

Präsenzübungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2013/2014
Prof. Dr. Jens Stoye · Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2013winter/SequenzAnalyse>

Präsenzübungsblatt 5, Woche 47/2013

Aufgabe 1 (Balanced De-Bruijn-Subgraph)

Zeichne den balancierten De-Bruijn-Subgraph für die Sequenz $x = \text{AAGTTTATATTAGTATG}$ mit $q = 4$. Beantworte folgende Fragen zu den Eigenschaften des vorliegenden Graphen:

1. Wann nennt man einen Knoten in einem Graph balanciert?
2. Wann ist ein gerichteter Graph balanciert?
3. Was ist ein Euler-Pfad in einem Graph?
4. Was unterscheidet den De-Bruijn-Subgraph von einem De-Bruijn-Graph?
5. Welche anderen Worte haben das gleiche q-gram-Profil wie x ?

Wenn ihr in folgenden Aufgaben einen „De-Bruijn-Graphen“ zeichnen sollt, ist der balancierte Subgraph gemeint, außer es wird explizit anders angegeben.

Aufgabe 2 (Filter)

1. Was ist mit *Filtern* im Zusammenhang mit Sequenzvergleichen gemeint? Warum verwendet man Filter?
2. Welche Distanzmodelle kann man beim Filtern benutzen?
3. Worin unterscheidet sich eine Heuristik von einem Filter?

Aufgabe 3 (Maximal-Matches-Distanz als Filter)

Entscheide mithilfe der Maximal-Matches-Distanz, ob die Sequenzen y_1 bis y_5 eine Edit-Distanz von max. 1 zur Sequenz x haben können.

$x = \text{AATCGAGGTAC}$

$y_1 = \text{TCGAGCGAGGT}$

$y_2 = \text{AAGATCGGACC}$

$y_3 = \text{AATCGCGGTAC}$

$y_4 = \text{AATCGGGAGGT}$

$y_5 = \text{CTGAACGTCTG}$