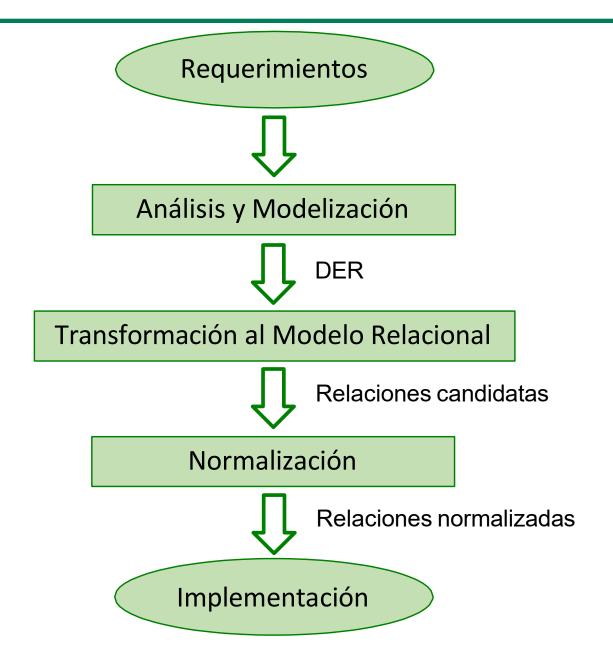


Universidad Nacional de La Matanza Catedra de Base de Datos

Clase de Normalización (parte 1 de 2)





Proceso de Diseño de Base de Datos



La Normalización es un proceso mediante el cual se puede <u>depurar</u> un diseño de base de datos.

<u>Permite eliminar ciertos defectos y características</u> <u>indeseables de los esquemas.</u>

Para entender bien cuáles son esos defectos que pueden tener los esquemas de datos, vamos a ver a continuación algunos ejemplos de problemas que se pueden presentar.



Problemas de un esquema NO Normalizado

Existen muchos problemas o inconvenientes, nosotros solo vamos a ver algunos de ellos:

- REDUNDANCIA
- ANOMALIAS DE ACTUALIZACIÓN
- ANOMALIAS DE INSERCIÓN
- ANOMALIAS DE ELIMINACIÓN
- PERDIDA DE INFORMACIÓN



Supongamos el siguiente esquema:

UNIVERSIDAD (dni_alumno, nya_alumno, materia, fecha_inscripción, fecha_examen, nota)

A simple vista se puede ver que este esquema no es correcto. Vamos a analizar sus problemas...



UNIVERSIDAD (dni_alumno, nya_alumno, materia, fecha_inscripción, fecha_examen, nota)

dni_alumno	nya_alumno	materia	fecha_inscripción	fecha_examen	nota
45	Florian Mayer	Bases de Datos	1/3/2014	1/4/2021	2
45	Florian Mayer	Bases de Datos	1/3/2014	1/4/2022	4
45	Florian Mayer	Fisica I	1/3/2014	1/4/2023	2
45	Florian Mayer	Fisica I	1/3/2014	1/7/2023	2
45	Florian Mayer	Fisica I	1/3/2014	1/12/2023	6

REDUNDANCIA

Si un alumno rindió 10 exámenes, vamos a repetir 10 veces su nombre cuando bastaría con tener solo el DNI.

Si un alumno cursó una materia y luego rindió 3 veces el examen final (hasta aprobarla), vamos a repetir la fecha de inscripción 3 veces. Esto ocurre porque estamos mezclando información de <u>hechos</u> <u>independientes</u>.



UNIVERSIDAD (dni_alumno, nya_alumno, materia, fecha_inscripción, fecha_examen, nota)

dni_alumno	nya_alumno	materia	fecha_inscripción	fecha_examen	nota
45	Florian Mayer	Bases de Datos	1/3/2014	1/4/2021	2
	Florian Mayer	Bases de Datos			
45			1/3/2014	1/4/2022	4
	Florian Mayer				
45		Fisica I	1/3/2014	1/4/2023	2
	Florian Mayer	Fisica I			
45			1/3/2014	1/7/2023	2
	Florian Mayer	Fisica I			
45			1/3/2014	1/12/2023	6

ANOMALIAS DE ACTUALIZACIÓN

Si hay que modificar el nombre de un alumno, tendremos que actualizar varias filas (en todas las que figure ese alumno).

