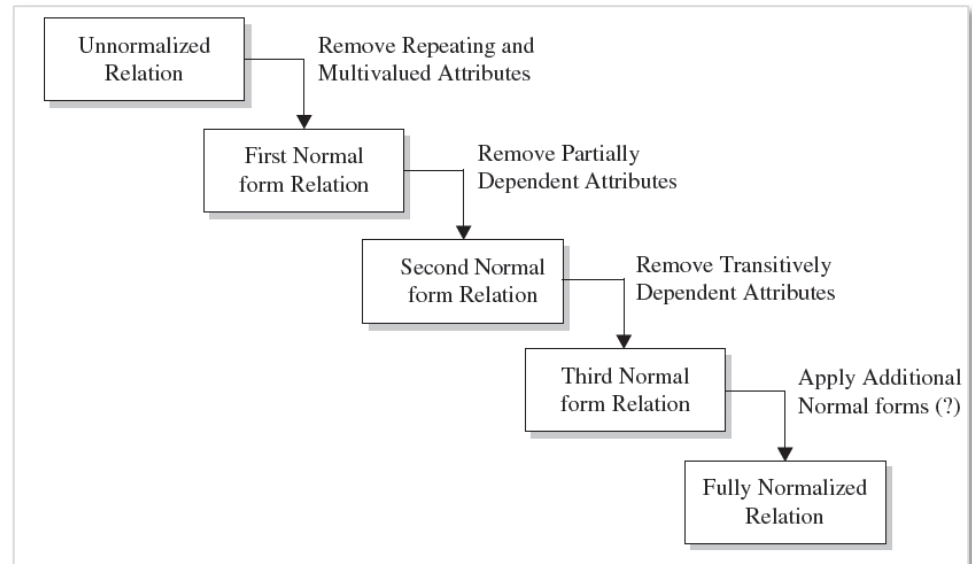




“La Ciencia sin Moral es Vana”

Objetivos:

- Conocer que es la normalización y qué papel desempeña en el proceso de diseño de bases de datos.
- Aplicar las formas normalizadas 1NF, 2NF, 3NF, BCNF Y 4NF.
- Utilizar la normalización y el modelo ER conjuntamente para producir un buen diseño de base de datos.
- Conocer las situaciones que requieren desnormalización.



Tema 6: Normalización de Bases de datos

Requerimiento: crear la base de datos de asignación de empleados a proyectos

A SAMPLE REPORT LAYOUT							
PROJECT NUMBER	PROJECT NAME	EMPLOYEE NUMBER	EMPLOYEE NAME	JOB CLASS	CHARGE/HOUR	HOURS BILLED	TOTAL CHARGE
15	Evergreen	103	June E. Arbough	Elec. Engineer	\$ 84.50	23.8	\$ 2,011.10
		101	John G. News	Database Designer	\$105.00	19.4	\$ 2,037.00
		105	Alice K. Johnson *	Database Designer	\$105.00	35.7	\$ 3,748.50
		106	William Smithfield	Programmer	\$ 35.75	12.6	\$ 450.45
		102	David H. Senior	Systems Analyst	\$ 96.75	23.8	\$ 2,302.65
				Subtotal			\$10,549.70
18	Amber Wave	114	Annelise Jones	Applications Designer	\$ 48.10	24.6	\$ 1,183.26
		118	James J. Frommer	General Support	\$ 18.36	45.3	\$ 831.71
		104	Anne K. Ramoras *	Systems Analyst	\$ 96.75	32.4	\$ 3,134.70
		112	Darlene M. Smithson	DSS Analyst	\$ 45.95	44.0	\$ 2,021.80
				Subtotal			\$ 7,171.47
22	Rolling Tide	105	Alice K. Johnson	Database Designer	\$105.00	64.7	\$ 6,793.50
		104	Anne K. Ramoras	Systems Analyst	\$96.75	48.4	\$ 4,682.70
		113	Delbert K. Joenbrood *	Applications Designer	\$48.10	23.6	\$ 1,135.16
		111	Geoff B. Wabash	Clerical Support	\$26.87	22.0	\$ 591.14
		106	William Smithfield	Programmer	\$35.75	12.8	\$ 457.60
				Subtotal			\$13,660.10
25	Starflight	107	Maria D. Alonzo	Programmer	\$ 35.75	24.6	\$ 879.45
		115	Travis B. Bawangi	Systems Analyst	\$ 96.75	45.8	\$ 4,431.15
		101	John G. News *	Database Designer	\$105.00	56.3	\$ 5,911.50
		114	Annelise Jones	Applications Designer	\$ 48.10	33.1	\$ 1,592.11
		108	Ralph B. Washington	Systems Analyst	\$ 96.75	23.6	\$ 2,283.30
		118	James J. Frommer	General Support	\$ 18.36	30.5	\$ 559.98
		112	Darlene M. Smithson	DSS Analyst	\$ 45.95	41.4	\$ 1,902.33
				Subtotal			\$17,559.82
				Total			\$48,941.09

Note: A * indicates the project leader.

¿Qué es la normalización de bases de datos?

- ❑ Es un proceso que consiste en aplicar una serie de reglas a las relaciones obtenidas tras el paso del modelo Entidad - Relación al modelo Relacional.
- ❑ Es un proceso que permite realizar un buen diseño de bases de datos, con tablas bien estructuradas para controlar las **redundancias de los datos** y para evitar las **anomalías** en los mismos.



Normalización de datos

La normalización de datos se refiere a modelar una base de datos relacional y saber cómo se genera una estructura relacional válida.

¿Por qué se normalizan las bases de datos?



- Para evitar la redundancia de los datos



- Para evitar problemas de actualización de los datos en las tablas



- Para proteger la integridad de los datos

Formas normales

- ❑ Las formas normales son aplicadas a las tablas de una base de datos.
- ❑ Decir que una base de datos está en la forma normal N es decir que todas sus tablas están en la forma normal N.

NORMAL FORMS	
NORMAL FORM	CHARACTERISTIC
First normal form (1NF)	Table format, no repeating groups, and PK identified
Second normal form (2NF)	1NF and no partial dependencies
Third normal form (3NF)	2NF and no transitive dependencies
Boyce-Codd normal form (BCNF)	Every determinant is a candidate key (special case of 3NF)
Fourth normal form (4NF)	3NF and no independent multivalued dependencies

Conversión a la primera forma normal

Una tabla está en primera forma normalizada (**1NF**) si:

1

- Todos los atributos de la clave primaria están definidos

2

- No existen **grupos repetidores** en la tabla

3

- Todos los atributos son dependientes de la clave primaria



Caso de ejemplo:

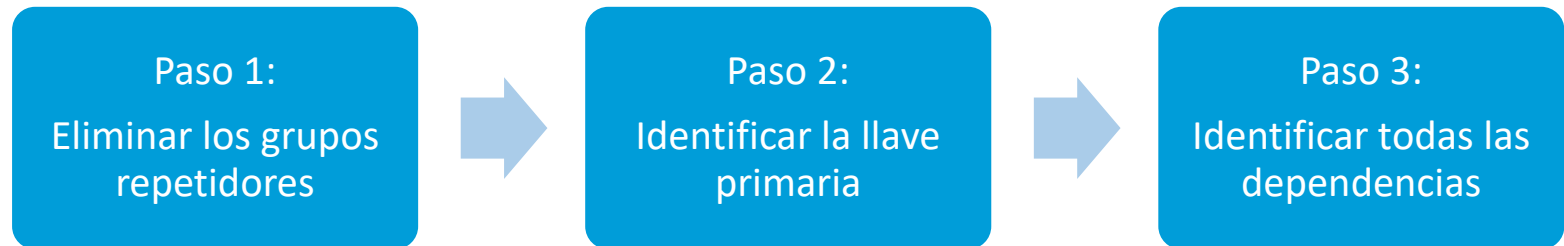
Base de datos para asignación de empleados a proyectos

