Hochschule Aalen
Fakultät Elektronik und Informatik
Studienbereich Informatik

Algorithmen und Datenstrukturen 2

Vorlesung im Wintersemester 2024/2025
Prof. Dr. habil. Christian Heinlein

6. Übungsblatt (16. Januar 2025)

Aufgabe 15: Editierdistanz

Gegeben seien die Zeichenfolgen OTTO, LOTTO und TOTO.

Bestimmen Sie die Editierdistanz zwischen jeweils zwei dieser Zeichenfolgen und geben Sie jeweils eine minimale Folge elementarer Editieroperationen an, um eine Folge in die andere zu überführen.

Elementare Editieroperationen sind:

- Entferne das Zeichen . . . an Position . . .
- Füge das Zeichen . . . nach Position . . . ein
- Ersetze das Zeichen . . . an Position . . . durch das Zeichen . . .

✎

a) OTTO ↔ LOTTO

		0	1	2	3	4
			O	T	T	O
0		**0**	1	2	3	4
1	L	**1**	1	2	3	4
2	O	2	**1**	2	3	3
3	T	3	2	**1**	2	3
4	T	4	3	2	**1**	2
5	O	5	4	3	2	**1**

OTTO → LOTTO

- Füge das Zeichen L nach Position 0 ein.

LOTTO → OTTO

- Entferne das Zeichen L an Position 1.

b) OTTO $\leftrightarrow$ TOTO

		0	1	2	3	4
			O	T	T	O
0		**0**	**1**	2	3	4
1	T	**1**	**1**	**1**	2	3
2	O	2	**1**	**2**	2	2
3	T	3	2	**1**	**2**	3
4	O	4	3	2	2	**2**

OTTO $\rightarrow$ TOTO

- Möglichkeit 1 ($D_{4,4} \rightarrow D_{3,3} \rightarrow D_{2,2} \rightarrow D_{1,1} \rightarrow D_{0,0}$):
  - Ersetze das Zeichen T an Position 2 durch das Zeichen O.
  - Ersetze das Zeichen O an Position 1 durch das Zeichen T.
- Möglichkeit 2 ($D_{4,4} \rightarrow D_{3,3} \rightarrow D_{3,2} \rightarrow D_{2,1} \rightarrow D_{1,0} \rightarrow D_{0,0}$):
  - Entferne das Zeichen T an Position 3.
  - Füge das Zeichen T nach Position 0 ein.
- Möglichkeit 3 ($D_{4,4} \rightarrow D_{3,3} \rightarrow D_{2,2} \rightarrow D_{1,2} \rightarrow D_{0,1} \rightarrow D_{0,0}$):
  - Füge das Zeichen O nach Position 2 ein.
  - Entferne das Zeichen O an Position 1.
- Möglichkeit 4 ($D_{4,4} \rightarrow D_{3,3} \rightarrow D_{2,2} \rightarrow D_{2,1} \rightarrow D_{1,0} \rightarrow D_{0,0}$):
  - Entferne das Zeichen T an Position 2.
  - Füge das Zeichen T nach Position 0 ein.

TOTO $\rightarrow$ OTTO

- Möglichkeit 1 ($D_{4,4} \rightarrow D_{3,3} \rightarrow D_{2,2} \rightarrow D_{1,1} \rightarrow D_{0,0}$):
  - Ersetze das Zeichen O an Position 2 durch das Zeichen T.
  - Ersetze das Zeichen T an Position 1 durch das Zeichen O.
- Möglichkeit 2 ($D_{4,4} \rightarrow D_{3,3} \rightarrow D_{3,2} \rightarrow D_{2,1} \rightarrow D_{1,0} \rightarrow D_{0,0}$):
  - Füge das Zeichen T nach Position 3 ein.
  - Entferne das Zeichen T an Position 1.
- Möglichkeit 3 ($D_{4,4} \rightarrow D_{3,3} \rightarrow D_{2,2} \rightarrow D_{1,2} \rightarrow D_{0,1} \rightarrow D_{0,0}$):
  - Entferne das Zeichen O an Position 2.
  - Füge das Zeichen O nach Position 0 ein.
- Möglichkeit 4 ($D_{4,4} \rightarrow D_{3,3} \rightarrow D_{2,2} \rightarrow D_{2,1} \rightarrow D_{1,0} \rightarrow D_{0,0}$):
  - Füge das Zeichen T nach Position 2 ein.
  - Entferne das Zeichen T an Position 1.

c) TOTO ↔ LOTTO

		0	1	2	3	4
			T	O	T	O
0		**0**	1	2	3	4
1	L	1	**1**	2	3	4
2	O	2	2	**1**	2	3
3	T	3	2	**2**	**1**	2
4	T	4	3	3	**2**	2
5	O	5	4	3	3	**2**

TOTO → LOTTO

- Möglichkeit 1 ($D_{5,4} \to D_{4,3} \to D_{3,2} \to D_{2,2} \to D_{1,1} \to D_{0,0}$):
  - Füge das Zeichen T nach Position 2 ein.
  - Ersetze das Zeichen T an Position 1 durch das Zeichen L.
- Möglichkeit 2 ($D_{5,4} \to D_{4,3} \to D_{3,3} \to D_{2,2} \to D_{1,1} \to D_{0,0}$):
  - Füge das Zeichen T nach Position 3 ein.
  - Ersetze das Zeichen T an Position 1 durch das Zeichen L.

