Introduction to Database Systems
(Einführung in Datenbanksysteme)

40 hours course, 30 hours practice, 7 credits

Prof. Dr. H. Schweppe
FU Berlin, SS 2010
hs@inf.fu-berlin.de

1 Introduction

1.1 Databases vs. files
1.2 Basic concepts and terminology
1.3 Brief history of databases
1.4 Architectures & systems
1.5 Technical Challenges
1.6 DB lifecycle

2 Conceptual Database Design

2.1 Requirement analysis
2.2 Modeling languages
Overview
Requirement Analysis (simple case study)
Basic Modeling Primitives
Modeling Languages: UML and Entity-Relationship Model (ERM)
Conceptual DB design: basics
From Requirements to Models
2.3 Integrity Constraints
Constraint types
Cardinality constraints
Weak entities
2.4 Modeling patterns
Modeling historical data
N-ary relationships
Generalization / specialization, recursive relations, aggregation

3 Relational Model: Logical Design using the Relational Data Model / Schema Definition

3.1 Logical Schema Design
The Relational Data Model – Basics
Keys, candidate keys and more
3.2 From Conceptual to Logical Schema: Mapping ER to RDM
Relationships to relations: a simple step
Mapping weak entities and multivalued attributes
Consolidation

3.2 SQL/DDL – first steps

Basis Schema Definition using SQL / DDL
SQL Data types, domains, user defined types
Creating simple tables

3.4 SQL/DDL – Constraints

Attribute and simple table constraints
Enforcing cardinality constraints and foreign keys
Deferred constraints
Assertions and triggers
Metadata management
Modifying and deleting definitions and more…

4 Normalization: - Quality of relational designs

4.1 Functional Dependencies

Design quality
Update anomalies
Functional Dependencies: definition
. Properties of Functional Dependencies

4.2 Normal forms

Informal introduction
Normal Forms and FDs
Normal forms (2NF, 3NF, BCNF,)
Lossless join and dependency preservation

4.4 Normal Forms: Critical review

5 Algebraic operations on tabular data

5.1 Basic idea of relational languages
5.2 Relational Algebra operations
3.3 Relational Algebra: Syntax and Semantics
5.4. More Operators
5.5 Special Topics of RA

Relational algebra operators in SQL,
relational completeness

Ch. 6 Predicate logic based query languages
(skipped, part of slide set, not relevant for exams)

7 SQL – Data Handling

7.1 Update, Deletion, Insertion and bulk load*
7.2 The query language SQL

Search predicates
Arithmetic expressions and functions in predicates
Different kinds of join
Output improvement

7.3 Advanced SQL

Subselects and Correlated subqueries
Quantified expressions, SOME, ANY
Grouping and Aggregation
Views

8 Views, PL/SQL / Triggers, Functions

8.1 Views and view updates
8.2 PL/SQL, stored procedures
8.3 User defined Functions
8.5 Triggers

9 Embedding SQL in Programming languages

9.1 Introduction: using SQL from programs
9.2 Embedded SQL

Static and dynamic embedding
Cursors
ESQL / C

9.3 Application programs and transaction
9.4 SQL and Java

JDBC, SQLJ

9.5 OR mapping and components

10 Physical schema design

10.1 Introduction
10.2 Technology

Disk technology
RAID

10.3 Index structures in DBS

Indexing concept
Primary and Secondary indexes
Types of indexes and index definition in SQL
Implementing indexes: search trees
- ISAM
- B+-tree

Height of B+-trees
Criteria for indexing

11 Transactions: models

11.1 Transactions in application programs

Definition
Isolation levels

11.2 Concepts: ACID properties

11.3 Modeling transactions: histories and schedules

Correctness criteria
Serial execution
History

11.4 Serializability

Conflict graph
Serializability theorem

12 Concurrency control

16.1 Serializability and Concurrency Control

16.2 Locking

Lock protocols
Two phase locking
Strict transactional protocols
Lock conflicts and deadlocks
Lock modes
Deadlock detection, resolution
Model for conflict / deadlock frequency

16.3 Nonlocking concurrency control

Optimistic CC

16.4 Multiversion CC

Read-only transactions (MVCC)
Snapshot isolation: lock based / first committer wins
2 Version 2PLMVCC

13 Principles of recovery

13.1 Failsafe system

13.2 Undo / redo recovery

13.3 Commit rule / WAL principle

13.4 Logging

14 Data Warehouses in a nutshell

14.1 Introduction OLTP vs. OLAP

14.2 DWH methodology

14.3 Stars and Stripes

14.4 OLAP operators: Roll up and Drill down, SQL operators ROLLUP, CUBE

14.5 ROLAP and MOLAP ... and more

15 Data mining – an overview

15.1 Motivation and goals

15.2 Association rules: how to find them –
support, confidence
Frequent item sets and A priori algorithm

16 A short introduction to managing unstructured data: Information Retrieval (not in course)

16.1 The general model
16.2 Similarity
16.3 Boolean Model
16.4 Vector space Model
Zipf empiricism
tf / idf ,Cosinus measure
16.5 Implementation issues
16.6 Evaluation of Retrieval effectiveness
16.7 Page Rank basics

17 What next?

17.1 Trends in Data management
17.3 What next in research?
17.4 What next in teaching?

For slides (pdf), book recommendations etc see http://w3.inf.fu-berlin.de/lehre/SS10/DBS-Intro/