

Práctica 1: Principios de las Bases de Datos

Dr. Héctor Selley

16 de agosto de 2017

Lea el Capítulo 1 (página 4) del libro “[Coronel, C.; Morris, S.; Rob, P.; Database Systems: Design, Implementation and Management](#)”, resuelva el cuestionario y los problemas siguientes.

1. CUESTIONARIO

1. Defina cada uno de los siguientes términos
 - a)* **Dato:** Los datos son hechos en *crudo*, esto significa que estos hechos aún no han sido procesados para revelar su significado. Los datos tienen poco significado hasta que se hayan organizado de alguna forma lógica.
 - b)* **Campo:** Un caracter o un grupo de caracteres que tienen un significado específico. Un campo se utiliza para definir y almacenar datos.
 - c)* **Registro:** Un conjunto conectado lógicamente a uno o más campos que describen una persona, lugar o cosa. Por ejemplo, un campo que constituye un registro de un cliente consiste de: nombre, dirección, número telefónico, fecha de nacimiento, etc.
 - d)* **Archivo:** Un conjunto de registros que tienen alguna relación entre si. Por ejemplo, un archivo puede contener datos acerca de los estudiantes inscritos a una Universidad.
2. ¿Qué es redundancia y qué características de un sistema pueden ocasionarla?. La redundancia de datos ocurre cuando un dato se encuentra almacenado innecesariamente en más de un lugar a la vez. La redundancia de datos puede ocasionar diversos problemas, tales como: Baja seguridad en los datos, inconsistencia de datos o Anomalías en los datos.

La redundancia ocurre debido a una estructura organizacional que lo fomenta. Para solucionarlo debe modificarse la estructura.

3. ¿Qué es un DBMS (SGBD en español) y cuáles son sus funciones?

Un Sistema de Gestión de Bases de Datos (SGBD) es un conjunto de programas que administra la estructura de la base de datos y controla el acceso a los datos almacenados en la base.

El SGBD sirve como un intermediario entre el usuario y la base de datos. La estructura de la base de datos se almacena como un conjunto de archivos y la única forma de tener acceso a esos archivos es a través del SGBD. El SGBD recibe todas las solicitudes y las traduce en las operaciones requeridas para completar esas solicitudes.

4. ¿Cuáles son las ventajas y desventajas de utilizar un DBMS?

Ventajas:

- **Compartición de datos mejorada.** Los SGBD ayudan a crear un ambiente en el cual los usuarios tienen mejor acceso a datos mejor administrados.
- **Seguridad de los datos mejorada.** Mientras más usuarios tienen acceso a los datos, mayor será el riesgo de que haya una brecha de seguridad. Un SGBD proporciona una infraestructura para una mejor privacidad de los datos y mejores políticas de seguridad.
- **Mejor integración de los datos.** Amplio acceso a datos mejor administrados permite una vista clara e integrada de las operaciones en una escala mayor.
- **Inconsistencia de datos minimizada.** La inconsistencia de datos ocurre cuando diferentes versiones de un mismo dato se encuentran en lugares distintos. La probabilidad de una inconsistencia de datos se reduce dramáticamente cuando se utiliza una base de datos diseñada adecuadamente.
- **Acceso mejorado a los datos.** Los SGBD producen respuestas rápidas a consultas específicas. Esta es una característica implícita en el diseño e implementación de los SGBD.
- **Toma de decisiones mejorada.** Datos mejor administrados y acceso mejorado a los datos, hacen posible generar información de mejor calidad, necesaria para tomar una mejor decisión.
- **Productividad del usuario mejorada.** La disponibilidad de los datos, combinado con las herramientas que transforman los datos en información útil, facilitan al usuario el tomar decisiones rápidamente.

Desventajas:

- **Complejidad.** Los SGBD son conjuntos de programas que pueden llegar a ser complejos con una gran funcionalidad. Es preciso comprender muy bien esta funcionalidad para poder realizar un buen uso de ellos.

