	ADA89A	Fiat
	LBF78G	Toyota
	XSA67D	Ford

Q	***marca***	modelo	color
	Ford	Ka	verde
	Toyota	corollaXL	blanco
	Fiat	siena	gris
	Ford	Ka	rojo

Carro	placa	marca	modelo	color
	MBO34L	Ford	Ka	verde
	MBO34L	Ford	Ka	rojo
	LDA75K	Toyota	corollaXL	blanco
	ADA89A	Fiat	siena	gris
	LBF78G	Toyota	corollaXL	blanco
	XSA67D	Ford	Ka	verde
	XSA67D	Ford	Ka	rojo

$Carro = \sigma_{(R.marca=Q.marca)}(R \times Q)$

o bien:

$Carro = R |x| Q$

Descomposición

- Es el reemplazo de una relación $R(A_1, A_2, ..., A_n)$, por una colección de relaciones $R_1, R_2, ..., R_n$ obtenidas de las ***proyecciones*** de R y tal que la relación resultado de las ***reuniones*** $R_1 |x| R_2 |x| ... |x| R_n$ tiene el mismo esquema que R.
- Ej: Si tenemos:

  $R_1 = \Pi_{placa,\ color,\ modelo}(Carro)$

  $R_2 = \Pi_{modelo,\ marca}(Carro)$

  ...resulta que:

  $R_1 |x| R_2 = Carro$

R |x| Q ≠ Carro

UNIVERSIDAD DE LOS ANDES VENEZUELA

¿¿¿ Carro = R|x|$_{marca}$Q ???

Carro	placa	marca	modelo	color
	MBO34L	Ford	Ka	verde
	MBO34L	Ford	Ka	rojo
	LDA75K	Toyota	corollaXL	blanco
	ADA89A	Fiat	siena	gris
	LBF78G	Toyota	corollaXL	blanco
	XSA67D	Ford	Ka	verde
	XSA67D	Ford	Ka	rojo

La relación original

¿Qué sucedió aquí?

Carro	placa	marca	modelo	color
	MBO34L	Ford	Ka	verde
	LDA75K	Toyota	corollaXL	blanco
	ADA89A	Fiat	siena	gris
	LBF78G	Toyota	corollaXL	blanco
	XSA67D	Ford	Ka	rojo

R |x| Q ≠ Carro

UNIVERSIDAD DE LOS ANDES VENEZUELA

¿¿¿ Carro = R|x|$_{marca}$Q ???

Carro	placa	marca	modelo	color
	MBO34L	Ford	Ka	verde
	MBO34L	Ford	Ka	rojo
	LDA75K	Toyota	corollaXL	blanco
	ADA89A	Fiat	siena	gris
	LBF78G	Toyota	corollaXL	blanco
	XSA67D	Ford	Ka	verde
	XSA67D	Ford	Ka	rojo

La relación original

...En el fondo, se perdió información...

Carro	placa	marca	modelo	color
	MBO34L	Ford	Ka	verde
	LDA75K	Toyota	corollaXL	blanco
	ADA89A	Fiat	siena	gris
	LBF78G	Toyota	corollaXL	blanco
	XSA67D	Ford	Ka	rojo

Descomposición sin Pérdida

- Es la descomposición de una relación R en R_1, R_2, ..., R_n tal que ***para toda extensión de R*** se tiene que $R = R_1 |x| R_2 |x| \ldots |x| R_n$.
- El problema de la concepción de bases de datos relacionales se reduce a la descomposición sin pérdida de las relaciones universales con todos sus atributos, en subrelaciones que no tengan anomalías (Tal como veremos más adelante).

Descomposición sin Pérdida

¿A dónde nos puede llevar una mala descomposición?

¿Y una descomposición insuficiente?

Anomalías de Actualización

UNIVERSIDAD DE LOS ANDES VENEZUELA

Prof_Depto	Cédula	Nombre_Prof	Fecha_Nac	Código_Depto	Nombre_Depto
	9.980.623	Pedro Pérez	01/06/73	01	Computación
	10.334.890	Luis García	01/06/76	02	Control
	12.334.222	Mario Lobo	01/06/77	01	Computación
	13.434.122	José Rivero	01/06/78	03	Investigación
	13.566.002	Frank Chacón	01/12/78	***NULL***	***NULL***
	17.544.672	Miguel Bravo	01/06/84	02	Control
	18.244.670	Andrés León	01/06/85	01	Computación

¿Qué problemas o anomalías se pueden producir en esta relación?

¿Qué "cosas malas" pueden ocurrir?

Anomalías de Actualización

UNIVERSIDAD DE LOS ANDES VENEZUELA

Prof_Depto	Cédula	Nombre_Prof	Fecha_Nac	Código_Depto	Nombre_Depto
	9.980.623	Pedro Pérez	01/06/73	01	Computación
	10.334.890	Luis García	01/06/76	02	Control
	12.334.222	Mario Lobo	01/06/77	01	Computación
	13.434.122	José Rivero	01/06/78	03	Investigación
	13.566.002	Frank Chacón	01/12/78	***NULL***	***NULL***
	17.544.672	Miguel Bravo	01/06/84	02	Control
	18.244.670	Andrés León	01/06/85	01	Computación

Anomalías de Inserción: Cada vez que se inserta un profesor es necesario repetir los datos del departamento

Anomalías de Modificación: Cada vez se actualiza un departamento es necesario hacer los cambios en cada una de las tuplas correspondientes a ese departamento

Anomalías de Eliminación: Si se elimina el último empleado asociado a un departamento se pierde la información del departamento

Normalización

UNIVERSIDAD DE LOS ANDES VENEZUELA

El esquema relacional es un **modelo de la realidad** bajo la forma de una colección de relaciones para que:

- La creación, modificación y supresión de datos de ***forma eficaz***. Para esto es indispensable eliminar toda ***redundancia innecesaria***. Idealmente, ante la ocurrencia de ***un evento*** se desea que éste se traduzca en el manejo de ***una única tupla*** en la extensión del esquema relacional.
- La modificación del esquema relacional por la evolución de la percepción de la realidad, sea lo más simple posible.
- La comprensión de la realidad sea facilitada por el esquema relacional (Ver el esquema -> Comprender la Realidad).

Descomposición sin Pérdida

UNIVERSIDAD DE LOS ANDES VENEZUELA

¿Cómo generamos buenas descomposiciones?

¿Qué relaciones tiene esto con el modelado semántico: ERE, OO, etc?

UNIVERSIDAD DE LOS ANDES VENEZUELA

¿Qué es normalizar una base de datos?

Es encontrar una **descomposición adecuada** de la "relación universal" de la base de datos que nos permite **cumplir con los criterios de** eficacia, ausencia de redundancia, evolución, comprensión, flexibilidad enunciados anteriormente

UNIVERSIDAD DE LOS ANDES VENEZUELA

Técnica formal para organizar datos basada en formas normales

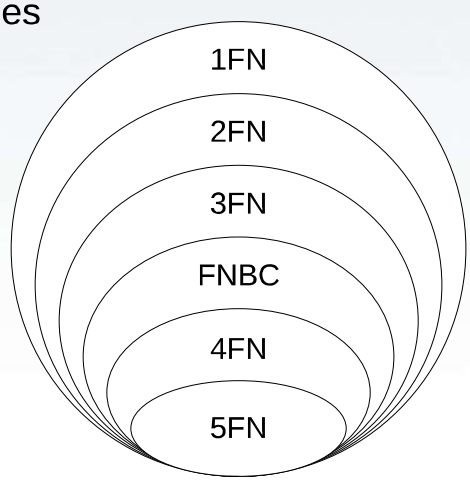
- Primera, segunda y tercera formas normales

  Codd 1970
- Forma normal de Boyce y Codd (FNBC)

  Boyce – Codd 1974
- Cuarta forma normal (4FN)

  Fagin 1977
- Quinta forma normal (5FN)

  Fagin 1979



UNIVERSIDAD DE LOS ANDES VENEZUELA

NUEVAMENTE:

¿Qué es normalizar una base de datos?

Es llevar el esquema de la base de datos a alguna de las formas normales anteriormente mencionadas

...y para eso, necesitamos conocer y comprender el concepto de "dependencia funcional"

UNIVERSIDAD DE LOS ANDES VENEZUELA

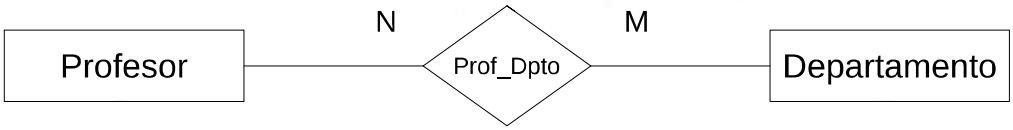
- Son ***restricciones de integridad*** que permiten conocer que ***interrelaciones*** existen entre dos o más atributos del mundo real.
- Son propiedades ***inherentes*** al ***contenido semántico*** de los datos, que se han de cumplir para cualquier extensión del esquema de relación.
- Informalmente, Y **depende funcionalmente** de X si ***cada valor*** de X tiene asociado ***siempre el mismo*** valor de Y en una relación R que contiene a X y Y como atributos.

UNIVERSIDAD DE LOS ANDES VENEZUELA

El resultado de una consulta cualquiera

(por ejemplo, de un producto entre la tabla profesor y departamento):

Cédula	Fecha_Nac	Sexo	Código_Depto	Nombre_Depto
9.980.623	06/01/73	M	01	Computación
10.334.890	**06/01/76**	**F**	01	Computación
17.544.672	**06/01/84**	**M**	03	Investigación
12.334.222	**06/01/77**	M	02	Control
13.566.002	**12/01/78**	**F**	02	Control
10.334.890	**06/01/76**	**F**	02	Control
12.334.222	**06/01/77**	**M**	01	Computación
13.434.122	06/01/78	F	03	Investigación
13.566.002	**12/01/78**	**F**	03	Investigación
17.544.672	**06/01/84**	**M**	02	Control
18.244.670	06/01/85	M	01	Computación



UNIVERSIDAD DE LOS ANDES VENEZUELA

Cédula	Fecha_Nac	Sexo	Código_Depto	Nombre_Depto
9.980.623	06/01/73	M	01	Computación
10.334.890	**06/01/76**	**F**	01	Computación
17.544.672	**06/01/84**	**M**	03	Investigación
12.334.222	**06/01/77**	M	02	Control
13.566.002	**12/01/78**	**F**	02	Control
10.334.890	**06/01/76**	**F**	02	Control
12.334.222	**06/01/77**	**M**	01	Computación
13.434.122	06/01/78	F	03	Investigación
13.566.002	**12/01/78**	**F**	03	Investigación
17.544.672	**06/01/84**	**M**	02	Control
18.244.670	06/01/85	M	01	Computación

¿Qué características destacan en la información de los profesores?

¿Se repite la información de los profesores?

¿Que sucede con los atributos Fecha_Nac y Sexo con respecto a la cédula?

¿Qué relaciones existen?

Dependencias Funcionales

UNIVERSIDAD DE LOS ANDES VENEZUELA

Cédula	Fecha_Nac	Sexo	Código_Depto	Nombre_Depto
9.980.623	06/01/73	M	01	Computación
10.334.890	06/01/76	F	01	Computación
17.544.672	06/01/84	M	03	Investigación
12.334.222	06/01/77	M	02	Control
13.566.002	12/01/78	F	02	Control
10.334.890	06/01/76	F	02	Control
12.334.222	06/01/77	M	01	Computación
13.434.122	06/01/78	F	03	Investigación
13.566.002	12/01/78	F	03	Investigación
17.544.672	06/01/84	M	02	Control
18.244.670	06/01/85	M	01	Computación

¿Y con respecto a la información de los departamentos?

