Musterlösung zur Aufgabe A3.1

a) Setzt man faire Würfel voraus, so ergibt sich für die Wahrscheinlichkeit, dass

- mit dem roten Würfel eine „6" geworfen wird:

$$\underline{\mathrm{Pr}(R=6)=1/6}=0.1667\,,$$

- mit dem blauen Würfel eine „1" oder eine „2" geworfen wird:

$$\underline{\mathrm{Pr}(B\le 2)=1/3}=0.3333\,,$$

- beide Würfel die gleiche Augenzahl anzeigen:

$$\underline{\mathrm{Pr}(R=B)=6/36}=0.1667\,.$$

Letzteres basiert auf der 2D–Darstellung auf dem Augenblatt sowie auf der **Klassischen Definition** der Wahrscheinlichkeit entsprechend K/M, wobei K = 6 der insgesamt M = 36 gleichwahrscheinlichen Elementarereignisse „$R \cap B$" dem hieraus abgeleiteten Ereignis „$R = B$" zugeordnet werden können, die auf der Diagonalen liegen. Würfelspieler sprechen in diesem Fall von einem Pasch.

b) Die Lösung basiert wieder auf der Klassischen Definition der Wahrscheinlichkeit:

- In K = 2 der M = 36 Elementarfelder steht eine „3": $\underline{\mathrm{Pr}(S=3)=2/36}=0.0556$.
- In K = 6 der M = 36 Elementarfelder steht eine „7": $\underline{\mathrm{Pr}(S=7)=6/36}=0.1667$.
- In K = 18 der M = 36 Felder steht eine ungerade Zahl ⇒ $\underline{\mathrm{Pr}(S\text{ ist ungerade})=18/36}=0.5$.

Dieses letzte Ergebnis könnte man auch auf anderem Wege erhalten:

$$\mathrm{Pr}(S\text{ ist ungerade}) =$$
$$= \mathrm{Pr}\big[(R\text{ ist ungerade})\cap(B\text{ ist gerade})\big]+\mathrm{Pr}\big[(R\text{ ist gerade})\cap(B\text{ ist ungerade})\big]\,.$$

Mit Pr(R gerade) = Pr(R ungerade) = Pr(B gerade) = Pr(B ungerade) = 1/2 folgt daraus ebenfalls:

$$\mathrm{Pr}(S\text{ ist ungerade}) = 1/2\cdot 1/2+1/2\cdot 1/2=1/2\,.$$

c) Die Wahrscheinlichkeit für das Ereignis, dass mindestens einer der beiden Würfel eine „6" zeigt, ist:

$$\mathrm{Pr}\big[(R=6)\cup(B=6)\big] = K/M = 11/36\;\underline{=0.3056}\,.$$

Die zweite Wahrscheinlichkeit steht für den „Sechser–Pasch":

$$\mathrm{Pr}\big[(R=6)\cap(B=6)\big] = K/M = 1/36\;\underline{=0.0278}\,.$$

d) Das Ergebnis für L = 1 wurde bereits in der Teilaufgabe (c) ermittelt:

$$p_1=\mathrm{Pr}\big[(R=6)\cup(B=6)\big] = 11/36\;\underline{=0.3056}\,.$$

Die Wahrscheinlichkeit p_2 lässt sich mit p_1 wie folgt ausdrücken:

$$p_2=(1-p_1)\cdot p_1=\frac{25}{36}\cdot\frac{11}{36}\;\underline{=0.2122}\,.$$

In anderen Worten: Die Wahrscheinlichkeit, dass im zweiten Wurf erstmals eine „6" geworfen wird, ist gleich der Wahrscheinlichkeit, dass im ersten Wurf keine „6" geworfen wurde ⇒ Wahrscheinlichkeit

$(1-p_1)$, aber im zweiten Wurf mindestens eine „6" dabei ist $\Rightarrow$ Wahrscheinlichkeit p_1. Entsprechend gilt für die Wahrscheinlichkeit „erste 6 im dritten Wurf":

$$p_3 = (1-p_1)^2 \cdot p_1 = \frac{25}{36} \cdot \frac{25}{36} \cdot \frac{11}{36} \underline{= 0.1474}\,.$$

e) Durch Erweiterung der Musterlösung zur Teilaufgabe (d) erhält man:

$$\begin{aligned}
\Pr(L \text{ ist geradzahlig}) &= p_2 + p_4 + p_6 + \ldots = \\
&= (1-p_1) \cdot p_1 + (1-p_1)^3 \cdot p_1 + (1-p_1)^5 \cdot p_1 + \ldots = \\
&= (1-p_1) \cdot p_1 \cdot \left[1 + (1-p_1)^2 + (1-p_1)^4 + \ldots\right]\,.
\end{aligned}$$

Entsprechend erhält man für die Wahrscheinlichkeit des Komplementärereignisses:

$$\begin{aligned}
\Pr(L \text{ ist ungeradzahlig}) &= p_1 + p_3 + p_5 + \ldots = \\
&= p_1 \cdot \left[1 + (1-p_1)^2 + (1-p_1)^4 + \ldots\right]
\end{aligned}$$

$$\Rightarrow \quad \frac{\Pr(L \text{ ist ungeradzahlig})}{\Pr(L \text{ ist geradzahlig})} = \frac{1}{1-p_1}.$$

Weiter muss gelten:

$$\Pr(L \text{ ist geradzahlig}) + \Pr(L \text{ ist ungeradzahlig}) = 1$$

$$\Rightarrow \quad \Pr(L \text{ ist geradzahlig}) \cdot \left[1 + \frac{1}{1-p_1}\right] = 1$$

$$\Rightarrow \quad \Pr(L \text{ ist geradzahlig}) = \frac{1-p_1}{2-p_1} = \frac{25/36}{61/36} = \frac{25}{61} \underline{= 0.4098}\,.$$

Musterlösung zur Zusatzaufgabe Z3.1

a) Werden die Karten nach dem Ziehen zurückgelegt, so ist zu jedem Zeitpunkt die Wahrscheinlichkeit für ein Ass gleich groß (1/8):

$$p_{\rm a} = {\rm Pr}(3\ {\rm Asse}) = {\rm Pr}(A_1) \cdot {\rm Pr}(A_2) \cdot {\rm Pr}(A_3) = \left(1/8\right)^3 \ \underline{\approx 0.002}.$$

b) Nun erhält man mit dem allgemeinen Multiplikationstheorem:

$$p_{\rm b} = {\rm Pr}(A_1 \cap A_2 \cap A_3) = {\rm Pr}(A_1) \cdot {\rm Pr}(A_2|A_1) \cdot {\rm Pr}(A_3|(A_1 \cap A_2)).$$

Die bedingten Wahrscheinlichkeiten können nach der klassischen Definition berechnet werden. Man erhält somit das Ergebnis „k/m" (bei m Karten sind noch k Asse enthalten).

$$p_{\rm b} = \frac{4}{32} \cdot \frac{3}{31} \cdot \frac{2}{30} \ \underline{\approx 0.0008}.$$

$p_{\rm b}$ ist kleiner als $p_{\rm a}$, da nun das zweite und dritte Ass unwahrscheinlicher sind als zuvor.

c) Analog zu Punkt (b) erhält man hier:

$$p_{\rm c} = {\rm Pr}(\overline{A_1}) \cdot {\rm Pr}(\overline{A_2}\,|\,\overline{A_1}) \cdot {\rm Pr}(\overline{A_3}\,|\,(\overline{A_1} \cap \overline{A_2})) = \frac{28}{32} \cdot \frac{27}{31} \cdot \frac{26}{30} \ \underline{\approx 0.6605}.$$

d) Diese Wahrscheinlichkeit kann man als die Summe dreier Wahrscheinlichkeiten ausdrücken, da die zugehörigen Ereignisse disjunkt sind:

$$p_{\rm d} = {\rm Pr}(D_1 \cup D_2 \cup D_3)\ {\rm mit:}$$

$${\rm Pr}(D_1) = {\rm Pr}(A_1 \cap \overline{A_2} \cap \overline{A_3}) = \frac{4}{32} \cdot \frac{28}{31} \cdot \frac{27}{30} = 0.1016,$$

$${\rm Pr}(D_2) = {\rm Pr}(\overline{A_1} \cap A_2 \cap \overline{A_3}) = \frac{28}{32} \cdot \frac{4}{31} \cdot \frac{27}{30} = 0.1016,$$

$${\rm Pr}(D_3) = {\rm Pr}(\overline{A_1} \cap \overline{A_2} \cap A_3) = \frac{28}{32} \cdot \frac{27}{31} \cdot \frac{4}{30} = 0.1016.$$

Diese Wahrscheinlichkeiten sind alle gleich – warum sollte es auch anders sein? Wenn man bei 3 Karten genau ein Ass zieht, ist es genau so wahrscheinlich, ob man dieses als erste, als zweite oder als dritte Karte zieht. Damit erhält man für die Summe $p_{\rm d}\ \underline{= 0.3048}$.

e) Definiert man die Ereignisse E_i = „es werden genau i Asse gezogen" mit den Indizes i = 0, 1, 2, 3, so beschreiben E_0, E_1, E_2 und E_3 ein vollständiges System. Deshalb gilt:

$$p_{\rm e} = {\rm Pr}(E_2) = 1 - p_{\rm b} - p_{\rm c} - p_{\rm d} \ \underline{= 0.0339}.$$

Musterlösung zur Aufgabe A3.2

a) Allgemein gilt für den Erwartungswert der Funktion $F(X)$ hinsichtlich der Zufallsvariablen X:

$$\mathrm{E}\left[F(X)\right] = \sum_{x\,\in\,\mathrm{supp}(P_X)} P_X(x) \cdot F(x)\,.$$

Im vorliegenden Beispiel gilt $X = \{0,\ 1,\ 2,\ 3\}$ und $P_X(X) = [1/2,\ 1/8,\ 0,\ 3/8]$. Wegen $P_X(X = 2) = 0$ ergibt sich somit für die zu berücksichtigende Menge („Support") in obiger Summation:

$$\mathrm{supp}(P_X) = \{0\,,1\,,3\}\,.$$

Mit $F(X) = 1/P_X(X)$ erhält man weiter:

$$\mathrm{E}\left[1/P_X(X)\right] = \sum_{x\,\in\,\{0,1\,,3\}} P_X(x) \cdot 1/P_X(x) = \sum_{x\,\in\,\{0,1\,,3\}} 1 \underline{= 3}\,.$$

Der zweite Erwartungswert liefert mit $\mathrm{supp}(P_Y) = \{0,\ 1,\ 2\}$ das gleiche Ergebnis: $\mathrm{E}[1/P_Y(Y)] \underline{= 3}$.

b) In beiden Fällen ist der Index der Wahrscheinlichkeitsfunktion mit der Zufallsvariablen (X bzw. Y) identisch und man erhält

$$\mathrm{E}\left[P_X(X)\right] = \sum_{x\,\in\,\{0,1\,,3\}} P_X(x) \cdot P_X(x) = (1/2)^2 + (1/8)^2 + (3/8)^2 = 13/32 \underline{\approx 0.406}\,,$$

$$\mathrm{E}\left[P_Y(Y)\right] = \sum_{y\,\in\,\{0,1\,,2\}} P_Y(y) \cdot P_Y(y) = (1/2)^2 + (1/4)^2 + (1/4)^2 \underline{= 0.375}\,.$$

c) Hier gelten folgende Gleichungen:

$$\mathrm{E}\left[P_Y(X)\right] = \sum_{x\,\in\,\{0,1\,,3\}} P_X(x) \cdot P_Y(x) = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{8}\cdot\frac{1}{4} + \frac{3}{8}\cdot 0 = 9/32 \underline{\approx 0.281}\,,$$

Die Erwartungswertbildung bezieht sich hier auf $P_X(\cdot)$, also auf die Zufallsgröße X. $P_Y(\cdot)$ ist dabei die formale Funktion ohne (direkten) Bezug zur Zufallsgröße Y.

Für den zweiten Erwartungswert erhält man den gleichen Zahlenwert (das muss nicht so sein):

$$\mathrm{E}\left[P_X(Y)\right] = \sum_{y\,\in\,\{0,1\,,2\}} P_Y(y) \cdot P_X(y) = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{4}\cdot\frac{1}{8} + \frac{1}{4}\cdot 0 = 9/32 \underline{\approx 0.281}\,.$$

d) Wir berechnen zunächst die drei Erwartungswerte:

$$\mathrm{E}\left[-\log_2 P_U(U)\right] = \frac{1}{2}\cdot\log_2\frac{2}{1} + \frac{1}{2}\cdot\log_2\frac{2}{1} \underline{= 1\,\mathrm{bit}}\,,$$

$$\mathrm{E}\left[-\log_2 P_V(V)\right] = \frac{3}{4}\cdot\log_2\frac{4}{3} + \frac{1}{4}\cdot\log_2\frac{4}{1} \underline{= 0.811\,\mathrm{bit}}\,,$$

$$\mathrm{E}\left[-\log_2 P_V(U)\right] = \frac{1}{2}\cdot\log_2\frac{4}{3} + \frac{1}{2}\cdot\log_2\frac{4}{1} \underline{= 1.208\,\mathrm{bit}}\,.$$

Richtig sind demnach die <u>beiden ersten Aussagen</u>:

- Die Entropie $H(U) = 1$ bit kann entsprechend der ersten Gleichung berechnet werden. Sie gilt für die binäre Zufallsgröße U mit gleichen Wahrscheinlichkeiten.

