



Bases de Datos

Tema 2: Modelos de datos

Dpto. de Ingeniería Informática

Contenidos

- Introducción
- Funciones de un SGBD
- Componentes del entorno de un SGBD
- Arquitectura de una BD
- Modelos de Datos
- Aspectos de un MD
- Referencias

Introducción

- SGBD: Conjunto de programas que permiten a los usuarios crear y mantener una(s) BD
 - Hace de interfaz entre el usuario y la BD
- Facilita las tareas de:
 - Definición: los tipos de datos, y las restricciones existentes entre ellos
 - Construcción: proceso de introducción de los datos en la BD previamente definida
 - Manipulación: insertar, actualizar, eliminar y consultar (generación de informes) datos - CRUD

Sistemas Gestores de BD

- Todo SGBD se construye en base a un modelo de datos
- Un modelo de datos (MD) es un conjunto de conceptos y reglas que permiten estructurar los datos resultantes del estudio de las necesidades de información de un sistemas informático
 - Ese estudio lo realiza un humano que decide qué (y cómo) se incluye en la BD y qué se descarta
 - El MD se basa en estructuras de datos y operadores que soporta el SGBD

Arquitectura de una BD

- Concepto de *independencia con respecto a los datos*: capacidad de modificar un esquema de la BD sin necesidad de modificar el esquema del nivel inmediato superior
 - Independencia *lógica*: capacidad de modificar el esquema conceptual sin alterar los esquemas externos ni los programas de aplicación
 - Independencia *física*: capacidad de modificar el esquema interno sin tener que alterar el esquema conceptual o los externos

Sistemas Gestores de BD

- Objetivos de un SGBD:
 - Independencia física: capacidad de modificar las estructuras de almacenamiento sin afectar las definiciones lógicas de los datos
 - Independencia lógica: capacidad de modificar las estructuras lógicas de los datos sin modificar las aplicaciones de usuario (mediante *vistas*)
 - No redundancia de los datos (*no deseada*)
 - Se comparten los datos por aplicaciones y/o usuarios
 - Manipulación: mediante el uso de lenguajes procedimentales o no procedimentales

Sistemas Gestores de BD

- Objetivos de un SGBD:
 - Eficacia en acceso a los datos: acceder a los datos de la manera más adecuada para cada usuario
 - Administración (centralizada) de los datos
 - Coherencia de los datos: mediante reglas de integridad
 - Compartición de los datos entre usuarios y aplicaciones (con acceso concurrente)
 - Seguridad de los datos: protección de los datos frente a accesos no autorizados

Funciones de un SGBD

- El SGBD debe satisfacer a todos los usuarios y a todas las aplicaciones que acceden a la BD
- Las operaciones típicas de un SGBD son:
 - Afectan a la totalidad de los datos de una BD:
 - Creación
 - Reestructuración
 - Consulta a la totalidad
 - Afectan a ciertos datos:
 - Actualización: inserción, borrado y modificación
 - Consulta selectiva
- Se organizan en tres funciones:
 - Descripción, Manipulación y Control

