



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

ELEMENTOS DE BASES DE DATOS

Segundo Cuatrimestre 2015

Clase 8: Diseño Relacional (Parte I)

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



Sobre los temas de hoy

- Diseño Relacional

base de datos = { relaciones }

- Visión crítica: ¿Qué podemos decir de nuestro diseño relacional?

- ¿Cómo es cada esquema de relación?
- ¿Qué **dependencias funcionales** hay para cada esquema?

También nos interesa responder

- ¿Hay un diseño mejor?
- ¿Cómo encontrarlo?



Sobre los temas de hoy

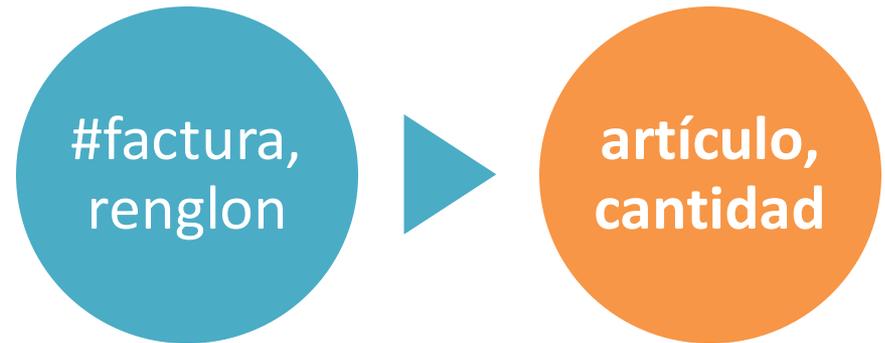
- Dependencia Funcional (df):
 - Definición
 - Satisfacibilidad
 - La PK como df.
- El conjunto F de dependencias funcionales de R
 - Implicancia
 - Clausura de F, F⁺
- Implicancia: Axiomas de Armstrong
 - Reflexividad
 - Aumento
 - Transitividad
 - + Reglas de inferencia
 - Unión
 - Pseudotransitividad
 - Descomposición
- Ejemplos

Sobre los temas de hoy

- Clausura de un conjunto de atributos: X_F^+

¿Cómo calcular la X_F^+ ?

- Deducir por los Axiomas de Amstrong
- Algoritmo Clausura
- Algoritmo es superllave
- Algoritmo es llave



¿Por qué calcular la X_F^+ ?

- Para obtener los atributos lógicamente implicados por X.
- Para responder $F \models X \rightarrow Y$.
- Para calcular F^+

Modelo Relacional - Framework

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

- Vamos a estudiar cómo respondernos:

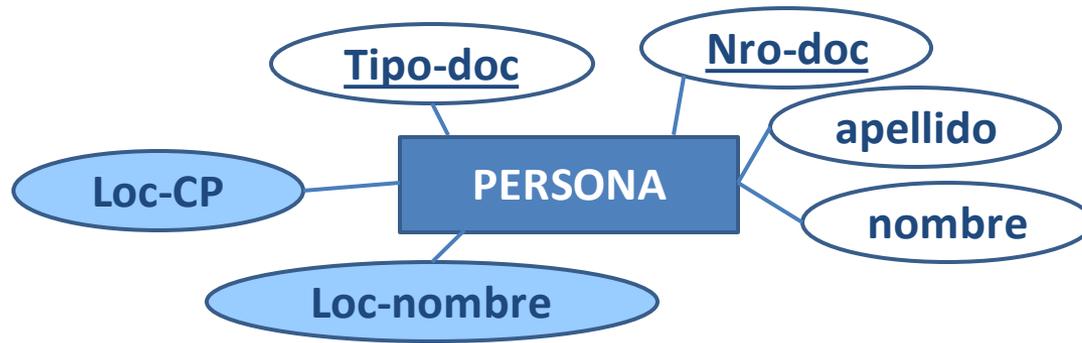
¿éste es un buen diseño?

¿existe un diseño relacional de mayor calidad?