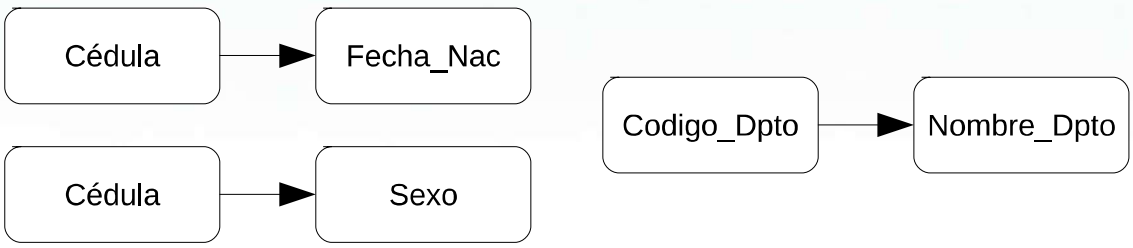
¿Se repite?

¿Qué ocurre con los atributos de las distintas filas?

Dependencias Funcionales

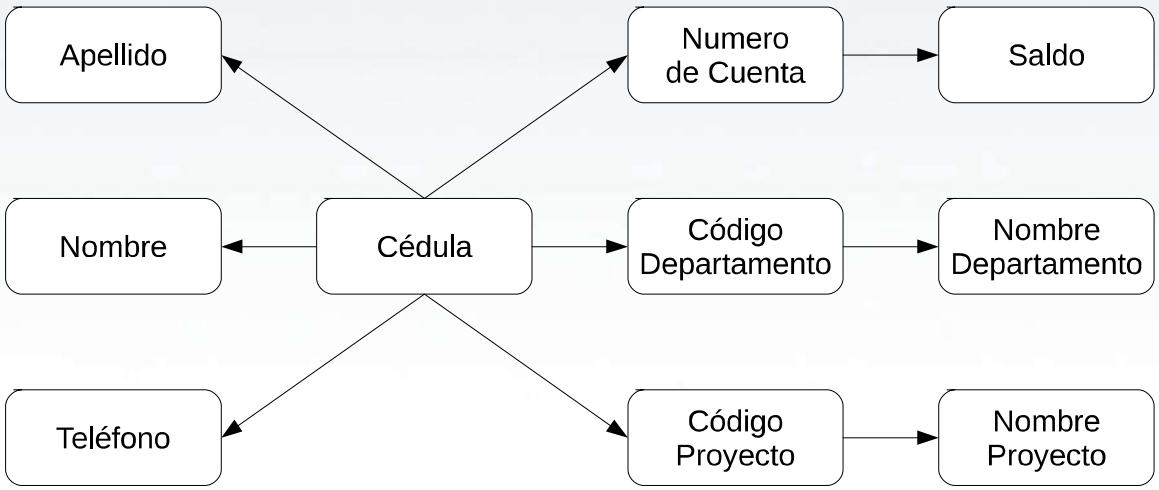
UNIVERSIDAD DE LOS ANDES VENEZUELA

Cédula	Fecha_Nac	Sexo	Código_Depto	Nombre_Depto
9.980.623	06/01/73	M	01	Computación
10.334.890	06/01/76	F	01	Computación
17.544.672	06/01/84	M	03	Investigación
12.334.222	06/01/77	M	02	Control
13.566.002	12/01/78	F	02	Control
10.334.890	06/01/76	F	02	Control
12.334.222	06/01/77	M	01	Computación
13.434.122	06/01/78	F	03	Investigación
13.566.002	12/01/78	F	03	Investigación
17.544.672	06/01/84	M	02	Control
18.244.670	06/01/85	M	01	Computación



Dependencias Funcionales

Son restricciones de integridad que permiten conocer que interrelaciones existen entre los atributos del mundo real

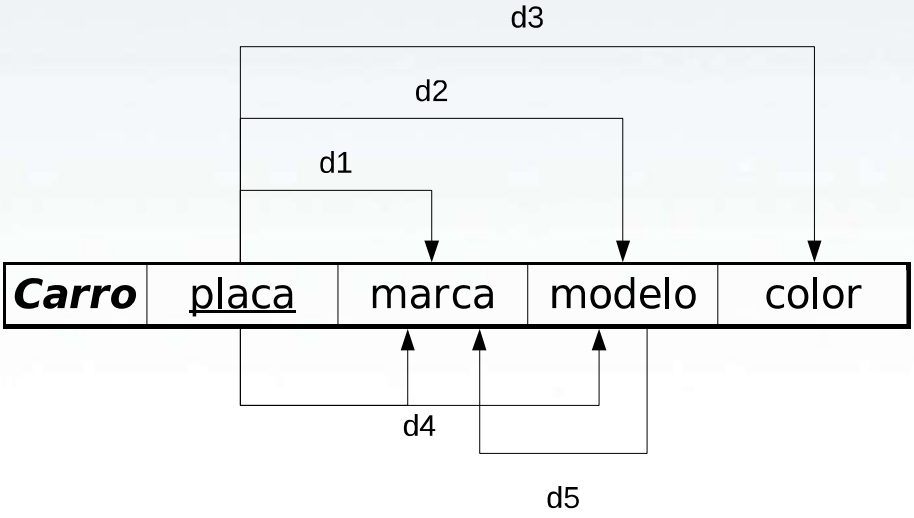


Dependencias Funcionales

- Sea R(A1, A2, ..., An) y X y Y dos subconjuntos del conjunto formado por {A1, A2, ..., An}. Se dice que X→Y (X implica a Y o Y depende funcionalmente de X) si para toda extensión **r** de R y para toda tupla t1 y t2 de **r** en la que t1[X]=t2[X] implica que t1[Y]=t2[Y]

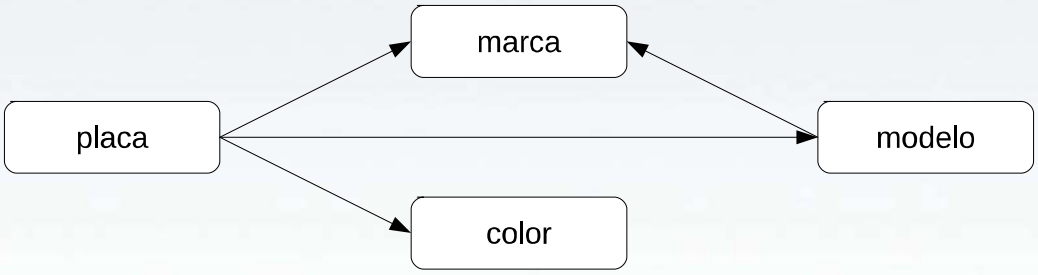
Dependencias Funcionales

- Ejemplos: *placa*→*marca*, *placa*→*modelo*, *placa*→*color*, *placa*→(*marca*, *modelo*), *modelo*→*marca*.



Grafos / Redes de DF

Los nodos representan atributos y las aristas representan DF

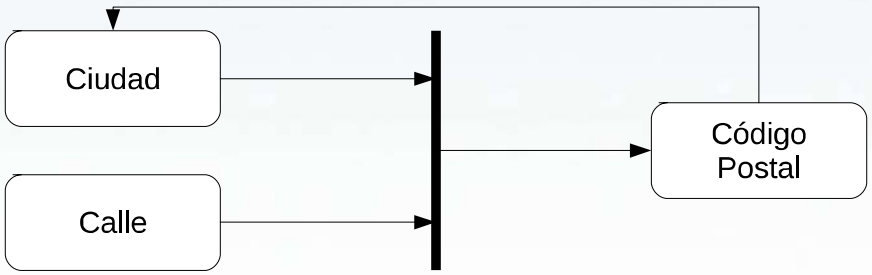


placa → marca
placa → modelo
placa → color
modelo → marca

Grafos / Redes de DFE

UNIVERSIDAD DE LOS ANDES VENEZUELA

Si hay más de un atributo del lado izquierdo se utiliza una línea que sirva para unir todos los atributos de la parte izquierda, y de ella sale una línea al atributo de la parte derecha



ciudad, calle → código_postal
código_postal → ciudad

Dependencias Funcionales

UNIVERSIDAD DE LOS ANDES VENEZUELA

Se viola la dependencia funcional *placa→color*

Carro	placa	marca	modelo	color
	MBO34L	Ford	Ka	**verde**
	MBO34L	Ford	Ka	**rojo**
	LDA75K	Toyota	corollaXL	blanco
	ADA89A	Fiat	siena	gris
	LBF78G	Toyota	corollaXL	blanco
	XSA67D	Ford	Ka	**verde**
	XSA67D	Ford	Ka	**rojo**

Carro	placa	marca	modelo	color
	MBO34L	**Ford**	**Ka**	verde
	LDA75K	**Toyota**	**corollaXL**	blanco
	ADA89A	Fiat	siena	gris
	LBF78G	**Toyota**	**corollaXL**	blanco
	XSA67D	**Ford**	**Ka**	rojo

Se cumplen todas las dependencias funcionales

Se viola la dependencia funcional *modelo→marca*

Carro	placa	marca	modelo	color
	MBO34L	**Ford**	**Ka**	verde
	XXR34L	**Chrysler**	**Ka**	rojo
	LDA75K	**Toyota**	**corollaXL**	blanco
	ADA89A	Fiat	siena	gris
	LBF78G	**Toyota**	**corollaXL**	blanco

Propiedades de las Dependencias Funcionales

1. Reflexibidad:

   Si $Y \subseteq X \Rightarrow X \rightarrow Y$

   Ej: color $\rightarrow$ color, ó bien: (marca, modelo) $\rightarrow$ marca

2. Aumento:

   Si $X \rightarrow Y \Rightarrow XZ \rightarrow YZ$

   Ej: modelo $\rightarrow$ marca $\Rightarrow$ (modelo, color) $\rightarrow$ (marca, color)

3. Transitividad:

   Si $X \rightarrow Y$ y $Y \rightarrow Z \Rightarrow X \rightarrow Z$

   Ej: placa $\rightarrow$ modelo y modelo $\rightarrow$ marca $\Rightarrow$ placa $\rightarrow$ marca

A estas primeras reglas se las con como las ***reglas inferencia (o axio de Armstron***

Propiedades de las Dependencias Funcionales

UNIVERSIDAD DE LOS ANDES VENEZUELA

4. Descomposición:

Si $X \rightarrow YZ \Rightarrow X \rightarrow Y$ y $X \rightarrow Z$

Ej: placa $\rightarrow$ (modelo, marca) $\Rightarrow$
placa $\rightarrow$ modelo y placa $\rightarrow$ marca

5. Unión:

Si $X \rightarrow Y$ y $X \rightarrow Z \Rightarrow X \rightarrow YZ$

Ej: placa $\rightarrow$ modelo y placa $\rightarrow$ marca $\Rightarrow$
placa $\rightarrow$ (modelo, marca)

Estas tres últimas reglas se pueden inferir o demostrar a partir de las tres primeras

