

Facultad de **Ingeniería**



Normalización

Tema V

Semestre 2024-1



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						
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 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						
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 buscar información?

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

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



1FN

EmpID	Last Name	First Name	Project	Time	Project	Time	Project	Time
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		



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		



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



- **Descomponer atributos compuestos**
- **Atributos con el mismo dominio**
- **Evitar atributos que tengan el mismo nombre**

¿Qué vamos a conseguir al aplicar la 1FN?

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



- **Evitar anomalías de actualización y pérdidas de información**



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?

¿Cómo conseguir 1FN en una tabla?

- Opción 1: Crear una nueva tabla, la cual contendrá el atributo multivaluado (PK)* y la llave primaria de la tabla que se está normalizando (FK)



Ejemplo

PK

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

FK

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

PK

¿Cumple la 1FN?

¿Cómo conseguir 1FN en una tabla?

- **Opción 2: Crear una llave primaria compuesta, con el atributo multivaluado y el atributo(s) más cercano a ser llave primaria**



Ejemplo

PK

PK

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?



Ejemplo

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



Ejemplo

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

- ¿Hay atributos multivaluados o grupos de repetición?
- ¿Existe llave primaria?



Ejemplo

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



Ejemplo

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

EN1-26	SEAN...	30-452-T3	0.25
EN1-26	SEAN...	30-457-T3	0.40
EN1-26	SEAN...	32-244-T3	0.30
EN1-33	AMY	30-452-T3	0.05
EN1-33	AMY	30-382-TC	0.35
EN1-33	AMY	32-244-T3	0.60
...



$DF : X \longrightarrow Y$

$$t_1 X = t_2 X$$

$$t_1 Y = t_2 Y$$

num_Cuenta	nombre	calificacion	academia	materia
1	Juan	7	BD	Bases de datos
2	Maria	6	Redes	Criptografía
3	Juan	7	BD	Bases de datos avanzadas
4	Maria	6	Redes	Arq. C/S
5	Carlos	8	Software	Sistemas embebidos
6	Karla	8	PM	Admon. de proyectos

num_cuenta -> nombre

nombre -> num_cuenta

num_cuenta -> calificacion

academia -> materia

materia -> academia

{num_cuenta, nombre} -> calificacion

nombre -> calificacion

{nombre, calificacion} -> academia

num_cuenta -> {nombre, calificacion}
{academia, materia} -> nombre
{num_cuenta, calificacion} ->
academia
nombre -> materia
{nombre, calificacion, academia} ->
num_cuenta

- Dependencia funcional trivial
- Dependencia funcional no trivial



Notación

Notación en el proceso de normalización:

- Notación de dependencia funcional:

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

$A \rightarrow C$



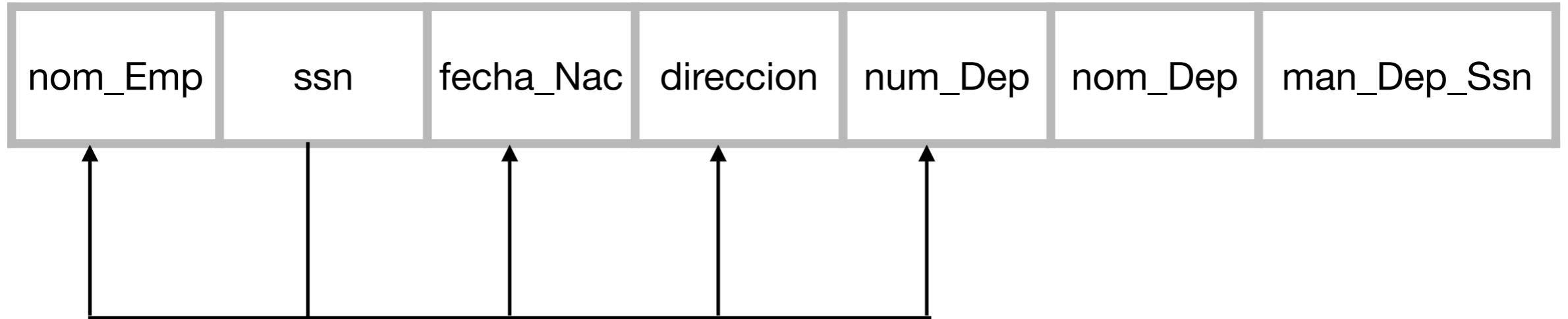
Notación

- Diagramas de dependencias:

