

Bases de Datos Unidad 2

Instructor: M.C. Luis Basto Díaz
Email: luis.basto@uady.mx

Modelo Relacional

- Definiciones
 - Relación
 - Tabla
 - Dominio
 - Tupla
 - Esquemas
- Álgebra Relacional
- Lenguaje SQL
- Integridad Referencial
- Normalización
 - Primera, segunda y tercera Forma Normal
 - Forma Normal de Boyce - Codd
 - Criterios de Normalización

Dependencias funcionales

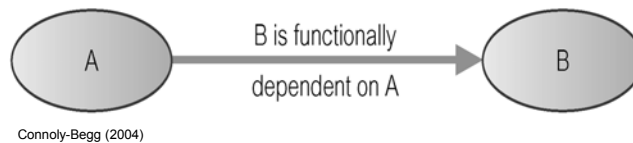
- La dependencia funcional describe la relación entre atributos.
- Ejemplo:
 - Si A y B son atributos de la relación R, B será funcionalmente dependiente de A ($A \rightarrow B$) si cada valor de A está asociado con exactamente un valor de B.
 - A y B pueden consistir cada uno de ellos de uno o más atributos.
 - A determina funcionalmente a B
- Es una propiedad del significado de los atributos de una relación.

M.C. Luis R. Basto - Base de Datos
CEL

3

Dependencias funcionales

- Sean A y B dos atributos en donde B es funcionalmente dependiente de A.
 - Si se conoce el valor de A, el valor de B es el mismo en todas las tuplas que tienen el valor de A dado.
 - Cuando dos tuplas tienen el mismo valor de A, también tiene el mismo valor de B.
 - Para un valor de B puede haber varios valores de A.



Connolly-Begg (2004)

M.C. Luis R. Basto - Base de Datos
CEL

4

Dependencias funcionales

- **Determinante:** Es un atributo o grupo de atributos en el lado izquierdo de la flecha que describe una dependencia funcional.

M.C. Luis R. Basto - Base de Datos
CEL

5

Dependencias funcionales

Staff

staffNo	sName	position	salary	branchNo
SL21	John White	Manager	30000	B005
SG37	Ann Beech	Assistant	12000	B003
SG14	David Ford	Supervisor	18000	B003
SA9	Mary Howe	Assistant	9000	B007
SG5	Susan Brand	Manager	24000	B003
SL41	Julie Lee	Assistant	9000	B005

Branch

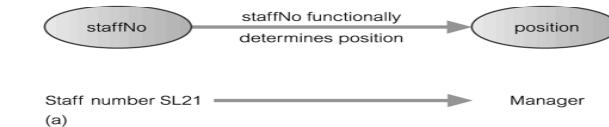
branchNo	bAddress
B005	22 Deer Rd, London
B007	16 Argyll St, Aberdeen
B003	163 Main St, Glasgow

M.C. Luis R. Basto - Base de Datos
CEL

6

Dependencias funcionales

- Sean los atributos staffNo y Position de la relación staff.
- Dado un valor para staffNo=SL21, se puede determinar la categoría de dicho empleado (Manager).
- StaffNo determina funcionalmente el valor de position.



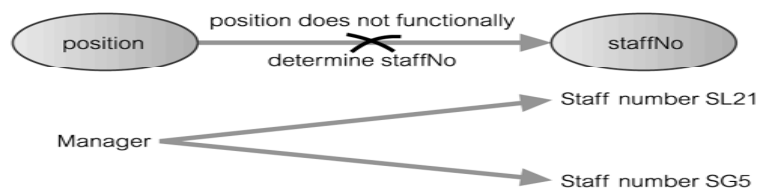
Connolly-Begg (2004)

M.C. Luis R. Basto - Base de Datos
CEL

7

Dependencias funcionales

- Position no determina funcionalmente staffNo. Un empleado tiene una categoría laboral, pero puede haber varios empleados con la misma categoría.



Connolly-Begg (2004)

M.C. Luis R. Basto - Base de Datos
CEL

8

Dependencias funcionales

- La relación entre staffNo y position es (1:1).
- La relación entre position y staffNo es de tipo uno a muchos (1:*).
- staffNo es el determinante de position.
- Para la normalización es importante identificar las dependencias funcionales entre atributos de una relación que tengan una relación uno a uno entre atributo(s) que forman el determinante en el lado izquierdo y atributo(s) en el lado derecho de la dependencia.