PROJ_NUM	PROJ_NAME	EMP_NUM	EMP_NAME	JOB_CLASS	CHG_HOUR	HOURS
15	Evergreen	103	June E. Arbough	Elect. Engineer	84.50	23.80
		101	John G. News	Database Designer	105.00	19.40
		105	Alice K. Johnson *	Database Designer	105.00	35.70
		106	William Smithfield	Programmer	35.75	12.60
		102	David H. Senior	Systems Analyst	96.75	23.80
18	Amber Wave	114	Annelise Jones	Applications Designer	48.10	24.60
		118	James J. Frommer	General Support	18.36	45.30
		104	Anne K. Ramoras *	Systems Analyst	96.75	32.40
		112	Darlene M. Smithson	DSS Analyst	45.95	44.00
22	Rolling Tide	105	Alice K. Johnson	Database Designer	105.00	64.70
		104	Anne K. Ramoras	Systems Analyst	96.75	48.40
		113	Delbert K. Joenbrood *	Applications Designer	48.10	23.60
		111	Geoff B. Wabash	Clerical Support	26.87	22.00
		106	William Smithfield	Programmer	35.75	12.80
25	Starflight	107	Maria D. Alonzo	Programmer	35.75	24.60
		115	Travis B. Bawangi	Systems Analyst	96.75	45.80
		101	John G. News *	Database Designer	105.00	56.30
		114	Annelise Jones	Applications Designer	48.10	33.10
		108	Ralph B. Washington	Systems Analyst	96.75	23.60
		118	James J. Frommer	General Support	18.36	30.50
		112	Darlene M. Smithson	DSS Analyst	45.95	41.40



¿Se encuentra esta tabla en la primera forma normal (1NF)?

Solución 1NF





Paso 1: eliminar los grupos repetidores

PROJ_NUM	PROJ_NAME	EMP_NUM	EMP_NAME	JOB_CLASS	CHG_HOUR	HOURS
15	Evergreen	103	June E. Arbough	Elect. Engineer	84.50	23.80
		101	John G. News	Database Designer	105.00	19.40
		105	Alice K. Johnson *	Database Designer	105.00	35.70
		106	William Smithfield	Programmer	35.75	12.60
		102	David H. Senior	Systems Analyst	96.75	23.80
18	Amber Wave	114	Annelise Jones	Applications Designer	48.10	24.60
		118	James J. Frommer	General Support	18.36	45.30
		104	Anne K. Ramoras *	Systems Analyst	96.75	32.40
		112	Darlene M. Smithson	DSS Analyst	45.95	44.00
22	Rolling Tide	105	Alice K. Johnson	Database Designer	105.00	64.70
		104	Anne K. Ramoras	Systems Analyst	96.75	48.40
		113	Delbert K. Joenbrood *	Applications Designer	48.10	23.60
		111	Geoff B. Wabash	Clerical Support	26.87	22.00
		106	William Smithfield	Programmer	35.75	12.80
25	Starflight	107	Maria D. Alonzo	Programmer	35.75	24.60
		115	Travis B. Bawangi	Systems Analyst	96.75	45.80
		101	John G. News *	Database Designer	105.00	56.30
		114	Annelise Jones	Applications Designer	48.10	33.10
		108	Ralph B. Washington	Systems Analyst	96.75	23.60
		118	James J. Frommer	General Support	18.36	30.50
		112	Darlene M. Smithson	DSS Analyst	45.95	41.40

Grupos repetidores: se debe iniciar por representar los datos en formato tabular, donde cada celda tenga un solo valor y donde no hay grupos repetidores. Para eliminar éstos, se deben eliminar los valores nulos asegurando que el atributo de cada uno de los grupos repetidores contiene un valor apropiado de datos



Paso 1: eliminar los grupos repetidores (Solución)

PROJ_NUM	PROJ_NAME	EMP_NUM	EMP_NAME	JOB_CLASS	CHG_HOUR	HOURS
15	Evergreen	103	June E. Arbough	Elect. Engineer	84.50	23.80
15	Evergreen	101	John G. News	Database Designer	105.00	19.40
15	Evergren	105	Alice K. Johnson *	Database Designer	105.00	35.70
15	Evergreen	106	William Smithfield	Programmer	35.75	12.60
15	Evergreen	102	David H. Senior	Systems Analyst	96.75	23.80
18	Amber Wave	114	Annelise Jones	Applications Designer	48.10	24.60
18	Amber Wave	118	James J. Frommer	General Support	18.36	45.30
18	Amber Wave	104	Anne K. Ramoras *	Systems Analyst	96.75	32.40
18	Amber Wave	112	Darlene M. Smithson	DSS Analyst	45.95	44.00
22	Rolling Tide	105	Alice K. Johnson	Database Designer	105.00	64.70
22	Rolling Tide	104	Anne K. Ramoras	Systems Analyst	96.75	48.40
22	Roling Tide	113	Delbert K. Joenbrood *	Applications Designer	48.10	23.60
22	Rolling Tide	111	Geoff B. Wabash	Clerical Support	26.87	22.00
22	Rolling Tide	106	William Smithfield	Programmer	45.75	12.80
25	Starflight	107	Maria D. Alonzo	Programmer	35.75	24.60
25	Starfligt	115	Travis B. Bawangi	Systems Analyst	96.75	45.80
25	Starflight	101	John G. News *	Database Designer	105.00	56.30
25	Starflight	114	Annelise Jones	Applications Designer	48.10	33.10
25	Starflight	108	Ralph B. Washington	Systems Analyst	96.75	23.60
25	Starflight	118	James J. Frommer	General Support	18.36	30.50
25	Starflight	112	Darlene M. Smithson	DSS Analyst	45.95	41.40



¿Cuál es la diferencia entre datos redundantes y grupos repetidores?

Otro ejemplo: grupos repetidores en filas y columnas



Cliente

ID Cliente	Nombre	Apellido	Teléfono
123	Rachel	Ingram	555-861-2025
456	James	Wright	555-403-1659, 555-776-4100
789	Cesar	Dure	555-808-9633

Cliente

ID Cliente	Nombre	Apellido	Teléfono 1	Teléfono 2	Teléfono 3
123	Rachel	Ingram	555-861-2025		
456	James	Wright	555-403-1659	555-776-4100	
789	Cesar	Dure	555-808-9633		



¿Qué problemas causan los grupos repetidores en filas y/o columnas?