Mientras que si hubiese estado bien normalizado, el nombre debería estar solo en una fila y solo sería necesario modificar un único valor.



UNIVERSIDAD (dni_alumno, nya_alumno, materia, fecha_inscripción, fecha_examen, nota)

dni_alumno	nya_alumno	materia	fecha_inscripción	fecha_examen	nota
45	Florian Mayer	Bases de Datos	1/3/2014	1/4/2021	2
45	,	Bases de Datos	1/3/2014	1/4/2021	2
45	Florian Mayer	Bases de Datos	1/3/2014	1/4/2022	4
	Florian Mayer				
45		Fisica I	1/3/2014	1/4/2023	2
	Florian Mayer	Fisica I			
45			1/3/2014	1/7/2023	2
	Florian Mayer	Fisica I			
45			1/3/2014	1/12/2023	6

ANOMALIAS DE INSERCIÓN

Si tenemos que dar de alta un nuevo alumno, que aun no cursó ni rindió ninguna materia, vamos a tener que dejar con valor NULO los otros atributos.

Idem si tenemos que dar de alta una materia sin alumnos.



UNIVERSIDAD (dni_alumno, nya_alumno, materia, fecha_inscripción, fecha_examen, nota)

dni_alumno	nya_alumno	materia	fecha_inscripción	fecha_examen	nota
45	Florian Mayer	Bases de Datos	1/3/2014	1/4/2021	2
	Florian Mayer	Bases de Datos			
45	,		1/3/2014	1/4/2022	4
	Florian Mayer				
45	·	Fisica I	1/3/2014	1/4/2023	2
	Florian Mayer	Fisica I			
45	,		1/3/2014	1/7/2023	2
	Florian Mayer	Fisica I			
45	ĺ		1/3/2014	1/12/2023	6

ANOMALIAS DE ELIMINACIÓN

Si tenemos que eliminar un Alumno y no queda otra fila con esa Materia, se pierde.

O bien, si se decide eliminar a todos los exámenes desaprobados (que tienen nota menor a 4) se eliminaran muchas filas y tal vez se pierdan alumnos que solo tienen exámenes desaprobados.

UNLaM – Base de Datos



UNIVERSIDAD (dni_alumno, nya_alumno, materia, fecha_inscripción, fecha_examen, nota)

PÉRDIDA DE INFORMACIÓN

Este problema aparece en el caso de una descomposición errónea de un esquema en varios subesquemas.

Mas adelante veremos detalladamente en que consiste la Pérdida de Información.



DEPENDENCIA FUNCIONAL

Definición:

Una dependencia funcional es una restricción entre dos conjuntos de atributos de una relación.

Sea R un esquema de relación, sea r una instancia cualquiera de R y sean X e Y dos subconjuntos de R. La dependencia funcional $X \rightarrow Y$ (léase X determina a Y) se cumple si y sólo si para cualesquiera dos tuplas t1 y t2 de r, tales que:

t1[X] = t2[X], entonces necesariamente t1[Y] = t2[Y]

Es decir, si dos tuplas cualesquiera tienen igual valor en X, deben tener igual valor en Y, para que se cumpla la dependencia funcional.

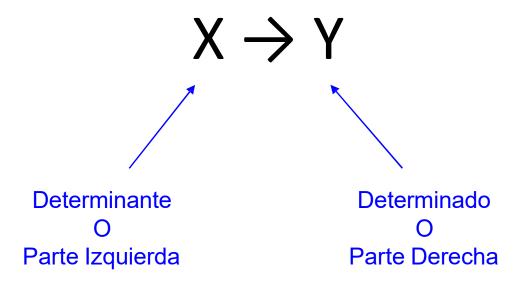
Esto significa que los valores componentes de Y de una tupla de r dependen de los valores del componente X, o están determinados por ellos; o bien, que los valores del componente X de una tupla determinan de manera única los valores del componente Y.



$$X \rightarrow Y$$



X determina a Y





EJEMPLO:

Tipo Documento, Nro Documento -> Nombre, Apellido (?)

EXAMEN

Tipo Documento	Nro Documento	Nombre	Apellido	Materia	Fecha
DNI	22.345.890	Diego	Perez	•••	•••
DNI	24.333.222	Martin	Malacia	•••	•••
Pasaporte	93.290.781	Aníbal	Marcato	•••	•••
DNI	22.345.890	•••	••••	•••	•••

Nombre, Apellido -> Tipo Documento, Nro Documento (?)



