



Proseminar

“Wissenschaftliches Arbeiten”

Ulf Leser

Proseminar

- We want to teach you how to
 - approach a **scientific topic**
 - find **scientific** literature and discern relevant from irrelevant one
 - systematically write a **scientific** seminar thesis
 - 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 a good start, but not enough

Who should be here

- Bachelor / Kombibachelor / Infomit Informatik
- Good knowledge in algorithms, programming and theoretical computer science (e.g. 4th semester)
- Ability to read [English papers](#) / books / web pages

How it will work - Overview

Termine im Überblick

27.04.2018, 11-13 Uhr	Leser	Themenvorstellung, Seminaridee, Themenvergabe
04.05.2018, 11-13 Uhr	Leser	Wissenschaftliches Recherchieren
11.05.2018	Alle Studierende (per Mail)	Abgabe Literaturliste (Feedback per Mail): <ul style="list-style-type: none">■ Kurze Zusammenfassung des Themas (ca. 20 Zeilen)■ Top-10 relevanteste Artikel (komplette Referenzen, 3-zeilige Inhaltsangabe)■ Zu den Top-3 jeweils eine kurze Zusammenfassung (ca. 10 Zeilen): Was steht drin, warum ist es für ihre Arbeit wichtig?
18.05.2018, 11-13 Uhr	Leser	Wissenschaftliche Vorträge halten
25.05.2018, 11-13 Uhr	Alle Studierende	5-Minuten "Teaser Talks" (Feedback nach dem Seminar)
01.06.2018, 11-13 Uhr	Leser	Wissenschaftliches Schreiben
15.06.2018	Alle Studierende (per Mail)	Abgabe 4-Seiten Fassung der Seminararbeit
29.06.2018, 11-13 Uhr	Leser und Studierende	Feedbackrunde 4-Seiten Arbeiten
13.07.2018, 11-13 Uhr	Ca 9 Studierende	15-Minuten Seminarvorträge (mit anonymer Peer-Bewertung)
20.07.2018, 11-13 Uhr	Ca 9 Studierende	15-Minuten Seminarvorträge (mit anonymer Peer-Bewertung)
30.08.2018	Alle Studierende	Abgabe 10-seitige Seminararbeit

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
 - Select most important ones and describe how you will use them
 - Due: 11.5.2018 (mail)

First Deliverable: Abstract and literature list

- An **abstract** of your topic (~20 lines)
 - Draft of the abstract of your thesis work. Say what your thesis will be about, why this is exciting, whether there are particular challenges, and what solutions exist
- Top-10 articles with short summary per article (~5 lines)
- Top-3 articles with longer summary; why did you chose them; what are you going to use them for (~10 lines)
 - Thus: 3 articles with ~10 lines description, 7 with ~5 lines
- Submit as one PDF
- Give **complete references** (details later)

Schlechtes Abstract

Bei „Mining Frequent Itemsets“ handelt es sich um eine Warenkorb Analyse, mit der das Einkaufsverhalten von Verbrauchern analysiert und ausgewertet werden kann. Mit Hilfe des Warenkorb Inhalts kann herausgefunden werden, welche Produkte zusammen eingekauft werden und wo man diese im Markt platzieren soll, damit sie öfters verkauft werden. Das Verfahren wird auch **auf Webseiten**, Mit Hilfe sogenannter wenn-dann Regeln können die einzelnen Warenkörbe beschrieben werden. Diese Regeln nennt man „association rules“ und bedeuten, wenn ein Warenkorb mit Produkten gekauft wurde, dann ist x ein Produkt davon. Daraus kann man schlussfolgern, dass **wenn ein ein andere Warenkorb die gleichen Produkte (bis auf x) beinhaltet ist es sehr wahrscheinlich dass auch x** noch in den Warenkorb gelegt wird. Problem ist, dass es riesige Datenmengen an Warenkörben mit unterschiedlichem Inhalt (sogenannte „item sets“) gibt, für die gute und **effiziente** Algorithmen benötigt werden. Deshalb werde ich auf Apriori-Algorithmus und FP-Tress eingehen.

Schlechtes Abstract

Nein. Das ist nur eine Anwendung. MFIS heisst einfach, in einer Menge von Mengen solche Mengen von Elementen zu finden, die häufig zusammen in den Mengen vorkommen.

