



Dpto. Ciencias e Ingeniería de la Computación  
 Universidad Nacional del Sur

## ELEMENTOS DE BASES DE DATOS

Segundo Cuatrimestre 2013

### Clase 9: Teoría Modelo Relacional – Problemas de Diseño – Descomposiciones – Cubrimientos



Mg. María Mercedes Vitturini  
 [mvitturi@uns.edu.ar]

## Modelo Relacional - Framework

Sabemos que el **modelo relacional** define una “*colección de tablas*” para representar los datos de un problema y sus relaciones.

- Se espera diseñar *relaciones bien estructuradas*, esto es, con **mínima redundancia y que permitan agregar, borrar y modificar tuplas sin causar errores o inconsistencias.**
- Estamos analizando cómo respondernos:  
**¿éste es un buen diseño?**  
**¿existe un diseño relacional de mayor calidad?**




EBD2013\_8 - Mg. Mercedes Vitturini

## Teoría Relacional

- Una forma de encontrar un diseño relacional de calidad es partir de un buen diseño conceptual, i.e. Entidad-Relación.
- Por su lado, *la teoría relacional analiza la calidad de los esquemas de relación en función de los conceptos:*
  - Dependencias Funcionales**
  - Llave primaria**

*Sabemos que:*

- Dado un **esquema de relación R**, es posible definir un **conjunto F de dependencias funcionales  $X \rightarrow Y$** , con  $X, Y \subseteq R$ . **F es el conjunto de restricciones esperadas para cualquier  $r(R)$**  (F puede ser vacío!).



EBD2013\_8 - Mg. Mercedes Vitturini

## Hasta ahora vimos que...

- Dado un esquema R, se define el conjunto F de dependencias funcionales o restricciones sobre las  $r(R)$ .
- Dado F se define como  $F^+$ , al conjunto de dependencias funcionales lógicamente implicadas de F.
  - Toda relación  $r(R)$  que satisface F, entonces satisface  $F^+$ .
- Dados R y F, sea  $X \subseteq R$ ,  $X^+$  es el conjunto de atributos lógicamente implicados por X bajo las restricciones F.
- Vimos:
  - un algoritmo simple para calcular  $X^+$ .
  - los usos de  $X^+$ .

EBD2013\_8 - Mg. Mercedes Vitturini

## Analizamos el esquema...

¿Una clave para PersHob?



Documento	Apellido	Nombre	Domicilio	Hobby	Hob. Categ.
11111111	Lorda	Ana	San Juan 25	Bailar	Arte
11111111	Lorda	Ana	San Juan 25	Cantar	Arte
11111111	Lorda	Ana	San Juan 25	Cocinar	Manualidad
22222222	Cuna	Juan	Mendoza 12	Cantar	Arte
22222222	Cuna	Juan	Mendoza 12	Pescar	Deporte
33333333	Donus	Carlos	Salta 332	Jugar al Tenis	Deporte
44444444	Pico	Clara	Vieytes 54	Bailar	Arte
44444444	Pico	Clara	Vieytes 54	Jugar al Tenis	Deporte
55555555	Pico	José	Vieytes 54	Cocinar	Manualidad
55555555	Pico	José	Vieytes 54	Jugar al Tenis	Deporte

**Redundancia – Posibilidad de Inconsistencias**  
**Anomalías de Inserción**  
**Anomalías de Borrado**

12\_9 - Mg. Mercedes Vitturini

## ¿Dependencias Funcionales?

Documento	Apellido	Nombre	Domicilio	Hobby	Hob. Categ.
11111111	Lorda	Ana	San Juan 25	Bailar	Arte
11111111	Lorda	Ana	San Juan 25	Cantar	Arte
11111111	Lorda	Ana	San Juan 25	Cocinar	Manualidad
22222222	Cuna	Juan	Mendoza 12	Cantar	Arte
22222222	Cuna	Juan	Mendoza 12	Pescar	Deporte
33333333	Donus	Carlos	Salta 332	Jugar al Tenis	Deporte
44444444	Pico	Clara	Vieytes 54	Bailar	Arte
44444444	Pico	Clara	Vieytes 54	Jugar al Tenis	Deporte
55555555	Pico	José	Vieytes 54	Cocinar	Manualidad
55555555	Pico	José	Vieytes 54	Jugar al Tenis	Deporte

**PersHob(documento, apellido, nombre, domicilio, hobby, hob\_categoria)**

Dependencias funcionales:  $F = \{$

- documento  $\rightarrow$  documento, apellido, nombre, domicilio,
- hobby  $\rightarrow$  hobby, hob\_categoria,
- documento, hobby  $\rightarrow$  documento, hobby }

## ¿Qué constituye un mal diseño de base de datos?

- Un **mal diseño de la base de datos** es aquel que tiene uno o más de los siguientes problemas de:
  - Redundancia en la información.
  - Inconsistencia en los datos.
  - Anomalías de inserción.
  - Anomalías de borrado.

EBD2012\_9 - Mg. Mercedes Vitturini

## Problemas de diseño

### REDUNDANCIA

- Si una persona tiene más de un hobby existen más de una fila para la misma persona, junto con sus datos personales (documento, apellido, nombre y domicilio).

### INCONSISTENCIA

- ¿Pueden existir para el mismo número de documento domicilios distintos? ¿Cuál es el correcto?

