



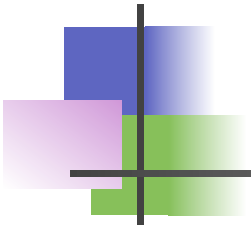
**Departamento de
Ciencias de la Computación y Tecnologías de Información
Universidad del Bío-Bío
Sede Chillán**

Bases de Datos

Normalización de Esquemas

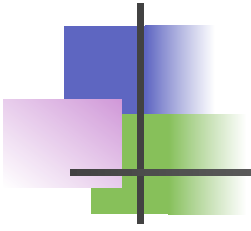
M^a Angélica Caro Gutiérrez

<http://www.face.ubiobio.cl/~mcaro/>
mcaro@ubiobio.cl



Normalización

- ➔ ■ Introducción al refinamiento de esquemas
- Dependencias Funcionales (DFs)
- Razonamiento sobre DFs
- Dependencias Multivaluadas, de Proyección y Unión



Introducción

- **Estructuración de la Información:**
 - Diseño de la BD -> Esquema Relacional
 - El diseño busca establecer la mejor manera de organizar la información.
 - **Aspectos que caracterizan un diseño**
 - Creativo
 - Basado en la experiencia (conocimiento previo)
 - Se aprende haciendo
 - Importante seguir un cierto método de trabajo



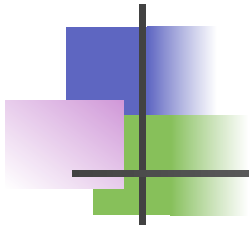
Introducción

- **Estructuración de la Información:**
 - PRINCIPIOS BÁSICOS de cualquier diseño
 - Ley del mínimo esfuerzo
 - Lo bueno, si es breve, dos veces bueno.
 - Entre dos opciones correctas, elegir la más simple.
 - Factor de Escala
 - La dificultad crece de forma exponencial al crecer el tamaño.
 - Metáfora del edificio: construir un cerco, un muro, una casa, un edificio de 3 pisos, una torre.



Introducción

- **Objetivos de la estructuración:**
 - **Evitar redundancias**
 - Duplicidad: repetición de un mismo dato.
 - Redundancia: duplicidad innecesaria.
 - Fuente de inconsistencias al poder tener varios valores diferentes para un mismo dato.
 - Dificultan las consultas y actualizaciones.
 - **Evitar valores desconocidos (nulos).**
 - Difíciles de manejar por las personas.
 - Producen paradojas lógicas en los computadores.
 - **Facilitar el manejo de los datos.**
 - Futuras consultas y modificaciones.



Introducción

- Problemas de un mal diseño, ejemplo:
 - Relación ESCRIBE

NOMBRE	NACIONALIDAD	COD_LIBRO	TITULO	EDITORIAL	AÑO
Date, C.	Norteamericana	989	Database Systems	Addison-W	1990
Date, C.	Norteamericana	989	Database Systems	Addison-W	1986
Date, C.	Norteamericana	989	Database Systems	Addison-W	1988
Codd, E.	Norteamericana	989	Database Systems	Addison-W	1990
Gardarin	Francesa	112	O-O Databases	Eyrolles	1986
Gardarin	Francesa	112	O-O Databases	Eyrolles	1984
Valduriez	Francesa	112	O-O Databases	Eyrolles	1984
Kim, W.	Norteamericana	112	O-O Databases	ACM Press	1989
Lochovski	Canadiense	112	O-O Databases	ACM Press	1989

Redundancia

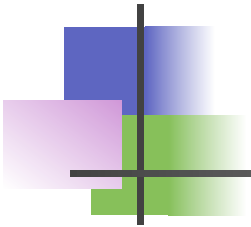
“Hechos distintos se deben almacenar en objetos distintos”

Prob. del borrado

Prob. de inserción

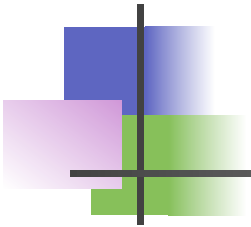
Prob. al modificar

Inserción de un autor del que no se tiene libro
Insertar un libro anónimo



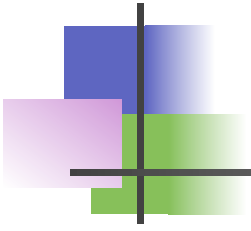
Introducción

- Resumiendo, algunos problemas que se pueden presentar por un mal diseño:
 - Incapacidad de almacenar ciertos hechos
 - Redundancias, i.e., posibilidad de incoherencias
 - Ambigüedades
 - Pérdida de información
 - Pérdida de dependencias funcionales, i.e., de ciertas restricciones de integridad (interdependencias entre los datos).
 - Aparición en la BD de estados no válidos en el mundo real.



