Database Systems



Thomas Neumann

Introduction

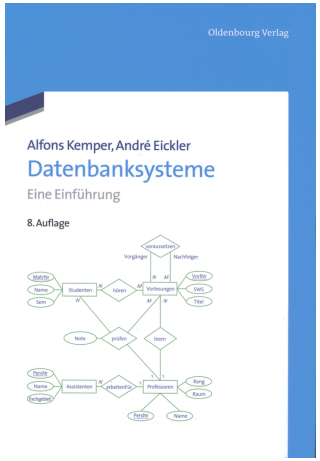
Textbook

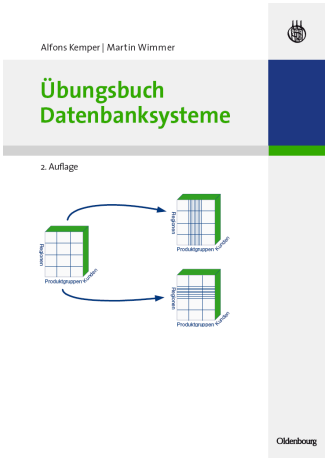
Alfons Kemper and Andre Eickler
Datenbanksysteme – Eine Einführung
10th edition
Oldenbourg Verlag, Munich
(approx. 50 euros)

http://www-db.in.tum.de/research/publications/books/DBMSeinf

http://www-db.in.tum.de

Textbook


Oldenbourg Verlag
Alfons Kemper, André Eickler
Datenbanksysteme
Eine Einführung
8. Auflage


Alfons Kemper | Martin Wimmer
Übungsbuch
Datenbanksysteme
2. Auflage
Oldenbourg

Literature: Alternative and Further Reading

- **A. Kemper, A. Eickler: Datenbanksysteme – Eine Einführung. Oldenbourg Verlag, 2015. 10th edition.**
- **A. Kemper, M. Wimmer: Übungsbuch Datenbanksysteme. Oldenbourg Verlag, 3rd edition, 2011.**
- A. Silberschatz, H. F. Korth and S. Sudarshan: Database System Concepts, 4th edition, McGraw-Hill Book Co., 2002.
- R. Elmasri, S.B. Navathe: Fundamentals of Database Systems, Benjamin Cummings, Redwood City, Ca, USA, 2nd edition, 1994
- R. Ramakrishnan, J. Gehrke: Database Management Systems, 3rd edition, 2003.
- G. Vossen: Datenmodelle, Datenbanksprachen und Datenbank-Management-Systeme. Oldenbourg, 2001.

- D. Maier: The Theory of Relational Databases. Computer Science Press. 1983.
- S. M. Lang, P.C. Lockemann: Datenbankeinsatz. Springer Verlag, 1995.
- C. Batini, S. Ceri, S.B. Navathe: Conceptual Database Design, Benjamin Cummings, Redwood City, Ca, USA, 1992.
- C. J. Date: An Introduction to Database Systems. McGraw-Hill, 8th edition, 2003.
- J.D. Ullman, J. Widom: A First Course in Database Systems, McGraw Hill, 2nd edition, 2001.

- A. Kemper, G. Moerkotte: Object-Oriented Database Management: Applications in Engineering and Computer Science, Prentice Hall, 1994
- E. Rahm: Mehrrechner-Datenbanksysteme. Addison-Wesley, 1994.
- P. Dadam: Verteilte Datenbanken und Client/Server Systeme. Springer Verlag, 1996
- G. Weikum, G. Vossen: Transactional Information Systems: Theory, Algorithms, and the Practice of Concurrency Control. Morgan Kaufmann, 2001.
- T. Härder, E. Rahm: Datenbanksysteme – Konzepte und Techniken der Implementierung, 2001.

Introduction

- What is a database system (DBS)?

  A system for storing and managing data.

- Why not use a conventional file system?

  Fault tolerance and scalability are achievable only with great effort.

