

@author: aar.velasquez@gmail.com

@date: 2020/10/16

## Capítulo 7: Diseño de base de datos relacionales.

---

- En este capítulo se considera el problema de diseñar el esquema de una base de datos relacional.
- En general, el objetivo del diseño de una base de datos relacional es la generación de un conjunto de esquemas de relación que permita almacenar la información sin redundancias innecesarias, pero que también permita recuperarla fácilmente.

### 7.1 Características de los buenos diseños relacionales

- Alternativas de diseño: Esquemas grandes:
- Alternativas de diseño: Esquemas pequeños:

### 7.2 Dominios atómicos y la primera forma normal

El modelo E-R permite que los conjuntos de entidades y los de relaciones tengan atributos con algún tipo de subestructura.

- permite atributos multivalorados y atributos compuestos.
- Cuando se crean tablas a partir de los diseños E-R que contienen ese tipo de atributos, se elimina esa subestructura.
- Para los atributos compuestos, se deja que cada atributo componente sea un atributo de pleno derecho.
- Para los atributos multivalorados, se crea una tupla por cada elemento del conjunto multivalorado.

- Un dominio es atómico si se considera que los elementos de ese dominio son unidades indivisibles.

Los conjuntos de nombres son ejemplos de valores no atómicos. Por ejemplo, si el esquema de la relación empleado incluyera el atributo hijos, los elementos de cuyo dominio son conjuntos de nombres, el esquema no estaría en la primera forma normal.

- Se da por supuesto que los números enteros son atómicos, por lo que el conjunto de los números enteros es un dominio atómico

### 7.3 Descomposición mediante dependencias funcionales

### Claves y dependencias funcionales:

- Las claves y, más en general, las dependencias funcionales son restricciones de la base de datos que exigen que las relaciones cumplan determinadas propiedades. Las relaciones que cumplen todas esas restricciones son legales.
- Las dependencias funcionales se emplean de dos maneras:
  - Para probar las relaciones y ver si son legales de acuerdo con un conjunto dado de dependencias funcionales.
  - Para especificar las restricciones del conjunto de relaciones legales.

### Forma normal de Boyce-Codd ( FNBC ):

-Elimina todas las redundancias que se pueden descubrir a partir de las dependencias funcionales aunque, puede que queden otros tipos de redundancia.

- Los diseños de bases de datos están en la FNBC si cada miembro del conjunto de esquemas de relación que constituye el diseño se halla en la FNBC .

### FNBC y la conservación de las dependencias:

#### Tercera norma formal

#### Formas normales superiores

## 7.4 Teoría de las dependencias funcionales

### Cierre de los conjuntos de dependencias funcionales

No es suficiente considerar el conjunto dado de dependencias funcionales. También hay que considerar todas las dependencias funcionales que se cumplen. Se verá que, dado un conjunto  $F$  de dependencias funcionales, se puede probar que también se cumple alguna otra dependencia funcional. Se dice que  $F$  "implica lógicamente" esas dependencias funcionales.

Se pueden utilizar las tres reglas siguientes para hallar las dependencias funcionales implicadas lógicamente. Aplicando estas reglas repetidamente se puede hallar todo  $F^+$ , dado  $F$ . Este conjunto de reglas se denomina **axiomas de Armstrong** en honor de la persona que las propuso por primera vez.

- Regla de la reflexividad. Si  $\alpha$  es un conjunto de atributos y  $\beta \subseteq \alpha$ , entonces se cumple que  $\alpha \rightarrow \beta$ .
- Regla de la aumentatividad. Si se cumple que  $\alpha \rightarrow \beta$  y  $\gamma$  es un conjunto de atributos, entonces se cumple que  $\gamma\alpha \rightarrow \gamma\beta$ .

- Regla de la transitividad. Si se cumple que  $\alpha \rightarrow \beta$  y que  $\beta \rightarrow \gamma$ , entonces se cumple que  $\alpha \rightarrow \gamma$ .

Aunque los axiomas de Armstrong son completos, resulta pesado utilizarlos directamente para el cálculo de  $F^+$ . Para simplificar más las cosas se dan unas reglas adicionales. Se pueden utilizar los axiomas de Armstrong para probar que son correctas.

- Regla de la unión. Si se cumple que  $\alpha \rightarrow \beta$  y que  $\alpha \rightarrow \gamma$ , entonces se cumple que  $\alpha \rightarrow \beta\gamma$ .
- Regla de la descomposición. Si se cumple que  $\alpha \rightarrow \beta\gamma$ , entonces se cumple que  $\alpha \rightarrow \beta$  y que  $\alpha \rightarrow \gamma$ .
- Regla de la pseudotransitividad. Si se cumple que  $\alpha \rightarrow \beta$  y que  $\gamma\beta \rightarrow \delta$ , entonces se cumple que  $\alpha\gamma \rightarrow \delta$ .

## Cierre de los conjuntos de atributos

### Recubrimiento canonico

El recubrimiento canónico  $F_c$  de  $F$  es un conjunto de dependencias tal que  $F$  implica lógicamente todas las dependencias de  $F_c$  y  $F_c$  implica lógicamente todas las dependencias de  $F$ . Además,  $F_c$  debe tener las propiedades siguientes:

- Ninguna dependencia funcional de  $F_c$  contiene atributos raros.
- El lado izquierdo de cada dependencia funcional de  $F_c$  es único. Es decir, no hay dos dependencias  $\alpha_1 \rightarrow \beta_1$  y  $\alpha_2 \rightarrow \beta_2$  de  $F_c$  tales que  $\alpha_1 = \alpha_2$ .

