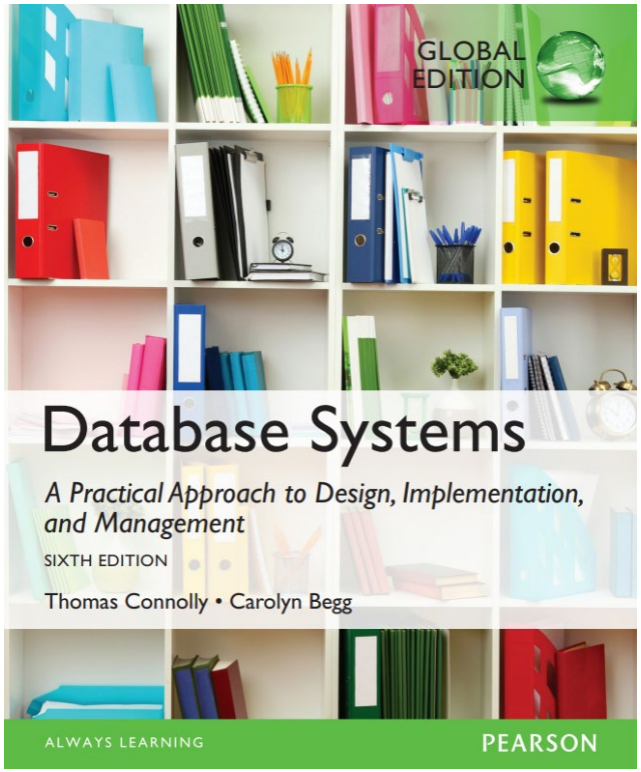
Database Design and Development

Unit 4



Lecture 1

Background on database design

BTEC Alliance with FPT Education

Pearson

Chapter outcomes

- By the end of this chapter you will be able to:
  - Define relational databases
  - Identify major RDBMSs
  - Identify main characteristics of RDBs
  - Understand SQL's role in RDBs
  - Recognize where a database could be useful

Databases

- Databases
  - Set of logically related data designed to meet organization's need
  - It doesn't have to be electronic, but usually is
- There are different main types of databases
  - Flat File Databases
  - Hierarchical Databases
  - Relational Databases

