11. Übung aus Informations- und Codierungstheorie SS 2016

1. Aus 2 Kanälen $(X_i, Y_i, \pi_i(y_i|x_i)),\ i = 1, 2$ mit den Kapazitäten C_1, C_2 wird der Kanal $((X_1, X_2), (Y_1, Y_2), \pi((y_1, y_2)|(x_1, x_2)) := \pi_1(y_1|x_1)\, \pi_2(y_2|x_2))$ gebildet. Welche Kapazität hat der neue Kanal?
Lösung: Es gilt $H(Y_1, Y_2) \leq H(Y_1) + H(Y_2)$ und

$$\begin{aligned}
&H(Y_1, Y_2|X_1, X_2)\\
&\quad = -\sum_{x_1,x_2} p(x_1, x_2) \sum_{y_1,y_2} \pi_1(y_1|x_1)\, \pi_2(y_2|x_2)\ \log(\pi_1(y_1|x_1)\, \pi_2(y_2|x_2))\\
&\quad = -\sum_{x_1,x_2} p(x_1, x_2) \sum_{y_1} \pi_1(y_1|x_1)\ \log \pi_1(y_1|x_1) \sum_{y_2} \pi_2(y_2|x_2)\\
&\qquad -\sum_{x_1,x_2} p(x_1, x_2) \sum_{y_2} \pi_2(y_2|x_2)\ \log \pi_2(y_2|x_2) \sum_{y_1} \pi_1(y_1|x_1)\\
&\quad = -\sum_{x_1} p(x_1) \sum_{y_1} \pi_1(y_1|x_1)\ \log \pi_1(y_1|x_1) \sum_{x_2} p(x_2|x_1)\\
&\qquad -\sum_{x_2} p(x_2) \sum_{y_2} \pi_2(y_2|x_2)\ \log \pi_2(y_2|x_2) \sum_{x_1} p(x_1|x_2)\\
&\quad = H(Y_1|X_1) + H(Y_2|X_2).
\end{aligned}$$

Daher gilt

$$\begin{aligned}
I((X_1, X_2), (Y_1, Y_2)) &= H(Y_1, Y_2) - H(Y_1, Y_2|X_1, X_2)\\
&\leq H(Y_1) - H(Y_1|X_1) + H(Y_2) - H(Y_2|X_2)\\
&= I(X_1, Y_1) + I(X_2, Y_2) \leq C_1 + C_2. \qquad (0.1)
\end{aligned}$$

In der letzten Zeile von Ungleichung (0.1) gilt Gleichheit genau dann, wenn $X_1 \sim P_1^*$ mit $I(P_1^*, \Pi_1) = C_1$ und $X_2 \sim P_2^*$ mit $I(P_2^*, \Pi_2) = C_2$. Gilt nun $X_i \sim P_i^*\ i = 1, 2$ und sind X_1 und X_2 unabhängig, d.h. $p((x_1.x_2) = p_1^*(x_1)\, p_2^*(x_2)$, so gilt auch

$$\begin{aligned}
P(Y_1 = y_1, Y_2 = y_2) &= \sum_{x_1,x_2} p_1^*(x_1)\, p_2^*(x_2)\, \pi_1(y_1|x_1)\, \pi_2(y_2|x_2)\\
&= \sum_{x_1} p_1^*(x_1)\, \pi_1(y_1|x_1) \sum_{x_2} p_2^*(x_2)\, \pi_2(y_2|x_2) = P(Y_1 = y_1)\, P(Y_2 = y_2),
\end{aligned}$$

d.h. Y_1, Y_2 sind unabhängig. Somit gilt auch Gleichheit zwischen der 1-ten und der 2-ten Zeile von (0.1). Die neue Kanalkapazität ist also $C_1 + C_2$.
2. Ein Kanal heißt symmetrisch, wenn jede Zeile der Übergangsmatrix eine Permutation der ersten Zeile dieser Matrix ist und wenn auch die Spalten

