

Datenkompression

Sommersemester 2015

Übungsblatt 1

Prof. Dr. E. Ohlebusch

Institut für Theoretische Informatik

T. Beller

Ausgegeben am 23.04.2015

Besprechung am 30.04.2015

Aufgabe 1.1

Wir betrachten die folgenden Mengen von Codewörtern:

$$\begin{aligned} M_1 &:= \{01, 10, 000, 011, 100, 111\}, & M_2 &:= \{0, 10, 11, 101\}, \\ M_3 &:= \{00, 01, 11, 100, 101\}, & M_4 &:= \{00, 11, 010, 0100\}. \end{aligned}$$

- Welche Mengen erfüllen die Kraft-McMillan-Ungleichung?
- Welche Mengen sind Bildbereich eines eindeutig dekodierbaren Codes, und warum?
- Geben Sie für jede Menge M_i einen Präfixcode an, welcher dieselben Codewortlänge wie M_i hat. Für welche Mengen ist dies nicht möglich?

Aufgabe 1.2

Auf dem Alphabet $\{a, e, i, o, u\}$ seien zwei Wahrscheinlichkeitsfunktionen P, P' definiert:

x	a	e	i	o	u
$P(x)$	1/4	7/16	1/8	1/8	1/16
$P'(x)$	5/32	3/8	5/32	3/16	1/8

- Bestimmen Sie für beide Fälle einen Präfixcode nach dem Shannon-Verfahren.
- Bestimmen Sie für beide Fälle einen Präfixcode nach dem Shannon-Fano-Verfahren (wie am Anfang von Aufgabe 1.3 erklärt).
- Bestimmen Sie für beide Fälle einen Präfixcode nach dem Huffman-Verfahren.
- Berechnen Sie die erwarteten Codewortlängen der Codes aus den Teilaufgaben a) bis c).

Aufgabe 1.3

Beim Shannon-Fano-Algorithmus werden wiederholt Listen (c_1, \dots, c_n) von Symbolen in zwei Teile $M_1 = (c_1, \dots, c_k)$ und $M_2 = (c_{k+1}, \dots, c_n)$ gespalten, sodass $|P(M_1) - P(M_2)|$ minimal ist. Unter Umständen existiert aber eine andere Partitionierung $\tilde{M}_1 \cup \tilde{M}_2 = \{c_1, \dots, c_n\}$, $\tilde{M}_1 \cap \tilde{M}_2 = \emptyset$, die „gleichmäßiger“ ist, also: $|P(\tilde{M}_1) - P(\tilde{M}_2)| < |P(M_1) - P(M_2)|$. Wir betrachten eine Variante des Algorithmus, die unter *allen* Partitionierungen stets eine „gleichmäßigste“ wählt.

- Geben Sie eine Wahrscheinlichkeitsfunktion P an, bei der der modifizierte Algorithmus einen Code mit größerer erwarteter Codewortlänge generiert als der Shannon-Fano-Algorithmus.
- (*schwieriger!*) Geben Sie ein Beispiel an, bei dem der modifizierte Algorithmus einen Code mit kleinerer erwarteter Codewortlänge generiert als der Shannon-Fano-Algorithmus.