Examples

Traditional application areas:

- Business data
- Accounting
- Administration
- ...

Today much broader:

- Scientific/medical data
- Data Mining
- Information integration
- Web search
- ...

Examples(2)

Indirectly, we use databases all the time:

- Web search on Google, Yahoo, ...
- Queries on Amazon, EBay, ...
- Backend of many large websites

Many variants (DB/IR, centralized/decentralized, etc.)

Databases are almost always used when

- the data volumes are large
- the data is valuable

Examples(3)

The major commercial database systems:

- Oracle
- IBM DB2
- Microsoft SQL Server
- Amazon Aurora
- Snowflake

Some free database systems:

- PostgreSQL
- MySQL
- DuckDB

Many more, some highly specialized systems.

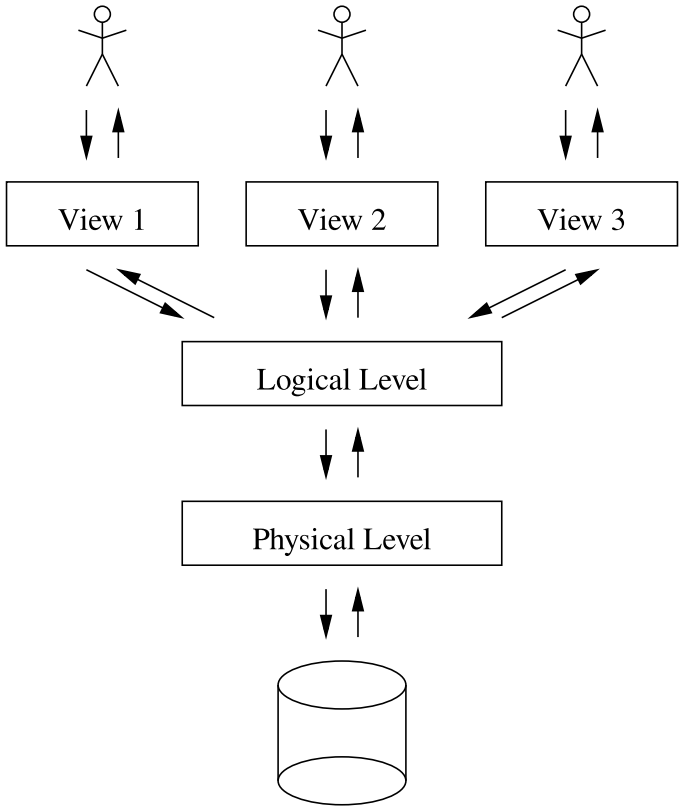
Typical Problems without a DBS

- Redundancy and inconsistency
- Different data formats
- Problems with multi-user operation
- Loss of data
- Integrity violations
- Security problems
- High development costs for application programs

Reasons for Using a DBS

- Data independence
- Declarative query languages
- Concurrency control
- Error handling (recovery)
- Ensuring data integrity
- Efficiency and scalability

Data Independence



Data Independence(2)

- View: describes how a user sees the data
- Logical level: describes how the data is structured
- Physical level: describes how the data is stored

Data Independence(3)

- A DBS decouples applications from the actual structure and storage of the data
- Logical data independence
  - Changes at the logical level have no effect on applications
- Physical data independence
  - Changes at the physical level have no effect on applications
  - Enforced in almost all modern DBSs

Declarative Query Language

- The user tells the DBS *what* data should be retrieved …
- …and not *how* the data should be retrieved
- Less error-prone (when formulating queries/developing applications), since no knowledge of the lower layers of the DBS is needed

Concurrency Control

- If multiple users can modify data concurrently without any control, there are big problems
- A DBS allows concurrent access and prevents harmful side effects

Error Handling

- A DBS can reconstruct the state at the time of a crash
- For this, log files are created and maintained by the DBS