- Una de las maneras para **encontrar un diseño relacional de calidad es partir de un buen diseño conceptual, i.e. Entidad-Relación**
- Para ello vamos a estudiar y analizar:
 - Dependencias funcionales.
 - Llaves de relación



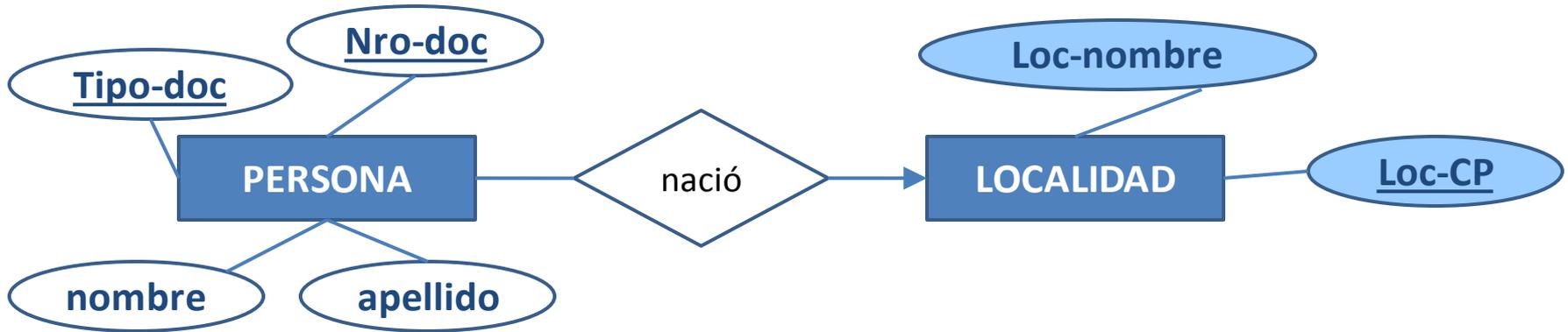
¿Es un buen diseño?



PERSONA (tipo-Doc, nro-Doc, apellido, nombre, **loc-CódigoPostal**, **loc-Nombre**)

Tipo Doc	Nro. Doc	Apellido	Nombre	Cod. Postal	Nombre Localidad
DNI	111111	Díaz	Alberto	8000	Bahía Blanca
DNI	222222	Vélez	Patricia	8000	Bahía Blanca
DNI	333333	Lino	Lucas	8000	Coronel Suárez
DNI	444444	Cané	María	8109	Punta Alta

Otro Diseño Posible



PERSONA (tipo-Doc, nro-Doc, apellido, nombre , loc-CódigoPostal)

LOCALIDAD (loc-CódigoPostal, ciudad-Nombre)

Tipo Doc	Nro. Doc	Apellido	Nombre	Cod. Posta
DNI	111111	Díaz	Alberto	8000
DNI	222222	Vélez	Patricia	8000
DNI	333333	Lino	Lucas	8000
DNI	444444	Cané	María	8109

Cod. Postal	Nombre Localidad
8000	Bahía Blanca
8109	Punta Alta
8302	Coronel Suárez

Dependencias Funcionales (DF's)

- Sea $R (A_1, \dots, A_n)$ un esquema de relación, con:
 - X e Y subconjuntos de $\{A_1, \dots, A_n\}$.
 - Se dice que " **X determina funcionalmente a Y** " o " **Y depende funcionalmente de X** " si para cualquier relación $r(R)$, no es posible que r contenga dos tuplas que coincidan en todos los componentes de X , pero no coincidan en uno o más componentes de Y :

$$\forall t_1, t_2 \in r \text{ si } t_1[X] = t_2[X] \text{ entonces } t_1[Y] = t_2[Y]$$

- Esta relación se representa como **$X \rightarrow Y$** .

Dependencia Funcional

“ $X \rightarrow A$ is an assertion about a relation R that whenever two tuples of R agree on all the attributes of X , then they must also agree on the attribute A .

- Say “ $X \rightarrow A$ holds in R .”

J. Ullman

Given an instance of $r(R)$, we say that the instance **satisfies** the **functional dependency $X \rightarrow Y$** if for all pairs of tuples t_1 and t_2 in the instance such that $t_1[X] = t_2[X]$, it is also the case that $t_1[Y] = t_2[Y]$.

- We say that the functional dependency $X \rightarrow Y$ **holds** on schema $r(R)$ if, in every legal instance of $r(R)$ it satisfies the functional dependency.

