Informationstheorie, FS 2017 Dr. Luis Haug
ETH Zürich, Dept. Informatik Nico Gorbach

Lösungen zu Übungsblatt 4

4.1

Ein möglicher Huffman-Code für X ist gegeben durch

X	x_1	x_2	x_3	x_4	x_5	x_6	x_7
C	0	10	110	11100	11101	11110	11111

Die erwartete Länge ist $L_C = 0.49 + 0.26 \cdot 2 + 0.12 \cdot 3 + 0.13 \cdot 5 = 2.02$ bits pro Symbol.

4.2

Eine mögliche Verteilung für $\{0, 10, 11\}$ ist $\mathbf{p} = (0.5, 0.3, 0.2)$. Weder $\{00, 01, 10, 110\}$ noch $\{01, 10\}$ sind die Mengen der Codewörter eines Huffman-Codes, denn unter den längsten Codewörtern eines Huffman-Codes gibt es immer zwei, die sich nur im letzten Bit unterscheiden (ausserdem sieht man zumindest bei der letzten Menge sofort, dass sie nicht optimal ist, was die erwartete Länge angeht).

4.3

$(1, 2, 2)$ sind die Längen des Huffman-Codes $\{0, 10, 11\}$ z.B. für $\mathbf{p} = (0.5, 0.3, 0.2)$. Für $(2, 2, 3, 3)$ und $(1, 3, 3, 3, 4, 5)$ gibt es keinen Huffman-Code mit diesen Wortlängen. Im letzten Fall sieht man das direkt daran, dass es ein Code mit diesen Wortlängen ein eindeutiges längstes Wort haben müsste, was bei einem Huffman-Code nicht der Fall ist. Wären $(2, 2, 3, 3)$ die Wortlängen eines Huffman-Codes, dann müsste das auch für $(2, 2, 2)$ der Fall sein (aufgrund der rekursiven Definition von Huffman-Codes), und ebenfalls für $(2, 1)$; im letzten Fall gibt es aber wiederum ein eindeutiges längstes Wort.

Allgemein sind $(l_1, \ldots, l_n)$ mit $n \geq 2$ genau dann die Wortlängen eines Huffman-Codes, wenn entweder (a) $n = 2$ und $(l_1, l_2) = (1, 1)$, oder (b) $n > 2$ und $l_{n-1} = l_n$ und $(l_1, \ldots, l_{n-2}, l_n - 1)$ die Wortlängen eines Huffman-Codes sind (wobei wir hier davon ausgehen, dass die Wörter der Länge nach geordnet sind, d.h. $l_{i-1} \leq l_i$). Das kann man leicht mit einem rekursiven Algorithmus überprüfen.

4.4

Z.B. für $\mathbf{p} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{9}, \frac{1}{18})$ sind sowohl $(2, 2, 2, 3, 3)$ und $(1, 2, 3, 4, 4)$ Wortlängen möglicher Huffman-Codes. Die erwartete Länge ist jeweils gleich $\frac{39}{18}$.

4.5

Sei zunächst C' irgendein optimaler präfixfreier Code für X. Dann ist klar, dass es für jedes Codewort w_0 mit maximaler Länge ein weiteres Codewort w_1 geben muss, so dass sich w_0

und w_1 nur im letzten Bit unterscheiden (sonst könnte man w_0 um ein Bit kürzen und hätte damit einen Code mit geringerer Länge als C', im Widerspruch zur Optimalität). Ausserdem ist klar, dass ein Paar $x_{i_0} \neq x_{i_1}$ mit kleinster Wahrscheinlichkeit $p_{i_0} + p_{i_1}$ existiert, so dass $C'(x_{i_0})$ und $C'(x_{i_1})$ maximale Länge haben (wäre nämlich für alle x_i, x_j, für die $C'(x_i)$, $C'(x_j)$ maximale Länge haben, die Wahrscheinlichkeit $p_i + p_j > p_{i_0} + p_{i_1}$, dann könnte man durch Vertauschen von Codewörtern wiederum einen kürzeren Code erzeugen). Indem wir C' durch Vertauschen von Codewörtern mit maximaler Länge so modifizieren, dass x_{i_0} auf w_0 und x_{i_1} auf w_1 abgebildet wird, erhalten wir einen Code C^* mit der geforderten Eigenschaft.

4.6

Sei z.B. (X, Y) eine Zufallsvariable mit der gemeinsamen Verteilung $P_{XY}(0,0) = \frac{1}{10}$, $P_{XY}(1,0) = \frac{1}{10}$, $P_{XY}(0,1) = \frac{4}{5}$, $P_{XY}(1,1) = 0$. Die Marginalverteilung von X ist dann gegeben durch $P_X(0) = \frac{1}{5}$ und $P_X(1) = \frac{4}{5}$, und damit ist $H(X) = \frac{1}{5}\log 5 + \frac{4}{5}\log\frac{5}{4} \approx 0.72$; andererseits gilt $P(X = 0|Y = 0) = \frac{1}{2} = P(X = 1|Y = 0)$, und damit ist $H(X|Y = 0) = 1$.