- Die Entropie $H(V)$ = 0.811 bit berechnet sich entsprechend der zweiten Gleichung. Aufgrund der Wahrscheinlichkeiten 3/4 und 1/4 ist die Entropie (Unsicherheit) kleiner als für die Zufallsgröße U.
- Der dritte Erwartungswert kann schon allein vom Ergebnis her (1.208 bit) nicht die Entropie einer binären Zufallsgröße angeben, die stets auf 1 bit begrenzt ist.

Musterlösung zur Zusatzaufgabe Z3.2

a) Man kommt von $P_{XY}(X, Y)$ zur 1D–Wahrscheinlichkeitsfunktion $P_X(X)$, indem man alle Y–Wahrscheinlichkeiten aufsummiert:

$$P_X(X = x_\mu) = \sum_{y \in Y} P_{XY}(x_\mu, y)$$

$$\Rightarrow \quad \begin{aligned} P_X(X = 0) &= 1/4 + 1/8 + 1/8 = 1/2 \underline{= 0.500}\,, \\ P_X(X = 1) &= 0 + 0 + 1/8 = 1/8 \underline{= 0.125}\,, \\ P_X(X = 2) &= 0 + 0 + 0 \underline{= 0}\,, \\ P_X(X = 3) &= 1/4 + 1/8 + 0 = 3/8 \underline{= 0.375} \end{aligned}$$

$$\Rightarrow \quad P_X(X) = [1/2\,, 1/8\,, 0\,, 3/8\,]\,.$$

b) Analog zur Teilaufgabe (a) gilt nun:

$$P_Y(Y = y_\kappa) = \sum_{x \in X} P_{XY}(x, y_\kappa)$$

$$\Rightarrow \quad \begin{aligned} P_Y(Y = 0) &= 1/4 + 0 + 0 + 1/4 = 1/2 \underline{= 0.500}\,, \\ P_Y(Y = 1) &= 1/8 + 0 + 0 + 1/8 = 1/4 \underline{= 0.250}\,, \\ P_Y(Y = 2) &= 1/8 + 1/8 + 0 + 0 = 1/4 \underline{= 0.250} \end{aligned}$$

$$\Rightarrow \quad P_Y(Y) = [1/2\,, 1/4\,, 1/4\,]\,.$$

c) Bei Unabhängigkeit sollte $P_{XY}(X, Y) = P_X(X) \cdot P_Y(Y)$ sein. Dies trifft hier nicht zu ⇒ <u>Antwort Nein</u>.

d) Ausgehend von $P_{XY}(X, Y)$ ⇒ linke Tabelle kommt man zu $P_{UY}(U, Y)$ ⇒ mittlere Tabelle, indem man gewisse Wahrscheinlichkeiten entsprechend $U = X$ mod 2 zusammenfasst. Berücksichtigt man noch $V = Y$ mod 2, so erhält man die gesuchten Wahrscheinlichkeiten entsprechend der rechten Tabelle:

$$\begin{aligned} P_{UV}(U = 0, V = 0) &= P_{UV}(U = 0, V = 1) = 3/8 \underline{= 0.375}\,, \\ P_{UV}(U = 1, V = 0) &= P_{UV}(U = 1, V = 1) = 1/8 \underline{= 0.125}\,. \end{aligned}$$

$P_{XY}(X, Y)$

Y \ X	0	1	2	3
0	1/4	0	0	1/4
1	1/8	0	0	1/8
2	1/8	1/8	0	0

$P_{UY}(U, Y)$

Y \ U	0	1
0	1/4	1/4
1	1/8	1/8
2	1/8	1/8

$P_{UV}(U, V)$

V \ U	0	1
0	3/8	3/8
1	1/8	1/8

e) Die 1D–Wahrscheinlichkeitsfunktionen lauten nun:

$$P_U(U) = [1/2\,, 1/2\,]\,, \quad P_V(V) = [3/4\,, 1/4\,]\,.$$

Damit gilt $P_{UV}(U, V) = P_U(U) \cdot P_V(V)$ ⇒ U und V sind statistisch unabhängig ⇒ <u>Antwort Ja</u>.

Musterlösung zur Aufgabe A3.3

a) Beide Aussagen sind richtig. Setzt man $p_3 = 0$ und formal $p_1 = p \Rightarrow p_2 = 1 - p$, so ergibt sich die binäre Entropiefunktion

$$H_{\rm bin}(p) = p \cdot \log_2 \frac{1}{p} + (1-p) \cdot \log_2 \frac{1}{1-p}\,.$$

b) Man kann die binäre Entropiefunktion wegen $\log_2(x) = \ln(x)/\ln(2)$ auch in die folgende Form bringen:

$$H_{\rm bin}(p) = \frac{-1}{\ln(2)} \cdot \left[p \cdot \ln(p) + (1-p) \cdot \ln(1-p) \right]\,.$$

Die erste Ableitung führt zum Ergebnis

$$\begin{aligned}\frac{{\rm d}H_{\rm bin}(p)}{{\rm d}p} &= \frac{-1}{\ln(2)} \cdot \left[\ln(p) + p \cdot \frac{1}{p} - \ln(1-p) - (1-p) \cdot \frac{1}{1-p} \right] = \\ &= \frac{1}{\ln(2)} \cdot \left[\ln(1-p) - \ln(p) \right] = \log_2 \frac{1-p}{p}\,.\end{aligned}$$

Durch Nullsetzen dieser Ableitung erhält man den Abszissenwert $p = 0.5$, der zum Maximum der Entropiefunktion führt: $H_{\rm bin}(p = 0.5) = 1$ bit $\Rightarrow$ Lösungsvorschlag 2 ist falsch..

Durch nochmaliges Differenzieren erhält man für die zweite Ableitung:

$$\frac{{\rm d}^2 H_{\rm bin}(p)}{{\rm d}p^2} = \frac{1}{\ln(2)} \cdot \left[\frac{-1}{1-p} - \frac{1}{p} \right] = \frac{-1}{\ln(2) \cdot p \cdot (1-p)}\,.$$

Diese Funktion ist im gesamten Definitionsgebiet $0 \le p \le 1$ negativ $\Rightarrow$ $H_{\rm bin}(p)$ ist konkav $\Rightarrow$ Richtig ist dementsprechend (allein) der Lösungsvorschlag 1.

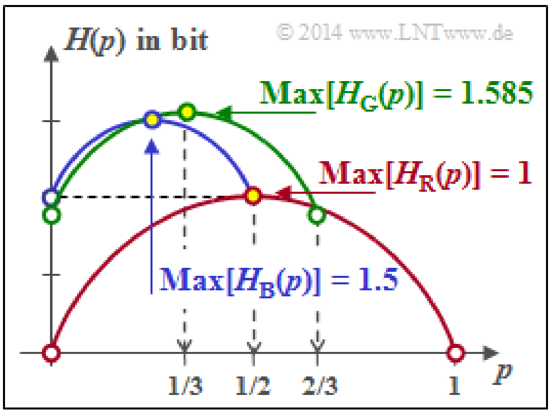
c) Richtig sind hier die Aussagen 1 und 2:

- Für $p = 0$ erhält man die Wahrscheinlichkeitsfunktion $P_X(X) = [0,\ 1/2,\ 1/2]$ $\Rightarrow$ $H(X) = 1$ bit.
- Das Maximum unter der Voraussetzung $p_3 = 1/2$ ergibt sich für $p_1 = p_2 = 1/4$:

$$P_X(X) = [\,1/4\,,\,1/4\,,\,1/2\,] \quad \Rightarrow \quad {\rm Max}[H_{\rm B}(p)] = 1.5\,{\rm bit}\,.$$

In kompakter Form lässt sich $H_{\rm B}(p)$ mit der Einschränkung $0 \le p \le 1/2$ wie folgt darstellen:

$$H_{\rm B}(p) = 1.0\,{\rm bit} + 1/2 \cdot H_{\rm bin}(2p)\,.$$



d) Richtig sind die erste und letzte Aussage. Der grüne Kurvenzug beinhaltet mit $p = 1/3$ auch die Gleichverteilung aller Wahrscheinlichkeiten $\Rightarrow$ ${\rm Max}[H_{\rm G}(p)] = \log_2 (3)$ bit. Allgemein lässt sich der gesamte Kurvenverlauf im Bereich $0 \le p \le 2/3$ wie folgt ausdrücken:

$$H_{\rm G}(p) = H_{\rm G}(p = 0) + 2/3 \cdot H_{\rm bin}(3p/2)\,.$$

Aus der Grafik auf der Angabenseite erkennt man auch, dass folgende Bedingung erfüllt sein muss:

$$H_\mathrm{G}(p=0) + 2/3 = \log_2(3) \quad \Rightarrow \quad H_\mathrm{G}(p=0) = 1.585 - 0.667 = 0.918\,\mathrm{bit}\,.$$

Der Lösungsvorschlag 2 ist hier somit falsch. Zum gleichen Ergebnis gelangt man über die Gleichung

$$H_\mathrm{G}(p=0) = 1/3 \cdot \log_2(3) + 2/3 \cdot \log_2(3/2) = \log_2(3) - 2/3 \cdot \log_2(2)\,.$$

Die Grafik zeigt nochmals die Ausgangsgrafik, aber nun mit Bemaßungen.

Musterlösung zur Zusatzaufgabe Z3.3

a) Mit $P_X(X)$ = [0.1, 0.2, 0.3, 0.4] erhält man für die Entropie:

$$H_\mathrm{a}(X) = 0.1 \cdot \log_2 \frac{1}{0.1} + 0.2 \cdot \log_2 \frac{1}{0.2} + 0.3 \cdot \log_2 \frac{1}{0.3} + 0.4 \cdot \log_2 \frac{1}{0.4} \underline{= 1.846}\,.$$

Hier (und bei den anderen Aufgaben) ist jeweils die Pseudo–Einheit „bit" anzufügen.

b) Die Entropie $H_\mathrm{b}(X)$ lässt sich als Summe zweier Anteile $H_\mathrm{b1}(X)$ und $H_\mathrm{b2}(X)$ darstellen, mit

$$H_\mathrm{b1}(X) = 0.1 \cdot \log_2 \frac{1}{0.1} + 0.2 \cdot \log_2 \frac{1}{0.2} = 0.797\,,$$
$$H_\mathrm{b2}(X) = p_3 \cdot \log_2 \frac{1}{p_3} + (0.7 - p_3) \cdot \log_2 \frac{1}{0.7 - p_3}\,.$$

Die zweite Funktion ist für $p_3 = p_4 = 0.35$ maximal. Ein ähnlicher Zusammenhang hat sich bei der **binären Entropiefunktion** ergeben. Damit erhält man:

$$H_\mathrm{b2}(X) = 2 \cdot p_3 \cdot \log_2 \frac{1}{p_3} = 0.7 \cdot \log_2 \frac{1}{0.35} = 1.060$$
$$\Rightarrow \quad H_\mathrm{b}(X) = H_\mathrm{b1}(X) + H_\mathrm{b2}(X) = 0.797 + 1.060 \underline{= 1.857}\,.$$

c) Analog zur Aufgabe (b) ergibt sich mit $p_1 = 0.1$, $p_4 = 0.4$ das Maximum für $p_2 = p_3 = 0.25$:

$$H_\mathrm{c}(X) = 0.1 \cdot \log_2 \frac{1}{0.1} + 2 \cdot 0.25 \cdot \log_2 \frac{1}{0.25} + 0.4 \cdot \log_2 \frac{1}{0.4} \underline{= 1.861}\,.$$

d) Die maximale Entropie für den Symbolumfang $M = 4$ ergibt sich bei gleichen Wahrscheinlichkeiten ($p_1 = p_2 = p_3 = p_4 = 0.25$):

$$H_\mathrm{max}(X) = \log_2 M \underline{= 2}\,.$$

Die Differenz der Entropien entsprechend (d) und (c) ergibt $\Delta H(X) = 0.139$ bit. Hierbei gilt:

$$\Delta H(X) = 1 - 0.1 \cdot \log_2 \frac{1}{0.1} - 0.4 \cdot \log_2 \frac{1}{0.4}\,.$$

Mit der binären Entropiefunktion

$$H_\mathrm{bin}(p) = p \cdot \log_2 \frac{1}{p} + (1-p) \cdot \log_2 \frac{1}{1-p}$$

lässt sich hierfür auch schreiben:

$$\Delta H(X) = 0.5 \cdot [1 - H_\mathrm{bin}(0.2)] = 0.5 \cdot [1 - 0.722] = 0.139\,.$$

