

# Präsenzübungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, SS 2017

Prof. Dr. Jens Stoye · M.Sc. Tizian Schulz

<https://gi.cebitec.uni-bielefeld.de/teaching/2017summer/sa>

## Präsenzübungsblatt 4, Woche 20/2017

### 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. Welche anderen Worte haben das gleiche q-gram-Profil wie  $x$ ?

### Aufgabe 2 (Maximal-Matches Distanz)

Berechne  $\delta$  und gib jeweils die links-nach-rechts und rechts-nach-links Partitionen an.

1.  $\delta(\text{montreal} \parallel \text{monmarelen})$
2.  $\delta(\epsilon \parallel \text{aaaa})$
3.  $\delta(\text{aaaa} \parallel \epsilon)$

Welche Beobachtung machst du bei 2. und 3.? Was sagt das über die Maximal-Matches Distanz aus?

### 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.

$x = \text{AATCGAGGTAC}$

$y_1 = \text{TCGAGCGAGGT}$

$y_2 = \text{AAGATCGGACC}$

$y_3 = \text{AATCGCGGTAC}$

$y_4 = \text{AATCGGGAGGT}$

$y_5 = \text{CTGAACGTCTG}$