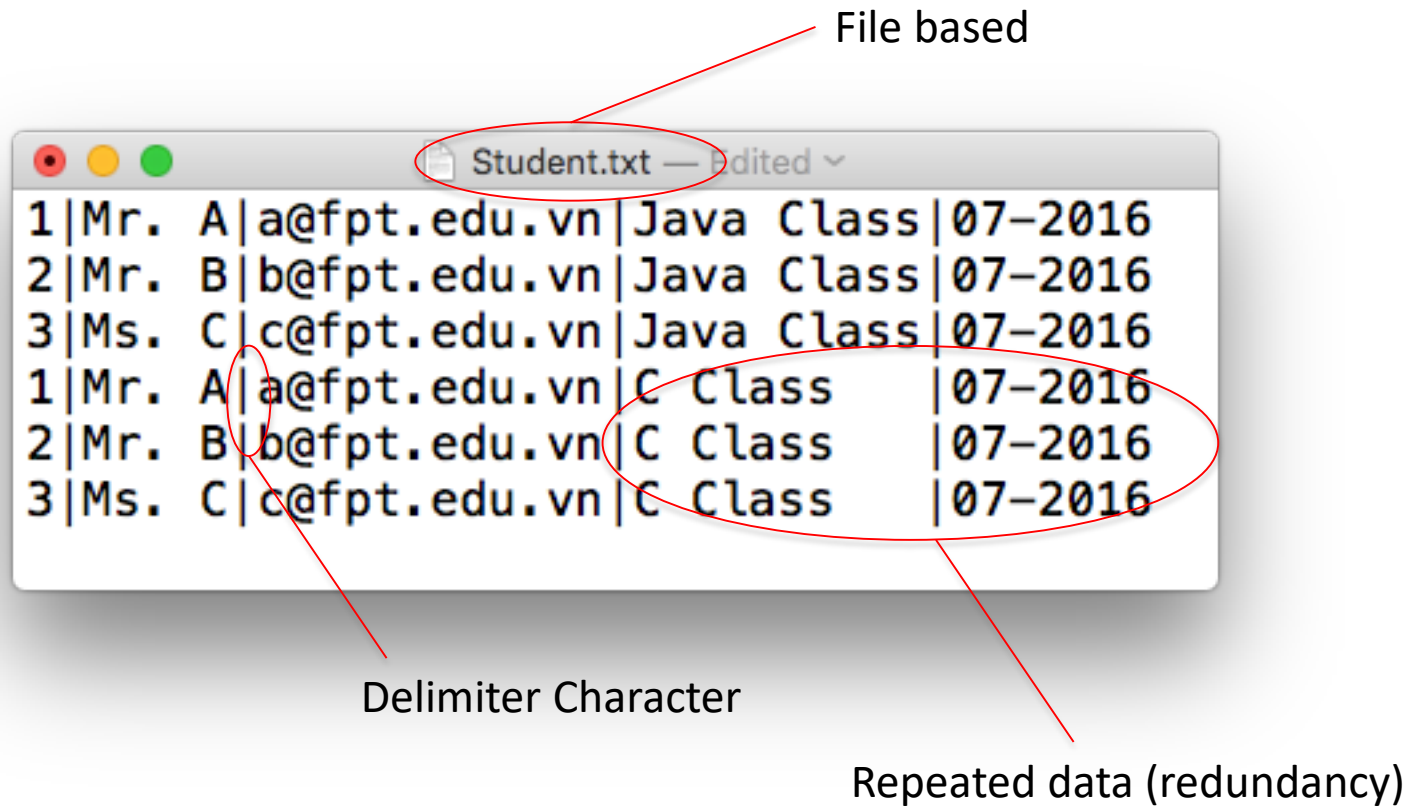
Flat file databases

- Flat file databases
  - Simplest form of an electronic database is the flat file database
  - Consist of a file which stores data in a structured way
  - A common format is the delimited file
- Delimited files
  - These have some sort of character separating columns of data
  - Delimiter is a coma/tab (or any non-alphanumeric character)
- Disadvantages
  - Have almost no protection for data **integrity** and security
  - Often contain many **redundancy** (repeated data)

Data Integrity, Redundancy?

- Data integrity
  - Accuracy/correctness of the data in the database
  - E.g., age must be integers
- Redundancy
  - Storing the same data in more than one place

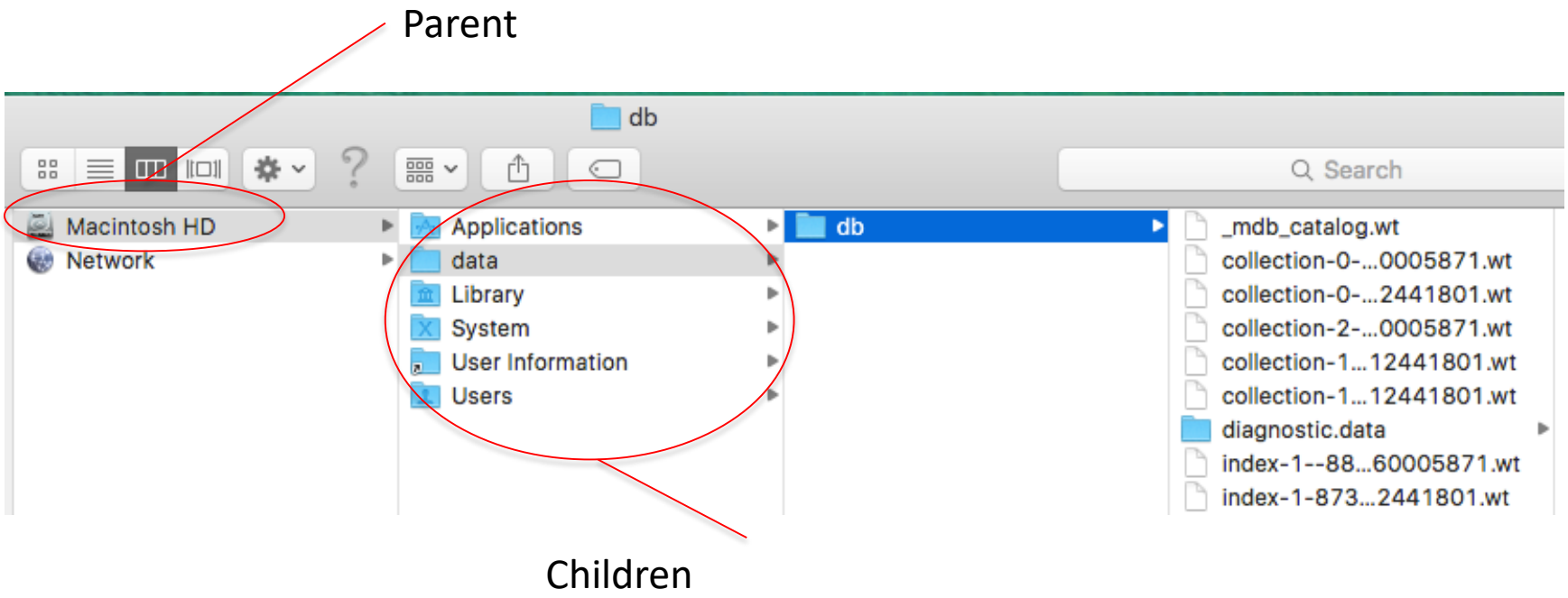
Flat File Databases



Hierarchical Databases

- Hierarchical databases are organized in a tree-like structure
  - Parent table can have many child tables
  - No child table can have more than one parent
  - They are connected to one another through links
- E.g.,
  - Directories/Sub directories/Files hierarchies in OS
- Disadvantages
  - It does present the same problems of redundancy, data integrity, and comparability of data

