

Facultad de **Ingeniería**



Normalización

Tema V

Semestre 2022-1



Objetivo

El alumno comprenderá y aplicará los conceptos del proceso de Normalización de Bases de Datos con la finalidad de implementar mejores diseños, estableciendo un equilibrio entre los niveles de redundancia y desempeño establecidos en los requerimientos no funcionales de casos de estudio.



Propiedades:

- **No pueden existir dos relaciones que se llamen igual**
- **No pueden existir tuplas iguales**
- **No pueden existir atributos que tengan el mismo nombre**
- **No hay orden en tuplas ni en atributos**
- **Los valores de los atributos deben ser atómicos**



Definición

Concepto introducido por Frank Codd en su artículo “A relational model of data for large shared data banks”, motivado por las anomalías que se presentaban en los conjuntos de datos.



Definición

Técnica empleada en el diseño de bases de datos que busca reducir redundancias e inconsistencias en los datos -> Organización en los datos



Ventajas

- Datos más consistentes
- Reducir datos duplicados
(reducir tamaño)
- Integridad de la base de datos
- Tablas más pequeñas
- Buenas prácticas



Desventajas

- Hay que tener nociones sólidas del concepto para llevar a cabo con éxito este proceso
- Se van a presentar más datos para ser unidos (joins entre tablas), lo que puede tomar tiempos considerables en ciertos casos -> Rendimiento



Desventajas

- **Las tablas contendrán códigos en vez de datos reales, lo que implica buscar ese código en otra(s) tabla(s)**



Desempeño

SalesStaff						
EmployeeID	SalesPerson	SalesOffice	OfficeNumber	Customer1	Customer2	Customer3
1003	Mary Smith	Chicago	312-555-1212	Ford	GM	
1004	John Hunt	New York	212-555-1212	Dell	HP	Apple
1005	Martin Hap	Chicago	312-555-1212	Boeing		

¿Qué problema(s) notan?



Desempeño

SalesStaff						
<u>EmployeeID</u>	<u>SalesPerson</u>	<u>SalesOffice</u>	<u>OfficeNumber</u>	<u>Customer1</u>	<u>Customer2</u>	<u>Customer3</u>
1003	Mary Smith	Chicago	312-555-1212	Ford	GM	
1004	John Hunt	New York	212-555-1212	Dell	HP	Apple
1005	Martin Hap	Chicago	312-555-1212	Boeing		

Los
Angeles

¿Qué problema(s) pueden pasar
cuando se agregue información?



Desempeño

SalesStaff						
<u>EmployeeID</u>	<u>SalesPerson</u>	<u>SalesOffice</u>	<u>OfficeNumber</u>	<u>Customer1</u>	<u>Customer2</u>	<u>Customer3</u>
1003	Mary Smith	Chicago	312-555-1212	Ford	GM	
1004	John Hunt	New York	212-555-1212	Dell	HP	Apple
1005	Martin Hap	Chicago	312-555-1212	Boeing		

¿Qué problema(s) pueden surgir al actualizar información?



Desempeño

SalesStaff						
EmployeeID	SalesPerson	SalesOffice	OfficeNumber	Customer1	Customer2	Customer3
1003	Mary Smith	Chicago	312-555-1212	Ford	GM	
1004	John Hunt	New York	212-555-1212	Dell	HP	Apple
1005	Martin Hap	Chicago	312-555-1212	Boeing		

**¿Qué problema(s) pueden surgir
al borrar información?**



Desempeño

SalesStaff						
<u>EmployeeID</u>	<u>SalesPerson</u>	<u>SalesOffice</u>	<u>OfficeNumber</u>	<u>Customer1</u>	<u>Customer2</u>	<u>Customer3</u>
1003	Mary Smith	Chicago	312-555-1212	Ford	GM	
1004	John Hunt	New York	212-555-1212	Dell	HP	Apple
1005	Martin Hap	Chicago	312-555-1212	Boeing		

¿Qué problema(s) pueden surgir al buscar información?

Un grupo de repetición es un conjunto de valores que ocurren varias veces en un registro.



1FN

EmployeeID	Name	Project	Time
EN1-26	Sean O'Brien	30-452-T3, 30-457-T3, 32-244-T3	0.25, 0.40, 0.30
EN1-33	Amy Guya	30-452-T3, 30-382-TC, 32-244-T3	0.05, 0.35, 0.60
EN1-35	Steven Baranco	30-452-T3, 31-238-TC	0.15, 0.80
EN1-36	Elizabeth Roslyn	35-152-TC	0.90
EN1-38	Carol Schaaf	36-272-TC	0.75
EN1-40	Alexandra Wing	31-238-TC, 31-241-TC	0.20, 0.70



1FN

EmpID	Last Name	First Name	Project1	Time1	Project2	Time2	Project3	Time3
EN1-26	O'Brien	Sean	30-452-T3	0.25	30-457-T3	0.40	32-244-T3	0.30
EN1-33	Guya	Amy	30-452-T3	0.05	30-382-TC	0.35	32-244-T3	0.60
EN1-35	Baranco	Steven	30-452-T3	0.15	31-238-TC	0.80		
EN1-36	Roslyn	Elizabeth	35-152-TC	0.90				
EN1-38	Schaaf	Carol	36-272-TC	0.75				
EN1-40	Wing	Alexandra	31-238-TC	0.20	31-241-TC	0.70		

¿Cómo identificar la llave primaria?

¿Cómo identificar llaves candidatas?

**Una relación está en
1FN si no presenta
grupos de repetición y
cada columna contiene
valores atómicos.**

¿Qué vamos a conseguir al aplicar la 1FN?

