Databases

Computer Literacy 1 Lecture 17
28/10/2008

Topics

- What is a Database?
- What can it do?
- User views
- Layers of the Database
- Example MySQL

What is a Database?

- A collection of interrelated data items that are managed as a single unit
- Broad definition because there is so much variation across the software vendors that provide database systems, e.g. Microsoft, Oracle, Access etc…

Databases

- Databases allow people to store, organise, retrieve, communicate and manage information
- “Information at your fingertips” can make life easier and more efficient: instant cash, airline reservation, websearches, etc…
- Also handy device for Big Brother (referring to George Orwell’s novel “1984” and not to Channel 4)

Electronic File Cabinet

- Database programs are applications
  - Programs for turning computers into productive tools
- If a word processor is a computerised typewriter and a spreadsheet is a computerised ledger, then a database is a computerised file cabinet

What can a Database do?

- Easier to store large quantities of information
- Easier to retrieve information quickly
- Easier to organise and reorganise information
- Easier to distribute

Database Taxonomy

- A database is a collection of information items managed as a single unit
- On the filing cabinet analogy, a database is composed of one or more files
- A file is a collection of related information
  - Related in terms of some selected organizing principles

Files and Databases

- Term file could cause confusion because of its multiple meanings
- A disk can contain application programs, system programs, documents, etc. All of which are files from the computer's point of view
- But for database users the term means a file that's part of a database

Jargon Abuse

- For database users a file is part of a database
- But an entire database can be just one file, from the computer's point of view
- Even the term database has more then one use, e.g. various Unix vendors call their password file a database, even though experts would say that it isn't one.

Some Database Taxonomy

- In order to clarify the issue we need to look at characteristics that distinguish databases from "ordinary" files
- These include:
- Database Management System (DBMS) e.g. MySQL
- Layers of abstraction architecture

DBMS

- The Data Base Management System is software provided by the database vendor
- The DBMS provides the services required to organise and maintain your database

DBMS

- Moving data to and from the physical data files as needed
- Managing concurrent data access by multiple users
- Support for a query language: a system of commands that a database user employs to retrieve data
- Provisions for backing up the database and recovering from failures
- Security mechanisms

Different Views

- Databases have the unique capability of presenting multiple users with distinct views of the data while storing the underlying information only once
  - These are called user views
- A user in this context is any person or application that signs on to the database to store or retrieve data

User Views

- With a spreadsheet application such as Excel, all users must share a common view of the data
  - View must match way it is
- With database systems the view can be tailored or costumised to the needs of individual users

User Views

- For example managers, programmers and costumers can all have different views of a DVD rental website's database
  - Costumers can browse through listings
  - Managers has control over pricing policies, view of inventory history
  - Programmer has control over EVERYTHING

Layers of Abstraction

- Most DBMSs follow a layer of abstraction architecture
- Which makes different user views possible
- The three primary layers are:
  - The physical layer
  - The logical layer
  - The external layer

The Physical Layer

- The physical layer contains the data files that hold all the information for the database
- Most DBMSs allow the database to be stored in multiple files, usually spread over multiple disk drives

More Physical Layer

- The database user doesn't need to have any knowledge of how the data is actually stored in these files
- Doesn't need to refer to physical files when using the database
- This is in **contrast** to *spreadsheets* and *word processing* programs

The Logical Layer

- The logical layer is the first level of abstraction
- The physical layer has a concrete existence in the operating system file
- In contrast the logical layer exists only as abstract data structure assembled from the physical layer as needed

More Logical Layer

- The DBMS transforms the data in the data files into a common structure, sometimes called **schema**
- Depending on the particular DBMS this can be a set of 2-dimensional tables, a hierarchical structure, a networkmodel etc…

The External Layer

- The external layer is the second level of abstraction
- This is the layer where users and applications issue queries against the database obtain the resulting user views
- The DBMS transforms data structures in the logical layer to form each user view

Example for DMBS = MySQL

- MySQL is a (R)DMBS
  - R = Relational
- It is written in C/C++
- It works on many different system platforms (Windows, Mac, Unix, etc)
- The libraries for accessing MySQL databases are available in all major programming languages with language specific APIs

Key Points

- What is a database and what can it do?
- What's the difference between a database and ordinary file?
- Layers of Abstractions architecture
  - Allows different user views
- MySQL RDMBS