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

Notación

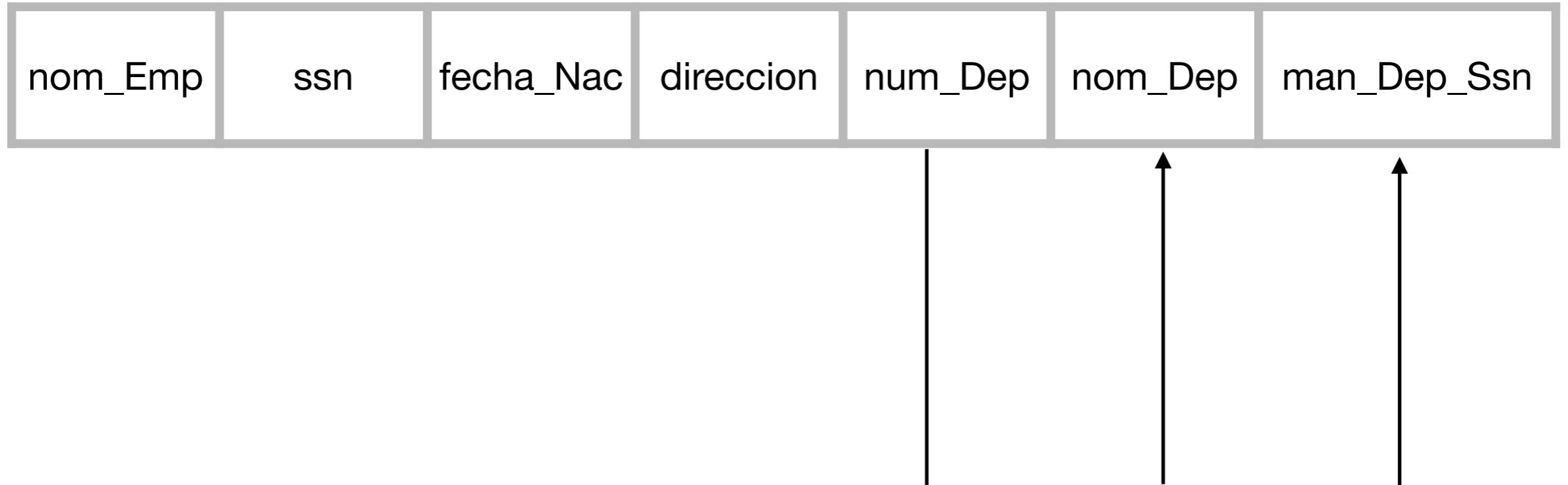
- Diagramas de dependencias:





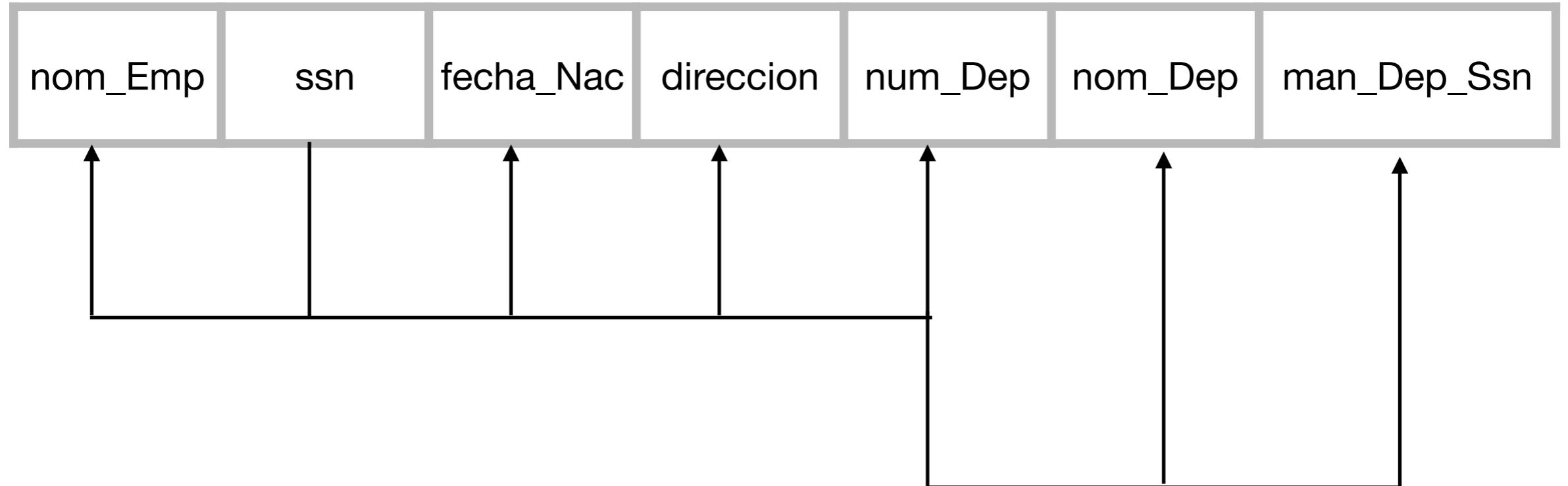
Notación

- Diagramas de dependencias:



Notación

- Diagramas de dependencias:



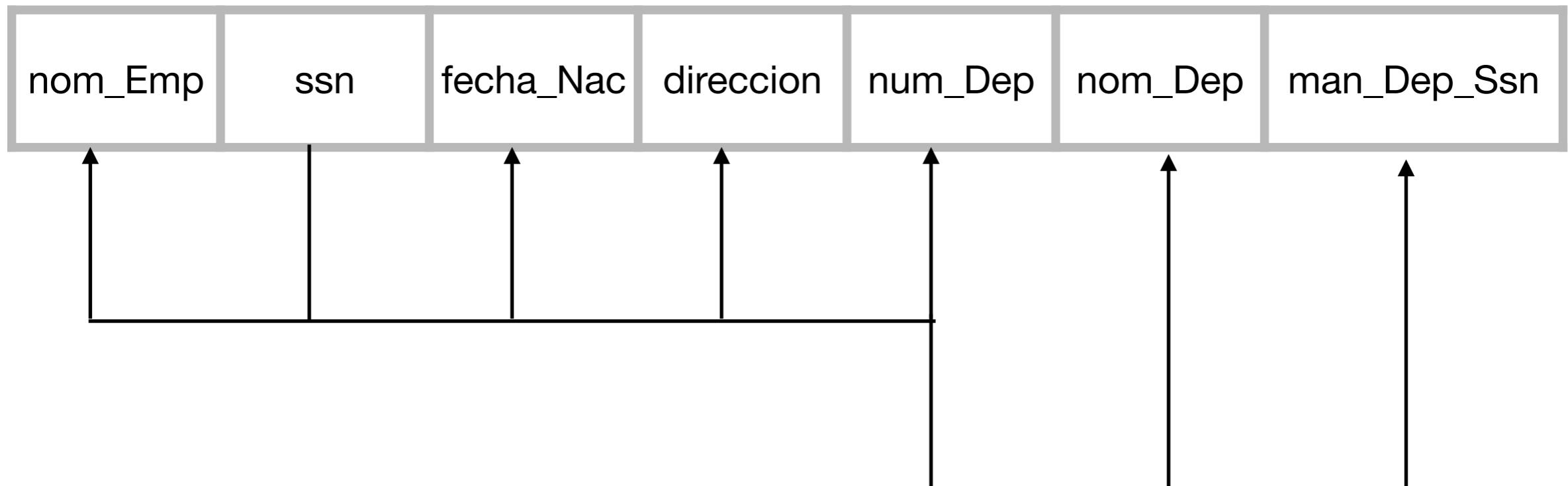


Notación



B → {A, C, D, E}

E → {F, G}





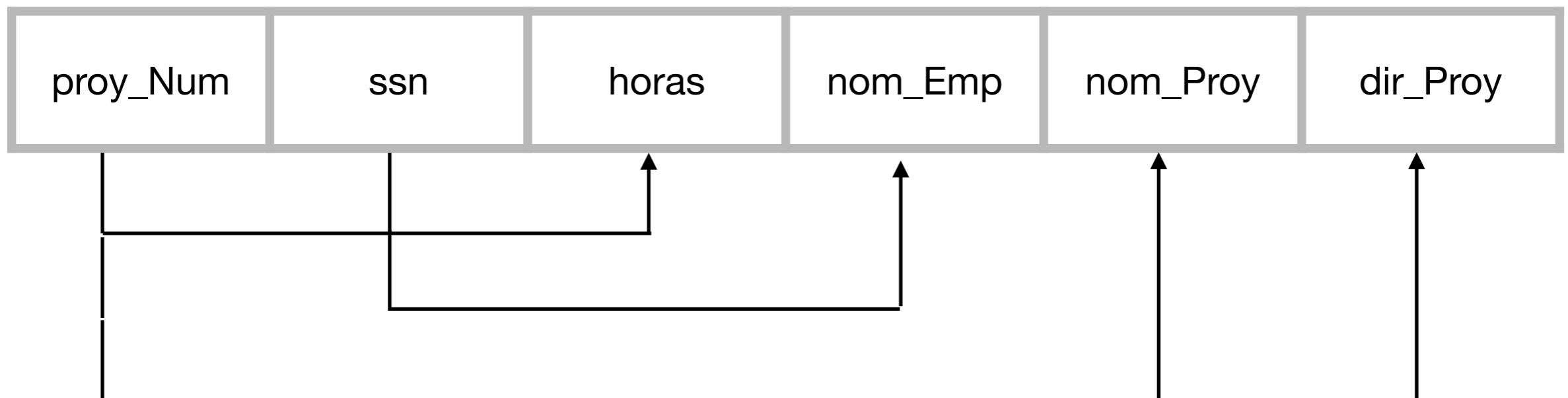
Ejemplo

- Indicar dependencias

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

Ejemplo

- Diagramas de dependencias:



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**

Dependencias 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



Dependencias parciales:

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

$$X \rightarrow Z$$

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



- **Llave**
- **Super llave (SK)**
- **Llave candidata (CK)**
- **Llave primaria (PK)**



2FN

Sea la relación estudiante:

nombre	calificacion	academia	materia
ana	8	datos	Bases de datos
carlos	6	redes	seguridad I
ana	8	datos	minería
carlos	6	redes	criptografia
rodrigo	9	hardware	Sistemas operativos

¿Existe algún posible identificador?



2FN

id_alumno	nombre	calificacion	academia	materia
1	ana	8	datos	Bases de datos
2	carlos	6	redes	seguridad I
3	ana	8	datos	minería
4	carlos	6	redes	criptografia
5	rodrigo	9	hardware	Sistemas



id_alumno	nombre	calificacion	academia	materia
1	ana	8	datos	Bases de datos
2	carlos	6	redes	seguridad I
3	ana	8	datos	minería
4	carlos	6	redes	criptografia
5	rodrigo	9	hardware	Sistemas

- **Encontrar SKs, CKs y PK**
- **¿Cuántas posibles SK's puedo tener en la relación?**



Closure set

La cerradura de un conjunto de atributos X es el conjunto de aquellos atributos que pueden ser funcionalmente dependientes de X .

$$X^+$$



Closure set

Procedimiento:

1) Agregar los atributos que componen a X como conjunto resultante de X^+



Procedimiento:

**2) Agregar aquellos atributos que
pueden ser funcionalmente
dependientes de los atributos que
ya estén contenidos en X^+**



Closure set

Procedimiento:

3) Repetir el paso 2 hasta que no puedan agregarse más atributos al conjunto X^+



Closure set - Ejemplo

Sea $R(A, B, C, D, E)$

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

Determinar:

A^+

B^+

AD^+

CD^+



Closure set - Ejemplo 2

Sea $R(A, B, C, D, E)$

DF: $\{A \rightarrow B, D \rightarrow E\}$

Determinar:

$$A^+$$

$$ABCDE^+$$

$$ACDE^+$$

$$BC^+$$

$$ABDE^+$$

$$ACD^+$$

¿Cuál(es) es SK? ¿Cuál(es) es CK?