- Identificar la PK de cada tabla
- Evitar atributos multivaluados o atributos que representan lo mismo
- Evitar duplicidad de registros



Ejemplo

Sea la tabla CLIENTE:

ID_Cliente	nombre	apellido	telefono
123	Juan	López	12123212, 43556786
547	Irma	Arriaga	56567364
490	Pablo	Juárez	12334657, 65784532

¿Cumple la 1FN?



Ejemplo

ID_Cliente	nombre	apellido	telefono1	telefono2
123	Juan	López	12123212	43556786
547	Irma	Arriaga	56567364	56567364
490	Pablo	Juárez	12334657	65784532

¿Cumple la 1FN?



Ejemplo

ID_Cliente	nombre	apellido	telefono
123	Juan	López	12123212
123	Juan	López	43556786
547	Irma	Arriaga	56567364
490	Pablo	Juárez	12334657
490	Pablo	Juárez	65784532

**¿Cumple la 1FN?
Sí, pero...**



Ejemplo

ID_Cliente	nombre	apellido
123	Juan	López
547	Irma	Arriaga
490	Pablo	Juárez

ID_Cliente	telefono
123	12123212
123	43556786
547	56567364
490	12334657
490	65784532

¿Cumple la 1FN?



Ejercicio

Partiendo de:

DNI	Nombre	Codigo_Tienda	Direccion_Tienda	turno	fecha
33445566	Paola Martin	100A	Transmisiones Miliares 70	M	02/01/2020
44552345	Laura Sanz	100A	Transmisiones Miliares 70	M	02/01/2020
86923456	Daniel Diaz	100A	Transmisiones Miliares 70	T	02/01/2020
33445566	Paola Martin	200B	Periférico Norte 80	T	03/01/2020
12234456	Emiliano López	300C	Av. Universidad 3000	M	03/01/2020
45678367	Francisco Monte	200B	Periférico Norte 80	M	03/01/2020
12234456	Emiliano López	300C	Av. Universidad 3000	M	04/01/2020
45678367	Francisco Monte	100A	Transmisiones Miliares 70	M	04/01/2020
44552345	Laura Sanz	100A	Transmisiones Miliares 70	T	04/01/2020
33445566	Paola Martin	200B	Periférico Norte 80	M	05/01/2020



Ejercicio

PK: {DNI}

Creamos relación con dni y nombre

Creamos otra con dni,

codigo_tienda, dirección tienda,

turno, fecha

Relaciones parciales:

Una dependencia funcional (DF) que ocurre en una relación es parcial cuando la eliminación de uno de los atributos determinantes genera una DF que sigue ocurriendo en la relación



Relaciones parciales:

$\{X, Y\} \rightarrow Z$

$X \rightarrow Z$

**Lo anterior implica que Z
depende parcialmente de $\{X, Y\}$**

Sea la tabla estudiante_Proyecto

id_Estudiante	id_Proyecto	nom_Est	nom_Proy
S01	P10	Laura	Bases de datos
S02	P32	Juan	Clúster

- Identificar PK y candidatas
- Validar las dependencias con los atributos restantes



Sea la tabla estudiante_Proyecto

id_Estudiante	id_Proyecto	nom_Est	nom_Proy
S01	P10	Laura	Bases de datos
S02	P32	Juan	Clúster

¿Qué tipo de dependencia hay?

X → Y

{id_Estudiante, id_proyecto} →
nom_Est, nom_Proy



id_Estudiante	id_Proyecto	nom_Est	nom_Proy
S01	P10	Laura	Bases de datos
S02	P32	Juan	Clúster

X -> Y

{id_Estudiante, id_proyecto} ->

nom_Est, nom_Proy

id_Estudiante -> nom_Est

id_Proy -> nom_Proy



2FN

staffNo	branchNo	branchAddress	name	position	hoursPerWeek
S4555	B002	City Center Plaza, Seattle, WA 98122	Ellen Layman	Assistant	16
S4555	B004	16 – 14th Avenue, Seattle, WA 98128	Ellen Layman	Assistant	9
S4612	B002	City Center Plaza, Seattle, WA 98122	Dave Sinclair	Assistant	14
S4612	B004	16 – 14th Avenue, Seattle, WA 98128	Dave Sinclair	Assistant	10

A B

C

D

E

F

PK: {A,B}

{A,B} -> C, D, E, F

A -> D,E

B -> C

{A, B} -> F

1 FN: Cumple 1FN ya que no hay atributos multivaluados ni grupos de repetición.

2 FN: No cumple porque hay atributos que dependen parcialmente de la PK



Ejercicio

Solución:

staffNo	name	position
S4555	Ellen Layman	Assistant
S4612	Dave Sinclair	Assistant

staffNo	branchNo	hoursPerWeek
S4555	B002	16
S4555	B004	9
S4612	B002	14
S4612	B004	10

branchNo	branchAddress
B002	City Center Plaza, Seattle, WA 98122
B004	16 – 14th Avenue, Seattle, WA 98128



Notación en el proceso de normalización:

- Notación de dependencia funcional:

