

Bases de Datos 1

Alejandra Lliteras

alejandra.lliteras@lifa.info.unlp.edu.ar



En la clase anterior...

Teoría de diseño de BBDD relacionales – EJERCICIO–

RECITALES(idRecital, idGrupo, idIntegranteGrupo, marcaInstrumento, vocalista, idOrganizador, nombreGrupo, nombreOrganizador)

Donde

- ▶ en un recital se presentan diversos grupos y un grupo se presenta en diversos recitales
- ▶ cada grupo tiene diversos integrantes. Los integrantes del grupo pueden variar para diferentes recitales
- ▶ cada grupo tiene solamente un vocalista, el vocalista del grupo no varía para los diferentes recitales en los que el grupo se presenta
- ▶ de un grupo se conoce el nombre, pero puede suceder que dos grupos se llamen de igual manera
- ▶ de cada integrante de un grupo y para cada recital en los que el grupo se presenta, se sabe que marca de instrumento uso el integrante
- ▶ un integrante de un grupo puede pertenecer a más de un grupo
- ▶ un vocalista de un grupo puede ser vocalista de otros grupos
- ▶ cada recital tiene diversos organizadores y un organizador puede organizar diversos recitales
- ▶ el nombre del organizador se puede repetir para diferentes idOrganizador. El idOrganizador es único

Realizar el proceso de normalización para llevar RECITALES a 4FN

Teoría de diseño de BBDD relacionales – EJERCICIO–

- ▶ RECITALES (idRecital, idGrupo, idIntegranteGrupo, marcaInstrumento, vocalista, idOrganizador, nombreGrupo, nombreOrganizador)

Dep. funcionales:

df1) idGrupo → nombreGrupo, vocalista

df2) idRecital, idGrupo, idIntegranteGrupo → marcaInstrumento

df3) idOrganizador → nombreOrganizador

Clave candidata:

{idRecital, idGrupo, idIntegranteGrupo, idOrganizador}

Una vez realizados los puntos mencionados como 1 y 2 del proceso, llevar RECITALES a BCNF

Una vez explicado y justificado el proceso hasta BCNF, se deben hallar las dependencias multivaluadas sobre la última partición en la que quedo la clave primaria

Teoría de diseño de BBDD relacionales – EJERCICIO–

- ▶ RECITALES (idRecital, idGrupo, idIntegranteGrupo, marcaInstrumento, vocalista, idOrganizador, nombreGrupo, nombreOrganizador)

Dep. funcionales:

df1) idGrupo → nombreGrupo, vocalista

df2) idRecital, idGrupo, idIntegranteGrupo → marcaInstrumento

df3) idOrganizador → nombreOrganizador

Clave candidata:

{idRecital, idGrupo, idIntegranteGrupo, idOrganizador}

Dep. Multivaluadas:

DM1) idRecital, idGrupo →> idIntegranteGrupo

DM2) idRecital →> idOrganizador

Una vez halladas las dependencias multivaluadas, realizar el proceso de normalización partiendo de un esquema en BCNF y llevando a 4FN

Ejercicio Grupal

EMPRESA(#Empleado, nombreEmpleado, #Area, nombreArea, tipoMaquina, nombreMaquina)

Sobre el esquema propuesto valen las siguientes restricciones:

- ▶ Cada área tiene varios empleados.
- ▶ Cada empleado trabaja en una sola área.
- ▶ Cada área tiene un solo tipo de máquina.
- ▶ Dado el nombre de una máquina puedo recuperar el tipo de la maquina
- ▶ Un tipo de máquina puede tener muchos nombres
- ▶ Por área existe al menos un empleado asignado

EMPRESA(#Empleado, nombreEmpleado, #Area, nombreArea, tipoMaquina, nombreMaquina)

Dependencias Funcionales

- ▶ df1 – nombreMaquina \rightarrow tipoMaquina
- ▶ df2 – #Empleado \rightarrow nombreEmpleado, #Area
- ▶ df3 – #Area \rightarrow nombreArea, tipoMaquina

Claves candidatas:

cc1 : {#Empleado, nombreMaquina}

EMPRESA(#Empleado, nombreEmpleado, #Area, nombreArea, tipoMaquina, nombreMaquina)

Dependencias Funcionales

- ▶ df1 – nombreMaquina → tipoMaquina
- ▶ df2 – #Empleado → nombreEmpleado, #Area
- ▶ df3 – #Area → nombreArea, tipoMaquina

Claves candidatas:

cc1: {#Empleado, nombreMaquina}

¿EMPRESA está en BCNF?

