

## **Bases de Datos I**

**14/10/2020**

### **Preguntas Cap 14 14/10**

IS-501

José Inestroza

#### **Capítulo 14 Fundamentals**

- **¿Qué es la información redundante en los SGBD?**

**R/** Es información repetida en la base de datos, dado que es uno de los objetivos implícitos en la actividad de diseño de la base de datos, minimizar la redundancia implica minimizar el almacenamiento redundante de la misma información y reducir la necesidad de múltiples actualizaciones para mantener la coherencia entre múltiples copias de la misma información.

- **¿Qué son las tuplas falsas?**

**R/** Se denominan tuplas falsas porque a las tuplas que representan información falsa que no es válida, para lo cual se deben evitar las relaciones que contengan atributos coincidentes que no sean combinaciones (clave externa, clave primaria) porque la unión de dichos atributos puede producir tuplas falsas.

- **¿Qué es la semántica en las bases de datos?**

**R/** La semántica de una relación se refiere a su significado resultante de la interpretación de los valores de los atributos en una tupla, por lo tanto la semántica en las bases de datos hace referencia al significado de cada una de las tuplas que la componen.

- **¿Qué son las anomalías de actualización?**

**R/** Son las anomalías o problemas que genera el almacenamiento de combinaciones naturales de relaciones base.

- **¿Qué son las anomalías de eliminación?**

**R/** Son los problemas generados al vaciar los registros de una base de datos que al no poder todos los valores de los atributos en NULL por violaciones de integridad a la información se pierden datos inadvertidamente de la base de datos.

- **¿Qué son las anomalías de modificación?**

**R/** Son los problemas que genera actualizar o modificar un registro dado que si tiene relaciones asociadas o dependencias, es necesario también modificar cada uno de los registros que representen una relación o dependencia.

- **¿Por qué la implementación de NULL no es una buena práctica?**

**R/** Porque Si muchos de los atributos no se aplican a todas las tuplas en la

relación, terminamos con muchos NULL en esas tuplas. Esto puede desperdiciar espacio a nivel de almacenamiento y también puede dar lugar a problemas para comprender el significado de los atributos y para especificar operaciones JOIN en el nivel lógico.

Las operaciones SELECT y JOIN implican comparaciones; si hay valores NULL presentes, los resultados pueden volverse impredecibles. Además, los NULL pueden tener múltiples interpretaciones.

- **¿Cuáles son las 4 guías de diseño dictadas en el capítulo y qué problemáticas resuelven?**

**R/ Directriz 1.** Diseñar un esquema de relación para que sea fácil de explicar su significado. No combine atributos de varios tipos de entidades y tipos de relaciones en una sola relación. Intuitivamente, si un esquema de relación corresponde a un tipo de entidad o un tipo de relación, es sencillo explicar su significado. De lo contrario, si la relación corresponde a una mezcla de múltiples entidades y relaciones, se producirán ambigüedades semánticas y la relación no podrá explicarse fácilmente.

**Directriz 2.** Diseñar los esquemas de relación base de modo que no existan anomalías de inserción, eliminación o modificación en las relaciones. Si se presentan anomalías, anótelas claramente y asegúrese de que los programas que actualizan la base de datos funcionarán correctamente. Es importante tener en cuenta que, en ocasiones, es posible que sea necesario infringir estas directrices para mejorar el rendimiento de determinadas consultas.

**Directriz 3.** En la medida de lo posible, evite colocar atributos en una relación base cuyos valores puedan ser NULL con frecuencia. Si los NULL son inevitables, asegúrese de que se apliquen solo en casos excepcionales y no se apliquen a la mayoría de las tuplas en la relación. Usar el espacio de manera eficiente y evitar uniones con valores NULL son los dos principales criterios que determinan si incluir las columnas que pueden tener NULL en una relación o tener una relación separada para esas columnas (con las columnas clave apropiadas).

**Directriz 4.** Diseñar esquemas de relaciones para que puedan unirse con condiciones de igualdad en atributos que estén apropiadamente relacionados (clave primaria, clave externa) pares de manera que garantice que no se generen tuplas falsas. Evite las relaciones que contengan atributos coincidentes que no sean combinaciones (clave externa, clave primaria) porque la unión de dichos atributos puede producir tuplas falsas.

**Problemas que resuelve:** Resuelve la ambigüedad semántica de las relaciones, permitiendo que se expliquen fácilmente; también resuelve las principales anomalías de inserción, eliminación o modificación en las relaciones teniendo en cuenta que se puede infringir en estas anomalías por únicamente conseguir mejor rendimiento; también resuelve la problemática de los atributos múltiples con valor NULL y por último también resuelve la problemática de crear tuplas falsas al evitar las relaciones que contengan atributos coincidentes que no sean combinaciones.