$\{A, B\} \rightarrow C$

$A \rightarrow C$

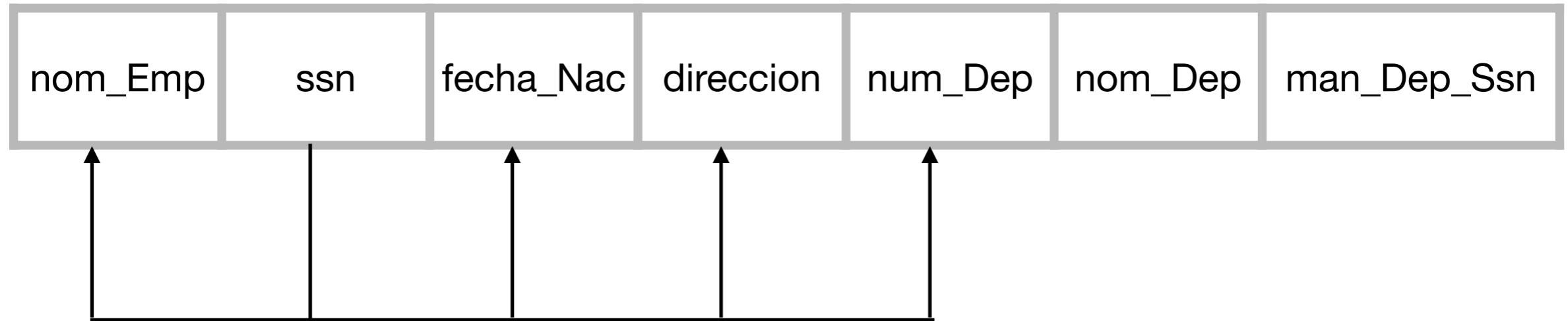


2FN

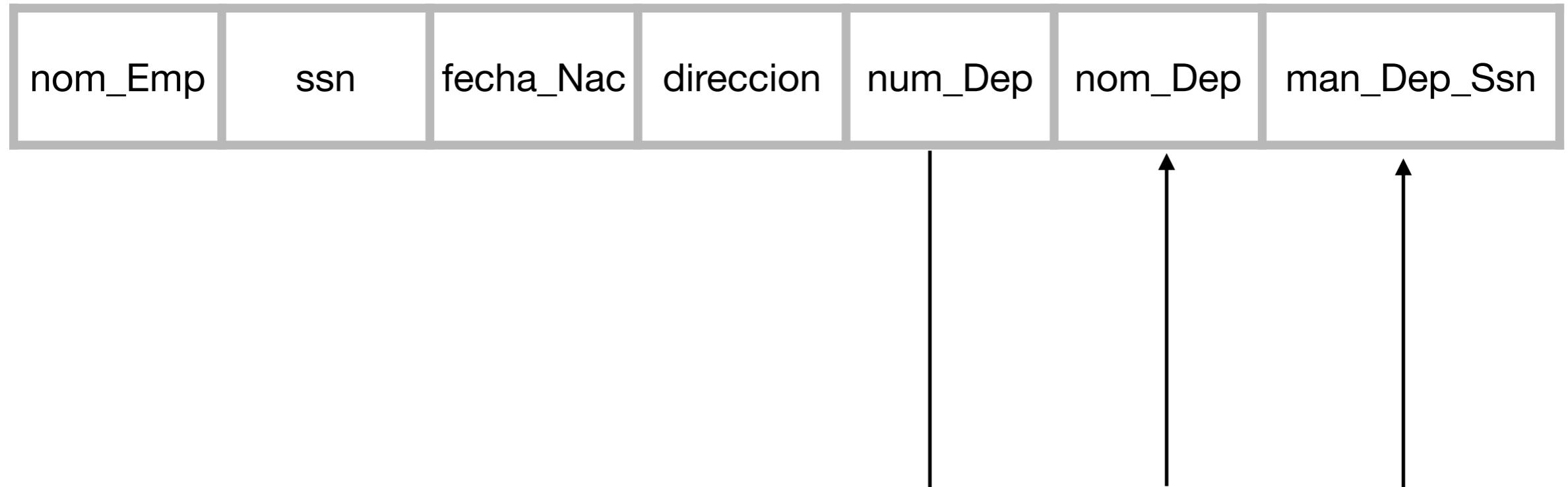
- Diagramas de dependencias:

nom_Emp	ssn	fecha_Nac	direccion	num_Dep	nom_Dep	man_Dep_Ssn
---------	-----	-----------	-----------	---------	---------	-------------

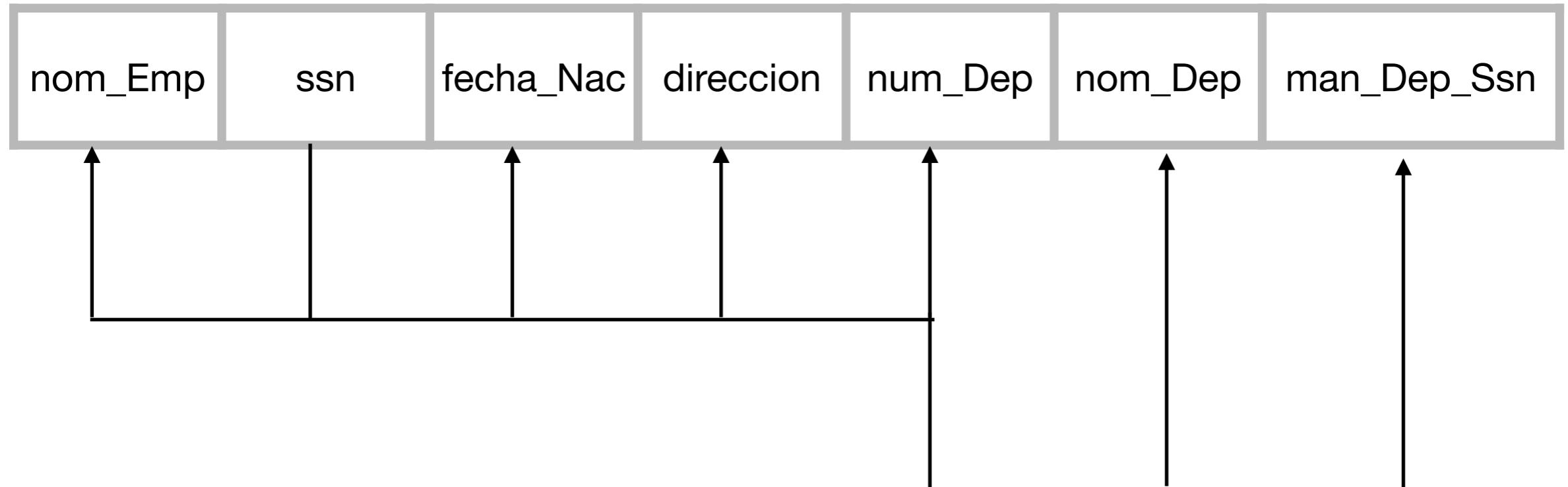
- Diagramas de dependencias:



- Diagramas de dependencias:



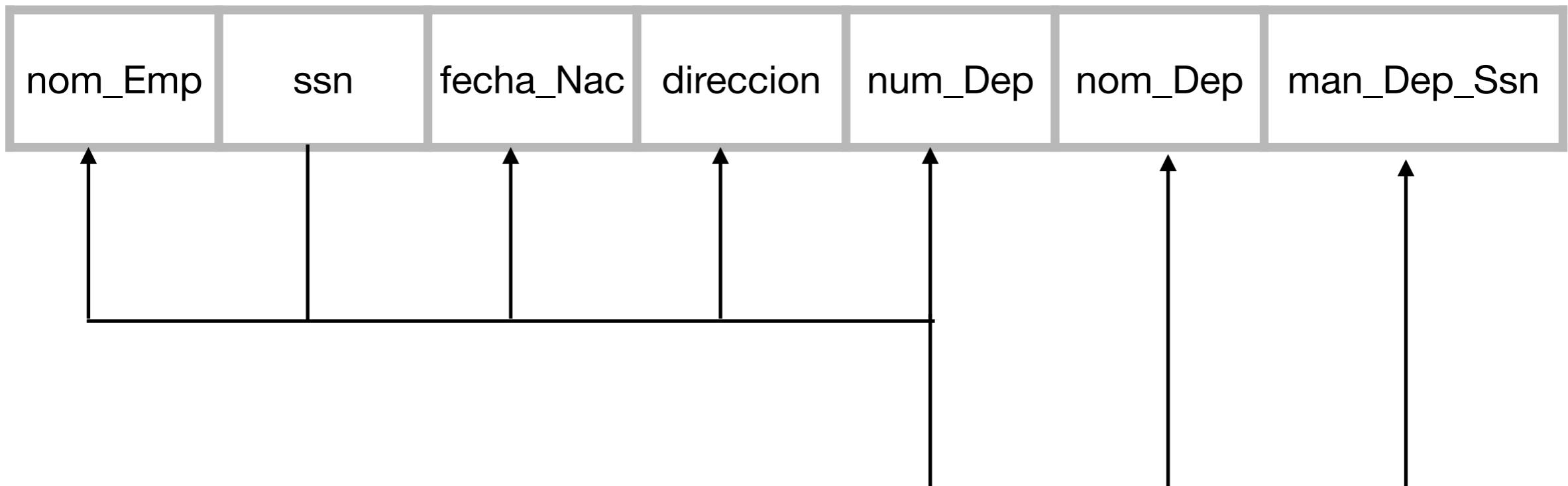
- Diagramas de dependencias:





B -> {A, C, D, E}

E -> {F, G}



Una tabla se encuentra en 2FN si y sólo si:

- **Se encuentra en 1FN**
- **Ninguno de sus atributos no-principales son funcionalmente dependientes en una parte de una llave primaria/candidata**



Ejemplo

Sea la tabla EMPLEADO:

ID_Emplead	ID_Dpto	Ubicacion	Nombre
EMP-01	DEP_BD	Yucatán	Guillermo
EMP-02	DEP_SD	CDMX	Fernando
EMP-03	DEP_CB	Guanajuato	Aldo
EMP-04	DEP_CT	Torreón	Carlos

$$\{X, Y\} \rightarrow \{Z, W\}$$

$$Y \rightarrow Z$$

$$X \rightarrow W$$



Ejemplo

Sea la tabla EMPLEADO:

¿Cumple la 1FN? -> Sí, ya que no hay valores multivaluados

¿Cumple la 2FN? No, porque existen dependencias funcionales parciales



Ejemplo

{id_Epleado, id_Depto} ->
{Ubicacion, Nombre}

id_Depto -> Ubicacion

id_Epleado -> Nombre



Ejemplo

Normalizando:

COLABORA

ID Empleado	ID Depto
EMP-01	DEP BD
EMP-02	DEP SD
EMP-03	DEP CB
EMP-04	DEP CT

ID Depto	Ubicacion
DEP BD	Yucatán
DEP SD	CDMX
DEP CB	Guanajuato
DEP CT	Torreón

DEPARTAMENTO

ID Empleado	Nombre
EMP-01	Guillermo
EMP-02	Fernando
EMP-03	Aldo
EMP-04	Carlos

EMPLEADO



Ejemplo

Sea la tabla EMPLEADO:

<u>ID Empleado</u>	<u>ID Depto</u>	Horas
EMP-01	DEP RD	10
EMP-02	DEP SD	5
EMP-03	DEP CB	15
EMP-04	DEP CT	13
EMP-01	DEP SD	7

$$\{x, y\} \rightarrow z$$

¿Cumple la 1FN? \rightarrow Sí, ya que todos los valores son atómicos y no hay grupos de repetición



Ejemplo

Sea la tabla EMPLEADO:

¿Cumple la 1FN? -> Sí, ya que todos los valores son atómicos y no hay grupos de repetición

