

NORMALIZACION

Peña Reyes Jocelyn Karina

Rivera Lima Carlos Eduardo

¿Qué es?

La normalización es una técnica para diseñar la estructura lógica de los datos de un sistema de información en el modelo relacional, desarrollada por E. F. Codd en 1972. Es una estrategia de diseño de abajo a arriba: se parte de los atributos y éstos se van agrupando en relaciones (tablas). Las ventajas de la normalización son las siguientes:

1. Evita anomalías en inserciones, modificaciones y borrados.
2. Mejora la independencia de datos.
3. No establece restricciones artificiales en la estructura de los datos.

Uno de los conceptos fundamentales en la normalización es el de dependencia funcional. Una dependencia funcional es una relación entre atributos de una misma relación (tabla). Cada dependencia funcional es una clase especial de regla de integridad y representa una relación de uno a muchos. En el proceso de normalización se debe ir comprobando que cada relación (tabla) cumpla una serie de reglas que se basan en la clave primaria y las dependencias funcionales. Cada regla que se cumple aumenta el grado de normalización. Si una regla no se cumple, la relación se debe descomponer en varias relaciones que sí la cumplan. Por lo tanto la normalización es el proceso mediante el cual se transforman datos complejos a un conjunto de estructuras de datos más pequeñas, que además de ser más simples y más estables, son más fáciles de mantener. Otra ventaja de la normalización de base de datos es el consumo de espacio. Una base de datos normalizada ocupa menos espacio en disco que una no normalizada. Hay menos repetición de datos, lo que tiene como consecuencia un mucho menor uso de espacio en disco.

Grados de normalización

Existen básicamente tres niveles de normalización: Primera Forma Normal (1NF), Segunda Forma Normal (2NF) y Tercera Forma Normal (3NF). Cada una de estas formas tiene sus propias reglas. Cuando una base de datos se conforma a un nivel, se considera normalizada a esa forma de normalización.

Primera Forma Normal

Establece que las columnas repetidas deben eliminarse y colocarse en tablas separadas. Una tabla está en 1FN cuando todos los atributos de clave están definidos y cuando todos los restantes dependen de la clave primaria. Sin embargo, una tabla en 1FN aún puede contener tanto dependencias parciales como transitivas (una dependencia parcial es aquella en la que un atributo es funcionalmente dependiente de una parte de una clave primaria de atributos múltiples. Una dependencia transitiva es aquella en la que un atributo es funcionalmente dependiente de otro atributo no de clave). Naturalmente una tabla con una clave primaria de un solo atributo no puede exhibir dependencias parciales.

Segunda Forma Normal

La regla de la Segunda Forma Normal establece que todas las dependencias parciales se deben eliminar y separar dentro de sus propias tablas. Una dependencia parcial es un término que describe a aquellos datos que no dependen de la llave primaria de la tabla para identificarlos. Podemos insertar un registro sin un exceso de datos en la mayoría de las tablas. Una tabla se encuentra en 2FN cuando está en 1FN y no contiene dependencias parciales. Por consiguiente, una tabla 1FN automáticamente está en 2FN si su clave primaria está basada solamente en un atributo simple. Una tabla en 2FN aún puede contener dependencias transitivas.

Tercera Forma Normal

Una tabla está normalizada en esta forma si todas las columnas que no son llave son funcionalmente dependientes por completo de la llave primaria y no hay dependencias transitivas. Cuando las tablas están en la Tercera Forma Normal se previenen errores de lógica cuando se insertan o borran registros. Cada columna en una tabla está identificada de manera única por la llave primaria, y no deben haber datos repetidos.

Referencias

1. <https://programas.cuaed.unam.mx>