Musterlösung zur Aufgabe A3.4

a) Bei der Binomialverteilung sind alle Wahrscheinlichkeiten $\Pr(X > I) = 0$ $\Rightarrow$ $\underline{I = 5}$. Damit ergibt sich für die Wahrscheinlichkeit, dass X gleich $I = 5$ ist:

$$\Pr(X = 5) = \binom{5}{5} \cdot p^5 = p^5 \approx 0.0003\,.$$

Damit erhält man für

- die charakteristische Wahrscheinlichkeit:

  $$p = (0.0003)^{1/5} = 0.1974 \underline{\approx 0.2}\,,$$

- den linearen Mittelwert (Erwartungswert):

  $$m_X = I \cdot p \underline{= 1}\,,$$

- die Varianz:

  $$\sigma_X^2 = I \cdot p \cdot (1-p) \underline{= 0.8}\,.$$

b) Richtig ist der Lösungsvorschlag 2. Bei Verwendung der Kullback–Leibler–Distanz $D(P_Y || P_X)$ würde sich unabhängig von λ stets ein unendlicher Wert ergeben, da für $\mu \ge 6$ gilt:

$$P_X(X = \mu) = 0\,, \quad P_Y(Y = \mu) \ne 0\,.$$

Auch wenn die Wahrscheinlichkeiten $P_Y(Y = \mu)$ für große μ sehr klein werden, sind sie doch „unendlich viel größer" als $P_X(X = \mu)$.

c) Wir verwenden die erste Kullback–Leibler–Distanz:

$$D = D(P_X \,||\, P_Y) = \sum_{\mu=0}^{5} P_X(\mu) \cdot \log_2 \frac{P_X(\mu)}{P_Y(\mu)}\,.$$

Bei Verwendung des Zehnerlogarithmus („lg") erhalten wir für die Poisson–Näherung mit $\lambda = 1$:

$$D' = 0.3277 \cdot \lg \frac{0.3277}{0.3679} + A' = -0.016468 + 0.021944 = 0.005476\,.$$

Nach Umrechnung auf den Zweierlogarithmus („$\log_2$") erhält man:

$$D = D(P_X \,||\, P_Y) = \frac{0.005476}{\lg(2)} \underline{\approx 0.0182\,(\text{bit})}\,.$$

d) Unter Verwendung des Zehnerlogarithmus lautet die Entropie der Poisson–Näherung ($\lambda = 1$):

$$\begin{aligned} H'(Y) &= -\text{E}\left[\lg P_Y(Y)\right] = -2 \cdot 0.3679 \cdot \lg(0.3679) - B' = \\ &= 0.31954 + 0.24717 = 0.56126 \end{aligned}$$

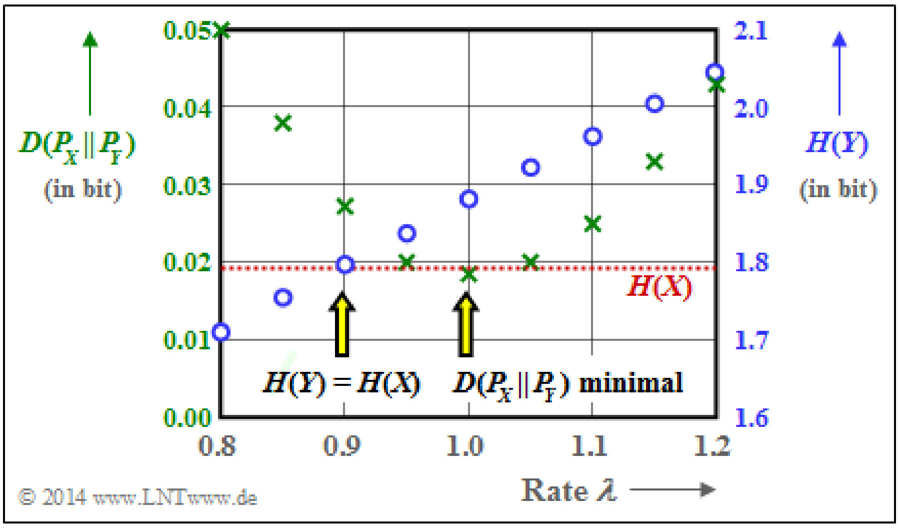
Die Umrechnung in „bit" liefert das gesuchte Ergebnis:

$$H(Y) = \frac{0.56126}{\lg(2)} \underline{= 1.864\,(\text{bit})}\,.$$

e) Richtig ist die Aussage 1. Bei der numerischen Berechnung der Kullback–Leibler–Distanz ist

- der Beitrag des μ–ten Terms positiv, falls $P_Y(\mu) > P_X(\mu)$,
- der Beitrag des μ–ten Terms negativ, falls $P_Y(\mu) < P_X(\mu)$.

f) Zutreffend ist der Lösungsvorschlag 1. Auch aus der nachfolgenden Grafik ist ersichtlich, dass $D(P_X||P_Y) = 0.0182$ bit von keinem anderen λ–Wert als $\lambda = 1$ unterschritten wird (grüne Kreuze).



Weiter erkennt man aus dieser Darstellung, dass man mit $\lambda = 0.9$ eine bessere Entropie–Approximation als mit $\lambda = 1$ erreicht (blaue Kreise):

$$H(Y) = 1.795\,\text{bit} \;\approx\; H(X) = 1.793\,\text{bit}\,.$$

Der zweite Lösungsvorschlag ist also falsch. Anzumerken ist:

- Mit $\lambda = 1$ stimmen die linearen Mittelwerte der beiden Zufallsgrößen überein: $m_X = m_Y = 1$.
- Mit $\lambda = 0.9$ stimmen die quadratischen Mittelwerte überein: $m_X + \sigma_X^2 = m_Y + \sigma_Y^2 = 1.8$.

Ob diese Aussage relevant ist, lasse ich dahingestellt. Denn: Aufgrund der stetigen Zunahme von $H(Y)$ mit zunehmendem λ ist klar, dass für irgendeinen λ–Wert tatsächlich $H(Y) = H(X)$ gelten muss.

Musterlösung zur Zusatzaufgabe Z3.4

a) Bei gleichen Wahrscheinlichkeiten gilt mit $M = 4$:

$$H(X) = \log_2 M \underline{= 2\,(\text{bit})}\,.$$

b) Die Wahrscheinlichkeiten für die empirisch ermittelten Zufallsgrößen Y weichen im Allgemeinen (nicht immer!) von der Gleichverteilung um so mehr ab, je kleiner der Parameter N ist. Man erhält

- für $N = 1000$ $\Rightarrow$ $P_Y(Y) = [0.225, 0.253, 0.250, 0.272]$:

$$\begin{aligned} H(Y) &= 0.225 \cdot \log_2 \frac{1}{0.225} + 0.253 \cdot \log_2 \frac{1}{0.253} + \\ &+ 0.250 \cdot \log_2 \frac{1}{0.250} + 0.272 \cdot \log_2 \frac{1}{0.272} \underline{= 1.9968\,(\text{bit})}\,, \end{aligned}$$

- für $N = 100$ $\Rightarrow$ $P_Y(Y) = [0.24, 0.16, 0.30, 0.30]$:

$$H(Y) = \ldots \underline{= 1.9410\,(\text{bit})}\,,$$

- für $N = 10$ $\Rightarrow$ $P_Y(Y) = [0.5, 0.1, 0.3, 0.1]$:

$$H(Y) = \ldots \underline{= 1.6855\,(\text{bit})}\,.$$

c) Die Gleichung für die gesuchte Kullback–Leibler–Distanz lautet:

$$\begin{aligned} D(P_X \,||\, P_Y) &= \sum_{\mu=1}^{4} P_X(\mu) \cdot \log_2 \frac{P_X(\mu)}{P_Y(\mu)} = \\ &= \frac{1/4}{\lg\,(2)} \cdot \left[\lg \frac{0.25}{P_Y(1)} + \frac{0.25}{P_Y(2)} + \frac{0.25}{P_Y(3)} + \frac{0.25}{P_Y(4)} \right] = \\ &= \frac{1}{4 \cdot \lg\,(2)} \cdot \left[\lg \frac{0.25^4}{P_Y(1) \cdot P_Y(2) \cdot P_Y(3) \cdot P_Y(4)} \right]. \end{aligned}$$

Der Logarithmus zur Basis 2 $\Rightarrow$ $\log_2(\cdot)$ wurde zur einfachen Nutzung des Taschenrechners durch den Zehnerlogarithmus $\Rightarrow$ $\lg(\cdot)$ ersetzt. Man erhält die folgenden numerischen Ergebnisse:

- für $N = 1000$:

$$D(P_X \,||\, P_Y) = \frac{1}{4 \cdot \lg\,(2)} \cdot \left[\lg \frac{0.25^4}{0.225 \cdot 0.253 \cdot 0.250 \cdot 0.272} \right] \underline{= 3.28 \cdot 10^{-3}\,(\text{bit})}\,,$$

- für $N = 100$:

$$D(P_X \,||\, P_Y) = \frac{1}{4 \cdot \lg\,(2)} \cdot \left[\lg \frac{0.25^4}{0.24 \cdot 0.16 \cdot 0.30 \cdot 0.30} \right] \underline{= 4.42 \cdot 10^{-2}\,(\text{bit})}\,,$$

- für $N = 10$:

$$D(P_X \,||\, P_Y) = \frac{1}{4 \cdot \lg\,(2)} \cdot \left[\lg \frac{0.25^4}{0.5 \cdot 0.1 \cdot 0.3 \cdot 0.1} \right] \underline{= 3.45 \cdot 10^{-1}\,(\text{bit})}\,.$$

d) Richtig ist <u>Nein</u>, wie am Beispiel $N = 100$ gezeigt werden soll:

$$D(P_Y\,||\,P_X) \;=\; \sum_{\mu=1}^{M} P_Y(\mu)\cdot\log_2\frac{P_Y(\mu)}{P_X(\mu)} = 0.24\cdot\log_2\frac{0.24}{0.25} + $$

$$+\; 0.16\cdot\log_2\frac{0.16}{0.25} + 2\cdot 0.30\cdot\log_2\frac{0.30}{0.25} = 0.0407\;(\text{bit})\,.$$

In der Teilaufgabe (c) haben wir stattdessen $D(P_X||P_Y)$ = 0.0442 erhalten. Das bedeutet auch: Der Name „Distanz” ist etwas irreführend. Danach würde man eigentlich $D(P_Y||P_X) = D(P_X||P_Y)$ erwarten.

e) Mit $P_Y(X)$ = [0, 0.25, 0.5, 0.25] erhält man:

$$D(P_X\,||\,P_Y) = 0.25\cdot\log_2\frac{0.25}{0} + 2\cdot 0.25\cdot\log_2\frac{0.25}{0.25} + 0.25\cdot\log_2\frac{0.25}{0.50}\,.$$

Aufgrund des ersten Terms ergibt sich für $D(P_X||P_Y)$ ein unendlich großer Wert. Für die zweite Kullback–Leibler–Distanz gilt:

$$D(P_Y\,||\,P_X) = 0\cdot\log_2\frac{0}{0.25} + 2\cdot 0.25\cdot\log_2\frac{0.25}{0.25} + 0.50\cdot\log_2\frac{0.5}{0.25}\,.$$

Nach einer Grenzwertbetrachtung erkennt man, dass der erste Term das Ergebnis 0 liefert. Auch der zweite Term ergibt sich zu 0, und man erhält als Endergebnis:

$$D(P_Y\,||\,P_X) = 0.50\cdot\log_2\,(2)\;\underline{= 0.5\,(\text{bit})}\,.$$

Richtig sind somit die Aussagen 3 und 5. Auch aus diesem Extrembeispiel wird deutlich, dass sich $D(P_Y||P_X)$ stets von $D(P_X||P_Y)$ unterscheidet. Nur für den Sonderfall $P_Y = P_X$ sind beide Kullback–Leibler–Distanzen gleich, nämlich 0. Die folgende Tabelle zeigt das vollständige Ergebnis dieser Aufgabe.

f) Die richtige Antwort ist Nein. Die Tendenz ist zwar eindeutig: Je größer N ist,

- desto mehr nähert sich $H(Y)$ im Prinzip dem Endwert $H(X)$ = 2 bit an.
- um so kleiner werden die Distanzen $D(P_X||P_Y)$ und $D(P_Y||P_X)$.

Man erkennt aus obiger Tabelle aber auch, dass es Ausnahmen gibt:

- Die Entropie $H(Y)$ ist für N = 1000 kleiner als für N = 400,
- Die Distanz $D(P_X||P_Y)$ ist für N = 1000 größer als für N = 400.

N	$P_Y(1)$	$P_Y(2)$	$P_Y(3)$	$P_Y(4)$	$H(Y)$ (in bit)	$D(P_X\|\|P_Y)$ (in bit)
4	0.0000	0.2500	0.5000	0.2500	1.5000	∞
10	0.5000	0.1000	0.3000	0.1000	1.6855	$3.45\cdot 10^{-1}$
40	0.2750	0.1750	0.2250	0.3250	1.9634	$3.76\cdot 10^{-2}$
100	0.2400	0.1600	0.3000	0.3000	1.9410	$4.42\cdot 10^{-2}$
400	0.2550	0.2450	0.2675	0.2325	1.9981	$1.92\cdot 10^{-3}$
1 000	0.2250	0.2530	0.2500	0.2720	1.9968	$3.28\cdot 10^{-3}$
4 000	0.2560	0.2463	0.2398	0.2580	1.9994	$4.87\cdot 10^{-4}$
10 000	0.2471	0.2570	0.2548	0.2411	1.9995	$4.63\cdot 10^{-4}$
40 000	0.2448	0.2529	0.2489	0.2534	1.9999	$1.40\cdot 10^{-4}$
100 000	0.2518	0.2511	0.2492	0.2479	2.0000	$2.74\cdot 10^{-6}$
400 000	0.2488	0.2504	0.2492	0.2516	2.0000	$1.38\cdot 10^{-6}$

Der Grund hierfür ist, dass das hier dokumentierte empirische Experiment mit N = 400 eher zu einer Gleichverteilung geführt hat als das Experiment mit N = 1000.

