4. Übung aus Informations- und Codierungstheorie SS 2016

1. Sind $X_1, \dots, X_n \quad n$ binäre Zufallsvariable, so versteht man unter einem *run* eine Teilfolge $X_i, \dots, X_j$, $i \leq j$, sodass gilt $X_i = X_k$, $\forall\ i \leq k \leq j$ und $i = 1 \ \vee \ X_{i-1} \neq X_i$ und $j = n \ \vee \ X_j \neq X_{j+1}$. Sei $\mathbf{R} := (R_1, \dots, R_N)$ der Vektor der *run*-Längen in $\mathbf{X} := (X_1, \dots, X_n)$ (d.h. N ist die Anzahl der *runs* und daher auch eine Zufallsvariable). Man zeige $H(\mathbf{R}) \leq H(X_1, \dots, X_n) \leq H(\mathbf{R}) + 1$.

   Lösung : $\mathbf{R} = \varphi(\mathbf{X}) \ \Rightarrow \ H(\mathbf{R}) = H(\varphi(\mathbf{X})) \leq H(\mathbf{X})$.

   Kennt man umgekehrt $\mathbf{R}$, so entspricht $\mathbf{X} = (X_1, \dots, X_n)$ einer Folge $(a_1, \dots, a_{R_1}, b_{R_1+1}, \dots, b_{R_1+R_2}, \dots)$, von der man allerdings nicht weiß, ob den a-s eine Eins und den b-s eine Null entspricht oder umgekehrt. Kennt man ein einziges X_i , so kann auch dies beantwortet werden und daher gilt $\mathbf{X} = \psi(\mathbf{R}, X_i)$. Somit gilt

$$H(\mathbf{X}) = H(\psi(\mathbf{R}, X_i)) \leq H(\mathbf{R}, X_i) \leq H(\mathbf{R}) + H(X_i) \leq H(\mathbf{R}) + 1\,.$$

   Die letzte Ungleichung gilt, da X_i eine binäre Zufallsvariable ist.

2. Man zeige, dass für Markoffketten $X_0, \dots, X_n$ gilt

$$H(X_0|X_n) \geq H(X_0|X_{n-1}). \tag{1}$$

   Beweis: Kann man zeigen, dass $X_0 \to X_{n-1} \to X_n$ eine Markoffkette ist, so ist auch $X_n \to X_{n-1} \to X_0$ eine, und daraus folgt bekanntlich $H(X_0|X_{n-1}) = H(X_0|X_{n-1}, X_n) \leq H(X_0|X_n)$, also Ungleichung (1).

   Dass $X_0 \to X_{n-1} \to X_n$ wirklich eine Markoffkette ist, sieht man mit $p(x_0^n) := P(X_0^n = x_0^n)$ und analogen Bezeichnungen so

$$\begin{aligned}
p(x_n|x_{n-1}, x_0) &= \frac{p(x_n, x_{n-1}, x_0)}{p(x_{n-1}, x_0)} = \frac{1}{p(x_{n-1}, x_0)} \sum_{x_1^{n-2}} \frac{p(x_0^n)}{p(x_0^{n-1})}\, p(x_0^{n-1}) \\
&= \frac{1}{p(x_{n-1}, x_0)} \sum_{x_1^{n-2}} p(x_n|x_0^{n-1})\, p(x_0^{n-1}) \\
&= \frac{1}{p(x_{n-1}, x_0)} \sum_{x_1^{n-2}} p(x_n|x_{n-1})\, p(x_0^{n-1}) = \frac{p(x_n|x_{n-1})}{p(x_{n-1}, x_0)} \sum_{x_1^{n-2}} p(x_0^{n-1}) \\
&= \frac{p(x_n|x_{n-1})}{p(x_{n-1}, x_0)}\, p(x_{n-1}, x_0) = p(x_n|x_{n-1}) \quad \forall\ x_o, x_{n-1}, x_n.
\end{aligned}$$

3. Man zeige: $I((X_1,\dots,X_n),Y) = I(X_1,Y) + \sum\limits_{i=2}^{n} I(X_i,Y|X_1,\dots,X_{i-1})$.

   Beweis: $I(X_1^n,Y) = H(X_1^n) - H(X_1^n|Y)$. Aber es gilt bekanntlich $H(X_1^n) = H(X_1) + \sum\limits_{i=2}^{n} H(X_i|X_1^{i-1})$. Weiters gilt

$$
\begin{aligned}
H(X_1,X_2|Y) &= \sum_{x_1,x_2,y} p(x_1,x_2,y)\ \log p(x_1,x_2|y) \\
&= \sum_{x_1,x_2,y} p(x_1,x_2,y)\ \log(p(x_1|y)\,p(x_2|x_1,y)\,) \\
&= \sum_{x_1,y} \log(p(x_1|y) \sum_{x_2} p(x_1,x_2,y) + \sum_{x_1,x_2,y} p(x_1,x_2,y) \log p(x_2|x_1,y) \\
&= H(X_1|Y) + H(X_2|X_1,Y), \quad \text{und vollständige Induktion liefert}
\end{aligned}
$$