Funciones de un SGBD

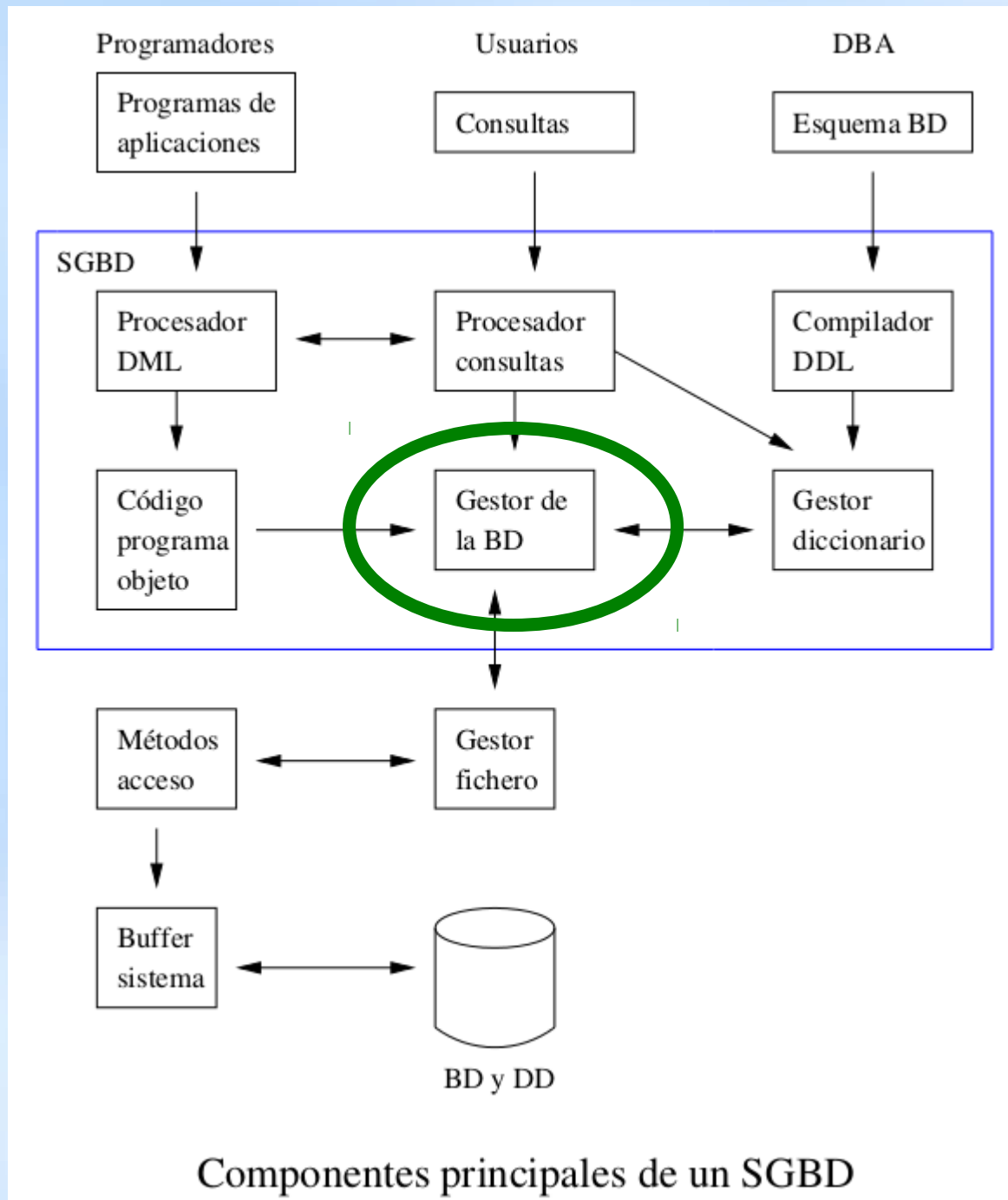
- Función de Descripción o Definición:
 - Debe permitir al diseñador especificar:
 - Los elementos de datos de la BD
 - Estructura de los datos
 - Las relaciones que existen entre los datos
 - Las reglas de integridad semántica
 - Las características de tipo físico
 - Las vistas lógicas de los usuarios
 - Se realiza con un Lenguaje de Definición de Datos (DDL) del SGBD que permite definir los tres niveles de la arquitectura y se almacena en el Dicc. de Datos
 - El SGBD se ocupa de la correspondencia entre estos niveles