Bei „Mining Frequent Itemsets“ handelt es sich um eine Warenkorb Analyse, mit der das Einkaufsverhalten von Verbrauchern analysiert und ausgewertet werden kann. Mit Hilfe des Warenkorb Inhalts kann herausgefunden werden, welche Produkte zusammen eingekauft werden und wo man diese im Markt platzieren soll, damit sie öfters verkauft werden. Das Verfahren wird auch auf Webseiten, Mit Hilfe sogenannter wenn-dann Regeln können die einzelnen Warenkörbe beschrieben werden. Diese Regeln nennt man „association rules“ und bedeuten, wenn ein Warenkorb mit Produkten gekauft wurde, dann ist x ein Produkt davon. Daraus kann man schlussfolgern, dass wenn ein ein andere Warenkorb die gleichen Produkte (bis auf x) beinhaltet ist es sehr wahrscheinlich dass auch x noch in den Warenkorb gelegt wird. Problem ist, dass es riesige Datenmengen an Warenkörben mit unterschiedlichem Inhalt (sogenannte „item sets“) gibt, für die gute und effiziente Algorithmen benötigt werden. Deshalb werde ich auf Apriori-Algorithmus und FP-Tress eingehen.

Not very good

Das ist unverständlich. Was ist x ? Es fehlt die Antezedenz und Konsequenz. Ala: Die Regel sei "wenn X , dann Y ". Diese Regel bedeutet: Und wovon ist X ein Produkt?

Bei „Mining Frequent Itemsets“ handelt es sich um eine Warenkorb Analyse, mit der das Einkaufsverhalten von Verbrauchern analysiert und ausgewertet werden kann. Mit Hilfe des Warenkorb Inhalts kann herausgefunden werden, welche Produkte zusammen eingekauft werden und wo man diese im Markt platzieren soll, damit sie öfters verkauft werden. Das Verfahren wird auch auf Webseiten, Mit Hilfe sogenannter wenn-dann Regeln können die einzelnen Warenkörbe beschrieben werden. Diese Regeln nennt man „association rules“ und bedeuten, wenn ein Warenkorb mit Produkten gekauft wurde, dann ist x ein Produkt davon. Daraus kann man schlussfolgern, dass wenn ein ein andere Warenkorb die gleichen Produkte (bis auf x) beinhaltet ist es sehr wahrscheinlich dass auch x noch in den Warenkorb gelegt wird. Problem ist, dass es riesige Datenmengen an Warenkörben mit unterschiedlichem Inhalt (sogenannte „item sets“) gibt, für die gute und effiziente Algorithmen benötigt werden.

Deshalb werde ich auf Apriori-Algorithmus und FP-Tress eingehen.

Not very good

"sehr wahrscheinlich" ist ziemlich ungenau. Eine Regel hat eine Konfidenz und einen Support. Das kann man auch im Abstrakt schon einführen.

Bei „Mining Frequent Itemsets“ handelt es sich um eine Warenkorb Analyse, mit der das Einkaufsverhalten von Verbrauchern analysiert und ausgewertet werden kann. Mit Hilfe des Warenkorb Inhalts kann herausgefunden werden, welche Produkte zusammen eingekauft werden und wo man diese im Markt platzieren soll, damit sie öfters verkauft werden. Das Verfahren wird auch auf Webseiten, Mit Hilfe sogenannter wenn-dann Regeln können die einzelnen Warenkörbe beschrieben werden. Diese Regeln nennt man „association rules“ und bedeuten, wenn ein Warenkorb mit Produkten gekauft wurde, dann ist x ein Produkt davon. Daraus kann man schlussfolgern, dass wenn ein ein andere Warenkorb die gleichen Produkte (bis auf x) beinhaltet ist es sehr wahrscheinlich dass auch x noch in den Warenkorb gelegt wird. Problem ist, dass es riesige Datenmengen an Warenkörben mit unterschiedlichem Inhalt (sogenannte „item sets“) gibt, für die gute und effiziente Algorithmen benötigt werden.

Deshalb werde ich auf Apriori-Algorithmus und FP-Tress eingehen.

Not very good

Da es in der Arbeit vor allem um die Algorithmen geht, sollte auch das Abstract vor allem Algorithmen vorstellen. Sie reden aber fast nur über Anwendungen - irreführend.