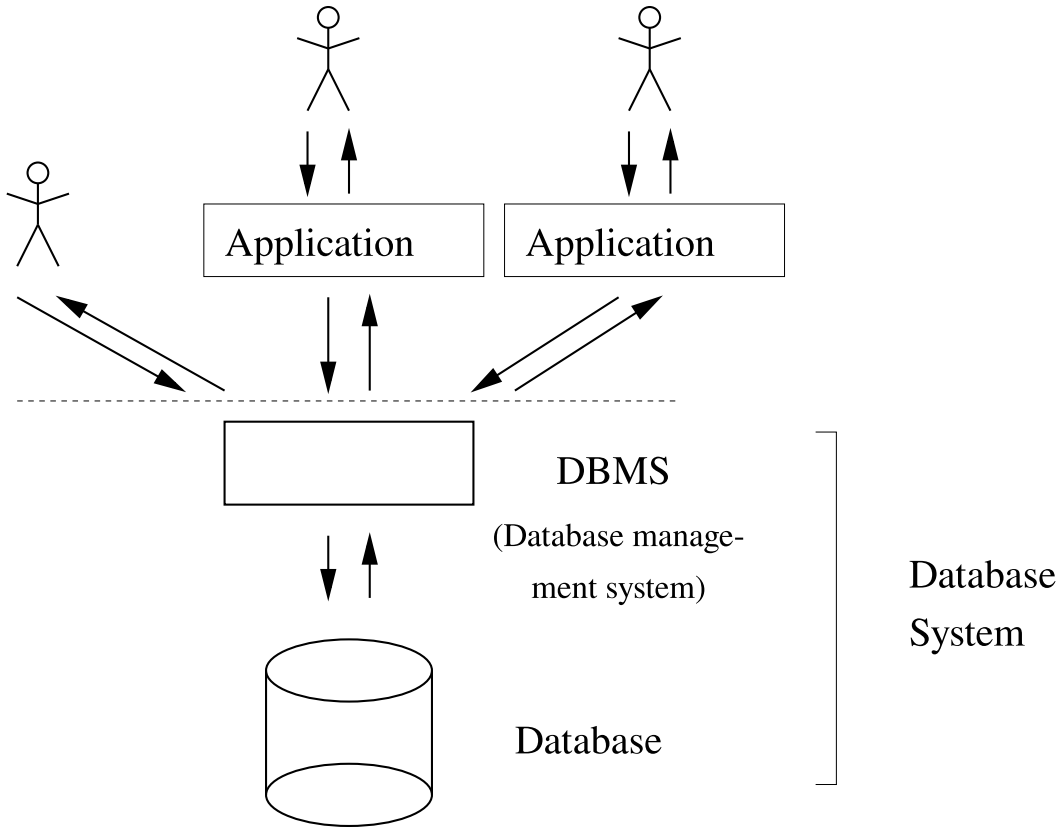
Data Integrity

- Data processing in an application does not happen completely at random, but follows certain principles
- A DBS enforces (specified) principles and thus protects against:
  - user errors
  - program errors