### ANOMALÍAS DE INSERCIÓN

- El problema de conocer todos los datos de una persona, pero no su hobby. Por ser hobby parte de la clave primaria, no es posible ingresar la persona a la tabla.

### ANOMALÍAS DE BORRADO

- Si para una persona deseamos borrar la información del único hobby que tiene cargado, debemos borrar todos sus datos

EBD2012\_9 - Mg. Mercedes Vitturini

## ¿Cuál es la solución?

- ¿Cuál es la solución si existen problemas de diseño en el modelo relacional?
  - Encontrar una **descomposición u organización mejor para los atributos.**
- ¿Cualquier descomposición de R es buena?
  - **NO!**
  - Se buscan descomposiciones con las propiedades:
    - **Join sin pérdida,**
    - **preservan dependencias,**
    - **respeten una buena forma normal.**

EBD2012\_9 - Mg. Mercedes Vitturini

## Descomposición #1

Propiedades:

- **Join con pérdida**
- **Preserva dependencias**
- **FNBC**

Hobby	
Hobby	Hob. Categ
Bailar	Arte
Cantar	Arte
Cocinar	Manualidad
Pescar	Deporte
Jugar al Tenis	Deporte

Hobby (hobby, hob\_categoria)  
 hobby → hobby,  
 hob\_categoria

Persona			
Documento	Apellido	Nombre	Domicilio
11111111	Lorda	Ana	San Juan 25
22222222	Cuna	Juan	Mendoza 12
33333333	Donus	Carlos	Salta 332
44444444	Pico	Clara	Vieytes 54
55555555	Pico	José	Vieytes 54

Persona (documento, apellido, nombre, domicilio)  
 documento → documento, apellido, nombre, domicilio

EBD2012\_9 - Mg. Mercedes Vitturini

¿Qué hobbies tiene Ana?

## Descomposición #2

Propiedades:

- **Join con pérdida.**
- **No preserva dependencias.**
- **1FN.**

Persona				
Documento	Apellido	Nombre	Domicilio	Hob. Categ
11111111	Lorda	Ana	San Juan 25	Arte
11111111	Lorda	Ana	San Juan 25	Manualidad
22222222	Cuna	Juan	Mendoza 12	Arte
22222222	Cuna	Juan	Mendoza 12	Deporte
33333333	Donus	Carlos	Salta 332	Deporte
44444444	Pico	Clara	Vieytes 54	Arte
44444444	Pico	Clara	Vieytes 54	Deporte
55555555	Pico	José	Vieytes 54	Manualidad
55555555	Pico	José	Vieytes 54	Deporte

Persona (documento, nombre, domicilio, hob\_categoria)

F= (documento → documento, apellido, nombre, hob\_categoria, documento → hob\_categoria, documento)

EBD2012\_9 - Mg. Mercedes Vitturini

Per-Hobby	
Documento	Hobby
11111111	Bailar
11111111	Cantar
11111111	Cocinar
22222222	Cantar
22222222	Pescar
33333333	Jugar al Tenis
44444444	Bailar
44444444	Jugar al Tenis
55555555	Cocinar

Per-Hob (documento, hobby)

Sin dependencias funcionales

¿Cuáles son los hobbies de Ana?

## Descomposición #3

Propiedades:

- **Join sin pérdida.**
- **Preserva dependencias.**
- **FNBC**

Hobby	
Hobby	Hob. Categ
Bailar	Arte
Cantar	Arte
Cocinar	Manualidad
Pescar	Deporte
Jugar al Tenis	Deporte

hobby → hob\_categoria

Persona				Per-Hobby	
Documento	Apellido	Nombre	Domicilio	Documento	Hobbies
11111111	Lorda	Ana	San Juan 25	11111111	Bailar
22222222	Cuna	Juan	Mendoza 123	11111111	Cantar
33333333	Donus	Carlos	Salta 332	11111111	Cocinar
44444444	Pico	Clara	Vieytes 54	22222222	Cantar
55555555	Pico	José	Vieytes 54	22222222	Pescar
				33333333	Jugar al Tenis
				44444444	Bailar
				44444444	Jugar al Tenis
				55555555	Cocinar

documento → apellido, nombre, domicilio

EBD2012\_9 - Mg. Mercedes Vitturini

## Descomposiciones

**Definición** – una **descomposición** de un esquema de relación  $R(A_1, A_2, \dots, A_n)$  es una **colección de subconjuntos**  $R_i$  de  $R$  de la forma  $\rho = (R_1, R_2, \dots, R_k)$  tal que  $R = R_1 \cup R_2 \cup \dots \cup R_k$ .

Ejemplo:

- PersHob (dni, apellido, nombre, domicilio, hobby, hob\_categ)
  - $R_1 = \text{PERSONA}$  (dni, apellido, nombre, domicilio)
  - $R_2 = \text{PERSONA\_HOBBY}$  (dni, hobby)
  - $R_3 = \text{HOBBY}$  (hobby, hob\_categ)
- **Observación:** la descomposición no exige que los  $R_i$ 's sean disjuntos.

