



# Proseminar

## “Wissenschaftliches Arbeiten”

Ulf Leser

# Proseminar

- We want to teach you how
  - to approach a **scientific topic**
  - to find **scientific** literature and discern relevant from irrelevant
  - to systematically write a **scientific** seminar thesis
  - to give good **scientific** presentations
- Topics are
  - Problems solved by rather simple **algorithms**
  - Problems in graphs (with simple algorithms)
  - Problems from data management (with simple algorithms)
- **Text-Book** knowledge is usually a good start, but never enough

# Who should be here

- Bachelor Informatik
- Basic skills in programming and theoretical computer science
- Algorithms and data structures
- Ability to read [English papers](#) / books / web pages

# How it will work 4 Seiten Sem Thesis

19.04.2016	Prof	Overview; Selection of topics
26.04.2016	Prof	Presentation: Finding and assessing literature
03.05.2016	Prof	Presentation: Giving scientific presentations
10.05.2016	Students	Submission: Literature list (mail)
17.05.2016	Students	5 min teaser talks
24.05.2016	Students	5 min teaser talks
31.05.2016	Prof	Presentation: Scientific writing
07.06.2016		Submission: 5-page seminar thesis (mail)
14.06.2016	Prof, ~8 studs	Feedback
21.06.2016	Prof, ~8 studs	Feedback
28.06.2016		Free
05.07.2016	Students	15-min seminar talks (with course assessment)
12.07.2016	Students	15-min seminar talks (with course assessment)
19.07.2016	Students	15-min seminar talks (with course assessment)
30.08.2016	Students	Submission: Full, 10-page seminar thesis (mail)

# Literature List

- Your topic essentially will consist of a single phrase
- First task: Create a **literature list**
  - Start from text books giving the basic idea
  - Find **current scientific articles** with extensions, variations, ...
  - Assess their quality and suitability
  - Create a top-10 list
  - Select most important ones and describe how you will use them
- First deliverable: 10.5.2016
  - An abstract of your topic
  - Top-10 articles with short summary
  - Top-3 articles with intended usage

# Teaser Talk: 5 minutes Presentation (17/24.5)

- Present **your topic in 5 minutes**
  - 3 content slides at most
- What is your topic about?
- Why is this topic/problem important? Applications?
- What is the standard way of solving it
- What are extensions – and why do they exist?
- What is **cool about your topic?**

# First Thesis: 4 Pages (7.6)

- Introduction: What is the problem?
  - What is it relevant for?
  - What are the important computational aspects?
  - What will this thesis describe?
- Formal problem statement
  - Express your problem in a formal, CS-based language
- List of solutions you will present with short description
  - Why these? What makes them interesting, what are their mutual differences / properties?
- What you will not include – and why not
- Reference list

# Second Talk: 15 Minutes (5/12/19.7)

- Final presentation
- 15 min + 2 min discussion
  
- Formal problem statement
- A running example
- Overview of solutions
- **One interesting solution** in more detail
- Applications
- Comparison of approaches
- Discussion: What is missing, open, not yet researched?



# Final Thesis: 10 Pages (30.8)

- Abstract
- Introduction
- Examples
- Applications
- Background and problem statement
- Most important solutions
- Comparison (qualitative, quantitative)
- Summary

- Questions?

# Topics

Topic	Assigned to
Travel planning in maps	T. Kieseling
Centrality in (social) networks	P. Holy
The edit distance of two strings	D. Zyla
Vector Space Model for Information Retrieval	Tran Gia
Suffix Trees and Suffix Arrays	R. Möller
Adaptive radix trees (*)	C. Hagmeister
Hierarchical clustering	B. Stelter
Decision-Tree induction	L. Kohler
Counting frequent subgraphs (*)	D. Hartmann
Cocoo hashing (*)	K. Sabel
Exact substring search using Karp-Rabin	F. Berger
Huffmann codes and compression (*)	L. Kolmetz
GRID files for multidimensional indexing	L. Beurer-Kellner
Bipartites matching using stable marriage (*)	R. Gützkow

# More Topics

Topic	Assigned to
Subgraph isomorphism (*)	
Traveling Salesman Problem	
Tree Edit Distance (*)	
Graph/Network Alignment mit Graphlets	
Mining Frequent Itemsets (association rules)	
Communities in (social) networks	
Set Containment Joins	
Fast Near-Duplicate Detection in Web Archives	

# How it will work

19.04.2016	Prof	Overview; Selection of topics
26.04.2016	Prof	Presentation: Finding and assessing literature
03.05.2016	Prof	Presentation: Giving scientific presentations
10.05.2016	Students	Submission: Literature list (mail)
17.05.2016	Students	5 min teaser talks
24.05.2016	Students	5 min teaser talks
31.05.2016	Prof	Presentation: Scientific writing (with course assessment)
07.06.2016		Submission: 5-page seminar thesis (mail)
14.06.2016	Prof, ~8 studs	Feedback
21.06.2016	Prof, ~8 studs	Feedback
28.06.2016		Free
05.07.2016	Students	15-min seminar talks (with course assessment)
12.07.2016	Students	15-min seminar talks (with course assessment)
19.07.2016	Students	15-min seminar talks (with course assessment)
30.08.2016	Students	Submission: Full, 10-page seminar thesis (mail)

# Literatur



15-20 Euro



4 Euro (!)



15 Euro



25-30 Euro