¿Cumple la 2FN? -> Sí, porque no hay dependencias funcionales parciales

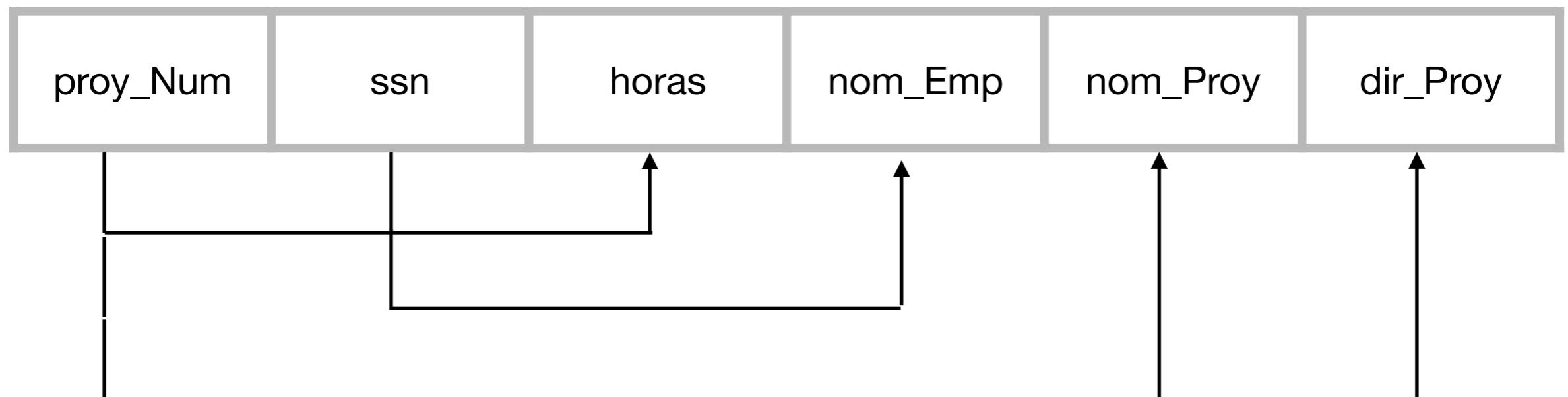


2FN

- Diagramas de dependencias:

proy_Num	ssn	horas	nom_Emp	nom_Proy	dir_Proy
----------	-----	-------	---------	----------	----------

- Diagramas de dependencias:





Ejemplo

¿está en 2FN?

Employee Id	Department Id	Office Location
1EDU001	ED-T1	Pune
1EDU002	ED-S2	Bengaluru
1EDU003	ED-M1	Delhi
1EDU004	ED-T3	Mumbai



Ejercicio

Partiendo de:

DNI	Nombre	Codigo_Tienda	Direccion_Tienda	turno	fecha
33445566	Paola Martin	100A	Transmisiones Miliares 70	M	02/01/2020
44552345	Laura Sanz	100A	Transmisiones Miliares 70	M	02/01/2020
86923456	Daniel Diaz	100A	Transmisiones Miliares 70	T	02/01/2020
33445566	Paola Martin	200B	Periférico Norte 80	T	03/01/2020
12234456	Emiliano López	300C	Av. Universidad 3000	M	03/01/2020
45678367	Francisco Monte	200B	Periférico Norte 80	M	03/01/2020
12234456	Emiliano López	300C	Av. Universidad 3000	M	04/01/2020
45678367	Francisco Monte	100A	Transmisiones Miliares 70	M	04/01/2020
44552345	Laura Sanz	100A	Transmisiones Miliares 70	T	04/01/2020
33445566	Paola Martin	200B	Periférico Norte 80	M	05/01/2020



3FN

Dependencias transitivas:

$X \rightarrow Y$

$Y \rightarrow Z$

$\text{¿}X \rightarrow Z?$

Relaciones transitivas:

Supongamos que los estudiantes solo pueden estar matriculados en un solo curso y supongamos que los profesores solo pueden dar un curso

Relaciones transitivas:

Entonces tenemos que ID_Estudiante determina a Curso_Tomado y el Curso_Tomado determina a Profesor_Asignado, indirectamente podemos saber a través del ID_estudiante el Profesor_Asignado.

Una tabla se encuentra en 3FN si y sólo si:

- Se encuentra en 2FN
- Cualquier atributo no-principal de la tabla sea no transitivamente dependiente de cada clave candidata de la misma



3FN

Student Id	Student Name	Subject Id	Subject	Address
1DT15ENG01	Alex	15CS11	SQL	Goa
1DT15ENG02	Barry	15CS13	JAVA	Bengaluru
1DT15ENG03	Clair	15CS12	C++	Delhi
1DT15ENG04	David	15CS13	JAVA	Kochi

$$A \rightarrow \{B, C, D, E\}$$

$$C \rightarrow D$$



3FN

Identificar PK: student_ID

Identificar candidatas: no hay

Identificar TODAS las dependencias:

$A \rightarrow \{B, C, D, E\}$

$C \rightarrow D$



3FN

Tanto C como D son atributos no-principales, por lo que la tabla no esta en 3FN

$A \rightarrow \{B, C, D, E\}$

$C \rightarrow D$



Quitamos la dependencia transitiva

Student Id	Student Name	Subject Id	Subject	Address
1DT15ENG01	Alex	15CS11	SQL	Goa
1DT15ENG02	Barry	15CS13	JAVA	Bengaluru
1DT15ENG03	Clair	15CS12	C++	Delhi
1DT15ENG04	David	15CS13	JAVA	Kochi



Sin perder información ni la relación con la tabla base

Subject Id	Subject
15CS11	SQL
15CS13	JAVA
15CS12	C++
15CS13	JAVA



3FN

FK

Student Id	Student Name	Subject Id	Address
1DT15ENG01	Alex	15CS11	Goa
1DT15ENG02	Barry	15CS13	Bengaluru
1DT15ENG03	Clair	15CS12	Delhi
1DT15ENG04	David	15CS13	Kochi

