

Präsenzübungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2015/2016

Prof. Dr. Jens Stoye · M.Sc. Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2015winter/SequenzAnalyse>

Präsenzübungsblatt 6, Woche 8.12.–14.12.2015

Aufgabe 1 (\$ – Der Wächter)

Bei der Erstellung eines Suffixbaums wird am Ende des Strings ein eindeutiges Zeichen, z.B. $\$ \in \Sigma$, angehängt. Warum?

Aufgabe 2 (Verallgemeinerter Suffixbaum)

Gegeben seien die Strings $s=GAAG$ und $t=AAGG$.

1. Zeiche den generalisierten Suffixbaum von s und t (mit $\# < \$ < A < G$).
2. Wie kann man den längsten gemeinsamen Substring von s und t finden? Gib seine Vorkommen an.
3. Formuliere, wie man allgemein längste gemeinsame Substrings zweier Strings im generalisierten Suffixbaum finden kann.
4. Wie kann man das längste palindromische Teilwort in einem Wort mit einem generalisierten Suffixbaum finden? Verwende dazu das Beispiel $x = BANANAS$. Beschreibe deine Idee. Es ist nicht nötig, den Baum explizit zu zeichnen.

Aufgabe 3 (Multiples Alignment)

Gibt jeweils eine bioinformatische Anwendung an, in der multiple Alignments verwendet werden, um (a) gemeinsame Abschnitte bzw. (b) unterschiedliche Regionen einer Menge von Sequenzen hervorzuheben.