Ejemplo (Identificar DF):

EXAMEN (DNI_Alumno, Nombre, Apellido, Cod_Materia, Nombre_Materia, Fecha_Examen, Nota)

Se cumplen las siguientes dependencias funcionales:

F= {



Ejemplo:

EXAMEN (DNI_Alumno, Nombre, Apellido, Cod_Materia, Nombre_Materia, Fecha_Examen, Nota)

Se cumplen las siguientes dependencias funcionales:

F= { DNI_Alumno → Nombre, Apellido



Ejemplo:

EXAMEN (DNI_Alumno, Nombre, Apellido, Cod_Materia, Nombre_Materia, Fecha_Examen, Nota)

Se cumplen las siguientes dependencias funcionales:

F= { DNI_Alumno → Nombre, Apellido ; Cod_Materia → Nombre_Materia



```
Ejemplo:
```

```
EXAMEN (DNI_Alumno, Nombre, Apellido, Cod_Materia, Nombre_Materia, Fecha_Examen, Nota)
```

Se cumplen las siguientes dependencias funcionales:

```
F= { DNI_Alumno → Nombre, Apellido ;

Cod_Materia → Nombre_Materia ;

DNI_Alumno, Cod_Materia, Fecha_Examen → Nota
```



Ejercicio 1





DEPENDENCIA FUNCIONAL TRIVIAL

Decimos que una dependencia funcional es **Trivial** cuando es obvia.

Por ejemplo $X \rightarrow X$

Todo conjunto de atributos se determina a si mismo o a un conjunto de atributos menor (que este contenido en el primero).

$$XY \rightarrow X$$



CUESTIONES A TENER EN CUENTA:

Las dependencias funcionales vinculan dos conjuntos de atributos. Según la teoría de conjuntos:

- El orden de los elementos de un conjunto no importa.

 $XY \rightarrow Z$ es exactamente igual que $YX \rightarrow Z$

- Los conjuntos no pueden tener elementos repetidos.

no podemos tener $XX \rightarrow Y$ de las dos X automáticamente se suprime una y queda $X \rightarrow Y$



CLAVE Y SUPERCLAVE

Dado un esquema de relación R(A1, A2, ... An) y un conjunto de dependencias funcionales F asociado, se dice que el conjunto de atributos X perteneciente a R es una CLAVE CANDIDATA de R si se cumplen las siguientes dos condiciones:

- 1) X determina a todos los atributos de R, X \rightarrow A1, A2, ... An
- 2) No existe ningún Z (subconjunto de X) que determine a todos los atributos de R (condición de minimalidad).

Por otra parte, decimos que el conjunto de atributos Y es **SUPERCLAVE** de R si cumple la condición 1.

Se cumple que toda clave es superclave, pero no a la inversa.



CLAVE Y SUPERCLAVE

Ejemplo:

EMPLEADO (Legajo, Nombre, DNI)

Claves = {



CLAVE Y SUPERCLAVE



CLAVE Y SUPERCLAVE

```
Ejemplo:
EMPLEADO (Legajo, Nombre, DNI)

Claves Candidatas = { Legajo ;
DNI }
```



CLAVE Y SUPERCLAVE

Ejemplo:

EMPLEADO (Legajo, Nombre, DNI)

Claves Candidatas = { Legajo ; DNI }

Clave Primaria Primary Key (PK)

El Diseñador de la Base de Datos tiene que elegir a una de las claves candidatas como Clave Primaria



CLAVE Y SUPERCLAVE



CLAVE Y SUPERCLAVE

```
Ejemplo:
EMPLEADO (Legajo, Nombre, DNI)
Claves Candidatas = { Legajo ;
                     DNI }
Superclaves = { Legajo, Nombre ;
               DNI, Nombre;
               Legajo, DNI;
               Legajo, Nombre, DNI;
               Legajo;
               DNI }
```