Solución



Cliente

ID Cliente	Nombre	Apellido
123	Rachel	Ingram
456	James	Wright
789	Cesar	Dure

Teléfono del cliente

ID Cliente	Teléfono
123	555-861-2025
456	555-403-1659
456	555-776-4100
789	555-808-9633



Paso 2: identificar la llave primaria

PROJ_NUM	PROJ_NAME	EMP_NUM	EMP_NAME	JOB_CLASS	CHG_HOUR	HOURS
15	Evergreen	103	June E. Arbough	Elect. Engineer	84.50	23.80
15	Evergreen	101	John G. News	Database Designer	105.00	19.40
15	Evergren	105	Alice K. Johnson *	Database Designer	105.00	35.70
15	Evergreen	106	William Smithfield	Programmer	35.75	12.60
15	Evergreen	102	David H. Senior	Systems Analyst	96.75	23.80
18	Amber Wave	114	Annelise Jones	Applications Designer	48.10	24.60
18	Amber Wave	118	James J. Frommer	General Support	18.36	45.30
18	Amber Wave	104	Anne K. Ramoras *	Systems Analyst	96.75	32.40
18	Amber Wave	112	Darlene M. Smithson	DSS Analyst	45.95	44.00
22	Rolling Tide	105	Alice K. Johnson	Database Designer	105.00	64.70
22	Rolling Tide	104	Anne K. Ramoras	Systems Analyst	96.75	48.40
22	Roling Tide	113	Delbert K. Joenbrood *	Applications Designer	48.10	23.60
22	Rolling Tide	111	Geoff B. Wabash	Clerical Support	26.87	22.00
22	Rolling Tide	106	William Smithfield	Programmer	45.75	12.80
25	Starflight	107	Maria D. Alonzo	Programmer	35.75	24.60
25	Starfligt	115	Travis B. Bawangi	Systems Analyst	96.75	45.80
25	Starflight	101	John G. News *	Database Designer	105.00	56.30
25	Starflight	114	Annelise Jones	Applications Designer	48.10	33.10
25	Starflight	108	Ralph B. Washington	Systems Analyst	96.75	23.60
25	Starflight	118	James J. Frommer	General Support	18.36	30.50
25	Starflight	112	Darlene M. Smithson	DSS Analyst	45.95	41.40



¿Qué atributo o atributos podrían componer la llave primaria (PK)?



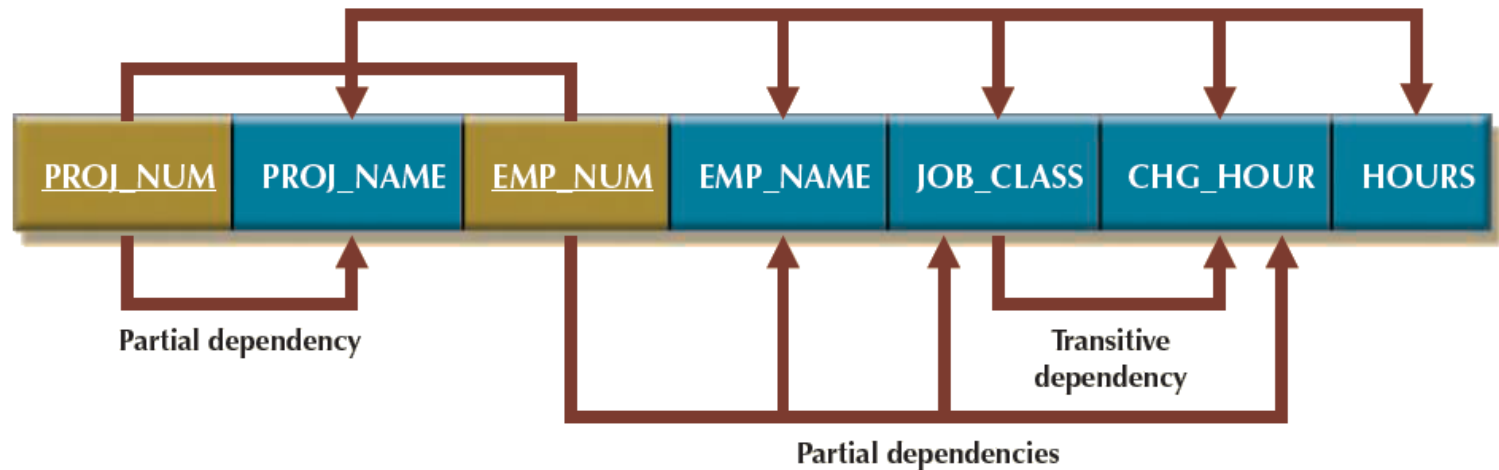
Paso 3: identificar todas las dependencias

PROJ_NUM	PROJ_NAME	EMP_NUM	EMP_NAME	JOB_CLASS	CHG_HOUR	HOURS
15	Evergreen	103	June E. Arbough	Elect. Engineer	84.50	23.80
15	Evergreen	101	John G. News	Database Designer	105.00	19.40
15	Evergren	105	Alice K. Johnson *	Database Designer	105.00	35.70
15	Evergreen	106	William Smithfield	Programmer	35.75	12.60
15	Evergreen	102	David H. Senior	Systems Analyst	96.75	23.80
18	Amber Wave	114	Annelise Jones	Applications Designer	48.10	24.60
18	Amber Wave	118	James J. Frommer	General Support	18.36	45.30
18	Amber Wave	104	Anne K. Ramoras *	Systems Analyst	96.75	32.40
18	Amber Wave	112	Darlene M. Smithson	DSS Analyst	45.95	44.00
22	Rolling Tide	105	Alice K. Johnson	Database Designer	105.00	64.70
22	Rolling Tide	104	Anne K. Ramoras	Systems Analyst	96.75	48.40
22	Roling Tide	113	Delbert K. Joenbrood *	Applications Designer	48.10	23.60
22	Rolling Tide	111	Geoff B. Wabash	Clerical Support	26.87	22.00
22	Rolling Tide	106	William Smithfield	Programmer	45.75	12.80
25	Starflight	107	Maria D. Alonzo	Programmer	35.75	24.60
25	Starfligt	115	Travis B. Bawangi	Systems Analyst	96.75	45.80
25	Starflight	101	John G. News *	Database Designer	105.00	56.30
25	Starflight	114	Annelise Jones	Applications Designer	48.10	33.10
25	Starflight	108	Ralph B. Washington	Systems Analyst	96.75	23.60
25	Starflight	118	James J. Frommer	General Support	18.36	30.50
25	Starflight	112	Darlene M. Smithson	DSS Analyst	45.95	41.40



¿Qué dependencias parciales y transitivas posee la tabla?

Paso 3: elaborar el diagrama de dependencias (Solución)



1NF (PROJ_NUM, EMP_NUM, PROJ_NAME, EMP_NAME, JOB_CLASS, CHG_HOURS, HOURS)

PARTIAL DEPENDENCIES:

(PROJ_NUM \Rightarrow PROJ_NAME)

(EMP_NUM \Rightarrow EMP_NAME, JOB_CLASS, CHG_HOUR)

TRANSITIVE DEPENDENCY:

(JOB_CLASS \Rightarrow CHG_HOUR)

Dependencia parcial: condición en la que un atributo depende sólo de una parte (subconjunto) de la clave primaria (compuesta)

Dependencia transitiva: condición en la que un atributo depende de otro atributo que no forma parte de la clave primaria.

Conversión a segunda forma normal

Una tabla está en segunda forma normalizada (**2NF**) si:

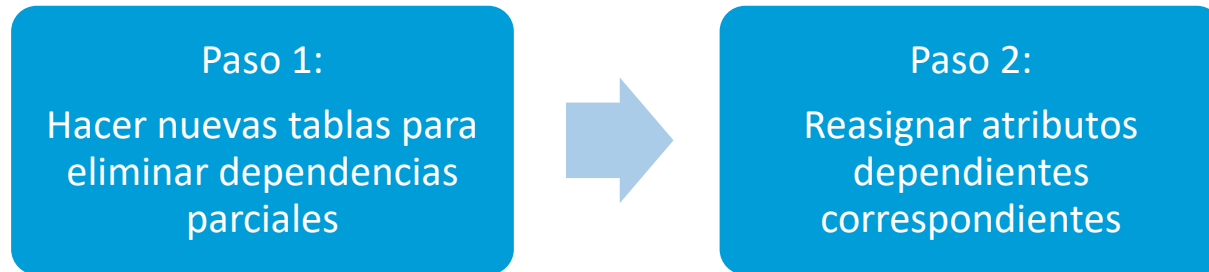
1

- Está en 1NF

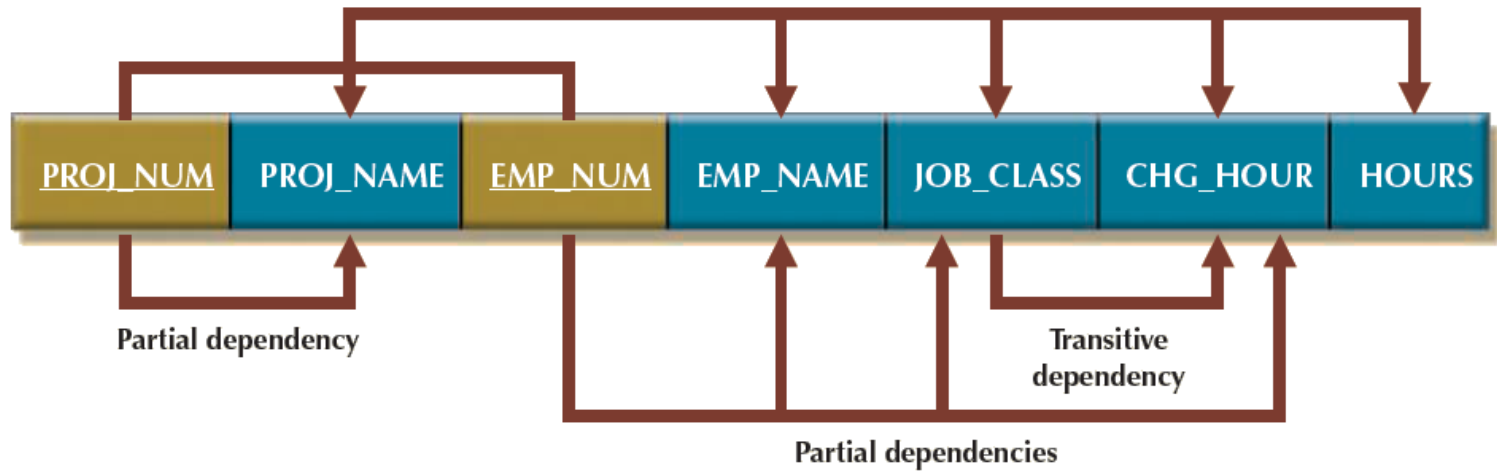
2

- No incluye **dependencias parciales**; esto es, ningún atributo es dependiente de sólo una parte de la llave primaria

Solución 2NF



Paso 1: hacer nuevas tablas para eliminar dependencias parciales



1NF (PROJ_NUM, EMP_NUM, PROJ_NAME, EMP_NAME, JOB_CLASS, CHG_HOURS, HOURS)

PARTIAL DEPENDENCIES:

(PROJ_NUM \Rightarrow PROJ_NAME)

(EMP_NUM \Rightarrow EMP_NAME, JOB_CLASS, CHG_HOUR)

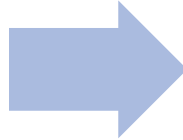
TRANSITIVE DEPENDENCY:

(JOB_CLASS \Rightarrow CHG_HOUR)

Paso 1: hacer nuevas tablas para eliminar dependencias parciales (Solución)

Generar nuevas tablas por cada componente de la llave primaria que actúe como determinante

- Llave Primaria (PK): **PROJ_NUM+EMP_NUM**
- Nuevas tablas: **PROJECT, EMPLOYEE**



Utilizar dichos componentes como llave primaria en las nuevas tablas y mantenerlos en la original

- Tabla: **PROJECT**
- Llave primaria: **PROJ_NUM**
- Tabla: **EMPLOYEE**
- Llave primaria: **EMP_NUM**
- Tabla: **ASSIGNMENT**
- Llave primaria: **PROJ_NUM+EMP_NUM**

Paso 2: reasignar atributos dependientes correspondientes (Solución)

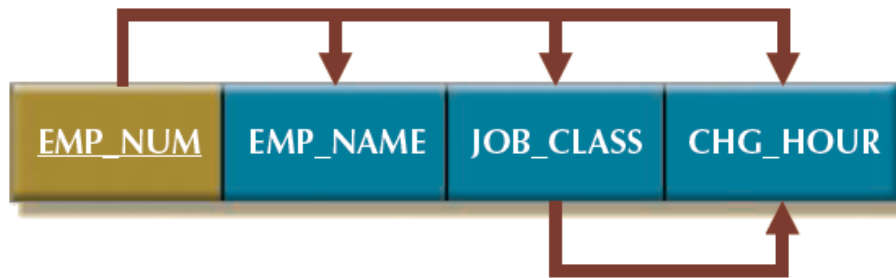
Table name: PROJECT

PROJECT (PROJ_NUM, PROJ_NAME)



Table name: EMPLOYEE

EMPLOYEE (EMP_NUM, EMP_NAME, JOB_CLASS, CHG_HOUR)

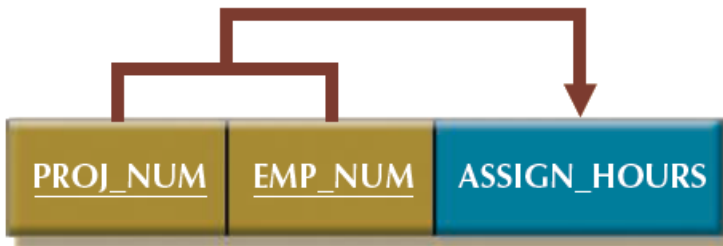


TRANSITIVE DEPENDENCY
(JOB_CLASS → CHG_HOUR)

Transitive
dependency

Table name: ASSIGNMENT

ASSIGNMENT (PROJ_NUM, EMP_NUM, ASSIGN_HOURS)



Conversión a tercera forma normal

Una tabla está en tercera forma normalizada (**3NF**) si:

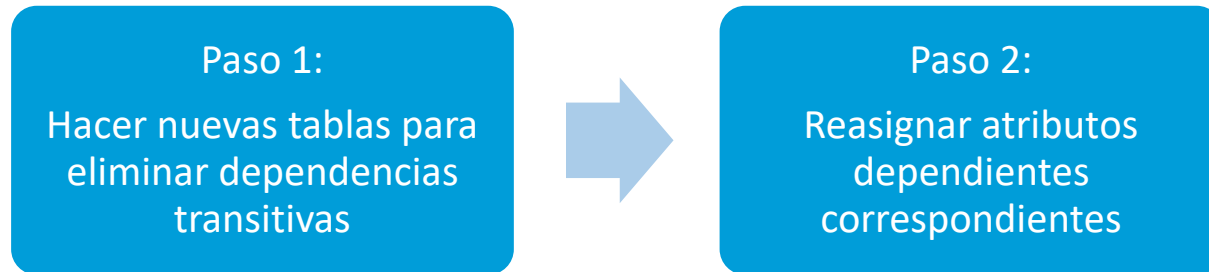
1

- Está en 2NF

2

- Y no contiene **dependencias transitivas**

Solución 3NF



Paso 1: hacer nuevas tablas para eliminar dependencias transitivas

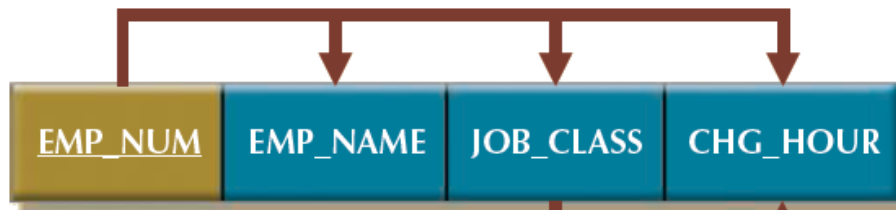
Table name: PROJECT

PROJECT (PROJ_NUM, PROJ_NAME)



Table name: EMPLOYEE

EMPLOYEE (EMP_NUM, EMP_NAME, JOB_CLASS, CHG_HOUR)

