

Informationstheorie

Lösung 7

7.1 Optimale Codes und Huffman-Algorithmus

a) Der Huffman-Baum ist in Abbildung 1 dargestellt. Die Buchstaben werden codiert durch

a	b	c	d	e	f	g	h
0100	0101	1110	1111	011	110	00	10

und $w = c e b g a h d f g$.

b) Für $D > 2$ fasst der Algorithmus jeweils D Knoten zusammen. Dabei kann es geschehen, dass für die Wurzel nicht mehr D Knoten zur Verfügung stehen. In diesem Fall wird der Code sicher nicht optimal sein. Die Lösung ist, zu Beginn Dummy-Blätter mit WSK 0 hinzuzufügen. Da in jedem Schritt des Algorithmus $D - 1$ Knoten wegfallen, müssen total $1 + k(D - 1)$ Knoten vorhanden sein, wobei k die Anzahl Schritte ist. Also werden soviele Dummy-Knoten dazugefügt, bis die Anzahl Knoten von der gewünschten Form ist. Die Dummy-Knoten treten nie auf (WSK 0), weshalb sie die mittlere Codelänge nicht verändern.

c) Wir benötigen $1 + k(D - 1) = 1 + 2k$ Blätter, d.h. wir müssen ein Dummyblatt einführen. Zwei mögliche optimale Codes sind

WSK	0.3	0.2	0.2	0.1	0.1	0.1	0
Code 1	0	10	11	12	20	21	22
Code 2	0	1	20	21	220	221	222

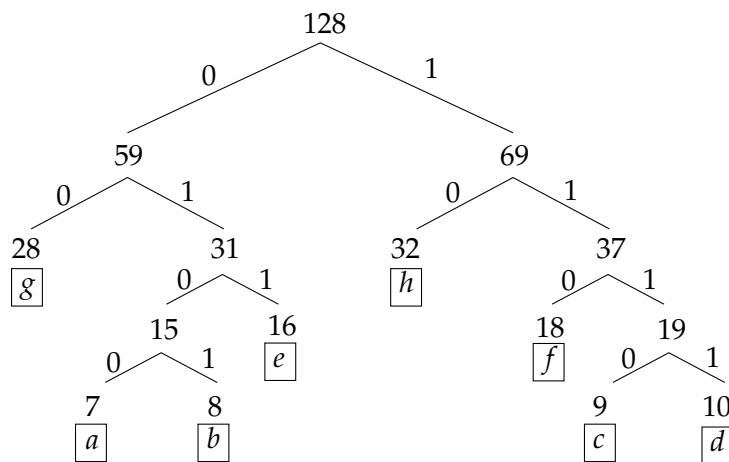
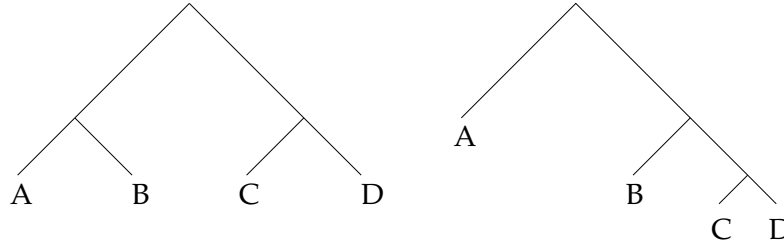


Abbildung 1: Der Huffman-Baum zu Aufgabe 7.1.a). Wahrscheinlichkeiten sind in $\frac{1}{128}$ angegeben.

- d) Damit beide Codes optimal sind, muss die Verteilung so gewählt werden, dass der Huffman-Algorithmus beide erzeugen könnte. Bei der zweiten Codierung (und somit wahrscheinlich auch bei der ersten) werden zuerst C und D zusammengefasst. Das heißt sie müssen die kleinsten Wahrscheinlichkeiten haben. Die erste Codierung fasst dann zuerst A und B zusammen, die zweite Codierung hingegen B und den neuen Knoten, welcher aus C und D entstanden ist.



Das heißt aber nichts anderes, als dass $P(B) \leq P(A) = P(C) + P(D)$ gelten muss. Eine mögliche Verteilung ist $P(A) = P(B) = 1/3$ und $P(C) = P(D) = 1/6$.