AXIOMAS DE ARMSTRONG

Reflexividad

Si Y
$$\subseteq$$
 X => se cumple X \rightarrow Y

Aumento

Dada $X \rightarrow Y$ se puede inferir $XZ \rightarrow YZ$

Transitividad

Dadas $X \rightarrow Y$ y $Y \rightarrow Z$ se puede inferir $X \rightarrow Z$



REGLAS ADICIONALES/DERIVADAS

Descomposición

Dada $X \rightarrow YZ$ se pueden inferir $X \rightarrow Y$ y $X \rightarrow Z$

Unión

Dadas $X \rightarrow Y$ y $X \rightarrow Z$ se puede inferir $X \rightarrow YZ$

Pseudotransitividad

Dadas $X \rightarrow Y$ y $WY \rightarrow Z$ se puede inferir $WX \rightarrow Z$



REGLAS ADICIONALES/DERIVADAS

Las Reglas Adicionales se pueden demostrar a partir de los Axiomas de Armstrong.

Demostración de la Pseudotransitividad:

- 1) $X \rightarrow Y$ Dada
- 2) $WX \rightarrow WY$ Aumento con W en 1
- 3) WY→Z Dada
- 4) WX→Z Transitividad entre 3 y 2

Ver las otras demostraciones en el apunte de Normalización...



CLAUSURA DE UN CONJUNTO DE DEPENDENCIAS (F+)

Dado un conjunto de dependencias llamado F, denominamos clausura de F (F+) al conjunto de **todas las dependencias posibles** que se pueden inferir de F.



CLAUSURA DE UN CONJUNTO DE DEPENDENCIAS (F+)

```
Ejemplo: R(A, B, C) F=\{A \rightarrow B, BC \rightarrow A\}
```

F+={

Todas las dependencias Triviales

+

Las dependencias de F

+

Dependencias inferidas (Armstrong)



CLAUSURA DE UN CONJUNTO DE DEPENDENCIAS (F+)

```
Ejemplo: R(A, B, C) F=\{A\rightarrow B, BC\rightarrow A\}

F^{+}=\{A\rightarrow A,

B\rightarrow B,

C\rightarrow C,
```



CLAUSURA DE UN CONJUNTO DE DEPENDENCIAS (F+)

```
Ejemplo: R(A, B, C) F=\{A\rightarrow B, BC\rightarrow A\}

F^{+}=\{A\rightarrow A,

B\rightarrow B,

C\rightarrow C,

AB\rightarrow AB
```



