

# Übung zu Algorithmen auf Sequenzen

## Blatt 8

Ausgabe: 13.12.2018    Besprechung: 20.12.2018

### Aufgabe 8.1

Zeigen Sie, dass die Edit-Distanz eine Metrik auf  $\Sigma^*$  ist.

### Aufgabe 8.2

Berechnen Sie die Edit-Distanz für folgende Sequenzpaare:

- jaegerboot, jamesbond
- pflanze, panzer
- haenschen, haenchen
- atlantik, cyborg

Wie kann das Ergebnis des letzten Paares verallgemeinert werden?

### Aufgabe 8.3

Wie viele Edit-Sequenzen gibt es zwischen zwei Strings der Länge 5? Schreiben Sie eine Routine, die alle Edit-Sequenzen systematisch als sog. "Cigar-String" auflistet. Dabei steht M für Match/Mismatch (also zwei alignierte Zeichen), I für Insertion und D für Deletion. Beispiele sind also MMMMM oder IIIIDDDDDD.

### Zusatzaufgabe 8.4

Angenommen, Sie kennen eine (korrekte) obere Schranke von  $K$  für die Edit-Distanz zwischen zwei gegebenen Sequenzen. Wie kann dann der DP-Algorithmus zur Berechnung der Edit-Distanz verbessert werden, so dass die worst-case Laufzeit nur noch  $\mathcal{O}(Kn)$  beträgt?

Konstruieren Sie aus dieser Idee einen Algorithmus zur Berechnung der Edit-Distanz mit der Laufzeit  $\mathcal{O}(kn)$ , wobei  $k$  die (unbekannte) tatsächliche Edit-Distanz ist.