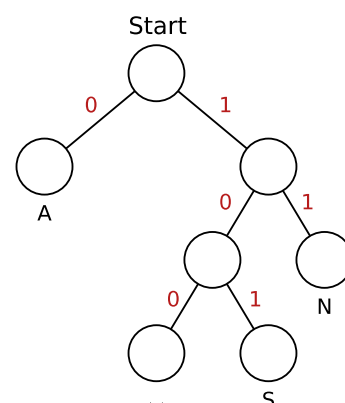


Arbeitsblatt Huffman-Codierung

David Huffman hat 1952 ein Verfahren entwickelt, mit welchem Zeichen platzsparender codiert werden können. Seine Idee ist, dass Zeichen, welche häufig im Text vorkommen, einen kürzeren Code erhalten, als Zeichen, welche selten im Text vorkommen.



Codebaum

Ein Codebaum ist eine Struktur mit einem Startknoten. Von diesem aus geht es entweder nach links oder rechts unten weiter. Eine 0 im Code bedeutet nach links gehen, eine 1 nach rechts gehen. Wenn ein Knoten mit einem Buchstaben erreicht wird, hat man ein Zeichen decodiert, man beginnt wieder von vorne.

Aufgabe 1 – Decodieren

Decodiere diese Bitfolge mit dem obenstehenden Codebaum. Das Symbol `_` steht für das Leerzeichen.

0111101011000110110101

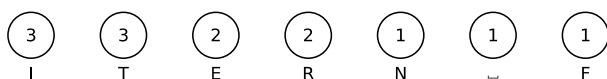
Erstellen eines Huffman-Baumes

Am Beispiel der Codierung des Texts «EINTRITT FREI» soll der Huffman-Algorithmus erläutert werden.

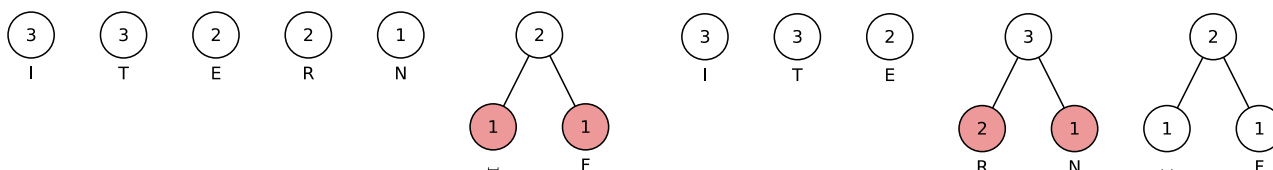
Zuerst zählt man, wie oft jedes Zeichen im Text vorkommt und erstellt eine Häufigkeitstabelle (siehe rechts).

Nun geht es darum, einen Codierungsbaum zu erstellen. Die Häufigkeiten der Buchstaben bilden je einen Knoten. Die Häufigkeit steht im Knoten, der Buchstaben darunter. Die Knoten werden nach aufsteigender Häufigkeit sortiert:

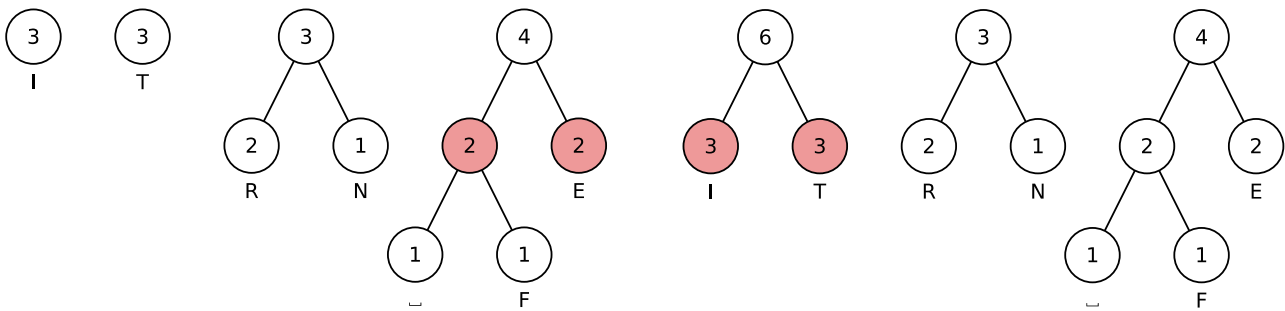
Zeichen	Häufigkeit
E	2
I	3
N	1
T	3
R	2
_	1
F	1