A. Silberschatz

Ejemplos

Algunos ejemplos de dependencias funcionales:

- artículo_código → artículo_descripción
- artículo_código, fecha → valor_venta
- nro_reserva → nro_vuelo, fecha_salida_vuelo, cliente_número
- tipo_documento, nro_documento → apellido, nombre, apellido, sexo, fecha_nacimiento
- nro_factura, nro_renglón → artículo, nro_unidades.
- ...

Dependencias Funcionales (DF's)

Una **dependencia funcional** es una forma de expresar restricciones que establecen para $r(R)$ que sólo un subconjunto de las relaciones posibles del dominio es “válido”.

Ejemplo: bajo la dependencia funcional **Loc_CP \rightarrow Loc_Nombre**, la tupla *persona* ('DNI', 333333, 'Lino', 'Lucas', 8000, 'Coronel Suárez') no sería válida

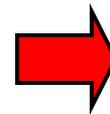
Tipo Doc	Nro. Doc	Apellido	Nombre	Cod. Postal	Nombre Localidad
DNI	111111	Díaz	Alberto	8000	Bahía Blanca
DNI	222222	Vélez	Patricia	8000	Bahía Blanca
DNI	333333	Lino	Lucas	8000	Coronel Suárez
DNI	444444	Cané	María	8109	Punta Alta

Propiedad: Satisfacibilidad

- Dado un esquema R , las dependencias funcionales **se definen sobre R y deben ser satisfechas por toda $r(R)$.**
- Formalmente, una relación r **satisface la df $X \rightarrow Y$** si para cualquier par de tuplas t_1 y t_2 en r vale que:
si $t_1[X] = t_2[X]$ entonces $t_1[Y] = t_2[Y]$.
 - Si r es una instancia del esquema R y hemos asegurado que $X \rightarrow Y$ vale para R entonces esperamos que r satisfaga $X \rightarrow Y$ **siempre.**

Satisfacibilidad

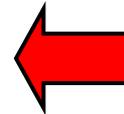
X			...	Y	
A	B	C	...	D	E
...
a1	b2	c3	...	d1	e2
...
a1	b2	c3	...	d1	e2
...



Vale $X \rightarrow Y$

X			...	Y	
A	B	C	...	D	E
...
a1	b2	c3	...	d1	e2
...
a1	b2	c3	...	d2	e1
...

No vale $X \rightarrow Y$



Llave como dependencia funcional

- Sea $R(A_1, \dots, A_n)$ un esquema de relación y X la llave de R entonces $X \rightarrow A_i$ vale para cualquier atributo A_i de R .
- Si R representa una relación muchos a uno de E_1 a E_2 y supongamos que entre los A_i 's existen suficientes atributos para formar una llave X para E_1 e Y para E_2 .
 - Entonces vale $X \rightarrow Y$.
 - Sin embargo, $Y \rightarrow X$ vale solo si la relación es uno a uno.
- ¿Cómo se determinan las dependencias funcionales?
 - A partir de la semántica del problema.*

Ejemplo

- Consideremos los atributos **M** (Materia), **P** (Profesor), **H** (Día-Hora), **A** (Aula), **E** (Estudiante) definidos en **R (MPHAE)**
- Podríamos tener, entre otras, las siguientes *dependencias funcionales*:
 - **HA** → **M** Un horario y aula una única materia.
 - **HP** → **AM** Un profesor en un horario está dando clases en una única aula y de una única materia.
 - **HE** → **A** Un alumno en un horario está en un único aula.
 - **EM** → **P** Un estudiante y una materia la cursa con un único profesor.

R(MPHAE) con,
 $F = \{HA \rightarrow M, HP \rightarrow AM, HE \rightarrow A, EM \rightarrow P\}$

Propiedad: Implica Lógicamente (\models)

- Sea F el conjunto de df 's correspondientes al esquema R , se dice que **F implica lógicamente $X \rightarrow Y$** , notado por **$F \models X \rightarrow Y$** si cada relación $r (R)$ que satisface las dependencias funcionales en F también satisface $X \rightarrow Y$.

