

Projektpraktikum und Seminar Graphen & Algorithmen

SS 2005 & WS 05/06

Prof. Dr. Ulrik Brandes, Martin Hoefler

Lehrstuhl Praktische Informatik

12. April 2005

<http://www.inf.uni-konstanz.de/algo/lehre/ss05/pp>

Allgemeines

- ▶ Themengebiet: Grundlagen der Informatik/Angewandte Informatik
- ▶ Adressaten: Information Engineering im Bachelor-Vertiefungsstudium
- ▶ Vorkenntnisse: "Datenstrukturen & Algorithmen", Programmierkenntnisse
- ▶ Veranstaltung über 2 Semester:
 - ▶ SS 05: Projektpraktikum und Seminar
 - ▶ WS 05/06: Erstellung einer Bachelorarbeit
- ▶ Projektpraktikum: Pro Semester 6 ECTS credits
- ▶ Seminar: 3 ECTS credits

Ablauf Praktikum

1. Teil: Programmierung von grundlegenden Werkzeugen (3-4 Wochen)

Erarbeiten von gemeinsamen Grundlagen

2. Teil: Arbeit an einem eigenen Projekt (Rest des Semesters)

Thema erarbeiten/verstehen

Eigene Ideen

Java-Implementation und Dokumentation (z.B. mit javadoc)

Test/Korrektheit

Präsentation/Vortrag

Ablauf Seminar

Erarbeiten/Verstehen

Eigene Literatursuche

Eigene Ideen

Ausarbeitung (10-20 Seiten)

Vortrag (60 Minuten)

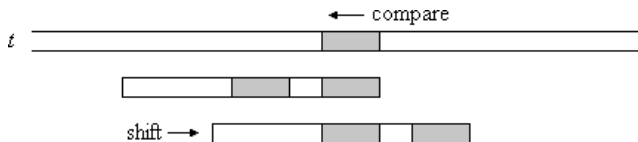
Worum gehts ?

Pattern Matching und Indizierung

Probleme:

Gegeben einen Text T und ein Pattern P ,

- ▶ entscheide, ob P in T vorkommt
- ▶ finde die Anzahl der Vorkommnisse von P in T
- ▶ finde den Anfang aller Vorkommnisse von P in T
- ▶ wobei P evtl. leicht verändert werden darf



On-line Suche

Ein **On-line Algorithmus** bekommt Text und Pattern ohne Zusatzinformationen geliefert, d.h. Preprocessing ist Teil des Algorithmus.

Ein **Real-Time Algorithmus** hat zusätzlich nur einen konstanten Arbeitsaufwand bevor er zum nächsten Zeichen des Textes vorrückt.

Wichtige Einsatzgebiete:

- ▶ Dynamische Inhalte, wenig Anfragen
- ▶ Textverarbeitungssoftware
- ▶ Extrem große Texte (wenig Speicherplatz)
- ▶ Suche bei Datenübertragung

Indizierte Suche

Ein Text darf vorher bearbeitet und ein Index erstellt werden, der es ermöglicht, die Anfragen schneller zu bearbeiten.

Wichtige Einsatzgebiete:

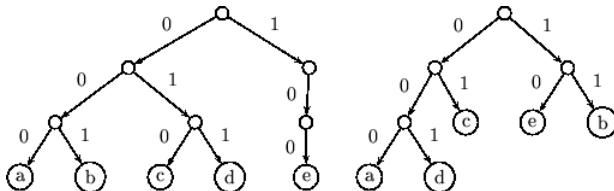
- ▶ Stabile Inhalte, sehr viele Anfragen
- ▶ In der Bioinformatik bei der Gensequenzsuche
- ▶ Suchmaschinen im Internet
- ▶ Große Datenbanken

Worum gehts noch?

Kompression und Codierung von Texten

Speichere einen Text T mit möglichst wenig Bytes, so daß...

- ▶ er dekomprimiert werden kann
- ▶ effizient nach Patterns gesucht werden kann
- ▶ effizient nach veränderten Patterns gesucht werden kann



Zeitplan

14.04.2005: Vorbesprechung

bis 29.04.2005: Entgeltige Themenvergabe

bis 15.05.2005: Grundlegende Werkzeuge implementieren

bis Ende Mai: Beschreibung der Projektimplementierung

Implementierung, Dokumentation, Tests, Ausarbeitung

letzte Semesterwoche Vorträge

Seminare am Semesterende (1 pro Woche)

Termine & Infos

- ▶ Wöchentlicher Termin
- ▶ Jürgens Gruppe: dienstags Nachmittag

- ▶ Verwaltung mit CVS:
CVSROOT = /net/algo/cvs/lehre/ss05/pp/
- ▶ Allgemeine Werkzeuge in \$CVSROOT/java_root/stringtools/
- ▶ Persönliche Implementation in \$CVSROOT/java_root/<username>

- ▶ Anmeldung mit dem Account-Tool
Gruppenname: ppaalgo_S05_W05

Projektpraktikum Graphen & Algorithmen

- ▶ Materialien:
<http://www.inf.uni-konstanz.de/algo/lehre/ss05/pp>
- ▶ Dozenten: Prof. Dr. Ulrik Brandes, Martin Hoefler
- ▶ Martin Hoefler
Raum E 215
Tel.: 07531 - 88 - 4263
hoefler@inf.uni-konstanz.de

Bis zum nächsten Mal...

- ▶ CVS klarmachen, bestehende Klassen auschecken & ausprobieren
- ▶ Informieren über Standard-Suchverfahren in Java (`java.lang.String`, `java.util.regex`)
- ▶ Gemeinsam Klassenstruktur festlegen
- ▶ Ein Verfahren als Klasse implementieren
- ▶ Testen gegen die beiden bisherigen Verfahren
- ▶ Nächstes Mal: Jeder 5-10 Minuten Vortrag über Verfahren & Ergebnisse