DATABASE DESIGN I
1DL300 Spring 2013
Sobhan Badiozamany
Silvia Stefanova
Course organization

L1 Introduction
L2 ER modeling
L3 HI covering
L4 Relational model
L5 Normalization
L6 Assessment example
Assignment 1 computer lab
L7 Relational algebra
L8 SQL query language
L9 SQL (continued)
Assignment 2 computer lab II
L10 Database programming
L11 Transactions and concurrency control
Assignment 3 computer lab III
L12 Recovery
Assignment 3 computer lab IV
L13 Guest lectures
L14 Introduction to physical DB design
L15 Feedback on study questions, general questions

What is a database?

DATABASE DESIGN I
1DL300 Spring 2013
Sobhan Badiazamany
Silvia Stefanova

DB is like a huge iceberg
Course organization

L1 Intro to database terminology
L2 ER modeling
L3 ER example
L4 Relational model
L5 normalization
L6 Normalization example

**Assignment 1 computer lab**
- ER modeling & database.
- Conceptual ER Bring exercise, no computer is really needed.

**L7 Relational algebra**

**L8 SQL query language** Prepare/bring USB sticks.

**L9 SQL [continued]** Bring USB sticks.

**Assignment 2 computer lab #1**
- Establishing a workload designed database.
- Imparability in SQL.

**L10 Database programming**

**L11 Transactions and concurrency control**

**Assignment 2 computer lab #2**

**L12 Recovery**

**Assignment 3 computer lab**
- Database recovery (MIS451)
- More SQL queries.

Lx Guest lecture

**L13 Introduction to physical DB design**

**L14 Feedback on study questions, general questions**
Literature and resources

- Databasteknik, Studentlitteratur, 2005: Padron-McCarthy, T. & Risch, T.

- Both books are available at Akademibokhandeln.

Course home page:
https://www.it.uu.se/edu/course/homepage/dbastekn/vt13/
Personnel

• Examiner
  • Kjell Orsborn
    • email: kjell.orsborn@it.uu.se, phone: 471 1154, room: 1321, ITC building 1, floor 3

• Lecturers
  • Sobhan Badiozamany
    • email: sobhan.badiozamany@it.uu.se, phone: 471 6345, room 1354, ITC building 1, floor 3
  • Silvia Stefanova
    • email: silvia.stefanova@it.uu.se, phone: 471 2846, room 1319, ITC building 1, floor 3

• Course assistants
  • Thanh Truong
    • thanh.truong@it.uu.se, phone: 471 6345, room 1354, ITC building 1, floor 3
  • Andrej Andrejev
    • andrej.andrejev@it.uu.se, phone: 471 7345, room 1306, ITC building 1, floor 3
  • Lars Melander
    • lars.melander@it.uu.se, phone: 471 7345, room 1306, ITC building 1, floor 3

All questions regarding registration is handled on Studentportalen, by the Student Office, or by your study counselor. None of us can handle any such administrative issues.
Prezi!

- Cloud-based presentation tool with a …
- Presentations are always accessible through Prezi web site, but,
  - You can also download them and play them offline (outside the browser).
  - iPhone/iPad free Prezi viewer.
- Good for explaining ideas in a lecture, but not necessarily good in studying for the exam.
- Downside: The PDF files generated from Prezi are images, no text search possible.
- Generally, you will need to read the book if you need more detailed information.
zooming user interface.
Prezi!

- Cloud-based presentation tool with a ... accessing user interface.
- Presentations are always accessible through Prezi web site, but,
  - You can also download them and play them offline (outside the browser).
  - iPhone/iPad free Prezi viewer.
- Good for explaining ideas in a lecture, but not necessarily good in studying for the exam.
- Downside: The PDF files generated from Prezi are images, no text search possible.
- Generally, you will need to read the book if you need more detailed information.
Examination

- Three assignments [plus possibly a shorter terminology based one]
  - The assignment home page is here: https://www.it.uu.se/edu/course/homepage/dbastekn/ht12/dbt1-ht2012-assignments.html
  - Form groups of 2 individuals in Studentportalen.
- Have already done some of the assignments from previous years?
  - You can skip what you have done.
- The written exam
  - Don’t forget to sign up for it.
  - To pass the course, you need to pass the exam and all three assignments.
  - The written exam will have a combination of tasks to solve and technical terminology to explain.
Assignments practicality [1]