Separar la transitividad

PK

Subject Id	Subject
15CS11	SQL
15CS13	JAVA
15CS12	C++
15CS13	JAVA



Ejercicio

Partiendo de:

staffNo	name	position	salary	branchNo	branchAddress	telNo
S1500	Tom Daniels	Manager	46000	B001	8 Jefferson Way, Portland, OR 97201	503-555-3618
S0003	Sally Adams	Assistant	30000	B001	8 Jefferson Way, Portland, OR 97201	503-555-3618
S0010	Mary Martinez	Manager	50000	B002	City Center Plaza, Seattle, WA 98122	206-555-6756
S3250	Robert Chin	Supervisor	32000	B002	City Center Plaza, Seattle, WA 98122	206-555-6756
S2250	Sally Stern	Manager	48000	B004	16 – 14th Avenue, Seattle, WA 98128	206-555-3131
S0415	Art Peters	Manager	41000	B003	14 – 8th Avenue, New York, NY 10012	212-371-3000

PK: {staffNo}

PK: {branchNo, staffNo}



Ejercicio

staffNo	name	position	salary	branchNo	branchAddress	telNo
S1500	Tom Daniels	Manager	46000	B001	8 Jefferson Way, Portland, OR 97201	503-555-3618
S0003	Sally Adams	Assistant	30000	B001	8 Jefferson Way, Portland, OR 97201	503-555-3618
S0010	Mary Martinez	Manager	50000	B002	City Center Plaza, Seattle, WA 98122	206-555-6756
S3250	Robert Chin	Supervisor	32000	B002	City Center Plaza, Seattle, WA 98122	206-555-6756
S2250	Sally Stern	Manager	48000	B004	16 – 14th Avenue, Seattle, WA 98128	206-555-3131
S0415	Art Peters	Manager	41000	B003	14 – 8th Avenue, New York, NY 10012	212-371-3000

$A \rightarrow \{B, C, D, E, F, G\}$

$A \rightarrow \{B, C, D, E\}$

$E \rightarrow \{F, G\}$

1FN: Sí cumple , ya que no hay atributos multivaluados y/o grupos de repetición

2FN: Sí cumple, ya que la PK es simple

3FN: No cumple, ya que hay transitividad entre atributos

$A \rightarrow \{B, C, D, E\}$

$E \rightarrow \{F, G\}$



Ejercicio

Ordenes

Id_orden	Fecha	Id_cliente	Nom_cliente	Estado	Num_art	nom_art	cant	Precio
2301	23/02/11	101	Martin	Caracas	3786	Red	3	35,00
2301	23/02/11	101	Martin	Caracas	4011	Raqueta	6	65,00
2301	23/02/11	101	Martin	Caracas	9132	Paq-3	8	4,75
2302	25/02/11	107	Herman	Coro	5794	Paq-6	4	5,00
2303	27/02/11	110	Pedro	Maracay	4011	Raqueta	2	65,00
2303	27/02/11	110	Pedro	Maracay	3141	Funda	2	10,00

A -> {B, C, D, E, F, G, H, I}

A -> {B, C}

C -> {D, E}

F -> {G, I}

{A, F} -> H

Ejercicio

Ordenes

<u>Id_orden</u>	<u>Fecha</u>	<u>Id_cliente</u>	<u>Nom_cliente</u>	<u>Estado</u>	<u>Num_art</u>	<u>nom_art</u>	<u>cant</u>	<u>Precio</u>
2301	23/02/11	101	Martin	Caracas	3786	Red	3	35,00
2301	23/02/11	101	Martin	Caracas	4011	Raqueta	6	65,00
2301	23/02/11	101	Martin	Caracas	9132	Paq-3	8	4,75
2302	25/02/11	107	Herman	Coro	5794	Paq-6	4	5,00
2303	27/02/11	110	Pedro	Maracay	4011	Raqueta	2	65,00
2303	27/02/11	110	Pedro	Maracay	3141	Funda	2	10,00

PK -> A (id_Orden)

R1 = (A, B, C, D, E)

Grado: 5

A -> {B,C,D,E}

A -> {B,C}

C -> {D,E}

R2 = (A, F, G, H, I)

Grado: 5

{A,F} -> {G,H,I}

F -> {G, I}

{A,F} -> H

1FN: ¿Cumple?

- ¿Existe pk? -> Sí

- ¿Hay atts multivaluados? -> No

- ¿Hay grupos de repetición? -> Sí

No está en 1FN



Ejercicio

Para 2FN:

Orden = (A, B, C, D, E)

Grado: 5

$A \rightarrow \{B, C, D, E\}$

$A \rightarrow \{B, C\}$

$C \rightarrow \{D, E\}$

R1 cumple 2FN ya que su PK es SIMPLE

R2 = (A, F, G, H, I)

Grado: 5

$\{A, F\} \rightarrow \{G, H, I\}$

$F \rightarrow \{G, I\}$

$\{A, F\} \rightarrow H$

No cumple 2FN, ya que hay dependencias parciales

Articulo = (F, G, I)

Grado: 3

$F \rightarrow \{G, I\}$

Detalle_Orden = (A, F, H)

Grado: 3

$\{A, F\} \rightarrow H$



Ejercicio

Para 3FN:

¿Cumple 3FN? No, ya que hay transitividad entre A y C

Orden = (A, B, C, D, E)

Grado: 5

A -> {B,C,D,E}

A -> {B, C}

C -> {D,E}

Orden = {A, B, C}

Grado: 3

A -> {B, C}

Cliente = {C, D, E}

Grado: 3

C -> {D,E}

Articulo = (F, G, I)

Grado: 3

F -> {G, I}

Detalle_Orden = (A, F, H)

Grado: 3

{A, F} -> H

¿Cumple 3FN? sí, ya que no hay transitividad

¿Cumple 3FN? sí, ya que no hay transitividad



Ejercicio

- La PK es id_Orden
- Dibujar el diagrama de dependencias
- Normalizar hasta 3FN



Ejercicio

num_Cuenta	id_Curso	nombre_Estudiante	nombre_Curso	calificacion	profesor	email_Profesor
12345	10	Javier	BD	8	Fernando	prof1@...
90908	10	Elena	BD	7	Fernando	prof1@...
12345	20	Javier	POO	9	Luis	prof2@...
90908	20	Elena	POO	10	Luis	prof2@...
67700	10	Sonia	BD	7	Fernando	prof1@...
45645	10	Luis	BD	5	Fernando	prof1@...
45645	30	Luis	SO	8	Laura	prof3@...