Würde man dagegen sehr (unendlich) viele Versuche mit N = 400 und N = 1000 starten und über diese mitteln, ergäbe sich tatsächlich der eigentlich erwartete monotone Verlauf.

Musterlösung zur Aufgabe A3.5

a) Für die Kullback–Leibler–Distanz (KLD) gilt:

$$D(P_X \,||\, P_Y) = \mathrm{E}\left[\log_2 \frac{P_X(X)}{P_Y(X)}\right] = \sum_{x \in X} P_X(x) \cdot \log_2 \frac{P_X(x)}{P_Y(x)} =$$
$$= \frac{1}{2} \cdot \log_2 \frac{1/2}{3/4} + 2 \cdot \frac{1}{4} \cdot \log_2 \frac{1/4}{1/8} = \frac{1}{2} \cdot \log_2 \frac{2}{3} + \frac{1}{2} \cdot \log_2 (2) =$$
$$= 1 - \frac{1}{2} \cdot \log_2 (3) \;\underline{= 0.2075\,(\mathrm{bit})}\,.$$

b) **Partitionierung A ⇒ $A_1 = \{0\}, A_2 = \{1, 2\}$**: Man erhält die Wahrscheinlichkeitsfunktionen $P_X^{(A)}(X) = [1/4, 3/4]$ und $Q_X^{(A)}(X) = [1/8, 7/8]$. Daraus folgt:

$$D(P_X^{(A)} \,||\, Q_X^{(A)}) = \frac{1}{4} \cdot \log_2 \frac{1/4}{1/8} + \frac{3}{4} \cdot \log_2 \frac{3/4}{7/8} =$$
$$= \frac{1}{4} \cdot \log_2 (2) + \frac{3}{4} \cdot \log_2 \frac{6}{7} \;\underline{= 0.0832\,(\mathrm{bit})}\,.$$

Es ergibt sich also eine kleinere KLD als in Teilaufgabe (a).

c) **Partitionierung B ⇒ $B_1 = \{1\}, B_2 = \{0, 2\}$**: Man erhält die Wahrscheinlichkeitsfunktionen $P_X^{(B)}(X) = [1/2, 1/2]$ und $Q_X^{(B)}(X) = [3/4, 1/4]$. Analog zur Teilaufgabe (b) erhält man nun:

$$D(P_X^{(B)} \,||\, Q_X^{(B)}) = \frac{1}{2} \cdot \log_2 \frac{1/2}{3/4} + \frac{1}{2} \cdot \log_2 \frac{1/2}{1/4} \;\underline{= 0.2075\,(\mathrm{bit})}\,.$$

⇒ gleiches Ergebnis wie in Aufgabe (a) ⇒ Bei Partitionierung (B) gilt das Gleichheitszeichen.

d) **Partitionierung C ⇒ $C_1 = \{2\}, C_2 = \{0, 1\}$**: Man erhält $P_x^{(C)} = [1/4, 3/4]$, $Q_x^{(C)} = [1/8, 7/8]$, also die gleichen Funktionen wie bei der Partitionierung A ⇒ <u>Lösungsvorschlag 1</u>.

e) Partitionierung B hat zum Ergebnis $D(P_x^{(B)} \,||\, Q_x^{(B)}) = D(P_x \,||\, Q_X)$ geführt. Für diesen Fall ist

$$\frac{P_X(1)}{Q_X(1)} = \frac{1/2}{3/4} = \frac{2}{3}, \quad \frac{P_X(B_1)}{Q_X(B_1)} = \frac{1/2}{3/4} = \frac{2}{3},$$
$$\frac{P_X(0)}{Q_X(0)} = \frac{1/4}{1/8} = 2, \quad \frac{P_X(B_2)}{Q_X(B_2)} = \frac{1/2}{1/4} = 2,$$
$$\frac{P_X(2)}{Q_X(2)} = \frac{1/4}{1/8} = 2, \quad \frac{P_X(B_2)}{Q_X(B_2)} = \frac{1/2}{1/4} = 2.$$

Es muss also für alle $x \in X$ gelten:

$$\frac{P_X(x)}{Q_X(x)} = \frac{P_X(B_1)}{Q_X(B_1)}, \quad \text{falls } x \in B_1,$$
$$\frac{P_X(x)}{Q_X(x)} = \frac{P_X(B_2)}{Q_X(B_2)}, \quad \text{falls } x \in B_2.$$

Durch Verallgemeinerung erhält man, dass <u>beide Lösungsvorschläge</u> richtig sind.

Musterlösung zur Aufgabe A3.6

a) Aus der gegebenen Verbundwahrscheinlichkeit erhält man

$$H(XY) = 0.18 \cdot \log_2 \frac{1}{0.18} + 0.16 \cdot \log_2 \frac{1}{0.16} + \\ + 0.02 \cdot \log_2 \frac{1}{0.02} + 0.64 \cdot \log_2 \frac{1}{0.64} \underline{= 1.393\,(\text{bit})}\,.$$

b) Die 1D–Wahrscheinlichkeitsfunktionen lauten $P_X(X)$ = [0.2, 0.8] und $P_Y(Y)$ = [0.34, 0.66]. Daraus folgt:

$$H(X) = 0.2 \cdot \log_2 \frac{1}{0.2} + 0.8 \cdot \log_2 \frac{1}{0.8} \underline{= 0.722\,(\text{bit})}\,,$$
$$H(Y) = 0.34 \cdot \log_2 \frac{1}{0.34} + 0.66 \cdot \log_2 \frac{1}{0.66} \underline{= 0.925\,(\text{bit})}\,.$$

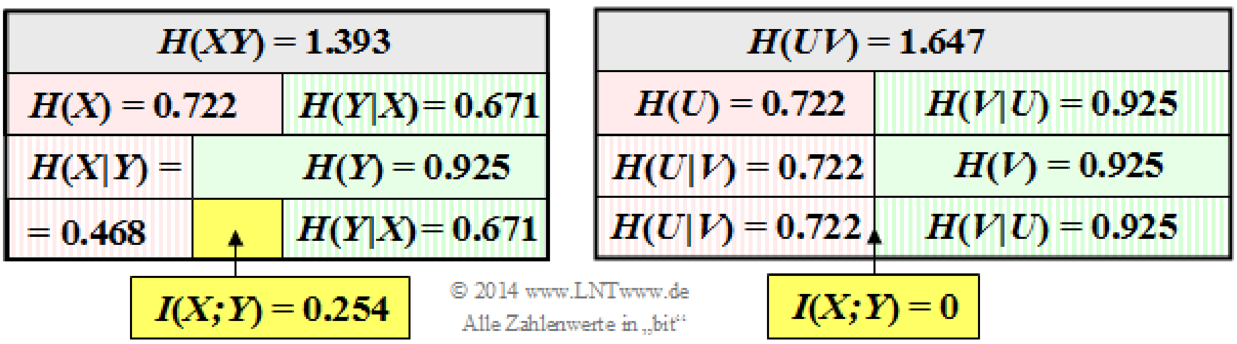
c) Aus der **Grafik** auf der Angabenseite erkennt man den Zusammenhang:

$$I(X;Y) = H(X) + H(Y) - H(XY) = \\ = 0.722\,(\text{bit}) + 0.925\,(\text{bit}) - 1.393\,(\text{bit}) \underline{= 0.254\,(\text{bit})}\,.$$

d) Ebenso gilt entsprechend der **Grafik** auf der Angabenseite:

$$H(X|Y) = H(XY) - H(Y) = 1.393 - 0.925 \underline{= 0.468\,(\text{bit})}\,,$$
$$H(Y|X) = H(XY) - H(X) = 1.393 - 0.722 \underline{= 0.671\,(\text{bit})}\,.$$

Die linke Grafik fasst die Ergebnisse der Teilaufgaben (a), ... , (d) maßstabsgetreu zusammen. Grau hinterlegt ist die Verbundentropie und gelb die Transinformation. Eine rote Hinterlegung bezieht sich auf die Zufallsgröße X, eine grüne auf Y. Schraffierte Felder deuten auf eine bedingte Entropie hin.



Die rechte Grafik beschreibt den gleichen Sachverhalt für die Zufallsgröße UV ⇒ Teilaufgabe (e).

e) Man erkennt die Gültigkeit von $P_{UV}(\cdot) = P_U(\cdot) \cdot P_V(\cdot)$ ⇒ Transinformation $I(U; V) = 0$ daran, dass die zweite Zeile der P_{UV}–Matrix sich von der ersten Zeile nur durch einen konstanten Faktor (4) unterscheidet. Richtig sind demzufolge die Aussagen <u>1, 2 und 4</u>. Weiter ist zu erwähnen:

- Es ergeben sich die gleichen 1D–Wahrscheinlichkeitsfunktiionen wie für die Zufallsgröße XY ⇒ $P_U(U)$ = [0.2, 0.8] und $P_V(V)$ = [0.34, 0.66].
- Deshalb ist auch $H(U) = H(X)$ = 0.722 bit und $H(V) = H(Y)$ = 0.925 bit.
- Hier gilt aber nun für die Verbundentropie: $H(UV) = H(U) + H(V) \neq H(XY)$.

Musterlösung zur Aufgabe A3.7

a) Mit $X = \{0,\ 1,\ 2\}$, $Y = \{0,\ 1,\ 2\}$ gilt $X + Y = \{0,\ 1,\ 2,\ 3,\ 4\}$ und auch die Wahrscheinlichkeiten stimmen mit der vorgegebenen Wahrscheinlichkeitsfunktion überein. Die Überprüfung der beiden anderen Vorgaben zeigt, dass auch $W = X - Y + 2$ möglich ist $\Rightarrow$ Lösungsvorschläge 1 und 2.

b) Aus der 2D–Wahrscheinlichkeitsfunktion $P_{XW}(X, W)$ auf der Angabenseite erhält man für

- die Verbundentropie:

$$H(XW) = \log_2(9) = 3.170\,\text{(bit)}\,,$$

- die Wahrsacheinlichkeitsfunktion der Zufallsgröße W:

$$P_W(W) = \left[\,1/9\,,\,2/9\,,\,3/9\,,\,2/9\,,\,1/9\,\right],$$

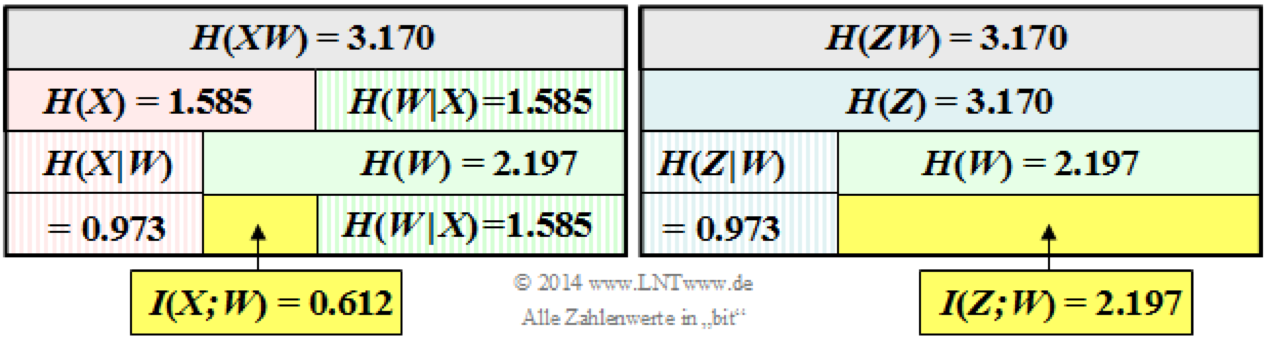
- die Entropie der Zufallsgröße W:

$$H(W) \;=\; 2 \cdot \frac{1}{9} \cdot \log_2 \frac{9}{1} + 2 \cdot \frac{2}{9} \cdot \log_2 \frac{9}{2} + \frac{3}{9} \cdot \log_2 \frac{9}{3} = 2.197\,\text{(bit)}\,.$$

Mit $H(X) = 1.585$ bit (wurde angegeben) ergibt sich somit für die *Mutual Information*:

$$\begin{aligned} I(X;W) \;&=\; H(X) + H(W) - H(XW) = \\ &=\; 1.585 + 2.197 - 3.170 \;\underline{= 0.612\,\text{(bit)}}\,. \end{aligned}$$

Das linke Schaubild verdeutlicht die Berechnung der Transinformation $I(X;\ W)$ zwischen der ersten Komponente X und der Summe W.



c) Die Grafik zeigt die Verbundwahrscheinlichkeit $P_{ZW}(\cdot)$. Das Schema besteht aus $5 \cdot 9 = 45$ Feldern im Gegensatz zur Darstellung von $P_{XW}(\cdot)$ auf der Angabenseite mit $3 \cdot 9 = 27$ Feldern.