- A written report is expected for each assignment.
- Attendance to labs are highly recommended.
- Progress will be reflected in Studentportal.
- Always keep signed assignment papers as a proof that you have passed an assignment.
- All labs need sign-up, sign up sheets will be posted before labs.
- Read instructions before going to labs.
Assignments practicality [2]

• Respect the deadlines:
  • Soft deadlines: approx. one week after each lab.
  • The hard deadline is 27th March, After which:
    • Submissions (and re-submissions) are not considered.
    • To get unfinished assignments corrected, Students need to re-register to the next course instance.
• Each group will be assigned a course assistant before the first lab.
• Respect your assistant's office hours. [They are your virtual deadlines]
portable MySQL.

- The whole database system is on a USB stick.
- Each group is given a flash disk, and yes, you can keep it.
- Required for:
  - Assignments 2 and 3.
  - Interactive lectures in SQL.
- Installation takes 30-40 minutes, do it before labs/lectures.
L1 Intro to database terminology
L2 ER modeling
L3 ER example
L4 Relational model
L5 normalization
L6 Normalization example

Assignment 1 computer lab
- ER modeling & Normalization
- Conceptual DB Design exercise, no computer is really needed.

L7 Relational algebra
L8 SQL query language  Prepare/bring USB sticks.
L9 SQL [continued]  Bring USB sticks.

Assignment 2 computer lab #1
- Extending a previously designed database
- Implementation in SQL

L10 Database programming
L11 Transactions and concurrency control

Assignment 2 computer lab #2
L12 Recovery

Assignment 3 computer lab
- JDBC API access to RDBMS
- More SQL query

Lx Guest lecture
L13 Introduction to physical DB design
L14 Feedback on study questions, general questions
L1 Intro to database terminology
L2 ER modeling
L3 ER example
L4 Relational model
L5 normalization
L6 Normalization example

Assignment 1 computer lab
- ER modeling & Normalization
- Conceptual DB Design exercise, no computer is really needed.

L7 Relational algebra
Assignment 1 computer lab

- ER modeling & Normalization
- Conceptual DB Design exercise, no computer is really needed.

L7 Relational algebra

L8 SQL query language

L9 SQL [continued]
Assignment 1 computer lab

- ER modeling & Normalization
- Conceptual DB Design exercise, no computer is really needed.

L7 Relational algebra
L8 SQL query language
L9 SQL [continued]

Prepare/bring USB sticks.
Bring USB sticks.

Assignment 2 computer lab #1

- Extending a previously designed database
- Implementation in SQL

L10 Database programming
L11 Transactions and concurrency
Relational algebra
SQL query language
SQL [continued]

Prepare/bring USB sticks.

Assignment 2 computer lab #1

- Extending a previously designed database
- Implementation in SQL

L10 Database programming
L11 Transactions and concurrency
Assignment 2 computer lab #1
- Extending a previously designed database
- Implementation in SQL

L10 Database programming
L11 Transactions and concurrency control

Assignment 2 computer lab #2

L12 Recovery

Assignment 3 computer lab
- JDBC API access to RDBMS
- More SQL queries

Lx Guest lecture
L13 Introduction to phy
L14 Feedback on study
2 Recovery

Assignment 3 computer lab

- JDBC API access to RDBMS
- More SQL queries

Lx Guest lecture
L13 Introduction to php
L14 Feedback on student
Lx  Guest lecture
L13 Introduction to physical DB design
L14 Feedback on study questions, general questions

- JDBC API access to RDBMS
- More SQL queries
Course organization

L1  Intro to database terminology
L2  ER modeling
L3  ER example
L4  Relational model
L5  Normalization
L6  Normalization example

Assignment 1 computer lab
- ER modeling, database example
- Convert HR data to a conceptual model

L7  Relational algebra
L8  SQL query language  Prepare/bring USB sticks.
L9  SQL [continued]  Bring USB sticks.