M.C. Luis R. Basto - Base de Datos
CEL

9

Dependencias funcionales

- Considere los atributos staffNo y sName de Staff.
- Para un valor de staffNo=SL21, se puede determinar el nombre de dicho empleado=John White.
- Para un valor de sName=Jhon white, se puede determinar el número de empleado=SL21.
- ¿El atributo staffNo depende funcionalmente de sName y sName depende funcionalmente de staffNo?

M.C. Luis R. Basto - Base de Datos
CEL

10

Dependencias funcionales

- Si los valores mostrados en la relación Staff representan el conjunto de todos los posibles valores de los atributos staffNo y sName, entonces se cumplen.

$staffNo \rightarrow sName$

$sName \rightarrow staffNo$

- Lo que se requiere es identificar dependencias funcionales que se cumplan para todos los posibles valores de los atributos de una relación, ya que representan los tipos de restricciones de integridad que se quiere identificar.

M.C. Luis R. Basto - Base de Datos
CEL

11

Dependencias funcionales

- Comprendamos el propósito de cada atributo.
 - staffNo identifica de manera única a un empleado.
 - sName sirve para almacenar los nombres de los empleados.
- ¿Qué pasa si hay empleados con el mismo nombre?
 - Para algunos empleados no se podrá determinar el número staffNo.

M.C. Luis R. Basto - Base de Datos
CEL

12

Dependencias funcionales

- La relación entre staffNo y sName es de tipo uno a uno (1:1).
 - Para cada número de empleado sólo hay un nombre.
- La relación entre sName y staffNo es de tipo uno a uno (1:*).
 - Puede haber varios números de empleados asociados con un nombre concreto.

$staffNo \rightarrow sName$

Dependencias funcionales

- Dependencia funcional completa
 - Indica que si A y B son atributos de una relación, B depende funcionalmente de A pero no de ningún subconjunto propio de A.
 - Una dependencia funcional $A \rightarrow B$ es una dependencia funcional completa si la eliminación de cualquier atributo de A hace que la dependencia deje de existir.
 - Una dependencia funcional $A \rightarrow B$ es una dependencia parcial si existe algún atributo que puede eliminarse de A y la dependencia continua verificándose.

Dependencias funcionales

- Consideremos la siguiente dependencia funcional en la relación Staff:

$staffNo, sName \rightarrow branchNo$

- Cada valor de (staffNo, sName) está asociado con un único valor de branchNo.
- branchNo depende funcionalmente de un subconjunto de (staffNo, sName), de staffNo.
- La dependencia entonces, es parcial no completa.

$staffNo \rightarrow branchNo$

M.C. Luis R. Basto - Base de Datos
CEL

15

Características de las dependencias funcionales

- Las dependencias funcionales que se utilizan en el proceso de normalización, tienen las sigs. Características:
 - Hay una relación uno a uno entre los atributos del lado izquierdo (determinante) y el lado derecho.
 - Se cumplen en todo instante de tiempo.
 - El determinante tiene el número mínimo de atributos necesarios para mantener la dependencia con los atributos del lado derecho (dependencia funcional completa).

M.C. Luis R. Basto - Base de Datos
CEL

16

Características de las dependencias funcionales

- Dependencia transitiva:
 - Una condición en la que A, B y C son atributos de una relación tales que si $A \rightarrow B$ y $B \rightarrow C$, entonces C depende transitivamente de A.