Introducción

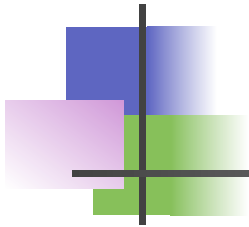
- La teoría de la normalización
 - Consiste en obtener esquemas relacionales que cumplan unas determinadas condiciones, y se centra en lo que se conoce como formas normales.
 - Existen 5 formas normales: 1FN, 2FN, 3FN (Codd) y FNBC, 4FN y 5FN (Fagin).
 - Se basa en ciertas restricciones definidas sobre los atributos de una relación, las cuales se conocen como dependencias:
 - 1FN, 2FN, 3FN y FNBC -> dependencias *funcionales*
 - 4FN -> dependencias *multivaluadas*
 - 5FN -> dependencias de *proyección-combinación*



Introducción

■ Primera forma normal (1FN)

- Es una restricción inherente al modelo relacional
- Prohíbe los atributos multivaluados, compuestos o combinaciones
- El *Dominio* de un atributo debe incluir sólo valores atómicos



Introducción

■ Primera forma normal (1FN): Ejemplo

NOMBRED	NUMEROD	NSS_JEFED	LOCALIZACIONES
Investigación	5	333445555	{ Bellaire, Sugarland, Houston}
Administración	4	987654321	{ Stafford}
Dirección	1	88665555	{ Houston}



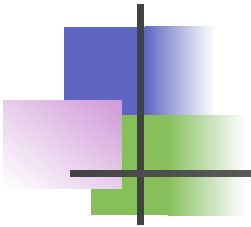
**No
está
en
1FN**

NOMBRED	NUMEROD	NSS_JEFED	LOCALIZACIONES
Investigación	5	333445555	Bellaire
Investigación	5	333445555	Sugarland
Investigación	5	333445555	Houston
Administración	4	987654321	Stafford
Dirección	1	88665555	Houston



**Si
está
en
1FN**





Introducción

■ Segunda forma normal (2FN)

- Está en 1FN
- Todos los atributos que no forman parte de alguna clave candidata dependen completamente de la clave (y no de parte de la clave)



Introducción

■ Segunda forma normal (2FN): Ejemplo

■ Sea el esquema:

- PRESTAMOS(numSocio, nombreSocio, codLibro, fechaPrestamo, editorial, país)

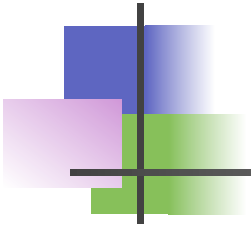


■ **claves candidatas:**

- (numSocio, codLibro)
- (nombreSocio, codLibro)

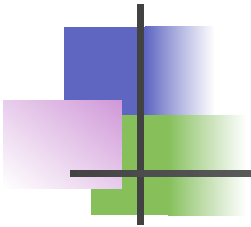
■ **Solución (2FN)**

- PRESTAMO1(numSocio, nombreSocio, codLibro, fechaPrestamo)
- LIBROS(codLibro, editorial, país)



■ Tercera forma normal (3FN)

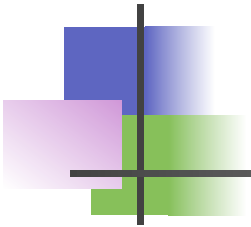
- Está en 2FN
- Todo atributo que no pertenece a alguna clave candidata depende de la clave y sólo de la clave candidata y no de otros atributos.



Introducción

■ Tercera forma normal (3FN): Ejemplo

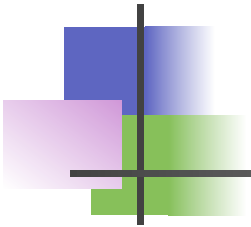
- PRESTAMO1(numSocio, nombreSocio, codLibro, fechaPrestamo) 😊
- LIBROS(codLibro, editorial, país) 😞
- No está en 3FN, ya que *país* depende de *editorial*
- **Solución en 3FN:**
 - LIBROS1(codLibro, editorial)
 - EDITORIALES(editorial, país)



Introducción

■ Forma normal de Boyce y Codd (FNBC)

- Una relación está en FNBC si y sólo si cada atributo depende de una clave candidata.
- La FNBC es una redefinición más estricta de la 3FN.
- Sólo pueden existir dependencias entre cada atributo no clave y una clave candidata.



