Aufgaben zum Tut am 18.06.2007

Thomas Pajor

17. Juni 2007

Aufgabe 1.

„Ein Ausgleich von Wahrscheinlichkeiten führt immer zu einer Erhöhung des Informationsgehalts.“

Gegeben seien die Wahrscheinlichkeiten $p_1, \ldots, p_m$ mit $p_1 > p_2$. Definiere dazu

$$\begin{aligned} p'_1 &= p_1 - \Delta p \\ p'_2 &= p_2 + \Delta p \\ p'_i &= p_i & \text{für } i = 3, \ldots, m \end{aligned}$$

wobei $\Delta p > 0$ und $p_1 - \Delta p \geq p_2 + \Delta p$.

Zeigen Sie, dass gilt

$$H(p'_1, p'_2, \ldots, p'_m) > H(p_1, p_2, \ldots, p_m)$$

Lösung.

Wir lösen die Aufgabe durch Nachrechnen.

$$
\begin{aligned}
H(p'_1,\ldots,p'_m) &= -(p_1-\Delta p)\log(p_1-\Delta p)-(p_2+\Delta p)\log(p_2+\Delta p)\\
&\quad -\sum_{i=3}^{m} p_i\log p_i\\
&= -p_1\log(p_1-\Delta p)-p_2\log(p_2+\Delta p)-\sum_{i=3}^{m} p_i\log p_i\\
&\quad \underbrace{+\,\Delta p\log\underbrace{\frac{p_1-\Delta p}{p_2+\Delta p}}_{\geq 1}}_{\geq 0}\\
&\geq -p_1\log(p_1-\Delta p)-p_2\log(p_2+\Delta p)-\sum_{i=3}^{m} p_i\log p_i\\
&> -\sum_{i=1}^{m} p_i\log p_i \qquad \text{nach Lemma aus Vorlesung}\\
&= H(p_1,\ldots,p_m)
\end{aligned}
$$

Damit erhalten wir insgesamt

$$H(p'_1,\ldots,p'_m) > H(p_1,\ldots,p_m).$$

Anmerkung. Das Lemma auf das hier Bezug genommen wird ist das Folgende:

Lemma. *Sind* $p_1,\ldots,p_m,q_1,\ldots,q_m \geq 0$ *mit*

$$\sum_{i=1}^{m} q_i \leq \sum_{i=1}^{m} p_i = 1$$

Dann gilt

$$-\sum_i p_i\log p_i \leq -\sum_i p_i\log q_i$$

wobei Gleichheit nur gilt falls $p_i = q_i$ *für alle* $i = 1,\ldots,m$.

Das Lemma lässt sich für unseren Fall instanziieren wenn wir $q_1 := p_1-\Delta p$ und $q_2 := p_2+\Delta p$ wählen. Hier gilt sogar

$$\sum q_i = \sum p_i = 1,$$

wegen $\Delta p > 0$ jedoch nicht $p_i = q_i$ für alle i. Somit sind alle Voraussetzungen an das Lemma erfüllt und wir erhalten im obigen Beweis das echte „>“-Zeichen.

Aufgabe 2.

Gegeben sei folgender binärer, asymetrischer Kanal über den Alphabeten $X = Y = \{0, 1\}$:

$$\begin{array}{ll} P(0|0) = 1 - \beta & P(1|0) = \beta \\ P(0|1) = 0 & P(1|1) = 1 \end{array}$$

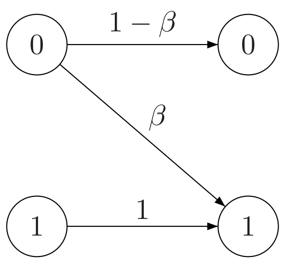
(a) Für welches β hat der Kanal maximale Kapazität?

(b) Für welches $H(X)$ wird bei gegebener Wahrscheinlichkeitsverteilung β der Wert von $H(Y)$ maximal?

(c) Berechnen Sie Fehlinformation, Äquivokation und Transinformation für $\beta = 0.9$ unter Annahme von Gleichverteilung auf X.

Lösung.