Von den 45 Feldern sind aber auch nur neun mit Wahrscheinlichkeiten $\neq 0$ belegt. Für die Verbundentropie gilt:

$$H(ZW) = 3.170\,\text{(bit)}\,.$$

Mit den weiteren Entropien

$$H(Z) = 3.170\,\text{(bit)}\,,$$
$$H(W) = 2.197\,\text{(bit)}$$

$P_{ZW}(Z, W)$

W \ Z	00	01	02	10	11	12	20	21	22	$P_W(W)$
0	1/9									1/9
1		1/9		1/9						2/9
2			1/9		1/9		1/9			3/9
3						1/9		1/9		2/9
4									1/9	1/9

© 2014 www.LNTwww.de

entsprechend der Aufgabe Z3.7 bzw. der Teilaufgabe (b) erhält man für die Transinformation:

$$I(Z;W) = H(Z) + H(W) - H(ZW) \underline{= 2.197\,(\text{bit})}\,,$$

wie auch aus dem rechten oberen Schaubild hervorgeht.

d) <u>Alle drei Aussagen</u> treffen zu, wie auch aus dem oberen Schaubild ersichtlich ist. Wir versuchen eine Interpretation dieser numerischen Ergebnisse:

- Die Verbundwahrscheinlichkeit P_{ZW} setzt sich ebenso wie P_{XW} aus neun gleichwahrscheinlichen Elementen $\neq 0$ zusammen. Damit ist offensichtlich, dass auch die Verbundentropien gleich sind:

  $$H(ZW) = H(XW) = 3.170\,(\text{bit})\,.$$

- Wenn ich das Tupel $Z = (X, Y)$ kenne, kenne ich natürlich auch die Summe $W = X + Y$. Damit ist $H(W|Z) = 0$. Dagegen ist $H(Z|W)$ ungleich 0. Vielmehr gilt $H(Z|W) = H(X|W) = 0.973$ bit.
- Die Zufallsgröße W liefert also die genau gleiche Information hinsichtlich des Tupels Z wie für die Einzelkomponente X. Dies ist die verbale Interpretation für die Aussage $H(Z|W) = H(X|W)$
- Die gemeinsame Information von Z und $W \Rightarrow I(Z;\, W)$ ist größer als die von X und $W \Rightarrow I(X;\, W)$, weil $H(W|Z)$ gleich 0 ist, während $H(W|X)$ ungleich 0 ist, nämlich genau so groß ist wie $H(X)$:

  $$I(Z;W) = H(W) - H(W|Z) = 2.197 - 0 = 2.197\,(\text{bit})\,,$$
  $$I(X;W) = H(W) - H(W|X) = 2.197 - 1.585 = 0.612\,(\text{bit})\,.$$

Musterlösung zur Zusatzaufgabe Z3.7

a) Bei den beiden Zufallsgrößen $X = \{0, 1, 2\}$ $\Rightarrow$ $|X| = 3$ und $Y = \{0, 1, 2\}$ $\Rightarrow$ $|Y| = 3$ liegt jeweils eine Gleichverteilung vor. Damit erhält man für die Entropien:

$$H(X) = \log_2(3) \underline{= 1.585\,(\text{bit})}\,,$$
$$H(Y) = \log_2(3) \underline{= 1.585\,(\text{bit})}\,.$$

Die 2D–Zufallsgröße $XY = \{00, 01, 02, 10, 11, 12, 20, 21, 22\}$ $\Rightarrow$ $|XY| = |Z| = 9$ weist ebenfalls gleiche Wahrscheinlichkeiten auf: $p_{00} = p_{01} = \ldots = p_{22} = 1/9$. Daraus folgt:

$$H(XY) = \log_2(9) \underline{= 3.170\,(\text{bit})}\,.$$

b) Die Zufallsgrößen X und Y sind wegen $P_{XY}(\cdot) = P_X(\cdot) \cdot P_Y(\cdot)$ statistisch unabhängig $\Rightarrow$ $\underline{I(X, Y) = 0}$. Zum gleichen Ergebnis kommt man über die Gleichung $I(X; Y) = H(X) + H(Y) - H(XY)$.

c) Interpretiert man $I(X; Z)$ als die verbleibende Unsicherheit hinsichtlich des Tupels Z, wenn die erste Komponente X bekannt ist, so gilt offensichtlich $I(X; Z) = H(Y) \underline{= 1.585 \text{ bit}}$.

Rein formal lässt sich diese Aufgabe auch wie folgt lösen:

- Die Entropie $H(Z)$ ist gleich $H(XY) = 3.170$ bit.
- Die Verbundwahrscheinlichkeit $P_{XZ}(X, Z)$ beinhaltet neun Elemente der Wahrscheinlichkeit 1/9, alle anderen sind mit Nullen belegt (linke Grafik) $\Rightarrow$ $H(XZ) = \log_2(9) = 3.170$ bit.
- Damit gilt für die Transinformation (gemeinsame Information der Zufalsgrößen X und Z):

$$I(X; Z) = H(X) + H(Z) - H(XZ) =$$
$$= 1.585 + 3.170 - 3.170 \underline{= 1.585\,(\text{bit})}\,.$$

$P_{XZ}(X, Z)$

X \ Z	00	01	02	10	11	12	20	21	22
0	1/9	1/9	1/9						
1				1/9	1/9	1/9			
2							1/9	1/9	1/9

$P_X(X)$
1/3
1/3
1/3

Alle Zahlenwerte in „bit“

$H(XZ) = 3.170$	
$H(X) = 1.585$	$H(Z\vert X) = 1.585$
$H(Z) = 3.170$	
$I(X;Z) = 1.585$	$H(Z\vert X) = 1.585$

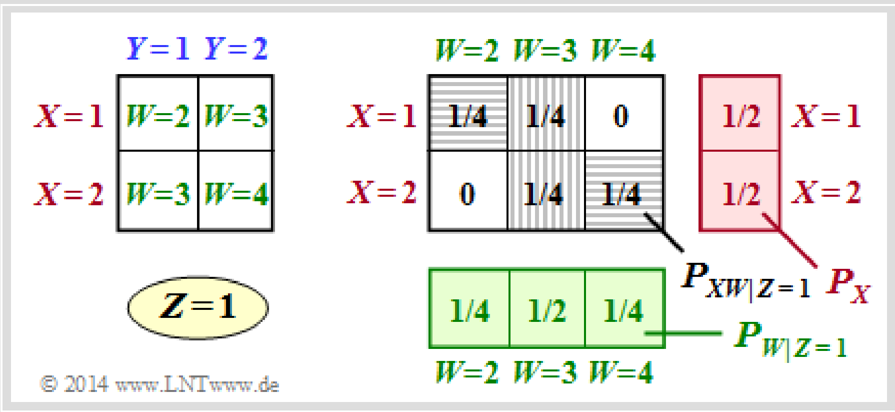
d) Entsprechend der rechten Grafik gilt:

$$H(Z|X) = H(XZ) - H(X) = 3.170 - 1.585 \underline{= 1.585\,(\text{bit})}\,,$$
$$H(X|Z) = H(XZ) - H(Z) = 3.170 - 3.170 \underline{= 0\,(\text{bit})}\,.$$

- $H(Z|X)$ gibt die Restunsicherheit hinsichtlich des Tupels Z an, wenn man die erste Komponente X kennt. Die Unsicherheit hinsichtlich des Tupels Z ist $H(Z) = 2 \cdot \log_2(3)$ bit, bei Kenntnis der Komponente X halbiert sich die Unsicherheit auf $H(Z|X) = \log_2(3)$ bit.
- $H(X|Z)$ gibt die verbleibende Unsicherheit hinsichtlich der Komponente X an, wenn man das Tupel $Z = (X, Y)$ kennt. Diese Unsicherheit ist natürlich 0: Kennt man Z, so kennt man auch X.

Musterlösung zur Aufgabe A3.8

a) Die folgende Grafik gilt für $Z = 1 \Rightarrow W = X + Y$. Unter den Voraussetzungen $P_X(X) = [1/2,\ 1/2]$ sowie $P_Y(Y) = [1/2,\ 1/2]$ ergeben sich somit die Verbundwahrscheinlichkeiten $P_{XW|Z=1}(X,\ W)$ entsprechend der rechten Grafik (graue Hinterlegung).



Damit gilt für die Transinformation unter der festen Bedingung $Z = 1$:

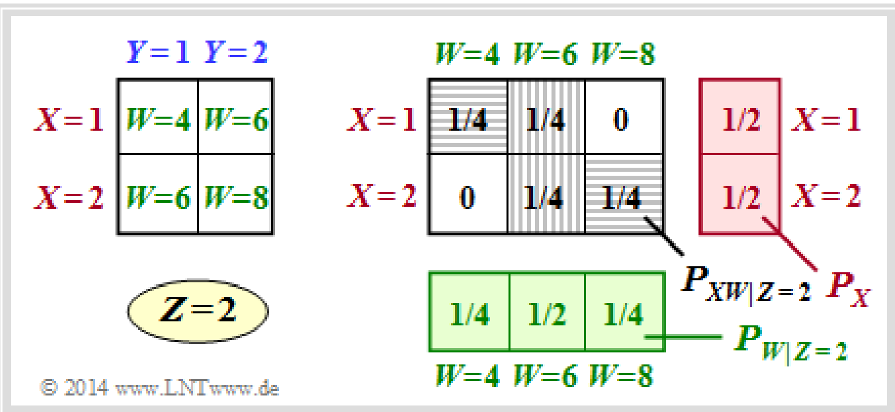
$$I(X;W\,|\,Z=1) = \sum_{(x,w)\,\in\,\text{supp}(P_{XW}\,|Z=1)} P_{XW\,|\,Z=1}(x,w) \cdot \log_2 \frac{P_{XW\,|\,Z=1}(x,w)}{P_X(x)\cdot P_{W\,|\,Z=1}(w)} =$$

$$= 2 \cdot \frac{1}{4} \cdot \log_2 \frac{1/4}{1/2 \cdot 1/4} + 2 \cdot \frac{1}{4} \cdot \log_2 \frac{1/4}{1/2 \cdot 1/2} \underline{= 0.5\,(\text{bit})}\,.$$

Der erste Term fasst die beiden horizontal schraffierten Felder in obiger Grafik zusammen, der zweite Term die vertikal schraffierten Felder. Letztere liefern wegen $\log_2(1) = 0$ keinen Beitrag.

b) Für $Z = 2$ gilt zwar $W = \{4,\ 6,\ 8\}$, aber hinsichtlich der Wahrscheinlichkeitsfunktionen ändert sich gegenüber der Teilaufgabe (a) nichts. Demzufolge erhält man auch die gleiche bedingte Transinformation:

$$I(X;W\,|\,Z=2) = I(X;W\,|\,Z=1) \underline{= 0.5\,(\text{bit})}\,.$$



c) Die angegebene Gleichung lautet für $Z = \{1, 2\}$ mit $\Pr(Z = 1) = p$ und $\Pr(Z = 2) = 1 - p$:

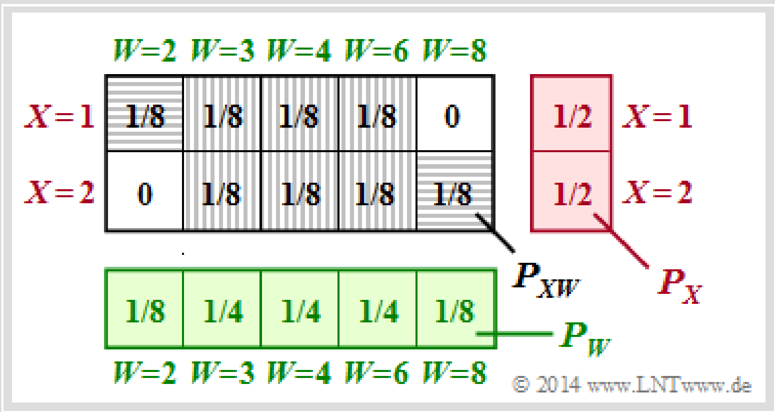
$$I(X;W\,|\,Z) = p \cdot I(X;W\,|\,Z=1) + (1-p) \cdot I(X;W\,|\,Z=2) \underline{= 0.5\,(\text{bit})}\,.$$

Es ist berücksichtigt, dass entsprechend den Teilaufgaben (a) und (b) die bedingten Transinformationen für gegebenes $Z = 1$ und gegebenes $Z = 2$ gleich sind. Damit ist $I(X;\ W|Z)$, also unter der Bedingung einer stochastischen Zufallsgröße $Z = \{1, 2\}$ mit $P_Z(Z) = [p,\ 1 - p]$, unabhängig von p. Das Ergebnis gilt insbesondere auch für <u>p = 1/2 und p = 3/4</u>.

d) Die Verbundwahrscheinlichkeiten $P_{XW}(\cdot)$ hängen auch von den Z–Wahrscheinlichkeiten p und $1 - p$ ab. Für Pr(Z = 1) = Pr(Z = 2) = 1/2 ergibt sich das nachfolgend skizzierte Schema. Zur Transinformation tragen nur wieder die beiden horizontal schraffierten Felder bei:

$$I(X;W) = 2 \cdot \frac{1}{8} \cdot \log_2 \frac{1/8}{1/2 \cdot 1/8} \underline{= 0.25\,(\mathrm{bit})} \quad < \quad I(X;W\,|\,Z)\,.$$

Das Ergebnis $I(X;\ W|Z) > I(X;\ W)$ trifft für dieses Beispiel, aber auch für viele andere Anwendungen zu:
Kenne ich Z, so weiß ich mehr über die 2D–Zufallsgröße XW als ohne diese Kenntnis. Man darf dieses Ergebnis aber nicht verallgemeinern.
Manchmal gilt tatsächlich $I(X;\ W) > I(X;\ W|Z)$, so wie im **Beispiel** im Theorieteil.