6. Pseudotransitividad:

Si $X \rightarrow Y$ y $WY \rightarrow Z \Rightarrow WX \rightarrow Z$

Ej: placa $\rightarrow$ modelo y (marca, modelo) $\rightarrow$ potencia $\Rightarrow$
(marca, placa) $\rightarrow$ potencia

Propiedades de las Dependencias Funcionales (RESUMEN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

1. Reflexibidad:
   Si $Y \subseteq X \Rightarrow X \rightarrow Y$
   Ej: color $\rightarrow$ color,
   (marca, modelo) $\rightarrow$ marca
2. Aumento:
   Si $X \rightarrow Y \Rightarrow XZ \rightarrow YZ$
   Ej: modelo $\rightarrow$ marca $\Rightarrow$ (modelo, color) $\rightarrow$ (marca, color)
3. Transitividad:
   Si $X \rightarrow Y$ y $Y \rightarrow Z \Rightarrow X \rightarrow Z$
   Ej: placa $\rightarrow$ modelo y modelo $\rightarrow$ marca $\Rightarrow$ placa $\rightarrow$ marca
4. Descomposición:
   Si $X \rightarrow YZ \Rightarrow X \rightarrow Y$ y $X \rightarrow Z$
   Ej: placa $\rightarrow$ (modelo, marca) $\Rightarrow$
   placa $\rightarrow$ modelo y placa $\rightarrow$ marca
5. Unión:
   Si $X \rightarrow Y$ y $X \rightarrow Z \Rightarrow X \rightarrow YZ$
   Ej: placa $\rightarrow$ modelo y placa $\rightarrow$ marca $\Rightarrow$
   placa $\rightarrow$ (modelo, marca)
6. Pseudotransitividad:
   Si $X \rightarrow Y$ y $WY \rightarrow Z \Rightarrow WX \rightarrow Z$
   Ej: placa $\rightarrow$ modelo y (marca, modelo) $\rightarrow$ potencia $\Rightarrow$
   (marca, placa) $\rightarrow$ potencia

Dependencias Funcionales Elementales

Dependencias Funcionales Elementales (DFE)

- Es una dependencia funcional de la forma $X \rightarrow A$, donde A es un ***atributo único no incluido en X*** y donde no existe un $X' \subseteq X$ tal que $X' \rightarrow A$
- Ejemplos:
  - placa $\rightarrow$ modelo (*Es una DFE*)
  - placa $\rightarrow$ (modelo, marca) (*No es DFE*)
  - (A, B, C) $\rightarrow$ A (*No es DFE*) (A incluido en X)
  - Si (A, B) $\rightarrow$ X luego (A, B, C) $\rightarrow$ X (*No es DFE*)

Dependencias Funcionales Elementales (DFE)

UNIVERSIDAD DE LOS ANDES VENEZUELA

En parte, la importancia de las DFE es que no es posible simplificarlas más. Por ejemplo la siguiente DFNE:

(A, B, C) →A

Se puede simplificar por reflexividad en la siguiente DFE:

(B, C) →A

O bien la siguiente DFNE:

A→(B, C)

Se puede simplificar en las siguientes DFE:

A→B y A→C

Dependencias Funcionales Elementales (DFE)

UNIVERSIDAD DE LOS ANDES VENEZUELA

Pero por ejemplo, las siguiente DFE:

(B, C) →A

K →J

No se pueden simplificar más

Como veremos más adelante esto es particularmente útil en ciertos casos...

Cierre Transitivo de un Conjunto de DF

- Dado un conjunto C de dependencias funcionales, el cierre transitivo de C, denotado por C^+, es un nuevo conjunto formado por las ***dependencias funcionales elementales*** más ***todas las calculadas*** aplicando los axiomas de Armstrong y sus reglas inferidas.
- ***Dos conjuntos de dependencias funcionales son equivalentes si tienen el mismo C^+.*** (o si a partir de los conjuntos iniciales se puede llegar a un mismo conjunto para ambos casos)

Cierre Transitivo de un Conjunto de DF

01) A →BC
02) B →E
03) CD→EF

```
04) CD→E  (de 3 por Descomposición)
05) CD→F  (de 3 por Descomposición)

06) A→EC  (con 1 y 2 por Transitividad)

07) A→B   (de 1 por Descomposición)
08) A→C   (de 1 por Descomposición)

09) AD→CD (de 8 por Aumento)

10) AD→C  (de 9 por Descomposición)
11) AD→D  (de 9 por Descomposición)

12) AD→EF (de 3 y 9 por Transitividad)

13) AD→E  (de 12 por Descomposición)
14) AD→F  (de 12 por Descomposición)

...

¿Hay más?
```

C+ = {A→BC, B→E, CD→EF, CD→E, CD→F, A→EC, A→B, A→C, AD→CD, AD→C, AD→D, AD→EF, AD→E, AD→F}

Cierre Transitivo de un Conjunto de DF

UNIVERSIDAD DE LOS ANDES VENEZUELA

01) A →BC
02) B →E
03) CD→EF

C+ = {A→BC, B→E, CD→EF, CD→E, CD→F, A→EC, A→B, A→C, AD→CD, AD→C, AD→D, AD→EF, AD→E, AD→F}

¿Son todas "Dependencias Funcionales Elementales"? ¿Hay alguna dependencia que esté de más?

Cierre Transitivo de un Conjunto de DF

UNIVERSIDAD DE LOS ANDES VENEZUELA

C,D,F->E;
C,D->E;
C,D->F;
C,D,E->F;
B,F->E;
B->E;
B,D,F->E;
B,D->E;
B,C,F->E;
B,C->E;
B,C,D,F->E;
B,C,D->E;
B,C,D->F;
B,C,D,E->F;
A,C,D,E,F->B;
A,D,E,F->B;
A,D,E,F->C;

A,E,F->B;
A,E,F->C;
A,F->B;
A,F->C;
A,F->E;
A->B;
A->C;
A->E;
A,E->B;
A,E->C;
A,D,F->B;
A,D,F->C;
A,D,F->E;
A,D->B;
A,D->C;
A,D->E;
A,D->F;
A,D,E->B;

A,D,E->C;
A,D,E->F;
A,C,E,F->B;
A,C,F->B;
A,C,F->E;
A,C->B;
A,C->E;
A,C,E->B;
A,C,D,F->B;
A,C,D,F->E;
A,C,D->B;
A,C,D->E;
A,C,D->F;
A,C,D,E->B;
A,C,D,E->F;
A,B,D,E,F->C;
A,B,E,F->C;

A,B,F->C;
A,B,F->E;
A,B->C;
A,B->E;
A,B,E->C;
A,B,D,F->C;
A,B,D,F->E;
A,B,D->C;
A,B,D->E;
A,B,D->F;
A,B,D,E->C;
A,B,D,E->F;
A,B,C,F->E;
A,B,C->E;
A,B,C,D,F->E;
A,B,C,D->E;
A,B,C,D->F;
A,B,C,D,E->F;

¿Existirá una manera más fácil y sistemática de calcular C+? (...más adelante...)

Cierre Transitivo de un Conjunto de DF (Utilidad)

UNIVERSIDAD DE LOS ANDES VENEZUELA

- Dos conjuntos de dependencias funcionales F y G son equivalentes si ambos tienen el mismo cierre transitivo, es decir, si: $F^+ = G^+$ (O si partiendo de F y G podemos obtener un mismo conjunto de dependencias funcionales H, o si todas las dependencias de F y todas las de G están en F/G^+)

Cierre de un Conjunto de Atributos

UNIVERSIDAD DE LOS ANDES VENEZUELA

- Dado un conjunto C de dependencias funcionales y un conjunto X de atributos, es posible determinar ***el conjunto X^+ de todos los atributos que dependen funcionalmente de X***. El conjunto X^+ se conoce como el cierre de X bajo C.

resultado := X
mientras (cambios en *resultado*) **hacer**
  para cada dependencia funcional $A \rightarrow B$ **en** C **hacer**
    si $A \subseteq$ *resultado* **entonces** *resultado* := *resultado* $\cup$ B

Cierre de un Conjunto de Atributos

UNIVERSIDAD DE LOS ANDES VENEZUELA

01) $A \rightarrow BC$
02) $B \rightarrow E$
03) $CD \rightarrow EF$

$A^+ = A$
$A^+ = A\ BC$
$A^+ = A\ BC\ E$

$B^+ = B$
$B^+ = B\ E$

$CD^+ = CD$
$CD^+ = CD\ EF$

Es decir, el cálculo de X^+ se hace fundamentalmente por Transitividad

¿Cuál el cierre de C? ¿y el de D?

Clave de una Relación

UNIVERSIDAD DE LOS ANDES VENEZUELA

- Una ***clave candidata*** es el conjunto X de atributos de una relación $R(A_1, A_2, ..., A_n)$ tal que:
  - $X \rightarrow A_1, A_2, ..., A_n$
  - No existe un $Y \subseteq X$ tal que $Y \rightarrow A_1, A_2, ..., A_n$, De ser así, entonces X sería una ***super-clave***, y Y ***podría ser (o no) una clave candidata***

Clave de una Relación

UNIVERSIDAD DE LOS ANDES VENEZUELA

- Es posible que existan varios atributos que cumplan con esta definición dentro de una misma relación (***claves candidatas***)
- La clave primaria dentro de una relación se denota subrayando los atributos que forman parte del conjunto X

Clave de una Relación

UNIVERSIDAD DE LOS ANDES VENEZUELA

R (A, B, C, D, E, F)

+

01) A →BC
02) B →E
03) CD→EF

Recordando que:

$A^+ = A\ BC\ E$

$B^+ = B\ E$

$CD^+ = CD\ EF$

¿Cuál es la clave de la relación R?

revisar concepto anterior

... sigue ...

Clave de una Relación

UNIVERSIDAD DE LOS ANDES VENEZUELA

R (A, B, C, D, E, F)

+

01) A →BC
02) B →E
03) CD→EF

Recordando que:

$A^+ = A\ BC\ E$

$B^+ = B\ E$

$CD^+ = CD\ EF$

¿A determina todos los atributos de R?

¿y B? ¿y C? ¿y D?

¿y CD?

¿Entonces cuál es la clave?

...sigue...

Clave de una Relación

UNIVERSIDAD DE LOS ANDES VENEZUELA

R (A, B, C, D, E, F)

+

01) A →BC
02) B →E
03) CD→EF

Recordando que:

$A^+ = A\ BC\ E$

$B^+ = B\ E$

$CD^+ = CD\ EF$

¿AB determina todos los atributos de R?

¿y ABCD?

¿y ABC?

¿y ACD?

Otro ejemplo...

Clave de una Relación

R (A, B, C, D, E) + F =

01) B →CD
02) D →E
03) B →A
04) E →C
05) AD →B

¿Está B→ E en F?

¿Es D una clave para R?

¿Es AD una clave para R?

¿Es ADE una super-clave o una clave candidata?

¿Es AD una super-clave o una clave candidata?