Hierarchical Databases



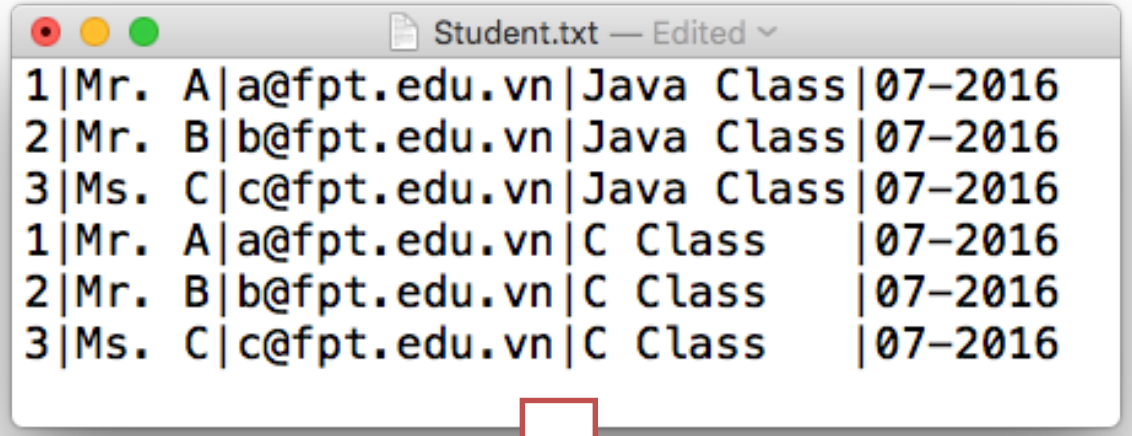
Relational Databases

- In relational model, data is organized into tables
  - Even the info about the tables is stored in tables
- Relationship among tables
  - Defined by repeating column(s) from one in another table
  - These repeating columns are called “keys”
- This solved many problems
  - One of those is data redundancy

Keys?

- In RDBs, each table usually has one (or more) column(s) designated as a **primary key (khóa chính)**
- A **key** uniquely identifies each row in a table
  - Giving one of its values, you can find exactly one row in the table
- This key becomes a **foreign key** when it is repeated in another table
  - To create **relationship** between the tables

Relational Databases



Primary keys

Repeating only key column
(to keep the relationship)

sId	sName	sEmail
1	Mr. A	a@fpt.edu.vn
2	Mr. B	b@fpt.edu.vn
3	Ms. C	c@fpt.edu.vn

sId	cId	cStartDate
1	1	Jul-16
2	1	Jul-16
3	1	Jul-16
1	2	Jul-16
2	2	Jul-16
3	2	Jul-16

cId	cName
1	Java Class
2	C Class

Foreign keys

Practical Tutorial 1

- Guide student how to create a database specified in previous slide using the Designer
  - Show the key
  - Try to insert some data
  - Try to insert duplicated key

Drawback of RDBMS

- It’s the complexity of the design
- So you should follow a design process that allows you to design a DB to achieve
  - Integrity
  - Flexibility

Situations where databases are for

- Due to the complexity of the RDB design
  - RDB is good when there is a large amount of complex data to handle

Activity

- Here are some situations where DB is good for
  - Student Management System
  - Supermarket System
  - Library Management System
  - Etc.,
- Please
  - Give some further situations to use DBs
- Explain why they are suitable to use DBs for these

Scenario – for this course

- There are 2 paper pages each week (AM and PM)
- Beginning of a month, tutors enter:
  - Their availability for each day
  - Duration for each day
  - Courses they can tutor
- Students sign up for particular sessions
  - It's free as long as they are registered in the class for which they are getting tutored
- Tutoring sessions are 30 minutes each
- A tutor can do maximum 15 hours a week
- As long as tutor shows up when scheduled, tutor is paid about $10.5/hour

Scenario – Current system and opportunity

- Current system
  - Paper based records are taken back to the office every 2 weeks and typed into Excel
- Opportunity
  - A DB is proposed to replace this

DBMS

- A DBMS is a system for managing DBs
- It supports
  - Creating/Manipulating DBs
  - Maintaining DB (backing up/recovering, etc.)
  - Security for the DB
  - Etc.