Musterlösung zur Aufgabe A3.9

a) Für die gesuchten Größen gilt allgemein bzw. mit den Zahlenwerten p_0 = 0.2 und ε = 0.1:

$$P_{XY}(0,0) = p_0 \cdot (1-\varepsilon) \underline{= 0.18}, \qquad P_{XY}(0,1) = p_0 \cdot \varepsilon \underline{= 0.02},$$
$$P_{XY}(1,0) = p_1 \cdot \varepsilon \underline{= 0.08}, \qquad P_{XY}(1,1) = p_1 \cdot (1-\varepsilon) \underline{= 0.72}.$$

b) Es gilt:

$$P_Y(Y) = \big(\Pr(Y=0)\,,\Pr(Y=1)\big) = (p_0\,, p_1) \cdot \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix}$$

$$\Rightarrow \quad \Pr(Y=0) = p_0 \cdot (1-\varepsilon) + p_1 \cdot \varepsilon = 0.2 \cdot 0.9 + 0.8 \cdot 0.1 \underline{= 0.26},$$
$$\Pr(Y=1) = p_0 \cdot \varepsilon + p_1 \cdot (1-\varepsilon) = 0.2 \cdot 0.1 + 0.8 \cdot 0.9 \underline{= 0.74}.$$

c) Für die Transinformation gilt gemäß der Definition mit p_0 = 0.2, p_1 = 0.8 und ε = 0.1:

$$I(X;Y) = \mathrm{E}\left[\log_2 \frac{P_{XY}(X,Y)}{P_X(X)\cdot P_Y(Y)}\right] = 0.18 \cdot \log_2 \frac{0.18}{0.2\cdot 0.26} + 0.02 \cdot \log_2 \frac{0.02}{0.2\cdot 0.74} +$$
$$+\, 0.08 \cdot \log_2 \frac{0.08}{0.8\cdot 0.26} + 0.72 \cdot \log_2 \frac{0.72}{0.8\cdot 0.74} \underline{= 0.3578\,\text{bit}}.$$

d) Mit der angegebenen Quellenentropie $H(X)$ erhält man für die Äquivokation:

$$H(X|Y) = H(X) - I(X;Y) = 0.7219 - 0.3578 \underline{= 0.3642\,\text{bit}}.$$

Man könnte auch die allgemeine Definition mit den Rückschlusswahrscheinlichkeiten $P_{X|Y}(\cdot)$ anwenden:

$$H(X|Y) = \mathrm{E}\left[\log_2 \frac{1}{P_{X|Y}(X|Y)}\right] = \mathrm{E}\left[\log_2 \frac{P_Y(Y)}{P_{XY}(X,Y)}\right].$$

Im Beispiel erhält man auch nach dieser Berechnungsvorschrift das gleiche Ergebnis $H(X|Y)$ = 0.3642 bit:

$$H(X|Y) = 0.18 \cdot \log_2 \frac{0.26}{0.18} + 0.02 \cdot \log_2 \frac{0.74}{0.02} + 0.08 \cdot \log_2 \frac{0.26}{0.08} + 0.72 \cdot \log_2 \frac{0.74}{0.72}.$$

e) Richtig ist der Lösungsvorschlag 2. Bei gestörter Übertragung ($\varepsilon > 0$) ist die Unsicherheit hinsichtlich der Sinke stets größer als die Unsicherheit bezüglich der Quelle. Man erhält hier als Zahlenwert:

$$H(Y) = H_{\text{bin}}(0.26) = 0.8268\,\text{bit}.$$

Bei fehlerfreier Übertragung (ε = 0) würde dagegen $P_Y(\cdot) = P_X(\cdot)$ und $H(Y) = H(X)$ gelten.

f) Auch hier ist der zweite Lösungsvorschlag richtig. Wegen

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$$

ist $H(Y|X)$ um den gleichen Betrag größer als $H(X|Y)$, um den auch $H(Y)$ größer ist als $H(X)$:

$$H(Y|X) = H(Y) - I(X;Y) = 0.8268 - 0.3578 = 0.4690\,\text{bit}.$$

Bei direkter Berechnung erhält man das gleiche Ergebnis $H(Y|X)$ = 0.4690 bit:

$$
\begin{aligned}
H(Y\,|X) \;&=\; \mathrm{E}\big[\log_2 \frac{1}{P_{Y|X}(Y\,|\,X)}\big] = \\
&=\; 0.18 \cdot \log_2 \frac{1}{0.9} + 0.02 \cdot \log_2 \frac{1}{0.1} + 0.08 \cdot \log_2 \frac{1}{0.1} + 0.72 \cdot \log_2 \frac{1}{0.9}\,.
\end{aligned}
$$

Musterlösung zur Zusatzaufgabe Z3.9

a) Aus der Tabelle auf der Angabenseite erkennt man, dass beim BSC–Modell (blaue Hinterlegung) wie auch beim allgemeineren (unsymmetrischen) BC–Modell (rote Hinterlegung) die Äquivokation $H(X|Y)$ von den Quellensymbolwahrscheinlichkeiten p_0 und p_1 abhängt. Daraus folgt, dass der Lösungsvorschlag 1 nicht richtig sein kann. Dagegen ist die Irrelevanz $H(Y|X)$ unabhängig von der Quellenstatistik, so dass auch der Lösungsvorschlag 3 ausgeschlossen werden kann.

Richtig ist vielmehr der Lösungsvorschlag 2, wie die folgende Rechnung zeigt:

$$\begin{aligned} H(Y|X) &= p_0 \cdot (1-\varepsilon) \cdot \log_2 \frac{1}{1-\varepsilon} + p_0 \cdot \varepsilon \cdot \log_2 \frac{1}{\varepsilon} + \\ &= p_1 \cdot \varepsilon \cdot \log_2 \frac{1}{\varepsilon} + p_1 \cdot (1-\varepsilon) \cdot \log_2 \frac{1}{1-\varepsilon} = \\ &= (p_0 + p_1) \cdot \left[\varepsilon \cdot \log_2 \frac{1}{\varepsilon} + (1-\varepsilon) \cdot \log_2 \frac{1}{1-\varepsilon} \right]. \end{aligned}$$

Mit $p_0 + p_1 = 1$ und der **binären Entropiefunktion** ⇒ $H_{\rm bin}$ erhält man das vorgeschlagene Ergebnis:

$$H(Y|X) = H_{\rm bin}(\varepsilon)\,.$$

Für $\varepsilon = 0.1$ ergibt sich $H(Y|X) = 0.4690$ bit. Der gleiche Wert steht für alle p_0 in der gegebenen Tabelle.

b) Zutreffend sind hier alle vorgegebenen Lösungsalternativen. Die Kanalkapazität ist definiert als die maximale Transinformation, wobei die Maximierung hinsichtlich $P_X = (p_0, p_1)$ zu erfolgen hat:

$$C = \max_{P_X(X)} \; I(X;Y)\,.$$

Diese Gleichung gilt allgemein, also auch für den rot hinterlegten **unsymmetrischen Binärkanal** (BC).

Die Transinformation kann zum Beispiel allgemein wie folgt berechnet werden:

$$I(X;Y) = H(Y) - H(Y|X)\,,$$

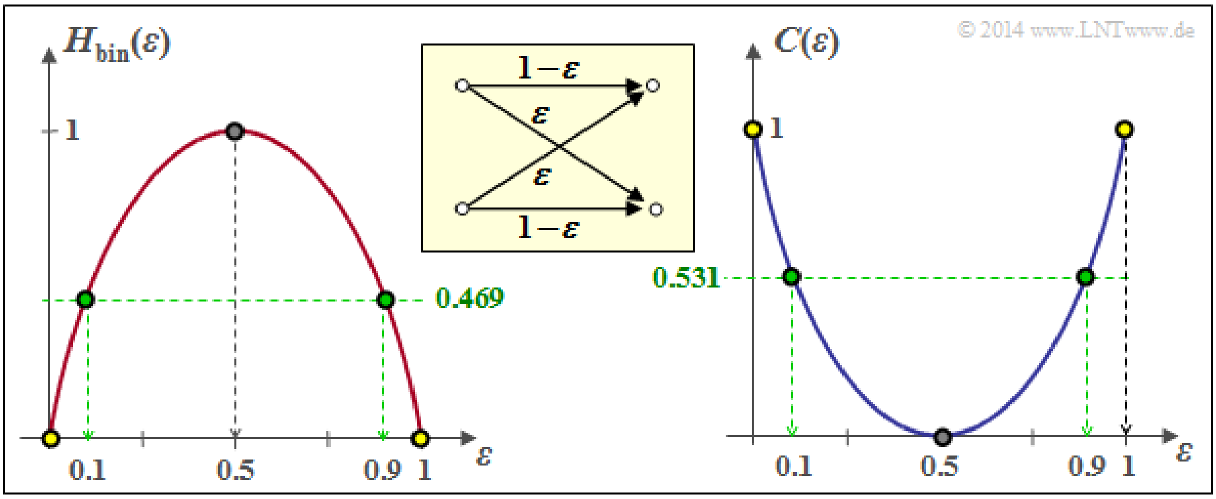
wobei entsprechend Teilaufgabe (a) der Term $H(Y|X)$ unabhängig von p_0 bzw. $p_1 = 1 - p_0$ ist. Die maximale Transinformation ergibt sich genau dann, wenn die Sinkenentropie $H(Y)$ maximal ist. Dies ist der Fall für $p_0 = p_1 = 1/2$ ⇒ $H(X) = H(Y) = 1$ bit.

c) Entsprechend den Teilaufgaben (a) und (b) erhält man für die BSC–Kanalkapazität:

$$C_{\rm BSC} = 1\,({\rm bit}) - H_{\rm bin}(\varepsilon)\,.$$

Die Grafik zeigt links die binäre Entropiefunktion und rechts die Kanalkapazität. Man erhält:

- für $\varepsilon = 0$ (fehlerfreier Kanal): $C = 1$ (bit) ⇒ Punkt mit gelber Füllung,
- für $\varepsilon = 0.1$ (bisher betrachtet): $C = 0.531$ (bit) ⇒ Punkt mit grüner Füllung,
- für $\varepsilon = 0.5$ (vollkommen gestört): $C = 0$ (bit) ⇒ Punkt mit grauer Füllung.



d) Aus der Grafik erkennt man, dass aus informationstheoretischer Sicht $\varepsilon = 1$ gleich ist wie $\varepsilon = 0$:

$$C_{\rm BSC}\Big|_{\varepsilon=1} = C_{\rm BSC}\Big|_{\varepsilon=0} \underline{= 1\,({\rm bit})}\,.$$

Der Kanal nimmt hier nur eine Umbenennung vor. Man spricht von *Mapping*. Aus jeder „0” wird eine „1” und aus jeder „1” eine „0”. Entsprechend gilt:

$$C_{\rm BSC}\Big|_{\varepsilon=0.9} = C_{\rm BSC}\Big|_{\varepsilon=0.1} \underline{= 0.531\,({\rm bit})}\,.$$

Musterlösung zur Aufgabe A3.10

a) Aufgrund der Symmetrie der Übergangswahrscheinlichkeiten $P_{Y|X}(Y|X)$ ist offensichtlich, dass eine Gleichverteilung zur maximalen Transinformation $I(X;\, Y)$ und damit zur Kanalkapazität C führen wird:

$$P_X(X) = P_X\big(X = 1, X = 2, \ldots, X = M\big) = \big[1/M, 1/M, \ldots, 1/M\big].$$

Richtig ist also der Lösungsvorschlag 2. Im Sonderfall $M = 2$ wäre auch $P_X(X) = (0.5, 0.5)$ richtig.

b) Von jedem Quellensymbol $X = \mu$ kommt man sowohl zum Sinkensymbol $Y = \mu$ als auch zum *Erasure* $Y = \text{E}$ ⇒ Zutreffend ist dementsprechend der Lösungsvorschlag 3 ⇒ Genau $2M$ Verbindungen.

c) Alle Wahrscheinlichkeiten $\Pr(Y = 1), \ldots, \Pr(Y = M)$ sind gleich groß. Man erhält für $\mu = 1, \ldots, M$:

$$\Pr(Y = \mu) = \frac{1-\lambda}{M}.$$

Außerdem kommt man von jedem Quellensymbol $X = 1, \ldots, X = M$ auch zum *Erasure* $Y = \text{E}$:

$$\Pr(Y = \text{E}) = \lambda.$$

Die Kontrolle ergibt, dass die Summe aller $M + 1$ Sinkensymbolwahrscheinlichkeiten tatsächlich 1 ergibt. Daraus folgt für die Sinkenentropie:

$$H(Y) = M \cdot \frac{1-\lambda}{M} \cdot \log_2 \frac{M}{1-\lambda} \;+\; \lambda \cdot \log_2 \frac{1}{\lambda}.$$

Zusammengefasst ergibt dies mit der **binären Entropiefunktion:**

$$H(Y) = (1-\lambda) \cdot \log_2 M \;+\; H_{\rm bin}(\lambda),$$

und mit $M = 4$ sowie $\lambda = 0.2$:

$$H(Y) = 1.6\,\text{bit} + H_{\rm bin}(0.2) \;\underline{= 2.322\,\text{bit}}.$$

d) Für die $2M$ Verbundwahrscheinlichkeiten ⇒ $p_{\mu\kappa} = \Pr\,[(X = \mu) \cap (Y = \kappa)] \ne 0$ und die zugehörigen bedingten Wahrscheinlichkeiten ⇒ $p_{\kappa|\mu} = \Pr(Y = \kappa|X = \mu)$ gilt:

- Die Kombination $p_{\mu\kappa} = (1 - \lambda)/M$ und $p_{\kappa|\mu} = 1 - \lambda$ kommt M mal vor.
- Die Kombination $p_{\mu\kappa} = \lambda/M$ und $p_{\kappa|\mu} = \lambda$ kommt ebenfalls M mal vor.

