

Diseño de Base de Datos Relacional

Temas:

- Pautas informales para diseño de esquemas
- Dependencias Funcionales
- Formas Normales
- Algoritmos de diseño
- Dependencias multivaluadas y cuarta forma normal
- Panorama del proceso de diseño de Base de Datos

Diseño de Base de Datos Relacional

◆ Referencia:

- Capítulos 12, 13 y 14 del [EN 2E].
- Capítulos 14, 15, 16. 1 y 16.2 del [EN 3E].

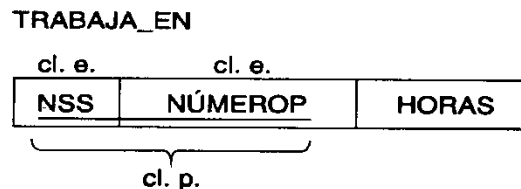
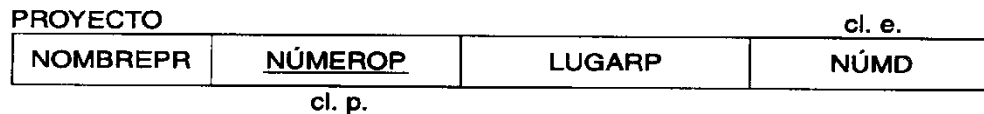
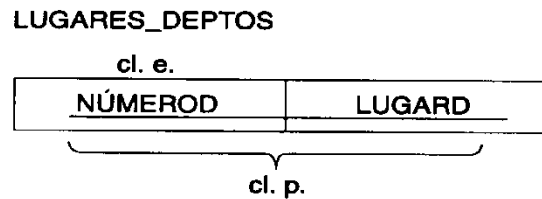
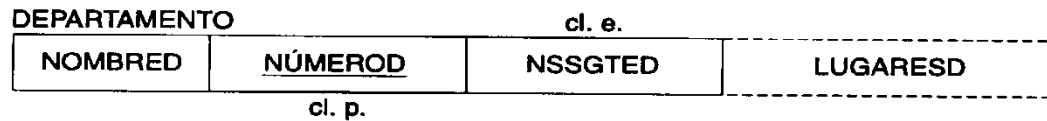
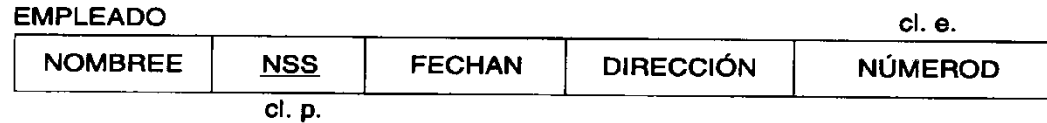
Pautas informales para el diseño

◆ 4 medidas informales de la calidad

- Semántica de los atributos
- Reducción de los valores redundantes en las tuplas
- Reducción de los valores nulos en las tuplas
- No generación de tuplas erróneas

