Chapter 1 Introduction

Tian Tang
United International College

February 14, 2019

Instructor Contact Information

Section No.	Instructor	Office	Email
1001	Prof. Weifeng Su	T3-602-R3	wfsu@uic.edu.hk
1002	Dr. Zhiyuan Li	T3-502-R6	goliathli@uic.edu.hk
1003	Dr. Zhiyuan Li	T3-502-R6	goliathli@uic.edu.hk

Course Information

- TextBook:

Database System Concepts
Sixth Edition

Avi Silberschatz
Henry F. Korth
S. Sudarshan

McGraw-Hill
ISBN 0-07-352332-1

Face The Real World of Database Systems Fully Equipped

Course Information

- Lecture:

Section	Time	Venue
1001	Wed. 9:00-10:50	T6-603
	Tue. 16:00-16:50	T4-601
1002	Mon. 10:00 -11:50	T7-303
	Tue. 15:00-15:50	T5-405
1003	Fri. 10:00 -11:50	T4-603
	Tue. 13:00-13:50	T4-404

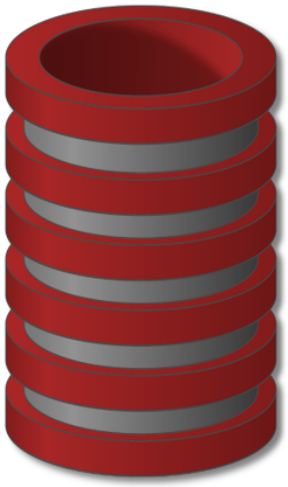
Course assessment

- 4 Assignments: 10%
- Midterm: 20%
- Project: 30%
- Final Exam: 40%

Course assessment

Week	Topics	Readings
Week 1	Chapter 1:Introduction	Textbook (Chap.1)
Week 2&3	Chapter 2:Relational Model	Textbook (Chap.2)
Week 4&5	Chapter 3:SQL	Textbook (Chap.3)
Week 6	Chapter 4:Advanced SQL	Textbook (Chap.4)
Week 7	Midterm	
Week 8&9	Chapter 6:ER model	Textbook (Chap.6)
Week 10&11	Chapter 7:Database Design	Textbook (Chap.7)
Week 12	Chapter 10:XML database	Textbook (Chap.10)
Week 13	Final	

Introduction



Whether you know it or not, you're using a database evetyday.

Database Management System (DBMS)

- Database Applications
  - Banking: transactions
  - Airlines: reservations, schedules
  - Universities: registration, grades
  - Sales: customers, products, purchases
  - Online retailers: order tracking, customized recommendations
  - Manufacturing: production, inventory, orders, supply chain
  - Human resources: employee records, salaries, tax deductions
  - Social media sites: posts, friendship relationship, activities
- Database can be very large.
- Database touch all aspects of our lives.

Why we need database: University Database Example

- Course Management
  - Add students, instructors, and courses
  - Register students for courses, and generate class rosters
  - Assign grades to students, compute grade point averages (GPA) and generate transcripts
- Provide basic features necessary for data access
  - Shared access by a community of uses
  - Well-defined schema for data access
  - Support query language
- In the early days, applications were built directly on top of file systems.

Drawbacks of using file systems to store data

- Data redundancy and inconsistency
  - Multiple file formats, duplication of information in different files
- Difficulty in accessing data
  - Need to write a new program to carry out each new task
- Data isolation
  - Multiple files and formats
- Integrity problems
  - Integrity constraints (e.g., account balance ≥ 0) become buried in program code rather than being stated explicitly
  - Hard to add new constraints or change existing ones

Drawbacks of using file systems to store data (Cont.)

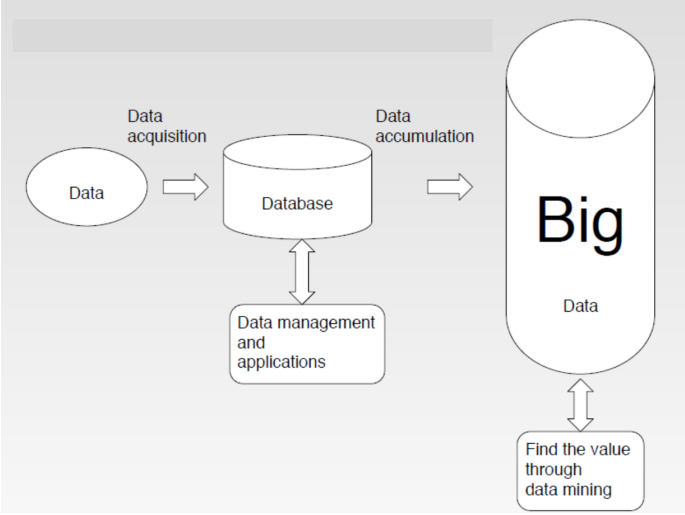
- Atomicity of updates
  - Failures may leave database in an inconsistent state with partial updates carried out.
  - Example: Transfer of funds from one account to another should either complete or not happen at all.
- Concurrent access by multiple users
  - Concurrent access needed for performance
  - Uncontrolled concurrent accesses can lead to inconsistencies
    - Example: Two students want to choose the same course that has only one vacancy left.
- Security problems
  - Hard to provide user access to some, but not all, data

Database system offer solutions to all the above problems.