- Ejemplo:

$$R (ABC) \text{ y } F = \{A \rightarrow B, B \rightarrow C\} \models A \rightarrow C.$$

Teoría del Modelo Relacional

Axiomas de Armstrong

- Sea U el conjunto universal de atributos, con X, Y, Z subconjunto de atributos de R :
 - **Reflexividad**: Si $Y \subseteq X \subseteq R$ entonces $X \rightarrow Y$.
 - **Aumento**: Si $X \rightarrow Y$ y $Z \subseteq R$ entonces $XZ \rightarrow YZ$.
 - **Transitividad**: Si $X \rightarrow Y$ e $Y \rightarrow Z$ entonces $X \rightarrow Z$.

Reglas de Inferencia

- **Unión**: $\{X \rightarrow Y, X \rightarrow Z\} \models X \rightarrow YZ$.
- **Pseudotransitividad**: $\{X \rightarrow Y, WY \rightarrow Z\} \models XW \rightarrow Z$.
- **Descomposición**: Si $X \rightarrow Y$ y $Z \subseteq Y$ entonces $X \rightarrow Z$.

Sean $R(\text{ABEGHI})$ y

$F = \{ \text{AB} \rightarrow \text{E}, \text{AG} \rightarrow \text{J}, \text{BE} \rightarrow \text{I}, \text{E} \rightarrow \text{G}, \text{GI} \rightarrow \text{H} \}.$

Veamos si $F \models \text{AB} \rightarrow \text{GH} ?$

1. $\text{AB} \rightarrow \text{E}$ Hipótesis.
2. $\text{AB} \rightarrow \text{B}$ Reflexividad.
3. $\text{AB} \rightarrow \text{BE}$ Unión sobre 1, 2.
4. $\text{BE} \rightarrow \text{I}$ Hipótesis.
5. $\text{AB} \rightarrow \text{I}$ Transitividad sobre 3, 4.
6. $\text{E} \rightarrow \text{G}$ Hipótesis.
7. $\text{AB} \rightarrow \text{G}$ Transitividad sobre 1, 6.
8. $\text{AB} \rightarrow \text{GI}$ Unión sobre 5, 7.
9. $\text{GI} \rightarrow \text{H}$ Hipótesis.
10. $\text{AB} \rightarrow \text{H}$ Transitividad sobre 8, 9.
11. $\text{AB} \rightarrow \text{GH}$ Unión sobre 7, 10.

Clausura de un Conjunto de df's: F^+

- Dado un conjunto de df's F sobre un esquema R existen otras df's que son *lógicamente implicadas* por F (\models).

– Ejemplo:

$$R(ABC), \text{ con } F = \{A \rightarrow B, B \rightarrow C\} \models A \rightarrow C$$

- El conjunto de *todas* las df's lógicamente implicadas por F es la **clausura** de F ó F^+ .
- ¿Cómo se obtienen las df's lógicamente implicadas?
 - Las df's que se implican lógicamente aplicando los Axiomas de Armstrong

Clausura de Conjuntos de DF's

Ejemplo: Sea $R=(ABC)$ y $F = \{AB \rightarrow C, C \rightarrow B\}$.

Su clausura incluye a las siguientes df's:

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

*No es práctico
calcular la
clausura de F*

Clausura de Atributos (X^+)

Dados R y F , sea $X \subseteq R$ un conjunto de atributos se define como la **clausura del conjunto de atributos X bajo F** (notada como X^+_F) al conjunto de los atributos funcionalmente determinados por X bajo F .

- Esto es, $\{A_i\}$ tal que $X \rightarrow A_i$ se puede deducir de F usando los Axiomas de Amstrong.

Ejemplo

- Sea $R=(ABCDE)$ y $F = \{A \rightarrow C, C \rightarrow B, AD \rightarrow E \}$
- $(A^+)_F = ACB$
 - $A \rightarrow A$ (reflexiva),
 - $A \rightarrow C$ (dada) y
 - $A \rightarrow B$ (transitiva en $A \rightarrow C$ y $C \rightarrow B$)

¿Cómo calcular X^+ ?