Introducción

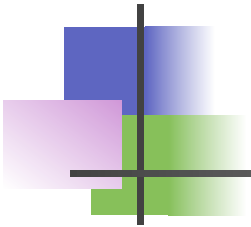
■ Forma normal de Boyce y Codd (FNBC)

■ Ejemplo:

- PRESTAMO1(numSocio, nombreSocio, codLibro, fechaPrestamo)
- No está en FNBC ya que se presenta la anomalía que numSocio y nombreSocio se repiten innecesariamente por cada libro, y además existe dependencia entre ellas.

■ **Solución en 3FN:**

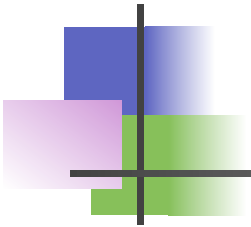
- SOCIOS(numSocio, nombreSocio)
- PRESTAMO2(numSocio, codLibro, fechaPrestamo)



Introducción

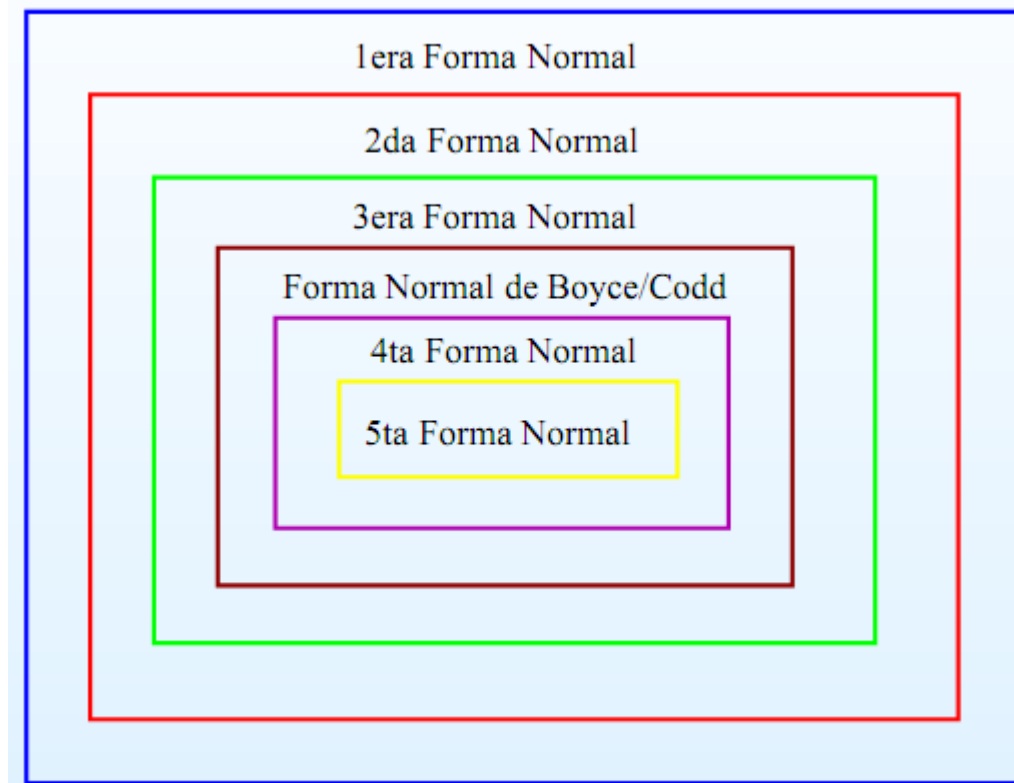
■ Forma normal de Boyce y Codd (FNBC)

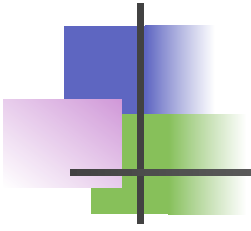
- De esta forma la relación:
 - PRESTAMOS(numSocio, nombreSocio, codLibro, fechaPrestamo, editorial, país)
- se descompone en las relaciones en FNBC:
 - LIBROS1(codLibro, editorial)
 - EDITORIALES(editorial, país)
 - SOCIOS(numSocio, nombreSocio)
 - PRESTAMOS2(numSocio, codLibro, fechaPrestamo)



