

Algorithmen und Komplexität  
TIF 21A/B  
Dr. Bruno Becker

Übungsblatt 8: Datenkompression

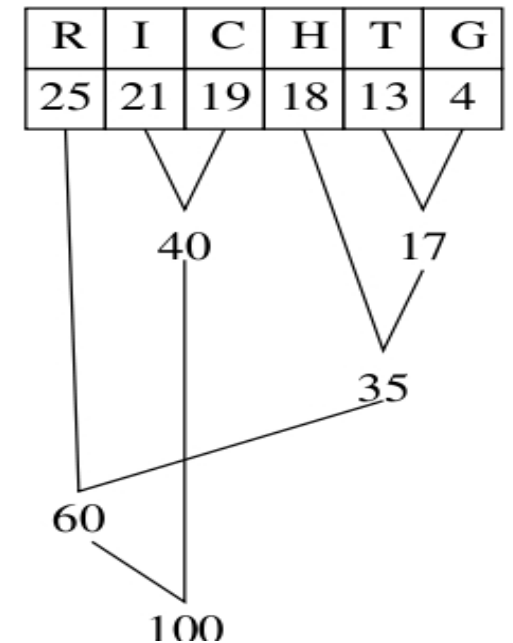
# Übungsblatt 8 – Aufgabe 1

Gegeben ist folgende Häufigkeitsverteilung für eine Menge von Zeichen

- Generieren Sie einen Huffman-Code für die angegebene Menge von Zeichen
- Decodieren Sie die folgende Nachricht (von links nach rechts) mit dem in a) generierten Code:

0010110100110100111

Zeichen	R	I	C	H	T	G
Häufigkeit	25	21	19	18	13	4
Code	00	10	11	010	0110	0111



b) 00 10 11 010 0110 10 0111  
R I C H T I G

# Übungsblatt 8 – Aufgabe 2

Gegeben sei die folgende Häufigkeitsverteilung für eine Menge von Zeichen:

Zeichen	E	L	F	G	I	O	R	V
Häufigkeit	2	2	2	2	2	1	1	1

- a) Generieren Sie einen Huffman-Code für die angegebene Menge von Zeichen und geben Sie entsprechenden Baum und Code-Tabelle an.
- b) Überführen Sie die Zeichenfolge: *VIELERFOLG* mittels des in a) generierten Code in eine Huffman-codierte Bitfolge.

Zeichen	E	L	F	G	I	O	R	V
Häufigkeit	2	2	2	2	2	1	1	1
	000	001	010	011	100	101	110	111

b) 111100000001000110010101001011

## Übungsblatt 8 – Aufgabe 3

Gegeben sei der Algorithmus für den Huffman-Code aus der Vorlesung:

- a. Welche Datenstruktur sollte man für den Aufbau des Baumes verwenden?
  - b. Welche Laufzeit hat in O-Notation der Algorithmus bei Verwendung der Datenstruktur aus a)?
- 
- a. Priority Queue (PQ) -> Min-Heap
  - b. Aufwand: N Zeichen:
    - Einfügen in PQ:  $N * O(\log N)$
    - (N-1) mal:
      - Die beiden Minima aus PQ entfernen  $O(\log N)$ ;
      - Neuen Baum bauen  $O(1)$ ,
      - Baum wieder in PQ einfügen  $O(\log N)$
    - Codebuch aufbauen  $N * O(\log N)$
    - Aufwand Gesamt  $O(N \log N)$

## Übungsblatt 8 – Aufgabe 4

Beweisen Sie, dass die beiden längsten Codewörter in einem Huffman-Code die gleiche Länge haben.

- Das längste Codewort  $C(w)$  gehört zu einem Eingangsbuchstaben  $w$  mit der geringsten Häufigkeit.
  - Daher wird dieser Eingangsbuchstabe  $w$  zusammen mit einem anderen Eingangsbuchstaben  $w'$  bei Huffman-Algorithmus im 1. Schritt zu einem Teilbaum zusammengefasst.
    - Die beiden Eingangsbuchstaben  $w$  und  $w'$  haben gleichen Vater.
    - Das Codewort des einen Wortes  $w$  ist  $\text{Codewort}(\text{Vater}) + 0$  (bzw.  $+1$ ),
    - Das Codewort von  $w'$  ist  $\text{Codewort}(\text{Vater}) + 1$  (bzw.  $0$ )
- ➔ Beide Codewörter sind gleich lang.