
Algorithmische Bioinformatik I

Aufgabe 1

Wie viel Speicherplatz (in MB) ist notwendig, um das menschliche Genom abzuspeichern? Finden Sie dafür zunächst heraus, aus wie vielen Basenpaaren (bp) es besteht.

Aufgabe 2

Bestimmen Sie für die Zeichenreihe

$$s = 01011010100$$

die Tabelle $border[]$ für den Knuth-Morris-Pratt-Algorithmus.

Aufgabe 3

Modifizieren Sie den KMP-Algorithmus so, dass er für jedes Vorkommen einer Zeichenreihe s in t die Position des Vorkommens ausgibt.

Aufgabe 4

Geben Sie eine Optimierung des KMP-Algorithmus für den Fall an, dass das vorliegende Alphabet binär ist ($\Sigma = \{0, 1\}$). Definieren Sie hierzu die $border[]$ -Tabelle geeignet um und modifizieren Sie den KMP-Algorithmus.

Wie sieht die Tabelle für das Beispiel in Aufgabe 2 aus?

Aufgabe 5

Konstruieren Sie einen Aho-Corasick-Automaten für folgende Menge von Suchwörtern:

$$S = \{IGEL, ELF, IST, IM, MIST, ES, IST, EIN, ESSEN, KEIN\}$$