CLAUSURA DE UN CONJUNTO DE DEPENDENCIAS (F+)

```
Ejemplo: R(A, B, C) F=\{A\rightarrow B, BC\rightarrow A\}

F^{+}=\{A\rightarrow A,

B\rightarrow B,

C\rightarrow C,

AB\rightarrow AB, AB\rightarrow A, AB\rightarrow B,
```



```
Ejemplo: R(A, B, C) F=\{A\rightarrow B, BC\rightarrow A\}

F^{+}=\{A\rightarrow A,

B\rightarrow B,

C\rightarrow C,

AB\rightarrow AB, AB\rightarrow A, AB\rightarrow B,

AC\rightarrow AC, AC\rightarrow A, AC\rightarrow C,

BC\rightarrow BC, BC\rightarrow B, BC\rightarrow C,
```



```
Ejemplo: R(A, B, C) F=\{A\rightarrow B, BC\rightarrow A\}

F^{+}=\{A\rightarrow A,

B\rightarrow B,

C\rightarrow C,

AB\rightarrow AB, AB\rightarrow A, AB\rightarrow B,

AC\rightarrow AC, AC\rightarrow A, AC\rightarrow C,

BC\rightarrow BC, BC\rightarrow B, BC\rightarrow C,

ABC\rightarrow ABC
```



```
Ejemplo: R(A, B, C) F=\{A\rightarrow B, BC\rightarrow A\}

F^{+}=\{A\rightarrow A, B\rightarrow B, C\rightarrow C, AB\rightarrow AB, AB\rightarrow A, AB\rightarrow B, AC\rightarrow AC, AC\rightarrow A, AC\rightarrow C, BC\rightarrow BC, BC\rightarrow B, BC\rightarrow C, ABC\rightarrow ABC, ABC\rightarrow AB, ABC\rightarrow AB, ABC\rightarrow C, ABC\rightarrow C,
```



```
Ejemplo: R(A, B, C) F={ A\rightarrowB, BC\rightarrowA }
F+={ A\rightarrowA,
B\rightarrowB,
C\rightarrowC,
AB\rightarrowAB, AB\rightarrowA, AB\rightarrowB,
AC\rightarrowAC, AC\rightarrowA, AC\rightarrowC,
BC\rightarrowBC, BC\rightarrowB, BC\rightarrowC,
ABC\rightarrowABC, ABC\rightarrowAB, ABC\rightarrowAC, ABC\rightarrowBC, ABC\rightarrowA, ABC\rightarrowB, ABC\rightarrowC, A\rightarrowB,
BC\rightarrowA
```



```
Ejemplo: R(A, B, C) F={ A\rightarrowB, BC\rightarrowA }
F+={ A\rightarrowA,
B\rightarrowB,
C\rightarrowC,
AB\rightarrowAB, AB\rightarrowA, AB\rightarrowB,
AC\rightarrowAC, AC\rightarrowA, AC\rightarrowC,
BC\rightarrowBC, BC\rightarrowB, BC\rightarrowC,
ABC\rightarrowABC, ABC\rightarrowAB, ABC\rightarrowAC, ABC\rightarrowBC, ABC\rightarrowA, ABC\rightarrowB, ABC\rightarrowC, A\rightarrowB, A\rightarrowAB
```



```
Ejemplo: R(A, B, C) F={ A\rightarrowB, BC\rightarrowA }
F+={ A\rightarrowA,
B\rightarrowB,
C\rightarrowC,
AB\rightarrowAB, AB\rightarrowA, AB\rightarrowB,
AC\rightarrowAC, AC\rightarrowA, AC\rightarrowC,
BC\rightarrowBC, BC\rightarrowB, BC\rightarrowC,
ABC\rightarrowABC, ABC\rightarrowAB, ABC\rightarrowAC, ABC\rightarrowBC, ABC\rightarrowA, ABC\rightarrowB, ABC\rightarrowC, A\rightarrowB, A\rightarrowAB
```



```
Ejemplo: R(A, B, C) F=\{A\rightarrow B, BC\rightarrow A\}

F^{+}=\{A\rightarrow A, B\rightarrow B, C\rightarrow C, AB\rightarrow AB, AB\rightarrow A, AB\rightarrow B, AC\rightarrow AC, AC\rightarrow A, AC\rightarrow C, BC\rightarrow BC, BC\rightarrow B, BC\rightarrow C, ABC\rightarrow ABC, ABC\rightarrow AB, ABC\rightarrow AC, ABC\rightarrow BC, ABC\rightarrow AB, ABC
```





```
Ejemplo: R(A, B, C) F=\{A\rightarrow B, BC\rightarrow A\}

F^{+}=\{A\rightarrow A, B\rightarrow B, C\rightarrow C, AB\rightarrow AB, AB\rightarrow A, AB\rightarrow B, AC\rightarrow AC, AC\rightarrow A, AC\rightarrow C, BC\rightarrow BC, BC\rightarrow B, BC\rightarrow C, ABC\rightarrow ABC, ABC\rightarrow AB, ABC\rightarrow AB, ABC\rightarrow AB, ABC\rightarrow AB, AC\rightarrow B, ABC\rightarrow AB, AC\rightarrow BB, AC\rightarrow AB, AC\rightarrow AC, AC\rightarrow AB, AC\rightarrow AC, AC\rightarrow AC, AC\rightarrow AC, AC\rightarrow AC, AC\rightarrow AC, AC
```



```
Ejemplo: R(A, B, C) F=\{A \rightarrow B, BC \rightarrow A\}
F^{+}=\{A\rightarrow A,
         B \rightarrow B,
         C \rightarrow C
         AB \rightarrow AB, AB \rightarrow A, AB \rightarrow B,
         AC \rightarrow AC, AC \rightarrow A, AC \rightarrow C,
         BC \rightarrow BC, BC \rightarrow B, BC \rightarrow C,
         ABC \rightarrow ABC, ABC \rightarrow AB, ABC \rightarrow AC, ABC \rightarrow BC, ABC \rightarrow A, ABC \rightarrow B, ABC \rightarrow C,
         A \rightarrow B, A \rightarrow AB, AC \rightarrow BC, AC \rightarrow ABC, AC \rightarrow AB, AC \rightarrow B,
         BC \rightarrow A, BC \rightarrow AB, BC \rightarrow AC, BC \rightarrow ABC
```



```
Ejemplo:
                     R(A, B, C) F=\{A \rightarrow B, BC \rightarrow A\}
F^+=\{A\rightarrow A,
                                                                                                         Azul: Triviales
        B \rightarrow B
                                                                                                          Rojo: Provenientes de F
                                                                                                         Verde: Inferidas
        C \rightarrow C
        AB \rightarrow AB, AB \rightarrow A, AB \rightarrow B,
        AC \rightarrow AC, AC \rightarrow A, AC \rightarrow C,
        BC \rightarrow BC, BC \rightarrow B, BC \rightarrow C,
        ABC \rightarrow ABC, ABC \rightarrow AB, ABC \rightarrow AC, ABC \rightarrow BC, ABC \rightarrow A, ABC \rightarrow B, ABC \rightarrow C,
        A \rightarrow B, A \rightarrow AB, AC \rightarrow BC, AC \rightarrow ABC, AC \rightarrow AB, AC \rightarrow B,
         BC \rightarrow A, BC \rightarrow AB, BC \rightarrow AC, BC \rightarrow ABC
```