LOTTO → TOTO

- Möglichkeit 1 ($D_{5,4} \to D_{4,3} \to D_{3,2} \to D_{2,2} \to D_{1,1} \to D_{0,0}$):
  - Entferne das Zeichen T an Position 3.
  - Ersetze das Zeichen L an Position 1 durch das Zeichen T.
- Möglichkeit 2 ($D_{5,4} \to D_{4,3} \to D_{3,3} \to D_{2,2} \to D_{1,1} \to D_{0,0}$):
  - Entferne das Zeichen T an Position 4.
  - Ersetze das Zeichen L an Position 1 durch das Zeichen T.

✐

Aufgabe 16: Blocksatz

Gegeben sei folgender Text mit $n = 12$ Wörtern:

```
Eins, zwei, drei, vier, fünf, sechs, sieben, in der Schule wird geschrieben.
```

Ein Satzzeichen gehört dabei immer zum vorhergehenden Wort.

Der Einfachheit halber wird „Schreibmaschinenschrift“ verwendet, d. h. jedes Zeichen einschließlich Satzzeichen und Leerzeichen hat Breite 1. Zwischen zwei Wörtern muss jeweils mindestens ein Leerzeichen stehen, d. h. der Mindestabstand σ_0 zwischen zwei Wörtern ist ebenfalls 1, und der zum Auffüllen einer Zeile zusätzlich benötigte Zwischenraum σ muss ganzzahlig aufgeteilt werden.

Die Zeilenbreite λ sei 23.

a) Formatieren Sie den Text gemäß der Nächstbest-Methode, indem Sie jede Zeile mit möglichst vielen Wörtern füllen!

✎

```
Eins, zwei, drei, vier,
fünf, sechs, sieben, in
der     Schule     wird
geschrieben.
```

✐

b) Berechnen Sie den Malus aller resultierenden Zeilen sowie den Malus des gesamten Absatzes!

✎

- Zeile 1 enthält die Wörter 1 bis 4 mit $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 5$.
  Damit ergibt sich für den zusätzlich benötigten Zwischenraum in dieser Zeile:
  $\sigma = \lambda - \sum_{k=1}^{4} \lambda_k - m\,\sigma_0 = 23 - (5 + 5 + 5 + 5) - 3 \cdot 1 = 0$
  und somit $\mu(1, 4) = 0$.
- Zeile 2 enthält die Wörter 5 bis 8 mit $\lambda_5 = 5$, $\lambda_6 = 6$, $\lambda_7 = 7$ und $\lambda_8 = 2$.
  Damit ergibt sich für den zusätzlich benötigten Zwischenraum in dieser Zeile:
  $\sigma = \lambda - \sum_{k=5}^{8} \lambda_k - m\,\sigma_0 = 23 - (5 + 6 + 7 + 2) - 3 \cdot 1 = 0$
  und somit $\mu(5, 8) = 0$.
- Zeile 3 enthält die Wörter 9 bis 11 mit $\lambda_9 = 3$, $\lambda_{10} = 6$ und $\lambda_{11} = 4$.
  Damit ergibt sich für den zusätzlich benötigten Zwischenraum in dieser Zeile:
  $\sigma = \lambda - \sum_{k=9}^{11} \lambda_k - m\,\sigma_0 = 23 - (3 + 6 + 4) - 2 \cdot 1 = 8$
  und somit $q = \left\lfloor \frac{\sigma}{m} \right\rfloor = \left\lfloor \frac{8}{2} \right\rfloor = 4$, $r = \sigma \bmod m = 8 \bmod 2 = 0$
  und $\mu(9, 11) = (m - r) \cdot q^2 + r \cdot (q + 1)^2 = (2 - 0) \cdot 4^2 + 0 \cdot (4 + 1)^2 = 32$.
- Zeile 4 enthält lediglich das Wort 12 mit $\lambda_{12} = 12$.
  Damit ergibt sich für den zusätzlich benötigten Zwischenraum in dieser Zeile:
  $\sigma = \lambda - \sum_{k=12}^{12} \lambda_k - m\,\sigma_0 = 23 - 12 - 0 \cdot 1 = 11$.
  Da es sich um die letzte Zeile handelt und $\sigma \geq 0$ ist, ist $\mu(12, 12) = 0$.
- Damit ergibt sich als Malus des gesamten Absatzes:
  $\mu\big(\{ 4, 8, 11 \}\big) = \mu(1, 4) + \mu(5, 8) + \mu(9, 11) + \mu(12, 12) = 0 + 0 + 32 + 0 = 32$.