   $H(X_1^n|Y) = H(X_1|Y) + \sum\limits_{i=2}^{n} H(X_i|X_1^{i-1},Y)$. Oben eingesetzt ergibt das

$$
\begin{aligned}
I(X_1^n,Y) &= H(X_1) - H(X_1|Y) + \sum_{i=2}^{n} \left[H(X_i|X_1^{i-1}) - H(X_i|X_1^{i-1},Y)\right] \\
&= I(X_1,Y) + \sum_{i=2}^{n} I(X_i,Y|X_1^{i-1}).
\end{aligned}
$$

4. Für eine Wahrscheinlichkeitsverteilung $P = (p_1,\dots,p_m)$ zeige man

   (a) sind $p_{(1)} \le \dots \le p_{(m)}$ die der Größe nach geordneten p_i so gilt für $0 \le f_1 \le \dots \le f_m$ und jede Permutation $\pi_1,\dots\pi_m$ von $1,\dots,m$

$$
\sum_{i=1}^{m} f_i\, p_{(m+1-i)} \le \sum_{i=1}^{m} f_i\, p_{\pi_i}, \tag{2}
$$

   (b) durch vollständige Induktion

$$
H^*(P) \le \sum_{i=1}^{m-1} i\, p_{(m+1-i)} + (m-1)\, p_{(1)} = \sum_{i=1}^{m} i\, p_{(m+1-i)} - p_{(1)}, \tag{3}
$$

   und geben Sie einen Binärcode an, dessen mittlere Wortlänge mit der oberen Schranke in Ungleichung (3) übereinstimmt.

Beweis: Mit $d_1 := f_1,\ d_i := f_i - f_{i-1},\ 2 \le i \le m$ gilt

$$\begin{aligned}
\sum_{i=1}^{m} f_i\, p_{(m+1-i)} &= \sum_{i=1}^{m} p_{(m+1-i)} \sum_{j=1}^{i} d_j = \sum_{j=1}^{m} d_j \sum_{i=j}^{m} p_{(m+1-i)} \\
&= \sum_{j=1}^{m} d_j \sum_{k=1}^{m+1-j} p_{(k)} \le \sum_{j=1}^{m} d_j \sum_{k=1}^{m+1-j} p_{\pi_k} = \sum_{j=1}^{m} d_j \sum_{i=j}^{m} p_{\pi_{m+1-i}} \\
&= \sum_{i=1}^{m} p_{\pi_{m+1-i}} \sum_{j=1}^{i} d_j = \sum_{i=1}^{m} f_i\, p_{\pi_{m+1-i}}. \qquad (4)
\end{aligned}$$

Der Code mit $c(\omega_{(m)}) := 0, c(\omega_{(m+1-i)}) = (\mathbf{1}_1^{i-1}, 0), 2 \le i \le m-1$ und $c(\omega_{(1)}) := (\mathbf{1}_1^{m-1})$ hat die mittlere Länge $\sum_{i=1}^{m-1} i\, p_{(m+1-i)} + (m-1)\, p_{(1)}$.

Für $m = 1$ gilt $0 = H^*(P) = p_1 - p_1$, und aus der Induktionsannahme für $m-1$ folgt mit $\widetilde{P} = (p_{(1)} + p_{(2)}, p_{(3)}, \dots, p_{(m)})$ und den der Größe nach geordneten Wahrscheinlichkeiten $\widetilde{p}_{(1)} \le \dots \le \widetilde{p}_{(m-1)}$ von $\widetilde{P}$ unter Verwendung von (3) mit $f_i := i,\ 1 \le i \le m-2, f_{m-1} := m-2$ und dem Satz von Huffman

$$\begin{aligned}
H^*(P) &= H^*(\widetilde{P}) + (p_{(1)} + p_{(2)}) \\
&\le \sum_{i=1}^{m-2} i\, \widetilde{p}_{(m-i)} + (m-2)\, \widetilde{p}_{(1)} + (p_{(1)} + p_{(2)}) \\
&\le \sum_{i=1}^{m-2} i\, p_{(m+1-i)} + (m-2)\, (p_{(1)} + p_{(2)}) + p_{(1)} + p_{(2)} \\
&= \sum_{i=1}^{m-1} i\, p_{(m+1-i)} + (m-1)\, p_{(1)} = \sum_{i=1}^{m} i\, p_{(m+1-i)} - p_{(1)}.
\end{aligned}$$

