

IBD – Introducción

Carlos Lombardi

2009

- 1 Sistemas de información
- 2 Bases de Datos
- 3 DB y DBMS — características deseadas
- 4 Modelo de datos
- 5 En la materia

Sistema de información

¿Qué es?

- Cualquier pieza informática que resuelva necesidades de una organización, típicamente manejando información de interés para la organización.

Ejemplos para una universidad

- Seguimiento del legajo académico de cada alumno
- Organización de cursos de cada semestre
- Pago de sueldos a empleados
- Confección de los padrones para las elecciones

Un sistema puede incluir uno o varios programas, no todo programa es parte de un sistema de información (pensar en juegos).

SI - ¿Qué se hace con la información?

¿Qué hay que hacer respecto de la información que maneja?

- Almacenarla (tener registrada cada acta con su detalle: cuatrimestre, materia, alumnos con sus notas).
- Modificarla ante ciertos eventos (entrega de un acta).
- Manipularla (p.ej. obtener el promedio de un alumno, o qué materias le faltan).
- Presentarla (por pantalla, impresora, informar a otro sistema).

SI - ¿Cómo organizo la información?

Necesito definir un **modelo** basado en conceptos formales

- Listas como las que se vieron en Intro a la Programación.
- Modelo de planilla con filas y columnas (como un Excel).
- Fichas como las de una biblioteca.

SI - modelando la información - con fichas

Sistema para legajos académicos, podemos tener

Una ficha por alumno

Legajo	11804
Nombres	Juan Alfredo
Apellido	Pérez
Carreras	TPI

Una ficha por curso

Materia	IBD	
Cuatrim.	02/2008	
Notas	Legajo	Nota
	11804	8
	14302	5
	15321	9

Cómo hacemos para

- obtener la foja académica de un alumno.
- agregar una nota en un curso de 25 y que queden ordenadas por legajo.
- buscar un curso entre 2000.
- almacenar la info si tuviéramos 20 millones de alumnos.

SI - ¿Cómo almaceno la información?

Si está en la memoria del programa . . . varios problemas

- ¿cómo hacer para que sobrevivan si cierro el programa y lo vuelvo a abrir?
- ¿qué pasa si muchos programas necesitan acceder a la misma información?
- ¿qué pasa si se apaga la máquina?
- ¿qué pasa si dos usuarios quieren modificar los mismos datos?
- ¿qué pasa si la info no entra en la RAM de la máquina?
- etc.

Base de datos

En inglés DB = Data Base

¿Qué es una base de datos?

Un conjunto de datos organizados de acuerdo a un modelo y **relacionados** entre sí.

No pueden ser datos sueltos e inconexos, tienen que tener cierta lógica (que es provista por el modelo) y relación entre sí.

Sistema de gestión de Base de Datos

En inglés: DataBase Management System = DBMS

- Es un programa (complejo) que sabe administrar bases de datos.
- Se encarga del almacenamiento, la manipulación y la modificación de la información.
- Los programas que hacemos nosotros acceden a las DB hablándole a un DBMS
 - le puedo hacer consultas, devuelve información.
 - le puedo indicar modificaciones en la información, afectan el resultado de consultas posteriores.
 - tiene muchas otras capacidades.

SGBD = DBMS

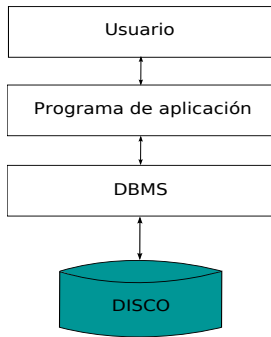
Resuelven los problemas de almacenamiento

- la información queda en almacenamiento persistente (disco, cinta).
- muchos programas pueden acceder a la misma DB.
- tiene mecanismos explícitos de recuperación ante cortes.
- tiene mecanismos para controlar la modificación simultánea de la misma información.
- tiene mecanismos especiales para backup.
- algunas están especialmente preparadas para grandes volúmenes de datos.
- etc.

¿Cómo se relacionan con los sistemas de información?