✐

c) Ermitteln Sie dann mit dem in der Vorlesung vorgestellten Algorithmus die optimale Formatierung des Textes, indem Sie für $j = 0, \ldots, n$ jeweils $\mu(j)$ sowie $\pi(j)$ berechnen!

✎

j	i $\pi(j)$	Letzte Zeile (Wort $i + 1$ bis j)	$\mu(i + 1, j)$	$\mu(i)$	$\mu(i) + \mu(i + 1, j)$ $\mu(j)$
0					**0**
1	**0**	`Eins,++++++++++++++++++`	$18^2 = 324$	0	**324**
2	1	`zwei,++++++++++++++++++`	$18^2 = 324$	324	648
	0	`Eins,-++++++++++++zwei,`	$12^2 = 144$	0	**144**
3	2	`drei,++++++++++++++++++`	$18^2 = 324$	144	468
	1	`zwei,-++++++++++++drei,`	$12^2 = 144$	324	468
	0	`Eins,-+++zwei,-+++drei,`	$3^2 + 3^2 = 18$	0	**18**
4	3	`vier,++++++++++++++++++`	$18^2 = 324$	18	342
	2	`drei,-++++++++++++vier,`	$12^2 = 144$	144	288
	1	`zwei,-+++drei,-+++vier,`	$3^2 + 3^2 = 18$	324	342
	0	`Eins,-zwei,-drei,-vier,`	0	0	**0**
5	4	`fünf,++++++++++++++++++`	$18^2 = 324$	0	324
	3	`vier,-++++++++++++fünf,`	$12^2 = 144$	18	**162**
	2	`drei,-+++vier,-+++fünf,`	$3^2 + 3^2 = 18$	144	162
	1	`zwei,-drei,-vier,-fünf,`	0	324	324
	0	`Eins,-zwei,-drei,-vier,-fünf,`	∞		
6	5	`sechs,+++++++++++++++++`	$17^2 = 289$	162	451
	4	`fünf,-+++++++++++sechs,`	$11^2 = 121$	0	121
	3	`vier,-++fünf,-+++sechs,`	$2^2 + 3^2 = 13$	18	**31**
	2	`drei,-vier,-fünf,-sechs,`	∞		
7	6	`sieben,++++++++++++++++`	$16^2 = 256$	31	287
	5	`sechs,-+++++++++sieben,`	$9^2 = 81$	162	243
	4	`fünf,-+sechs,-++sieben,`	$1^2 + 2^2 = 5$	0	**5**
	3	`vier,-fünf,-sechs,-sieben,`	∞		
8	7	`in+++++++++++++++++++++`	$21^2 = 441$	5	446
	6	`sieben,-+++++++++++++in`	$13^2 = 169$	31	200
	5	`sechs,-+++sieben-,+++in`	$3^2 + 3^2 = 18$	162	180
	4	`fünf,-sechs,-sieben,-in`	0	0	**0**
9	8	`der++++++++++++++++++++`	$20^2 = 400$	0	400
	7	`in-+++++++++++++++++der`	$17^2 = 289$	5	294
	6	`sieben,-++++in-+++++der`	$4^2 + 5^2 = 41$	31	**72**
	5	`sechs,-sieben,-+in-+der`	$1^2 + 1^2 = 2$	162	164
	4	`fünf,-sechs,-sieben,-in-der`	∞		
10	9	`Schule+++++++++++++++++`	$17^2 = 289$	72	361
	8	`der-+++++++++++++Schule`	$13^2 = 169$	0	169
	7	`in-+++++der-+++++Schule`	$5^2 + 5^2 = 50$	5	55
	6	`sieben,-in-+der-+Schule`	$1^2 + 1^2 = 2$	31	**33**
	5	`sechs,-sieben,-in-der-Schule`	∞		
11	10	`wird+++++++++++++++++++`	$19^2 = 361$	33	394
	9	`Schule-++++++++++++wird`	$12^2 = 144$	72	216
	8	`der-++++Schule-++++wird`	$4^2 + 4^2 = 32$	0	32
	7	`in-+der-++Schule-++wird`	$1^2 + 2^2 + 2^2 = 9$	5	**14**
	6	`sieben,-in-der-Schule-wird`	∞		
12	**11**	`geschrieben.`	0	14	**14**
	10	`wird-geschrieben.`	0	33	33
	9	`Schule-wird-geschrieben.`	∞		

- Erläuterungen zur Tabelle: siehe Vorlesungsfolien.
- Die optimalen Umbruchstellen sind: $n = 12$, $\pi(12) = 11$, $\pi(11) = 7$, $\pi(7) = 4$, $\pi(4) = 0$
- Damit ergibt sich folgende optimale Formatierung:

```
Eins, zwei, drei, vier,
fünf,  sechs,   sieben,
in  der   Schule   wird
geschrieben.
```