



Blatt 4

Aufgabe 4.1 Im Folgenden sollen die Laufzeiten des *naiven Algorithmus* (Übung 1), des *Z-Algorithmus* (Übung 2) und des *naiven Algorithmus mit bad character rule* (Übung 3) in Abhängigkeit von Alphabetgröße und Musterlänge untersucht werden.

Zur Analyse der Laufzeiten soll jeweils die Anzahl durchgeführter Vergleiche herangezogen werden. Inkrementieren Sie dazu einen Zähler vor oder nach **jedem** erfolgten Zeichenvergleich. Der Grund für diese (und keine möglicherweise effizientere) Zählung ist, dass wir nun die Operation *Zeichenvergleich und Zählerinkrementierung* als eine komplexere Vergleichsoperation betrachten können und somit die Laufzeiten für die verschiedenen Algorithmen in gleichem Maße beeinflusst werden. Außerdem soll die Laufzeit (Realtime oder Usertime, je nach Möglichkeit) gemessen werden.

Auf der Internetseite zur Vorlesung finden sich Texte über einem Alphabet von 40 bzw. 80 Zeichen. Suchen Sie in diesen beiden Texten jeweils nach Mustern der Längen 10, 20, 50, 100, 200, 500, 1000, 2000 und 5000. Erstellen Sie anschließend für beide Alphabetgrößen Diagramme, die die benötigte Anzahl an Vergleichen pro Musterlänge zeigen, und Diagramme, die die Laufzeit pro Musterlänge zeigen, jeweils in Abhängigkeit der Musterlänge. Was lässt sich daraus ersehen? Welcher Algorithmus braucht für jede Musterlänge am wenigsten Vergleiche? Welcher Algorithmus braucht real am wenigsten Zeit? (5 Punkte)

Aufgabe 4.2 Gegeben sei ein String S , in dem das Abschlußzeichen $\$$ nicht vorkommt.

- Warum wird der String S mit einem Abschlußzeichen versehen, wenn der zugehörige Suffixbaum aufgebaut wird? (1 Punkt)
- Zeigen Sie, dass für den Suffixbaum $ST(S\$)$ gilt: Anzahl der inneren Knoten $\leq |S\$|$. (2 Punkte)
- Zeigen Sie, dass für den Suffixbaum $ST(S\$)$ gilt: Anzahl der Kanten $\leq 2|S\$| - 1$. (2 Punkte)

Aufgabe 4.3 Skizzieren Sie die Konstruktion der impliziten Suffixbäume für den String AGGTAGG mit Hilfe des Ukkonen-Algorithmus. (2 Punkte)

Abgabe per E-Mail an grau@informatik.uni-halle.de oder in Raum 4.12.