Daraus folgt:

$$\begin{aligned} H(Y|X) &= M \cdot \frac{1-\lambda}{M} \cdot \log_2 \frac{1}{1-\lambda} \;+\; M \cdot \frac{\lambda}{M} \cdot \log_2 \frac{1}{\lambda} = \\ &= (1-\lambda) \cdot \log_2 \frac{1}{1-\lambda} \;+\; \lambda \cdot \log_2 \frac{1}{\lambda} = H_{\rm bin}(\lambda). \end{aligned}$$

Das Ergebnis ist unabhängig von M. Mit $\lambda = 0.2$ erhält man:

$$H(Y|X) = H_{\rm bin}(0.2) \;\underline{= 0.722\,\text{bit}}.$$

e) Die Kanalkapazität C ist gleich der maximalen Transinformation $I(X;\, Y)$, wobei die Maximierung hinsichtlich $P_X(X)$ bereits durch den symmetrischen Ansatz berücksichtigt wurde:

$$\begin{aligned} C &= \max_{P_X(X)} I(X;Y) = H(Y) - H(Y|X) = \\ &= (1-\lambda)\cdot \log_2 M + H_{\rm bin}(\lambda) - H_{\rm bin}(\lambda) = (1-\lambda)\cdot \log_2 M \end{aligned}$$

$$\Rightarrow \quad M = 4: \quad \underline{C = 1.6\ \text{bit}},$$
$$M = 2: \quad \underline{C = 0.8\ \text{bit}}.$$

f) Der BEC–Kanal ist ein Sonderfall des hier betrachteten allgemeinen Modells mit $M = 2$:

$$C_{\rm BEC} = 1 - \lambda.$$

Richtig ist somit der Lösungsvorschlag 1. Der zweite Lösungsvorschlag gilt dagegen für den **Binary Symmetric Channel** (BSC) mit der Verfälschungswahrscheinlichkeit λ.

Musterlösung zur Zusatzaufgabe Z3.10

a) Die Quellenentropie ergibt sich entsprechend der binären Entropiefunktion:

$$H(X) = H_{\rm bin}(p_0) = H_{\rm bin}(0.4) \ \underline{= 0.971 \ {\rm bit}}\,.$$

b) Die Wahrscheinlichkeiten der Sinkensymbole sind:

$$P_Y(1) = p_1/2 = (1-p_0)/2 = 0.3\,,\ \ P_Y(0) = 1 - P_Y(1) = p_1/2 = (1-p_0)/2 = 0.7$$

$$\Rightarrow \quad H(Y) = H_{\rm bin}(\frac{1+p_0}{2}) = H_{\rm bin}(0.7) \ \underline{= 0.881 \ {\rm bit}}\,.$$

c) Die Verbundwahrscheinlichkeiten $p_{\mu\kappa} = \Pr[(X = \mu) \cap (Y = \kappa)]$ ergeben sich zu:

$$p_{00} = p_0\,,\quad p_{01} = 0\,,\quad p_{10} = (1-p_0)/2\,,\quad p_{11} = (1-p_0)/2$$

$$\begin{aligned}\Rightarrow \quad H(XY) &= p_0 \cdot \log_2 \frac{1}{p_0} + 2 \cdot \frac{1-p_0}{2} \cdot \log_2 \frac{2}{1-p_0} = \\ &= p_0 \cdot \log_2 \frac{1}{p_0} + (1-p_0) \cdot \log_2 \frac{1}{1-p_0} + (1-p_0) \cdot \log_2 (2) = \\ &= H_{\rm bin}(p_0) + 1 - p_0\,.\end{aligned}$$

Das numerische Ergebnis für $p_0 = 0.4$ ist somit:

$$H(XY) = H_{\rm bin}(0.4) + 0.6 = 0.971 + 0.6 \ \underline{= 1.571 \ {\rm bit}}\,.$$

d) Eine (mögliche) Gleichung zur Berechnung der Transinformation lautet:

$$I(X;Y) = H(X) + H(Y) - H(XY)\,.$$

Daraus erhält man mit den Ergebnissen der Teilaufgaben (a), (b) und (c):

$$I(X;Y) = H_{\rm bin}(p_0) + H_{\rm bin}(\frac{1+p_0}{2}) - H_{\rm bin}(p_0) - 1 + p_0 = H_{\rm bin}(\frac{1+p_0}{2}) - 1 + p_0$$

$$\Rightarrow \quad p_0 = 0.4\text{:} \quad I(X;Y) = H_{\rm bin}(0.7) - 0.6 = 0.881 - 0.6 \ \underline{= 0.281 \ {\rm bit}}\,.$$

e) Die Kanalkapazität C ist die Transinformation $I(X;\,Y)$ bei bestmöglichen Wahrscheinlichkeiten p_0 und p_1 der Quellensymbole. Nach Differentiation erhält man die Bestimmungsgleichung:

$$\frac{\rm d}{{\rm d}p_0}\, I(X;Y) = \frac{\rm d}{{\rm d}p_0}\, H_{\rm bin}(\frac{1+p_0}{2}) + 1 \stackrel{!}{=} 0\,.$$

Mit dem Differentialquotienten der binären Entropiefunktion,

$$\frac{\rm d}{{\rm d}p}\, H_{\rm bin}(p) = \log_2 \frac{1-p}{p}\,,$$

und entsprechendes Nachdifferenzieren erhält man:

$$\frac{1}{2} \cdot \log_2 \frac{(1-p_0)/2}{1-(1-p_0)/2} + 1 \stackrel{!}{=} 0 \quad \Rightarrow \quad \frac{1}{2} \cdot \log_2 \frac{(1-p_0)/2}{(1+p_0)/2} + 1 \stackrel{!}{=} 0$$

$$\Rightarrow \quad \log_2 \frac{1+p_0}{1-p_0} \stackrel{!}{=} 2 \quad \Rightarrow \quad \frac{1+p_0}{1-p_0} \stackrel{!}{=} 4 \quad \Rightarrow \quad p_0 \ \underline{= 0.6}\,.$$

f) Für die Kanalkapazität gilt dementsprechend:

$$C = I(X;Y)\big|_{p_0\,=\,0.6} = H_{\rm bin}(0.8) - 0.4 = 0.722 - 0.4 \underline{= 0.322\,{\rm bit}}\,.$$

In **Aufgabe A3.13** wird dieses Ergebnis im Vergleich zum BSC–Kanalmodell interpretiert.

g) Für die Äquivokation gilt:

$$H(X\,|Y) = H(X) - I(X;Y) = 0.971 - 0.322 \underline{= 0.649\,{\rm bit}}\,.$$

Wegen $H_{\rm bin}(0.4) = H_{\rm bin}(0.6)$ ergibt sich die gleiche Quellenentropie $H(X)$ wie in Teilaufgabe (a). Die Sinkenentropie muss neu berechnet werden. Mit $p_0 = 0.6$ erhält man $H(Y) = H_{\rm bin}(0.8) = 0.722$ bit, und damit ergibt sich für die Irrelevanz:

$$H(Y\,|X) = H(Y) - I(X;Y) = 0.722 - 0.322 \underline{= 0.400\,{\rm bit}}\,.$$

Musterlösung zur Aufgabe A3.11

a) Teilkanal A ist ein BSC mit der Verfälschungswahrscheinlichkeit q ⇒ Lösungsvorschlag 1.

b) Teilkanal B ist ein Auslöschungskanal. Sowohl die Sinkenentropie als auch die Irrelevanz dieses Teilkanals sind 0 ⇒ $C_{\rm B} = 0$ ⇒ Lösungsvorschlag 3.

c) Die Kapazität C des Gesamtkanals kann mit der angegebenen Gleichung berechnet werden:

$$C = p_{\rm A} \cdot C_{\rm A} + p_{\rm B} \cdot C_{\rm B} = p_{\rm A} \cdot [1 - H_{\rm bin}(q)]\,.$$

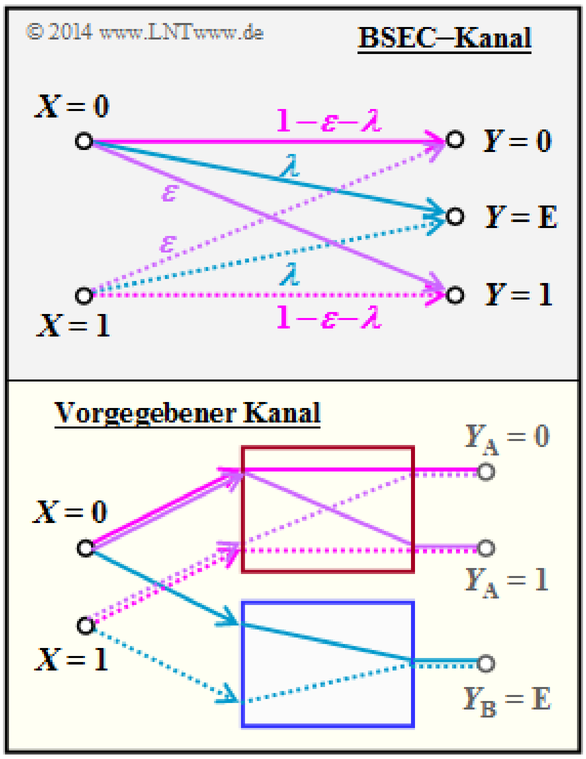
Hier stimmt somit der Lösungsvorschlag 2.

d) Beim bisher betrachteten Modell ergeben sich folgende Übergangswahrscheinlichkeiten:

$${\rm Pr}(Y = {\rm E}\,|\,X = 0) = {\rm Pr}(Y = {\rm E}\,|\,X = 1) = p_{\rm B}\,.$$

Beim BSEC–Modell sind die entsprechenden bedingten Wahrscheinlichkeiten gleich λ ⇒ siehe **Grafik** auf der Angabenseite. Richtig ist also der Lösungsvorschlag 2:

$$p_{\rm B} = \lambda = 1 - p_{\rm A} \quad \Rightarrow \quad p_{\rm A} = 1 - \lambda\,.$$



e) Beim BSEC–Modell gilt beispielsweise:

$${\rm Pr}(Y = 1\,|\,X = 0) = \varepsilon\,.$$

Dagegen ergibt sich bei unserem Hilfsmodell

$${\rm Pr}(Y = 1\,|\,X = 0) = (1 - \lambda) \cdot q\,.$$

Damit erhält man $q = \varepsilon/(1 - \lambda)$ ⇒ Lösungsvorschlag 4. Die Grafik verdeutlicht anhand von Farben und Strichart (durchgezogen/gepunktet) den Zusammenhang zwischen beiden Modellen.

f) Mit den Ergebnissen der Teilaufgaben (c), (d) und (e) erhält man allgemein:

$$C_{\rm BSEC} = (1 - \lambda) \cdot \left[1 - H_{\rm bin}\left(\frac{\varepsilon}{1 - \lambda}\right)\right]\,,$$

bzw. für $\varepsilon = 0.08$ und $\lambda = 0.2$:

$$C_{\rm BSEC} = 0.8 \cdot [1 - H_{\rm bin}(0.1)] = 0.8 \cdot [1 - 0.469] \;\underline{= 0.425\,{\rm bit}}\,.$$

g) Der BSC ist ein Sonderfall des BSEC mit $\lambda = 0$:

$$C_{\rm BSC} = 1 - H_{\rm bin}(\varepsilon) = 1 - H_{\rm bin}(0.08) = 1 - 0.402 \;\underline{= 0.598\,{\rm bit}}\,.$$

h) Der BEC ist ein Sonderfall des BSEC mit $\varepsilon = 0$:

$$C_{\rm BEC} = (1 - \lambda) \cdot [1 - H_{\rm bin}(0)] = 1 - \lambda\,.$$

Mit $\lambda = 0.2$ ergibt sich hierfür $C_{\rm BEC}$ $\underline{= 0.8\,{\rm bit}}$.

