
Algorithmen auf Sequenzen

Abgabetermin: Freitag, den 22. Januar, 09⁰⁰ in Moodle

Aufgabe 1

Gib eine möglichst speicherplatzsparende Implementierung eines *stabilen* Bucket-Sorts nach dem ersten Zeichen aller Suffixe von $t\$$ an. Das Feld t für $t\$$ soll dabei als nur-lesbares Feld verwendet werden. Versuche möglichst mit nur zwei weiteren Feldern (von Integers oder Characters) der Länge höchstens $|t\$|$ auszukommen. Die Implementierung soll als Pseudo-Code angegeben werden.

Eine Sortierung nach dem ersten Zeichen heißt *stabil*, wenn für die gegebene Eingabefolge $(1, \dots, n+1) \hat{=} (t^1, \dots, t^{|t\$|})$ mit $t^i := t_i \cdots t_{|t\$|} \in \Sigma^+$ nach der Sortierung nach dem ersten Zeichen in ein Feld A für alle $i \in [1 : |t\$| - 1]$ gilt, dass entweder $t_1^{A[i]} < t_1^{A[i+1]}$ oder $t_1^{A[i]} = t_1^{A[i+1]}$ und $A[i] < A[i+1]$. Umgangssprachlich bedeutet dies, dass für Zeichenreihen mit demselben ersten Buchstaben innerhalb dieses Buckets die ursprüngliche Reihenfolge beibehalten werden soll.

Aufgabe 2

Erstelle für das Wort $t\$ = t_0 \cdots t_{10}\$ = \text{baabaabbaab}\$$ ein Suffix-Array je einmal nach dem Algorithmus von Manber und Myers und dem Algorithmus von Kärkkäinen und Sanders.

Gib dabei alle Zwischenschritte an, wobei der rekursive Aufruf beim Algorithmus von Kärkkäinen und Sanders von Hand sortiert werden darf.

*Wir wünschen allen
ein erfolgreiches und gesundes
neues Jahr 43*47!*