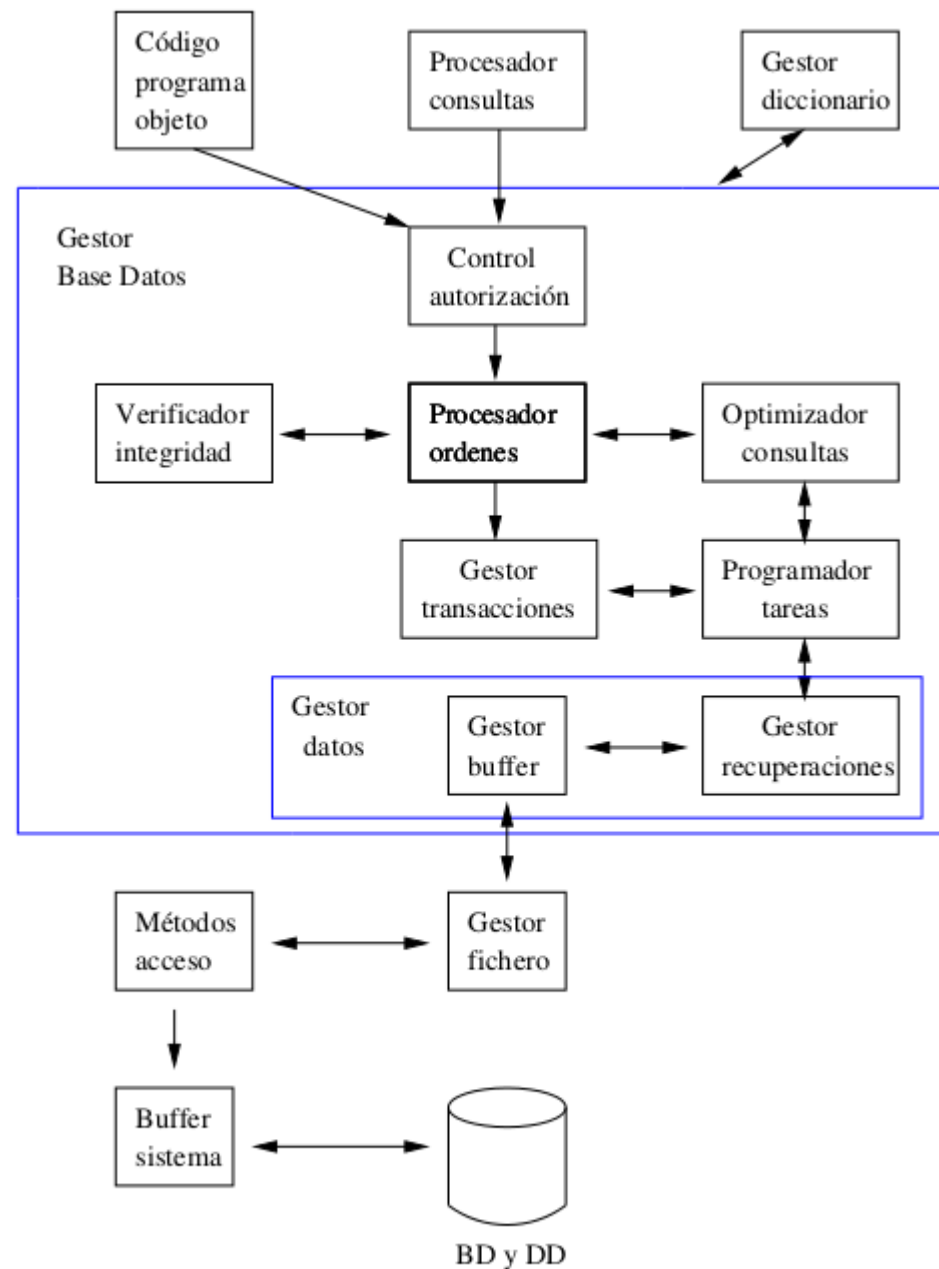
Funciones de un SGBD

- Función de Manipulación:
 - La consulta a la BD puede ser de dos tipos:
 - Totalidad de los datos: se recuperan todos los datos o los de un determinado tipo
 - Consulta selectiva: se recuperan los datos que cumplan un criterio de selección
 - Actualización: inserción, borrado o modificación
 - Esta función se lleva a cabo con el Lenguaje de Manipulación de datos (DML):
 - Lenguaje incorporado en el SGBD: SQL
 - Lenguajes huéspedes sobre un lenguaje anfitrión: SQL embebido en C, Pascal, PHP, ...

Funciones de un SGBD

- Función de control:
 - Reúne todas las interfaces que necesitan los usuarios para comunicarse con la BD
 - Mediante Lenguaje de Control de Datos (DCL)
 - Procedimientos para la administración de la BD, tales como:
 - Servicio:
 - Creación y mantenimiento de ficheros
 - Obtención de estadísticas de utilización de la BD
 - Gestión de usuarios
 - Seguridad física:
 - Copias de seguridad
 - Arranque y parada de la BD en casos de fallos
 - Protección contra accesos no autorizados





Componentes de un gestor de BD

Componentes del entorno de un SGBD

- Clasificación:
 - Máquina:
 - Software
 - Hardware
 - Humana:
 - Procedimientos
 - Personas
 - Datos: lazo de unión entre máquina y personas
- Profundizamos en cada uno