EBD2012\_9 - Mg. Mercedes Vitturini

## Join sin Pérdida (JSP)

Un esquema de relación  $R(A_1, A_2, \dots, A_n)$  descompuesto en

$$\rho = (R_1, R_2, \dots, R_k)$$

es una **descomposición join sin pérdida** con respecto al conjunto de df's  $F$  si cada relación  $r$  de  $R$  satisfaciendo  $F$  es tal que:

$$r = \Pi_{R_1}(r) \bowtie \Pi_{R_2}(r) \bowtie \dots \bowtie \Pi_{R_k}(r)$$

- Esto es, toda  $r(R)$  es igual al join natural de sus proyecciones en los  $R_i$ 's.
- Cuando  $r \subset \Pi_{R_1}(r) \bowtie \Pi_{R_2}(r) \bowtie \dots \bowtie \Pi_{R_k}(r)$  entonces hay más información en la descomposición que en la relación original y se dice que es una **descomposición join con pérdida**.

EBD2012\_9 - Mg. Mercedes Vitturini

## Join con Pérdida

$r$			
A	B	C	D
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>	d <sub>1</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>2</sub>	d <sub>1</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>3</sub>	d <sub>1</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>	d <sub>2</sub>

$r_1$		
A	B	D
a <sub>1</sub>	b <sub>1</sub>	d <sub>1</sub>
a <sub>1</sub>	b <sub>1</sub>	d <sub>2</sub>

$r_2$	
B	C
b <sub>1</sub>	c <sub>1</sub>
b <sub>1</sub>	c <sub>2</sub>
b <sub>1</sub>	c <sub>3</sub>

$r_1 \bowtie r_2$			
A	B	C	D
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>	d <sub>1</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>	d <sub>2</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>2</sub>	d <sub>1</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>2</sub>	d <sub>2</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>3</sub>	d <sub>1</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>3</sub>	d <sub>2</sub>

$r \subset r_1 \bowtie r_2$   
Es JCP

EBD2012\_9 - Mg. Mercedes Vitturini

## Test Join sin Pérdida

### Teorema

• Sea  $\rho = (R_1, R_2)$  una descomposición en dos subesquemas para  $R$  y sea  $F$  el conjunto de df's, entonces  $\rho$  es una descomposición jsp con respecto a  $F$  **si y solo si se verifica que:**

- $(R_1 \cap R_2) \rightarrow R_2$ , o bien:
- $(R_1 \cap R_2) \rightarrow R_1$ .

• **Importante:** es suficiente con encontrar a una de tales dependencias en  $F$ .

EBD2012\_9 - Mg. Mercedes Vitturini

## Ejemplo

- Dado:
  - **PersHob** (documento, apellido, nombre, domicilio, hobby, hob\_categ)
  - $F = \{\text{documento} \rightarrow \text{documento, apellido, nombre, domicilio}, \text{hobby} \rightarrow \text{hobby, hob\_categ}\}$
- La descomposición#1 en los subesquemas:
  - $R_1 = \text{Personas}$  (documento, apellido, nombre, domicilio)
  - $R_2 = \text{Hobbies}$  (hobby, hob\_categ)
$$(R_1 \cap R_2) \not\rightarrow R_1 \quad (R_1 \cap R_2) \not\rightarrow R_2.$$

EBD2012\_9 - Mg. Mercedes Vitturini

## Ejemplo

- La descomposición# 3 en los sub-esquemas:
  - $R_1 = \text{PERSONA}$  (documento, apellido, nombre, domicilio), con  $F_1 = \{\text{documento} \rightarrow \text{apellido, nombre, domicilio}\}$
  - $R_2 = \text{HOBBY}$  (hobby, hob\_categ), con  $F_2 = \{\text{hobby} \rightarrow \text{hob\_categ}\}$
  - $R_3 = \text{PER-HOBBY}$  (documento, hobby)
$$(R_1 \cap R_3) \rightarrow R_1 \quad ((R_1 \cup R_3) \cap R_2) \rightarrow R_2, \text{ luego la descomposición es jsp}$$

EBD2012\_9 - Mg. Mercedes Vitturini

### Testeo de una descomposición

- Una descomposición válida debería cumplir con la propiedad jsp.
- Para testear si la descomposición de  $R, \rho=(R_1, R_2)$  en *dos subesquemas* tiene la propiedad j.s.p
  - Usar el Teorema.
- Para testear si la descomposición de  $R, \rho=(R_1, R_2 \dots R_n)$  tiene la propiedad j.s.p
  - Usar repetidamente el teorema ó
  - Usar el algoritmo Anexo.

EBD2012\_9 - Mg. Mercedes Vitturini

### Preserva Dependencias (PD)

Sea  $R$  un esquema de relación y  $F$  el conjunto de df's para  $R, R_i \subseteq R$ . La *proyección de  $F$  en  $R_i$* , denotado por  $\Pi_{R_i}(F)$ , es el conjunto de df's  $X \rightarrow Y$  en  $F^+$  tales que  $XY \subseteq R_i$ .

