

7. Übung zur Vorlesung “Einführung in die Bioinformatik I, 1. Teil”

Wintersemester 2019/2020

Prof. Peter Dittrich, Emanuel Barth, Marcus Ludwig

Ausgabe: 4. Dezember 2019,
Abgabe: 11. Dezember 2019 zu Beginn der Übung

Boyer-Moore-Algorithmus

Aufgabe 1 (5 Punkte): Berechnen Sie für die Strings `abcabcabc`, `aabababab` und `abbabab` jeweils die Werte $L(i)$, $L'(i)$ und $l'(i)$.

Aufgabe 2 (5 Punkte): Wenden Sie den Boyer-Moore-Algorithmus auf den Text `acababacabababa` und das Pattern `abacaba` an und geben Sie für jeden Schritt an, wie groß die Verschiebung des Patterns jeweils nach der Good-Suffix-Regel und nach der Bad-Character-Regel wäre.

Aufgabe 3 (5 Punkte): Beweisen Sie: $l'(i)$ ist der größte Index $j \leq m-i+1$, sodass $N_j(P) = j$.
Anmerkung: $N_j(P)$ ist die Länge des längsten Suffixes von $P[1 \dots j]$, der gleichzeitig Suffix von P ist.

Aufgabe 4 (5 Punkte): Konstruieren Sie ein Beispiel (Text und Pattern), für das der Boyer-Moore-Algorithmus weniger Vergleiche benötigt, wenn nur die Bad-Character-Regel angewendet wird, als wenn diese mit der Good-Suffix-Regel kombiniert wird.