Componentes del entorno de un SGBD

- Máquina:
 - Software: herramientas de 4ª generación (4GL) específicas de dominio
 - Con el lenguaje SQL podemos manipular datos relacionales muy fácilmente
 - Pero no es sencillo hacer un programa para calcular fibonacci
 - Lenguaje NO procedimental (demasiado) fácil de entender
 - Se basa en álgebra y cálculo relacional (potente)
 - Además suelen incorporar: generadores de gráficas, asistentes para informes, formularios para aplicaciones ...
 - Hardware: los SGBD tienden a ser independientes del hardware y de los SO, aunque los fabricantes suelen tener sus SSOO favoritos/certificados

Componentes del entorno de un SGBD

- Humana:
 - Procedimientos: para realizar copias de seguridad, herramientas del SGBD para ayuda a los usuarios, etc.
 - Tienen que estar documentados
 - Personas:
 - Administrador de Datos (DA)
 - Administrador de la BD (DBA)
 - Diseñadores de la BD
 - Programadores de aplicaciones
 - Usuarios finales
 - En organizaciones pequeña los roles se agrupan (hasta llegar al *informático orquesta*)

Componentes del entorno de un SGBD

- Datos:
 - Diccionario de Datos (DD): es un repositorio de metadatos. Almacena información sobre los datos de nuestra BD: tipos de datos, relaciones entre ellos, etc
 - Ficheros de datos: contienen los datos en sí

Clasificación de los SGBD

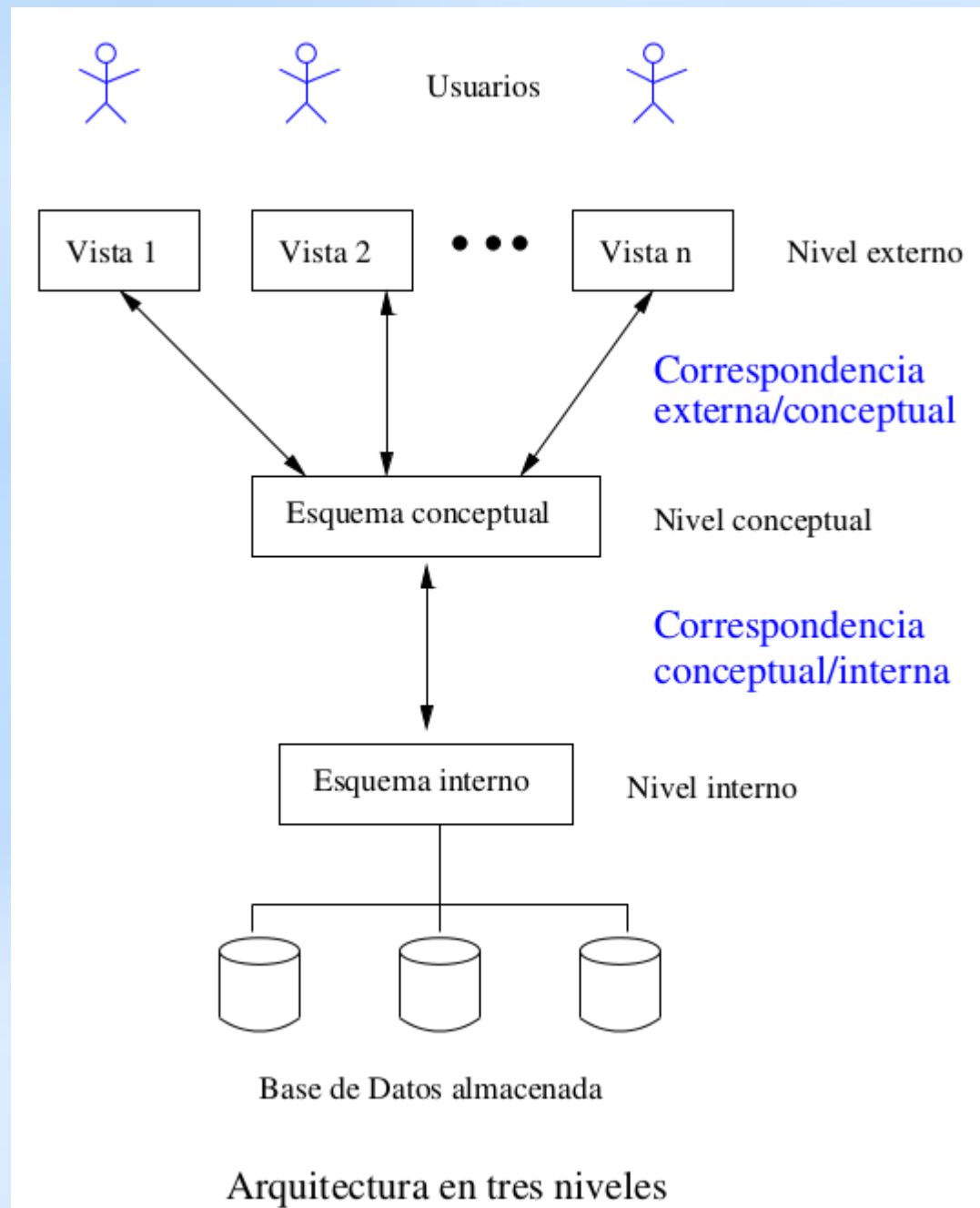
- Existen diversas clasificaciones según:
 - El MD en que está basado:
 - Relacional
 - Orientado a objetos
 - En red
 - ...
 - Número de usuarios: monousuario o multiusuario
 - Según el almacenamiento de los datos: centralizado o distribuido
 - Propósito: general o específico (geográficos, ...)