Permutationen voneinander sind. Man berechne die Kanalkapazität eines symmetrischen Kanals.
Lösung: Ist π die 1-te Zeile der Übergangsmatrix $\Pi = (p_{x,y})_{x\in A, y\in B}$, d.h. $\pi(y) = p_{x_1,y} \quad \forall\, y \in B$ und q die 1-te Spalte, also $q(x) = p_{x,y_1} \quad \forall\, x \in B$, so gilt $H(Y|X = x) = -\sum\limits_y p_{x,y} \log p_{x,y} = -\sum\limits_y \pi(y) \log \pi(y) \quad \forall\, x \in A$. Daraus folgt $H(Y|X) = H(\pi)$. Klarerweise gilt stets $H(Y) \leq \log|B|$. Daraus folgt $C = \sup I(X,Y) = H(Y) - H(Y|X) \leq \log|B| - H(\pi)$. Aber, wenn X gleichverteilt ist, d.h. $P(X = x) = \frac{1}{|A|} \quad \forall\, x \in A$, so gilt für Y $P(Y = y) = \sum\limits_x \frac{1}{|A|}\, p_{x,y} = \frac{1}{|A|} \sum\limits_x q(x) \quad \forall\, y \in B$. Somit ist dann Y gleichverteilt auf B, also gilt $H(Y) = \log|B| \;\Rightarrow\; C = \log|B| - H(\pi)$.

3. Für einen Kanal mit dem Eingangsalphabet $\{0,1\}$, dem Ausgangsalphabet $\{0,e,1\}$ und den Übergangswahrscheinlichkeiten $p(0|0) = p(1|1) = 1-\varepsilon$, $p(e|0) = p(e|1) = \varepsilon, \quad 0 < \varepsilon < 1$ berechne man die Kanalkapazität.
   Lösung: Offensichtlich gilt $P(X = 0|Y = 0) = P(X = 1|Y = 1) = 1$. Daraus folgt $H(X|Y) = P(Y = e)\, H(X|Y = e)$. Für $X \sim B_p$ gilt nun $P(Y = e) = (1-p)\,\varepsilon + p\,\varepsilon = \varepsilon$, $P(X = 1|Y = e) = \frac{P(X=1,Y=e)}{P(Y=e)} = \frac{p\,\varepsilon}{\varepsilon} = p$, bzw. $P(X = 0|Y = e) = 1-p$. Somit gilt $H(X|Y) = \varepsilon\, H(p, 1-p)$ und $I(X,Y) = H(X) - H(X|Y) = (1-\varepsilon)\, H(p, 1-p) \leq (1-\varepsilon)\, H\left(\frac{1}{2}, \frac{1}{2}\right) = 1-\varepsilon$. Somit wird die Kapazität $C = \sup I(X,Y) = 1-\varepsilon$ bei $p = \frac{1}{2}$ erreicht.
4. Man zeige, dass n hintereinander geschaltete binäre, symmetrische Kanäle mit Fehlerwahrscheinlichkeit $0 < \varepsilon < 1$ einem binären, symmetrischen Kanal mit Fehlerwahrscheinlichkeit $\varepsilon_n := \frac{1}{2}\,[1 - (1-2\,\varepsilon)^n]$ entsprechen und, dass deshalb gilt $\lim\limits_n I(X_0, X_n) = 0$.
$$X_0 \to [\text{BSC}_1] \to X_1 \to \ldots \to X_{n-1} \to [\text{BSC}_n] \to X_n$$
   Lösung: $n = 1$: $\quad \varepsilon_1 = \frac{1}{2}\,[1 - (1-2\,\varepsilon)] = \varepsilon$
   $n-1 \to n$: Mit $\delta := \varepsilon + \varepsilon_{n-1} - 2\,\varepsilon\,\varepsilon_{n-1}$ gilt
$$\begin{pmatrix} 1-\varepsilon_{n-1}, & \varepsilon_{n-1} \\ \varepsilon_{n-1}, & 1-\varepsilon_{n-1} \end{pmatrix} \begin{pmatrix} 1-\varepsilon, & \varepsilon \\ \varepsilon, & 1-\varepsilon \end{pmatrix} = \begin{pmatrix} 1-\delta, & \delta \\ \delta, & 1-\delta \end{pmatrix}.$$
   und
$$\begin{aligned} \delta &= \varepsilon + \frac{1}{2}\,[1 - (1-2\,\varepsilon)^{n-1}] - \frac{2\,\varepsilon}{2}\,[1 - (1-2\,\varepsilon)^{n-1}] \\ &= \varepsilon + \frac{1-2\,\varepsilon}{2}\,\left[1 - (1-2\,\varepsilon)^{n-1}\right] = \frac{1}{2}\,[1 - (1-2\,\varepsilon)^n] = \varepsilon_n\,. \end{aligned}$$
5. Zu einer Quelle W, die gleichverteilt eine von m Nachrichten erzeugt, welche durch eine gegebene Codierung f in m verschiedene Codeworte $f(i) \in A^n$ abgebildet werden, suche man jene Decodierung φ, die die Fehlerwahrscheinlichkeit $P_e := P(W \neq \varphi(\mathbf{Y}_1^n))$ minimiert.
   Hinweis: Betrachten Sie die Decodierung $\hat{\varphi}$, für die gilt $\hat{\varphi}(\mathbf{y}_1^n) = i$, wenn $P(\mathbf{Y}_1^n = \mathbf{y}_1^n | \mathbf{X}_1^n = f(i)) = \max\limits_{1\leq j\leq m} P(\mathbf{Y}_1^n = \mathbf{y}_1^n | \mathbf{X}_1^n = f(j))$.

