

Algorithmen auf Sequenzen

Vorbereitung zur Prüfung

Dominik Kopczynski

Lehrstuhl für Algorithm Engineering (LS11)
Fakultät für Informatik
TU Dortmund

Überblick

- Zusammenfassung aller Themen
- Fragen, die in der Prüfung drankommen können
- Vorbereitung zur Prüfung
- Anmeldung zur Prüfung

Zusammenfassung

Bitsequenzen:

- Bitoperationen
- Popcount
- Rank-Datenstruktur

Prüfungsfragen

Bitsequenzen:

- Wie wird Popcount in $\mathcal{O}(\log(W))$ berechnet?
- Was sind Rank-Datenstrukturen, wie funktionieren sie und wo können sie eingesetzt werden?
- Wie kann der Speicherplatz der Rank-Datenstruktur verbessert werden?

Zusammenfassung

Exakte Online Pattern-Matching Algorithmen:

- NFA / DFA
- Knuth-Morris-Pratt
- Shift-And
- Horspool
- Backward-Nondeterministic DAWG-Matching

Prüfungsfragen

Exakte Online Pattern-Matching Algorithmen:

- Beschreiben Sie den EOPM Algorithmus.
- Führen Sie den EOPM Algorithmus mit gegebenem Beispiel durch.
- Wie kommen Best / Worst / Average Case Laufzeiten zustande?
- Wie hängen DFA und KMP zusammen?
- Was sind die Vorteile von Horspool und BNDM?

Zusammenfassung

Exakte Online Multi-Pattern-Matching Algorithmen:

- Naiv
- Shift-And
- Aho-Corasick

Prüfungsfragen

Exakte Online Multi-Pattern-Matching Algorithmen:

- Beschreiben Sie den EOMPM Algorithmus.
- Führen Sie den EOMPM Algorithmus mit gegebenem Beispiel durch.
- Wie stehen Aho-Corasick und KMP im Zusammenhang?
- Erstellen Sie einen Aho-Corasick zu den Pattern ...

Zusammenfassung

Erweiterte Patternklassen:

- Verallgemeinerte Strings
- Gaps beschränkter Länge
- Optionale Zeichen

Prüfungsfragen

Erweiterte Patternklassen:

- Beschreiben Sie die Erweiterungen des Shift-And für erweiterte Patternklassen.
- Erstellen Sie einen Automaten für VS / GbL / OZ für das Pattern ...
- Welche Laufzeiten haben die Algorithmen.

Zusammenfassung

Volltext-Indizes:

- Suffix-Trie
- Suffix-Tree, Ukkonen-Algorithmus
- Suffix-Array, SAIS
- FM-Index, BWT, Backward Search

Prüfungsfragen

Volltext-Indizes:

- Erklären Sie was ein Suffix-* ist.
- Wie werden die Indizes erstellt.
- Welche Fragestellungen können beantwortet werden.
- Was sind die Tricks um die Laufzeit auf $\mathcal{O}(n)$ zu senken?
- Erstellen Sie eine BWT und invertieren Sie sie.
- Wie kann man den FM-Index effizient abspeichern?

Zusammenfassung

Fehlertolerantes Pattern-Matching:

- DP-Verfahren für Alignment
- LCS, LCF
- Edit-Graph
- Ukkonen-Algorithmus für semiglob. Alignment
- Fehlertoleranter Shift-And
- Fehlertolerante Backward Search

Prüfungsfragen

Fehlertolerantes Pattern-Matching:

- Welche Distanzen gibt es?
- Was ist ein (semi)globales Alignment?
- Wie kann es berechnet werden?
- Welche Tricks nutzt der Ukkonen-Algorithmus?
- Erklären Sie die Funktionsweise des fehlertoleranten Shift-And / Backward Search

Zusammenfassung

Paarweises Sequenzalignment:

- Traceback Verfahren
- Verschiedene Alignment-Varianten
- Affine Gapkosten
- Hirschberg-Algorithmus
- Four-Russians Trick

Prüfungsfragen

Paarweises Sequenzalignment:

- Erstellen Sie ein gl, semigl, feg, lok Alignment für P, T mit Scorewerten S .
- Berechnen Sie das optimale glob Alignment mit affinen Gapkosten.
- Erklären Sie die Funktionsweise des Hirschberg-Algorithmus / Four-Russians Tricks

Vorbereitung zur mündl. Prüfung

- Im Team lernen!!
- Prüfung nachstellen: einer spielt den Prüfer einer den Prüfling und umgekehrt.
- Die gelisteten Fragen sollten im ausreichenden Maße beantwortet werden können.

Anmeldung zur Prüfung

- 1 Termin mit Prüfer festlegen, der Termin kann im Prinzip zu jedem Zeitpunkt sein.
- 2 Anmeldung für mündliche Prüfungen besorgen: Vordrucke gibt es auf der TU-Seite¹ oder auf der Veranstaltungsseite.
- 3 Dokument ausfüllen (Modul-Nr: Inf-BSc-315) und vom Prüfer unterschreiben lassen.
- 4 Anmeldung im LS11-Sekretariat (OH14/R2.39) abgeben.

¹http://www.tu-dortmund.de/uni/studierende/pruefungsangelegenheiten/infopoint_team1/vordrucke_team1/index.html

Ablauf einer Prüfung

- Gewöhnlich sind nur 3 Personen im Prüfungsraum: Prüfling, Prüfer und Protokollant.
- Die Prüfung dauert (zumindest für diese Veranstaltung) 20-25 Minuten.
- Der Prüfer stellt Fragen oder Aufgaben, die der Prüfling dann beantworten bzw. bearbeiten muss.
- Stift und Zettel werden bereit gestellt.
- Der Protokollant hat kein Fragerecht.

Ablauf einer Prüfung

- Nach der Prüfung beraten sich Prüfer und Protokollant unter Ausschluss des Prüflings.
- Die Note wird dann anschließend sofort dem Prüfling mitgeteilt mit evtl. Begründungen.
- Den weiteren Papierkram wird (zumindest für diese Veranstaltung) der Prüfer erledigen.

Tipps während der Prüfung

- Erst denken, dann antworten.
- Ruhig ein paar Sekunden Zeit nehmen um auf die Frage geachtet zu antworten.
- Es ist nicht zielführend wenn man sofort mit dem ersten Gedanken losprescht: *“Der Ukkonenalgorithmus nutzt Suffix-Links ...”*
- Nur beantworten was auch gefragt ist.

Tipps während der Prüfung

- Vor Allem sollte hinreichend gut erklärt werden, wie Laufzeiten und Speicherplatzangaben zustande kommen.
- Bei einer Wissenslücke lieber *“Weiß ich nicht, nächstes Thema ...”* sagen als sich irgendwas aus den Fingern zu saugen. So wird nur unnötig Zeit verschwendet.
- Formeln und Definitionen nicht informell erklären, sondern präzise aufschreiben oder wiedergeben.
- Wenn etwas unklar ist, nachfragen!

In diesem Sinne ...

Frohes Lernen und viel Erfolg bei der Prüfung.