Semántica de los atributos

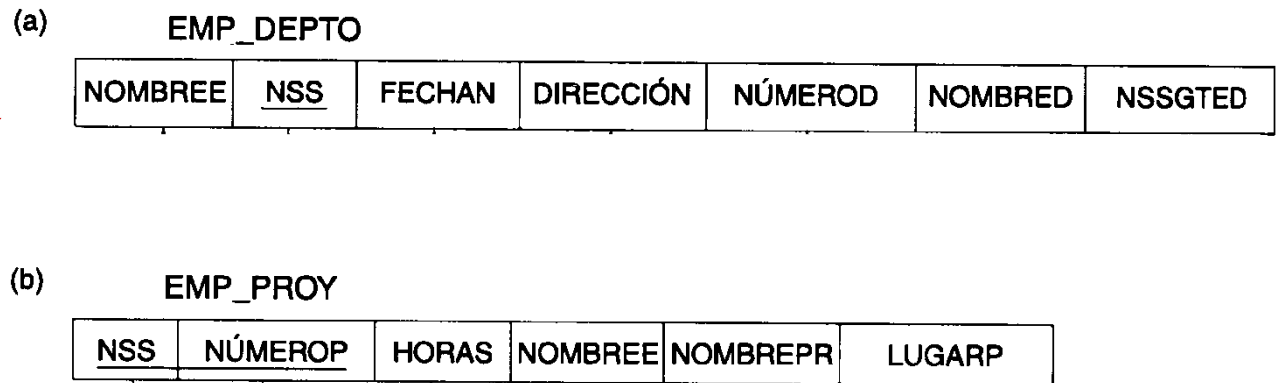
◆ Ejemplo



Semántica de los atributos

◆ Ejemplo

**Semánticamente
confusas**



Semántica de los atributos

◆ Pauta 1

- » Diseñe un esquema de relación de modo que sea fácil explicar su significado. No combine atributos de varios tipos de entidades y tipos de vínculos en una sola relación.

Reducción de los valores redundantes

◆ Información redundante en las tuplas

EMP_DEPTO

NOMBREE	NSS	FECHAN	DIRECCIÓN	NÚMEROD	NOMBRED	NSSGTED
Silva, José B.	123456789	09-ENE-55	Fresnos 731, Higueras, MX	5	Investigación	333445555
Vizcarra, Federico T.	333445555	08-DIC-45	Valle 638, Higueras, MX	5	Investigación	333445555
Zapata, Alicia J.	999887777	19-JUL-58	Castillo 3321, Sucre, MX	4	Administración	987654321
Valdés, Jazmín S.	987654321	20-JUN-31	Bravo 291, Belén, MX	4	Administración	987654321
Nieto, Ramón K.	666884444	15-SEP-52	Espiga 957, Heras, MX	5	Investigación	333445555
Esparza, Josefa A.	453453453	31-JUL-62	Rosas 5631, Higueras, MX	5	Investigación	333445555
Jabbar, Ahmed V.	987987987	29-MAR-59	Dalias 980, Higueras, MX	4	Administración	987654321
Botello, Jaime E.	888665555	10-NOV-27	Sorgo 450, Higueras, MX	1	Dirección	888665555

EMPLEADO

NOMBREE	NSS	FECHAN	DIRECCIÓN	NÚMEROD
Silva, José B.	123456789	09-ENE-55	Fresnos 731, Higueras, MX	5
Vizcarra, Federico T.	333445555	08-DIC-45	Valle 638, Higueras, MX	5
Zapata, Alicia J.	999887777	19-JUL-58	Castillo 3321, Sucre, MX	4
Valdés, Jazmín S.	987654321	20-JUN-31	Bravo 291, Belén, MX	4
Nieto, Ramón K.	666884444	15-SEP-52	Espiga 957, Heras, MX	5
Esparza, Josefa A.	453453453	31-JUL-62	Rosas 5631, Higueras, MX	5
Jabbar, Ahmed V.	987987987	29-MAR-59	Dalias 980, Higueras, MX	4
Botello, Jaime E.	888665555	10-NOV-27	Sorgo 450, Higueras, MX	1

DEPARTAMENTO

NOMBRED	NÚMEROD	NSSGTED
Investigación	5	333445555
Administración	4	987654321
Dirección	1	888665555

Reducción de los valores redundantes

◆ Anomalías de actualización

- Anomalías de inserción
- Anomalías de eliminación
- Anomalías de modificación

◆ Pauta 2

» Diseñe los esquemas de las relaciones de modo que no haya anomalías de inserción, eliminación o modificación en las relaciones. Si hay anomalías señálelas con claridad a fin de que los programas que actualicen la BD operen correctamente.

Valores nulos en las tuplas

◆ Posibles problemas

- Desperdicio de espacio
- Dificultad para entender el significado
- Aplicación de funciones agregadas (count,sum)
- Múltiples interpretaciones

◆ Pauta 3

- » Hasta donde sea posible, evite incluir en una relación atributos cuyos valores pueden ser nulos. Si no es posible, asegúrese de que se apliquen solo en casos excepcionales y no a la mayoría de las tuplas de una relación.