$staffNo \rightarrow sName, position, salary, branchNo, bAdress$

$branchNo \rightarrow bAdress$

Identificación de dependencias funcionales

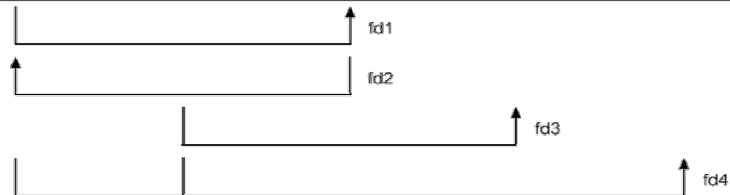
Sample Relation

A	B	C	D	E
a	b	z	w	q
e	b	r	w	p
a	d	z	w	t
e	d	r	w	q
a	f	z	s	t
e	f	r	s	t

Identificación de dependencias funcionales

Sample Relation

A	B	C	D	E
a	b	z	w	q
e	b	r	w	p
a	d	z	w	t
e	d	r	w	q
a	f	z	s	t
e	f	r	s	t



M.C. Luis R. Basto - Base de Datos
CEL

19

Identificación de dependencias funcionales

A → C (df1)
C → A (df2)
B → D (df3)
A, B → E (df4)

M.C. Luis R. Basto - Base de Datos
CEL

20

Identificación de dependencias funcionales

StaffBranch

staffNo	sName	position	salary	branchNo	bAddress
SL21	John White	Manager	30000	B005	22 Deer Rd, London
SG37	Azin Beech	Assistant	12000	B003	163 Main St, Glasgow
SG14	David Ford	Supervisor	18000	B003	163 Main St, Glasgow
SA9	Mary Howe	Assistant	9000	B007	16 Argyll St, Aberdeen
SG3	Susan Brend	Manager	24000	B003	163 Main St, Glasgow
SL41	Julie Lee	Assistant	9000	B005	22 Deer Rd, London

M.C. Luis R. Basto - Base de Datos
CEL

21

Identificación de dependencias funcionales

staffNo \rightarrow sName, position, salary, branchNo, bAddress

branchNo \rightarrow bAddress

bAddress \rightarrow branchNo

branchNo, position \rightarrow salary

bAddress, position \rightarrow salary

M.C. Luis R. Basto - Base de Datos
CEL

22

Identificación de la clave primaria

- Identificación de la clave primaria
 - Se trata de identificar, mediante el uso de dependencias funcionales, las claves candidatas y de ellas tomar la clave primaria.
 - La característica principal de una clave candidata es que los atributos de un determinante puedan determinar funcionalmente (individualmente o todos juntos) a **todos** los demás atributos de la relación.

Identificación de la clave primaria

- Identificación de la clave primaria
 - Relación StaffBranch

staffNo \rightarrow sName, position, salary, branchNo, bAddress
branchNo \rightarrow bAddress
bAddress \rightarrow branchNo
branchNo, position \rightarrow salary
bAddress, position \rightarrow salary

Identificación de la clave primaria

- Identificación de la clave primaria
 - Relación SampleRelation
- $A \rightarrow C$ (df1)
- $C \rightarrow A$ (df2)
- $B \rightarrow D$ (df3)
- $A, B \rightarrow E$ (df4)

Propósito de normalización

- Técnica para producir un conjunto de relaciones con una serie de propiedades deseables, partiendo de los requisitos de datos de una organización (Connolly, Begg).
- Técnica de diseño de bases de datos que examina las relaciones (dependencias funcionales) existentes entre los atributos.

Normalización

- Los atributos describen alguna propiedad de los datos o de las relaciones entre los datos.
- La normalización emplea una serie de pruebas (formas normales) para identificar el agrupamiento óptimo de los atributos, para identificar el conjunto de relaciones adecuados para un problema determinado.

M.C. Luis R. Basto - Base de Datos
CEL

27

Normalización

- Las características de un conjunto de relaciones adecuadas incluyen:
 - Número mínimo de atributos necesario para soportar los requisitos de datos de la organización.
 - Los atributos con relación lógica fuerte (dependencia funcional) se encuentran en la misma relación.
 - Redundancia mínima (excepción claves externas).

M.C. Luis R. Basto - Base de Datos
CEL

28

Normalización

- Ventajas
 - Fácil acceso
 - Mantenimiento adecuado y sencillo
 - Reduce la posibilidad de incoherencias
 - Se necesita menos espacio para almacenamiento.

M.C. Luis R. Basto - Base de Datos
CEL

29

Normalización

- Redundancia de datos y anomalías de actualización.