Arquitectura de una BD

- Las BD tienen las siguientes características:
 - Separación entre los programas y los datos: independencia (hay datos “hardcodeados” ni dependiente de ficheros concretos)
 - Manejo de múltiples vistas de para múltiples usuarios
 - Empleo de un catálogo para el almacenamiento de la descripción de la BD
- Los SGBD respetan la arquitectura propuesta por ANSI/PARC y distinguen tres niveles:
 - Externo, global e interno
 - La definición de la BD a cada uno de los niveles se denomina *esquema*.

Arquitectura de una BD

- Niveles estandarizados:
 - Nivel *interno o físico*: describe la estructura física de almacenamiento y los caminos de accesos a la BD (único)
 - Nivel *conceptual o lógico*: define la estructura de toda la BD sin especificar las estructuras necesarias (único)
 - Nivel *externo*: vista de los usuarios (el esquema externo es el conjunto de esquema parciales de cada usuario/aplicación)
- Los SGBD deben asegurar la independencia entre estos niveles, para ello tenemos:
 - Correspondencia externa/conceptual
 - Correspondencia conceptual/interna



Introducción a los MD

- El mecanismo que usan los SGBD para ocultar detalles del almacenamiento y gestión de datos a los usuarios no administradores son los Modelos de Datos (MD)
 - Respetan los tres niveles: interno, lógico y externo
- Los MD que permiten describir:
 - Estructura de la DB
 - Tipos de datos
 - Interrelaciones entre los datos
 - Restricciones de los datos

Introducción a los MD

- Nivel externo:
 - Es la “vista” que los usuarios tienen de la BD
 - Este nivel describe la parte de la BD que es relevante para cada usuario concreto
 - Solo se incluyen las entidades, atributos y relaciones que son necesarias para dicho usuario
 - Cada (tipo de) usuario tiene una representación de los datos personalizada

Introducción a los MD

- Nivel conceptual / global:
 - Es la vista intermedia entre los otros niveles
 - Describe qué datos están almacenados en la BD y las relaciones existentes entre ellos
 - Ofrece una vista completa de los requisitos de datos que cubre el sistema
 - Es independiente del tipo de almacenamiento físico
- Es el que nos centraremos a estudiar en la asignatura

Introducción a los MD

- Nivel interno:
 - Indica la representación física de los datos
 - Busca una implementación física óptima según los parámetros de rendimiento de uso de la BD
 - Considera aspectos como:
 - Asignación de espacio para datos e índices
 - Descripciones de los registros
 - Ubicación de los registros
 - Compresión de datos y técnicas de cifrado de los datos

Modelos de datos globales

- Los MD se pueden categorizar en:
 - De alto nivel (o conceptuales): cercanos al usuario final de la DB
 - MD representacionales o de implementación
 - De bajo nivel (o físicos): cercanos a los detalles de implementación

Modelos de datos conceptuales

- Los MD de alto nivel se basan en:
 - Entidades: que representan un concepto del mundo a modelar del que se almacenará alguna información en la base de datos
 - Atributos: describen características de interés de una entidad
 - Relaciones: indican las interdependencias entre dos o más entidades
- El MD conceptual más usado es el E/R

Modelos de datos representacionales

- Los MD representacionales:
 - Suelen ser los implementados en los SGBD
 - El más común es el MD relacional
 - Existen otros como jerárquico, en red, etc.