- ▶ El esquema **EMPRESA** no está en BCNF porque al menos uno de los antecedentes de las dependencias funcionales 1–3 no es superclave en el esquema **EMPRESA** (Por ejemplo: la dependencia funcional 1). Entonces divido la tabla utilizando la dependencia funcional 1 quedando lo siguiente:
- ▶ E1(nombreMaquina , tipoMaquina)
- ▶ E2(#Empleado, nombreEmpleado, #Area, nombreArea, nombreMaquina)

EMPRESA(#Empleado, nombreEmpleado, #Area, nombreArea, tipoMaquina, nombreMaquina)

- ▶ df1 – nombreMaquina → tipoMaquina
- ▶ df2 – #Empleado → nombreEmpleado, #Area
- ▶ df3 – #Area → nombreArea, tipoMaquina

- ▶ E1(nombreMaquina , tipoMaquina)
- ▶ E2(#Empleado, nombreEmpleado, #Area, nombreArea, nombreMaquina)

- ▶ El esquema E1 está en BCNF ya que el antecedente de la dependencia funcional 1 es superclave en E1 y vale la dependencia funcional 1 la cual se uso en el proceso de división.
- ▶ El problema que se plantea con este particionamiento es si la dependencia funcional 3 sigue valiendo en el esquema E2. Para analizar esto debo usar el algoritmo para verificar pérdida de dependencias funcionales.

- Se quiere analizar si se perdió la dependencia funcional **#Area \rightarrow nombreArea, tipoMaquina**
Para esto aplico el algoritmo para determinar pérdida de dependencias funcionales

RES: (#Area)

Mientras Res cambia

Para $i = 1$ to cant_de_particiones_realizadas

$$\text{Res} = \text{Res} \cup ((\text{Res} \cap R_i)^+ \cap R_i)$$

#Area → nombreArea, tipoMaquina

- ▶ **E1(nombreMaquina , tipoMaquina)**
- ▶ **E2(#Empleado, nombreEmpleado, #Area, nombreArea, nombreMaquina)**

Desde $i = 1$ hasta 2

Paso 1)

Res =

(#Area)

\cup

((#Area)

\cap

(nombreMaquina, tipoMaquina))⁺

\cap (nombreMaquina , tipoMaquina))

Res = (#Area) \cup

((\emptyset)⁺

\cap (nombreMaquina , tipoMaquina))

Res = (#Area) \cup

(\emptyset)

Res = (#Area)

#Area → nombreArea, tipoMaquina

- ▶ **E1(nombreMaquina , tipoMaquina)**
- ▶ **E2(#Empleado, nombreEmpleado, #Area, nombreArea, nombreMaquina)**

Desde $i = 1$ hasta 2

Paso 2)

Res = (#Area)

Res = (#Area) \cup

((#Area)

\cap (#Empleado, nombreEmpleado, #Area, nombreArea, nombreMaquina))⁺

\cap (#Empleado, nombreEmpleado, #Area, nombreArea, nombreMaquina))

Res = (#Area) \cup

((#Area))⁺ \cap

(#Empleado, nombreEmpleado, #Area, nombreArea, nombreMaquina))

–Debo hallar la clausura del conjunto de atributos (#Area)

Result:= X

While (hay cambios en result) do

For (cada dependencia funcional $Y \rightarrow Z$ en F) do

if ($Y \subseteq \text{result}$) then

result := result \cup Z

result : (#Area)

para cada dependencia funcional del conjunto

{nombreMaquina \rightarrow tipoMaquina

#Empleado \rightarrow nombreEmpleado, #Area

#Area \rightarrow nombreArea, tipoMaquina}

Entra al while por primera vez y result= (#Area, nombreArea, tipoMaquina)

Como result cambio, entro una vez mas al while

result : (#Area, nombreArea, tipoMaquina)

para cada dependencia funcional del conjunto

{nombreMaquina \rightarrow tipoMaquina

#Empleado \rightarrow nombreEmpleado, #Area

#Area \rightarrow nombreArea, tipoMaquina}

En esta iteración, result queda igual y termino la iteración

–Una vez hallada la clausura del conjunto de atributos (#Area)