**Definición** – Sea  $\rho = (R_1, R_2, \dots, R_k)$  una descomposición de  $R$  y  $F$  el conjunto de df's para  $R$ . Se dice que  $\rho$  **preserva dependencias** si:

$$\bigcup_{i=1}^k \Pi_{R_i}(F^+) = \Pi_{R_1}(F^+) \cup \Pi_{R_2}(F^+) \cup \dots \cup \Pi_{R_k}(F^+) \mid= F$$

EBD2012\_9 - Mg. Mercedes Vitturini

### Ejemplo 1

- Sea  $R=(CSZ)$ ,
  - $F = \{CS \rightarrow Z, Z \rightarrow C\}$
  - $\rho = (CZ, SZ)$ .
- Luego:
  - $R_1 = (CZ)$  y  $F_1 = \Pi_{R_1}(F) = \{Z \rightarrow C\}$ ,
  - $R_2 = (SZ)$  y  $F_2 = \Pi_{R_2}(F) = \{ \}$ .
- ¿Preserva dependencias?
  - **No preserva dependencias**, es imposible recuperar la dependencia funcional  $CS \rightarrow Z$ .
- ¿Es join sin pérdida ?
  - $F \mid= (SZ \cap CZ) \rightarrow CZ = Z \rightarrow C$ .

EBD2012\_9 - Mg. Mercedes Vitturini

### Ejemplo 2

- Sea  $R=(ABCD), F = \{A \rightarrow B, C \rightarrow D\}$  y  $\rho = (AB, CD)$ .
- Esto es:
  - $R_1 = (AB)$  y  $F_1 = \Pi_{R_1}(F) = \{A \rightarrow B\}$ ,
  - $R_2 = (CD)$  y  $F_2 = \Pi_{R_2}(F) = \{C \rightarrow D\}$ .
- Esta descomposición **preserva dependencias** puesto que  $F = F_1 \cup F_2 = \Pi_{R_1}(F) \cup \Pi_{R_2}(F)$ .
- Sin embargo, **no es j.s.p** pues  $F$  no implica ninguna de las siguientes df's:
  - $(AB \cap CD) \rightarrow (AB - CD) = \emptyset \rightarrow AB$ .
  - $(AB \cap CD) \rightarrow (CD - AB) = \emptyset \rightarrow CD$ .

EBD2012\_9 - Mg. Mercedes Vitturini

### Preserva de Dependencias

- Sin embargo, para responder correctamente deberíamos previamente calcular  $F^+$  y proyectar todas las dependencias que se proyectan en cada  $R_i$
- **Ejemplo:**
  - Sea  $R=(ABCD), F = \{A \rightarrow BC, B \rightarrow AD\}$  y  $\rho = (AD, ABC)$ .
  - $\rho = (AD, ABC)$  preserva dependencias:
    - $R_1 = (AD)$  y  $F_1 = \Pi_{R_1}(F) = \{A \rightarrow D\}$
    - $R_2 = (ABC)$  y  $F_2 = \Pi_{R_2}(F) = \{A \rightarrow BC, B \rightarrow A\}$ .

EBD2012\_9 - Mg. Mercedes Vitturini

### Preservación de Dependencias

- Existe un algoritmo alternativo no requiere de  $F^+$ .

**Definición:** Una **R-operación** sobre un conjunto de atributos  $Z$  con respecto a una descomposición y conjunto de df's se define como:

$$Z \cup ((Z \cap R_i)^+ \cap R_i)$$

- Esta operación agrega a  $Z$  los atributos  $A$  tales que  $(Z \cap R_i) \rightarrow A$  está en  $\Pi_{R_i}(F)$ .

EBD2012\_9 - Mg. Mercedes Vitturini

## Algoritmo Clausura sobre $\rho$

### Algoritmo Clausura

Datos de Entrada: Un esquema de relación  $R=(A_1, \dots, A_n)$ , una descomposición  $\rho=(R_1, \dots, R_k)$ , un conjunto de  $df$ 's  $F$  y un conjunto de atributos  $X$ .

Datos de Salida:  $Z$  ( $X^+$ : la clausura de  $X$  en  $\rho$ ).

### Comienzo del Algoritmo

$Z \leftarrow X$

Repetir

$W \leftarrow Z$

Para cada esquema  $R_i$  Hacer

$Z \leftarrow Z \cup ((Z \cap R_i)^+ \cap R_i)$

{ Esta es la clausura de  $Z$  con respecto a  $F$  }

Hasta  $W = Z$

### Fin del Algoritmo

## Ejemplo 3

- Sea  $R=(ABCD)$ ,  $\rho=(AB, BC, CD)$  y:

$$F = \{A \rightarrow B, B \rightarrow C, C \rightarrow D, D \rightarrow A\}$$

- Veremos si  $\rho$  preserva dependencias, esto es:

$$\exists G = \Pi_{AB}(F) \cup \Pi_{BC}(F) \cup \Pi_{CD}(F) \models F?$$

- $A \rightarrow B, B \rightarrow C$  y  $C \rightarrow D$  se preservan trivialmente.
- Se preserva  $D \rightarrow A$ ?

EBD2012\_9 - Mg. Mercedes Vitturini

## ¿Se preserva $D \rightarrow A$ ?

- Dados:

$R=(ABCD)$

$\rho=(AB, BC, CD)$

$F = \{A \rightarrow B, B \rightarrow C, C \rightarrow D, D \rightarrow A\}$

- Sea  $Z = \{D\}$ . Aplicar la **AB-operación** produce:

$$\{D\} \cup ((\{D\} \cap \{A, B\})_F^+ \cap \{A, B\}) = \{D\}$$

- Similarmente, la **BC-operación** no produce cambios.