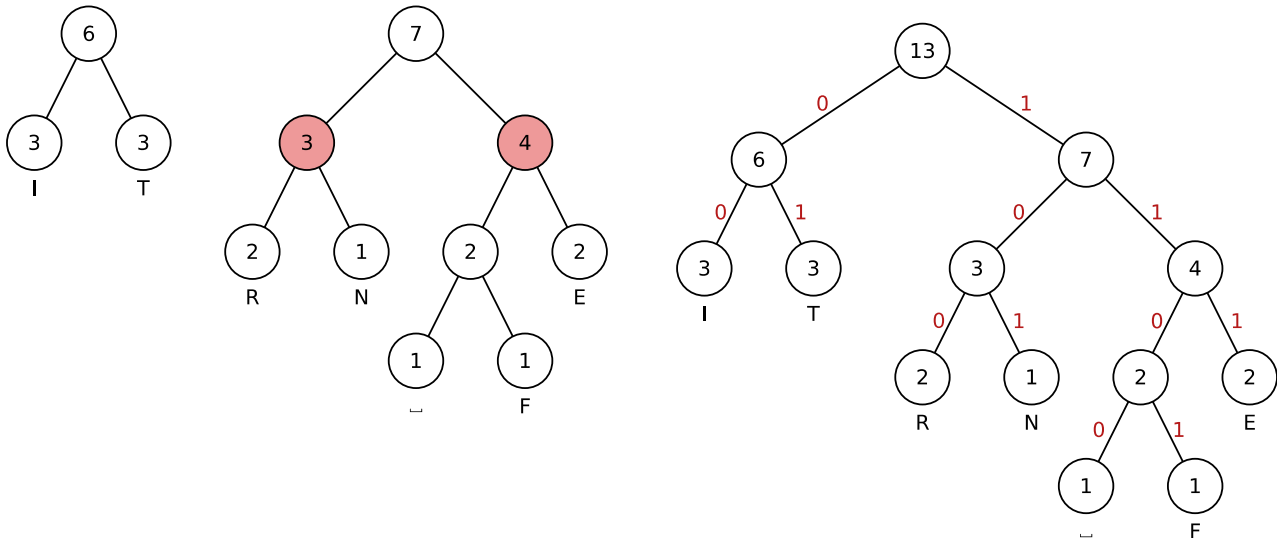
Nun werden die zwei Knoten mit den kleinsten Häufigkeiten an einen neuen Knoten angehängt. Der neue Knoten enthält die summierte Häufigkeit der ursprünglichen Knoten:



Die wird wiederholt, bis alle Knoten miteinander verbunden sind.



Wichtig ist, dass nach jedem Schritt die Knoten wieder neu sortiert werden.



Wenn der Baum fertig ist, werden alle Äste, welche nach links zeigen, mit einer «0» markiert, alle die nach rechts zeigen mit einer «1».

Nun kann eine Codierungstabelle erstellt werden, indem der Code für jedes Zeichen vom Baum abgelesen wird:

Zeichen	Code
I	00
T	01
R	100
N	101
E	111
⌵	1100
F	1101

Aufgabe Huffman-Codierung

Gehe für die untenstehenden Texte so vor:

- Erstelle eine Häufigkeitstabelle.
- Erstelle einen Huffman-Baum.
- Erstelle die Codierungstabelle.
- Codiere den Text mit dem Huffman-Baum.

Die Texte sind:

- «MISSISSIPPI»
- «EXTERNER EFFEKT»