

Übungen zur Vorlesung Sequenzanalyse I

Universität Bielefeld, WS 2011/2012

Dr. Alexander Sczyrba · Nina Luhmann · Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2011winter/SequenzAnalyse>

Übungsblatt 3 vom 18.11.2011

Abgabe in einer Woche vor Beginn der Vorlesung.

Aufgabe 1 (*Edit-Funktion*)

(2 Punkte)

Überführe den String $x = \text{ADVENT}$ in den String $y = \text{VERPENNT}$ durch Anwendung der *Edit-Funktion*.

1. Gib eine möglichst kurze Edit-Sequenz an, die x in y überführt.
2. Gib eine der längsten Edit-Sequenzen an, die x in y überführt.

Aufgabe 2 (*Edit-Distanz und Transformationen*)

(5 Punkte)

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.

Aufgabe 3 (*Edit-Matrix*)

(2 Punkte)

Die Edit-Distanz kann effizient mit Hilfe der *Edit-Matrix* berechnet werden.

1. Berechne von Hand die Edit-Matrix $D(x, y)$ für die Sequenzen $x = \text{FEIERTAG}$ und $y = \text{FREIERTACH}$.
 - Gib die Matrix, sowie die Edit-Distanz $d(x, y)$ der beiden Strings an. (Zur einfacheren Korrektur schreibe x links vertikal neben die Matrix und y oben horizontal an die Matrix.)