- La **CD-operación** produce:

$$\{D\} \cup ((\{D\} \cap \{C, D\})_F^+ \cap \{C, D\}) = \{D\} \cup (\{D\}^+ \cap \{C, D\})$$

$$\{D\} \cup (\{A, B, C, D\} \cap \{C, D\}) = \{D\} \cup \{C, D\} = \{C, D\}$$

EBD2012\_9 - Mg. Mercedes Vitturini

## ¿Se preserva $D \rightarrow A$ ?

- En el siguiente paso de la **BC-operación** produce  $\{B, C, D\}$ .
- En el último paso, la **AB-operación** produce  $\{A, B, C, D\}$ .
- Por lo tanto, la clausura de  $D$  con respecto a  $G$  es  $ABCD$ .
- Luego  $\rho = (AB, BC, CD)$  **preserva dependencias**.
- Por lo tanto, tenemos que:
  - $G = \Pi_{AB}(F) \cup \Pi_{BC}(F) \cup \Pi_{CD}(F)$ .
  - $G = \{A \rightarrow B, B \rightarrow A\} \cup \{B \rightarrow C, C \rightarrow B\} \cup \{C \rightarrow D, D \rightarrow C\}$ .

EBD2012\_9 - Mg. Mercedes Vitturini

## Descomposición: Propiedades

- Dados  $R$  un esquema de relación,  $F$  un conjunto de  $df$ 's definido sobre  $R$ , en ciertos casos  $R$  no responde a un buen diseño relacional: problemas de redundancia, inconsistencias, anomalías de inserción y borrado.
- En tal caso la solución será descomponer a  $R$  en  $\rho(R_1, R_2, \dots, R_k)$  una representación mejor para  $R$ .
- Una descomposición  $\rho$  debe satisfacer:
  - Ser una descomposición JSP **SIEMPRE!!**
  - Idealmente, además se espera que  $\rho$  preserve las dependencias funcionales de  $F$  **DESEABLE!!**
  - En una Forma Normal Alta (FNBC, 3FN).

EBD2012\_9 - Mg. Mercedes Vitturini

## Sobre el conjunto $F$ de Dependencias Funcionales

Ejemplo :

**PeliSocio** (nroSocio, nombreSocio, dniSocio, idPelicula, nombrePelicula, fechaAlquiler)

$F = \{nroSocio \rightarrow nombreSocio, dniSocio;$   
 $idPelicula \rightarrow nombrePelicula; idPelicula, fechaAlquiler \rightarrow nroSocio,$   
 $dniSocio \rightarrow nroSocio\}$

$G = \{dniSocio \rightarrow nombreSocio, nroSocio;$   
 $nroSocio \rightarrow dniSocio, nombreSocio; idPelicula \rightarrow nombrePelicula;$   
 $idPelicula, fechaAlquiler \rightarrow dniSocio\}$

¿Qué podemos decir de  $F$  y  $G$ ?

EBD2012\_9 - Mg. Mercedes Vitturini

## Cubrimientos de conjuntos de DF's

**Definición** – dos conjuntos de df's  $F$  y  $G$  sobre un mismo esquema  $R$  son **equivalentes** ( $F \equiv G$ ) si:

$$F^+ = G^+$$

- Si  $F \equiv G$  entonces,  $F$  es un cubrimiento para  $G$  (y viceversa  $G$  es un cubrimiento para  $F$ ).

**Lema:** Dados dos conjuntos de df's  $F$  y  $G$  sobre el esquema  $R$ , entonces  $F \equiv G$  si y solo si:

$$F \models G \text{ y } G \models F.$$

EBD2012\_9 - Mg. Mercedes Vitturini

## Ejemplos

### Ejemplo 1:

– Sea  $R(ABC)$

$$F = \{A \rightarrow BC, B \rightarrow A\} \text{ y } G = \{A \rightarrow B, B \rightarrow AC\}$$

–  $F \models G$  ? sí

–  $G \models F$  ? sí

$$F \equiv G$$

$F$  cubre a  $G$  y  $G$  cubre a  $F$

### Ejemplo 2:

– Sea  $R(ABCD)$

$$F = \{A \rightarrow BC, B \rightarrow AD\} \text{ y } G = \{A \rightarrow B, B \rightarrow AC\}$$

–  $F \models G$  ? sí

–  $G \models F$  ? no

$$F \not\equiv G$$

$F$  cubre a  $G$ ,  $G$  NO cubre a  $F$

EBD2012\_9 - Mg. Mercedes Vitturini

## Qué pasa en este caso

### Ejemplo 3:

– Sea  $R(ABCD)$

$$F = \{A \rightarrow BC, B \rightarrow CD, A \rightarrow D\} \text{ y } G = \{A \rightarrow B, B \rightarrow CD\}$$

–  $F \models G$  ? sí

–  $G \models F$  ? sí

$$F \equiv G$$

$F$  cubre a  $G$  y  $G$  cubre a  $F$

Analicemos en detalle los conjuntos  $F$  y  $G$ ,  $R(ABCD)$ :

$$F = \{A \rightarrow BC, B \rightarrow CD, A \rightarrow D\} \text{ y}$$

$$G = \{A \rightarrow B, B \rightarrow CD\}$$

¿Qué conjunto de restricciones elegirías? ¿Por qué?

EBD2012\_9 - Mg. Mercedes Vitturini

## Cubrimiento NoRedundante

Quando  $F \models G$ , entonces  $F$  cubre a  $G$ .