(#Area, nombreArea, tipoMaquina)

Se remplaza en la siguiente expresión

Res = (#Area) \cup

((#Area))⁺ \cap

(#Empleado, nombreEmpleado, #Area, nombreArea, nombreMaquina))

Res = (#Area) \cup

((#Area, nombreArea, tipoMaquina) \cap

(#Empleado, nombreEmpleado, #Area, nombreArea, nombreMaquina))

Res = (#Area) \cup

(#Area, nombreArea)

Res = (#Area, nombreArea)

Como se paso por las dos particiones y Res cambió, entonces se vuelve a iterar

Res = (#Area, nombreArea)

Trabajo con la primer partición realizada

Res = (#Area, nombreArea)

\cup

((#Area, nombreArea)

\cap

(nombreMaquina, tipoMaquina))⁺

\cap (nombreMaquina , tipoMaquina))

Res = (#Area , nombreArea) \cup

((\emptyset)⁺

\cap (nombreMaquina , tipoMaquina))

Res = (#Area , nombreArea) \cup

(\emptyset)

Res = (#Area , nombreArea)

Res = (#Area, nombreArea)

Trabajo con la segunda partición realizada

Res = (#Area, nombreArea)

Res = (#Area , nombreArea) \cup

((#Area , nombreArea)

\cap (#Empleado, nombreEmpleado, #Area, nombreArea, nombreMaquina))⁺

\cap (#Empleado, nombreEmpleado, #Area, nombreArea, nombreMaquina))

Res = (#Area nombreArea) \cup

((#Area, nombreArea))⁺ \cap

(#Empleado, nombreEmpleado, #Area, nombreArea, nombreMaquina))

–Debo hallar la clausura del conjunto de atributos (#Area, nombreArea)

Result:= X

While (hay cambios en result) do

For (cada dependencia funcional $Y \rightarrow Z$ en F) do

if ($Y \subseteq \text{result}$) then

result := result \cup Z

result : (#Area, nombreArea)

para cada dependencia funcional del conjunto

{nombreMaquina \rightarrow tipoMaquina

#Empleado \rightarrow nombreEmpleado, #Area

#Area \rightarrow nombreArea, tipoMaquina}

Entra al while por primera vez y result= (#Area, nombreArea)

–Una vez hallada la clausura del conjunto de atributos (#Area, nombreArea)
(#Area, nombreArea)

Se remplacea en la siguiente expresión

$$\text{Res} = (\#Area, \text{nombreArea}) \cup$$
$$((\#Area, \text{nombreArea}))^+ \cap$$
$$(\#Empleado, \text{nombreEmpleado}, \#Area, \text{nombreArea}, \text{nombreMaquina}))$$
$$\text{Res} = (\#Area, \text{nombreArea}) \cup$$
$$((\#Area, \text{nombreArea}) \cap$$
$$(\#Empleado, \text{nombreEmpleado}, \#Area, \text{nombreArea}, \text{nombreMaquina}))$$
$$\text{Res} = (\#Area, \text{nombreArea}) \cup$$
$$(\#Area, \text{nombreArea})$$
$$\text{Res} = (\#Area, \text{nombreArea})$$

Como se paso por las dos particiones y Res no cambi6, no se vuelve a iterar.

Luego de aplicar el algoritmo no se logro incorporar al conjunto res, el atributo tipoMaquina de la dependencia funcional que estamos validando si se perdi6

#Area → nombreArea, tipoMaquina

Luego de aplicar el algoritmo no se logro incorporar al conjunto res, el atributo tipoMaquina de la dependencia funcional que estamos validando si se perdió

#Area \rightarrow nombreArea, tipoMaquina

Se concluye que la dependencia funcional **#Area \rightarrow nombreArea, tipoMaquina** se pierde con la partición propuesta para llevar a BCNF.

Por lo anterior, se deja el esquema en 3FN

Las dependencias funcionales que aun no se trataron hasta el esquema **EMPRESA** son:

nombreMaquina \rightarrow tipoMaquina

#Empleado \rightarrow nombreEmpleado, #Area

#Area \rightarrow nombreArea, tipoMaquina

Particiones que surgen con el análisis de 3FN

R1_{3FN} (nombreMaquina, tipoMaquina)

R2_{3FN} (#Empleado, nombreEmpleado, #Area)

R3_{3FN} (#Area, nombreArea, tipoMaquina)

R4_{3FN} (#Empleado, nombreMaquina)

Particiones que quedan en 3FN

$R1_{3FN}$ (nombreMaquina, tipoMaquina)

$R2_{3FN}$ (#Empleado, nombreEmpleado, #Area)

$R3_{3FN}$ (#Area, nombreArea, tipoMaquina)

$R4_{3FN}$ (#Empleado, nombreMaquina)

Clave Primaria del esquema

cp: {#Empleado, nombreMaquina}

$R4_{3FN}$ cumple con 4FN?