- Un algoritmo para calcular la clausura de un conjunto de atributos X bajo un conjunto de df's F (X^+_F) es:

1. Partir de la **reflexiva**.

– $(X^+_F) = X$

2. **Repetir** mientras exista una df. $Y \rightarrow Z$ en F tal que $Y \subseteq X^+$

– $(X^+_F) = (X^+_F) \cup Z$

Hasta que no existan cambios en (X^+_F)

Ejemplo

– Sea $R=(ABCDE)$ y $F = \{A \rightarrow C, C \rightarrow B, AD \rightarrow E\}$

- $A \rightarrow A$
- De $A \rightarrow A$ y $A \rightarrow C$ se obtiene $A \rightarrow AC$
- De $A \rightarrow AC$ y $C \rightarrow B$ se obtiene $A \rightarrow ACB$

Algoritmo CLAUSURA_X

Algoritmo Clausura_X (R, F, X)

Datos de Entrada: un conjunto finito de atributos R, un conjunto de dependencias funcionales F y un conjunto $X \subseteq R$.

Datos de Salida: X^+ , la clausura de X con respecto a F.

Comienzo del Algoritmo

$i \leftarrow 0$

$X^{(i)} \leftarrow X$

Repetir

$X^{(i+1)} \leftarrow X^{(i)}$

Repetir mientras exista $Y \rightarrow Z$ en F tal que $Y \subseteq X^{(i)}$

$X^{(i+1)} \leftarrow X^{(i+1)} \cup Z$

$i \leftarrow i + 1$

Hasta $X^{(i)} = X^{(i-1)}$

$X^+ \leftarrow X^{(i)}$

Fin del Algoritmo

Clausura de X- Ejemplo

Sean F y X los siguientes conjuntos: $R(ABCDEG)$, $X=BD$ y $F = \{ AB \rightarrow C, CG \rightarrow BD, D \rightarrow EG, C \rightarrow A, BE \rightarrow C, BC \rightarrow D, ACD \rightarrow B, CE \rightarrow AG \}$,

$X=BD$

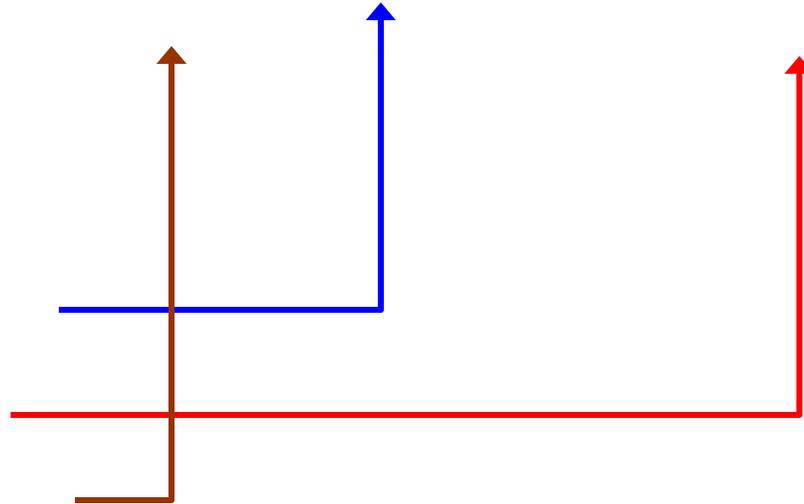
$X^{(0)} = BD$

$X^{(1)} = BDEG$

$X^{(2)} = BCDEG$

$X^{(3)} = ABCDEG$

$X^{(4)} = ABCDEG$



Por lo tanto, $X^+ = ABCDEG$.

Ejercicios para Clausura de X

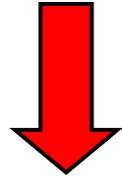
Sean:

$R(ABEGJIH)$ y $F = \{ AB \rightarrow E, AG \rightarrow J, BE \rightarrow I, E \rightarrow G, GI \rightarrow H \}$.

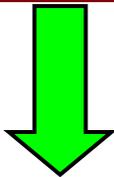
- Algunos ejemplos de clausuras de atributos con respecto a F:
 - $AB^+ = ABEIGHJ$
 - $AG^+ = AGJ$.
 - $A^+ = A$.
 - $BE^+ = BEIGH$.

Clausura de conjuntos de DF's

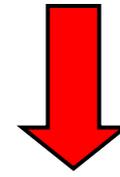
¿Cómo saber si vale $F \models X \rightarrow Y$?