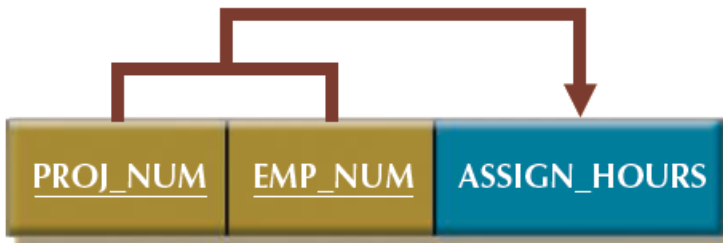


TRANSITIVE DEPENDENCY
(JOB_CLASS → CHG_HOUR)

Transitive
dependency

Table name: ASSIGNMENT

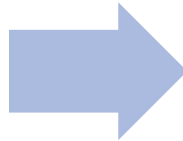
ASSIGNMENT (PROJ_NUM, EMP_NUM, ASSIGN_HOURS)



Paso 1: hacer nuevas tablas para eliminar dependencias transitivas (Solución)

Por cada dependencia transitiva, escriba una copia de su determinante como llave primaria para una nueva tabla

- Llave Primaria (PK):
JOB_CLASS



Utilizar dicho componente como llave primaria en la nueva tabla y mantenerlo en la original

- Tabla: JOB
- Llave primaria: **JOB_CLASS**

Un **determinante**: es cualquier atributo cuyo valor determina otros valores dentro de un renglón (fila o registro)

Paso 2: reasignar atributos dependientes correspondientes (Solución)



Table name: PROJECT

PROJECT (PROJ_NUM, PROJ_NAME)



Table name: JOB

JOB (JOB_CLASS, CHG_HOUR)

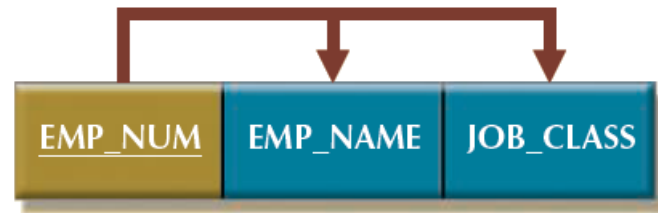


Table name: EMPLOYEE

EMPLOYEE (EMP_NUM, EMP_NAME, JOB_CLASS)

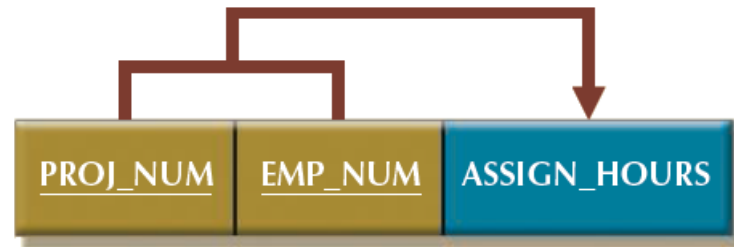


Table name: ASSIGNMENT

ASSIGNMENT (PROJ_NUM, EMP_NUM, ASSIGN_HOURS)

Mejoramiento del diseño

Evaluar asignaciones de llaves primarias

- Utilizar identificadores únicos que reduzcan la cantidad de errores al ser introducidos (Llaves primarias sustitutas), Ej. Es mejor usar **JOB_CODE** que **JOB_CLASS**

Evaluar convenciones para dar nombre

- Es muy recomendable usar convenciones para dar nombre a los atributos, facilita el trabajo a diseñadores y programadores, **TABLA_ATRIBUTO**

Refinar atomicidad de atributo

- Un atributo atómico es el que no puede subdividirse en más; se dice que tal atributo muestra atomicidad. Es preferible utilizar atributos sencillos, de un solo valor. Ej. **EMP_NAME** (dividirlo)

Identificar nuevos atributos

- Es posible agregar atributos para mejorar la capacidad de la base de datos para dar información y mejorar sus características operacionales, Ej. Dar información histórica

Mejoramiento del diseño (cont..)

Identificar nuevas relaciones

- Para dar cumplimiento a las reglas de negocio es posible incorporar **nuevas relaciones** entre las tablas, pero bajo los principios de normalización

Refinar llaves primarias

- Se deben refinar las llaves primarias según sea necesario para granularidad de datos. La **granularidad** se refiere al nivel de detalle representado por los valores guardados en cada fila de la tabla

Mantener precisión histórica

- Es apropiado incorporar nuevos atributos para mantener precisión histórica de los datos en la tabla. Ej. Historial de sueldos

Evaluar el uso de atributos derivados

- Guardar **atributos derivados** en la tabla facilita escribir el software de aplicación para producir los resultados deseados. También, si se requiere de muchas transacciones complejas ahorrará tiempo de informe

Base de datos mejorada



Table name: PROJECT



Table name: PROJECT

PROJ_NUM	PROJ_NAME	EMP_NUM
15	Evergreen	105
18	Amber Wave	104
22	Rolling Tide	113
25	Starflight	101

Table name: JOB



Table name: JOB

JOB_CODE	JOB_DESCRIPTION	JOB_CHG_HOUR
500	Programmer	35.75
501	Systems Analyst	96.75
502	Database Designer	105.00
503	Electrical Engineer	84.50
504	Mechanical Engineer	67.90
505	Civil Engineer	55.78
506	Clerical Support	26.87
507	DSS Analyst	45.95
508	Applications Designer	48.10
509	Bio Technician	34.55
510	General Support	18.36



Base de datos mejorada (cont..)

Table name: ASSIGNMENT

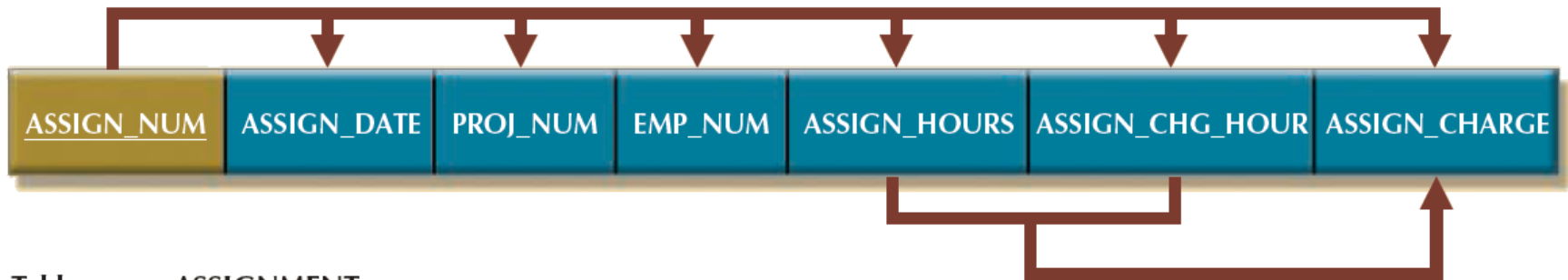


Table name: ASSIGNMENT

ASSIGN_NUM	ASSIGN_DATE	PROJ_NUM	EMP_NUM	ASSIGN_HOURS	ASSIGN_CHG_HOUR	ASSIGN_CHARGE
1001	04-Mar-16	15	103	2.6	84.50	219.70
1002	04-Mar-16	18	118	1.4	18.36	25.70
1003	05-Mar-16	15	101	3.6	105.00	378.00
1004	05-Mar-16	22	113	2.5	48.10	120.25
1005	05-Mar-16	15	103	1.9	84.50	160.55
1006	05-Mar-16	25	115	4.2	96.75	406.35
1007	05-Mar-16	22	105	5.2	105.00	546.00
1008	05-Mar-16	25	101	1.7	105.00	178.50
1009	05-Mar-16	15	105	2.0	105.00	210.00
1010	06-Mar-16	15	102	3.8	96.75	367.65
1011	06-Mar-16	22	104	2.6	96.75	251.55
1012	06-Mar-16	15	101	2.3	105.00	241.50
1013	06-Mar-16	25	114	1.8	48.10	86.58
1014	06-Mar-16	22	111	4.0	26.87	107.48



Base de datos mejorada (cont..)