Tuplas erróneas

◆ Ejemplo - Se aplica proyección a EMP-PROY

EMP_PROY

NSS	NÚMEROP	HORAS	NOMBREE	NOMBREPR	LUGARP
123456789	1	32.5	Silva, José B.	ProductoX	Belén
123456789	2	7.5	Silva, José B.	ProductoY	Sacramento
666884444	3	40.0	Nieto, Ramón K.	ProductoZ	Higueras
453453453	1	20.0	Esparza, Josefa A.	ProductoX	Belén
453453453	2	20.0	Esparza, Josefa A.	ProductoY	Sacramento
333445555	2	10.0	Vizcarra, Federico T.	ProductoY	Sacramento
333445555	3	10.0	Vizcarra, Federico T.	ProductoZ	Higueras
333445555	10	10.0	Vizcarra, Federico T.	Automatización	Santiago
333445555	20	10.0	Vizcarra, Federico T.	Reorganización	Higueras

LUGARES_EMP

NOMBREE	LUGARP
Silva, José B.	Belén
Silva, José B.	Sacramento
Nieto, Ramón K.	Higueras
Esparza, Josefa A.	Belén
Esparza, Josefa A.	Sacramento
Vizcarra, Federico T.	Sacramento
Vizcarra, Federico T.	Higueras
Vizcarra, Federico T.	Santiago

EMP_PROY1

NSS	NÚMEROP	HORAS	NOMBREPR	LUGARP
123456789	1	32.5	ProductoX	Belén
123456789	2	7.5	ProductoY	Sacramento
666884444	3	40.0	ProductoZ	Higueras
453453453	1	20.0	ProductoX	Belén
453453453	2	20.0	ProductoY	Sacramento
333445555	2	10.0	ProductoY	Sacramento
333445555	3	10.0	ProductoZ	Higueras
333445555	10	10.0	Automatización	Santiago
333445555	20	10.0	Reorganización	Higueras

Tuplas erróneas

- ◆ - Se aplica join natural a EMP-PROY1 y LUGARES-EMP

NSS	NÚMEROP	HORAS	NOMBREPR	LUGARP	NOMBREE
123456789	1	32.5	ProductoX	Belén	Silva, José B.
* 123456789	1	32.5	ProductoX	Belén	Esparza, Josefa A.
123456789	2	7.5	ProductoY	Sacramento	Silva, José B.
* 123456789	2	7.5	ProductoY	Sacramento	Esparza, Josefa A.
* 123456789	2	7.5	ProductoY	Sacramento	Vizcarra, Federico T.
666884444	3	40.0	ProductoZ	Higueras	Nieto, Ramón K.
* 666884444	3	40.0	ProductoZ	Higueras	Vizcarra, Federico T.
* 453453453	1	20.0	ProductoX	Belén	Silva, José B.
453453453	1	20.0	ProductoX	Belén	Esparza, Josefa A.
			⋮		
333445555	10	10.0	Automatización	Santiago	Vizcarra, Federico T.
* 333445555	20	10.0	Reorganización	Higueras	Nieto, Ramón K.
333445555	20	10.0	Reorganización	Higueras	Vizcarra, Federico T.

Tuplas erróneas

◆ Pauta 4

»Diseñe los esquemas de modo que puedan reunirse por condición de igualdad sobre atributos claves, para garantizar que no se formen tuplas erróneas.

Resumen

◆ Problemas a evitar

- Anomalías en inserción, modificación y eliminación de tuplas por *redundancia*
- Desperdicio de espacio y dificultad para operaciones por *valores nulos*
- Generación de datos erróneos por joins hechos *relacionando mal* las relaciones.

◆ Entonces se presentarán...

- Conceptos y teorías formales para detectar y evitar estos problemas.

Dependencias Funcionales

- Dependencias Funcionales - Definición
- Clausura de F
- Reglas de inferencia para las dfs
- Clausura de X
- Equivalencia de conjuntos de dfs
- Conjunto minimal de dfs

Dependencias Funcionales

◆ Definición

- Una df $X \rightarrow Y$, entre 2 conjuntos de atributos X e Y que son subconjuntos de R especifica una restricción sobre las posibles tuplas que formarían una instancia r de R . La restricción dice que, **para 2 tuplas cualesquiera t_1 y t_2 de r tales que $t_1[X]=t_2[X]$, debemos tener también $t_1[Y]=t_2[Y]$.**
- Observar:
 - » Si X es una clave candidata de R , entonces $X \rightarrow Y$ para cualquier subconjunto de atributos Y de R .
 - » Si $X \rightarrow Y$ en R , esto no nos dice si $Y \rightarrow X$ en R o no.
- Las dfs son propiedades de la semántica de los atributos.
- En el ejemplo de EMP_PROY, se cumplen:
 - » $NSS \rightarrow NOMBREE$, $NUMEROP \rightarrow \{NOMBREPR, LUGARP\}$,
 $\{NSS, NUMEROP\} \rightarrow HORAS$

Ejercicio 1 – Deducir atributos y dfs.