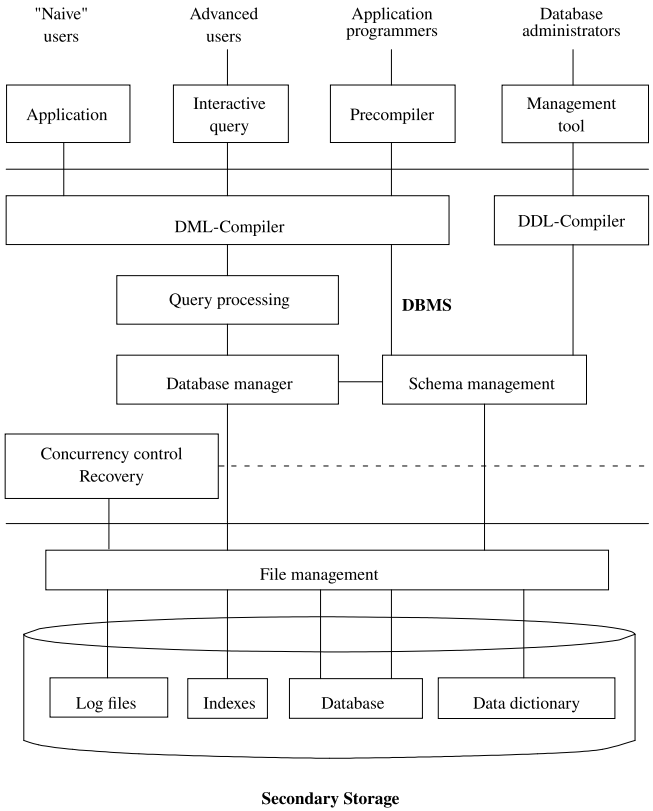
Efficiency and Scalability

- DBSs are designed for large-scale applications
- DBSs integrate techniques that can handle large data volumes

Coarse Architecture



More Detailed Overview



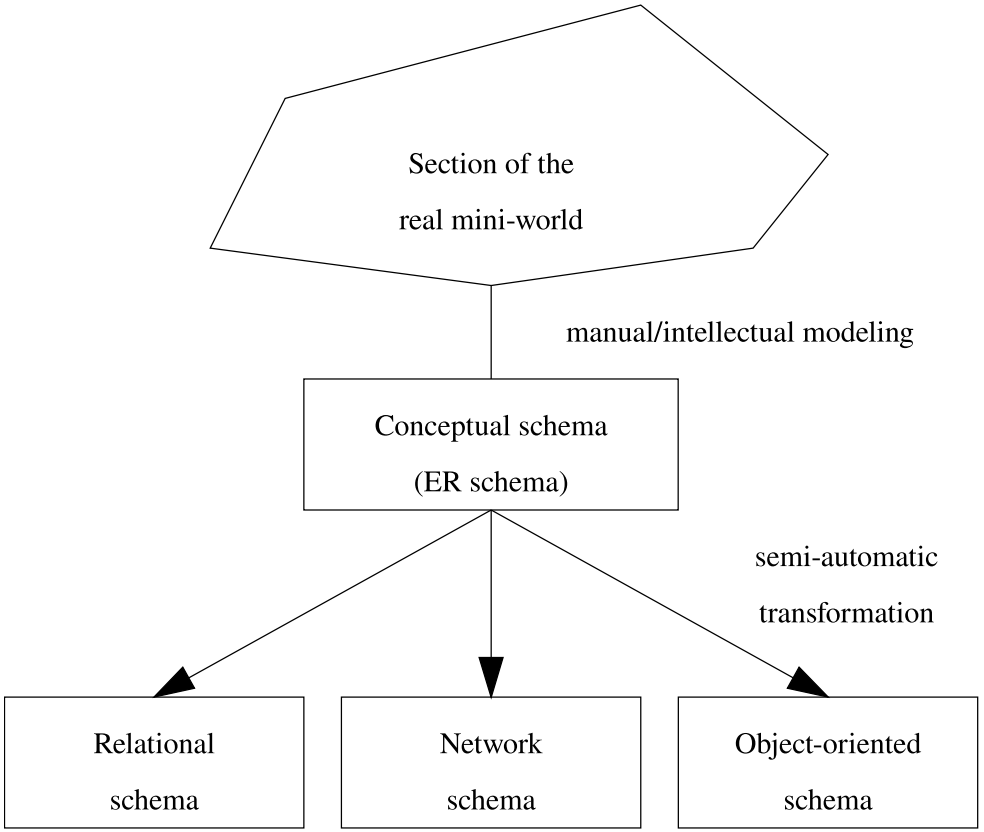
- Even though a DBS can do a lot, it cannot do everything!
- A user still has to specify the

  requirements of an application …
- …and the kind of data that is to be stored
- Two important concepts in design:
  - Data model: defines which constructs exist for describing the data
  - Schema: a concrete description of a particular data collection (using a data model)