R4_{3FN} cumple con 4FN?

EMPRESA(#Empleado, nombreEmpleado, #Area, nombreArea, tipoMaquina, nombreMaquina)

Sobre el esquema propuesto valen las siguientes restricciones:

- ▶ Cada área tiene varios empleados.
- ▶ Cada empleado trabaja en una sola área.
- ▶ Cada área tiene un solo tipo de máquina.
- ▶ Dado el nombre de una máquina puedo recuperar el tipo de la maquina
- ▶ Un tipo de máquina puede tener muchos nombres
- ▶ Por área existe al menos un empleado asignado

Dependencias Multivaluadas halladas en R4_{3FN}:

DM1) $\emptyset \twoheadrightarrow \#Empleado$

DM2) $\emptyset \twoheadrightarrow nombreMaquina$

Dependencias Multivaluadas halladas en $R4_{3FN}$:

DM1) $\emptyset \twoheadrightarrow \#Empleado$

DM2) $\emptyset \twoheadrightarrow nombreMaquina$

El esquema $R4_{3FN}$ no esta en 4NF porque existen dependencias multivaluadas que no son triviales en $R4_{3FN}$.

Entonces divido la tabla $R4_{3FN}$ utilizando la dependencia multivaluada 1 quedando lo siguiente:

$R4.1_{4FN}$ ($\#Empleado$)

$R4.2_{4FN}$ ($nombreMaquina$)

El esquema $R4.1_{4FN}$ esta en 4NF ya que sólo vale la dependencia multivaluada 1 que es trivial.

El esquema $R4.2_{4FN}$ esta en 4NF ya que sólo vale la dependencia multivaluada 2 que es trivial.



R4.1_{4FN} (#Empleado)

R4.2_{4FN} (nombreMaquina)

El esquema **R4.1_{4FN}** esta en 4NF ya que sólo vale la dependencia multivaluada 1 que es trivial.

El esquema **R4.2_{4FN}** esta en 4NF ya que sólo vale la dependencia multivaluada 2 que es trivial.

Una vez llevadas las particiones a 4FN, escribir de manera explicita al final de todo el proceso cuales son las tablas que considera que han quedado en 4FN

Tablas en 4FN: **R1_{3FN}** , **R2_{3FN}** , **R3_{3FN}** , **R4.1_{4FN}** , **R4.2_{4FN}**

Las particiones **R1_{3FN}** , **R2_{3FN}** , **R3_{3FN}** se encuentran en 4FN, ya que no tienen dependencias multivaluadas

Retomando el AR como lenguaje de consulta

Ejercicio

Empleado(dni, nombre, domicilio, salario)

¿Qué empleados cobran el máximo valor de salario?

Ejercicio

Empleado(dni, nombre, domicilio, salario)

¿Qué empleados cobran el máximo valor de salario?

- **Buscar el valor del salario máximo**
 - **Busco los salarios que están por debajo del resto**
 - **A todos los salarios le resto los mínimos**
- **Relaciono el máximo salario con las personas que lo perciben**

Empleado(dni, nombre, domicilio, salario)

¿Qué empleados cobran el máximo valor de salario?

- **Buscar el valor del salario máximo**
 - **Busco los salarios que están por debajo del resto**

$\rho_{EMP(dni,nom,sal)}$ (Empleado)

$EMPLE \leftarrow \Pi_{dni, nombre, salario} (Empleado)$

$EmplBajoSalario \leftarrow \sigma_{salario < sal} (EMPLE \times EMP)$

$SalarioDebajoMaximo \leftarrow \Pi_{salario} (EmplBajoSalario)$

EMPLE

dni	nombre	salario
123456	Juan	10200
456789	Pedro	15000
567890	María	22000
234567	Joaquín	28000
345678	Martina	25000
678901	Mario	28000