### Definiciones

- Un conjunto  $F$  de dependencias funcionales es no redundante si no contiene ningún subconjunto propio  $F'$  ( $F' \subset F$ ) tal que  $F' \equiv F$ . Caso contrario, se dice que  $F$  es redundante.
- $F$  es un **cubrimiento no redundante para  $G$**  si  $F$  es un cubrimiento para  $G$  y  $F$  es no redundante.

EBD2012\_9 - Mg. Mercedes Vitturini

## Cubrimientos

- **Definición:** Un conjunto de df's  $F$  es *no redundante* si no existe ningún subconjunto propio  $F'$  de  $F$  equivalente a  $F$ , esto es, no existe  $F'$  tal que  $F' \subset F$  y  $F' \equiv F$ .

- **Definición:** Un conjunto de df's  $F$  es *mínimo* si no existe un conjunto de df's  $G$  equivalente a  $F$  con menos df's, esto es, no existe  $G$  tal que  $\|G\| < \|F\|$  y  $G \equiv F$ .

- **Definición:** Un conjunto de df's  $F$  es *óptimo* si es mínimo y no existe un conjunto de df's  $G$  equivalente a  $F$  de un tamaño menor (el tamaño es la suma de los atributos que aparecen a izquierda y derecha de cada df).

EBD2012\_9 - Mg. Mercedes Vitturini

## Resultados

- Si esquema de relación  $R$  no es de calidad  $\Rightarrow$  tiene problemas de **asegurar integridad en los datos** que persiste.
- La **solución** es encontrar otro diseño relacional que resulte de **descomponer  $R$  en  $\rho(R_1, R_2, \dots, R_k)$** .
  - $\rho(R_1, R_2, \dots, R_k)$  debe ser una descomposición j.s.p.
  - $\rho(R_1, R_2, \dots, R_k)$  debe respetar alguna Forma Normal alta (3FN o FNBC).
  - Es deseable que  $\rho(R_1, R_2, \dots, R_k)$  preserve dependencias.
- Las **calidad de un esquema  $R$  está condicionada por el conjunto de restricciones** o dependencias funcionales definidas en  $R$ .
- Dados  $R$  y  $F$ , es deseable **encontrar el un conjunto de restricciones  $F'$  equivalente a  $F$  que sea un conjunto de restricciones óptimo o al menos mínimo.**

EBD2013\_8 - Mg. Mercedes Vitturini

### ¿Cómo encontrar las llaves de R?

Dado un esquema de relación  $R(A_1, \dots, A_n)$ , y un conjunto  $F$  de  $df$ 's mínimo reducido, entonces para cada atributo  $A_i$ :



- **Siempre** es parte de una llave si no aparece ni a izquierda ni a derecha en ninguna dependencia de  $F$ , o aparece solamente a izquierda en  $df$ 's de  $F$ .
- **Nunca** es parte de una llave si aparece solamente a derecha en las  $df$ 's de  $F$ .
- **Tal vez** es parte de una llave si aparece a izquierda y derecha en dos o más  $df$ 's de  $F$ .

Un conjunto  $X$  es llave de  $R$  si  $X^+ = R$ .

EBD2012\_9 - Mg. Mercedes Vitturini

### Ejemplo: encontrar todas las llaves

Sea  $R=(ABDEGHI)$ , y un conjunto mínimo  $F$  de  $df$ 's  $\{ A \rightarrow B, BG \rightarrow H, G \rightarrow DE, AE \rightarrow G \}$ .

- **Siempre** son parte de una llave: **A** (a izquierda solamente) e **I** (ni a izquierda ni a derecha).
- **Nunca** son parte de una llave: **D** y **H** (aparecen a derecha solamente).
- **Tal vez** son parte de esta llave: **B, E** y **G** (aparecen a izquierda y derecha).
- Posibles llaves:  $\{ AI, AIB, AIE, AIG, AIBE, AIBG, AIEG, AIBEG \}$ .
- Llaves de esta relación:  $\{ AIE, AIG \}$

EBD2012\_9 - Mg. Mercedes Vitturini

### Temas de la Clase de Hoy

- Descomposiciones
- Propiedades de las descomposiciones:
  - Join sin pérdida
  - Preserva dependencias.
- Cubrimientos de conjuntos de dependencias.

#### Bibliografía:

- *Principles of Database and Knowledge Based Systems.* Jeffrey Ullman. Capitulo 7
- *Database System Concepts.* A. Silberschatz. Capitulo 7 (edición 2005)

EBD2012\_9 - Mg. Mercedes Vitturini

### Algoritmo EsJoinSinPérdida

#### Algoritmo EsJoinSinPérdida

Datos de Entrada: Un esquema de relación  $R=(A_1, \dots, A_n)$ , una descomposición  $\rho=(R_1, \dots, R_k)$  y un conjunto de  $df$ 's  $F$ .

Datos de Salida: EsJSP (verdadero o falso).

#### Comienzo del Algoritmo

Se construye una tabla de  $n$  columnas y  $k$  filas.

La fila  $i$  corresponde al esquema  $R_i$ .

La columna  $j$  corresponde al atributo  $A_j$ .

En la fila  $i$  y columna  $j$  se coloca  $a$ , si  $A_j$  está en  $R_i$ . Sino va  $b_j$ .