- cada programa interactúa con una BD (eventualmente varias) mediante SGBD.
- cada BD incluye el formato de la información (definido de acuerdo a un modelo) y la información real.
P.ej. formato de la ficha de alumno + info de cada alumno.
- el SGBD se encarga de almacenar la información en el formato definido para cada BD.
- los programas consultan al SGBD cuando necesitan obtener información
 - ¿qué materias aprobó el alumno de legajo 148032?
 - ¿qué alumnos de apellido Iturbe hay?
- se informa al SGBD de los eventos que implican modificaciones en la información almacenada
 - se agrega un alumno a un acta

BD, SGBD y sistemas de información - esquema



Esquema en capas

BD y SGBD

En resumen

Base de datos la info + su formato

SGBD el programa que administra y permite acceder a
una (o varias) BD

La materia tratará básicamente sobre bases de datos, vamos a aprender a definir las, agregar información, y hacer consultas, todo esto aprendiendo cómo pedirle cada una de estas cosas a un SGBD.

¿DB o cada programa con sus datos?

Para el manejo de información de los sistemas se presentan dos opciones

- que cada aplicación maneje su información con mecanismos internos, que son parte del programa
- que las aplicaciones hablen con DBs separadas mediante DBMS (el esquema que vimos recién).

¿DB o cada programa con sus datos?

Si cada programa maneja sus propios datos, se presentan varias debilidades

- duplicación de información, más trabajo de mantenimiento
- posibilidad de inconsistencia
- es más difícil saber con qué información se cuenta, está entremezclada con los programas.
- es más difícil tener políticas uniformes de resguardo y seguridad.

¿DB o cada programa con sus datos?

Conclusiones

- las organizaciones tienden al uso de DBMS.
- las DB son activos valiosos, aparecen distintas formas de explotarlos.
- Expertos en manejar DB para lograr confiabilidad y eficiencia: DBA (Data Base Administrator).

DB y DBMS — características deseadas

Integridad

garantía de la calidad de la información que hay en una DB, no debería ser posible incluir:

- datos con valores inválidos
 - Un alumno tuvo nota -3 en un parcial.
- incoherencias
 - En una cursada aparece un legajo que no corresponde a ningún alumno.
 - Un alumno cursa dos veces la misma materia en el mismo cuatrimestre.

DB y DBMS — características deseadas

Independencia / transparencia

el acceso a la información es independiente de la forma en la que están guardados los datos:

- en un disco, repartida en varios discos.
- replicada
- cómo están organizados los archivos internamente.
- etc.

DB y DBMS — características deseadas

Control de acceso y modificación concurrente

que la DB funcione correctamente si hay varios usuarios accediendo simultáneamente.

- que distintos usuarios puedan acceder a la misma información al mismo tiempo.
- manejar el conflicto de que dos usuarios modifiquen y/o eliminen la misma información al mismo tiempo.
 - que un usuario pueda bloquear la información que va a modificar.
 - que ante dos modificaciones concurrentes, que tome (p.ej.) la primera que empezó, y descarte la otra avisando al usuario que la pidió.

DB y DBMS — características deseadas

Manejo transaccional

que se puedan agrupar muchas modificaciones en una unidad de trabajo, de forma tal que se hagan o todas o ninguna

- distintas acciones forman parte de una misma unidad.
- si una aplicación falla en el medio de una transacción, no queda registrada “por la mitad”.

DB y DBMS — características deseadas

Performance y escala

- Que las operaciones usuales se hagan con un tiempo de respuesta adecuado,
 - “ensuciando” lo menos posible a las aplicaciones
 - incluso si los volúmenes de datos a manejar son grandes.
- los DBMS usan algoritmos de ordenamiento y búsqueda.
 - esto no se ve en las aplicaciones (transparencia).
 - pueden darse ayudas a un DBMS para que ciertas búsquedas se aceleren (trabajo del DBA).

Otras características deseadas

- seguridad: no todos los usuarios pueden ver todos los datos.
- recuperación ante fallos - backup: backup en caliente / replicación.
- catálogo - metadata: poder contar con información sobre la estructura de la base.

Resumen de las características deseadas

- integridad
- independencia — transparencia.
- control de acceso y modificación concurrente.
- manejo transaccional.
- performance y escala.
- otras.

Definición

Modelo de datos

Conceptos que uso para dar forma que se le da a la información que se va a incluir en una BD.

Usando los conceptos de un modelo defino

- Qué información incluir
- Para cada cosa que pasa en la realidad y que afecta al modelo, qué hay que agregar / modificar / borrar.
- Para cada consulta que quiera hacer, cómo se resuelve.

Algunas decisiones

- ❶ ¿Qué necesito guardar de cada curso?
- ❷ ¿Con qué otras entidades está relacionada cada cursada?
- ❸ ¿Cómo voy a organizar la información en p.ej. una ficha?.
- ❹ ¿Cómo voy a relacionar de qué alumno es cada nota?.
- ❺ ¿Cuántos bits ocupa un número en memoria?
- ❻ ¿Cómo organizo la información de todos los cursos en el disco rígido?