Bei „Mining Frequent Itemsets“ handelt es sich um eine Warenkorb Analyse, mit der das Einkaufsverhalten von Verbrauchern analysiert und ausgewertet werden kann. Mit Hilfe des Warenkorb Inhalts kann herausgefunden werden, welche Produkte zusammen eingekauft werden und wo man diese im Markt platzieren soll, damit sie öfters verkauft werden. Das Verfahren wird auch auf Webseiten, Mit Hilfe sogenannter wenn-dann Regeln können die einzelnen Warenkörbe beschrieben werden. Diese Regeln nennt man „association rules“ und bedeuten, wenn ein Warenkorb mit Produkten gekauft wurde, dann ist x ein Produkt davon. Daraus kann man schlussfolgern, dass wenn ein ein andere Warenkorb die gleichen Produkte (bis auf x) beinhaltet ist es sehr wahrscheinlich dass auch x noch in den Warenkorb gelegt wird. Problem ist, dass es riesige Datenmengen an Warenkörben mit unterschiedlichem Inhalt (sogenannte „item sets“) gibt, für die gute und effiziente Algorithmen benötigt werden.

Deshalb werde ich auf Apriori-Algorithmus und FP-Tress eingehen.

Not very good

Ist MFIS, Warenkorbanalyse und Warenkorbinhalt das gleiche? Welches Verhältnis haben Warenkorb und Produkt? Von welchen Märkten sprechen wir hier?

Bei „Mining Frequent Itemsets“ handelt es sich um eine Warenkorb Analyse, mit der das Einkaufsverhalten von Verbrauchern analysiert und ausgewertet werden kann. Mit Hilfe des Warenkorbinhalts kann herausgefunden werden, welche Produkte zusammen eingekauft werden und wo man diese im Markt platzieren soll, damit sie öfters verkauft werden. Das Verfahren wird auch auf Webseiten, Mit Hilfe sogenannter wenn-dann Regeln können die einzelnen Warenkörbe beschrieben werden. Diese Regeln nennt man „association rules“ und bedeuten, wenn ein Warenkorb mit Produkten gekauft wurde, dann ist x ein Produkt davon. Daraus kann man schlussfolgern, dass wenn ein ein andere Warenkorb die gleichen Produkte (bis auf x) beinhaltet ist es sehr wahrscheinlich dass auch x noch in den Warenkorb gelegt wird. Problem ist, dass es riesige Datenmengen an Warenkörben mit unterschiedlichem Inhalt (sogenannte „item sets“) gibt, für die gute und effiziente Algorithmen benötigt werden.

Deshalb werde ich auf Apriori-Algorithmus und FP-Tress eingehen.

Gutes Abstract

Begriff erklären, kommt zu überraschend und ist total überladen.

Die Erforschung der Funktion von Proteinen hat in den Lebenswissenschaften eine sehr hohe Relevanz - die durchaus chaotisch anmutenden, komplexen Wechselwirkungen zwischen Proteinen sind die Grundlage zum Verständnis vieler biologischer Prozesse wie beispielsweise Krankheiten. Dabei lassen sich Protein-Protein-Interaktionen (PPI) als funktionales Netzwerk begreifen, in welchem Proteine die Knoten und Interaktionen die Kanten darstellen. Für ein noch unerforschtes PPI-Netzwerk lassen sich anhand von Ähnlichkeiten zu einem bereits erforschten Voraussagen über Modul- und Proteinfunktionen treffen. Für die Bioinformatik stellt sich hier ein graphentheoretisches Problem, das Network Alignment genannt wird. Mit der Einführung sogenannter Graphlets haben in der Bioinformatik neue Metriken Einzug gehalten, die es ermöglichen, Netzwerke bzw. Graphen auf Ähnlichkeit zu testen, indem herkömmliche Graphenkonzepte wie Knotengrad und Gradverteilung über dem Graphen zu Graphletgrad und deren Verteilung verallgemeinert sowie Häufigkeiten von Graphlettypen in zwei Graphen verglichen werden. Ich werde in meiner Arbeit zeigen, was beim Network Alignment beachtet werden muss, was Graphlets sind und wie die Metrik graphlet degree signature dazu verwendet wird, das Network Alignment Problem zu lösen. Dazu wird der noch sehr junge Algorithmus LGRAAL vorgestellt, charakterisiert und kurz mit anderen Algorithmen verglichen. Abschließend werde ich seine Effizienz und Komplexität auf Grundlage der vorliegenden Forschungsarbeiten unter Abwägung seiner möglichen Schwächen beurteilen.

Gutes Abstract

Eher umdrehen. Die Wechselwirkungen sind Grundlage ... daher hat Funktion hohe Relevanz ...

