
Algorithmische Bioinformatik I

Abgabetermin: Donnerstag, den 09. Juni, vor der Vorlesung

Aufgabe 1

Bestimme für das Wort $s = \text{babbaababbaababa}$ die Shift-Tabelle für den Boyer-Moore-Algorithmus aus der Vorlesung.

Gib dabei nicht nur die Tabelle, sondern auch alle Zwischenschritte für die Berechnung an.

Aufgabe 2

Betrachte den unten angegebenen Boyer-Moore-Algorithmus, der für die Bestimmung der Shifts bei einem Mismatch nur die Extended-Bad-Character-Rule berücksichtigt. Gib eine Konstante $c > 0$ sowie eine unendliche Familie $\mathcal{F} = \{(s, t) : s, t \in \Sigma^*\}$ an, die für jedes Paar $(m, n) \in \mathbb{N}^2$ ein Paar $(s, t) \in \mathcal{F}$ mit $|s| \geq m$, $|t| \geq n$ und $|s| \leq |t|$ enthält und bei dem diese Variante für eine erfolglose Suche mindestens $c \cdot (|s| \cdot |t|)$ Zeichenvergleiche ausführt.

BC-Boyer-Moore (char $t[]$, int n , char $s[]$, int m)

```
int  $i := 0$ ,  $j := m - 1$ ;  
while ( $i \leq n - m$ ) do  
  while ( $t[i + j] = s[j]$ ) do  
    if ( $j = 0$ ) then return TRUE;  
     $j--$ ;  
     $i := i + j - \max\{k : (k < j \wedge s[k] = t[i + j]) \vee (k = -1)\}$ ;  
     $j := m - 1$ ;  
  return FALSE;
```

Aufgabe 3

Bestimme die Z-Werte für das folgende Wort: bbabbababb .

Gib dabei nicht nur die Tabelle, sondern auch die Zwischenschritte (d.h. die verwendeten Z-Boxen) für die Berechnung an und insbesondere auch welcher Fall jeweils eingetreten ist.

Am Dienstag, den 7. Juni findet um 16ct im Hörsaal B006 in der Theresienstr. 39 eine Informationsveranstaltung für Studierende der Bioinformatik zum Hauptstudium (5. und 6. Fachsemester) im Bachelor und zum Übergang in den Master Bioinformatik statt.