Una empresa de alquiler de vehículos desea implementar una base de datos con la información de su negocio. Se tienen **vehículos** identificados por su **numero de matrícula** y de los que se conoce su **marca**, **color**, **modelo** y **año**. También se tienen **clientes** identificados por su **número de cédula de identidad** y de los que se conoce su **nombre**, **dirección** y **teléfono**. Un **contrato** de alquiler de vehículo está identificado por un **número de contrato** y se realiza en una **fecha** dada entre un **cliente** y un **vehículo**, registrándose el **periodo** de alquiler en días y el **precio** del servicio. Se considera que en una misma fecha no se puede alquilar más de una vez el mismo vehículo al mismo cliente.

Clausura de F - F+

◆ Definición

- **F** - conjunto de dfs que se especifican sobre el esquema relación R.
- **F+** - conjunto de todas las dfs que se cumplen en todas las instancias que satisfacen a F.

◆ Inferencia de dfs

– Ejemplo

$F = \{ \text{NSS} \rightarrow \{ \text{NOMBREE}, \text{FECHAN}, \text{DIRECCION}, \text{NUMEROD} \},$
 $\text{NUMEROD} \rightarrow \{ \text{NOMBRED}, \text{NSSGTED} \} \}$

Podemos inferir:

$\text{NSS} \rightarrow \{ \text{NOMBRED}, \text{NSSGTED} \}, \text{NUMEROD} \rightarrow \text{NOMBRED}$

Reglas de inferencia para las dfs

- ◆ **Reglas:** (siendo X, Y, W, Z conjuntos de atributos)
 - (RI1) reflexiva - Si $X \supseteq Y$, entonces $X \rightarrow Y$
 - (RI2) de aumento - $\{X \rightarrow Y\} \models XZ \rightarrow YZ$
 - (RI3) transitiva - $\{X \rightarrow Y, Y \rightarrow Z\} \models X \rightarrow Z$
 - (RI4) descomposición - $\{X \rightarrow YZ\} \models X \rightarrow Y$
 - (RI5) unión - $\{X \rightarrow Y, X \rightarrow Z\} \models X \rightarrow YZ$
 - (RI6) pseudotransitiva - $\{X \rightarrow Y, WY \rightarrow Z\} \models WX \rightarrow Z$
- ◆ **Reglas de Armstrong: RI1 a RI3**
 - Minimales: Las demás se pueden derivar a partir de estas tres.

Clausura de X bajo F - X^+

◆ Definición

- X^+ es el conjunto de atributos determinados funcionalmente por X

Clausura de X bajo F - X^+

♦ Algoritmo - Determinar X^+ bajo F

$X^+ := X$

repetir

 viejo $X^+ := X^+$;

 para cada df $Y \rightarrow Z$ en F hacer

 si $Y \subseteq X^+$ entonces $X^+ := X^+ \cup Z$;

hasta que (viejo $X^+ = X^+$);

Clausura de X bajo F - X+

◆ Ejemplo

Dado EMP_PROY(NSS, NUMEROP, HORAS, NOMBREE, NOMBREPR, LUGARP)

F = { NSS → NOMBREE
NUMEROP → NOMBREPR, LUGARP
NSS, NUMEROP → HORAS }

podemos calcular:

{ NSS }+ = {
NSS,
NOMBREE
}

{ NUMEROP }+ = {
NUMEROP,
NOMBREPR,
LUGARP
}

Clausura de X bajo F - X+

◆ Ejemplo (Cont)

$$\{ \text{NSS, NUMEROP} \}^+ = \{$$

NSS,
NUMEROP,
NOMBREE,
NOMBREPR,
LUGARP,
HORAS

$$\}$$

- Observar que no es simplemente la union de las clausuras de los elementos del conjunto.