Ejercicio

Partiendo de:

DNI	Nombre	Codigo_Tienda	Direccion_Tienda	turno	fecha
33445566	Paola Martin	100A	Transmisiones Miliares 70	M	02/01/2020
44552345	Laura Sanz	100A	Transmisiones Miliares 70	M	02/01/2020
86923456	Daniel Diaz	100A	Transmisiones Miliares 70	T	02/01/2020
33445566	Paola Martin	200B	Periférico Norte 80	T	03/01/2020
12234456	Emiliano López	300C	Av. Universidad 3000	M	03/01/2020
45678367	Francisco Monte	200B	Periférico Norte 80	M	03/01/2020
12234456	Emiliano López	300C	Av. Universidad 3000	M	04/01/2020
45678367	Francisco Monte	100A	Transmisiones Miliares 70	M	04/01/2020
44552345	Laura Sanz	100A	Transmisiones Miliares 70	T	04/01/2020
33445566	Paola Martin	200B	Periférico Norte 80	M	05/01/2020

Ejercicio



Repaso



Recordando...

num_Cuenta	nombre	direccion	telefono
123	gabriel	ayuntamiento, santo domingo, 888c	55800, 88900
456	sandra	pino, santa ursula, 999c	55643
234	gerardo	delta, bondojo, 777c	5563, 5567, 5523
523	marisol	gamma, narvarte, 666c	5590
245	ivan	universidad, copilco, 555c	556, 559



Recordando...

num_Cuenta	nombre	calle	colonia	cp	telefono
123	gabriel	ayuntamiento	santo domingo	888c	55800, 88900
456	sandra	pino	santa ursula	999c	55643
234	gerardo	delta	bondojo	777c	5563, 5567, 5523
523	marisol	gamma	narvarte	666c	cinco cinco nueve...
245	ivan	universidad	copilco	555c	556, 559



Recordando...

num_Cuenta	nombre	calle	colonia	cp	telefono
123	gabriel	ayuntamiento	santo domingo	888c	55800, 88900
456	sandra	pino	santa ursula	999c	55643
234	gerardo	delta	bondojo	777c	5563, 5567, 5523
523	marisol	gamma	narvarte	666c	5590
245	ivan	universidad	copilco	555c	556, 559



Recordando...

num_Cuenta	nombre	calle	colonia	cp	telefono
123	gabriel	ayuntamiento	santo domingo	888c	55800
123	gabriel	ayuntamiento	santo domingo	888c	88900
456	sandra	pino	santa ursula	999c	55643
234	gerardo	delta	bondojo	777c	5563
234	gerardo	delta	bondojo	777c	5567
234	gerardo	delta	bondojo	777c	5523
523	marisol	gamma	narvarte	555c	5590
245	ivan	universidad	copilco	555c	556
245	ivan	universidad	copilco	555c	559



¿Está o no está en 1FN?



Recordando...

num_Cuenta	nombre	calle	colonia	cp
123	gabriel	ayuntamiento	santo domingo	888c
456	sandra	pino	santa ursula	999c
234	gerardo	delta	bondojo	777c
523	marisol	gamma	narvarte	666c
245	ivan	universidad	copilco	555c



Recordando...

num_Cuenta	telefono
123	55800
123	88900
456	55643
234	5563
234	5567
234	5523
523	5590
245	556
245	559



Normalizar...

num_Cuenta	nombre	calle	colonia	cp	telefono
123	gabriel	ayuntamiento	santo domingo	888c	55800
123	gabriel	ayuntamiento	santo domingo	888c	88900
456	sandra	pino	santa ursula	999c	55643
234	gerardo	delta	bondojo	777c	5563
234	gerardo	delta	bondojo	777c	5567
234	gerardo	delta	bondojo	777c	5523
523	marisol	gamma	narvarte	555c	5590
245	ivan	universidad	copilco	555c	556
245	ivan	universidad	copilco	555c	559



Normalizar...



Su objetivo es que las dependencias funcionales sólo dependan de la clave primaria o de cualquier clave candidata



Partimos de la relación $R(A, B, C, D)$ y las dependencias:

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

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

$D \rightarrow B$



FNBC

Separamos en dos relaciones:

R1(A, C, D)

R2(D, B)



Ejercicio

student_id	subject	professor
101	Java	P.Java
101	C++	P.Cpp
102	Java	P.Java2
103	C#	P.Chash
104	Java	P.Java

**¿En qué FN se encuentra?
Dependencias**



Dependencias originales:

$$\{A, B\} \rightarrow C$$

$$C \rightarrow B$$



Solución

Respuesta simple:

B va a estar en una nueva relación ($C \rightarrow B$), pero en las dependencias originales C es determinado por A en conjunto con B.

Por lo tanto, no puede darse que A por si solo, determine a C



Solución

Respuesta con tablas:
A -> C

student_id	professor
101	p.java
101	p.cpp
102	p.java2
103	p.hash
104	p.java

C -> B

professor	subject
p.java	java
p.cpp	c++
p.java2	java
p.hash	c#

No cumple
dependencia, ya
que A no implica
un único C

Cumple
dependencia



Solución

Dependencias (ojo en los números):

- 1) $\{A, B\} \rightarrow C$
- 2) $C \rightarrow B$

Descripción de la solución:

- Por definición, C y B forman una relación nueva (obtenido de 2)
- B ya no puede aparecer en 1. Al ya no aparecer en 1, no puede determinarse C



Solución

Dependencias (ojo en los números):

- 1) $\{A, B\} \rightarrow C$
- 2) $C \rightarrow B$

Descripción de la solución (cont.):

Obteniéndose como resultado:

A

$C \rightarrow B$



Solución

Representación en tablas:

student_id
101
101
102
103
104

professor	subject
p.java	java
p.cpp	C++
p.java2	java
p.chash	C#

A

C -> B

Solución



En el slide anterior, se aprecia que no hay relación alguna entre las dos relaciones. Hasta ahora hemos cumplido con el proceso de FNBC, pero como hemos mencionado, no debe perderse información ni la relación entre ella.