Musterlösung zur Aufgabe A3.12

a) Entsprechend der vorgegebenen Gleichung gilt:

$$p_{\rm B,\,min} = H_{\rm bin}^{-1}\left(1 - \frac{C}{R}\right) = H_{\rm bin}^{-1}(0.75)\,.$$

Aus der Grafik auf der Angabenseite (blaue Markierung) lässt sich ablesen:

$$p_{\rm B,\,min} \;\underline{= 0.215}\,.$$

b) Durch Umstellung obiger Gleichung erhält man mit $H_{\rm bin}(p_{\rm B})$ aus der **Grafik:**

$$R_{\rm max} = \frac{C}{1 - H_{\rm bin}(p_{\rm B})} \quad\Rightarrow\quad p_{\rm B} = 0.10:\;\; R_{\rm max} = \frac{0.5}{1 - 0.469}\;\underline{= 0.942}\,,$$
$$\Rightarrow\quad p_{\rm B} = 0.05:\;\; R_{\rm max} = \frac{0.5}{1 - 0.286}\;\underline{= 0.700}\,.$$

c) Für die binäre Entropiefunktion gilt definitionsgemäß:

$$H_{\rm bin}(p) = p \cdot \log_2 \frac{1}{p} + (1-p)\cdot \log_2 \frac{1}{1-p}\,.$$

Beide Terme sind positiv. Daraus folgt der <u>Lösungsvorschlag 2</u>:

$$H_{\rm bin}(p) > p \cdot \log_2 \frac{1}{p}\,.$$

Die Differenz ist durch den zweiten Term der binären Entropiefunktion gegeben:

$$H_{\rm bin}(p) - p \cdot \log_2 \frac{1}{p} = (1-p)\cdot \log_2 \frac{1}{1-p}\,.$$

Dieser Term wird umso kleiner, je kleiner p ist. Zutreffend ist also <u>auch der Lösungsvorschlag 3</u>:

$$p_{\rm B} = 10^{-3}:\quad 0.999 \cdot \log_2 \frac{1}{0.999} = 1.4420 \cdot 10^{-3}\,,$$
$$p_{\rm B} = 10^{-4}:\quad 0.9999 \cdot \log_2 \frac{1}{0.9999} = 1.4427 \cdot 10^{-4}\,,$$
$$p_{\rm B} = 10^{-5}:\quad 0.99999 \cdot \log_2 \frac{1}{0.99999} = 1.4427 \cdot 10^{-5}\,.$$

d) Der Ausdruck R/C gibt die obere Schranke gemäß dem inversen Kanalcodierungstheorem an. Für $p_{\rm B} = 10^{-3}$ erhält man folgende Schranke: $R/C \le 1.01154 \;\;\Rightarrow\;\; R/C - 1 \le 1.154 \cdot 10^{-2}$.

- Beim Vorschlag 2 ist die Näherung gemäß (c) berücksichtigt. Da nun der Nenner vergrößert wird, ist $R'/C < R/C$, beispielsweise gilt für $p_{\rm B} = 10^{-3}$: $R'/C = 1.01007 \cdot 10^{-2}$. Es handelt sich auch hier um eine obere Schranke, die etwas sicherer ist als die erste Schranke R/C.
- Die Schranke $R'' < R'$ ergibt sich, wenn man $1/(1 - \varepsilon)$ durch $1 + \varepsilon$ ersetzt (ε ist positiv). Für $p_{\rm B} = 10^{-3}$ erhält man nun beispielsweise $R''/C = 1.00997 \;\;\Rightarrow\;\; R''/C - 1 \le 0.997 \cdot 10^{-2}$.
- R''' ist identisch mit R'' und für Überschlagsrechnungen bei ganzzahligem Exponenten i gut geeignet. Für den Zehnerlogarithmus gilt der Zahlenwert $\lg(2) \approx 0.30103$.

i	3	4	5	6	7	8
p_B	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
$R/C-1$	$1.154 \cdot 10^{-2}$	$1.475 \cdot 10^{-3}$	$1.896 \cdot 10^{-4}$	$2.138 \cdot 10^{-5}$	$2.472 \cdot 10^{-6}$	$2.832 \cdot 10^{-7}$
$R'''/C-1$	$0.997 \cdot 10^{-2}$	$1.328 \cdot 10^{-3}$	$1.661 \cdot 10^{-4}$	$1.993 \cdot 10^{-5}$	$2.325 \cdot 10^{-6}$	$2.655 \cdot 10^{-7}$

Alle Vorschläge sind richtig. Die Tabelle zeigt für verschiedene Bitfehlerwahrscheinlichkeiten p_B die Schranke R/C und die 3. Näherung. Man erkennt die gute Übereinstimmung der beiden Schranken.

Musterlösung zur Aufgabe A3.13

a) Die BSC–Fehlerwahrscheinlichkeit ist mit $p_0 = p_1 = 0.5$ bei uncodierter Übertragung ($R = 1$):

$$p_\mathrm{B} = 0.5 \cdot 0.25 + 0.5 \cdot 0.25 = 0.25\,.$$

Entsprechend gilt bei gleichen Randbedingungen für das EUC–Modell:

$$p_\mathrm{B} = 0.5 \cdot 0 + 0.5 \cdot 0.5 = 0.25\,.$$

Richtig ist demnach der Lösungsvorschlag 3.

b) Beim BSC–Modell mit der Verfälschungswahrscheinlichkeit $\varepsilon = 0.25$ gilt bei uncodierter Übertragung ($R = 1$) unabhängig von p_0 und p_1 für die Bitfehlerwahrscheinlichkeit: $p_\mathrm{B} = 0.25$. Dagegen erhält man beim EUC–Modell beispielsweise mit $p_0 = 0.6$ und $p_1 = 0.4$ eine kleinere Bitfehlerwahrscheinlichkeit:

$$p_\mathrm{B} = 0.6 \cdot 0 + 0.4 \cdot 0.5 = 0.2\,.$$

Richtig ist demnach der Lösungsvorschlag 3.

Zu beachten ist jedoch, dass nun die Quellenentropie nicht mehr $H(X) = 1$ (bit) beträgt, sondern nur mehr $H(X) = H_\mathrm{bin}(0.6) = 0.971$ (bit). Im Grenzfall $p_0 = 1$ werden nur noch Nullen übertragen und es gilt $H(X) = 0$. Für die Bitfehlerwahrscheinlichkeit gilt dann aber tatsächlich:

$$p_\mathrm{B} = 1 \cdot 0 + 0 \cdot 0.5 = 0\,.$$

Man überträgt also keine Information, diese aber mit der Bitfehlerwahrscheinlichkeit 0.

c) Aus der Grafik auf der Angabenseite lässt sich für die Kapazitäten der beiden Kanäle ablesen:

- $C_\mathrm{BSC} = 0.1887$ bit/use,
- $C_\mathrm{EUC} = 0.3219$ bit/use.

Nach dem Kanalcodierungstheorem kann unter der Bedingung $R \le C$ eine Kanalcodierung gefunden werden, mit der die Fehlerwahrscheinlichkeit zu 0 gemacht werden kann. Bei beiden Kanälen trifft dies mit $R = 0.16$ zu $\Rightarrow$ Lösungsvorschlag 1.

d) Richtig ist hier der Lösungsvorschlag 3: Beim EUC–Modell wird mit $R = 0.32$ und $C = 0.3219$ die notwendige Bedingung $R \le C$ für eine fehlerfreie Übertragung erfüllt. Voraussetzung hierfür ist allerdings die Wahrscheinlichkeitsfunktion $P_X(X) = (0.6,\ 0.4)$. Dagegen ergäbe sich für gleichwahrscheinliche Symbole $\Rightarrow$ $P_X(X) = (0.5,\ 0.5)$ die Transinformation $I(X;\,Y) = 0.3113 < R$.

Man erkennt: Das EUC–Modell bietet mehr Potenzial für die Anwendung einer Kanalcodierung als das BSC–Modell. Hier kann beispielsweise im Code ausgenutzt werden, dass eine gesendete „0” stets fehlerfrei übertragen wird.

e) Aus der Kommentierung der Aufgaben (c) und (d) geht hervor, dass der Lösungsvorschlag 4 zutrifft.

Musterlösung zur Aufgabe A3.14

a) Beide Prozessoren beschreiben **streng symmetrische Kanäle** ⇒ sowohl *gleichmäßig dispersiv* als auch *gleichmäßig fokussierend*. Für einen solchen Binärkanal gilt mit $Y = \{0, 1\}$ ⇒ $|Y| = 2$:

$$I(X;Y) = 1 + \sum_{y \in Y} P_{Y|X}(y|x) \cdot \log_2 P_{Y|X}(y|x).$$

Hierbei ist es völlig egal, ob man von $X = 0$ oder von $X = 1$ ausgeht. Für $X = 0$ erhält man mit $P_{Y|X}(Y = 1|X = 0) = p$ und $P_{Y|X}(Y = 0|X = 0) = 1 - p$:

$$I(X;Y) = 1 + p \cdot \log_2 (p) + (1-p) \cdot \log_2 (1-p) = 1 - H_{\rm bin}(p)\,,$$
$$\text{mit} \qquad H_{\rm bin}(p) = p \cdot \log_2 \frac{1}{p} + (1-p) \cdot \log_2 \frac{1}{1-p}\,.$$

Das Ergebnis gilt allerdings nur für $P_X(X) = [0.5, 0.5]$ ⇒ maximale Transinformation ⇒ Kanalkapazität. Andernfalls ist $I(X; Y)$ kleiner. Beispielsweise gilt für $P_X(X) = [1, 0]$: $H(X) = 0$ ⇒ $I(X; Y) = 0$.

Richtig ist also nur der <u>Lösungsvorschlag 1</u>. Die binäre Entropiefunktion ist zwar konkav, aber diese Eigenschaft wurde bei der Herleitung nicht benutzt.

b) Für den Prozessor 1 ergibt sich mit $p = 0.1$:

$$I(X;Y) = 1 - H_{\rm bin}(0.1) = 1 - 0.469 \; \underline{= 0.531\,{\rm (bit)}}\,.$$

c) Entsprechend gilt für den zweiten Prozessor mit $q = 0.2$:

$$I(Y;Z) = 1 - H_{\rm bin}(0.2) = 1 - 0.722 \; \underline{= 0.278\,{\rm (bit)}}\,.$$

d) Die Wahrscheinlichkeit für $Z = 0$ unter der Bedingung $X = 0$ ergibt sich über zwei Wege zu

$$P(Z=0\,|\,X=0) = (1-p) \cdot (1-q) + p \cdot q = 1 - p - q + 2pq \stackrel{!}{=} 1 - \varepsilon\,.$$

Das Gesamtsystem hat dann die genau gleiche BSC–Struktur wie die Prozessoren 1 und 2, aber nun mit der Verfälschungswahrscheinlichkeit

$$\varepsilon = p + q - 2pq\,.$$

Für $p = 0.1$ und $q = 0.2$ erhält man

$$\varepsilon = 0.1 + 0.2 - 2 \cdot 0.1 \cdot 0.2 = 0.26$$
$$\Rightarrow \quad I(X;Z) = 1 - H_{\rm bin}(0.26) = 1 - 0.827 \; \underline{= 0.173\,{\rm (bit)}}\,.$$

e) Die Antwort ist natürlich <u>JA</u>, da beim *Data Processing Theorem* genau von den hier gegebenen Voraussetzungen ausgegangen wird. Wir wollen aber zusätzlich einige numerische Ergebnisse auswerten:

- Gilt $0 \le p < 0.5$ und $0 \le q < 0.5$, so erhält man:

$$\varepsilon = p + q \cdot (1 - 2p) > p \quad \Rightarrow \quad I(X;Z) < I(X;Y)\,,$$
$$\varepsilon = q + p \cdot (1 - 2q) > q \quad \Rightarrow \quad I(X;Z) < I(Y;Z)\,.$$

- Für $p = 0.5$ gilt unabhängig von q, da $I(X; Z)$ nicht größer sein kann als $I(X; Y)$:

$$\varepsilon = 0.5 + q \cdot (1 - 1) = 0.5 \quad \Rightarrow \quad I(X;Y) = 0 \quad \Rightarrow \quad I(X;Z) = 0\,.$$

Ebenso erhält man mit $q = 0.5$ unabhängig von p:

$$\varepsilon = 0.5 + p \cdot (1-1) = 0.5 \quad \Rightarrow \quad I(Y;Z) = 0 \quad \Rightarrow \quad I(X;Z) = 0\,.$$

- Auch für $p < 0.5$ und $q > 0.5$ wird das *Data Processing Theorem* nicht verletzt, was hier nur an einem Beispiel gezeigt werden soll. Mit $p = 0.1$ und $q = 0.8$ erhält man das gleiche Ergebnis wie in Teilaufgabe (d):

  $$\varepsilon = 0.1 + 0.8 - 2 \cdot 0.1 \cdot 0.8 = 0.74$$
  $$\Rightarrow \quad I(X;Z) = 1 - H_{\rm bin}(0.74) = 1 - H_{\rm bin}(0.26) = 0.173\,({\rm bit})\,.$$