Table name: EMPLOYEE

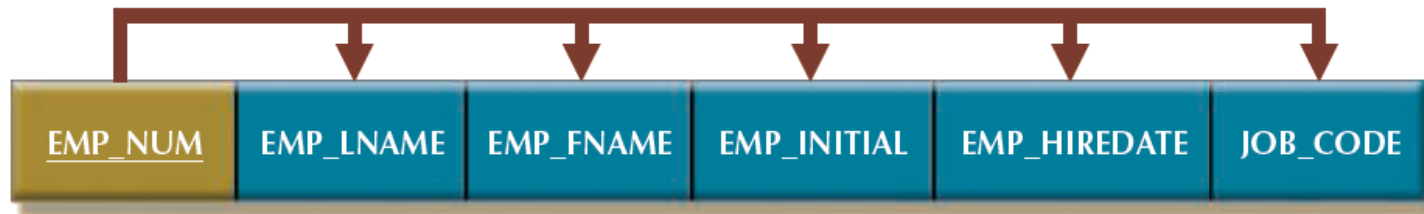
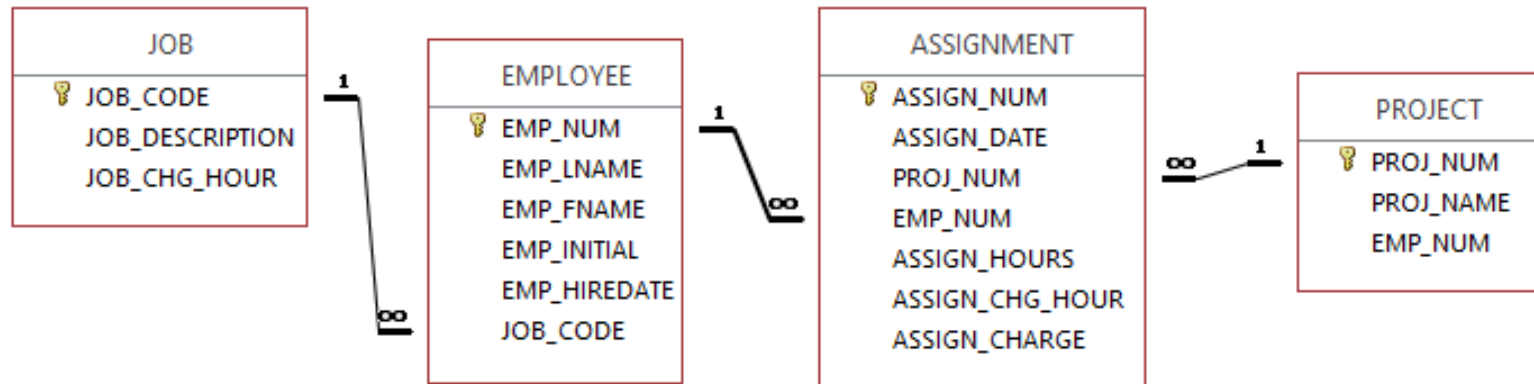


Table name: EMPLOYEE

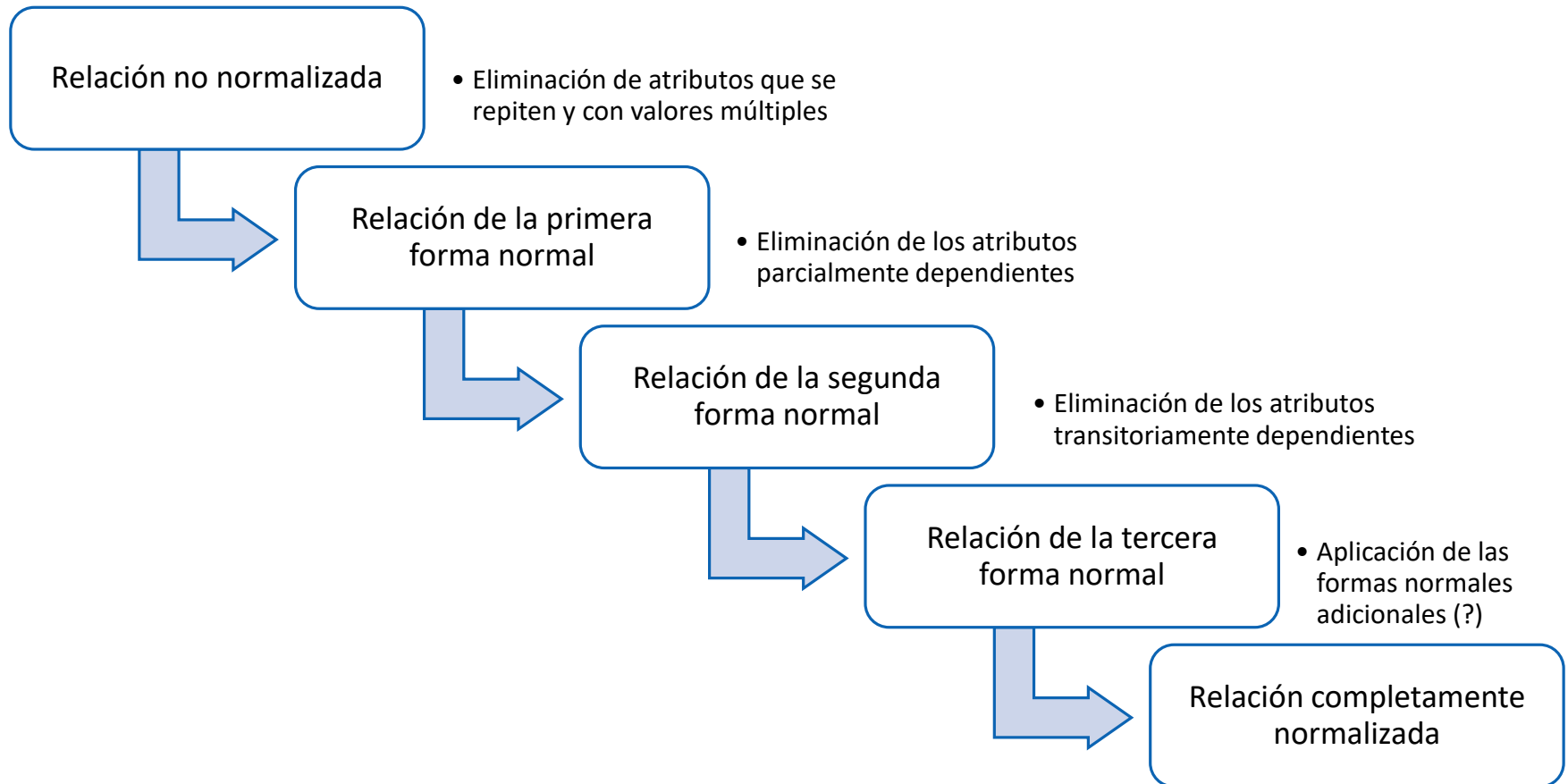
EMP_NUM	EMP_LNAME	EMP_FNAME	EMP_INITIAL	EMP_HIREDATE	JOB_CODE
101	News	John	G	08-Nov-00	502
102	Senior	David	H	12-Jul-89	501
103	Arbough	June	E	01-Dec-97	503
104	Ramoras	Anne	K	15-Nov-88	501
105	Johnson	Alice	K	01-Feb-94	502
106	Smithfield	William		22-Jun-05	500
107	Alonzo	Maria	D	10-Oct-94	500
108	Washington	Ralph	B	22-Aug-89	501
109	Smith	Larry	W	18-Jul-99	501
110	Olenko	Gerald	A	11-Dec-96	505
111	Wabash	Geoff	B	04-Apr-89	506