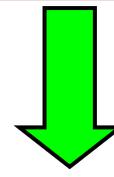
Verifico que $X \rightarrow Y$
pertenezca a la
*clausura del conjunto
de df's F (F^+)*



Poco práctico



Verifico que Y esté
contenido en la *clausura
del conjunto de atributos
 X (X^+)* con respecto a F



Algoritmo de Clausura_X

Usos de X^+

Resultados:

1. Para comprobar si un conjunto X de atributos es una **superllave** de R : $R \subseteq X^+$.
2. Para testear si F deduce lógicamente una dependencia funcional $X \rightarrow Y$ ($F \models \{X \rightarrow Y\}$)
 - $F \models \{X \rightarrow Y\} \equiv \{X \rightarrow Y\} \subseteq F^+ \equiv Y \subseteq (X^+)_F$.
3. Para calcular F^+
 - Por cada subconjunto de atributos $X \subseteq R$, calcular X^+ , y por cada subconjunto de atributos $S \subseteq X^+$, agregar a F^+ la dependencia funcional $X \rightarrow S$.

Ejercicios

- Sea $R = (ACDELMNPS)$ y F el siguiente conjunto de df's:

$$F = \{E \rightarrow NS, NL \rightarrow EMD, EN \rightarrow LCD, C \rightarrow S\} \cup \\ \{D \rightarrow M, M \rightarrow D, EPD \rightarrow AE, NLCP \rightarrow A\}$$

- Verificar que los siguientes conjuntos son superllaves de R según F :
 - DLMNP.
 - EMNPS.
 - ACDLNP.
 - ACEP.

Otros Resultados

Algoritmo Deduce (F, df)

Datos de Entrada: un conjunto de dependencias funcionales F y una dependencia funcional $X \rightarrow Y$.

Datos de Salida: Pertenece (*verdadero* si $F \models X \rightarrow Y$, *falso* en caso contrario).

Comienzo del Algoritmo

Si $Y \subseteq \text{Clausura}_X(R, F, X)$ entonces

Deduce \leftarrow Verdadero

Sino

Deduce \leftarrow Falso

Fin del Algoritmo

Algoritmo EsSuperLlave

Algoritmo EsSuperLlave (R, F, X)

Datos de Entrada: un conjunto de dependencias funcionales F, un esquema R y un conjunto de atributos $X \subseteq R$.

Datos de Salida: EsSuperLlave (*verdadero* si X es superllave de R, *falso* en caso contrario).

Comienzo del Algoritmo

Si $R \subseteq \text{Clausura}_X(R, F, X)$ entonces
EsSuperLlave \leftarrow Verdadero

Sino

EsSuperLlave \leftarrow Falso

Fin del Algoritmo

Algoritmo EsLlave

Algoritmo EsLlave (R, F, X)

Datos de Entrada: un conjunto de dependencias funcionales F, un esquema R y un conjunto de atributos $X \subseteq R$.

Datos de Salida: EsLlave (*verdadero* si X es llave de R, *falso* en caso contrario).

Comienzo del Algoritmo

Si $R \subseteq \text{Clausura}_X (R, F, X)$ entonces && X es superllave

Si existe $W \subset X$ tal que $R \subseteq \text{Clausura}_X (R, F, W)$ entonces

&& X tiene atributos demás

EsLlave \leftarrow Falso

Sino { X es llave }

EsLlave \leftarrow Verdadero

Sino

&& X no es llave ni superllave

EsLlave \leftarrow Falso

Fin del Algoritmo

Temas de la Clase de hoy

- Problemas de diseño
- Teoría Relacional
 - Dependencias Funcionales
 - Clausura de un conjunto de dependencias.
 - Clausura de un conjunto de atributos.
 - Algoritmos: Clausura_X, Pertence, esSuperllave, esLLave
- Bibliografía:
 - *Database System Concepts*. A. Silverschatz. Capítulo 8.
 - *Principles of Database and Knowledge Based Systems*. Jeffrey Ullman. Capítulo 8 (ed. 2010)
 - *DataBase System – The Complete Book* – H. Molina, J. Ullman. Capítulo 3.