Ludwig-Maximilians-Universität München
Institut für Informatik
Prof. Dr. Volker Heun

Wintersemester 2016/17
Übungsblatt 7
8. Dezember 2016

Algorithmische Bioinformatik II

Abgabetermin: Donnerstag, den 15. Dezember, vor der Vorlesung

Aufgabe 1

Betrachte folgende Sequenzen $s_1 = ACGTGC$, $s_2 = AGCC$, $s_3 = ACCTG$ und $s_4 = AGGCTT$. Der optimale Abstand für die paarweise Sequenzen-Alignments ist rechts angegeben. Hierbei gilt $w(a,b) = 1$ und $w(a,a) = 0$ für alle $a \neq b \in \overline{\Sigma}$. Konstruiere für diese Sequenzen ein mehrfaches Sequenzen-Alignment mit Hilfe der Center-Star-Methode.

d	s_1	s_2	s_3	s_4
s_1	0	3	2	4
s_2	3	0	3	3
s_3	2	3	0	3
s_4	4	3	3	0

Aufgabe 2

Zeige, dass es für jedes $n \in \mathbb{N}$ eine Menge von Sequenzen $S = \{s_1, \ldots, s_k\}$ mit $|s_i| \geq n$ gibt, so dass es eine Sequenz $s \in S$ gibt, die als Zentrum bei der Center-Star-Methode eine Approximationsgüte von $\Omega(k)$ liefert. Hierbei gilt $w(x,y) = 1$ und $w(x,x) = 0$ für alle $x \neq y \in \overline{\Sigma}$.

Hinweis: Die Menge ist hier als Multimenge zu verstehen, d.h. Sequenzen dürfen mehrfach in S vorkommen.

Aufgabe 3

Sei $S = \{s_1, \ldots, s_k\} \subseteq \Sigma^*$ und sei $M(i) = E_S(s_i)$. Dabei sei ohne Beschränkung der Allgemeinheit $M := M(1) \leq \cdots \leq M(k)$.

Zeige, dass $M(\lfloor \frac{k+1}{2} \rfloor) \leq 3M$.

Hinweis: Die Beziehung $\frac{1}{k} \sum_{i=1}^{k} M(i) < 2M$ kann hilfreich sein (siehe Beweis von Lemma 6.42).