Solución



¿Solución? Enlazar la información.



Solución

Representación en tablas:

student_id	id_professor
101	1
101	2
102	3
103	4
104	1

id_professor	professor	subject
1	p.java	java
2	p.cpp	c++
3	p.java2	java
4	p.chash	c#

La pk está marcada con color. Ya hay forma de enlazar ambas relaciones



Solución

Representación en tablas:

student_id	id_professor
101	1
101	2
102	3
103	4
104	1

id_professor	professor	subject
1	p.java	java
2	p.cpp	c++
3	p.java2	java
4	p.chash	c#

No olvidar que les comenté que un profesor imparte sólo una asignatura, y una asignatura la imparten varios profesores, con lo que queda resuelto el ejercicio.



**Su objetivo es que no haya
dependencias múltiples.**



¿Cómo identificar dependencias múltiples?

- La tabla contiene al menos, tres atributos
- A determina valores múltiples de B
- A determina valores múltiples de C
- B y C son independientes



4FN

Sea $R(A, B, C)$

$A \rightarrow\rightarrow B$

$A \rightarrow\rightarrow C$

Entonces:

$R1(A, B)$ y $R2(A, C)$



4FN

s_id	course	hobby
1	Science	Cricket
1	Maths	Hockey
2	C#	Cricket
2	Php	Hockey

Tabla con información de alumnos, los cursos que toman y sus hobbies



4FN

R1

s_id	course
1	Science
1	Maths
2	C#
2	Php

R2

s_id	hobby
1	Cricket
1	Hockey
2	Cricket
2	Hockey



Su objetivo es asegurar que, cuando separamos tablas para eliminar redundancia, al hacer una operación de join, no se pierde información.

SUBJECT	LECTURER	CLASS
Mathematics	Alex	SEMESTER 1
Mathematics	Rose	SEMESTER 1
Physics	Rose	SEMESTER 1
Physics	Joseph	SEMESTER 2
Chemistry	Adam	SEMESTER 1

Si queremos agregar a alguien de tercer semestre, pero no sabemos qué materia a va a impartir... ¿Qué pasa?



5FN

		5NF	
SUBJECT	LECTURER	CLASS	LECTURER
Mathematics	Alex	SEMESTER 1	Alex
Mathematics	Rose	SEMESTER 1	Rose
Physics	Rose	SEMESTER 1	Rose
Physics	Joseph	SEMESTER 2	Joseph
Chemistry	Adam	SEMESTER 1	Adam

CLASS	SUBJECT
SEMESTER 1	Mathematics
SEMESTER 1	Physics
SEMESTER 1	Chemistry
SEMESTER 2	Physics



Denormalización

Técnica de optimización en bases de datos, en la cual, de manera premeditada, se agrega redundancia de información a una o más tablas.



Denormalización

Técnica de optimización en bases de datos, en la cual, de manera premeditada, se agrega redundancia de información a una o más tablas.



Denormalización

Métodos:

- Agregar columnas redundantes
- Agregar columnas derivadas
- Juntar tablas

Denormalización

EMPLOYEE	EMP_ID	EMP_NAME	ADDRESS	DEPT_ID	PROJ_ID
	100	Joseph	Clinton Town	10	206
	101	Rose	Fraser Town	20	205
	102	Mathew	Lakeside Village	10	206
	103	Stewart	Troy	30	204
	104	William	Holland	30	202

DEPARTMENT	DEPT_ID	DEPT_NAME
	10	Accounting
	20	Quality
	30	Design

**Se te solicita
generar un
reporte de los
empleados y el
nombre de su
departamento.
¿Qué harías?**



Denormalización

EMPLOYEE						
EMP_ID	EMP_NAME	ADDRESS	DEPT_ID	PROJ_ID	DEPT_NAME	
100	Joseph	Clinton Town	10	206	Accounting	
101	Rose	Fraser Town	20	205	Quality	
102	Mathew	Lakeside Village	10	206	Accounting	
103	Stewart	Troy	30	204	Design	
104	William	Holland	30	202	Design	

DEPARTMENT	
DEPT_ID	DEPT_NAME
10	Accounting
20	Quality
30	Design



Denormalización

Ventajas:

- Obtención rápida de información
- Se evita el uso de joins
- Consultas más fáciles de leer y elaborar



Denormalización

Desventajas:

- Específica a un requerimiento en particular
- Implica mayor espacio de almacenamiento
- Redundancia de información
- Dificultad para agregar y actualizar información
- Pérdida de información



Ejercicio

Código/ alumno	Nombre/ alumno	Especialidad	Código/ curso	Nombre_curso	Nombre/ docente	Oficina	curso
382145A	Luis Zuloaga	Industrial	MA123	Matemática 2	Carlos Arambulo	CB-214	U
382145A	Luis Zuloaga	Industrial	QU514	Física Química	Petra Rondinel	CB-110	U
382145A	Luis Zuloaga	Industrial	AU521	Descriptiva	Víctor Moncada	CB-120	W
360247k	Raúl Rojas	Sistemas	PA714	Investigación 1	Cesar Fernandez	SC-220	V
360247k	Raúl Rojas	Sistemas	MA123	Matemática 2	Carlos Arambulo	CB-214	V
360247k	Raúl Rojas	Sistemas	AU511	Dibujo	Víctor Moncada	CB-120	U