Diagrama de la base de datos en 3NF



Resumen del proceso de normalización



Conversión a forma normal de Boyce-Codd

Una tabla está en forma normalizada de Boyce-Codd (**BCNF**) si:

1

- Está en 3NF

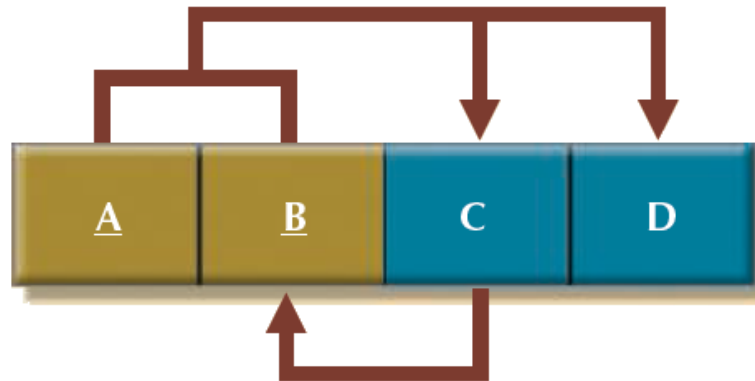
2

- Y no posee atributos que forman parte de una clave primaria compuesta **dependientes funcionalmente** de un atributo que no es una clave

Dependencia funcional: Una dependencia funcional es una conexión entre uno o más atributos. Por ejemplo si se conoce el valor de *FechaDeNacimiento* podemos conocer el valor de *Edad*.

Aquí a *FechaDeNacimiento* se le conoce como un determinante. Se puede leer de dos formas *FechaDeNacimiento* determina a *Edad* o *Edad* es funcionalmente dependiente de *FechaDeNacimiento*.

Tabla que está en 3NF, pero no en BCNF



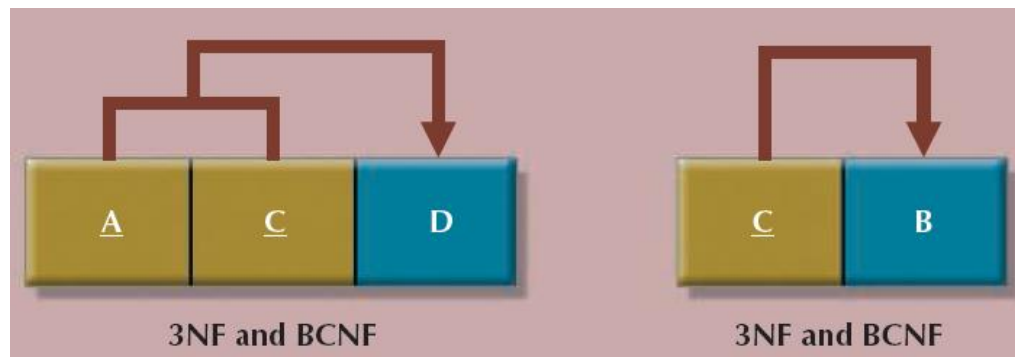
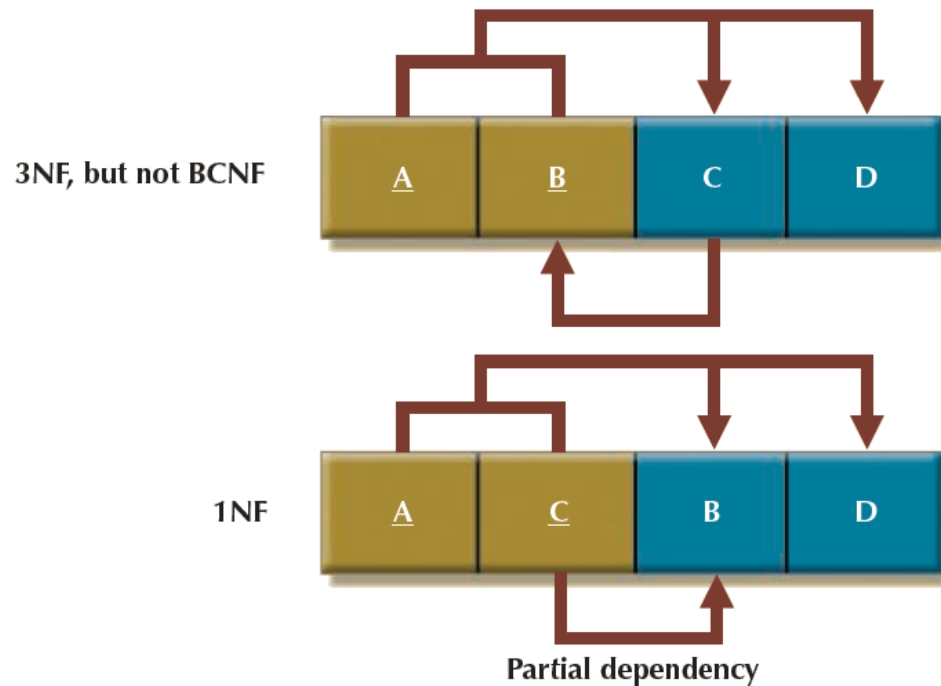
Dependencias:

$A + B \rightarrow C, D$

$C \rightarrow B$

La condición $C \rightarrow B$ indica que un atributo no de clave determina una parte de la clave primaria (y esta dependencia no es transitiva)

Descomposición de una estructura de tabla para satisfacer los requerimientos de BCNF





Otro ejemplo: BCNF

SAMPLE DATA FOR A BCNF CONVERSION			
STU_ID	STAFF_ID	CLASS_CODE	ENROLL_GRADE
125	25	21334	A
125	20	32456	C
135	20	28458	B
144	25	27563	C
144	20	32456	B

TABLA DE INSCRIPCIÓN DE CAPACITACIONES:

STU_ID: código de identificación del estudiante

STAFF_ID: código de identificación del capacitador

Una tabla está en forma normalizada de Boyce-Codd (**BCNF**) cuando todo determinante de la tabla es una llave candidata.