Staff Branch

staffNo	sName	position	salary	branchNo	bAddress
SL21	John White	Manager	30000	B005	22 Deer Rd, London
SG37	Ann Beech	Assistant	12000	B003	163 Main St, Glasgow
SG14	David Ford	Supervisor	18000	B003	163 Main St, Glasgow
SA9	Mary Howe	Assistant	9000	B007	16 Argyll St, Aberdeen
SG5	Susan Brand	Manager	24000	B003	163 Main St, Glasgow
SL41	Julie Lee	Assistant	9000	B005	22 Deer Rd, London

M.C. Luis R. Basto - Base de Datos
CEL

30

Normalización

Staff

staffNo	sName	position	salary	branchNo
SL21	John White	Manager	30000	B005
SG37	Ann Bech	Assistant	12000	B003
SG14	David Ford	Supervisor	18000	B003
SA9	Mary Howe	Assistant	9000	B007
SG5	Susan Brown	Manager	20000	B003
SL41	Julie Lee	Assistant	9000	B005

Branch

branchNo	bAddress
B005	22 Deer Rd, London
B007	16 Argyll St, Aberdeen
B003	163 Main St, Glasgow

M.C. Luis R. Basto - Base de Datos
CEL

31

Normalización

- Anomalías de actualización:
 - Anomalías de inserción
 - Anomalías de eliminación
 - Anomalías de modificación

M.C. Luis R. Basto - Base de Datos
CEL

32

Normalización

- Anomalías de inserción
 - Insertar un empleado dada una sucursal
 - Insertar una nueva sucursal que no tenga empleados.
- ¿Tienen el mismo problema las relaciones Staff y Branch?

Staff Branch

staffNo	sName	position	salary	branchNo	bAddress
SL21	John White	Manager	30000	B005	22 Deer Rd, London
SG37	Ann Beech	Assistant	12000	B003	163 Main St, Glasgow
SG14	David Ford	Supervisor	18000	B003	163 Main St, Glasgow
SA9	Mary Howe	Assistant	9000	B007	16 Argyll St, Aberdeen
SG5	Susan Brand	Manager	24000	B003	163 Main St, Glasgow
SL41	Julie Lee	Assistant	9000	B005	22 Deer Rd, London

M.C. Luis R. Basto - Base de Datos
CEL

33

Normalización

- Anomalías de eliminación
 - Eliminar una tupla de la tabla StaffBranch
- ¿Tienen el mismo problema las relaciones Staff y Branch?

Staff Branch

staffNo	sName	position	salary	branchNo	bAddress
SL21	John White	Manager	30000	B005	22 Deer Rd, London
SG37	Ann Beech	Assistant	12000	B003	163 Main St, Glasgow
SG14	David Ford	Supervisor	18000	B003	163 Main St, Glasgow
SA9	Mary Howe	Assistant	9000	B007	16 Argyll St, Aberdeen
SG5	Susan Brand	Manager	24000	B003	163 Main St, Glasgow
SL41	Julie Lee	Assistant	9000	B005	22 Deer Rd, London

M.C. Luis R. Basto - Base de Datos
CEL

34

Normalización

- Anomalías de modificación
 - Actualizar la sucursal B005
- ¿Tienen el mismo problema las relaciones Staff y Branch?

Staff Branch

staffNo	sName	position	salary	branchNo	bAddress
SL21	John White	Manager	30000	B005	22 Deer Rd, London
SG37	Ann Beech	Assistant	12000	B003	163 Main St, Glasgow
SG14	David Ford	Supervisor	18000	B003	163 Main St, Glasgow
SA9	Mary Howe	Assistant	9000	B007	16 Argyll St, Aberdeen
SG5	Susan Brand	Manager	24000	B003	163 Main St, Glasgow
SL41	Julie Lee	Assistant	9000	B005	22 Deer Rd, London

