

Präsenzübungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2024

Prof. Dr. Jens Stoye · Leonard Bohnenkämper

<https://gi.cebitec.uni-bielefeld.de/teaching/2024winter/sa>

Präsenzübungsblatt 2, vom 04.11.2024/05.11.2024

Aufgabe 1 (*Edit*-Distanz und Transformationen)

1. Wie ist die *Edit*-Distanz formal definiert?
2. Zeige formal, dass folgende Transformationen, wenn sie gleichzeitig auf Sequenzen u und v angewendet werden, den Wert der *Edit*-Distanz $d(\cdot, \cdot)$ nicht ändern:
 - (a) Umkehrung (d.h. Rückwärtslesen);
 - (b) Umbenennung der Zeichen;
 - (c) Reverses Komplement.

Sind also u und v DNA-Sequenzen, so folgt, dass $d(u, v) = d(u^{-1}, v^{-1}) = d(\bar{u}, \bar{v})$ ist, wobei x^{-1} die Rückwärts-Sequenz und \bar{x} das reverse Komplement von x sind.

Aufgabe 2 (*Edit*-Matrix)

Gegeben seien die Sequenzen $x = \text{CAG}$ und $y = \text{GA}$.

1. Berechne die *Edit*-Matrix $D(x, y)$ und gib die *Edit*-Distanz $d(x, y)$ der beiden Strings an.
2. Gib alle optimalen *Edit*-Sequenzen an.