Die Erforschung der Funktion von Proteinen hat in den Lebenswissenschaften eine sehr hohe Relevanz - die durchaus chaotisch anmutenden, komplexen Wechselwirkungen zwischen Proteinen sind die Grundlage zum Verständnis vieler biologischer Prozesse wie beispielsweise Krankheiten. Dabei lassen sich Protein-Protein-Interaktionen (PPI) als funktionales Netzwerk begreifen, in welchem Proteine die Knoten und Interaktionen die Kanten darstellen. Für ein noch unerforschtes PPI-Netzwerk lassen sich anhand von Ähnlichkeiten zu einem bereits erforschten Voraussagen über Modul- und Proteinfunktionen treffen. Für die Bioinformatik stellt sich hier ein graphentheoretisches Problem, das Network Alignment genannt wird. Mit der Einführung sogenannter Graphlets haben in der Bioinformatik neue Metriken Einzug gehalten, die es ermöglichen, Netzwerke bzw. Graphen auf Ähnlichkeit zu testen, indem herkömmliche Graphenkonzepte wie Kontengrad und Gradverteilung über dem Graphen zu Graphletgrad und deren Verteilung verallgemeinert sowie Häufigkeiten von Graphlettypen in zwei Graphen verglichen werden. Ich werde in meiner Arbeit zeigen, was beim Network Alignment beachtet werden muss, was Graphlets sind und wie die Metrik graphlet degree signature dazu verwendet wird, das Network Alignment Problem zu lösen. Dazu wird der noch sehr junge Algorithmus LGRAAL vorgestellt, charakterisiert und kurz mit anderen Algorithmen verglichen. Abschließend werde ich seine Effizienz und Komplexität auf Grundlage der vorliegenden Forschungsarbeiten unter Abwägung seiner möglichen Schwächen beurteilen.

Gutes Abstract

Logisch ist das redundant. Neue Metriken ermöglichen immer neue Messungen. Man erwartet eher "... die es ermöglichen, Netzwerke .. schneller / genauer ... zu vergleichen". Ein neues Maß ist ja auch kein Selbstzweck.

Die Erforschung der Funktion von Proteinen ist von großer Relevanz - die durchaus chaotisch anmuten. Proteinen sind die Grundlage zum Verständnis vieler biologischer Prozesse wie beispielsweise Krankheiten. Dabei lassen sich Protein-Protein-Interaktionen (PPI) als funktionales Netzwerk begreifen, in welchem Proteine die Knoten und Interaktionen die Kanten darstellen. Für ein noch unerforschtes PPI-Netzwerk lassen sich anhand von Ähnlichkeiten zu einem bereits erforschten Voraussagen über Modul- und Proteinfunktionen treffen. Für die Bioinformatik stellt sich hier ein graphentheoretisches Problem, das Network Alignment genannt wird. Mit der Einführung sogenannter Graphlets haben in der Bioinformatik neue Metriken Einzug gehalten, die es ermöglichen, Netzwerke bzw. Graphen auf Ähnlichkeit zu testen, indem herkömmliche Graphenkonzepte wie Kontengrad und Gradverteilung über dem Graphen zu Graphletgrad und deren Verteilung verallgemeinert sowie Häufigkeiten von Graphlettypen in zwei Graphen verglichen werden. Ich werde in meiner Arbeit zeigen, was beim Network Alignment beachtet werden muss, was Graphlets sind und wie die Metrik graphlet degree signature dazu verwendet wird, das Network Alignment Problem zu lösen. Dazu wird der noch sehr junge Algorithmus LGRAAL vorgestellt, charakterisiert und kurz mit anderen Algorithmen verglichen. Abschließend werde ich seine Effizienz und Komplexität auf Grundlage der vorliegenden Forschungsarbeiten unter Abwägung seiner möglichen Schwächen beurteilen.