CLAUSURA DE UN CONJUNTO DE ATRIBUTOS (X+)

Sea X un conjunto de atributos del esquema de relación R, y F un conjunto de dependencias funcionales que se cumplen en R, llamamos Clausura de X en F (X+F) al conjunto de **todos los** atributos determinados funcionalmente por X.



CLAUSURA DE UN CONJUNTO DE ATRIBUTOS (X+)

Ejemplo:

$$R(A, B, C)$$
 $F = \{A \rightarrow B, BC \rightarrow A\}$

$$A_F^+ = AB$$

$$B_F^+ = B$$

$$C_F^+ = C$$

$$AB_F^+ = AB$$

$$AC_F^+ = ACB$$

$$BC_F^+ = BCA$$

$$ABC_F^+ = ABC$$



CLAUSURA DE UN CONJUNTO DE ATRIBUTOS (X+)

Ejemplo:

$$R(A, B, C)$$
 $F = \{A \rightarrow B, BC \rightarrow A\}$

$$A_F^+ = AB$$

$$B_F^+ = B$$

$$C_F^+ = C$$

$$AB_F^+ = AB$$

$$AC_F^+ = ABC$$

$$BC_F^+ = ABC$$

$$ABC_F^+ = ABC$$

Con esto podemos calcular F+ fácilmente, (haciendo todas las combinaciones)

Introducción a la Normalización



Ejercicio 5





EQUIVALENCIA ENTRE DOS CONJUNTOS DE DEPENDENCIAS FUNCIONALES

Dos conjuntos de dependencias funcionales F y G son equivalentes cuando sus clausuras son iguales:

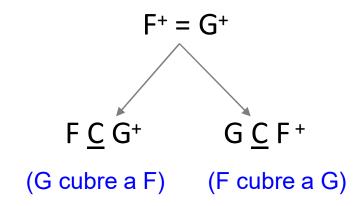
$$F \equiv G \iff F^+ = G^+$$

Calcular F⁺ y G⁺ puede ser muy largo y tedioso...



EQUIVALENCIA ENTRE DOS CONJUNTOS DE DEPENDENCIAS FUNCIONALES

Para probar que dos conjuntos F y G son equivalentes, sin necesidad de calcular F⁺ ni G⁺, se debería probar que toda dependencia de F puede inferirse en G (F \underline{C} G⁺), y recíprocamente, toda dependencia de G se puede inferir en F (G \underline{C} F⁺).



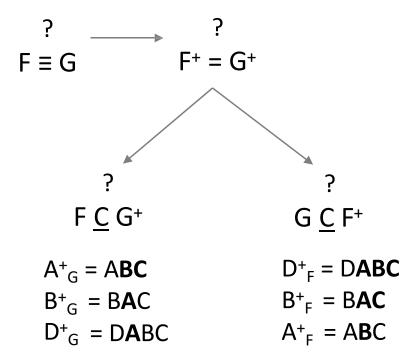
Podemos determinar si F cubre a G calculando X^+_F para cada dependencia $X \rightarrow Y$ de G, y comprobando que Y $C X^+_F$



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

$$G = \{ D \rightarrow ABC, B \rightarrow A, B \rightarrow C, A \rightarrow B \}$$

¿Son F y G equivalentes?



Respuesta: Sí, F y G son equivalentes.

Introducción a la Normalización



Ejercicio 8