Assignment 2 computer lab #1
- Developing a word bank database
- Importing data into SQL

L10 Database programming
L11 Transactions and concurrency control

Assignment 2 computer lab #2

L12 Recovery

Assignment 3 computer lab
- JDBC API (exercise 1 M3M34)
- More SQL queries

Lx  Guest lecture
L13 Introduction to physical DB design
L14 Feedback on study questions, general questions
Databases are everywhere

- When you are shopping
- When you make a phone call
- When you Google
What makes databases so widespread?

Imagine an ATM (Automated Teller Machine) . . .

- Persistent storage
- Access control
- Multi-user
  - Concurrency control
- Efficient
  - 1000’s of queries and/or updates per second
- Reliable
  - 99.99999%

Applications supported by databases are:

- Easier to implement and maintain
  - Physical data independence
  - High-level declarative query language
- Support massive data volumes
  - Google processes about 24 petabytes of data per day.
Many actors are involved in a database

- End Users
  - naive or parametric end users
  - Casual end users
  - sophisticated end users
  - Standalone users
- Database designers
- Database administrators
  - Operation and maintenance
  - Tuning and optimization
- Database-system designers and implementers
Early Database Applications

Hierarchical (IMS) Trees

Network model (CODASYL) Graph

1960 1970
Relational model based systems

(e.g. ORACLE)

Tables

1980
Object oriented DBMS
(e.g. ObjectStore)
OO data structures

Object-relational DBMS
(e.g. SQL:99)
Object model

1990 1997
Outline of a database system

A database system makes simple and efficient manipulation of large data sets possible.

A database system consists of . . .
- the physical database (instance)
- a database management system
- one or several database languages
  - means for communicating with the database
- one or several application program(s)

The term DB can refer to both the content and to the system (the answer to this ambiguity is governed by the context).
Database

A database (DB) is a more or less well-organized collection of related data.
The information in a database . . .
• represents information within some subarea of “the reality”,
  that is, objects, characteristics and relationships between objects
• is logically connected through the intended meaning
• has been organized for a specific group of users and applications
A database management system (DBMS) is one (or several) program that provides functionality for users to develop, use, and maintain a database.

Thus, a DBMS is a general software system for defining, populating (constructing), manipulating and sharing databases for different types of applications.

Also supports protection (system and security) and maintenance to evolve the system.
A database is a more or less well-organized collection of related data. The information in a database...
- represents information within some subset of “the reality”, that is, objects, characteristics, and relationships between objects
- is logically connected through the intended meaning
- has been organized for a specific group of users and applications
Databases are everywhere

- When you are shopping
- When you Google
- When you make a phone call
What is a database?

Tip
DB is like a huge iceberg

Outline of a database system:
- Databases are everywhere
  - When you shop
  - When you make a phone call
- Databases are everywhere
  - When you shop
  - When you make a phone call

DB is like a huge iceberg
DATABASE DESIGN I - 1DL300

Spring 2013

An Introductory Course on Database Systems

http://www.it.uu.se/edu/course/homepage/dbastekn/vt13/

Uppsala Database Laboratory
Department of Information Technology, Uppsala University,
Uppsala, Sweden
Introduction to Database Technology

Elmasri/Navathe ch 1-2
Padron-McCarthy/Risch ch 1

Sobhan Badiozamany
Silvia Stefanova

Department of Information Technology
Uppsala University, Uppsala, Sweden
Outline

1. Data Model, Database Schema
   • Categories of data models
   • Database schemas and instances

2. Data Independence
   • Three-schema architecture
   • Logical independence
   • Physical independence

3. Database Languages

4. Classification of DBMS

5. To Summarize
Outline

1. Data Model, Database Schema
   - Data models, categories of data models
   - Database schemas and instances

2. Data Independence
   - Three-schema architecture
   - Logical independence
   - Physical independence

3. Database Languages

4. Classification of DBMS

5. To Summarize
Data Model ???
Data Model

• **Data Model**: A collection of concepts that can be used to describe the structure of database

• Data Model consists of
  - Set of concepts: data types, relationships, constraints
  - Set of basic operations for specifying retrievals and updates on the database

• Every database has a data model which makes it possible to “hide” the physical representation of data
Categories of data models