Ejercicio: Clausuras de Atributos

- ◆ **Hallar la clausura de los siguientes conjuntos de atributos**
 - $\{\text{nro_mat}\}$, $\{\text{nro_mat}, \text{ci_cli}\}$, $\{\text{nro_contrato}\}$,
 $\{\text{marca}\}$, $\{\text{fecha}, \text{ci_cli}, \text{nro_mat}\}$

- ◆ **Dar alguna dependencia funcional que pertenezca a F^+ y no a F .**

Equivalencia de conjuntos de dfs

◆ Definición

- Dos conjuntos de dfs E y F son equivalentes si $E^+ = F^+$.

◆ Podemos decir...

- Todas las dfs en E se pueden inferir de F y todas las dfs en F se pueden inferir de E .
- E cubre a F y F cubre a E .

◆ Como determinamos si F cubre a E

- Para cada df $X \rightarrow Y \in E$, calculamos $X^+(F)$ y verificamos que X^+ incluya los atributos en Y .

Equivalencia de conjuntos de dfs

◆ Ejemplo:

- $F = \{AB \rightarrow C, B \rightarrow D, D \rightarrow GC, CG \rightarrow H\}$
- $F_1 = \{D \rightarrow H, B \rightarrow C, AD \rightarrow GH\}$
 - » F_1 cubre a F ?
 - » F cubre a F_1 ?
 - » F es equivalente a F_1 ?
- $F_2 = \{B \rightarrow D, D \rightarrow G, D \rightarrow C, CG \rightarrow H\}$
 - » Que pasa entre F_2 y F ?
 - » Qué pasa entre F_1 y F_2 ?
- Observar que F_2 es más “simple” que F . Dado F , siempre se puede encontrar un conjunto con estas características?.

Conjunto minimal de dfs

◆ **F es minimal sii**

- Toda df en F tiene un solo atributo a la derecha
- No podemos reemplazar ninguna df $X \rightarrow A \in F$ por una df $Y \rightarrow A$, donde $Y \subset X$, y seguir teniendo un conjunto de dfs equivalente a F
- No podemos quitar ninguna df de F y seguir teniendo un conjunto de dfs equivalente a F

Conjunto minimal de dfs

◆ Definición

- Un **cubrimiento minimal** de F es un conjunto minimal F_{\min} que es equivalente a F .

Encontrar un cubrimiento minimal

◆ Algoritmo

1. Hacer $G := F$;
2. Reemplazar cada df $X \rightarrow A_1, A_2, \dots, A_n$ en G por las n dfs $X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_n$;
3. Para cada df restante $X \rightarrow A$ en G
para cada atributo B que sea un elemento de X
{ calcular $(X - B)^+$ respecto a G ;
si $(X - B)^+$ contiene a A ,
reemplazar $X \rightarrow A$ por $(X - B) \rightarrow A$ en G };
4. Para cada df $X \rightarrow A$ en G
{ calcular X^+ respecto a $(G - (X \rightarrow A))$;
si X^+ contiene a A , eliminar $X \rightarrow A$ de G };

Encontrar un cubrimiento minimal

◆ Ejemplo:

- $F = \{AB \rightarrow C, B \rightarrow D, D \rightarrow GC, CG \rightarrow H\}$
- Paso 1: Cada dependencia que tiene varios atributos a la derecha, es sustituida por las dependencias a los atributos individuales.
 - » $F_1 = \{AB \rightarrow C, B \rightarrow D, D \rightarrow G, D \rightarrow C, CG \rightarrow H\}$
- Paso 2: Estudiamos atributos redundantes.
 - » $B^+ = \{B, D, G, C, H\}$ entonces $F_2 = \{B \rightarrow C, B \rightarrow D, D \rightarrow G, D \rightarrow C, CG \rightarrow H\}$
- Paso 3: Estudiamos dependencias redundantes.
 - » Con respecto a $F_2 - \{B \rightarrow C\}$, $B^+ = \{B, D, G, C, H\}$ entonces $F_3 = \{B \rightarrow D, D \rightarrow G, D \rightarrow C, CG \rightarrow H\}$ Minimal.