Cierre Transitivo de un Conjunto de DF
(Usando el cierre de un Conjunto de Atributos)

UNIVERSIDAD DE LOS ANDES VENEZUELA

R(A,B,C,D,E,F)

01) A →BC
02) B →E
03) CD→EF

Se calcula el cierre de todos los posibles conjuntos de atributos, Ej: $\{A\}^+$, $\{B\}^+$, $\{C\}^+$, $\{D\}^+$, $\{E\}^+$, $\{F\}^+$, $\{AB\}^+$, $\{AC\}^+$, $\{AD\}^+$, $\{AE\}^+$, $\{AF\}^+$, $\{BC\}^+$, $\{BD\}^+$, $\{BE\}^+$, $\{BF\}^+$, $\{CD\}^+$, $\{CE\}^+$, $\{CF\}^+$, $\{DE\}^+$, $\{DF\}^+$, $\{EF\}^+$, $\{ABC\}^+$, $\{ABD\}^+$, $\{ABE\}^+$, $\{ABF\}^+$, ..., ..., ..., ... hasta llegar a $\{ABCDEF\}^+$

Del resultado, y por descomposición se consideran todas las DFEs no triviales que se puedan obtener de los cierres anteriormente calculados, por ejemplo:

$\{ABF\}^+=\{ABCEF\}$ (Lo que representa en realidad ABF->ABCEF y se puede descomponer en:

ABF->A, **ABF->B**, **ABF->C**, **ABF->E**, **ABF->F**

Las resaltadas en rojo son triviales, por lo tanto se descartan, mientras que las marcadas en verde son no triviales y se añaden al cierre del conjunto de DFs

Descomposición de una Relación
(Propiedades Deseables)

UNIVERSIDAD DE LOS ANDES VENEZUELA

- Toda relación R tiene ***al menos una descomposición en 3FN*** tal que:
  - La descomposición sea sin pérdida (de producto sin pérdida)
  - La descomposición preserve (conserva) las DF

Descomposición de una Relación
(Descomposición de producto sin pérdida)

- Sea R una relación y F un conjunto de DF en R. La descomposición $A=\{R_1, R_2\}$ es una descomposición de producto sin pérdida ***si al menos una*** de las siguientes dependencias están en F^+:

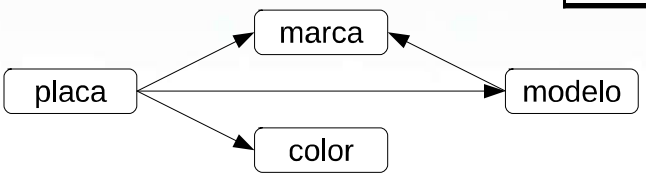
$R_1 \cap R_2 \rightarrow R_1$
$R_1 \cap R_2 \rightarrow R_2$

¿Cómo? ¿Qué tendrá esto que ver con las claves primarias y foráneas?

¿Se pierden dependencias?

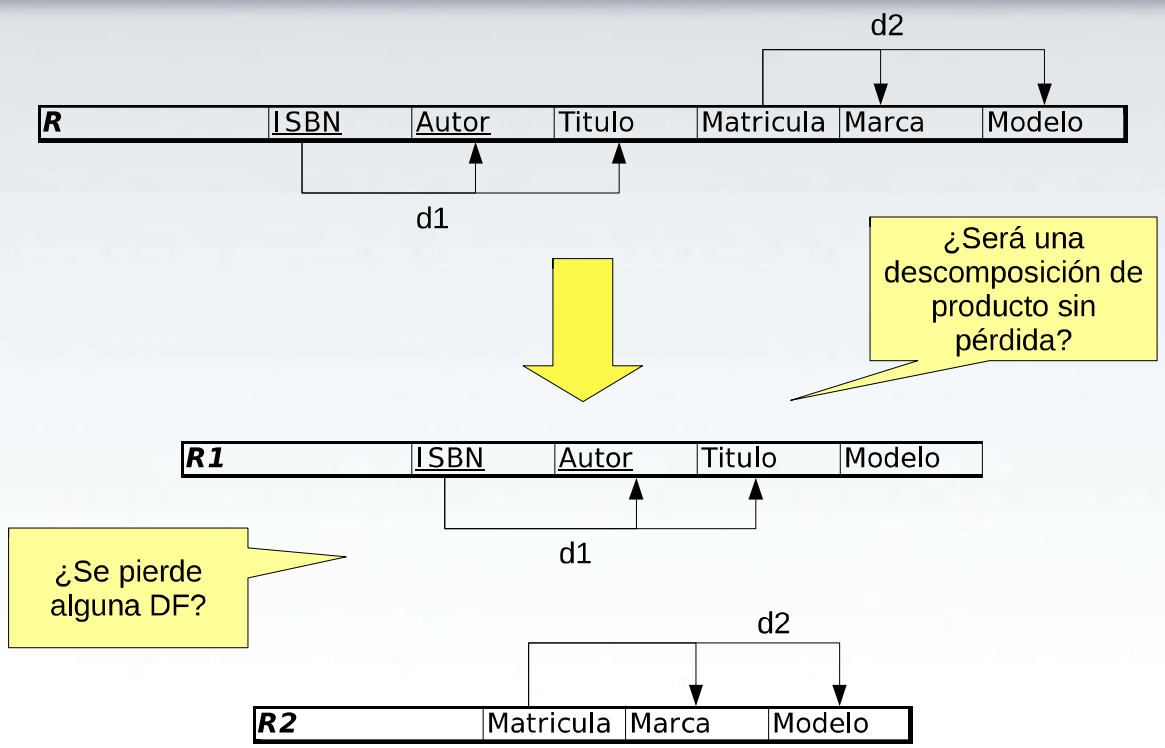
R	placa	***marca***
	MBO34L	Ford
	LDA75K	Toyota
	ADA89A	Fiat
	LBF78G	Toyota
	XSA67D	Ford

Q	***marca***	modelo	color
	Ford	Ka	verde
	Toyota	corollaXL	blanco
	Fiat	siena	gris
	Ford	Ka	rojo



Carro	placa	marca	modelo	color
	MBO34L	Ford	Ka	verde
	MBO34L	Ford	Ka	rojo
	LDA75K	Toyota	corollaXL	blanco
	ADA89A	Fiat	siena	gris
	LBF78G	Toyota	corollaXL	blanco
	XSA67D	Ford	Ka	verde
	XSA67D	Ford	Ka	rojo

Descomposición sin pérdida
(Descomposición de producto sin pérdida)



Descomposición de una Relación
(Descomposición de producto sin pérdida)

- $R_1 \cap R_2$ = Modelo
- Es decir que se debe cumplir que:
- **Modelo→ISBN, Autor, Titulo (1)**

  o bien que...
- **Modelo→Matricula, Marca (2)**
- Es decir, que las DFEs deducibles de (1) o de (2) por descomposición se encuentren en el cierre transitivo de las DFs asociadas a la relación R. ¿Cómo podemos calcular eso en base a d1 y d2 de la transparencia anterior?

Descomposición de una Relación
(Descomposición de producto sin pérdida)

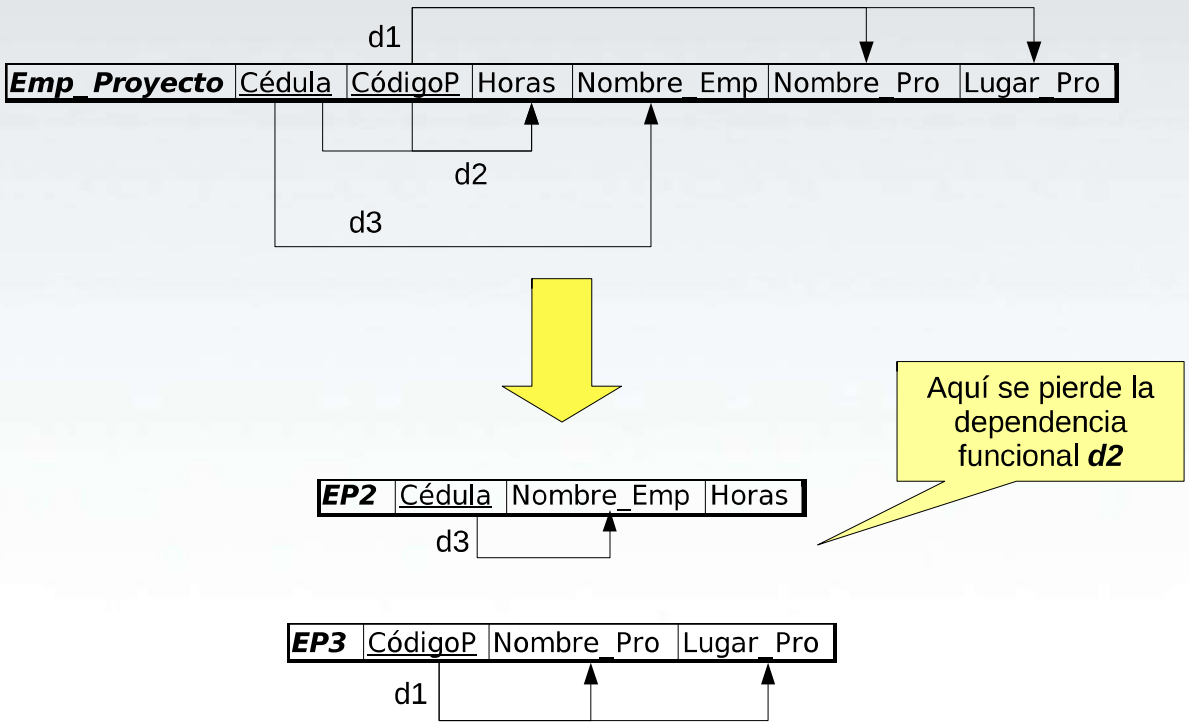
- Se puede calcular $\{Modelo\}^+$ en base a d1 y d2:
- **$\{Modelo\}^+=\{Modelo\}$**
- De modo que el cierre de las DFs asociado a R no contiene ninguna de las Dfs:
  - **Modelo→ISBN, Autor, Titulo (1)**
  - **Modelo→Matricula, Marca (2)**
- Por lo tanto la descomposición **NO ES** de producto sin pérdida.

Descomposición de una Relación
(Descomposición que preserva las DFs)

- Sea R una relación y F un conjunto de DF en R. La descomposición $A=\{R_1, R_2, ... R_n\}$ preserva las dependencias funcionales si el conjunto de DFs G que se obtiene de hacer $R_1 \cup R_2 \cup ... \cup R_n$ (es decir, que aporta cada relación R_i) es equivalente a G.
- Nota: Puede ser que a simple vista $F \neq G$, pero se cumple que $F^+ = G^+$.