5. Man zeige, dass $H(p_1, \dots, p_m) = -\sum_{i=1}^{m} p_i\ \operatorname{ld} p_i$ die einzige Funktion ist, die folgende Bedingungen erfüllt:

   (a) $H(p_1, \dots, p_m) = H(p_{\pi_1}, \dots, p_{\pi_m})$ für jede Permutation $\pi_1, \dots, \pi_m$ von $1, \dots, m$, d.h. H ist symmetrisch in $p_1, \dots, p_m$,

   (b) $H(p_1, \dots, p_m) = H(p_1{+}p_2, p_3, \dots, p_m) + (p_1{+}p_2)\, H\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2}\right)$,

   (c) $H(p, 1-p)$ ist stetig in p,

   (d) $H\left(\frac{1}{2}, \frac{1}{2}\right) = 1$,

   (e) $h(m) := H\left(\frac{1}{m}, \dots, \frac{1}{m}\right)$ ist monoton steigend in m.

Beweis: Wir zeigen zuerst durch vollständige Induktion nach k, dass mit $P := (p_1, \dots, p_m)$ und $s_i := \sum_{j=1}^{i} p_j$ gilt

$$H(P) = H(s_k, p_{k+1}, \dots, p_m) + s_k\, H\left(\frac{p_1}{s_k}, \dots, \frac{p_k}{s_k}\right). \tag{5}$$

Für $k = 2$ ist das Bedingung b). Aus der Induktionsannahme für $k-1$ folgt

$$\begin{aligned}
H(P) &= H(s_{k-1}, p_k, \dots, p_m) + s_{k-1}\, H\left(\frac{p_1}{s_{k-1}}, \dots, \frac{p_{k-1}}{s_{k-1}}\right) \\
&= H(s_k, p_{k+1}, \dots, p_m) + s_k\, H\left(\frac{s_{k-1}}{s_k}, \frac{p_k}{s_k}\right) + s_{k-1}\, H\left(\frac{p_1}{s_{k-1}}, \dots, \frac{p_{k-1}}{s_{k-1}}\right) \\
&= H\left(s_k, \mathbf{p}_{k+1}^m\right) + s_k \left[H\left(\frac{s_{k-1}}{s_k}, \frac{p_k}{s_k}\right) + \frac{s_{k-1}}{s_k}\, H\left(\frac{\frac{p_1}{s_k}}{\frac{s_{k-1}}{s_k}}, \dots, \frac{\frac{p_{k-1}}{s_k}}{\frac{s_{k-1}}{s_k}}\right)\right] \\
&= H\left(s_k, \mathbf{p}_{k+1}^m\right) + s_k\, H\left(\frac{p_1}{s_k}, \dots, \frac{p_k}{s_k}\right).
\end{aligned} \tag{6}$$

Wiederholte Anwendung von (5) ergibt

$$\begin{aligned}
h(m\,n) &= H\left(\frac{1}{n}, \frac{1}{m\,n}, \dots, \frac{1}{m\,n}\right) + \frac{1}{n}\, H\left(\frac{\frac{1}{m\,n}}{\frac{1}{n}}, \dots, \frac{\frac{1}{m\,n}}{\frac{1}{n}}\right) \\
&= H\left(\frac{1}{n}, \frac{1}{mn}, \dots, \frac{1}{mn}\right) + \frac{1}{n} h(m) = H\left(\frac{1}{n}, \frac{1}{n}, \frac{1}{mn}, \dots, \frac{1}{mn}\right) + \frac{2}{n} h(m) \\
&= \dots = H\left(\frac{1}{n}, \dots, \frac{1}{n}\right) + \frac{n}{n} h(m) = h(n) + h(m).
\end{aligned} \tag{7}$$

Damit gilt $h(1) = h(1 \cdot 1) = h(1) + h(1) \;\Rightarrow\; h(1) = 0$. Aus (7) folgt auch $h(2^n) = n\, h(2) = n = \operatorname{ld}(2^n)$. Ist $m \in \mathbb{N}$ und $2^{n-1} < m^k \leq 2^n$, so gilt $n - 1 \leq h(m^k) = k\, h(m) \leq n \;\Rightarrow\; \frac{n-1}{k} \leq h(m) \leq \frac{n}{k}$. Andererseits gilt auch $\frac{n-1}{k} \leq \operatorname{ld} m \leq \frac{n}{k} \;\Rightarrow\; h(m) = \operatorname{ld} m$. Der Rest geht wie im Skriptum.