### Descomposición sin pérdida

Se dice que la descomposición es una descomposición sin pérdidas si, para todos los ejemplares legales de la base de datos (es decir, los ejemplares de la base de datos que satisfacen las dependencias funcionales especificadas y otras restricciones)

Las descomposiciones que no son sin pérdidas se denominan descomposiciones con pérdidas. Los términos descomposición de reunión sin pérdidas y descomposición de reunión con pérdidas se utilizan a veces en lugar de descomposición sin pérdidas y descomposición con pérdidas.

### Conservación de las dependencias

## 7.5 Algoritmos de descomposición

### Descomposición en la FNBC

Se puede emplear la definición de la FNBC para comprobar directamente si una relación se halla en esa forma normal. Sin embargo, el cálculo de F + puede resultar una tarea tediosa. En primer lugar se van a describir pruebas simplificadas para verificar si una relación dada se halla en la FNBC . En caso de que no lo esté, se puede descomponer para crear relaciones que sí estén en la FNBC

### Descomposición en la 3FN

**Comparación de la FNBC y la 3FN** De las dos formas normales para los esquemas de las bases de datos relacionales, la 3FN y la FNBC , la 3FN es más conveniente, ya que se sabe que siempre es posible obtener un diseño en la 3FN sin sacrificar la ausencia de pérdidas ni la conservación de las dependencias.

## 7.6 Descomposición mediante dependencias multivaloradas

### Dependencias multivaloradas

Las dependencias funcionales impiden que ciertas tuplas estén en una relación dada. Si  $A \rightarrow B$ , entonces no puede haber dos tuplas con el mismo valor de A y diferentes valores de B. Las dependencias multivaloradas, por otro lado, no impiden la existencia de esas tuplas.

### Cuarta forma normal

Un esquema de relación R está en la cuarta forma normal ( 4FN ) con respecto a un conjunto D de dependencias funcionales y multivaloradas si, para todas las dependencias multivaloradas de D + de la forma  $\alpha \twoheadrightarrow \beta$ , donde  $\alpha \subseteq R$  y  $\beta \subseteq R$ , se cumple, como mínimo, una de las condiciones siguientes:

- $\alpha \twoheadrightarrow \beta$  es una dependencia multivalorada trivial.
- $\alpha$  es superclave del esquema R.

### Descomposición de la 4FN

## 7.7 Más formas normales

La cuarta forma normal no es, de ningún modo, la forma normal “definitiva”. Como ya se ha visto, las dependencias multivaloradas ayudan a comprender y a abordar algunas formas de repetición de la información que no pueden comprenderse en términos de las dependencias funcionales. Hay restricciones denominadas dependencias de reunión que generalizan las dependencias multivaloradas y llevan a otra forma normal denominada forma normal de reunión por proyección ( FNRP ) (la FNRP se denomina en algunos libros quinta forma normal). Hay una clase de restricciones todavía más generales, que lleva a una forma normal denominada forma normal de dominios y claves ( FNDC ).

## 7.8 Proceso de diseño de las base de datos

### El modelo ER y la normalización

Cuando se definen con cuidado los diagramas E-R , identificando correctamente todas las entidades, los esquemas de relación generados a partir de ellos no deben necesitar mucha más normalización. No obstante, puede haber dependencias funcionales entre los atributos de alguna entidad. Por ejemplo, supóngase que la entidad empleado tiene los atributos número\_departamento y dirección\_departamento, y que hay una dependencia funcional  $\text{número\_departamento} \rightarrow \text{dirección\_departamento}$ . Habrá que normalizar la relación generada a partir de empleado.

- Las dependencias funcionales pueden ayudar a detectar un mal diseño E-R .

### Denominación de los atributos y de las relaciones\*

Una característica deseable del diseño de bases de datos es la asunción de un rol único, lo que significa que cada nombre de atributo tiene un significado único en toda la base de datos. Esto evita que se utilice el mismo atributo para indicar cosas diferentes en esquemas diferentes.

En los esquemas de bases de datos de gran tamaño los conjuntos de relaciones (y los esquemas derivados) se suelen denominar mediante la concatenación de los nombres de los conjuntos de entidades a los que hacen referencia, quizás con guiones o caracteres de subrayado intercalados.

### Desnormalización para el rendimiento

A veces, los diseñadores de bases de datos escogen un esquema que tiene información redundante; es decir, que no está normalizado. Utilizan la redundancia para mejorar el rendimiento de aplicaciones concretas. La penalización sufrida por no emplear un esquema normalizado es el trabajo adicional (en términos de tiempos de codificación y de ejecución) de mantener consistentes los datos redundantes.

### Otros problemas del diseño

Hay algunos aspectos del diseño de bases de datos que la normalización no aborda y, por tanto, pueden llevar a un mal diseño de la base de datos. Los datos relativos al tiempo o a intervalos temporales presentan varios de esos problemas.

## 7.9 Modelo de datos temporales

En general, los datos temporales son datos que tienen asociado un intervalo de tiempo durante el cual son válidos<sup>4</sup>. Se utiliza el término instantánea de los datos para indicar su valor en un momento determinado.

El modelado de datos temporales es una cuestión interesante por varios motivos. Por ejemplo, supóngase que se tiene una entidad cliente con la que se desea asociar una dirección que varía con el tiempo. Para añadir información temporal a una dirección hay que crear un atributo multivalorado, cada uno de cuyos valores es un valor compuesto que contiene una dirección y un intervalo de tiempo. Además de los valores de los atributos que varían con el tiempo, puede que las propias entidades tengan un periodo de validez asociado.