

## Der Huffman-Algorithmus - Ein Beispiel

Am Beispiel der Codierung des Textes "MISSISSIPPI" soll der Huffman-Algorithmus erläutert werden.

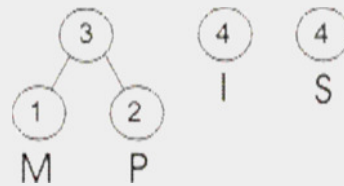
Zunächst zählt man, wie häufig jedes Zeichen im Text vorkommt und schreibt dies in eine Häufigkeitstabelle:

Zeichen	M	P	I	S
Häufigkeit	1	2	4	4

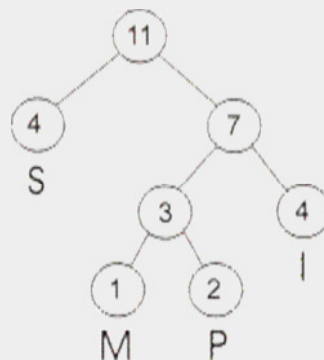
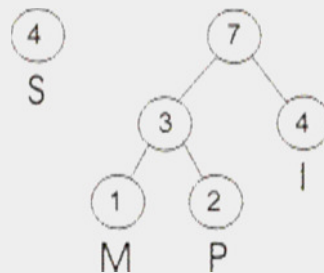
Nun wird jedes Zeichen zusammen mit seiner Häufigkeit als ein trivialer Baum, bestehend nur aus der Wurzel, aufgefasst und in einer Liste zusammengefasst.



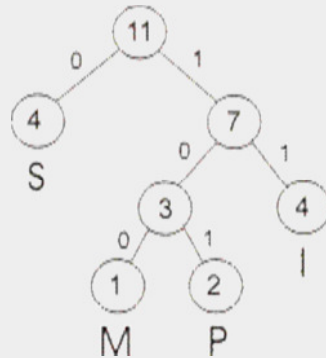
Nun werden die beiden Bäume mit den kleinsten Häufigkeiten aus der Liste zu einem neuen Baum zusammengefasst, bei dem die Wurzel mit der Summe der Häufigkeiten markiert wird.



Dies wird nun solange wiederholt, bis alle Zeichen in einem einzigen Baum sind.



Den nun entstandenen Baum nennt man den Huffman-Baum. Jede Verzweigung nach links wird mit einer "0", jede Verzweigung nach rechts mit einer "1" markiert. Anschließend kann man aus dem Codebaum den Code jedes Zeichens ablesen und eine Codetabelle mit dem Huffman-Code erstellen. Man sieht, dass tatsächlich das häufig vorkommende "S" und "I" einen kürzeren Code (1 bzw. 2 Binärzeichen) erhält als das selten vorkommende "M" (3 Binärzeichen).



#### Codetabelle:

Zeichen	Code
S	0
I	11
P	101
M	100

Mit Hilfe dieser Codetabelle kann nun der Text codiert werden. Statt 88 Bit bei einer Codierung mit ASCII benötigt man nun nur noch 21 Bit. Mit dem Huffman-Verfahren lässt sich ein Text typischerweise auf ca. 2/3 seiner ursprünglichen Größe reduzieren.

M	I	S	S	I	S	S	I	P	P	I
100	11	0	0	11	0	0	11	101	101	11