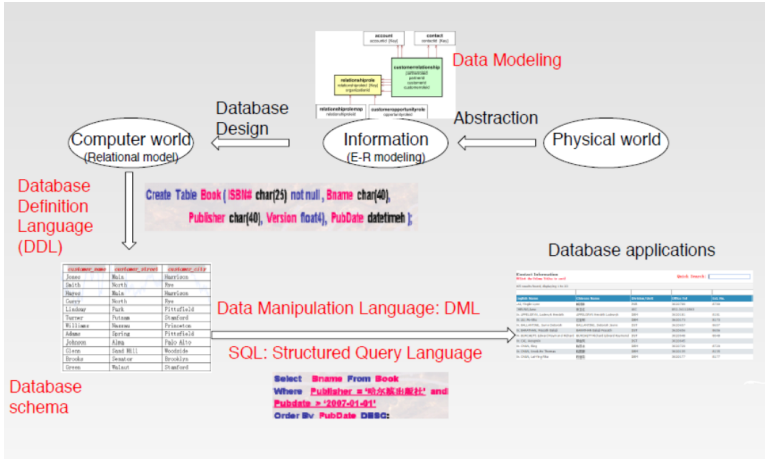
Database Management System (DBMS) provides...

- Efficient
- Reliable
- Convenient
- Safe
- Multi-user (Storage of and access to)
- Massive amounts of persistent data

In general: why we need database

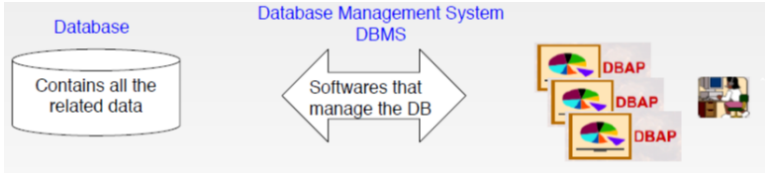


Key concepts



Key concepts

- Data modeling
- Schema vs. data instances
- Data definition language (DDL)
- Data manipulation or query language (DML)
- DBMS implementer
- Database designer
- Database application developer
- Database administrator



History of Database Systems

- 1950s and early 1960s:
  - Data processing uses magnetic tapes for storage.
    - Tapes provide only sequential access.
  - Punched cards are used for inputs.
- Late 1960s and 1970s:
  - Hard disks allow direct access to data.
  - Network and hierarchical data models are in widespread use.
  - Edgar Frank Codd invented the relational data model.
    - He won the ACM Turing Award in 1981.
    - IBM Research begins System R prototype.
    - UC Berkeley begins Ingres prototype.




History of Database Systems (Cont.)

- 1980s:
  - Research relational prototypes evolve into commercial systems.
    - SQL becomes an industrial standard.
  - Object-oriented database systems
- 1990s:
  - Large decision support and data-mining applications.
  - Large multi-terabyte data warehouses
  - Emergence of Web commerce
- 2000s:
  - XML and XQuery standards
  - Automated database administration

Big Names in Database Systems

Company	Product	Remarks
Oracle	Oracle 10i, 11i, 12i, ect. Berkeley DB, MySQL	World's 2nd largest software maker by revenue, after Microsoft
IBM	DB2	Since 1970s, when Edgar F. Codd described the theory of relational databases.
Microsoft	Access, SQL Server	Access comes with MS Office
Sybase	Adaptive Server	Relational model database server product for business
Informix	Informix Dynamic Server (IDS)	Acquired by IBM in 2001

Database Popularity Ranking-2018

Rank Jan 2018	Rank Dec 2017	Rank Jan 2017	DBMS	Database Model	Score Jan 2018	Score Dec 2017	Score Jan 2017
1.	1.	1.	Oracle	Relational DBMS	1341.94	+0.40	-74.78
2.	2.	2.	MySQL	Relational DBMS	1299.71	-18.36	-66.58
3.	3.	3.	Microsoft SQL Server	Relational DBMS	1148.07	-24.42	-72.89
4.	4.	↑ 5.	PostgreSQL	Relational DBMS	386.18	+0.75	+55.81
5.	5.	↓ 4.	MongoDB	Document store	330.95	+0.18	-0.96
6.	6.	6.	DB2	Relational DBMS	190.28	+0.70	+7.78
7.	7.	↑ 8.	Microsoft Access	Relational DBMS	126.70	+0.82	-0.75
8.	↑ 9.	↓ 7.	Cassandra	Wide column store	123.88	+0.67	-12.57
9.	↓ 8.	9.	Redis	Key-value store	123.14	-0.10	+4.44
10.	10.	↑ 11.	Elasticsearch	Search engine	122.55	+2.77	+16.38
11.	11.	↓ 10.	SQLite	Relational DBMS	114.25	-0.94	+1.88
12.	12.	12.	Teradata	Relational DBMS	72.63	-2.11	-1.54
13.	↑ 14.	13.	SAP Adaptive Server	Relational DBMS	65.46	-0.22	-3.64
14.	↓ 13.	14.	Solr	Search engine	64.37	-1.93	-3.71
15.	15.	↑ 16.	Splunk	Search engine	64.00	+0.21	+8.51
16.	16.	↓ 15.	HBase	Wide column store	61.64	-1.78	+2.50
17.	17.	↑ 20.	MariaDB	Relational DBMS	58.30	+1.56	+13.26
18.	↑ 19.	↑ 19.	Hive	Relational DBMS	55.49	+0.81	+4.35

End of Chapter