Closure set

Muchas combinaciones para encontrar las posibles llaves candidatas -> La complejidad es proporcional a la cantidad de atributos en la tabla/relación



Closure set - Ejemplo 3

Sea $R(A, B, C, D, E)$

DF: $\{A \rightarrow B, D \rightarrow E\}$

Encontrar todas las CKs y la PK



id_alumno	nombre	calificacion	academia	materia
1	ana	8	datos	Bases de datos
2	carlos	6	redes	seguridad I
3	ana	8	datos	minería
4	carlos	6	redes	criptografia
5	rodrigo	9	hardware	Sistemas

- **Encontrar SKs, CKs y PK**
- **¿Cuántas posibles SK's puedo tener en la relación?**

Sea la tabla estudiante_Proyecto

id_Estudiante	id_Proyecto	nom_Est	nom_Proy
S01	P10	Laura	Bases de datos
S02	P32	Juan	Redes
S02	P40	Juan	Sistemas operativos

- 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	Redes
S02	P40	Juan	Sistemas operativos

¿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	Redes
S02	P40	Juan	Sistemas operativos

X -> Y

{id_Estudiante, id_proyecto} ->

nom_Est, nom_Proy

id_Estudiante -> nom_Est

id_Proy -> nom_Proy



Ejemplo

Sea la tabla EMPLEADO:

ID_Epleado	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

¿Dependencias?



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_BD	Guanajuato	Aldo
EMP-04	DEP_CT	Torreón	Carlos

X

Y

Z

W

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

$$Y \rightarrow Z$$

$$X \rightarrow W$$



Ejemplo

Sea la tabla EMPLEADO:

¿Cumple la 1FN? -> Sí, ya que no hay valores multivaluados ni grupos de rep.

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



Ejemplo

$\{id_Empleado, id_Depto\} \rightarrow \{Ubicacion, Nombre\}$

$\{id_Empleado, id_Depto\} \rightarrow \{\}$

$id_Depto \rightarrow Ubicacion$

$id_Empleado \rightarrow Nombre$



Ejemplo

Normalizando:

COLABORA

ID Empleado	ID Depto
EMP-01	DEP BD
EMP-02	DEP SD
EMP-03	DEP DB
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



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



Ejercicio 1_5

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



Ejercicio 1_5

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



3FN

Dependencias transitivas:

$X \rightarrow Y$

$Y \rightarrow Z$

$\text{¿}X \rightarrow Z\text{?}$

Relaciones transitivas:

Supongamos que Juan es hermano de Pedro. Además, Pedro es hermano de Javier, por lo que podemos concluir que Juan es hermano de Javier.



Relaciones transitivas:

Entonces tenemos que Juan determina a Pedro y Pedro determina a Javier, indirectamente podemos saber que hay relación entre Juan y Javier.

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



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 TODAS las dependencias:
Identificar candidatas:
Determinar el conjunto de APs:
Identificar PK:



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++



Ejercicio 5_2

- Encontrar todas las CKs
- Determinar el conjunto de APs
- Indicar la PK
- Normalizar hasta 3FN

Ejercicio 5_2



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



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 5_3

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



Ejercicio 5_3

- Encontrar todas las CKs
- Determinar el conjunto de APs
- Indicar la PK
- Normalizar hasta 3FN
- Generar el MER



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



Ejemplo

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**

Solución



¿Dependencias?



Solución

Dependencias:

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

$$C \rightarrow B$$

¿En que forma normal se encuentra?



Solución

A -> C

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

C -> B

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

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

**Cumple
dependencia**



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



4FN

**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.



Ejercicio 5_4

Sea la relacion Ventas(NumVenta, NoRegCarro, NoVendedor, NoCliente, Precio, Modelo, Marca, NombreCliente, NombreVendedor, FechaVenta) de un sistema de captura de órdenes. Funciones de dependencia:

Numventa -> {NoRegCarro, NoVendedor, NoCliente, Precio, FechaVenta }

NoRegCarro -> {Modelo, Marca}

NoVendedor -> NombreVendedor

NoCliente -> NombreCliente



Ejercicio 5_4

- ¿Cuál de los atributos puede ser empleado como llave primaria y por qué?
- ¿En qué forma normal se encuentra la relación ventas? Justificar respuesta
- Normalizar hasta 3FN