Descomposición sin pérdida (que Preserva las DFs)

UNIVERSIDAD DE LOS ANDES VENEZUELA



Descomposición sin pérdida (que Preserva las DFs)

UNIVERSIDAD DE LOS ANDES VENEZUELA

Cierre de {d1, d2, d3}

CodigoP,Lugar_Pro>Nompre_Pro;
CodigoP>Nompre_Pro;
CodigoP>Lugar_Pro;
CodigoP,Nompre_Pro>Lugar_Pro;
CodigoP,Nombre_Emp,Lugar_Pro>Nompre_Pro;
CodigoP,Nombre_Emp>Nompre_Pro;
CodigoP,Nombre_Emp>Lugar_Pro;
CodigoP,Nombre_Emp,Nompre_Pro>Lugar_Pro;
CodigoP,Horas,Lugar_Pro>Nompre_Pro;
CodigoP,Horas>Nompre_Pro;
CodigoP,Horas>Lugar_Pro;
CodigoP,Horas,Nompre_Pro>Lugar_Pro;
CodigoP,Nombre_Emp,Lugar_Pro>Nompre_Pro;
CodigoP,Nombre_Emp>Nompre_Pro;
CodigoP,Nombre_Emp>Lugar_Pro;
CodigoP,Nombre_Emp,Nompre_Pro>Lugar_Pro;
CodigoP,Nombre_Emp,Horas,Lugar_Pro>Nompre_Pro;
CodigoP,Nombre_Emp,Horas>Nompre_Pro;
CodigoP,Nombre_Emp,Horas>Lugar_Pro;
CodigoP,Nombre_Emp,Horas,Nompre_Pro>Lugar_Pro;
Cedula,Horas,Nompre_Pro,Lugar_Pro>Nombre_Emp;
Cedula,Nompre_Pro,Lugar_Pro>Nombre_Emp;
Cedula,Lugar_Pro>Nombre_Emp;
Cedula>Nombre_Emp;
Cedula,Nompre_Pro>Nombre_Emp;
Cedula,Horas,Lugar_Pro>Nombre_Emp;
Cedula,Horas>Nombre_Emp;
Cedula,Horas,Nompre_Pro>Nombre_Emp;
Cedula,CodigoP,Horas,Nompre_Pro,Lugar_Pro>Nombre_E
Cedula,CodigoP,Nompre_Pro,Lugar_Pro>Nombre_Emp;
Cedula,CodigoP,Nompre_Pro,Lugar_Pro>Horas;
Cedula,CodigoP,Lugar_Pro>Nombre_Emp;

CodigoP,Lugar_Pro>Nompre_Pro;
CodigoP>Nompre_Pro;
CodigoP>Lugar_Pro;
CodigoP,Nompre_Pro>Lugar_Pro;
CodigoP,Nombre_Emp,Lugar_Pro>Nompre_Pro;
CodigoP,Nombre_Emp>Nompre_Pro;
CodigoP,Nombre_Emp>Lugar_Pro;
CodigoP,Nombre_Emp,Nompre_Pro>Lugar_Pro;
Cedula,Nompre_Pro,Lugar_Pro>Nombre_Emp;
Cedula,Lugar_Pro>Nombre_Emp;
Cedula>Nombre_Emp;
Cedula,Nompre_Pro>Nombre_Emp;
Cedula,CodigoP,Nompre_Pro,Lugar_Pro>Nombre_Emp;
Cedula,CodigoP,Lugar_Pro>Nombre_Emp;
Cedula,CodigoP,Lugar_Pro>Nompre_Pro;
Cedula,CodigoP>Nombre_Emp;
Cedula,CodigoP>Nompre_Pro;
Cedula,CodigoP>Lugar_Pro;
Cedula,CodigoP,Nompre_Pro>Nombre_Emp;
Cedula,CodigoP,Nompre_Pro>Lugar_Pro;
Cedula,CodigoP,Nombre_Emp,Lugar_Pro>Nompre_Pro;
Cedula,CodigoP,Nombre_Emp>Nompre_Pro;
Cedula,CodigoP,Nombre_Emp>Lugar_Pro;
Cedula,CodigoP,Nombre_Emp,Nompre_Pro>Lugar_Pro;

Cierre de {d1, d3}

CodigoP,Lugar_Pro>Nompre_Pro;
CodigoP>Nompre_Pro;
CodigoP>Lugar_Pro;
CodigoP,Nompre_Pro>Lugar_Pro;
CodigoP,Nombre_Emp,Lugar_Pro>Nompre_Pro;
CodigoP,Nombre_Emp>Nompre_Pro;
CodigoP,Nombre_Emp>Lugar_Pro;
CodigoP,Nombre_Emp,Nompre_Pro>Lugar_Pro;
Cedula,Nompre_Pro,Lugar_Pro>Nombre_Emp;
Cedula,Lugar_Pro>Nombre_Emp;
Cedula>Nombre_Emp;
Cedula,Nompre_Pro>Nombre_Emp;
Cedula,CodigoP,Nompre_Pro,Lugar_Pro>Nombre_Emp;
Cedula,CodigoP,Nombre_Emp;
Cedula,CodigoP,Lugar_Pro>Nombre_Emp;
Cedula,CodigoP,Lugar_Pro>Nompre_Pro;
Cedula,CodigoP>Nombre_Emp;
Cedula,CodigoP>Nompre_Pro;
Cedula,CodigoP>Lugar_Pro;
Cedula,CodigoP,Nompre_Pro>Nombre_Emp;
Cedula,CodigoP,Nompre_Pro>Lugar_Pro;
Cedula,CodigoP,Nombre_Emp,Lugar_Pro>Nompre_Pro;
Cedula,CodigoP,Nombre_Emp>Nompre_Pro;
Cedula,CodigoP,Nombre_Emp>Lugar_Pro;
Cedula,CodigoP,Nombre_Emp,Nompre_Pro>Lugar_Pro;

Es evidente por los cierres que ambos conjuntos de DFs NO SON equivalentes, por lo tanto se pierden DFs en la descomposición

Descomposición sin pérdida (que Preserva las DFs)

UNIVERSIDAD DE LOS ANDES VENEZUELA

- La descomposición $A=\{R_1, R_2, ..., R_n\}$ de una relación R preserva las DF de R, si ***C^+ de R es el mismo que la de la unión de las DF de $A=\{R_1, R_2, ..., R_n\}$***

Cobertura Mínima

UNIVERSIDAD DE LOS ANDES VENEZUELA

- Es el conjunto C de DFE asociado a un conjunto de atributos que verifican las propiedades siguientes:
  - Ninguna DF es redundante en C, es decir para toda DF denotada **f** en C, C – **f** no es equivalente a C.
  - Toda DFE de los atributos está dentro de C^+.
- Ejemplo: C = {placa $\rightarrow$ modelo, placa $\rightarrow$ color, modelo $\rightarrow$ marca}
- C es esencial para la descomposición sin pérdida (existen algoritmos de descomposición que dependen de esto).

Formas Normales

1ra Forma Normal (1FN)

- Se define para prohibir los atributos multivaluados, los atributos compuestos y sus combinaciones.
- Establece que los dominios de los atributos deben incluir sólo valores atómicos (simples e indivisibles) y que el valor de cualquier atributo en una tupla debe ser un valor individual proveniente del dominio de ese atributo.
- La primera forma normal se considera actualmente ***parte de la definición*** de relación.

1ra Forma Normal (1FN)

Estudiante	Cédula	Nombre	Apellido	Curso
	9.644.667	Pedro	Pérez	BD, IS
	10.133.212	Gabriel	Mendoza	PRI, CA10
	11.332.334	Luis	Gonzales	PRII, SR10, EST1
	14.126.112	Gilberto	Zapata	BD, IA

Atributo Multivaluado

No está en 1FN
(No es una relación)
Hay grupos repetitivos

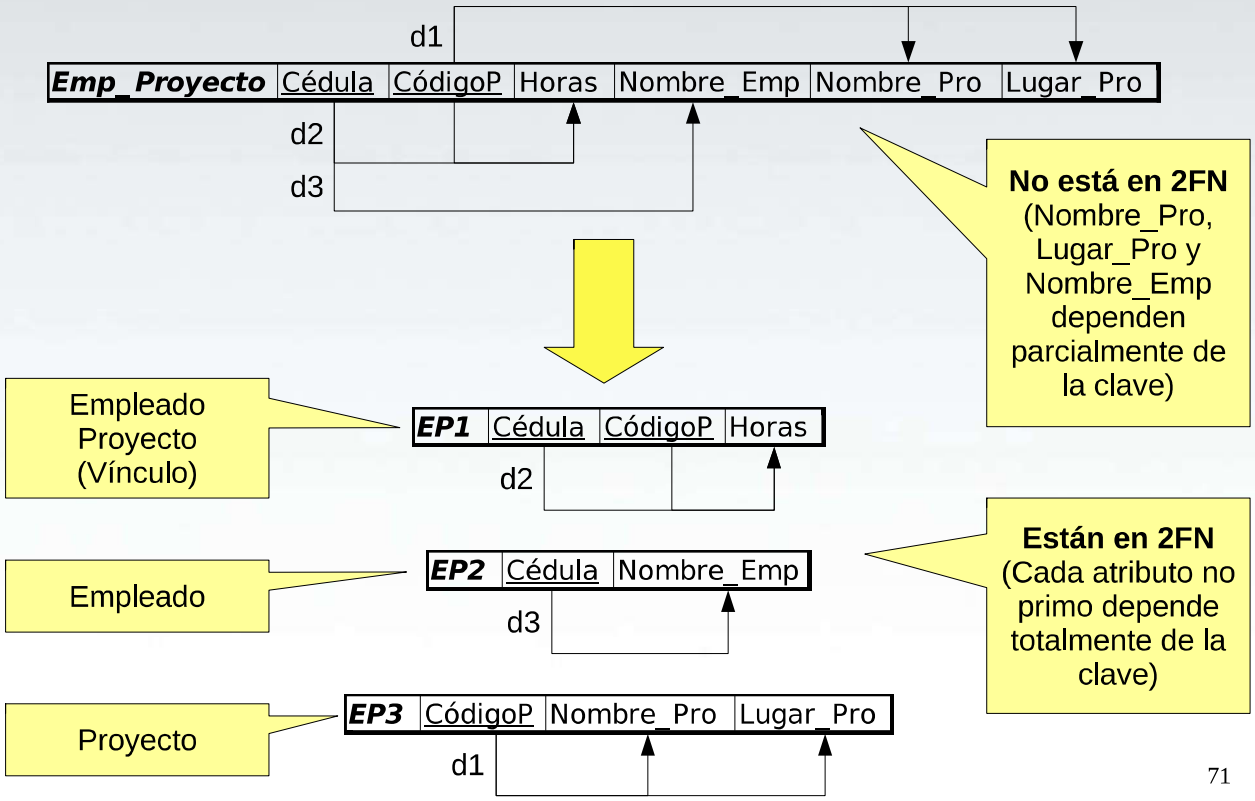
Estudiante	Cédula	Nombre	Apellido	Curso
	9.644.667	Pedro	Pérez	BD
	9.644.667	Pedro	Pérez	IS
	10.133.212	Gabriel	Mendoza	PRI
	10.133.212	Gabriel	Mendoza	CA10
	11.332.334	Luis	Gonzales	PRII
	11.332.334	Luis	Gonzales	SR10
	11.332.334	Luis	Gonzales	EST1
	14.126.112	Gilberto	Zapata	BD
	14.126.112	Gilberto	Zapata	IA

Está en 1FN
(Sin embargo, hay información repetida)

2da Forma Normal (2FN)

- Se basa en el concepto de **dependencia funcional total**, una dependencia funcional $X \rightarrow Y$ es total si la eliminación de cualquier atributo A de X rompe la dependencia.
- Una relación está en 2FN si está en 1FN y todo atributo que no sea clave depende de forma total de la clave.
- La 2FN permite eliminar las ***redundancias*** para que ningún atributo sea determinado ***sólo por una parte*** de una clave.