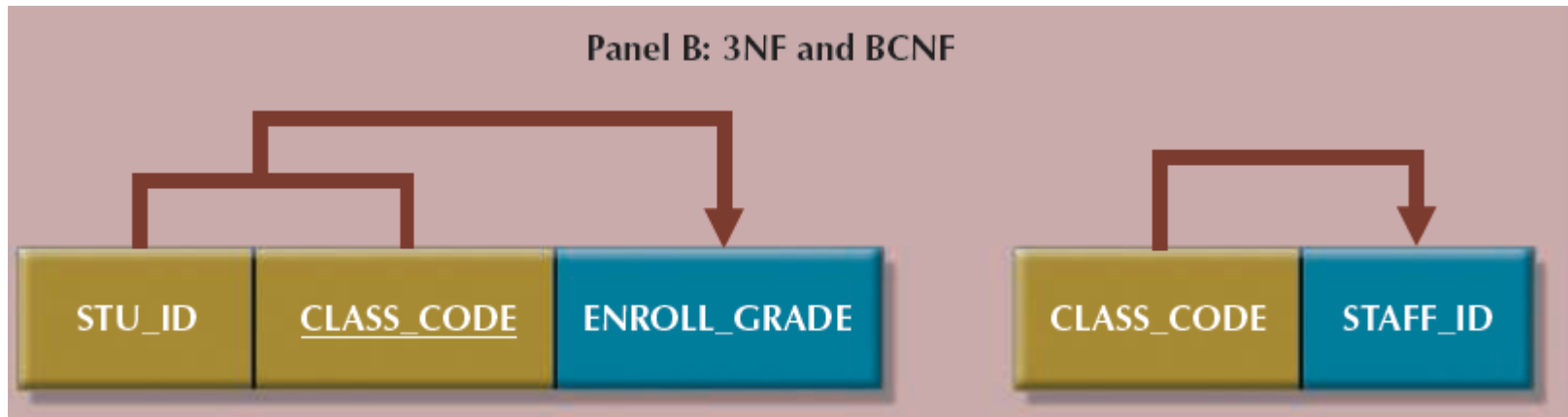
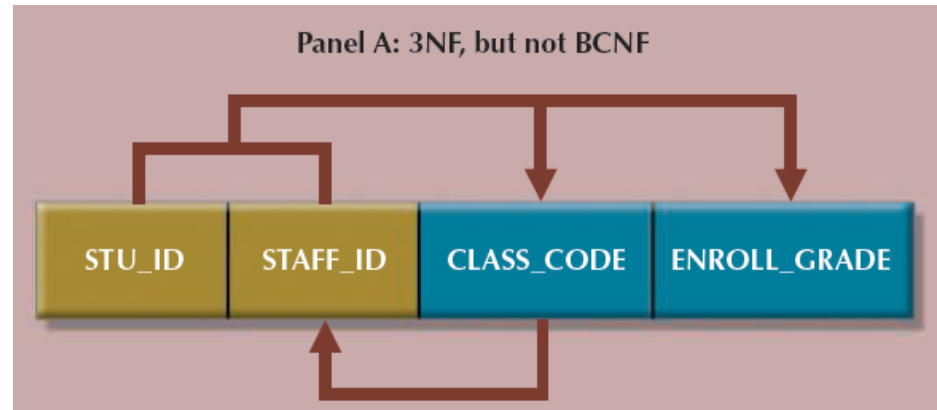
Llave candidata 1: STU_ID + STAFF_ID

Llave candidata 2: STU_ID + CLASS_CODE



Otro ejemplo: BCNF

La condición $CLASS_CODE \rightarrow STAFF_ID$ indica que un atributo no de clave determina una parte de la clave primaria (**dependencia funcional**)



Conversión a cuarta forma normal

Una tabla está en cuarta forma normalizada (**4NF**) si:

1

- Está en 3NF

2

- No tiene **dependencias de valores múltiples**

Dependencia multivaluada: es donde la existencia de dos o más relaciones independientes muchos a muchos (*entre los mismos atributos de la tabla*) causa redundancia; y es esta redundancia la que es suprimida por la cuarta forma normal.

Ejemplo con dependencias de valores múltiples

- Este ejemplo ilustra **una base de datos mal diseñada**. Por ejemplo, considere la posibilidad de que un empleado pueda tener múltiples asignaciones y también puede estar involucrado en múltiples organizaciones de servicio. Suponga que el empleado **10123** es voluntario de la Cruz Roja y United Way. Además, el mismo empleado podría ser asignado para trabajar en tres proyectos: 1, 3 y 4. La siguiente figura ilustra cómo se puede registrar ese conjunto de hechos de maneras muy diferentes.

Table name: VOLUNTEER_V1

EMP_NUM	ORG_CODE	ASSIGN_NUM
10123	RC	1
10123	UW	3
10123		4

Table name: VOLUNTEER_V3

EMP_NUM	ORG_CODE	ASSIGN_NUM
10123	RC	1
10123	RC	3
10123	UW	4

Table name: VOLUNTEER_V2

EMP_NUM	ORG_CODE	ASSIGN_NUM
10123	RC	
10123	UW	
10123		1
10123		3
10123		4

VOLUNTEER: posee dos relaciones muchos a mucho independientes entre sí.

1. Un **empleado** puede pertenecer a muchas **organizaciones** y una organización puede tener muchos empleados trabajando
2. Un **empleado** puede ser asignado a muchos **proyectos** y un proyecto puede tener muchos empleados asignados

Solución 4NF

Table name: PROJECT

PROJ_CODE	PROJ_NAME	PROJ_BUDGET
1	BeThere	1023245.00
2	BlueMoon	20198608.00
3	GreenThumb	3234456.00
4	GoFast	5674000.00
5	GoSlow	1002500.00

Table name: EMPLOYEE

EMP_NUM	EMP_LNAME
10121	Rogers
10122	O'Leery
10123	Panera
10124	Johnson

Table name: SERVICE_V1

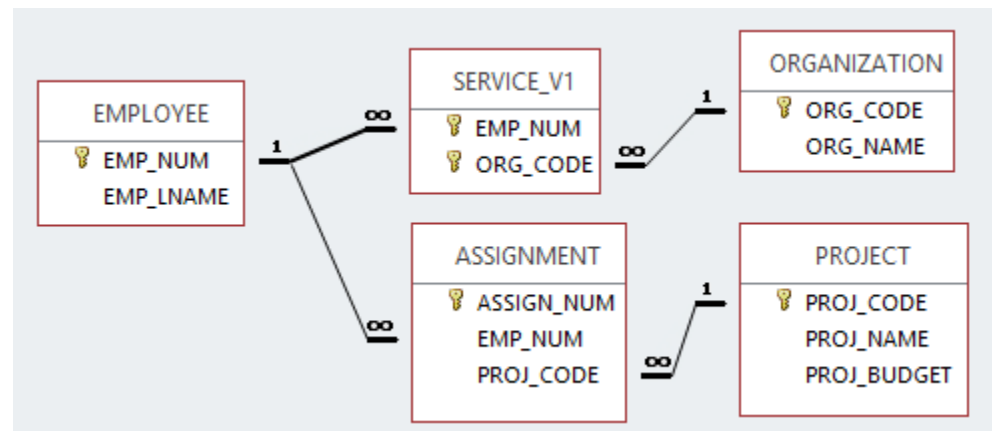
EMP_NUM	ORG_CODE
10123	RC
10123	UW
10123	WF

Table name: ASSIGNMENT

ASSIGN_NUM	EMP_NUM	PROJ_CODE
1	10123	1
2	10121	2
3	10123	3
4	10123	4
5	10121	1
6	10124	2
7	10124	3
8	10124	5

Table name: ORGANIZATION

ORG_CODE	ORG_NAME
RC	Red Cross
UW	United Way
WF	Wildlife Fund



¿Qué es la desnormalización?

- ❑ La normalización conduce a más relaciones, lo que se traduce en más tablas y combinaciones.
- ❑ Cuando los usuarios de una base de datos padecen **problemas de desempeño** que no se pueden resolver por otros medios, como afinar las bases de datos o actualizar el hardware en que funciona el DBMS, es posible que se requiera una desnormalización.
- ❑ La desnormalización debe utilizarse como último recurso.

Tomar en cuenta que la ventaja de una mayor velocidad de procesamiento debe evaluarse cuidadosamente contra la desventaja de la redundancia de datos y datos anómalos.

Desnormalización

Desnormalizar sería el proceso inverso de pasar una relación de mayor nivel a otra de menor nivel.

Tarea de la semana

Leer el capítulo 6 del libro: Bases de Datos Diseño, Implementación y Administración

Realizar la guía práctica 6

Bibliografía

1. CORONEL, MORRIS, ROB, 2017, Bases de Datos Diseño, Implementación y Administración, 12ª Ed., Cengage.
2. REINOSA ENRIQUE, 2012, Bases de datos, Alfaomega.