Introducción

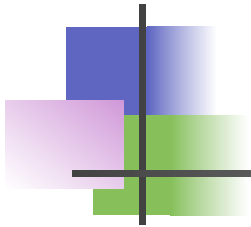
■ Noción intuitiva de las Formas Normales





Unidad 6: Normalización

- Introducción al refinamiento de esquemas
- ➡ ■ Dependencias Funcionales (DFs)
- Razonamiento sobre DFs
- Dependencias Multivaluadas, de Proyección y Unión



Dependencias Funcionales (DFs)

- Las dependencias constituyen una parte importante de la semántica del “negocio” y deben ser incluidas en los esquemas de relación.
- Así, un esquema de una relación será:
 - $R(A, DEP)$
 - donde A es el conjunto de atributos y DEP es el conjunto de todo tipo de dependencias que existen entre los atributos.
- Primero estudiaremos las dependencias funcionales (dfs), es decir: **$R(A, DF)$** .



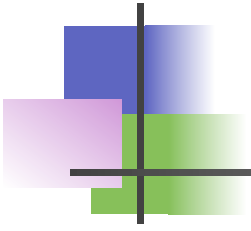
Dependencias Funcionales

■ Definiendo dependencia funcional

- Las dfs son un tipo especial de dependencias, el más extendido, y está relacionado con las tres primeras formas normales.
- Sea el esquema R definido sobre el conjunto de atributos A y sean X e Y subconjuntos de A llamados descriptores. Se dice que Y *depende funcionalmente de* X o que X *determina* o *implica* a Y lo que se denota por:

$$\mathbf{X \rightarrow Y}$$

- si y, sólo si, cada valor de X tiene asociado en todo momento un único valor de Y.

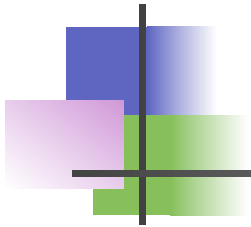


Dependencias Funcionales

■ Definiendo dependencia funcional

■ Ejemplo:

- $\text{codLibro} \rightarrow \text{titulo}$
- $\text{numSocio} \rightarrow \text{nombre, domicilio, telefono}$
- $(\text{codLibro}, \text{numSocio}) \rightarrow \text{fechaPrestamo, fechaDev}$
- ¿Otros ejemplos?



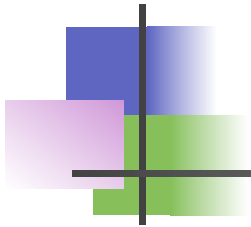
Dependencias Funcionales

■ Definiendo dependencia funcional

- Una dependencia funcional (DF) es una restricción de integridad que generaliza el concepto de clave.
- Sea R el esquema de una relación y sean X, Y conjuntos no vacíos de atributos de R . Se dice que una instancia r de R satisface la DF $X \rightarrow Y$ si para cada par de tuplas t_1 y t_2 en r se cumple:

Si $t_1.X = t_2.X$, entonces $t_1.Y = t_2.Y$

donde $t_1.X$ es la proyección de los atributos X en la tupla t_1



Dependencias Funcionales

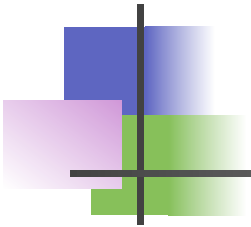
■ Definiendo dependencia funcional

- En otras palabras, $X \rightarrow Y$ indica básicamente que si dos tuplas de r coinciden en el valor de los atributos de X , también deben coincidir en el valor de los atributos de Y .

- Ejemplo: La siguiente relación satisface la DF $AB \rightarrow C$:

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_1	d_2
a_1	b_2	c_2	d_1
a_2	b_1	c_3	d_1

- Si se inserta la tupla $\langle a_1, b_1, c_2, d_1 \rangle$ a r , entonces r no satisface la DF.
- Las dos primeras tuplas muestran que las DF no son lo mismo que las restricciones de claves: **es evidente que AB no es clave de la relación**



Dependencias Funcionales

■ Definiendo dependencia funcional

- Ejercicio: Dada la relación que satisface la DF $AB \rightarrow C$:

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_1	d_2
a_1	b_2	c_2	d_1
a_2	b_1	c_3	d_1

- Dar un ejemplo de una tupla que satisfaga $AB \rightarrow C$
- ¿Se cumple la DF $A \rightarrow B$? ¿Porqué?
- De un ejemplo de otra DF que se cumpla en r.