Modelos de datos de bajo nivel

- Los MD de bajo nivel:
 - Se basan en estructuras de datos de bajo nivel
 - Por ejemplo: Árboles B+, estructuras con Hash, etc

Aspectos de un MD

- Todo modelo de datos tienen dos aspectos:
 - Estáticos: descripción de la base de datos (a.k.a. esquema de la BD)
 - No debe cambiar demasiado en el tiempo
 - Dinámicos: la base de datos, los datos concretos que en un momento del tiempo se almacenan (denominados instancias u ocurrencias). Deben cumplir el esquema
 - Pueden cambiar mucho en el tiempo
 - Si hay demasiados cambios en el esquema de una BD mejor ver alternativas a los modelos “clásicos”

Aspectos de un MD

- Limitaciones de aspectos estáticos de un MD:
 - El MD ofrece elementos que permite definir y manipular: objetos, asociaciones, propiedades, dominios, etc
 - No debe cambiar demasiado en el tiempo
 - Elementos que escapan del ámbito del MD usado:
 - Restricciones inherentes: por falta de expresividad en el MD (el humano sabe qué quiere poner pero el sistema no tiene constructores para ello)
 - Restricciones semánticas o de integridad: porque sean restricciones propias de los datos tratados (qué significan los datos en sí, son problemas humanos)

Aspectos de un MD

- Restricciones semánticas y de integridad:
 - Las semánticas atienden al significado de los datos en el mundo real
 - Las de integridad atienden a la corrección de los datos y su consistencia con el mundo real
 - Suelen ir ligadas: una correcta descripción semántica permite que comprobar la integridad de los datos (que puede hacer el SGBD)
- Ejemplo: toda población pertenece a una (y sólo una provincia)

Aspectos de un MD

- Limitaciones de aspectos dinámicos de un MD:
 - Los datos concretos que admite una BD están restringidos por su DM
 - En cada momento del tiempo pueden cambiar
 - Ojo: cuidado con la concurrencia (ejemplo: last_intert_id)
 - Evolucionan usando operadores que proporciona el DM para cada tipo de elemento
 - No son iguales los operadores para objetos que para atributos
 - Cada operación se define por un par (localización, acción)

Aspectos de un MD

- Ejemplo de Esquema de la BD (*estático*)
 - Alumno: DNI, nombre, apellidos, curso
 - Asignatura: código, nombre, créditos
- Ejemplo de base de datos (*dinámico*):
 - Alumno (12345678,Pedro,Pérez,1)
 - Alumno (12345679,Ana,López,2)
 - Asignatura (17101,Álgebra,6)
 - Asignatura (17103,Bases de datos,6)
- *Alumno y Asignatura* son constructores del esquema

Aspectos de un MD

Ventajas de integrar las restricciones en la definición de datos:

- Al compartir datos entre varios programas, las restricciones se almacenan una sola vez
- Las gestiona sólo el diseñador de la base de datos
- Si se incluyen en el esquema, el SGBD se encargará de velar porque los usuarios las cumplan

Aspectos de un MD

- Una reflexión:
 - Hasta ahora en la carrera hemos usado lenguajes de programación que alguien definió por nosotros
 - Con estructuras de control “clásicas”: if, while, etc
 - Con tipos de datos simples limitados
 - Pero esos tipos de datos se podían ampliar (por ejemplo, creando el tipo de dato “número complejo”)
 - Pero si guardo los número complejos en un fichero ¿puedo poner restricciones al uso que hacen de él varios programas?

Referencias

- Apuntes de bases de datos de la profesora Esther Gadeschi 2003/4
 - Apuntes de bases de datos del profesor Carlos Rioja 2006/7
- Libro “Fundamentals of Database Systems” Elmasri y Navathe (3a Ed.)

Gracias por la atención
¿Preguntas?