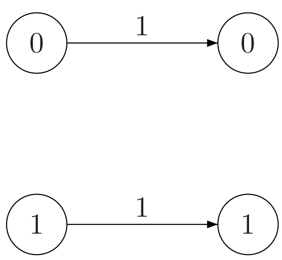
Wir machen uns zunächst den Übertragungskanal an einem Diagramm anschaulich klar:



(a) Die Kapazität des Kanals ist definiert durch

$$C := \max_{P(X)} \{H(X;Y)\}$$

Dabei ist $H(X;Y)$ die Transinformation, also die tatsächlich übertragene Information. Offensichtlich wird die Transinformation maximal, wenn keine Störungen auftreten. Wählen wir $\beta = 0$ so ergibt sich folgende Situation:



Jedes Zeichen wird mit Wahrscheinlichkeit 1 korrekt übertragen. Somit gibt es hier keine Störungen, und die Kapazität wird für $\beta = 0$ maximal.

(b) Unser Ziel ist es die Entropie $H(Y)$ zu maximieren. Machen wir uns zunächst klar, dass $H(Y) = H(p_0, p_1)$ immer bei Gleichverteilung einen maximalen Wert annimmt. Da $p_0 + p_1 = 1$ gelten muss, müssen die Werte p_1 und p_2 folgenden Bedingungen genügen:

a) $p_0 = \alpha$

b) $p_1 = (1 - \alpha)$

wobei $\alpha \in [0, 1]$. Die Entropie errechnet sich dann durch $H(Y) = -(\alpha \log \alpha + (1 - \alpha)\log(1 - \alpha))$. Trägt man $H(Y)$ in abhängigkeit von α über dem Intervall $[0, 1]$ auf, so ergibt sich das Schaubild aus Abbildung 1.

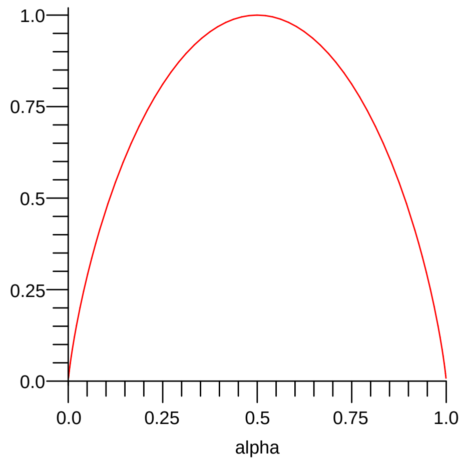


Abbildung 1: $H(Y)$ in Abhängigkeit von α bei einem binären Alphabet mit den Wahrscheinlichkeiten $p_0 = \alpha$ und $p_1 = 1 - \alpha$.

Es ist leicht ersichtlich, dass das Maximum hier bei $\alpha = \frac{1}{2}$ liegt. Im Allgemeinen gilt, dass die größte Entropie stets bei einer Gleichverteilung angenommen wird[1].

Wir fordern nun

$$P(Y = 0) = P(Y = 1) = \frac{1}{2}$$

Da wir jedoch die Entropie des Senders, also $H(X)$ berechnen wollen, müssen wir irgendwie $P(X)$ ins Spiel bringen. Wir können die Wahrscheinlichkeit $P(Y = y_i)$ auch schreiben

[1] Bei Gleichverteilung ist die „Unsicherheit“, welches Zeichen als nächstes kommt am größten, somit ist auch der Informationsgehalt jedes Zeichens am größten.

als

$$P(Y = y_i) = \sum_{x_i} P(X = x_i)P(Y|X)$$

denn die Wahrscheinlichkeit des Auftretens eines Zeichens y_i aus Y ist gerade die Summe der Wahrscheinlichkeiten jedes Auftreten von x_j multipliziert mit der Wahrscheinlichkeit, dass es als y_i übertragen wird.