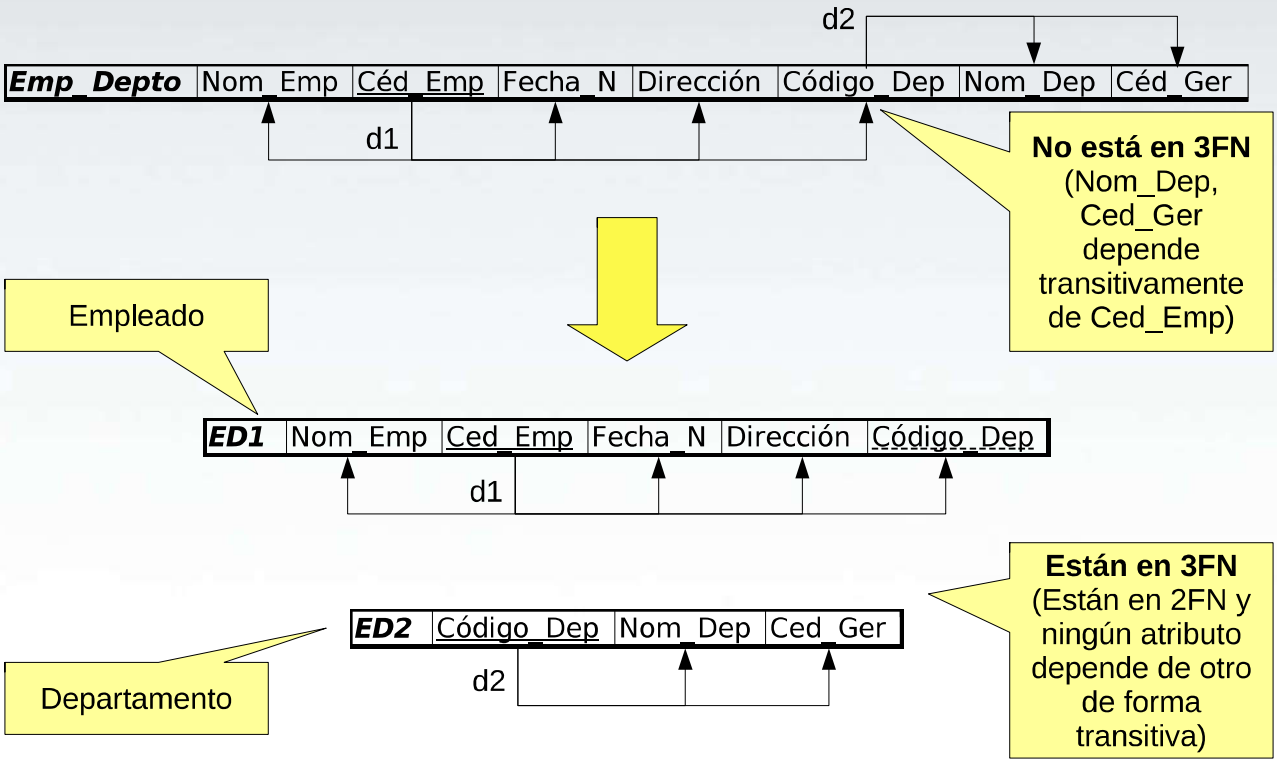
2da Forma Normal (2FN)



3ra Forma Normal (3FN)

- Se basa en el concepto de **dependencia transitiva**. Una dependencia funcional $X \rightarrow Z$ es transitiva si existen dos dependencias funcionales $X \rightarrow Y$ y $Y \rightarrow Z$ de la que puede deducirse por transitividad $X \rightarrow Z$.
- Una relación está en 3FN si está en 2FN y todo atributo que no sea clave no depende de un atributo que no sea clave.
- La 3FN permite eliminar las redundancias para que ningún atributo dependa de otro de forma transitiva.

3ra Forma Normal (3FN)



Algoritmo de Descomposición en 3FN

Propuesto por Bernstein en 1976, se basa en el siguiente principio:

- Se construye la cobertura mínima **C** y se editan los atributos aislados, considerándolos como claves.
- Se busca el conjunto más grande X de atributos que determinen a otros A_1, A_2, ..., A_n con $n \geq 1$ y como salida se genera la relación $R(X, A_1, A_2, ..., A_n)$.
- Las DFE utilizadas en la formación de esa relación se eliminan de C y todos los atributos aislados que no están en las DFE que quedaron en C.

Algoritmo de Descomposición en 3FN

UNIVERSIDAD DE LOS ANDES VENEZUELA

Procedimiento *Normalizar_3FN()*

C := Cobertura mínima de las DFE

Att := Atributos aislados que pertenecen a C

reducir(C, Att) /* Ver siguiente lámina */

Formar una relación con los atributos restantes en Att (Si los hay)

Fin

Algoritmo de Descomposición en 3FN

UNIVERSIDAD DE LOS ANDES VENEZUELA

Procedimiento *Reducir (C, Att)*

Repita mientras que una DFE no incluya todos los atributos o C esté vacío

Buscar el conjunto más grande de atributos X tal que $X \rightarrow A_1$, $X \rightarrow A_2$, ..., $X \rightarrow A_k$

Formar la relación $R(X, A_1, A_2, ..., A_k)$

Eliminar de C las DFE utilizadas en R

Eliminar de Att los atributos que ya no pertenezcan a C

Juramento Final!

Un esquema normalizado hasta 3FN debe cumplir con el juramento siguiente:

Forma Normal Boyce-Codd (FNBC)

- ***Navathe*** dice que un esquema está en 3FN si siempre que se tiene una dependencia funcional X->A, se cumple que:
  - X es una superclave, o bien:
  - *A es un atributo primo (**Parte de la Clave**) de R.*
- La FNBC ***elimina la segunda condición***, de manera que para que una relación esté en FNBC se tiene que cumplir que X sea una superclave de R y no basta con que A sea un atributo primo de R.

Forma Normal Boyce-Codd (FNBC)

UNIVERSIDAD DE LOS ANDES VENEZUELA

Una relación está en FNBC si y sólo si las solas DFE son aquellas dentro de las cuales una clave determina un atributo

Si cada profesor dicta una única materia, entonces se cumple:

cédula_est, codigo_mat->cedula_prof

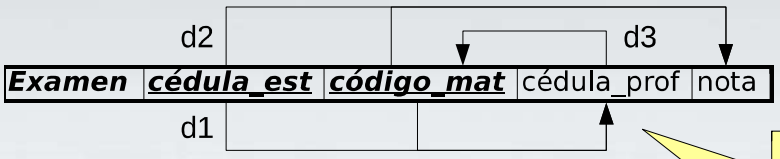
cédula_est, codigo_mat->nota

cedula_prof->codigo_mat

Entonces la relación está en 3FN pero no en FNBC

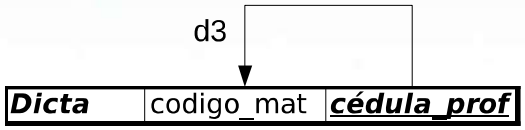
Forma Normal Boyce-Codd (FNBC)

UNIVERSIDAD DE LOS ANDES VENEZUELA



No está en FNBC (en código_mat depende de cédula_prof que no es primo)

¿Qué ocurrió con d1?



es sin pérdida, pero no preserva las DFE, que pueden obtenerse en conjunto por medio de una reunión o producto

Forma Normal Boyce-Codd (FNBC)

The definition of BCNF addresses certain (rather unlikely) situations which 3NF does not handle. The characteristics of a relation which distinguish 3NF from BCNF are given below. Since it is so unlikely that a relation would have these characteristics, in practical real-life design it is usually the case that relations in 3NF are also in BCNF. Thus many authors make a "fuzzy" distinction between 3NF and BCNF when it comes to giving advice on "how far" to normalize a design. Since relations in 3NF but not in BCNF are slightly unusual, it is a bit more difficult to come up with meaningful examples.

http://www.imsmca.in/forums/computer-softwares/normal-form-definitions-examples-t17.html

Conste que no lo digo yo solamente...

¿Dependencias Multivaluadas / 4FN?

Imagine la siguiente relación:

estudiante	cedula	asignatura	deporte
	12.324.334	PR1, CA10, FS11	natación, fútbol
	13.455.322	BD, IS, LC	tenis, esgrima
	15.443.767	BD, RD, SO	tenis, fútbol

¿Está en primera forma normal?
¿Por qué?

¿Dependencias Multivaluadas / 4FN?

UNIVERSIDAD DE LOS ANDES VENEZUELA

¿Está en primera forma normal?
¿Tiene atributos multivaluados?

estudiante	***cedula***	asignaturas	deportes
	12.324.334	PR1, CA10, FS11	natación, fútbol
	13.455.322	BD, IS, LC	tenis, esgrima
	15.443.767	BD, RD, SO	tenis, fútbol

Un estudiante puede cursar varias asignaturas y preferir varios deportes.

El atributo asignaturas es independiente del atributo deportes, sólo están correlacionados por medio del atributo cédula.

¿Dependencias Multivaluadas / 4FN?

UNIVERSIDAD DE LOS ANDES VENEZUELA

Si resolvemos el atributo multivaluado asignatura:

estudiante	cedula	asignatura	deporte
	12.324.334	PR1	natación, fútbol
	12.324.334	CA10	natación, fútbol
	12.324.334	FS11	natación, fútbol
	13.455.322	BD	tenis, esgrima
	13.455.322	IS	tenis, esgrima
	13.455.322	LC	tenis, esgrima
	15.443.767	BD	tenis, fútbol
	15.443.767	RD	tenis, fútbol

¿Está en primera forma normal?
¿Por qué?

Sigamos...

¿Dependencias Multivaluadas / 4FN?

UNIVERSIDAD DE LOS ANDES VENEZUELA

¿Y ahora, está en 1ra forma normal?

¿Existe alguna dependencia entre asignatura y deporte?

En realidad, lo que ocurre es que existe una *dependencia multivaluada* (DMV) entre ***cedula ->> asignatura*** lo que implica además la existencia de una DMV entre ***cedula ->> deporte***

estudiante	cedula	asignatura	deporte
	12.324.334	PR1	***natación***
	12.324.334	PR1	fútbol
	12.324.334	CA10	***natación***
	12.324.334	CA10	fútbol
	12.324.334	FS11	***natación***
	12.324.334	FS11	fútbol
	13.455.322	BD	***tenis***
	13.455.322	BD	esgrima
	13.455.322	IS	***tenis***
	13.455.322	IS	esgrima
	13.455.322	LC	***tenis***
	13.455.322	LC	esgrima
	15.443.767	BD	**tenis**
	15.443.767	BD	fútbol
	15.443.767	RD	**tenis**
	15.443.767	RD	fútbol
	15.443.767	SO	**tenis**
	15.443.767	SO	fútbol

¿Qué anomalías se producen?
¿Cómo resolvemos el problema?

