



Blatt 10

Aufgabe 10.1

- a) Ermitteln Sie drei unterschiedliche Alignments für die Strings “KUNST” und “KITSCH”, die jeweils mindestens ein *match*, ein *insert* oder *delete* und ein *replace* enthalten. Bestimmen Sie die Kosten für diese drei Alignments gemäß folgender Kostenfunktion $\delta(a, b)$:

$$\delta(a, b) = \begin{cases} 0, & a = b \\ 2, & a \neq b, \quad a, b \neq \lambda \\ 1, & a = \lambda \text{ oder } b = \lambda \end{cases}$$

(1 Punkt)

- b) Stellen Sie den Edit-Graphen für die Strings “KUNST” und “KITSCH” auf. Markieren Sie in diesem Edit-Graphen die Pfade, die den Alignments aus Aufgabe a) entsprechen. (1 Punkt)

Aufgabe 10.2 Diese Aufgabe beschäftigt sich mit dem Sequenzalignment und Edit-Transkripten für zwei Strings S_1 und S_2 . Zeigen Sie, dass jedes Alignment von S_1 und S_2 durch genau ein Edit-Transkript erzeugt werden kann. (3 Punkte)

Aufgabe 10.3 In der Vorlesung wurde der Satz eingeführt, dass zu jedem Edit-Transkript $T = (t_1, \dots, t_L)$ für zwei Strings S_1 und S_2 , $n = |S_1|$, $m = |S_2|$, genau ein Pfad von $(0, 0)$ nach (n, m) im Edit-Graphen $G(S_1, S_2)$ existiert. Zeigen Sie, dass dieser Satz gilt. Dazu müssen zwei Dinge gezeigt werden:

- Für jedes Edit-Transkript $T = (t_1, \dots, t_L)$ für S_1 und S_2 gibt es in $G(S_1, S_2)$ einen Pfad von $(0, 0)$ nach (n, m) mit der Kantenbeschriftung t_1, \dots, t_L . (3 Punkte)
- Jeder Pfad von $(0, 0)$ nach (n, m) in $G(S_1, S_2)$ repräsentiert ein Edit-Transkript von S_1 nach S_2 (3 Punkte)

Abgabe per E-Mail an grau@informatik.uni-halle.de oder in Raum 4.12.