Para organizar mi base de datos para legajos académicos, debo resolver todas estas cuestiones ... que no están todas al mismo nivel.

Dándole forma a una DB — niveles de análisis

Nivel **conceptual**

Qué info debo representar

- de qué cosas hay que tener información
(alumnos, cursos, mesas de integrador)
- qué información de cada una ...
(de cada curso: materia, cuatrimestre, alumnos inscriptos, etc.)
- ... que incluye relaciones
(en este caso entre cursos y alumnos)
- Rangos de valores válidos (p.ej. las notas son de 1 a 10)

Dándole forma a una DB — niveles de análisis

Nivel lógico

Darle una forma a la info que me permita decir, para cada posible consulta, dónde están los datos que la responden.

- organización de la información
(para cada curso una ficha, donde las notas son una tabla adentro de la ficha)
- cómo se manifiestan las relaciones
(anoto el legajo del alumno en la tabla de notas)

Dándole forma a una DB — niveles de análisis

Nivel lógico - ejemplos de consultas

- **fecha de nacimiento del alumno de legajo 13434**: la fecha que está bajo el nombre “nacimiento” para la ficha del alumno que para el nombre “legajo” tiene el valor 13434.
- **notas del alumno de legajo 13434**: para todas las fichas de curso que tienen en la tabla de notas el valor 13434 en la columna “legajo”, la materia de ese curso y la nota que corresponde a ese legajo.
- **notas de Juan López**: para todas las fichas de curso que tienen en la columna “legajo” de la tabla de notas el mismo valor que el que está bajo “legajo” en la ficha de alumno con nombre = “Juan” y apellido = “López”, la materia de ese curso y la nota que corresponde a ese legajo.

Dándole forma a una DB — niveles de análisis

Nivel físico

Organizar la información en el lugar en donde se va a guardar.

- cómo representamos cada valor
(número de X longitud, String de X caracteres, fecha).
- cómo manejamos los archivos donde vamos a guardar los datos.

Dándole forma a una DB — niveles de análisis

Cada modelo de datos apunta a uno de los tres niveles. Para las fichas podemos pensar en

un modelo lógico cada ficha es una tabla de nombre/valor, el valor puede ser una tabla con muchas entradas, para cada ficha elijo un identificador que lo identifica en otras fichas, hay fichas de distinto tipo.

un modelo físico se usan fichas de 9x5 cm, se pone un par nombre/valor en cada renglón, si no alcanza una ficha se le pone una indicación abajo a la derecha y se sigue en otra ficha, va un fichero por cada tipo de ficha, ordenado por el valor del identificador.

Dándole forma a una DB — niveles de análisis

Cada modelo de datos apunta a uno de los tres niveles. En la materia vamos a ver

- **Modelo de entidad-relación** o entidad-vínculo: es **conceptual**
- **Modelo relacional**: es **lógico**

En la materia no nos vamos a meter con el nivel físico, se lo dejamos a los que implementan los SGBD.

Distintos modelos para la misma realidad

Hay mucha información que podría incorporar en una BD
...¿incorporo toda?

- Cantidad de hermanos de cada alumno.
- Color de ojos de cada alumno.
- En qué banco se sentaba cada alumno en cada curso.
- Cantidad de mm llovidos en los horarios de cada curso.

Modelar implica **elegir** qué información incluir y qué información **no** incluir

Con **qué criterio** incluir la info que sea útil para los SI que van a trabajar con esta BD, o para las consultas que se quieran hacer.

De qué vamos a hablar en la materia

Mucho

- modelo de datos relacional
- cómo definir una BD armada según el modelo relacional
- cómo hacer actualizaciones y consultas sobre una BD armada según el modelo relacional

Bastante

- modelo de entidad-relación
- transacciones: marcar unidades de trabajo que se hacen “todo o nada”
- concurrencia: qué pasa si dos usuarios quieren tocar los mismos datos al mi

De qué vamos a hablar en la materia

Algo

- armar pequeños programas que se ejecutan dentro del SGBD (stored procedures)
- cuestiones de performance en una BD (índices, tal vez optimización de consultas)

Casi nada

- Otros modelos de datos (objetos, jerárquico, red).
- Cómo usar una BD desde un programa (p.ej. en lenguaje C).
- Cómo está hecho un DBMS por adentro.