Kodierungsverfahren

1. Kodierung mit Kodewörtern konstanter Länge

- Bsp.: 7-Bit-ASCII-Code, 8-Bit-ASCII-Code, Unicode(16 bzw. 32 Bit)

- Wie viele Zeichen kann ich mit Kodewörtern der Länge n kodieren bzw. wie lang müssen die Kodeworte sein, wenn ich insgesamt k Zeichen kodieren will?
  Es gilt: es gibt 2^n Kodewörter der Länge n
  Wenn für k gilt: $2^{n-1}+1 \leq k \leq 2^n$ brauche ich Kodewörter der Länge n

- Beispiel: siehe Vorlesung
  1. Kodierung für die Nachricht $I = abac\ bdbc\ babb\ abbd$
  2. Kodierung für die Nachricht $I = ABRAKADABRA$

- Aber: Wenn für eine Nachricht I gilt $n > H(I)$, dann ist diese Kodierung nicht optimal!
  $\Rightarrow$ die Kodierung enthält Redundanz $\Rightarrow$ eventuell Komprimierung möglich

2. Verfahren zur Optimierung

- Komprimierung: Verringerung der Redundanz

- Varianten:
  logische $\Leftrightarrow$ physikalische Komprimierung
  symmetrische $\Leftrightarrow$ asymmetrische Komprimierung
  adaptive $\Leftrightarrow$ semiadaptive $\Leftrightarrow$ nichtadaptive Komprimierung
  verlustfreie $\Leftrightarrow$ verlustbehaftete Komprimierung

- Beispiele für Redundanzfreie Codes:
  Läuflängenkodierung
  Huffmann Kodierung
  Wörterbuchbasierte Komprimierung
  Arithmetische Kodierung

2.1. Läuflängenkodierung

- Folgen von sich wiederholenden Zeichen zusammenfassen, indem man die Folge nur einmal angibt und dazu die Anzahl der jeweiligen Wiederholungen

- Beispiel: siehe Vorlesung

2.2. Huffman-Kodierung

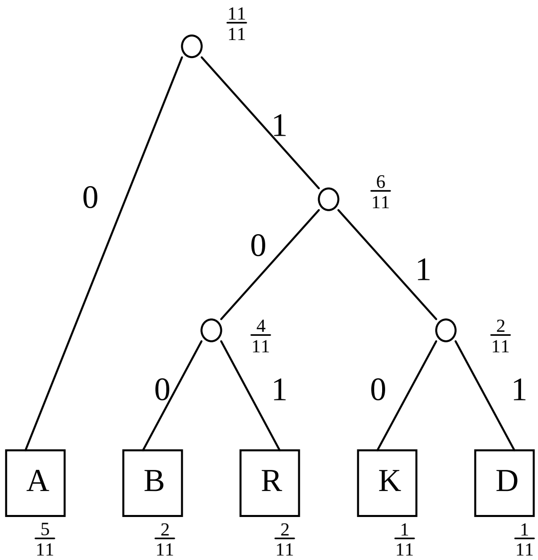
- nutzten, das die Zeichen mit unterschiedlicher Wahrscheinlichkeit auftreten
  ⇒ Platzeinsparung, wenn häufiger vorkommende Zeichen kürzer kodiert werden als seltenere

- Achtung: Kodierung sollte eindeutig dekodierbar sein
  ⇒ kein Code darf zugleich der Beginn eines anderen Codes sein
  ⇒ ein Verfahren: Huffman-Kodierung

- Algorithmus: Erzeugung durch Baumkodierung
  1. Ermittle die relative Häufigkeit der zu kodierenden Zeichen und sie als Blätter eines Baumes
  2. Fasse die beiden Zeichen c_i und c_j mit der geringsten Häufigkeit $f(c_i)$ und $f(c_j)$ zusammen zu einem neuen Knoten mit der Häufigkeit $f(c_i) + f(c_j)$
  3. Fahre fort, bis alle Blattknoten in einem gemeinsamen Baum verbunden sind
  4. Interpretiere Baum als Baumkodierung, d.h. Verzweigung nach links $\hat{=}$ 0
     Verzweigung nach rechts $\hat{=}$ 1

Beispiel: siehe auch Vorlesung
$I = ABRAKADABRA$

Zeichen	Absolute Häufigkeit	Relative Häufigkeit	Code
A	5	5/11	0
B	2	2/11	100
R	2	2/11	101
K	1	1/11	110
D	1	1/11	111



Aufgabe: Ermittle eine Huffman-Kodierung für folgende Nachrichten!
a) $I = KERNENERGIE$
b) $I = FISCHERS\ FRITZE$

2.3. Wörterbuchbasierte Komprimierung

- zu dem Text wird ein Wörterbuch erstellt
- z.B. mittels LZW-Algorithmus
- Beispiel: siehe Vorlesung

2.4. Arithmetische Kodierung

- Zeichen werden durch Häufigkeitsintervalle kodiert
- Zeichenfolgen werden durch geschachtelte Häufigkeitsintervalle kodiert

Beispiel: Gegebenen ist die folgende Tabelle mit Zeichen und den dazugehörigen Relativen Häufigkeiten (Wahrscheinlichkeit, dass das Zeichen gesendet wird). Wir geben Kodes für die Zeichenfolgen A, AC, ACC an.

Zeichen	Relative Häufigkeit
A	0,5
B	0,2
C	0,2
D	0,1

Zeichenfolge	Häufigkeitsintervall		Kodierung
A	$[0 ; 0{,}5]_{10}$	$[0{,}0 ; 0{,}1]_{2}$	1
AC	$[0{,}35 ; 0{,}45]_{10}$	$[0{,}010 ; 0{,}011]_{2}$	011
ACC	$[0{,}42 ; 0{,}44]_{10}$	$[0{,}0110 ; 0{,}0111]_{2}$	0111

Aufgabe: Nimm die Zeichen aus obiger Tabelle und berechne für die angegebenen Zeichnfolgen einen Kode mittels Arithmetischer Kodierung!

a) B b) D c) BD d) ACD

: