Aufgabe

Es werden die Ziffern $0, 1, \ldots, 9$, die Rechenzeichen $+, -$ und ein Pausenzeichen $\Box$ gesendet. Im statistischen Mittel sind die Ziffern gleich wahrscheinlich, während auf je vier Ziffern ein Rechenzeichen mit $P(+) = P(-)$ und auf je zwölf Ziffern ein Pausensymbol kommt.

1. Bestimmen Sie die Wahrscheinlichkeiten $P(0), \ldots, P(\Box)$!
2. Kodieren Sie die Quellensymbole $0, \ldots, \Box$ nach HUFFMAN unter Verwendung von Ternärzeichen $+U, 0, -U$ (Spannung 0 soll möglichst oft auftreten)!
3. Zeichnen Sie den Kodebaum; Test auf unverzögerte Dekodierbarkeit.
4. Bestimmen Sie die Koderedundanz!

1 Wahrscheinlichkeiten

Die einzelnen Wahrscheinlichkeiten lassen sich wie folgt klassifizieren:

$P_{\mathrm{Z}} \ldots$ Wahrscheinlichkeit einer Ziffer
$P_{\mathrm{R}} \ldots$ Wahrscheinlichkeit eines Rechenzeichens
$P_{\Box} \ldots$ Wahrscheinlichkeit des Pausenzeichens

Die Summe aller Wahrscheinlichkeiten muß Eins ergeben.

$$1 \stackrel{!}{=} 10P_{\mathrm{Z}} + 2P_{\mathrm{R}} + P_{\Box}$$

Weiterhin gilt:

$$P_{\mathrm{R}} = \frac{1}{8}P_{\mathrm{Z}} \quad \text{und} \quad P_{\Box} = \frac{1}{12}P_{\mathrm{Z}}.$$

Diese beiden Gleichungen in die Normierungsbedingung eingesetzt ergibt

$$P_{\mathrm{Z}} = \frac{24}{248}, \quad P_{\mathrm{R}} = \frac{3}{248} \quad \text{und} \quad P_{\Box} = \frac{2}{248}.$$

2 Kodierung

Die Symbole werden in einer Tabelle nach fallenden Wahrscheinlichkeiten angeordnet. Dann wird bei der Ternärkodierung den letzten drei Zeichen vorerst A, B, und C zugeordnet. Damit kann später dem häufigsten Buchstaben die Spannung 0 zugewiesen werden. Die letzten drei Zeichen werden nun zu einem Blocksymbol zusammengefaßt. Danach wird die gesamte Liste wieder nach fallenden Wahrscheinlichkeiten geordnet. Bei gleichen Wahrscheinlichkeiten wird wegen geringerer Varianz das Blocksymbol oben eingeordnet [Lo95]. Der Kodierungsvorgang ist in Tabelle 1 dargestellt. Es ist $P^* = 248P$.

Nun kann den einzelnen Symbolen ein Ternärkode zugewiesen werden. Dabei wird „hinten“ begonnen. Außerdem wird in Tabelle 2 noch die mittlere Kodewortlänge L, die Kodeentropie H_{C} sowie die relative Häufigkeit der Zeichen A, B und C bestimmt.

Für einen dreiwertigen Kode gilt:

$$L = \sum_{i=1}^{I} P(i) l_i, \tag{1}$$

i	$P^*(i)$	i_1	$P^*(i)$	i_2	$P^*(i)$	i_3	$P^*(i)$	i_4	$P^*(i)$	i_5	$P^*(i)$
0	24	0	24	$89+-\square$	56	567	72	234	72	$0189+-\square$	104
1	24	1	24	0	24	$89+-\square$	56	567	72	234	72
2	24	2	24	1	24	0	24	$89+-\square$	56	567	72
3	24	3	24	2	24	1	24	0	24		
4	24	4	24	3	24	2	24	1	24		
5	24	5	24	4	24	3	24				
6	24	6	24	5	24	4	24				
7	24	7	24	6	24						
8	24	8	24	7	24						
9	24	9	24								
$+$	3	$+-\square$	8								
$-$	3										
$\square$	2										