Richtig zitieren: Bibliographische Angaben

- Buch: Author, title, year, Verlag
- Journals: Author, title, year, journal, volume, (issue), pages
- Conference: Author, title, year, conference name, location
- Reports: Author, title, year, type of work, institution, number
- Buchkapitel (editiert): Author, title, year, Buch, Verlag, Editoren
- Beispiele
 - Rawald, T., Sips, M., Marwan, N. and Leser, U. (2015). "Massively Parallel Analysis of Similarity Matrices on Heterogeneous Hardware". Int. Workshop on Data (Co-)Processing on Heterogeneous Hardware Brussels, Belgium
 - Rheinländer, A., Heise, A., Hueske, F., Leser, U. and Naumann, F. (2013). "SOFA: An Extensible Logical Optimizer for UDF-heavy Dataflows". CoRR/abs:1311.6335.
 - Rheinländer, A., Heise, Hueske, Leser, and Naumann, (2015). "SOFA: An Extensible Logical Optimizer for UDF-heavy Data Flows " Information Systems 52: 96 - 125.
 - Hakenberg, J., Plake, C. and Leser, U. (2010). Ali Baba: A Text Mining Tool for Complex Biological Systems. In Lodhi, H. and Muggleton, S. (ed): Elements of Computational Systems Biology, Wiley & Sons

Zitation (Kurzbeleg)

- **Drei Autorennachnamen Anfangsbuchstaben**, dann Jahr
 - [RHH+15] Rheinländer, A., Heise, A., Hueske, F., Leser, U. and Naumann, F. (2015). "SOFA: An Extensible Logical Optimizer for UDF-heavy Data Flows " Information Systems 52: 96 - 125.
 - [Sea05] Searls, D. B. (2005). "Data Integration: Challenges for Drug Discovery." Nature Reviews Genetics 4: 45-58.

How it will work - Overview

Termine im Überblick

27.04.2018, 11-13 Uhr	Leser	Themenvorstellung, Seminaridee, Themenvergabe
04.05.2018, 11-13 Uhr	Leser	Wissenschaftliches Recherchieren
11.05.2018	Alle Studierende (per Mail)	Abgabe Literaturliste (Feedback per Mail): <ul style="list-style-type: none"> ■ Kurze Zusammenfassung des Themas (ca. 20 Zeilen) ■ Top-10 relevanteste Artikel (komplette Referenzen, 3-zeilige Inhaltsangabe) ■ Zu den Top-3 jeweils eine kurze Zusammenfassung (ca. 10 Zeilen): Was steht drin, warum ist es für ihre Arbeit wichtig?
18.05.2018, 11-13 Uhr	Leser	Wissenschaftliche Vorträge halten
25.05.2018, 11-13 Uhr	Alle Studierende	5-Minuten "Teaser Talks" (Feedback nach dem Seminar)
01.06.2018, 11-13 Uhr	Leser	Wissenschaftliches Schreiben
15.06.2018	Alle Studierende (per Mail)	Abgabe 4-Seiten Fassung der Seminararbeit
29.06.2018, 11-13 Uhr	Leser und Studierende	Feedbackrunde 4-Seiten Arbeiten
13.07.2018, 11-13 Uhr	Ca 9 Studierende	15-Minuten Seminarvorträge (mit anonymer Peer-Bewertung)
20.07.2018, 11-13 Uhr	Ca 9 Studierende	15-Minuten Seminarvorträge (mit anonymer Peer-Bewertung)
30.08.2018	Alle Studierende	Abgabe 10-seitige Seminararbeit

Deliverable two: 5 minutes Presentation

- On the 25th May
- 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**?

Deliverable 3: First Written Text (4 Pages)

- Due: 15.6.2018
- Introduction: **What is the problem?**
 - What is it relevant for?
 - What are the important **computational aspects**?
 - What will this seminar 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

Details

- Die 4-Seiten Fassung soll die Seminararbeit in Kurzform sein. Sie soll sowohl sehr sorgfältig formuliert als auch korrekt und umfassend im Sinne des Themas sein - aber es fehlen Details, für Beispiele wird kaum Platz sein, etc. Die Gliederung sollte im Grunde die gleiche wie bei der endgültigen Arbeit sein.
- Sinn dieser Arbeit ist es, dass Sie auf 4 Seiten gutes Schreiben und die Erfassung eines Themas üben, ohne vor zu viel Text zurückzuschrecken. Idealerweise ist die 10-Seiten Fassung am Ende einfach die 4-Seitenfassung, ergänzt um Beweise, Messungen, Beispiele und ausführlichere Erklärungen.
- Das Feedback zur 4-Seiten Fassung erfolgt im individuellen Gespräch
- Die 4 Seiten sind inkl. Bibliographie, aber zusätzlich zu Inhaltsverzeichnis und Titelseite

How it will work - Overview

Termine im Überblick

27.04.2018, 11-13 Uhr	Leser	Themenvorstellung, Seminaridee, Themenvergabe
04.05.2018, 11-13 Uhr	Leser	Wissenschaftliches Recherchieren
11.05.2018	Alle Studierende (per Mail)	Abgabe Literaturliste (Feedback per Mail): <ul style="list-style-type: none"> ■ Kurze Zusammenfassung des Themas (ca. 20 Zeilen) ■ Top-10 relevanteste Artikel (komplette Referenzen, 3-zeilige Inhaltsangabe) ■ Zu den Top-3 jeweils eine kurze Zusammenfassung (ca. 10 Zeilen): Was steht drin, warum ist es für ihre Arbeit wichtig?
18.05.2018, 11-13 Uhr	Leser	Wissenschaftliche Vorträge halten
25.05.2018, 11-13 Uhr	Alle Studierende	5-Minuten "Teaser Talks" (Feedback nach dem Seminar)
01.06.2018, 11-13 Uhr	Leser	Wissenschaftliches Schreiben
15.06.2018	Alle Studierende (per Mail)	Abgabe 4-Seiten Fassung der Seminararbeit
29.06.2018, 11-13 Uhr	Leser und Studierende	Feedbackrunde 4-Seiten Arbeiten
13.07.2018, 11-13 Uhr	Ca 9 Studierende	15-Minuten Seminarvorträge (mit anonymer Peer-Bewertung)
20.07.2018, 11-13 Uhr	Ca 9 Studierende	15-Minuten Seminarvorträge (mit anonymer Peer-Bewertung)
30.08.2018	Alle Studierende	Abgabe 10-seitige Seminararbeit

Deliverable 4: 15 Minutes Talk

- On 13.7.18 or 20.7.18 (block)
- 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?
- Do not think that students remember your teaser talk!
- The other students will evaluate your performance

Deliverable 5: Final Thesis (10 Pages)

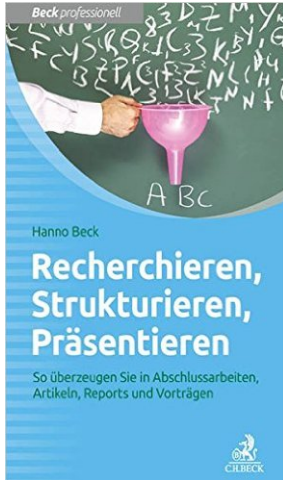
- Abstract
- Introduction
- Examples
- Applications
- Background and problem statement
- Most important solutions
- Comparison (qualitative, quantitative)
- Summary

- Questions?

Topics

Topic	Area	Assigned to
Travel planning in maps	Graph Algs	Vatandas
Centrality in (social) networks	Graph Algs	Görg
Suffix Trees and Suffix Arrays	String Algs	
Ähnlichkeitssuche und M-Trees	Indexing	Chmielewski ****
Decision-Tree induction	ML	Henning
Cocoo hashing (*)	Indexing	Neumann
Exact substring search using Karp-Rabin	String Algs	Spitzner
Huffmann codes and compression (*)	String Algs	Scheidt
GRID files for multidimensional indexing	Indexing	
Bipartites matching using stable marriage (*)	Graph Algs	Tavakoli
Subgraph isomorphism (*)	Graph Algs	
Traveling Salesman Problem	Graph Algs	Mayerhofer
Graph Alignment mit Graphlets	Graph Algs	Lange
Mining Frequent Itemsets (association rules)	Data Mining	Bussenius
Communities in (social) networks (*)	Data Mining	Yilmaz
Set Containment Joins	Database	Jandieri
Fast Near-Duplicate Detection in Web Archives (*)	String Algs	Plachetta
Flussmaximierung in Netzwerken	Graph Algs	Kotsev
Lineare und logische Regression als Classifier	ML	Rabuske
Steiner Bäume und deren Approximierbarkeit	Graph Algs	
Locality Sensitive Hashing (*)	Indexing	Wilke

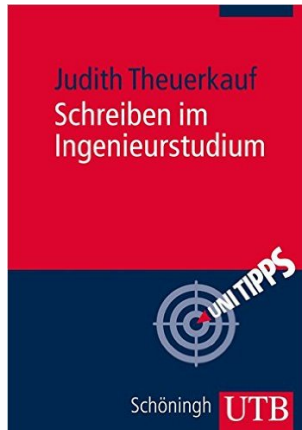
Literatur



15-20 Euro, sehr lebendig, BWL orientiert



4 Euro (!), etwas dröge und konservativ



15 Euro



25-30 Euro