

3. Übung „Algorithmen der Bioinformatik I“

1. Zum induktiven Beweis der Korrektheit des *Z-Algorithmus* (Notation: siehe Vorlesung):
 - a) Zeigen Sie, dass ein Schritt des *Z-Algorithmus* im Fall IIa die Werte $Z[i]$ für $i \geq 3$ korrekt berechnet, falls $Z[j]$ für $j = 1, \dots, i - 1$ und $r = r_{i-1}, l = l_{i-1}$ bereits korrekt berechnet wurden. (2 Punkte)
 - b) Beweisen Sie, dass im Fall IIb des *Z-Algorithmus* aus der Vorlesung die *Z-Boxen* $Z[i]$ korrekt berechnet werden, falls $Z[j]$ für $j = 1 \dots, i - 1$ und $r = r_{i-1}, l = l_{i-1}$ bereits korrekt berechnet wurden. (3 Punkte)
 - c) Zeigen Sie die Korrektheit der Berechnung von r, l im *Z-Algorithmus* unter der Voraussetzung, dass $Z[i]$ korrekt bestimmt wurde. (2 Punkte)
2. Erweitern Sie Ihre Implementierung des naiven Exact-Matching Algorithmus aus der ersten Übung um die *bad character rule* kombiniert mit Rechts-Links-Vergleichen. Überprüfen Sie die Korrektheit anhand des naiven Algorithmus. Testdaten finden Sie auf der Webpage zur Übung (http://www2.informatik.uni-halle.de/agprbio/AG/Lehre/AdB-I_SS05/uebungen.html). (4 Punkte)
3. Wir betrachten das *Good Suffix Preprocessing* im Boyer-Moore-Algorithmus.
 - a) Beweisen Sie, dass $N_j(P) = Z_{n-j+1}(P^r)$ gilt. (2 Punkte)
 - b) Beweisen Sie, dass der Algorithmus zur Berechnung der $L'(i)$ aus den $N_j(P)$ korrekt arbeitet. (2 Punkte)

Abgabe des Programms per E-Mail an { gohr | grau } @ informatik.uni-halle.de.