- **Coste del equipamiento adicional.** Tanto el SGBD, como la propia base de datos, pueden hacer que sea necesario adquirir más espacio de almacenamiento. Además, para alcanzar las prestaciones deseadas, es posible que sea necesario adquirir una máquina más grande o una máquina que se dedique solamente al SGBD. Todo esto hará que la implantación de un sistema de bases de datos sea más cara.
 - **Vulnerable a los fallos.** El hecho de que todo esté centralizado en el SGBD hace que el sistema sea más vulnerable ante los fallos que puedan producirse. Es por ello que deben tenerse copias de seguridad (Backup).
5. Explique la diferencia entre datos e información.
 Los **datos** son hechos en *crudo*, esto significa que estos hechos aún no han sido procesados para revelar su significado. Por otro lado, la **información** es el resultado de procesar los datos crudos, de manera que se revela su significado.
 Los datos sólo necesitan ser obtenidos mediante algún mecanismo, por otro lado, la información debe ser formateada, procesada y presentada.
 Los datos son la base de la información que, a su vez, es el cimiento del conocimiento.
6. Explique que son los metadatos.
 Los Metadatos son los datos acerca de los datos. A través de los metadatos los datos son administrados y almacenados.
 Los metadatos proporcionan una descripción de las características de los datos y un conjunto de relaciones que enlazan los datos encontrados en la base de datos. Por ejemplo, el metadato almacena información como el nombre de cada elemento, tipo de dato (numérico, texto, etc.) para cada elemento, puede ser nulo o no, etc. Dadas las características de los metadatos, las bases de datos suelen ser conocidas como "conjunto de datos auto-descriptivos".
7. Liste y describa los diferentes tipos de bases de datos (no DBMS).
 Un SGBD puede manejar muchos tipos diferentes de bases de datos. Las bases de datos pueden clasificarse por número de usuarios, ubicación de la base de datos y el tipo de uso.
 El número de usuarios determina si una base de datos es mono-usuario o multi-usuario. Una base de datos es **mono-usuario** si sólo permite que un usuario a la vez la utilice. Por otro lado, una base de datos es **multi-usuario** si permite que varios usuarios la utilicen al mismo tiempo. Una base de datos multi-usuario que permite un número máximo de 50 usuarios simultáneos, se le conoce como **base de datos de un grupo de trabajo**. Si la base multi-usuario permite más de 50 usuarios simultáneos, se le conoce como **base de datos empresarial**.
 La ubicación también se utiliza para clasificar las bases de datos. Si una base de datos de encuentra en un sólo lugar, se le conoce como **base de datos centralizada**. Si la base soporta que los datos se encuentren distribuidos en varios sitios, se le conoce como **base de datos distribuida**.

Una base de datos que está diseñada para almacenar las operaciones diarias de una organización, se le conoce como **base de datos transaccional** (también conocida como operacional o de producción). Por otro lado, una **base banco de datos** está enfocada en el almacenamiento de datos utilizados para generar información requerida para tomar decisiones tácticas o estratégicas.

Numero de usuarios			Ubicación		Tipo de Uso	
Mono-Usuario	Multi-Usuario		Centralizada	Distribuida	Transaccional	Banco de Datos
	Grupo de Trabajo	Empresarial				

Tabla 1.1: Tipos de Bases de Datos

8. ¿Cuáles son los principales componentes de un sistema de base de datos?

- **Hardware.** Se refiere a todos los dispositivos físicos del sistema. Por ejemplo: computadoras, discos duros, impresoras, dispositivos de red, etc.
- **Software.**
 - **Sistema Operativo.** Administra los componentes de hardware y hace posible la ejecución de programas en una computadora. Ejemplos: Windows, macOS, Linux, UNIX, BSD, OS2, AIX, Solaris, etc.
 - **SGBD.** Sistema de software que administra la base de datos. Ejemplos: Microsoft SQL Server, Oracle MySQL (antes Sun SQL), MariaDB (Open Source), DB2 de IBM, PostgreSQL (Open Source), etc.
 - **Aplicaciones y Utilidades.** Programas que permiten el acceso y la manipulación de datos del SGBD. Sus funciones pueden ser desde consulta de los datos para generar reportes o hasta para la administración y control del SGBD.
- **Personas.** Todos los usuarios de la base de datos. Pueden ser:
 - **Administradores del sistema.** Supervisan el funcionamiento del sistema.
 - **Diseñadores de la base de datos.** Se les conoce como arquitectos de la base de datos. Una base de datos bien diseñada permitirá tener un desempeño óptimo de la misma.
 - **Programadores y analistas de sistema.** Diseñan y crean las aplicaciones que interactúan con el sistema de base de datos.
 - **Usuario final.** Conjunto de personas que utilizan las aplicaciones para tener acceso a la base de datos. Utilizan la información para su trabajo, no tienen necesariamente conocimiento de la estructura de la base ni conocimientos técnicos avanzados.
- **Procedimientos.** Instrucciones y reglas que rigen el diseño y uso de un sistema de base de datos.

