

Präsenzübungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2012/2013

Prof. Dr. Jens Stoye · Nina Luhmann · Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2012winter/SequenzAnalyse>

Präsenzübungsblatt 2, Woche 44/2012

Aufgabe 1 (LCS-Distanz)

Gegeben seien die beiden Strings s und t mit Länge $|s|$ und $|t|$.

1. Schreibe die Rekurrenzformel und die Basisfälle zur Berechnung der Edit-Distanz so um, dass du die LCS-Distanz der beiden Strings s und t damit berechnen kannst.
2. Berechne mit Hilfe der Edit-Matrix $D(s, t)$ die LCS-Distanz für die beiden Strings $s = ACTT$ und $t = TCTA$. Gib zudem alle möglichen Pfade durch die Matrix an.
Wie kannst du nun mit Hilfe der d_{LCS} auf die längste Subsequenz schliessen?

Aufgabe 2 (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) Reverses Komplement

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