M.C. Luis R. Basto - Base de Datos
CEL

35

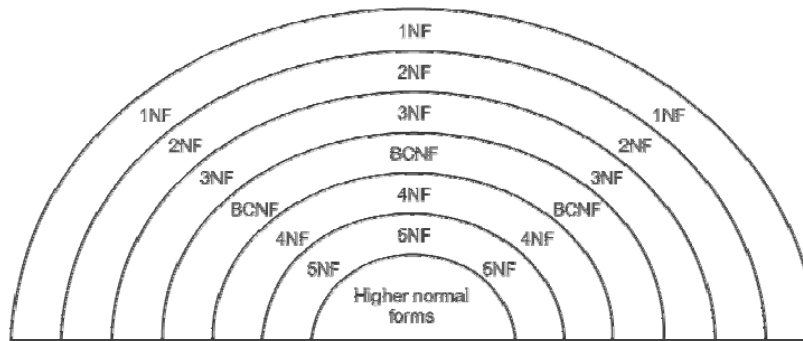
Normalización

- Descomposición de relaciones
 - Propiedad de combinación sin pérdidas
 - Garantiza que cualquier instancia de la relación original puede ser identificada a partir de las instancias correspondientes de las relaciones más pequeñas.
 - Propiedad de preservación de la dependencia
 - Garantiza que una restricción de la relación original puede mantenerse imponiendo alguna restricción a cada una de las relaciones más pequeñas.

M.C. Luis R. Basto - Base de Datos
CEL

36

Normalización



M.C. Luis R. Basto - Base de Datos
CEL

37

Primera Forma Normal

- Forma no normalizada: Tabla que contiene uno o más grupos repetidos.
- Primera Forma Normal (1FN)
 - Una relación en la que la intersección de toda fila y columna contiene un valor y sólo un valor.
- Para transformar una tabla no normalizada a 1FN se tiene que identificar y eliminar grupos repetidos.
 - Conjunto de atributos que contienen múltiples valores para un mismo valor de los atributos designados como clave principal.

M.C. Luis R. Basto - Base de Datos
CEL

38

Primera Form Normal

- Eliminar grupos repetidos de tablas no normalizadas:
 - Introducir datos apropiados en las columnas vacías de las filas que contienen los datos repetitivos.
 - Colocar los datos repetidos, junto con una copia de los atributos clave originales en una relación independiente.

M.C. Luis R. Basto - Base de Datos
CEL

39

Primera Form Normal

OrdNo	Date	CustNo	Name	Address	ProdNo	Desc	Price	Qty
1	5/1/05	22	Smith	London	A95	Jacket	55	4
					G17	Coat	120	8
					K10	Suit	90	5
2	19/1/05	47	Jones	Paris	G17	Coat	120	9
					D77	Shirt	35	20
3	27/3/05	25	West	Glasgow	E30	Tie	5	25
					D77	Shirt	35	4

M.C. Luis R. Basto - Base de Datos
CEL

40

Segunda Forma Normal

- Basada en la dependencia funcional completa.
- Se aplica a relaciones con claves compuestas.
- La relación con clave principal de un único atributo está automáticamente en forma 2FN.
- 2NF
 - Una relación que está en 1FN y en la que todo atributo que no sea clave principal depende funcionalmente de manera completa de la clave principal.

Segunda Forma Normal

- Para transformar las relaciones a 2FN implica la eliminación de las dependencias parciales.
- Si hay dependencia parcial, se elimina de la relación los atributos parcialmente dependientes, situándolos en una nueva relación junto con la copia del determinante.

Segunda Form Normal

Df1	clientNo,propertyNo→rentStart,rentFish	Cve. primaria
Df2	clientNo→cName	Dep. parcial
Df3	propertyNo→pAddress, rent, ownerNo, oName	Dep. parcial
Df4	ownerNo→oName	Dep. transitiva
Df5	clientNo, rentStart→propertyNo, pAddress, rentFinish, rent, ownerNo, oName.	Cve. Candidata
Df6	propertyNo, rentStar→clientNo, cName, rentFinish	Cve. candidata

M.C. Luis R. Basto - Base de Datos
CEL

43

Segunda Form Normal

- Client (clientNo, cName)
- Rental (clientNo, propertyNo, rentStar, rentFinish).
- PropertyOwner(propertyNo, pAddress, rent, ownerNo, oName)

