

Laboratorio de Base de Datos Prelaboratorio Nro. 8, Normalización

Prof. Solazver Solé
Preps. Victor Albornoz, Yenifer Ramirez
Semestre B-2018

1. Normalización

Cuando diseñamos una base de datos, el principal objetivo es crear una representación adecuada de los datos, relaciones entre los datos y restricciones sobre los datos relativas al problema.

La normalización es una técnica de base de datos, que comienza examinando las relaciones entre los atributos. (**Dependencias funcionales**).

La normalización usa una serie de pruebas (descritas como **formas normales**), para ayudar a identificar el mejor/optimo agrupamiento de los atributos y de esta manera identificar el conjunto de relaciones que permiten dar soporte a los requerimientos del problema o de la descripción del minimundo.

1.1. Proposito de la Normalizacion

- Producir un conjunto de relaciones con propiedades deseadas, dado los requerimientos de un problema.
- Identificar un conjunto propicio de relaciones que soporten los requerimientos de un problema.
- Identificar el mínimo conjunto de atributos necesarios para soportar los requerimientos
- Identificar los atributos con una relación cercana (dependencias funcionales) encontrados en una misma relación.
- Minimizar la redundancia de cada atributo, representándolo una sola vez, con la excepción de los atributos que forman parte de la claves foráneas

1.2. Beneficios de la normalización

Los beneficios usar un conjunto de relaciones propicio a los requerimientos (normalizado), es que permite una base de datos amigable en acceso y mantenimiento de los datos, tomando la mínima cantidad de espacio en la computadora.

No obstante, la normalización permite abordar dos problemas fundamentales en las bases de datos relacionales: **Redundancia de datos** y **anomalías de actualización**.

- Las actualizaciones de datos almacenados en la base de datos son logradas con el mínimo número de operaciones, reduciendo la probabilidad inconsistencias en la base de datos.
- Reducción del espacio de almacenamiento de archivos requeridos por la relaciones, minimizando costos.

1.3. Anomalías de actualización

Las anomalías de actualización se producen en tablas **desnormalizadas** con redundancia de datos y están calificadas en: inserción, eliminación y modificación

PERSONALSUCURSAL

NoPersonal	NombrePersonal	Cargo	Salario	NoOficina	DireccionOficina
SL21	Antonio Rojas	Gerente	30000	B005	22
SG37	José Lopez	Asistente	12000	B003	163
SG14	Raul Almeida	Supervisor	18000	B003	163
SA9	Georgios Botas	Asistente	9000	B007	16
SG5	Dimitri Perez	Gerente	24000	B003	163
SL41	Deiby Matos	Asistente	9000	B005	22

En el estado de la relación **PERSONALSUCURSAL** podemos observar las siguientes anomalías:

1.3.1. Anomalías de inserción:

- Para insertar los detalles de un nuevo personal en la relación **PERSONALSUCURSAL**, debemos incluir la sucursal donde la persona trabaja.
- Para insertar los detalles de una nueva sucursal que actualmente no tiene personal, debemos insertar valores nulos en los atributos para el personal.

1.3.2. Anomalías de eliminación:

- Si eliminamos una tupla de la relación **PERSONALSUCURSAL**, que represente el último miembro del personal en una sucursal dada, los detalles de la sucursal serían borrados de la base de datos.

1.3.3. Anomalías de modificación:

- Si quisiéramos cambiar el valor de un atributo de una sucursal particular en la relación **PERSONALSUCURSAL**, por ejemplo, el numero de oficina **B003**, debemos actualizar el numero de todo el personal que trabaja en esa sucursal.

2. Dependencias Funcionales

Una dependencia funcional es una propiedad de la semántica de los atributos relacionados con otros.