- Datos. El término dato se refiere al conjunto de todos los hechos almacenados en la base de datos.
9. ¿Cuáles son algunas funciones básicas que las bases de datos pueden realizar pero las hojas de cálculo (Excel) no pueden?
- Auto-Documentación a través de los metadatos
 - Aplicación de tipos de datos o dominios para asegurar la consistencia de datos en una columna
 - Relación entre tablas
 - Restricciones para asegurar la consistencia de los datos entre tablas relacionadas

2. PROBLEMAS

1. Dada la estructura presentada en la tabla 2.1 responda las preguntas

PROJ_CODE	PROJ_MANAGER	PROJ_PHONE	PROJ_ADDRESS	PROJ_DB_PRICE
21-5Z	Holly B. Parker	904-338-3416	3334 Lee Rd., Gainesville, FL 37123	16633460.00
25-2D	Jane D. Grant	615-098-9909	218 Clark Blvd., Nashville, TN 36362	12500000.00
25-5A	George F. Dorts	615-227-1245	124 River Dr., Franklin, TN 29185	32512420.00
25-3T	Holly B. Parker	904-338-3416	3334 Lee Rd., Gainesville, FL 37123	21563234.00
27-4Q	George F. Dorts	615-227-1245	124 River Dr., Franklin, TN 29185	10314545.00
29-2D	Holly B. Parker	904-338-3416	3334 Lee Rd., Gainesville, FL 37123	25559990.00
31-7P	William K. Moor	904-445-2719	216 Morton Rd., Stetson, FL 30155	56850000.00

Tabla 2.1: Datos de los proyectos para el Ejercicio 1

- a) ¿Cuántos registros contiene la estructura?
7
- b) ¿Cuántos campos existen por registro?
5
- c) ¿Qué problema encontraría si usted deseara ordenar la estructura por ciudad?, ¿Qué cambios debe hacer a la estructura para resolver ese problema?
Sería difícil hacer una búsqueda de información en la dirección dado que todos esos datos se encuentran en un solo campo.
Se puede resolver agregando un campo para cada componente de la dirección.
- d) Si deseara producir una lista de la estructura ordenado por apellido, clave lada, ciudad, estado o código postal, ¿qué cambios debería hacer a la estructura?
Se debería hacer un cambio estructural de manera que los datos contenidos en los campos sean homogéneos.

e) ¿Qué redundancia de datos observa en la estructura?

Hay datos que se repiten innecesariamente en la tabla, tal como Teléfono y Dirección.

2. Dada la estructura presentada en la tabla 2.2 responda las preguntas

PROJ_NUM	PROJ_NAME	EMP_NUM	EMP_NAME	JOB_CODE	JOB_CHG_HOUR	PROJ_HOURS	EMP_PHONE
1	Hurricane	101	John D. Newson	EE	85.00	13.3	653-234-3245
1	Hurricane	105	David F. Schwann	CT	60.00	16.2	653-234-1123
1	Hurricane	110	Anne R. Ramoras	CT	60.00	14.3	615-233-5568
2	Coast	101	John D. Newson	EE	85.00	19.0	653-234-3254
2	Coast	108	June H. Satbemeir	EE	85.00	17.5	905-554-7812
3	Satelite	110	Anne R. Ramoras	CT	62.00	11.6	615-233-5560
3	Satelite	105	David F. Schwann	CT	26.00	23.4	653-234-5432
3	Satelite	123	Mary D. Chen	EE	85.00	19.1	615-233-5432
3	Satelite	112	Allecia R. Smith	EE	85.00	20.7	615-678-6879

Tabla 2.2: Datos de los proyectos para el Ejercicio 2

a) Identifique y describa los problemas de redundancia.

Datos como nombre del proyecto, nombre del empleado y teléfono se repiten innecesariamente.

b) De acuerdo a los contenidos de EMP_NAME y EMP_PHONE, ¿qué cambios recomendaría hacer?

Crear una tabla independiente que contenga información detallada de los empleados.