M.C. Luis R. Basto - Base de Datos
CEL

44

Tercera Forma Normal

clienNo	cName
CR76	Jhon kay
CR56	Aline Stewart

clientNo	propertyNo	rentStart	rentFinish
CR76	PG4	01-Jul-03	31-Ago-04
CR76	PG16	02-Jul-03	01-Sep-04
CR56	PG4	03-Jul-03	02-Sep-04
CR56	PG36	04-Jul-03	03-Sep-04
CR56	PG16	05-Jul-03	04-Sep-04

propertyNo	pAddress	rent	ownerNo	oName
PG4	6 Lawrence St, Glasgow	350	CO40	Tina Murphy
PG16	5 Novar Dr, Glasgow	450	CO93	Tony Shaw
PG36	2 Manor Rd, Glasgow	375	CO93	Tony Shaw

M.C. Luis R. Basto - Base de Datos
CEL

45

Tercera Forma Normal

- Se necesita eliminar la dependencia transitiva para transformar a la 3FN.
- 3FN:
 - Una relación que está en primera y segunda formas normales y en la que ningún atributo que no sea de clave principal depende transitivamente de la clave principal.
- Si existe una dependencia transitiva, se elimina de la relación los atributos que dependen transitivamente, situándolos en una nueva relación junto con una copia del determinante.

M.C. Luis R. Basto - Base de Datos
CEL

46

Tercera Forma Normal

propertyNo	pAddress	rent	ownerNo
PG4	6 Lawrence St, Glasgow	350	CO40
PG16	5 Novar Dr, Glasgow	450	CO93
PG36	2 Manor Rd, Glasgow	375	CO93

ownerNo	oName
CO40	Tina Murphy
CO93	Tony Shaw

M.C. Luis R. Basto - Base de Datos
CEL

47

Forma Normal de Boyce-Codd (BCNF)

- La forma normal de Boyce-Codd (BCNF, por sus siglas en inglés) está basada en dependencias funcionales que toman en consideración todas las claves candidatas de una relación; sin embargo, BCNF tiene también restricciones adicionales, si se compara con la definición general de 3NF.

M.C. Luis R. Basto - Base de Datos
CEL

48

Forma Normal de Boyce-Codd (BCNF)

- La definición de BCNF dicta: Una relación está en BCNF, si y solo sí, todo determinante es una clave candidata.
- Para comprobar si una relación está en BCNF se identifican todos los determinantes y se comprueba que sean claves candidatas.

Forma Normal de Boyce-Codd (BCNF)

- La diferencia entre 3FN y BCNF es que para una dependencia funcional $A \rightarrow B$, la 3NF permite esta dependencia en una relación si B es un atributo de clave primaria y A no es una clave candidata.
- Mientras que, BCNF insiste en que para que esta dependencia permanezca en una relación, A debería ser una clave candidata.

Forma Normal de Boyce-Codd (BCNF)

- Toda relación en BCNF está también en 3NF. Sin embargo, una relación en 3NF no necesariamente está en BCNF.
- Consideremos de nuevo las relaciones Client, Rental, PropertyForRent y Owner.

Client(clientNo, cName)
Rental(clientNo, propertyNo, rentStart, rentFish)
PropertyForRent (propertyNo, pAddress, rent, ownerNo)
Owner(ownerNo, oName)

M.C. Luis R. Basto - Base de Datos
CEL

51

Forma Normal de Boyce-Codd (BCNF)

- Las relaciones PropertyForRent y Owner están en BCNF, ya que cada relación tiene un único determinante, que es la clave candidata.
- La relación Rental tiene a los tres determinantes (clientNo, propertyNo), (clientNo, rentStart) y (propertyNo, rentStart) tal y como sigue:

clientNo,propertyNo→rentStart,rentFish
clientNo, rentStart→propertyNo, rentFinish
propertyNo, rentStar→clientNo, rentFinish

M.C. Luis R. Basto - Base de Datos
CEL

52

Forma Normal de Boyce-Codd (BCNF)

- Como los tres determinantes de la relación Rental son también claves candidatas, la relación Rental está también en BCNF.
- La violación de BCNF es bastante raro que se dé, ya que sólo puede producirse bajo condiciones muy específicas.

Forma Normal de Boyce-Codd (BCNF)

- La posibilidad de violar las condiciones de BCNF puede aparecer cuando:
 - La relación contenga dos o más claves candidatas compuestas; o
 - Las claves candidatas solapen, es decir, tenga al menos un atributo en común.