Dependencias Multivaluadas

UNIVERSIDAD DE LOS ANDES VENEZUELA

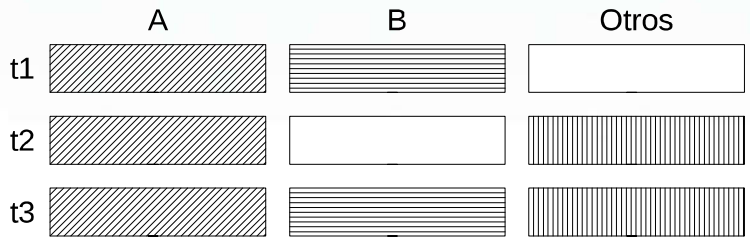
- Sea $R(A_1, A_2, ..., A_n)$ y X e Y dos subconjuntos de atributos de $\{A_1, A_2, ..., A_n\}$. Se dice que X ->> Y, si dados los valores de X hay un conjunto de valores Y asociados y este conjunto es ***independiente*** de otros atributos Z = R – X – Y de R.
- Las DM caracterizan la independencia entre Y y Z correlacionadas por X

Dependencias Multivaluadas

UNIVERSIDAD DE LOS ANDES VENEZUELA

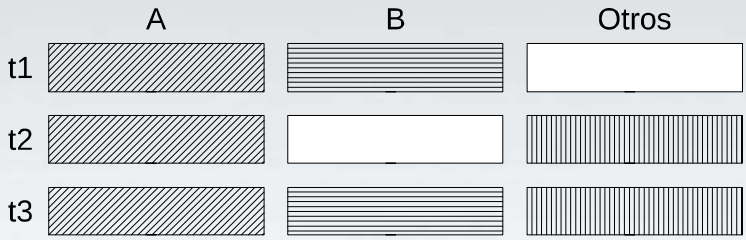
Existe una dependencia multivaluada (DMV) A ->> B si para cada par de tuplas t1 y t2 de la relación R que coinciden en todos los valores de A se puede encontrar una tupla t3 tal que:

- Coincida con t1 y t2 en A.
- Coincida con t1 en B.
- Coincida con t2 en todos los atributos de R que no pertenecen ni a A ni a B.



Dependencias Multivaluadas

UNIVERSIDAD DE LOS ANDES VENEZUELA



Razonando un poco, se puede llegar a la conclusión de que si A-->>B luego A-->>Otros

estudiante	cedula	asignatura	deporte
t1	12.324.334	PR1	***natación***
t3	12.324.334	PR1	fútbol
t4	12.324.334	CA10	***natación***
t2	12.324.334	CA10	fútbol
	12.324.334	FS11	***natación***
	12.324.334	FS11	fútbol
	...	...	...

O bien:
t2
t4
t3
t1

4ta Forma Normal (4FN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

- Está en 3FN y FNBC
- Si y sólo si las solas DMV (Elementales*) son aquellas donde ***una clave determina un atributo (a un bloque de atributos),*** es decir, no dos al mismo tiempo.

*Recordar Dependencias Funcionales Elementales (transparencia 22), pero también:

- Y no es vacío y es disjunto de X
- R no contiene otra DMV del tipo X' ->> Y' tal que X' $\subset$ X y Y' $\subset$ Y

4ta Forma Normal (4FN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

Otra forma de verlo: Una R está en 4FN si y solo si las solas DME son aquellas donde una clave determina un atributo (Es decir, no dos al mismo tiempo). Una R en 4FN está en 3FN y en FNBC.

4ta Forma Normal (4FN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

estudiante	***cedula***	***asignatura***
	12.324.334	PR1
	12.324.334	CA10
	12.324.334	FS11
	13.455.322	BD
	13.455.322	IS
	13.455.322	LC
	15.443.767	BD
	15.443.767	RD
	15.443.767	SO

estudiante	***cedula***	***deporte***
	12.324.334	natación
	12.324.334	fútbol
	13.455.322	tenis
	13.455.322	esgrima
	15.443.767	tenis
	15.443.767	fútbol

Esta descomposición no repite información y mantiene las DMV

Todas las dependencias funcionales:

cédula ->> asignatura

cédula ->> deporte

Son tales que cédula (atributo parte de la clave)

determina sólo a un atributo

4ta Forma Normal (4FN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

Si tenemos el siguiente caso:

estudiante	*cedula*	*nombre_e*	*cod_asig**	*desc_asig**	*cod_deporte**	desc_deporte*

d1 d2 d3

Y además:
cédula ->> cod_asig

Eso implica que:
cédula ->> cod_deporte

¿Cómo se normaliza?

* Representa atributos multivaluados

4ta Forma Normal (4FN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

1FN: Se eliminan atributos multivaluados:

estudiante	*cedula*	*nombre_e*	*cod_asig*	*desc_asig*	*cod_deporte*	desc_deporte

d1 d2 d3

Recordar que: ***cédula ->> cod_asig*** *y* ***cédula ->> cod_deporte***

2FN: Dependencia total de la clave:

estudiante	*cedula*	*nombre_e*	*cod_asig*	*cod_deporte*

d1

deporte	cod_deporte	desc_deporte

d2

materia	cod_asig	*desc_asig*

d3

¿3FN / FNBC? ¿4FN?

4ta Forma Normal (4FN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

deporte	cod_deporte	desc_deporte

d2

materia	cod_asig	*desc_asig*

d3

estudiante	*cedula*	*nombre_e*

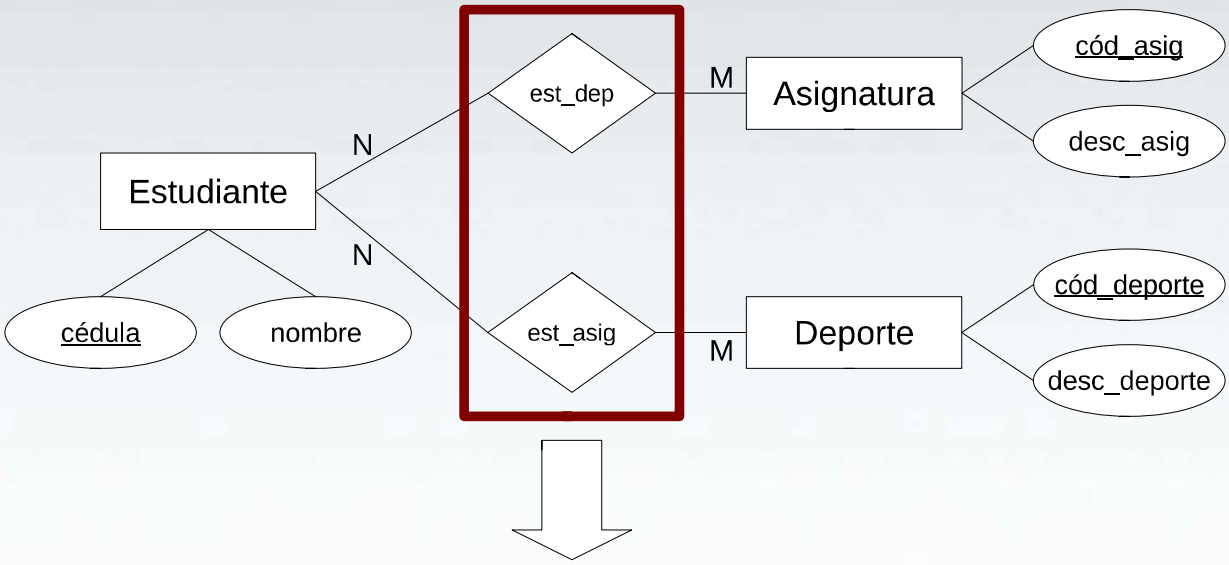
d1

est_dep	*cedula*	*cod_deporte*

est_dep	*cedula*	*cod_asig*

Finalmente está en 4FN. Las DMV
cedula ->> cod_deporte y cedula ->> cod_asig
están en *est_dep* y *est_asig* respectivamente y son triviales

4ta Forma Normal (4FN)



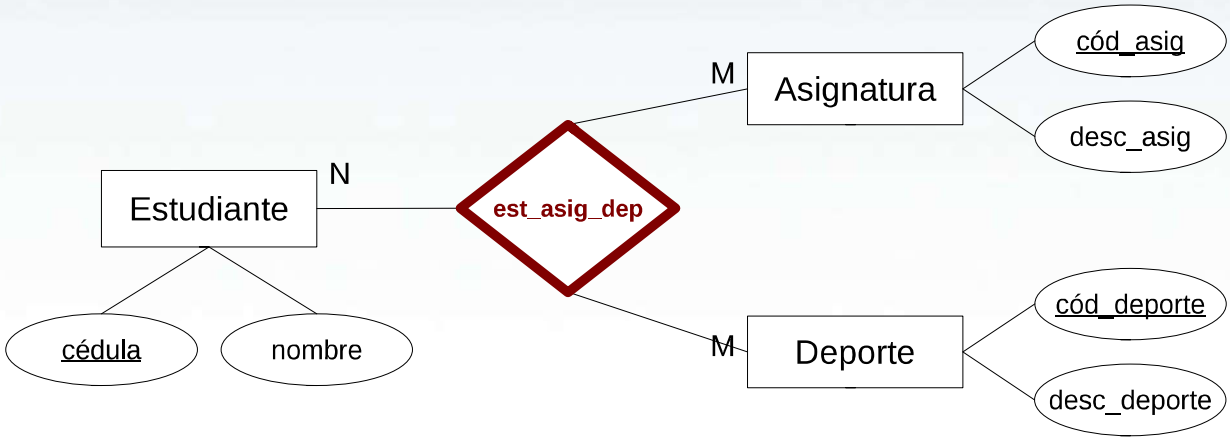
¿Será esto equivalente?
¿Será mejor o peor?

estudiante	cedula	asignatura	deporte

¿Está en 4FN?

4ta Forma Normal (4FN)

Es decir, la 4FN (y 5FN como veremos más adelante) SON LA RAZÓN por la que hay que evitar los vínculos n-arios



5ta Forma Normal (5FN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

Si suponemos el siguiente esquema

Suministro	Proveedor*	Pieza*	Proyecto*
	E1, E4, E6	PI3,PI6	PR2, PR4
	E2, E5	PI1, PI2	PR1, PR3
	E3, E7	PI4, PI5	PR5, PR6

Y asumimos que un Proveedor suministra ciertas Piezas en particular, un Proyecto usa ciertas Piezas, y un Proyecto es suplido por ciertos Proveedores, entonces tenemos las siguientes DMV:

Proveedor ->> Pieza

Pieza ->> Proyecto

Proyecto ->> Proveedor

Y sus respectivas complementarias...

5ta Forma Normal (5FN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

Suministro	Proveedor*	Pieza*	Proyecto*
	E1, E4, E6	PI3,PI6	PR2, PR4
	E2, E5	PI1, PI2	PR1, PR3
	E3, E7	PI4, PI5	PR5, PR6

Suministro	Proveedor*	Pieza*	Proyecto*
	E1	PI3, PI6	PR2, PR4
	E4	PI3, PI6	PR2, PR4
	E6	PI3, PI6	PR2, PR4
	E2	PI1, PI2	PR1, PR3
	E5	PI1, PI2	PR1, PR3
	E3	PI4, PI5	PR5, PR6
	E7	PI4, PI5	PR5, PR6

Suministro	Proveedor*	Pieza*	Proyecto*
	E1	PI3	**PR2, PR4**
	E1	PI6	**PR2, PR4**
	E4	PI3	**PR2, PR4**
	E4	PI6	**PR2, PR4**
	E6	PI3	**PR2, PR4**
	E6	PI6	**PR2, PR4**
	E2	PI1	**PR1, PR3**
	E2	PI2	**PR1, PR3**
	E5	PI1	**PR1, PR3**
	E5	PI2	**PR1, PR3**
	E3	PI4	**PR5, PR6**
	E3	PI5	**PR5, PR6**
	E7	PI4	**PR5, PR6**
	E7	PI5	**PR5, PR6**

Etcétera

5ta Forma Normal (5FN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

Suministro	Proveedor*	Pieza*	Proyecto*
	E1	PI3	PR2
	E1	PI3	PR4
	E1	PI6	PR2
	E1	PI6	PR4
	E4	PI3	PR2
	E4	PI3	PR4
	E4	PI6	PR2
	E4	PI6	PR4
	E6	PI3	PR2
	E6	PI3	PR4
	E6	PI6	PR2
	E6	PI6	PR4
	E2	PI1	PR1
	E2	PI1	PR3
	E2	PI2	PR1
	E2	PI2	PR3
	E5	PI1	PR1
	E5	PI1	PR3
	E5	PI2	PR1
	E5	PI2	PR3
	E3	PI4	PR5
	E3	PI4	PR6
	E3	PI5	PR5
	E3	PI5	PR6
	E7	PI4	PR5
	E7	PI4	PR6
	E7	PI5	PR5
	E7	PI5	PR6

¿Está la relación en 4FN?