Repetir para cada  $df$   $X \rightarrow Y$  en  $F$ :

Si existen dos filas que coincidan en todos los atributos de  $X$  se igualan todos los atributos de  $Y$  en ambas filas (en caso de igualar  $a$ 's y  $b$ 's tiene prioridad la  $a$ ).

Hasta que no haya más cambios o una fila tiene todos  $a$ 's.

Si alguna fila tiene todos los atributos  $a$ 's

Entonces EsJSP  $\leftarrow$  verdadero.

Sino EsJSP  $\leftarrow$  falso.

#### Fin del Algoritmo

EBD2012\_9 - Mg. Mercedes Vitturini

### Ejemplo

Sean  $R=(ABCDE)$ ,  $\rho=(AB, AD, AE, BE, CDE)$  y  $F = \{ A \rightarrow C, B \rightarrow C, C \rightarrow D, DE \rightarrow C, CE \rightarrow A \}$

	A	B	C	D	E
AB	a <sub>1</sub>	a <sub>2</sub>	b <sub>13</sub>	b <sub>14</sub>	b <sub>15</sub>
AD	a <sub>1</sub>	b <sub>22</sub>	b <sub>23</sub>	a <sub>4</sub>	b <sub>25</sub>
AE	a <sub>1</sub>	b <sub>32</sub>	b <sub>33</sub>	b <sub>34</sub>	a <sub>5</sub>
BE	b <sub>41</sub>	a <sub>2</sub>	b <sub>43</sub>	b <sub>44</sub>	a <sub>5</sub>
CDE	b <sub>51</sub>	b <sub>52</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>

EBD2012\_9 - Mg. Mercedes Vitturini

### Ejemplo

Sean  $R=(ABCDE)$ ,  $\rho=(AB, AD, AE, BE, CDE)$  y  $F = \{ A \rightarrow C, B \rightarrow C, C \rightarrow D, DE \rightarrow C, CE \rightarrow A \}$

	A	B	C	D	E
AB	a <sub>1</sub>	a <sub>2</sub>	b <sub>13</sub>	b <sub>14</sub>	b <sub>15</sub>
AD	a <sub>1</sub>	b <sub>22</sub>	b <sub>13</sub>	a <sub>4</sub>	b <sub>25</sub>
AE	a <sub>1</sub>	b <sub>32</sub>	b <sub>13</sub>	b <sub>34</sub>	a <sub>5</sub>
BE	b <sub>41</sub>	a <sub>2</sub>	b <sub>43</sub>	b <sub>44</sub>	a <sub>5</sub>
CDE	b <sub>51</sub>	b <sub>52</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>

EBD2012\_9 - Mg. Mercedes Vitturini

### Ejemplo

Sean  $R=(ABCDE)$ ,  $\rho=(AB, AD, AE, BE, CDE)$  y  $F = \{ A \rightarrow C, B \rightarrow C, C \rightarrow D, DE \rightarrow C, CE \rightarrow A \}$

	A	B	C	D	E
AB	a <sub>1</sub>	a <sub>2</sub>	b <sub>13</sub>	b <sub>14</sub>	b <sub>15</sub>
AD	a <sub>1</sub>	b <sub>22</sub>	b <sub>13</sub>	a <sub>4</sub>	b <sub>25</sub>
AE	a <sub>1</sub>	b <sub>32</sub>	b <sub>13</sub>	b <sub>34</sub>	a <sub>5</sub>
BE	b <sub>41</sub>	a <sub>2</sub>	b <sub>13</sub>	b <sub>44</sub>	a <sub>5</sub>
CDE	b <sub>51</sub>	b <sub>52</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>

EBD2012\_9 - Mg. Mercedes Vitturini

### Ejemplo

Sean  $R=(ABCDE)$ ,  $\rho=(AB, AD, AE, BE, CDE)$  y  $F = \{ A \rightarrow C, B \rightarrow C, C \rightarrow D, DE \rightarrow C, CE \rightarrow A \}$

	A	B	C	D	E
AB	a <sub>1</sub>	a <sub>2</sub>	b <sub>13</sub>	a <sub>4</sub>	b <sub>15</sub>
AD	a <sub>1</sub>	b <sub>22</sub>	b <sub>13</sub>	a <sub>4</sub>	b <sub>25</sub>
AE	a <sub>1</sub>	b <sub>32</sub>	b <sub>13</sub>	a <sub>4</sub>	a <sub>5</sub>
BE	b <sub>41</sub>	a <sub>2</sub>	b <sub>13</sub>	a <sub>4</sub>	a <sub>5</sub>
CDE	b <sub>51</sub>	b <sub>52</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>

EBD2012\_9 - Mg. Mercedes Vitturini

### Ejemplo

Sean  $R=(ABCDE)$ ,  $\rho=(AB, AD, AE, BE, CDE)$  y  $F = \{ A \rightarrow C, B \rightarrow C, C \rightarrow D, DE \rightarrow C, CE \rightarrow A \}$

	A	B	C	D	E
AB	a <sub>1</sub>	a <sub>2</sub>	b <sub>13</sub>	a <sub>4</sub>	b <sub>15</sub>
AD	a <sub>1</sub>	b <sub>22</sub>	b <sub>13</sub>	a <sub>4</sub>	b <sub>25</sub>
AE	a <sub>1</sub>	b <sub>32</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>
BE	b <sub>41</sub>	a <sub>2</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>
CDE	b <sub>51</sub>	b <sub>52</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>