• **High-level** (conceptual): close to how users perceive data
  - ER (Entity-Relationship) model

• **Low-level** (physical): details of how data is stored on computers

• **Representational** (implementation): in between conceptual and physical
  - Relational (ORACLE, DB2, SQL Server, MySQL, Mimer)
  - Object-oriented (ObjectStore, Objectivity, Versant, Poet)
  - Object-relational (Informix, Oderpreter, DB2)
  - Hierarchical (IMS)
  - Network (IDMS)
What is the difference between database and database description?
Database Schema and Instances

• **Database schema**: the database description specified during the database design and *not expected to change*

• The **actual data** in the database *may change frequently*

• **Database instance (state)**: the data in the database at a particular moment. Every change of data creates a new instance of the database.
  
  o **Valid state**: satisfies the structure and constraints in the database schema
  
  o **Meta-data**: description of the schema (data about data)
Outline

1. Data Model, Database Schema
   • Categories of data models
   • Database schemas and instances

2. Data Independence
   • Three-schema architecture
   • Logical independence
   • Physical independence

3. Database Languages

4. Classification of DBMS

5. To Summarize
Three-schema Architecture

End users

External level

Conceptual level

Internal level

Database instance

Conceptual schema

view1

view2

(...)

(...)

(...)

viewn

Internal schema
Three-schema Architecture

• *Internal level*
  o Described the **physical storage structure of the database**
  o Uses a physical data model (internal schema)

• *Conceptual level*
  o Describes the **structure of the whole database** for community of users; describes entities, data types, relationships, user operations
  o Uses a representation data model

• *External (view) level*
  o Describes the **part of the database** that a particular user group is interested in; hides the rest from that group
  o Has a number of external schemas
What is the goal with the three-schema architecture?
Data Independence

• **Data Independence**: The capacity to *change the schema* at one level of a database system *without* having to *change at the next higher level*
Data Independence

• **Logical Independence:** The capacity to *change the conceptual schema* without having to *change the external schema*
  
  - Example: Add another field to a conceptual schema

• **Physical Independence:** The capacity to *change the internal schema* without having to *change the conceptual schema*
  
  - Example: Put an access path(index) to reorganize the files
Outline

1. Data Model, Database Schema
   • Categories of data models
   • Database schemas and instances

2. Data Independence
   • Three-schema architecture
   • Logical independence
   • Physical independence

3. Database Languages

4. Classification of DBMS

5. To Summarize
Database Languages

- **Database Language**: Generic term for a class of languages used for defining, communicating with or manipulating a database.
- **Storage Definition Language (SDL)** – to specify internal schema
- **Data Definition Language (DDL)** – to specify conceptual schema
- **View Definition Language (VDL)** – to specify user views
- **Data Manipulation Language (DML)** – to manipulate the data: insert, delete, update)
Database Languages

• In the DDL the database administrator define the internal and conceptual schema and in this manner the database is designed. Subsequent modifications in the schema design is also made in DDL.

• The DML used by DB users and application programs retrieve, add, remove, or alter the information in the database.

• SQL represents a combination of DDL, VDL and DML
Outline

1. Data Model, Database Schema
   • Categories of data models
   • Database schemas and instances

2. Data Independence
   • Three-schema architecture
   • Logical independence
   • Physical independence

3. Database Languages

4. Classification of DBMS

5. To Summarize
Classification of DBMS

• Criteria *data model*:
  - Relational DBMS
  - Object-oriented DBMS
  - Object-relational DBMS
  - Graph DBMS

• Criteria *number of sites on which database is distributed*:
  - Centralized DBMS
  - Distributed DBMS
Classification of DBMS

• Criteria *number of users*:
  o Single-user DBMS
  o Multi-user DBMS

• Criteria *purpose*:
  o General purpose DBMS
  o Special-purpose DBMS
Outline

1. Data Model, Database Schema
   • Categories of data models
   • Database schemas and instances

2. Data Independence
   • Three-schema architecture
   • Logical independence
   • Physical independence

3. Database Languages

4. Classification of DBMS

5. To Summarize
Summary

• Data model
• Database schema, database instance
• Data independence
Components of a DBMS (Figure 2.3 Elmasri/Navathe)