

BASES DE DATOS II

¿Qué es la Normalización? ¿Qué Propósito Sirve?

James Martin lo pone esta manera:

"Normalización es un enfoque formal que examina datos, y grupos de datos de una manera que mejor pueda acomodar futuros cambios en el negocio y de minimizar el impacto de estos cambios en sistemas de aplicación."

"Normalización" se refiere al proceso de crear una estructura para guardar información "relacional" que sea eficaz, fiable, flexible, y apropiada. Datos Normalizados deben pertenecer a una estructura de datos "relacional".

Una estructura de datos relacional conecta múltiples, juegos definidos de información para que puedan ser manipulados por un solo programa y unidos si es necesario. Un banco de datos relacional comprende un programa y juegos de datos, llamados tablas, construidos por un RDBMS.

Para producir un diseño de una base de datos se tiene que comenzar con tablas sin normalizar y a través de una serie de pasos descomponer la estructura de datos hasta desarrollar el esquema final de la base de datos.

El proceso de normalización

Pasos para Normalizar los Datos

1. Listar los datos
2. Identificar por lo menos una llave por cada tabla. Cada tabla debe tener una llave primaria.
3. Identificar las llaves para las relaciones. Las llaves de una relación son las llaves de las dos tablas que junta.
4. Buscar datos calculados en su lista de datos de apoyo. Datos Calculados normalmente no se guardan en la base de datos.
5. Poner los datos en 1FN:
 - a. Remover los datos repetidos en las tablas y relaciones.
 - b. Crear una o más tablas y relaciones con los datos que quita.
6. Poner los datos en 2FN:
 - a. Identificar las tablas y relaciones con más de una llave.
 - b. Remover los datos que dependen solamente de una parte de la llave.
 - c. Crear una o más tablas y relaciones con los datos que quita.
7. Poner los datos en 3FN:
 - a. Remover datos que dependen de otros datos en la tabla o relación y no en la llave.
 - b. Crear una o más tablas y relaciones con los datos que quita.

Formas Normales

Las Formas normales son las pruebas que usualmente se usan para normalizar datos. Cuando los datos pasan la primera prueba, se considera estar en la primera forma normal (1FN), cuando pasan la segunda prueba, se está en la segunda forma normal (2FN) y así sucesivamente.

Normalizar

Inicialmente, los datos están sin normalizar y se dice que están en la forma normal cero. Los analistas de datos realizan el proceso de normalización para diseñar la base de datos en la mejor manera posible para la organización.

Esto incluiría:

- Relaciones dentro de la base de datos que son fáciles de entender y simples mantener.
- Nuevas demandas pueden ser acomodadas fácilmente.
- Se previene el almacenamiento de datos inválidos.
- Agrupaciones apropiadas de datos en columnas, dentro de las tablas relacionales
- Representación eficiente de los datos.

El proceso de normalización es un intento de superar varios problemas asociados con la manera en que se pueden almacenar los datos y desde el punto de vista de la implementación, planes mejores llevan a:

- Eliminación de la duplicación innecesaria de datos
- La creación de relaciones donde cada atributo depende funcionalmente de la llave de la relación.
- Simplificación de la entrada en vigor de restricciones de integridad.

Esencialmente, lo que ocurre con la normalización es que uno empieza con una visión de usuario compleja y reduce esta a un conjunto de estructuras de datos pequeñas y estables. Estas estructuras de datos han demostrado ser más flexibles, estables y simples de mantener, que datos de una estructura sin normalizar.

Anomalías de la Modificación

La "Anomalía de la tachadura" ocurre cuando uno deliberadamente anula una pieza de datos y en consecuencia accidentalmente pierde otros datos.

Otra anomalía que puede ocurrir es la "Anomalía de la Inserción" que se da porque una llave primaria no puede tener un valor nulo, en otras palabras cuando insertamos un nuevo valor en la llave tiene que haber por obligación datos en las llaves extranjeras.

La tercera forma de anomalía es la "Anomalía de actualización", esta anomalía concierne la presencia de redundancia innecesaria en nuestra base de datos. Cuando se cambian datos al registro maestro nos debemos de asegurar que todas las ocurrencias de este dato son actualizadas si no, estamos permitiendo inconsistencias en nuestra base de datos.

Consecuencias Adicionales de la Normalización

Mientras la normalización resuelve varios problemas relacionados con la estructuración de datos, también crea dos problemas nuevos:

1. Duplicación de datos
2. Ineficacia de recuperación

Duplicación de Datos

El proceso de normalización envuelve la descomposición de una estructura de datos en estructuras más pequeñas. El proceso de descomposición requiere que la llave primaria de la relación original se incluya, como una llave extranjera, en la relación (es) que se formó. Esto significa que a medida que se van creando estas llaves se va incrementando las probabilidades de poner en peligro la integridad referencial.

El trabajo extra de mantener restricciones de integridad referencial se debe balancear, por el analista de datos, contra la facilidad con que los usuarios puedan interactuar con el sistema.

Ineficacia de la Recuperación

Un efecto del adicional del número creciente de estructuras de datos en la base de datos, es que se afecta la eficacia de recuperación del sistema de la computadora. Este puede ser particularmente notable con sistemas basados en microordenadores. Aquí tenemos otro compromiso: la normalización por descomposición reducirá el espacio general de requerido para almacenar datos, pero incrementa el tiempo de recuperación de la información ya que relaciones numerosas necesitarían ser reunidas para extraer la información.

Bibliografía

Sitios de búsqueda

- **pcwebopaedia:** <http://www.sandybay.com>
- **altavista:** <http://www.altavista.digital.com>

Documentación

Reid Software Development: <http://www.accessdatabase.com>

Documentos: <http://www.accessdatabase.com/normalize.html>

Swinburne University of Technology: <http://swin.edu.au>

Documentos: <http://swin.edu.au/is/subjects/bt220/bt220s1.html>

sybase: <http://www.sybase.com>