EBD2012\_9 - Mg. Mercedes Vitturini

### Ejemplo

Sean  $R=(ABCDE)$ ,  $\rho=(AB, AD, AE, BE, CDE)$  y  $F = \{ A \rightarrow C, B \rightarrow C, C \rightarrow D, DE \rightarrow C, CE \rightarrow A \}$

	A	B	C	D	E
AB	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>	a <sub>4</sub>	b <sub>15</sub>
AD	a <sub>1</sub>	b <sub>22</sub>	a <sub>3</sub>	a <sub>4</sub>	b <sub>25</sub>
AE	a <sub>1</sub>	b <sub>32</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>
BE	b <sub>41</sub>	a <sub>2</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>
CDE	b <sub>51</sub>	b <sub>52</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>

EBD2012\_9 - Mg. Mercedes Vitturini

### Ejemplo

Sean  $R=(ABCDE)$ ,  $\rho=(AB, AD, AE, BE, CDE)$  y  $F = \{ A \rightarrow C, B \rightarrow C, C \rightarrow D, DE \rightarrow C, CE \rightarrow A \}$

	A	B	C	D	E
AB	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>	a <sub>4</sub>	b <sub>15</sub>
AD	a <sub>1</sub>	b <sub>22</sub>	a <sub>3</sub>	a <sub>4</sub>	b <sub>25</sub>
AE	a <sub>1</sub>	b <sub>32</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>
BE	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>
CDE	a <sub>1</sub>	b <sub>52</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>

EBD2012\_9 - Mg. Mercedes Vitturini

### Ejemplo

Sean  $R=(ABCDE)$ ,  $\rho=(AB, AD, AE, BE, CDE)$  y  $F = \{ A \rightarrow C, B \rightarrow C, C \rightarrow D, DE \rightarrow C, CE \rightarrow A \}$

	A	B	C	D	E
AB	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>	a <sub>4</sub>	b <sub>15</sub>
AD	a <sub>1</sub>	b <sub>22</sub>	a <sub>3</sub>	a <sub>4</sub>	b <sub>25</sub>
AE	a <sub>1</sub>	b <sub>32</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>
BE	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>
CDE	a <sub>1</sub>	b <sub>52</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>

JSP

EBD2012\_9 - Mg. Mercedes Vitturini



### Ejemplo

Sea  $R = (ABC)$ ,  $F = \{A \rightarrow B\}$  y  $\rho = (AB, AC)$ .

	A	B	C
AB	a <sub>1</sub>	a <sub>2</sub>	b <sub>13</sub>
AC	a <sub>1</sub>	b <sub>22</sub>	a <sub>3</sub>

➔

	A	B	C
AB	a <sub>1</sub>	a <sub>2</sub>	b <sub>13</sub>
AC	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>

$AB \cap AC = A$

¿  $F \models A \rightarrow AB$  ?

SI

¿  $F \models A \rightarrow AC$  ?

NO

➔  
**Es JSP**

EBD2012\_9 - Mg. Mercedes Vitturini

### Ejemplo 2

Sea  $R = (ABC)$ ,  $F = \{A \rightarrow B\}$  y  $\rho = (AB, BC)$ .

	A	B	C
AB	a <sub>1</sub>	a <sub>2</sub>	b <sub>13</sub>
BC	b <sub>21</sub>	a <sub>2</sub>	a <sub>3</sub>

➔

	A	B	C
AB	a <sub>1</sub>	a <sub>2</sub>	b <sub>13</sub>
BC	b <sub>21</sub>	a <sub>2</sub>	a <sub>3</sub>

$AB \cap BC = B$

¿  $F \models B \rightarrow AB$  ?

NO

¿  $F \models B \rightarrow BC$  ?

NO

➔  
**No es JSP**

EBD2012\_9 - Mg. Mercedes Vitturini

## Algoritmo PreservaDependencias

### Algoritmo PreservaDependencias

Datos de Entrada: Un esquema de relación  $R=(A_1, \dots, A_n)$ , una descomposición  $\rho=(R_1, \dots, R_k)$  y un conjunto de df's F.  
 Datos de Salida: PD (verdadero o falso).

#### Comienzo del Algoritmo

Sea  $G = \cup_{i=1}^k \Pi_{R_i}(F^+)$

$H \leftarrow F$

PD  $\leftarrow$  Verdadero

Repetir

Sea  $X \rightarrow Y$  la primera df de H

$H \leftarrow H - \{X \rightarrow Y\}$

Si  $Y \notin X^+_G$  entonces PD  $\leftarrow$  Falso

Hasta (No PD) o ( $H = \emptyset$ )

#### Fin del Algoritmo

## Sobre las descomposiciones

- Un esquema de relación R no es de calidad, i.e. tiene problemas de redundancia y posibilidad de inconsistencias, anomalías de inserción y borrado, la solución es encontrar otro diseño relacional que resulte de descomponer R en  $\rho(R_1, R_2, \dots, R_k)$ .
- $\rho(R_1, R_2, \dots, R_k)$  debe ser una descomposición j.s.p.
- $\rho(R_1, R_2, \dots, R_k)$  debe respetar alguna Forma Normal alta (3FN o FNBC)
- Es deseable que  $\rho(R_1, R_2, \dots, R_k)$  preserve dependencias

EBD2013\_8 - Mg. Mercedes Vitturini