Lösung: Für die Wahrscheinlichkeit einer fehlerfreien Decodierung gilt

$$P(W = \varphi(\mathbf{Y}_1^n)) = \sum_{\mathbf{y}_1^n \in B^n} P(\mathbf{Y}_1^n = \mathbf{y}_1^n)\, P(W = \varphi(\mathbf{y}_1^n) | \mathbf{Y}_1^n = \mathbf{y}_1^n).$$

In dieser Gleichung hängen nur die bedingten Wahrscheinlichkeiten in der rechten Summe von der Decodierung φ ab. Daher wird $P(W = \varphi(\mathbf{Y}_1^n))$ maximiert, wenn man die einzelnen $P(W = \varphi(\mathbf{y}_1^n) | \mathbf{Y}_1^n = \mathbf{y}_1^n)$ maximiert. Nun gilt bekanntlich

$$\begin{aligned} P(W = j | \mathbf{Y}_1^n = \mathbf{y}_1^n) &= \frac{P(W = j)\, P(\mathbf{Y}_1^n = \mathbf{y}_1^n | \mathbf{X}_1^n = f(j))}{P(\mathbf{Y}_1^n = \mathbf{y}_1^n)} \\ &= \frac{\frac{1}{m}\, P(\mathbf{Y}_1^n = \mathbf{y}_1^n | \mathbf{X}_1^n = f(j))}{P(\mathbf{Y}_1^n = \mathbf{y}_1^n)}. \end{aligned}$$

Daher gilt $P(W = i | \mathbf{Y}_1^n = \mathbf{y}_1^n) \geq P(W = j | \mathbf{Y}_1^n = \mathbf{y}_1^n) \quad \forall\, j$ genau dann, wenn $P(\mathbf{Y}_1^n = \mathbf{y}_1^n | \mathbf{X}_1^n = f(i)) \geq P(\mathbf{Y}_1^n = \mathbf{y}_1^n | \mathbf{X}_1^n = f(j)) \quad \forall\, j$. Die beste Decodierung ist also eine Maximum-Likelihood-Decodierung.