

Aufgabe 1 Suchen in Zeichenketten I

10 Punkte

Implementieren Sie den naiven Algorithmus und den Algorithmus von Rabin-Karp zur Suche in Zeichenketten. Finden Sie dann heraus, wie oft das Wort *whale* in Moby Dick vorkommt (ignorieren Sie dabei Groß- und Kleinschreibung). Wie schneiden Ihre Implementierungen im Vergleich ab?

Hinweis: Den Roman Moby Dick finden Sie unter
<http://www.gutenberg.org/files/2701/2701-0.txt>.

Aufgabe 2 Suchen in Zeichenketten II

10 Punkte

- (a) Der Algorithmus von Rabin-Karp lässt sich leicht auf *mehrere* Suchmuster verallgemeinern. Gegeben eine Zeichenkette s und Suchmuster t_1, \dots, t_k , bestimme die erste Stelle in s , an der eines der Muster t_1, \dots, t_k vorkommt. Beschreiben Sie, wie man den Algorithmus von Rabin-Karp für diese Situation anpassen kann. Was ist die heuristische Laufzeit Ihres Algorithmus (unter der Annahme, dass Kollisionen selten sind)?

Hinweis: Nehmen Sie zunächst an, dass alle Suchmuster die gleiche Länge haben. Der allgemeine Fall lässt sich darauf zurückführen, z.B., indem man zum Hashen alle Suchmuster auf die Minimallänge verkürzt.

- (b) Implementieren Sie Ihren Algorithmus aus (a). Beantworten Sie sodann folgende Frage: Was kommt öfter in dem Roman *Sense & Sensibility* vor: *sense* oder *sensibility/sensible*?

Hinweis: Siehe <http://www.gutenberg.org/files/161/161-0.txt>.

Aufgabe 3 Suche in Zeichenketten III

10 Punkte

Sei $\Sigma = C, G, T, A$. Sei $s = \text{CTTGGATTA}$ und $t = \text{TTA}$.

- (a) Verwenden Sie den naiven Algorithmus, um festzustellen, ob/wo das Muster t in der Zeichenkette s vorkommt.
Zeigen Sie die einzelnen Schritte.
- (b) Verwenden Sie den Algorithmus von Rabin-Karp, um festzustellen, ob/wo das Muster t in der Zeichenkette s vorkommt. Verwenden Sie $A : 0, T : 1, G : 2, C : 3$ und die Primzahl 5 als Modulus für die Hashfunktion.
Zeigen Sie die einzelnen Schritte.

- (c) Verwenden Sie den Algorithmus von Knuth-Morris-Pratt, um festzustellen, ob/wo das Muster t in der Zeichenkette s vorkommt. Zeigen Sie die einzelnen Schritte.