7.2 Bedingte Entropie und Fehlerwahrscheinlichkeit

- a) Sei $1 \leq j \leq N$. Zunächst stellen wir fest, dass X gegeben $Y = j$ uniform auf 2^j Elementen verteilt ist, was

$$H(X|Y = j) = j \quad \text{für } j = 1, \dots, N$$

impliziert. Daraus erhalten wir

$$H(X|Y) = \sum_{j=1}^N P_Y(j) H(X|Y = j) = \frac{1}{N} \sum_{j=1}^N j,$$

wobei wir verwendet haben, dass Y uniform auf $\{1, \dots, N\}$ ist. Daraus folgt

$$H(X|Y) = \frac{N+1}{2}. \quad (1)$$

- b) Die Fano-Ungleichung $h(P_e) + P_e \log_2(|\mathcal{X}| - 1) \geq H(X|Y)$ besagt in diesem Fall wegen (1) und $|\mathcal{X}| = |\{0, 1\}^N| = 2^N$

$$h(P_e) + P_e \log_2(2^N - 1) \geq \frac{N+1}{2}. \quad (2)$$

Diese Bedingung ist äquivalent zu

$$f_N(P_e) \geq 0$$

wobei $f_N : [0, 1] \rightarrow [0, 1]$ durch

$$f_N(p) := h(p) + p \log_2(2^N - 1) - \frac{N+1}{2}$$

definiert ist. Aus Abbildung 2 können wir daher ablesen, dass

$$\begin{aligned} P_e &\geq 0.39 && \text{für } N = 4 \\ P_e &\geq 0.43 && \text{für } N = 8. \end{aligned}$$

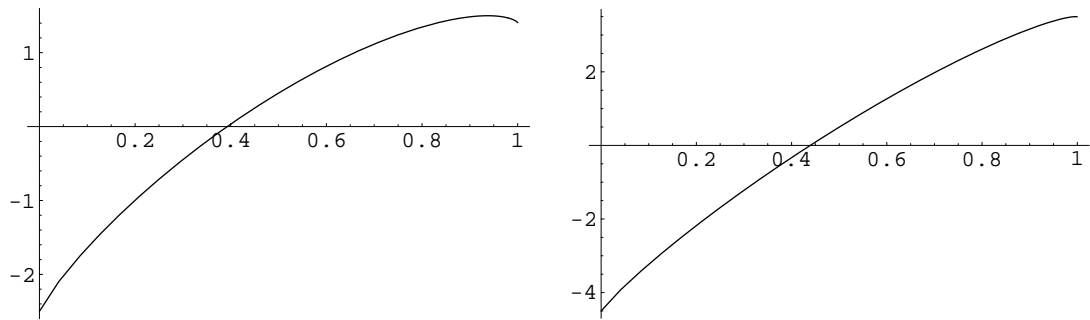


Abbildung 2: Die Funktionen f_4 (links) und f_8 (rechts)

Um den Grenzfall $N \rightarrow \infty$ zu untersuchen, formen wir (2) um zu

$$P_e \geq \frac{N + 1 - 2h(P_e)}{2 \log_2(2^N - 1)}.$$

Es ist leicht nachzuprüfen, dass der Term auf der rechten Seite dieser Ungleichung gegen $\frac{1}{2}$ strebt.

- c) Gegeben dass $Y = j$, ist die beste Strategie um X zu schätzen, zufällig eines der $x \in \{0, 1\}^N$ mit $P_{X|Y}(x|j) > 0$ zu wählen. Gegeben $Y = j$ ist dabei die Erfolgswahrscheinlichkeit gleich 2^{-j} . Die mittlere Erfolgswahrscheinlichkeit P_s über das ganze Experiment ist dann

$$P_s = \frac{1}{N} \sum_{j=1}^N 2^{-j} = \frac{1}{2N} \sum_{j=0}^{N-1} 2^{-j} = \frac{1}{2N} \cdot \frac{1 - (\frac{1}{2})^N}{1 - \frac{1}{2}}$$

Daraus erhalten wir schliesslich

$$P_e = 1 - P_s = 1 - \frac{1}{N} \left(1 - \left(\frac{1}{2}\right)^N\right)$$

und

$$P_e = 0.7656 \quad \text{für } N = 4$$

$$P_e = 0.8755 \quad \text{für } N = 8,$$

woraus wir sehen, dass die Fano-Ungleichung hier nur eine relativ schlechte untere Schranke an P_e gibt.