Dies liefert für unsere Werte die beiden Gleichungen (LGS)

$$\begin{aligned} P(Y = 0) = (1 - \beta)P(X = 0) &\overset{!}{=} \frac{1}{2} \\ P(Y = 1) = P(X = 1) + \beta P(X = 0) &\overset{!}{=} \frac{1}{2} \end{aligned}$$

Wir können nun dieses LGS auflösen um $P(X = 0)$ und $P(X = 1)$ zu bestimmen.

$$P(X = 0) = \frac{P(Y = 0)}{1 - \beta} = \frac{\frac{1}{2}}{1 - \beta} = \frac{1}{2 - 2\beta}$$

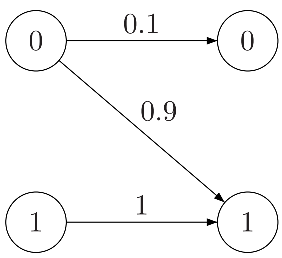
Da wir ein binäres Alphabet haben, und $P(X = 0) + P(X = 1) = 1$ ergeben muss kann man nun $P(X = 1)$ ohne viel Rechnerei hinschreiben

$$P(X = 1) = 1 - \frac{1}{2 - 2\beta}$$

Die Entropie des Senders lässt sich nun einfach berechnen:

$$\begin{aligned} H(X) &= H(P(X = 0), P(X = 1)) \\ &= H(\frac{1}{2 - 2\beta}, \underbrace{1 - \frac{1}{2 - 2\beta}}_{\frac{1-2\beta}{2-2\beta}}) \\ &= -\frac{1}{2 - 2\beta} \log \frac{1}{2 - 2\beta} - \frac{1 - 2\beta}{2 - 2\beta} \log \frac{1 - 2\beta}{2 - 2\beta} \end{aligned}$$

(c) Es ergibt sich nun folgende konkrete Situation für den Kanal:



Zunächst berechnen wir die Entropien $H(X)$, $H(Y)$ und $H(X,Y)$ (Totalinformation) aus dem Diagramm. Die restlichen Werte lassen sich dann sehr einfach ableiten.

Da wir eine Gleichverteilung seitens X gegeben haben, ist $P(X=0) = P(X=1) = \frac{1}{2}$. Abbildung 1 entnehmen wir eine Entropie $H(X) = 1\text{bit}$[2] Die Wahrscheinlichkeiten $P(Y)$ sind ebenfalls schnell berechnet. Es gilt nämlich

$$
\begin{aligned}
P(Y=0) &= P(X=0)\cdot P(Y=0|X=0) + P(X=1)\cdot \underbrace{P(Y=0|X=1)}_{=0} \\
&= \frac{1}{2}\cdot 0.1 \\
&= \frac{1}{20}
\end{aligned}
$$

Und damit folgt sofort $P(Y=1) = 1 - \frac{1}{20} = \frac{19}{20}$. Für die Entropie $H(Y)$ gilt daher

$$
\begin{aligned}
H(Y) &= H(\frac{1}{20}, \frac{19}{20}) \\
&= -\frac{1}{20}\log\frac{1}{20} - \frac{19}{20}\log\frac{19}{20} \\
&\approx 0.286\text{bit}
\end{aligned}
$$

Die Totalinformation, also $H(X,Y)$ berechnet sich durch

$$
\begin{aligned}
H(X,Y) = H(&P(X=0)P(Y=0|X=0), P(X=0)P(Y=1|X=0), \\
&P(X=1)P(Y=0|X=1), P(X=1)P(Y=1|X=1)) \\
&= H(\frac{1}{20}, \frac{9}{20}, \frac{1}{2}) \\
&\approx 1.234\text{bit}
\end{aligned}
$$

Mit diesen drei Werten lassen sich nun die gewünschten Entropien ableiten. Es gilt:

- Fehlinformation (Irrelevanz):

$$H(Y|X) = H(X,Y) - H(X) = 1.234 - 1 = 0.234\text{bit}$$

- Äquivokation (Verlustinformation):

$$H(X|Y) = H(X,Y) - H(Y) = 1.234 - 0.286 = 0.948\text{bit}$$

- Transinformation:

$$H(X;Y) = H(X) - H(X|Y) = 0.052\text{bit}$$

Man sieht dass der Kanal recht schlecht ist, und nur extrem wenig Information überträgt.

[2] Wer das nicht „glaubt“, mag das gerne zu Fuß nachrechnen.