

Komprimierung

Übersicht

- Der Begriff Komprimierung

Erklärung des Begriffs Komprimierung und dessen Vor- und Nachteile

- Das Run Length Encoding

Zeichenkette:

AAABBEEEEENNNNND

Codefolge:

3ABB4E6ND

Bei mehr als zwei aufeinanderfolgende gleiche Zeichen werden diese durch Zahlenwerte ausgedrückt.

- Die Lempel Ziv Welch Komprimierung

Zeichenkette:

das haus ist das haus

Tabelle:

das	haus	ist
0	1	2

Codefolge:

0 1 2 0 1

Zeichenfragmente innerhalb von Zeichenketten werden durch Zahlen abgekürzt, und bei Wiederholung derselben wird die entsprechende Zahl verwendet.

- Die Huffman Codierung

Zeichenkette:

AABBBEEEEENDDDD

Informationen:

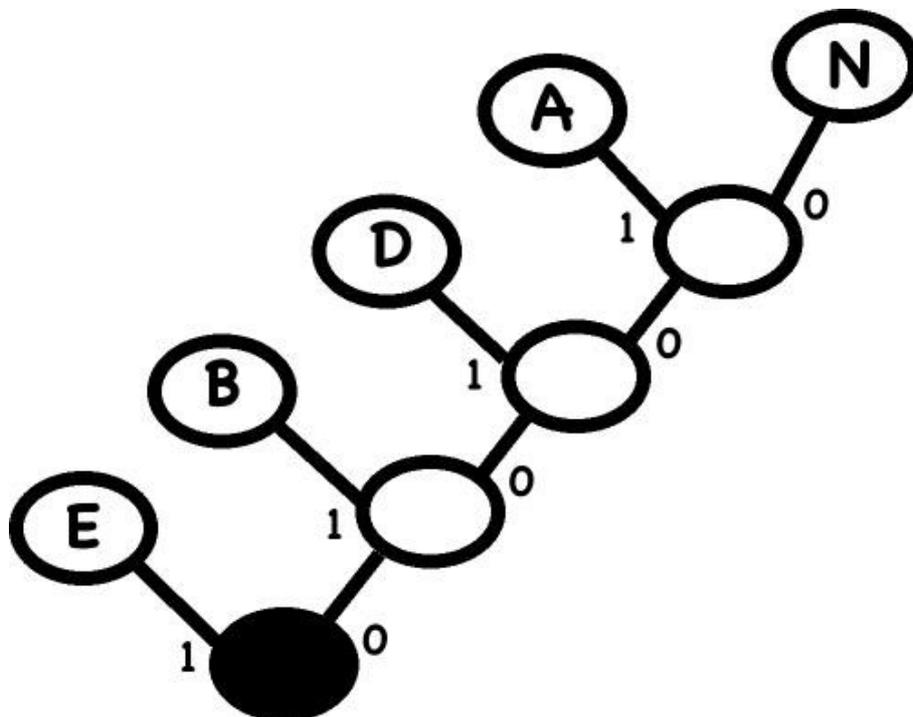
Anzahl Zeichen: 13

Buchstaben: A B E N D

Komprimierung:

Buchstabe:	Häufigkeit:
A	2
B	3
E	4
N	1
D	3

Huffman Baum:



Zeichen:

Buchstabe:	Kette:
A	0001
B	01
E	1
N	0000
D	001

Zeichenkette:

0001 0001 01 01 01 1 1 1 1 0000 001 001 001
(Leerzeichen sind übersichtshalber von uns hinzugefügt worden)

Nachdem analysieren der Zeichenkette, nach Anzahl und Häufigkeit der Zeichen, wird ein sogenannter Huffman - Baum mit den Informationen aus der Analyse erstellt. Dieser Baum richtet sich nach der Zeichenhäufigkeit so aus, dass das häufigste Zeichen ziemlich weit unten angesiedelt wird, und das seltenste nach ganz oben kommt. Durch „Abfahren“ des Baumes kann man mit Zahlenkombinationen aus 0 und 1 nun die Zeichen ausdrücken.

Komprimierung.

Referat von Martin Grabowski und Frieder Griebhammer

©2006