Technische Universität München
Fakultät für Informatik
Lehrstuhl für Effiziente Algorithmen
Dr. Hanjo Täubig
Johannes Nowak

Sommersemester 2008
Übungsblatt 2
06. Mai 2008

Algorithmische Bioinformatik I

Aufgabe 1

Konstruieren Sie einen AC-Automaten für die Menge $S = \{IGEL, ELF, IST, IM, MIST, ES, IST, EIN, ESSEN, KEIN\}$ von Suchwörtern.

Aufgabe 2

Bestimme für die Zeichenreihe

$$abaabbabaabaabbabaabbabab$$

die Shift-Tabelle für den Boyer-Moore-Algorithmus.

Aufgabe 3

Betrachte den Boyer-Moore Algorithmus, der für die Bestimmung der Shifts bei einem Mismatch nur die Bad-Character-Rule berücksichtigt:

```
BC-Boyer-Moore(char t[], int n, char s[], int m)
{
    int i = 0;
    int j = m − 1;
    while (i ≤ n − m)
    {
        while (t[i + j] == s[j])
        {
            if (j == 0) return TRUE;
            j−−;
        }
        i = i + j − max{k : (k < j ∧ s[k] = t[i + j]) ∨ (k = −1)};
        j = m − 1;
    }
    return FALSE;
}
```

Man gebe eine unendliche Familie von Beispielen an, bei dem diese Variante eine Laufzeit von $\Theta(n \cdot m)$ hat.

Aufgabe 4

Ein Wort $w \in \Sigma^*$ heißt *periodisch*, wenn es ein Wort $v \in \Sigma^*$ und ein $i \in \mathbb{N}$ gibt, so dass $|v| \geq 1$, $n > 1$ und $w = v^i$.

Konstruiere einen Algorithmus an, der mit linear vielen Vergleichen feststellt, ob ein Wort periodisch ist.

Aufgabe 5

Man gebe für jedes Paar $n, m \in \mathbb{N}$ jeweils einen Text t der Länge n und ein Suchmuster s der Länge m an, so dass der Algorithmus von Knuth, Morris und Pratt mindestens $2n-m$ Vergleiche ausführt (die Vergleiche zum Erstellen der Tabelle $border[]$ sind hierbei nicht zu berücksichtigen).