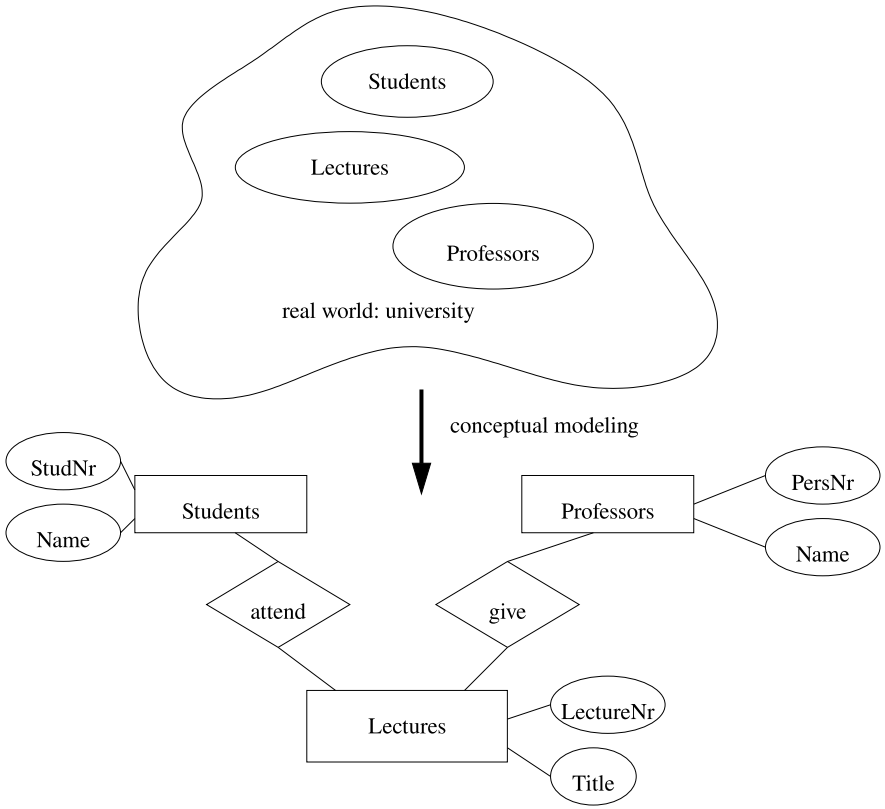
Data Models

- Conceptual models:
  - Entity-relationship model (ER model)
  - Unified Modeling Language (UML)
- Logical models:
  - Hierarchical model
  - Network model
  - Relational model
  - Object-oriented model
  - Object-relational model
  - Semistructured model
  - Graph-structured model

Modeling Steps



Example


Students
Lectures
Professors
real world: university
conceptual modeling
StudNr
Name
Students
attend
Lectures
give
Professors
PersNr
Name
LectureNr
Title

Example(2)

Relational Schema

Students	
StudNr	Name
26120	Fichte
25403	Jonas
...	...

attend	
StudNr	LectureNr
25403	5022
26120	5001
...	...

Lectures	
LectureNr	Title
5001	Fundamentals
5022	Faith and Knowledge
...	...

Summary

- Many applications have similar requirements regarding data management
- A DBS forms a common basis and provides an infrastructure that supports these applications
- For an overview of what is offered, see slide “Reasons for Using a DBS”

Lecture Overview

- We try to cover two sides:
  - How do you use a DBMS?
  - What happens behind the scenes?

Lecture Overview(2)

- Contents:
  - Database design (Chapter 2)
  - Relational model (Chapter 3)
  - SQL (Chapter 4)
  - Relational design theory (Chapter 5)
  - Physical data organization (Chapter 6)
  - Query processing (Chapter 7)
  - Transaction management (Chapter 8)
  - Recovery (Chapter 9)
  - Concurrency control (Chapter 10)
  - Distributed database systems (Chapter 11)