

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, SS 2021

Prof. Dr. Jens Stoye · Dr. Marília D. V. Braga

<https://gi.cebitec.uni-bielefeld.de/teaching/2021summer/sa>

Übungsblatt 12 vom 8.7.2021

Abgabe am 15.7.2021 bis 12:00 Uhr (mittags)

Aufgabe 1 (Suchraum multipler Alignments)

(8 Punkte)

Um den Wert $D(i_1, i_2, \dots, i_k)$ eines Knotens im k -dimensionalen Edit-Graphen mittels des grundlegenden Algorithmus (Abschnitt 11.2.1 im Skript) zu berechnen, muss über eine Anzahl an Werten minimiert werden, die der Anzahl Vorgängerknoten im Graphen entspricht. Die Zahl dieser Vorgängerknoten ist unterschiedlich, je nachdem, wo man sich im Edit-Graphen befindet; es gibt also verschiedene *Typen* von Knoten. Betrachte die folgenden Situationen:

- (i) $k = 2$ Sequenzen der Länge 3;
- (ii) $k = 3$ Sequenzen der Länge 3; und
- (iii) $k = 4$ Sequenzen der Länge 3.

Für die entsprechenden k -dimensionalen Edit-Graphen sind die folgenden Fragen zu beantworten:

1. Wie viele Knoten hat der k -dimensionale Edit-Graph? (Start- und Endknoten des Graphen können ignoriert werden.)
2. Welche verschiedenen Typen an Knoten gibt es, und wie viele eingehende Kanten hat jeder von ihnen? Wie viele Knoten jeden Typs gibt es?
3. Wie viele Berechnungsschritte werden also insgesamt durchgeführt?

Aufgabe 2 (Carrillo-Lipman-Heuristik)

(4 Punkte)

1. Wie viele Carrillo-Lipman-Schranken $U_{x,y}$ müssen berechnet werden, um k Sequenzen zu alignieren? Wie ist die asymptotische Laufzeit für die Berechnung der Schranken?
2. Für welche Sequenzen funktioniert die Carrillo-Lipman-Heuristik gut bzw. schlecht?

Aufgabe 3 (Carrillo-Lipman in der Praxis)

(4 Punkte)

Um den Suchraum mittels Carrillo-Lipman Schranken einzuschränken, verwendet man eine obere Schranke für die optimalen Alignmentkosten. Durch geschicktes „Mogeln“ kann diese Schranke verfeinert werden.

1. Diskutiere Vor- und Nachteile dieses „Mogelns“.
2. Überlege dir eine Strategie, wie man günstige Werte für die $\epsilon_{(x,y)}$ findet, um in möglichst kurzer Zeit ein garantiert optimales Alignment zu erhalten.