

# Seminar

# Query Processing in

# Database Systems

Winter Term 2009/2010

Lehrgebiet Informationssysteme

Andreas M. Weiner  
weiner@informatik.uni-kl.de



**AG Datenbanken und  
Informationssysteme**



**AG Heterogene  
Informationssysteme**

## Part 1: Introduction

- [Topic 1: Introduction](#) (Motivation, Introduction to Query Optimization, Query Rewrite, Cost Estimation, Cost Models, Optimization Goals)

## Part 2: Classical Query Optimization

- [Topic 2: Extensible and Rule-Based Query Optimizers](#) (Comparison of various optimization frameworks and strategies)
- [Topic 3: Selectivity Estimation and Cost Models](#) (Strengths, limitations, and weaknesses of cost estimation methods, histograms, cost modelling)
- [Topic 4: Search Strategies](#) (Comparison of various algorithms for search-space exploration)

## Part 3: Distributed Query Optimization

- [Topic 5: Client-Server Database Systems](#) (Challenges, data shipping vs. query shipping, place and time of optimization)
- [Topic 6: Heterogeneous Database Systems](#) (Semantic heterogeneity, Wrappers, Garlic, APIs)

## Part 4: Future Developments

- [Topic 7: In-Memory Query Processing](#) (Concepts, IO behaviour, scalability, algorithms, cost models)
- [Topic 8: Query Processing in XML Database Systems](#) (Challenges, differences to classical query optimization, operators, cost estimation, cost models)
- [Topic 9: Stream-Based Query Processing](#) (Concepts, algorithms, cost models, operators)
- [Topic 10: Learning Optimizers](#) (Adaptivity, Feedback-learning)

- Self-employed familiarization with a scientific topic
- Finding scientific literature on web or in the library (DBLP, Citeseer, Google)
- Written composition, presentation, and discussion
- Correct citations
- No copy & paste!
- Time Management

## Searching for Literature

- Some basic literature is provided by your supervisor
- You have to search for further literature
- Collect a list of related literature and send it to your supervisor

## Annotated Table of Contents (TOC)

- Prepare an annotated TOC for your supervisor (approx. 1.5 pages)

## Written Composition

- LNCS Layout
- PDF Format required
- Length: 6000–8000 Words (net.) $\cong$  15–20 Pages
- Correct and complete bibliography
- Deadline: two weeks before the presentations start

- Length: 60 Minutes (45 minutes for the presentation, 15 minutes for the discussion)
- Presentation:
  - Beamer (private or one of our notebooks)
  - Overhead foils
- You must submit your presentation electronically, one week after you gave your talk
- Two schedules expected
  - Friday, February 5<sup>th</sup>, 2010
  - Friday, February 12<sup>th</sup>, 2010
  - Room 36/336, 1:30 p.m.

- **“Unbenoteter Schein”:**
  - Meet the deadlines!
  - Decent presentation
  - Be present when others give their talks

- **“Benoteter Schein”**

- See above

## **Criteria for grading:**

- Quality of your written composition
- Quality of your presentation (including your foils)
- Discussion
- Meeting the deadlines
- Overall impression of your supervisor
- ...

- Monday, October 26<sup>th</sup>, 2009: Kick-off meeting
- Monday, November 9<sup>th</sup>, 2009: Deadline for literature list
- Monday, November 30<sup>th</sup>, 2009: Deadline annotated TOC
- Friday, January 15<sup>th</sup>, 2010: Deadline written composition (final version)
- Friday, January 22<sup>nd</sup>, 2010: Deadline corrected written composition
- Friday, February 5<sup>th</sup>, 2010: First presentation

**All deadlines are strict!**

## Part 1: Introduction

- [Topic 1: Introduction](#) (Motivation, Introduction to Query Optimization, Query Rewrite, Cost Estimation, Cost Models, Optimization Goals)

## Part 2: Classical Query Optimization

- [Topic 2: Extensible and Rule-Based Query Optimizers](#) (Comparison of various optimization frameworks and strategies)
- [Topic 3: Selectivity Estimation and Cost Models](#) (Strengths, limitations, and weaknesses of cost estimation methods, histograms, cost modelling)
- [Topic 4: Search Strategies](#) (Comparison of various algorithms for search-space exploration)

## Part 3: Distributed Query Optimization

- [Topic 5: Client-Server Database Systems](#) (Challenges, data shipping vs. query shipping, place and time of optimization)
- [Topic 6: Heterogeneous Database Systems](#) (Semantic heterogeneity, Wrappers, Garlic, APIs)

## Part 4: Future Developments

- [Topic 7: In-Memory Query Processing](#) (Concepts, IO behaviour, scalability, algorithms, cost models)
- [Topic 8: Query Processing in XML Database Systems](#) (Challenges, differences to classical query optimization, operators, cost estimation, cost models)
- [Topic 9: Stream-Based Query Processing](#) (Concepts, algorithms, cost models, operators)
- [Topic 10: Learning Optimizers](#) (Adaptivity, Feedback-learning)



## Homepage

- <http://www.lgis.informatik.uni-kl.de/cms/courses/seminar/>

## LNCS

- <http://www.springer.com/computer/lncs?SGWID=0-164-2-72376-0>

## DBLP

- <http://www.informatik.uni-trier.de/~ley/db/>

## Typography

- <http://www.zvisionwelt.de/typokurz.pdf>

# Thank your for your attention!

Questions