- **¿Qué es una dependencia funcional?**

**R/** Es una restricción entre dos conjuntos de atributos de la base de datos. Una dependencia funcional  $X \rightarrow Y$  es una dependencia funcional completa si la eliminación de cualquier atributo A de X significa que la dependencia ya no se mantiene; es decir, para cualquier atributo  $A \in X$ ,  $(X - \{A\})$  no determina funcionalmente Y. **Una dependencia funcional** es una propiedad del esquema de relación R, no de un estado de relación legal particular r de R. Por lo tanto, una FD no puede inferirse automáticamente de una extensión de relación dada r, sino que debe ser definida explícitamente por alguien que conozca la semántica de la relación.

- **¿Qué son las normas formales sobre las llaves primarias?**

**R/** Son en las que se da un conjunto de dependencias funcionales para cada relación, y que cada relación tiene una clave primaria designada. Esta información combinada con las pruebas (condiciones) para las formas normales impulsa el proceso de normalización para el diseño de esquemas relacionales.

- **¿Qué es la normalización de datos?**

**R/** Puede considerarse un proceso de análisis de los esquemas de relación dados en función de sus FD y claves primarias para lograr las propiedades deseables de (1) minimizar la redundancia y (2) minimizar las anomalías de inserción, eliminación y actualización, el procedimiento de normalización proporciona a los diseñadores de bases de datos lo siguiente:

- Un marco formal para analizar esquemas de relaciones basados en sus claves y en las dependencias funcionales entre sus atributos.
- Una serie de pruebas de forma normal que se pueden realizar en esquemas de relaciones individuales para que la base de datos relacional se pueda normalizar al grado deseado.

- **¿En qué consiste la primera forma normal?**

**R/** , se definió para no permitir atributos multivalor, atributos compuestos y sus combinaciones. Establece que el dominio de un atributo debe incluir solo valores atómicos (simples, indivisibles) y que el valor de cualquier atributo en una tupla debe ser un valor único del dominio de ese atributo. Por lo tanto, 1NF no permite tener un conjunto de valores, una tupla de valores o una combinación de ambos como valor de atributo para una sola tupla. En otras palabras, 1NF no permite relaciones dentro de relaciones o relaciones como valores de atributo dentro de tuplas. Los únicos valores de atributo permitidos por 1NF son valores atómicos únicos (o indivisibles). La primera forma normal también rechaza atributos multivalor que son compuestos en sí mismos.

- **¿En qué consiste la segunda forma normal?**

**R/** La segunda forma normal (2NF) se basa en el concepto de dependencia funcional completa. Una dependencia funcional  $X \rightarrow Y$  es una dependencia

funcional completa si la eliminación de cualquier atributo A de X significa que la dependencia ya no se mantiene; es decir, para cualquier atributo  $A \in X$ ,  $(X - \{A\})$  no determina funcionalmente Y. Una dependencia funcional  $X \rightarrow Y$  es una dependencia parcial si algún atributo  $A \in X$  puede eliminarse de X y la dependencia aún se mantiene; es decir, para algunos  $A \in X$ ,  $(X - \{A\}) \rightarrow Y$ .

Un esquema de relación R está en 2NF si cada atributo no principal A en R es completamente funcionalmente dependiente de la clave primaria de R.

- **¿En qué consiste la tercera forma normal?**

**R/** La tercera forma normal (3NF) se basa en el concepto de dependencia transitiva. Una dependencia funcional  $X \rightarrow Y$  en un esquema de relación R es una dependencia transitiva si existe un conjunto de atributos Z en R que no es una clave candidata ni un subconjunto de cualquier clave de R, 11 y tanto  $X \rightarrow Z$  como  $Z \rightarrow Y$  sostener.

De acuerdo con la definición original de Codd, un esquema de relación R está en 3NF si satisface 2NF y ningún atributo no primario de R depende transitivamente de la clave primaria.

Forma Normal	Prueba	Normalización
<b>Primera forma normal (1FN)</b>	La relación no debe tener atributos de valores múltiples ni relaciones anidadas.	Forme nuevas relaciones para cada atributo multivalor o relación anidada.
<b>Segunda forma normal (2FN)</b>	Para las relaciones en las que la clave principal contiene varios atributos, ningún atributo que no sea de clave debe ser funcionalmente dependiente de una parte de la clave principal.	Descomponga y configure una nueva relación para cada clave parcial con sus atributos dependientes. Asegúrese de mantener una relación con la clave primaria original y cualquier atributo que sea funcionalmente dependiente de ella.
<b>Tercera forma normal (3FN)</b>	La relación no debe tener un atributo no clave	Descomponer y establecer una relación que incluya los

	determinado funcionalmente por otro atributo no clave (o por un conjunto de atributos no clave). Es decir, no debería haber dependencia transitiva de un atributo que no sea de clave en la clave principal.	atributos no clave que determinan funcionalmente otros atributos no clave.
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------

- **¿Qué remedia/corriges la normalización?**  
**R/** Corrige la redundancia de datos, las tuplas falsas y los atributos con valores en NULL que dañan la integridad de la información.
- **¿Los ejercicios hechos en clase síncrona consideran tablas de bases normalizadas? ¿Por qué de su respuesta?**  
**R/** No, porque tenemos múltiples relaciones en una sola tabla y porque los datos presentan dependencias funcionales y transitivas entre relaciones.