EMP

dni	nom	sal
123456	Juan	10200
456789	Pedro	15000
567890	María	22000
234567	Joaquín	28000
345678	Martina	25000
678901	Mario	28000

EmplBajoSalario ←

$\sigma_{emple.salario < sal (EMPLEX EMP)}$

EMPLE

dni	nombre	salario
123456	Juan	10200
456789	Pedro	15000
567890	María	22000
234567	Joaquín	28000
345678	Martina	25000
678901	Mario	28000

EMP

dni	nom	sal
123456	Juan	10200
456789	Pedro	15000
567890	María	22000
234567	Joaquín	28000
345678	Martina	25000
678901	Mario	28000

EmplBajoSalario

EMPLE.dni	nombre	salario	EMP.dni	nom	sal
123456	Juan	10200	456789	Pedro	15000
123456	Juan	10200	567890	María	22000
123456	Juan	10200	234567	Joaquín	28000
123456	Juan	10200	345678	Martina	25000
123456	Juan	10200	678901	Mario	28000
456789	Pedro	15000	567890	María	22000
456789	Pedro	15000	234567	Joaquín	28000
456789	Pedro	15000	345678	Martina	25000
456789	Pedro	15000	678901	Mario	28000
567890	María	22000	234567	Joaquín	28000
567890	María	22000	345678	Martina	25000
567890	María	22000	678901	Mario	28000
345678	Martina	25000	234567	Joaquín	28000
345678	Martina	25000	678901	Mario	28000

SalarioDebajoMaximo $\leftarrow \Pi_{\text{salario}} (\text{EmplBajoSalario})$

EmplBajoSalario

EMPLE.dni	nombre	salario	EMP.dni	nom	sal
123456	Juan	10200	456789	Pedro	15000
123456	Juan	10200	567890	María	22000
123456	Juan	10200	234567	Joaquín	28000
123456	Juan	10200	345678	Martina	25000
123456	Juan	10200	678901	Mario	28000
456789	Pedro	15000	567890	María	22000
456789	Pedro	15000	234567	Joaquín	28000
456789	Pedro	15000	345678	Martina	25000
456789	Pedro	15000	678901	Mario	28000
567890	María	22000	234567	Joaquín	28000
567890	María	22000	345678	Martina	25000
567890	María	22000	678901	Mario	28000
345678	Martina	25000	234567	Joaquín	28000
345678	Martina	25000	678901	Mario	28000

SalarioDebajoMaximo

salario
10200
15000
22000
25000

Empleado(dni, nombre, domicilio, salario)

¿Qué empleados cobran el máximo valor de salario?

- **Buscar el valor del salario máximo**
 - Busco los salarios que están por debajo del resto
 - **A todos los salarios le resto los mínimos**

MaximoSalario $\leftarrow \Pi_{\text{salario}}(\text{Empleado}) - \text{SalarioDebajoMaximo}$

$\text{maximoSalario} \leftarrow \Pi_{\text{salario}}(\text{Empleado}) - \text{salarioDebajoMaximo}$

$\Pi_{\text{salario}}(\text{Empleado})$

salario
10200
15000
22000
28000
25000

SalarioDebajoMaximo

salario
10200
15000
22000
25000

MaximoSalario

salario
28000

Empleado(dni, nombre, domicilio, salario)

¿Qué empleados cobran el máximo valor de salario?

- **Buscar el valor del salario máximo**
 - Busco los salarios que están por debajo del resto
 - A todos los salarios le resto los mínimos
- **Relaciono el máximo salario con las personas que lo perciben**

Resultado ← Empleado |X| MaximoSalario

Empleado(dni, nombre, domicilio, salario)

¿Qué empleados cobran el máximo valor de salario?

- Buscar el valor del salario máximo
 - Busco los salarios que están por debajo del resto
 - A todos los salarios le resto los mínimos
- Relaciono el máximo salario con las personas que lo perciben

Resultado ← **Empleado** |X| **MaximoSalario**

Empleado

dni	nombre	domicilio	salario
123456	Juan	1 y 36	10200
456789	Pedro	154 y 78	15000
567890	María	22 y 61	22000
234567	Joaquín	12 y 62	28000
345678	Martina	58 y 7	25000
678901	Mario	120 y 50	28000

MaximoSalario

salario
28000

Resultado

dni	nombre	domicilio	salario
234567	Joaquín	12 y 62	28000
678901	Mario	120 y 50	28000