¿Es posible descomponer la relación en dos relaciones tales que estén en 4FN?

Aquí es donde aplica la 5FN, que no es más que una generalización de la 4FN

5ta Forma Normal (5FN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

Suministro	Proveedor*	Pieza*	Proyecto*
	E1	PI3	PR2
	E1	PI3	PR4
	E1	PI6	PR2
	E1	PI6	PR4
	E4	PI3	PR2
	E4	PI3	PR4
	E4	PI6	PR2
	E4	PI6	PR4
	E6	PI3	PR2
	E6	PI3	PR4
	E6	PI6	PR2
	E6	PI6	PR4
	E2	PI1	PR1
	E2	PI1	PR3
	E2	PI2	PR1
	E2	PI2	PR3
	E5	PI1	PR1
	E5	PI1	PR3
	E5	PI2	PR1
	E5	PI2	PR3
	E3	PI4	PR5
	E3	PI4	PR6
	E3	PI5	PR5
	E3	PI5	PR6
	E7	PI4	PR5
	E7	PI4	PR6
	E7	PI5	PR5
	E7	PI5	PR6

S_Prov_Pi	Proveedor*	Pieza*
	E1	PI3
	E1	PI6
	E4	PI3
	E4	PI6
	E6	PI3
	E6	PI6
	E2	PI1
	E2	PI2
	E5	PI1
	E5	PI2
	E3	PI4
	E3	PI5
	E7	PI4
	E7	PI5

S_Prov_Proy	Proveedor*	Proyecto*
	E1	PR2
	E1	PR4
	E4	PR2
	E4	PR4
	E6	PR2
	E6	PR4
	E2	PR1
	E2	PR3
	E5	PR1
	E5	PR3
	E3	PR5
	E3	PR6
	E7	PR5
	E7	PR6

S_Pi_Proy	Pieza*	Proyecto*
	PI3	PR2
	PI3	PR4
	PI6	PR2
	PI6	PR4
	PI1	PR1
	PI1	PR3
	PI2	PR1
	PI2	PR3
	PI4	PR5
	PI4	PR6
	PI5	PR5
	PI5	PR6

5ta Forma Normal (5FN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

S_Prov_Pi	Proveedor*	Pieza*
	E1	PI3
	E1	PI6
	E4	PI3
	E4	PI6
	E6	PI3
	E6	PI6
	E2	PI1
	E2	PI2
	E5	PI1
	E5	PI2
	E3	PI4
	E3	PI5
	E7	PI4
	E7	PI5

S_Pi_Proy	Pieza*	Proyecto*
	PI3	PR2
	PI3	PR4
	PI6	PR2
	PI6	PR4
	PI1	PR1
	PI1	PR3
	PI2	PR1
	PI2	PR3
	PI4	PR5
	PI4	PR6
	PI5	PR5
	PI5	PR6

S_Prov_Proy	Proveedor*	Proyecto*
	E1	PR2
	E1	PR4
	E4	PR2
	E4	PR4
	E6	PR2
	E6	PR4
	E2	PR1
	E2	PR3
	E5	PR1
	E5	PR3
	E3	PR5
	E3	PR6
	E7	PR5
	E7	PR6

Resulta que:

S_Prov_Pi |X| S_Pi_Proy != Suministro

S_Prov_Pi |X| S_Prov_Proy != Suministro

S_Pi_Proy |X| S_Prov_Proy != Suministro

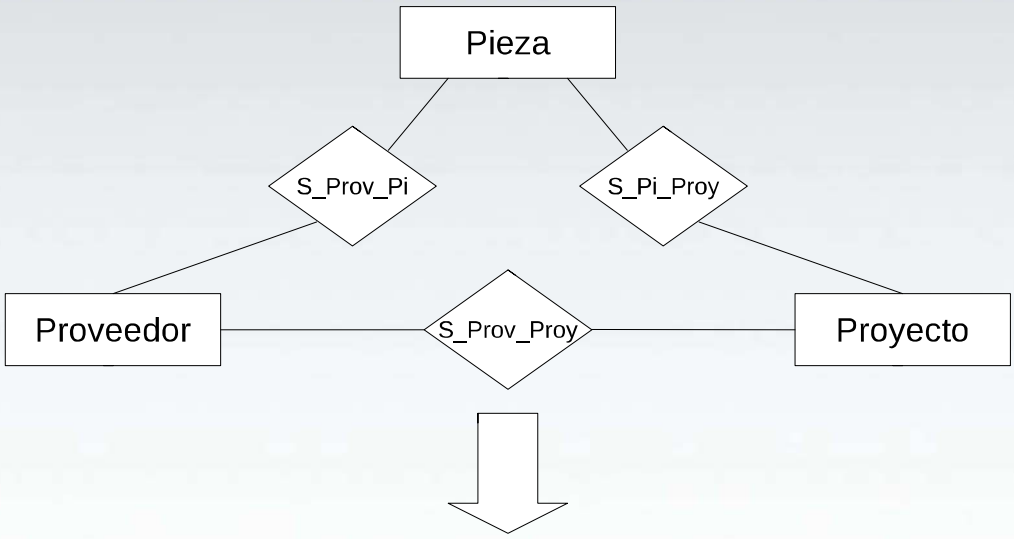
pero:

S_Pi_Proy |X| S_Pi_Proy |X| S_Prov_Proy = Suministro

A esto se le llama “Dependencia de Producto” y es lo que define la 5FN

5ta Forma Normal (5FN)

UNIVERSIDAD DE LOS ANDES VENEZUELA

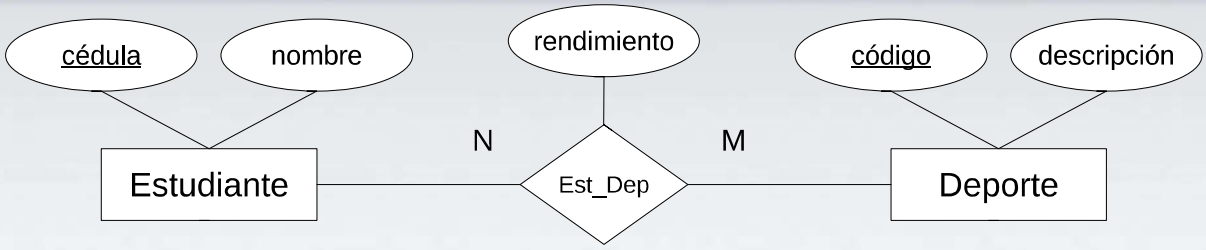


¿Será esto equivalente?
¿Será mejor o peor?

suministro	proveedor	pieza	proyecto

¿Está en 5FN?

¿Que ocurre con una relación N:M?



estudiante	cédula	nombre

deporte	código	descripción

est_dep	cédula	código	rendimiento

Se cumplen las siguientes DF:

cédula->nombre

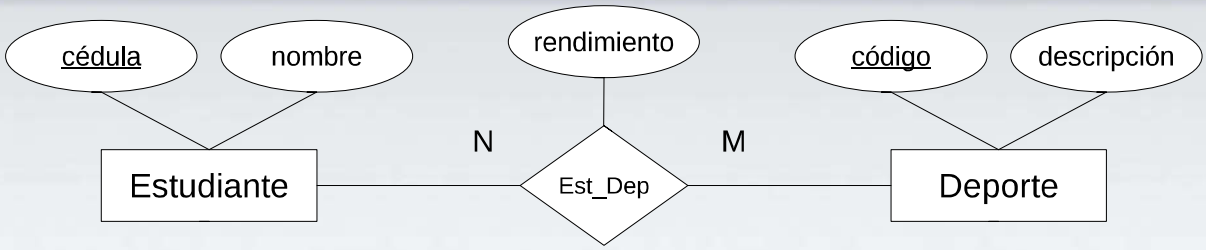
código->descripción

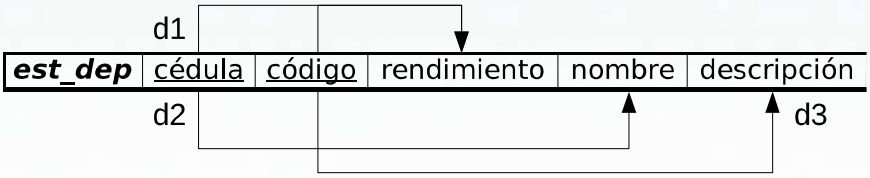
cédula. código->rendimiento

¿Está en 1FN, 2FN, 3FN, FNBC?

¿Se podrá "desnormalizar" la relación?

¿Que ocurre con una relación N:M?

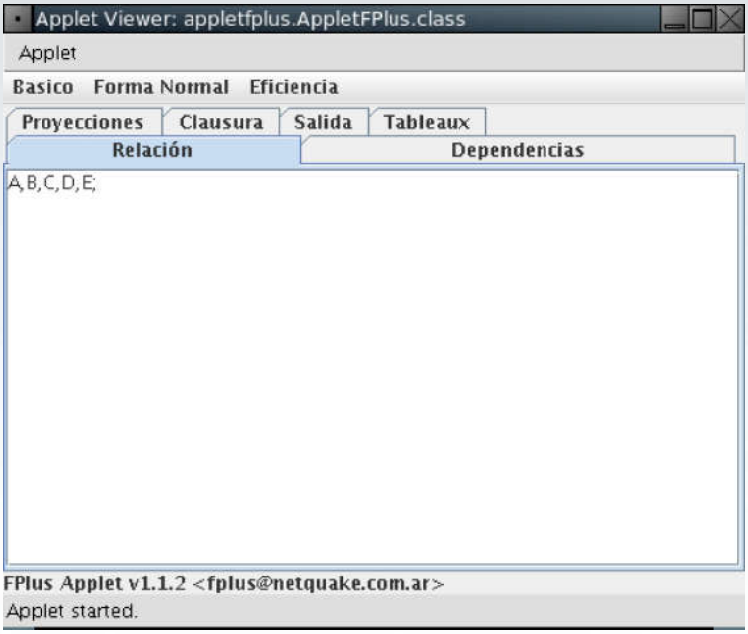




Conclusión

- El enfoque por descomposición funcional es más formal, pero resulta bastante complejo de utilizar en especial en bases de datos complejas con muchas dependencias funcionales.
- El enfoque de modelado semántico de datos (usando el modelo ERE o el modelo de clases) es mucho más simple y es el enfoque más comúnmente utilizado hoy en día.
- Es necesario asegurarse de que en las relaciones de las bases de datos modeladas no existan anomalías de ningún tipo causadas por la falta de normalización.

FPLUS!!!!



http://fplus.sourceforge.net/

UNIVERSIDAD DE LOS ANDES VENEZUELA

¡Gracias!