

Aufgaben zum Tut am 11.06.2007

Thomas Pajor

13. Juni 2007

Aufgabe 1.

Ein Beobachter beobachtet eine Quelle über dem Alphabet $X = \{A, B, C, D, R\}$, die folgende Zeichen aussendet:

$$S := ABRACADABRA$$

- (a) Berechnen sie die Wahrscheinlichkeitsverteilung von S
- (b) Welche Information hat das Zeichen A , welche das Zeichen C ?
- (c) Wieviel Bits sind mindestens nötig um S zu kodieren?

Lösung.

- (a) Um die Wahrscheinlichkeitsverteilung zu ermitteln, können wir einfach die relative Häufigkeit jedes Zeichens $x_i \in X$ ermitteln, also abzählen. Mit $|S| = 11$ folgt dann

$$P(X = A) = \frac{5}{11} \quad P(X = B) = \frac{2}{11} \quad P(X = C) = \frac{1}{11} \quad P(X = D) = \frac{1}{11} \quad P(X = R) = \frac{2}{11}$$

Wir vergewissern uns, dass die Summe der Wahrscheinlichkeiten gerade 1 ergibt.

- (b) Die Information $I(x_i)$ eines Zeichens $x_i \in X$ ist definiert durch

$$I(x_i) := -\log P(X = x_i)$$

Damit gilt

$$I(A) = -\log P(X = A) = -\log \frac{5}{11} \approx 1.137\text{bit}$$

und

$$I(C) = -\log P(X = C) = -\log \frac{1}{11} \approx 3.459\text{bit}$$

Das Zeichen C hat also viel höhere Information als das Zeichen A . Das kann man sich intuitiv dadurch klarmachen, dass das C eine viel höhere „Überraschung“ mit sich bringt, als ein A , da es statistisch viel seltener auftaucht. Das Auftauchen von C wird also sehr viel mehr Information mit sich bringen.

- (c) Um eine untere Schranke anzugeben wieviel Bit zur Kodierung von S nötig sind, berechnen wir den Informationsgehalt der kompletten Zeichenkette S . Dafür berechnen wir zunächst die Entropie für X und kriegen damit eine Größe wieviel bit Information pro Zeichen in S stecken.

Die Entropie $H(X)$ ist definiert durch

$$H(X) = H(p_A, p_B, p_C, p_D, p_R) = - \sum_{i=1}^{|X|} P(X = x_i) \log P(X = x_i)$$

In unserem Fall also

$$H(X) = - \left(\frac{5}{11} \log \frac{5}{11} + \frac{2}{11} \log \frac{2}{11} + \frac{1}{11} \log \frac{1}{11} + \frac{1}{11} \log \frac{1}{11} + \frac{2}{11} \log \frac{2}{11} \right) \\ \approx 2.040 \frac{\text{bit}}{\text{Zeichen}}$$

Multiplizieren wir diesen Wert mit der Gesamtlänge der Zeichenkette, so erhalten wir die Information die in S steckt:

$$H_S = |S| \cdot H(X) = 11 \cdot 2.040 = 22.444 \text{bit}$$

Die minimale Anzahl an Bit die wir benötigen um S zu kodieren beträgt also 23.

Aufgabe 2.

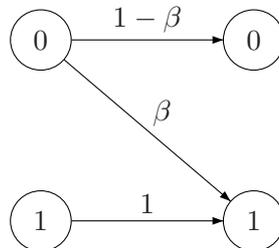
Gegeben sei folgender binärer, asymmetrischer Kanal über den Alphabeten $X = Y = \{0, 1\}$:

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

- (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.

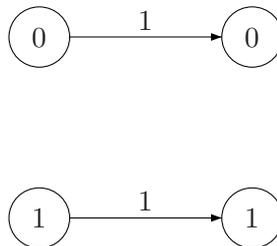
Wir machen uns zunächst den Übertragungskanal an einem Diagramm anschaulich klar:



(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_0 und p_1 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.

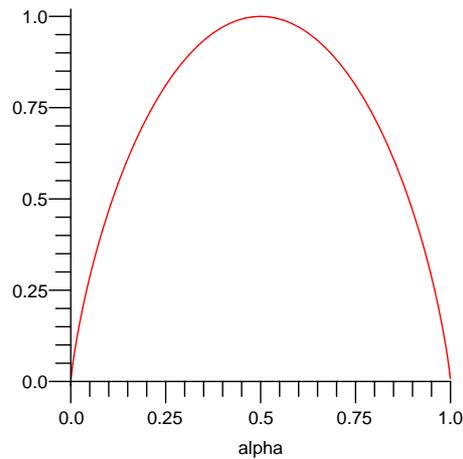


Abbildung 1: $H(Y)$ in Abhängigkeit von α bei einem binären Alphabet mit den Wahrscheinlichkeiten $p_0 = \alpha$ und $p_1 = 1 - \alpha$.

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 wird¹.

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 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)

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

$$P(Y = 1) = P(X = 1) + \beta P(X = 0) \stackrel{!}{=} \frac{1}{2}$$

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

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

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

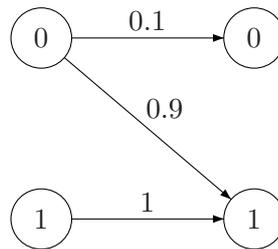
Da wir ein binäres Alphabet haben, und $P(X = 0) + P(X = 1) = 1$ ergeben muss kann man nun $P(X = 1)$ ohne viel Rechnerei hinschreiben

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

Die Entropie des Senders lässt sich nun einfach berechnen:

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

(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

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

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

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

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

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

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

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.