

Übungsblatt 14

Abgabe für ESE: bis Donnerstag, den 12. Februar um 10:00 Uhr

Abgabe für IEMS: bis Donnerstag, den **19. Februar** um 10:00 Uhr

(IEMS diesmal nur 2 Wochen Bearbeitungszeit, damit die Musterlösungen rechtzeitig zur Klausurvorbereitung online gestellt werden können)

Aufgabe 1 (10 Punkte)

Ergänzen Sie die Klasse *EditDistance* aus der Vorlesung um eine Methode *computeViaTable*. Diese soll die Editierdistanz zwischen zwei Strings x und y in Zeit und Platz $O(|x| \cdot |y|)$ berechnen, nach dem in der Vorlesung erklärten Schema.

Aufgabe 2 (10 Punkte)

Schreiben Sie ein Programm *ApproximateGrepMain*, das für eine gegebene Zeichenkette x alle Zeilen y einer gegebenen Datei ausgibt, für die $ED(|x|, |y|) \leq \max\{|x|, |y|\}/5$. (Das heißt, je länger die längere der beiden Zeichenketten ist, desto mehr Fehler sind erlaubt.)

Finden Sie mit Ihrem Programm die zu jeder der drei folgenden Suchanfragen ähnlichen Zeilen in dem auf der Homepage verlinkten AOL Query Log: *google*, *angelina jolie*, *united airlines*. Tragen Sie für jede dieser drei Anfragen die Anzahl der Treffer und die Gesamtlaufzeit in die auf der Homepage verlinkte Tabelle ein.

Achtung: das AOL Query Log enthält viele Anfragen mehrfach. Zählen Sie aber nur die verschiedenen Treffer. Sie können das auch nutzen um Ihr Programm effizienter zu machen, indem Sie sich Ergebnisse für bereits gesehene Anfragen merken.

Committen Sie wie gehabt Ihren Code in das SVN, in einen neuen Unterordner *uebungsblatt_14*, und ebenso Ihr Feedback in einer Textdatei *erfahrungen.txt*.