- Describe la relación entre atributos en una relación $R(A,B,C,..Z)$
 $A \longrightarrow B$.
 - B es funcionalmente dependiente de A.
 - Cada valor de A es asociado con un solo valor de B.
- La semántica indica como los atributos se relacionan unos con otros y especifica la dependencia funcional.
- Cuando una dependencia funcional esta presente, la dependencia es especificada como una **restricción entre atributos**.
- Si $A \longrightarrow B$.
 - Si conocemos el valor de **A** y examinamos el estado de la relación que mantiene esta dependencia, solo encontramos un valor de **B** en todas las tuplas que tengan el mismo valor de A en cualquier momento .
 - Cuando dos tuplas tiene el mismo valor de **A**, también tiene el mismo valor de **B**.

PERSONALSUCURSAL

NoPersonal	NombrePersonal	Cargo	Salario	NoOficina	DireccionOficina
SL21	Antonio Rojas	Generente	30000	B005	22
SG37	José Lopez	Asistente	12000	B003	163
SG14	Raul Almeida	Supervisor	18000	B003	163
SA9	Georgios Botas	Asistente	9000	B007	16
SG5	Dimitri Perez	Gerente	24000	B003	163
SL41	Deiby Matos	Asistente	9000	B005	22

- Identifique que otras dependencias funcionales existen en el estado de la relación **PERSONALSUCURSAL**

Respuestas:

- **NoPersonal** ->NombrePersonal,Cargo,Salario,NoOficina,DireccionOficina
- **NoOficina** ->DireccionOFicina.
- **DireccionOficina** ->NoOficina.
- **NoOficina, Cargo** ->Salario.
- **DireccionOficina, Cargo** ->Salario.

Para propósitos de la normalización, debemos identificar las dependencias funcionales entre los atributos de una relación, que tienen una relación **1:1** entre los atributos.

Una dependencia funcional, es una propiedad de un esquema relacional y no una propiedad de una instancia particular de un esquema(extensión).

2.1. Dependencia Funcional completa

- Una característica adicional de las dependencias funcionales, que es útil para la **normalización**, es que sus **determinantes** deben tener el mínimo numero de atributos necesarios para mantener la dependencia funcional con los atributos del lado derecho (**dependientes funcionalmente**). A esto se le conoce como **dependencia funcional completa**.
- Si $A \longrightarrow B$.
 - B es completamente y funcionalmente determinada por A, si B depende de A, pero no de ningún subconjunto propio de A.

2.2. Dependencia funcional transitiva

- Una característica adicional, útil para la normalización es la dependencia transitiva.
- Si A, B y C son atributos de una relación, tal que:
 $A \longrightarrow B$ y $B \longrightarrow C$, entonces:
 - C es transitivamente dependiente de A a traves de B . $A \longrightarrow C$

3. Conceptos importantes para realizar la normalización

- **Superclave:** Un atributo o conjunto de atributos, que identifican de manera única una tupla en una relación.
 - Una superclave identifica de manera única cada tupla dentro de una relación. Sin embargo, una superclave puede contener atributos adicionales que no son necesarios para la identificación única
- **Clave candidata:** Una superclave tal que ningún sub-conjunto propio es una superclave dentro de una relación.
- Una clave candidata K de una relacion R tiene dos propiedades:
 - **-Unicidad:** En cada tupla de la relación R, los valores de K identifican de manera única esa tupla.
 - **-Irreducibilidad:** Ningún subconjunto propio de K tiene la propiedad de unicidad.
- Pueden existir varias claves candidatas en una relación. Esto **influye en el proceso de normalización**.
- Si un esquema de relación tiene más de una clave, todas se denominan **claves candidatas**.
- Una de las claves candidatas se designa “arbitrariamente”(la que mejor represente de manera univoca la relación) como clave primaria y las demás se denominan **claves secundarias**.
- Todo esquema de relación debe tener una clave primaria.
- Un atributo es un esquema de relación R, se denomina **primo** de si es miembro de alguna clave de R.
- Un atributo es **no primo** si no es atributo primo, es decir, si no es miembro de ninguna clave candidata.