Some famous RDBMSs

Table 1-2 Some Relational Database Management Systems

RDBMS	Comments	URL
ORACLE	The biggest and the first commercial RDBMS. Powers many of the world's largest companies	http://www.Oracle.com
SQL Server	Microsoft's RDMS product. Ships in many versions designed for different company needs. Also powers many large enterprises	http://www.microsoft.com/sql/default.mspx
DB2	IBM's RDBMS	http://www306.ibm.com/software/data/db2/9/
MySQL	The most popular open source RDBMS, currently owned by SUN	http://www.MySql.com
PostGres SQL	Another free, open source RDBMS. It is older and some would say more powerful than MySQL	http://www.postgresql.org/
ACCESS	Microsoft's desktop database	http://office.microsoft.com/en-us/access/default.aspx?ofcresset=1

SQL(structured query language)

- Allows users to access data in relational database management systems
- Allows users to describe data
- Allows to embed within other languages SQL modules, libraries & pre-compilers
- Allows users to create and drop DBs and tables
- Allows users to create view, stored procedure, functions in a database
- Allows users to set permission on tables, procedures, and views

https://www.tutorialspoint.com/sql/sql-overview.htm

Practical Tutorial 2

- Use SQL Code to create the same database that you've made in Practical Tutorial 1

How do you get the requirements?

- One of the ways is to prepare an interview, in which you can ask about
  - What things the client is doing/recording?
  - What business constraints are required?
  - What reports are expected?
  - Etc.,

Identifying the big topics

- After the interview, first thing to do is to identify the big topic
  - What the database is about?
  - What are the major components going to be?
  - What does it include?
- Specifically, list the entities of the DB and specify the attributes inside them
- How to find these?
  - One way is to look at the **nouns** in your document

Entities and Attributes?

- An entity(thực thể) is something that the database is concerned with
  - Data is stored about this
  - It may have relationship with other entities
- Attributes (thuộc tính) define entities
  - The entity student has attributes like Id, name, DoB, email, etc.

Activity: Finding the entities

- Find the entities for this “tutoring” scenario

Activity Result

- Tutors
- Students
- Schedules
- Courses
- Requests
- Sessions

Getting the scope

- Statement of work
  - Is a short statement of one or more paragraphs
  - Says in clear, general terms what project will do
  - It's a more complete statement about the objectives and timeline of the project
- Why?
  - We are making a DB for a client not just ourselves
  - Not get trapped by preconceived notions
  - Need to get as clear as possible about what DB is intended to do

Elements of Statement of Work

- History: Reasons for the project
  - Problem of the current system or
  - Opportunity to provide new services
- Scope: Requirements and expectations
  - States high level requirements
  - It doesn't go into details about how things are done
  - May include some general constraints (time, budgets)
- Objectives: Things intended to achieve
  - What database is supposed to achieve
  - I.e., why the client wants the DB
- Tasks and deliverables:
  - Project is broken into discrete tasks with time and deliverables

Activity (Homework)

- Study statement of the work for the “tutoring” scenario in the textbook of this course

Documenting a DB

- It’s important
  - Imagine if you taken over a DB from another
  - How do you work with it if you don’t have a doc
- How to document a DB? There are two main aspects to describe:
  - Process by which the DB was developed
  - The structure of the DB

Documenting DB

- Describe processes by which DB will be developed
  - The first one is the Statement of the work
  - In the future, you will learn more documents
- Structures of the DB, describes “Data dictionary”:
  - Tables
  - Columns and their data types
  - Relationships among tables

Things we have done

- Identified situation in which a DB could prove valuable
- Reviewed briefly the history of DBs
- Identified some components of RDBMS
  - e.g., entities, attributes and key fields
- Using interview
  - To gather general information about a DB
- Developing Statement of work for a DB

Activity

Vocabulary

Match the definitions to the vocabulary words:

1. Attribute	— a. A type of database that uses "relations," tables, to store and relate tables.
2. Foreign key	— b. The process of organizing data into tables or entities and then determining the relations among them.
3. Statement of work	— c. The language relational databases use to create their objects and to modify and retrieve data.
4. Primary key	— d. These files have some sort of character separating columns of data. The delimiter is often a comma or tab, but it can be any non-alphanumeric character.
5. Data integrity	— e. Files where the length in characters of each column is the same.
6. Redundancy	— f. Refers to the accuracy and the correctness of the data in the database.
7. Delimited files	— g. Refers to storing the same data in more than one place in the database.
8. Relational database	— h. This key uniquely identifies each row in the table.
9. Entity	— i. This key is the primary key repeated in another table to create a link between the tables.
10. Relational design	— j. A short statement of one or more paragraphs that says in clear, but general, terms what the project will do.
11. SQL	— k. Something that the database is concerned with, about which data can be stored.
12. Constraints	— l. Things that define aspects of entities.
13. Fixed width files	— m. Limits on what the database will do.
	— n. A document including the scope, objectives, and timeline for a given project.

References

- Cogner, S., 2012. *Hands-on Database: An Introduction to Database Design and Development*. Prentice Hall.