Tabelle 1: HUFFMAN-Kodierung

i	$P^*(i)$	C(i)	$l_i P^*(i)$	$-P \,\mathrm{ld}\, P$	$n_A P^*(i)$	$n_B P^*(i)$	$n_C P^*(i)$
0	24	AB	48	0.326	24	24	0
1	24	AC	48	0.326	24	0	24
2	24	BA	48	0.326	24	24	0
3	24	BB	48	0.326	0	48	0
4	24	BC	48	0.326	0	24	24
5	24	CA	48	0.326	24	0	24
6	24	CB	48	0.326	0	24	24
7	24	CC	48	0.326	0	0	48
8	24	AAA	72	0.326	72	0	0
9	24	AAB	72	0.326	48	24	0
+	3	AACA	12	0.077	9	0	3
−	3	AACB	12	0.077	6	3	3
□	2	AACC	8	0.056	4	0	4
Σ	248		560	3.470	235	171	154

Tabelle 2: Ternärkode

$$H = -\sum_{i=1}^{I} P(i) \,\mathrm{ld}\, P(i) \,\mathrm{bit\ und} \tag{2}$$

$$H_C = L \,\mathrm{ld}\, 3 \,\mathrm{bit}. \tag{3}$$

Mit Zahlenwerten ergibt sich

$L = \frac{560}{248}$ Zeichen/Symbol ≈ 2.258 Zeichen/Symbol,

$H = 3.470$ bit/Symbol und

$H_C = 3.579$ bit/Symbol.

Da das Zeichen A im Kode am häufigsten vorkommt, wird ihm die Spannung Null zugewiesen. B wird dann z. B. durch $-U$ und C durch $+U$ ersetzt.

3 Kodebaum

In Abbildung 1 ist der Kodebaum dargestellt. Durch Rundungsfehler ist nicht immer die Blocksymbol-Wahrscheinlichkeit gleich der Summe der Einzelsymbol-Wahrscheinlichkeiten. Die Kanten, die jeweils nach rechts oben führen entsprechen der Spannung Null, usw.

Unverzögerte Dekodierbarkeit kann bei variabler Kodewortlänge z. B. durch folgende Bedingung gewährleistet werden. Kein Kodewort darf der Beginn eines anderen sein, d.h., von Knoten, die das Ende eines Kodeworts markieren, darf keine weitere Kante fortführen. Bei dem dargestellten Baum ist das der Fall. Der Kode kann also unverzögert dekodiert werden.

4 Koderedundanz

Die Koderedundanz bestimmt sich wie folgt:

$$R_C = H_C - H. \tag{4}$$

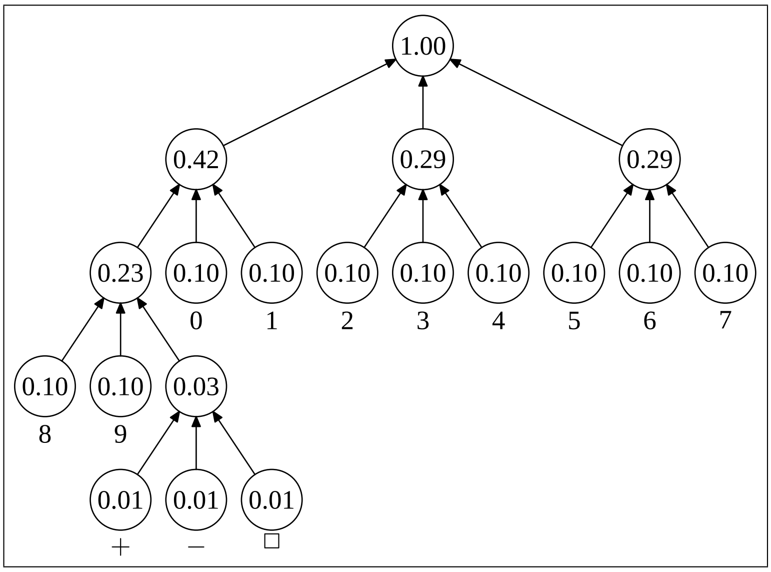

Abbildung 1: Kodebaum

Mit den oben errechneten Werten erhält man

$R_{\mathrm{C}} = (3.579 - 3.470)$ bit/Symbol $= 0.109$ bit/Symbol.

Literatur

[Lo95] Lochmann, D.: Digitale Nachrichtentechnik. Verlag Technik GmbH, Berlin 1995, S. 368.