Vorlesung am 21.04.06

2. V

IC
SS – 06

## 1.3: Entropie einer diskreten, gedächtnisfreien Informationsquelle

Gegeben sei eine IQ mit den Symbolen $X = \{x_1, x_2, \ldots, x_M\}$. Die entsprechende Wahrscheinlichkeitsverteilung sei $P = \{p_1, p_2, \ldots, p_M\}$. Die Informationsgehalte der Symbole sind $I = \{I_1, I_2, \ldots, I_M\}$

$$I_i = \log_2 \frac{1}{p_i}$$

Die Entropie der Quelle ist der mittlere Informationsgehalt der Symbole:

$$H(x) = \sum_{i=1}^{M} p_i \, I_i \quad [\text{Bit/Symbol}]$$

Der Beitrag des i-ten Symbols zur Entropie ist

$$H_i = p_i \, I_i = p_i \log_2 \frac{1}{p_i}$$

$$H(x) = \sum_{i=1}^{M} H_i$$

Für die zwei extreme Fälle $P_i = 1$ und $P_i = 0$ verschwindet $H_i$:

$$P_i = 1 \implies H_i = P_i \log_2 \frac{1}{P_i} = \log_2 1 = 0$$

$$P_i = 0 \implies H_i = \lim_{P_i \to 0} P_i \log_2 \frac{1}{P_i} = \frac{-1}{\ln 2} \lim_{P_i \to 0} P_i \ln P_i$$

$$= \frac{+1}{\ln 2} \lim_{\lambda \to \infty} \frac{\ln \lambda}{\lambda} \quad ; \quad \lambda = \frac{1}{P_i}$$

$$= \frac{1}{\ln 2} \lim_{\lambda \to \infty} \frac{\frac{1}{\lambda}}{1} = 0$$

- Die Informationsrate R

Falls die Symbole mit einer Symbolrate $r$ [Symbol/Sek] gesendet werden, dann ist die

Informationsrate $R = r\,H(x)$ [Bit/Sek]

$r$: [Symbol/Sek] $\quad H(x)$: [Bit/Symbol]

- Eigenschaften von $H(x) = \sum_{i=1}^{M} H_i$

Da $H_i = P_i \log_2 \frac{1}{P_i} = \frac{-1}{\ln 2} P_i \ln P_i \quad ; 0 \le P_i \le 1$

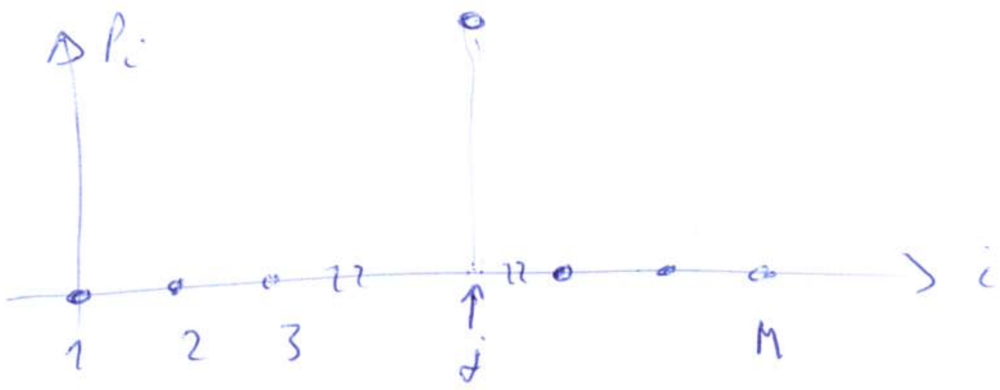
$\Longrightarrow H_i \geqslant 0 \quad ; \quad H(x) = \sum_{i=1}^{M} H_i \geqslant 0$

Wann ist $H(x) = 0$ ?

Da $H_i \geqslant 0$, dann $H(x) = 0$ wenn alle $H_i = 0$ ; $i = 1, 2, \ldots, M$. Dies ist nur möglich für eine Wahrscheinlichkeitsverteilung

$$P = \{0, 0, \ldots, 1, 0, 0 \ldots, 0\}$$

d.h., ein Einziges Symbol ist 100% wahrscheinlich, alle andere können nicht auftreten.

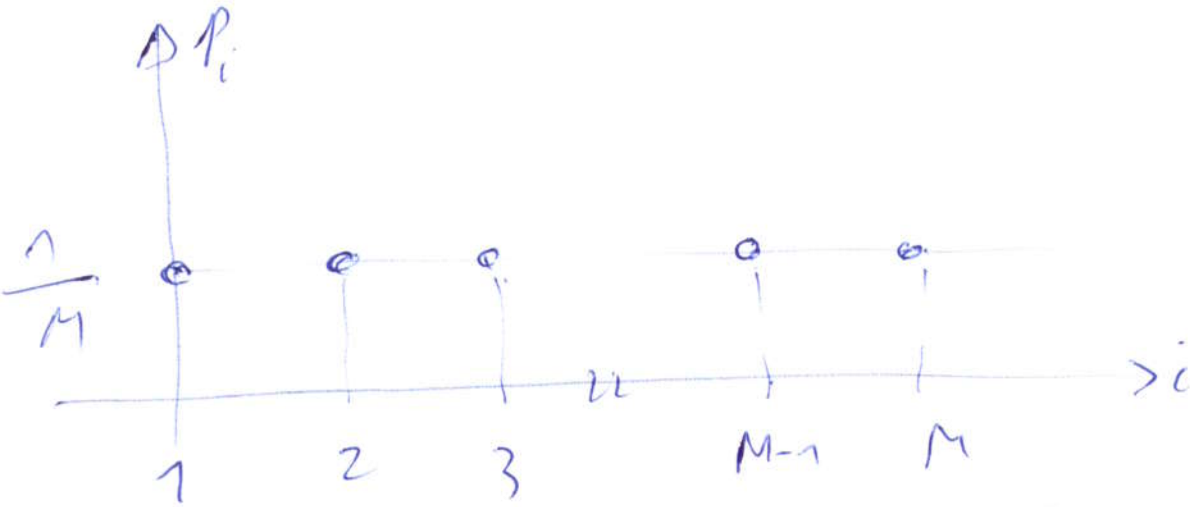



Wann erreicht $H(x)$ ihr Maximum

$$H(x) = \sum_{i=1}^{M} H_i = \sum_{i=1}^{M} P_i \log_2 \frac{1}{P_i} = \frac{-1}{\ln 2} \sum_{i=1}^{M} P_i \ln P_i$$

$H(x)$ erreicht ihr Maximum, wenn $\frac{\partial H}{\partial P_i} = 0$ für $i = 1, \ldots, M$. Wegen der symmetrischen Abhängigkeit von $H$ im Hinblick auf $P_i$, ist dies nur möglich, wenn $P_1 = P_2 = \ldots = P_M$

Da $\sum_{i=1}^{M} P_i = 1 \Longrightarrow P_i = \frac{1}{M}$




dies bedeutet, daß die gleichmäßige Wahrscheinlichkeitsverteilung zur maximalen Entropie führt

$$H_{max}(x) = \sum_{i=1}^{M} \frac{1}{M} \log_2 \frac{1}{\frac{1}{M}} = \sum_{i=1}^{M} \frac{\log_2 M}{M}$$

$$= \log_2 M$$

Beispiel: binäre Quelle

$$X = \{ x_1, x_2 \} \qquad P = \{ p, (1-p) \}$$

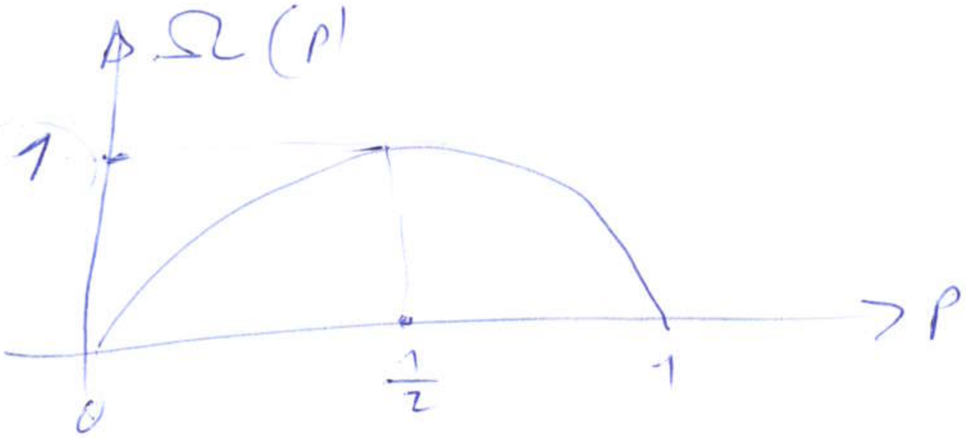
$$I = \left\{ \log_2 \frac{1}{p}, \log_2 \frac{1}{1-p} \right\}$$

$$H(X) = p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p} = \Omega(p)$$

$$\frac{\partial H}{\partial p} = \frac{-1}{\ln 2} \left[ p \frac{1}{p} + \ln p + (1-p) \cdot \frac{-1}{1-p} + \ln (1-p) \right]$$

$$= \frac{-1}{\ln 2} \ln \frac{p}{1-p} = 0 \quad , \text{wenn } \frac{p}{1-p} = 1$$

$$\Longrightarrow p = 1-p \Longrightarrow 2p = 1 \Longrightarrow p = \frac{1}{2} \; ; \; 1-p = \frac{1}{2}$$




$\log_2 2$

## – Die Redundanz der Quelle $R_Q(X)$

Diese ist die Differenz zwischen $H_{max} = \log_2 M$ und $H(X)$:

$$R_Q(X) = \log_2 M - H(X)$$

Solange $H(X) < \log_2 M$ ist, ist es immer möglich, eine Abbildung zu finden, die die Symbole von X auf Sequenzen von Symbolen anderer Quelle C, so daß die Entropie von C maximiert wird; d.h., die Symbole von C haben eine gleichmäßige Wahrscheinlichkeitsverteilung.

Wie bereits gezeigt, ist die Entropie ein Maß für die Homogenität der Wahrscheinlichkeitsverteilung. $H(X) = 0$ für die extrem heterogene Verteilung $P = \{0, 0, \ldots, 0, 1, 0, \ldots, 0\}$; und $H(X) = H_{max} = \log_2 M$ für die extrem homogene Verteilung $\left\{ \frac{1}{M}, \frac{1}{M}, \ldots, \frac{1}{M} \right\}$

**Beispiel**: Berechnen Sie die Entropie der Quelle $X = \{A, B, C, D\}$ mit $P = \left\{ \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8} \right\}$

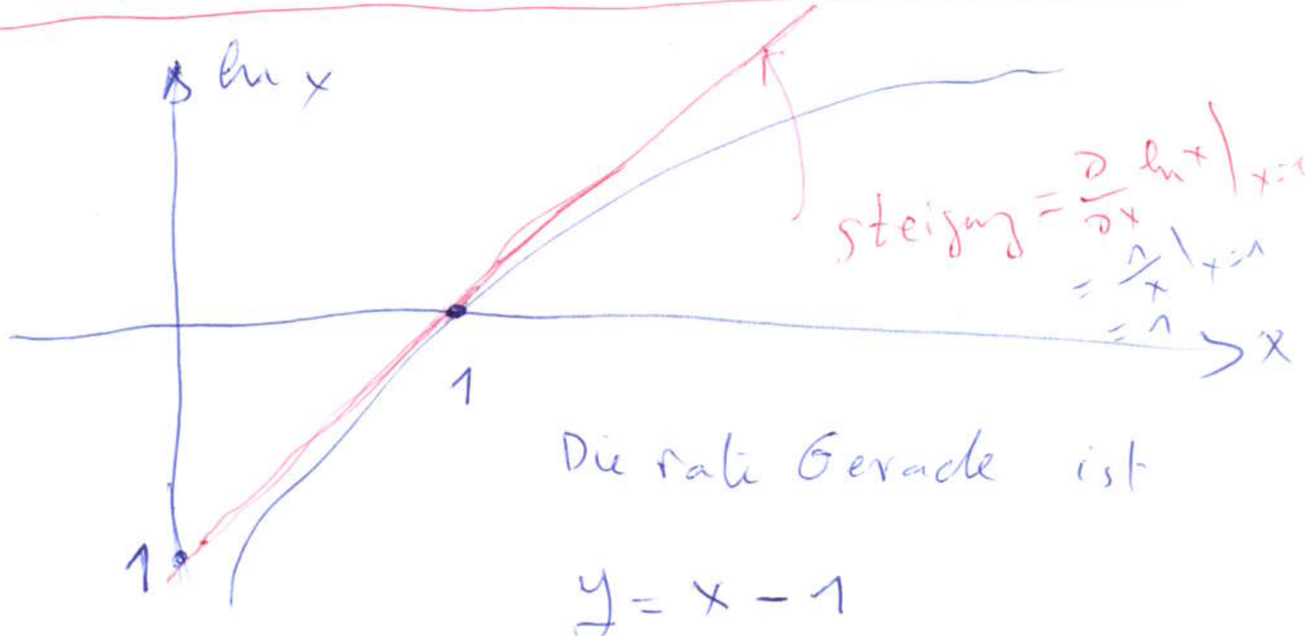
**Lösung**: $I = \{1, 2, 3, 3\}$

$$H(X) = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 3$$
$$= 1{,}75 \ \text{(Bit/Symbol)}$$

Die maximale Entropie ist $H_{max} = \log_2 4 = 2$

$R_Q(X) = 2 - 1{,}75 = 0{,}25$ (Bit/Symbol)

– Ein anderer Beweis für $H(X) \leq \log_2 M$




Die rote Gerade ist

$y = x - 1$

Aus dem Bild ist es e

Nun, $$H(X) = \sum_{i=1}^{M} P_i \log_2 \frac{1}{P_i}$$

$$= \frac{1}{\ln 2} \sum_{i=1}^{M} P_i \ln \frac{1}{P_i}$$

$$\log_2 M = \frac{\ln M}{\ln 2} = \frac{1}{\ln 2} \sum_{i=1}^{M} P_i \ln M$$

$$H(x) - \log_2 M = \frac{1}{\ln 2} \sum_{i=1}^{M} P_i \left[\ln \frac{1}{P_i} - \ln M\right]$$

$$= \frac{1}{\ln 2} \sum_{i=1}^{M} P_i \ln \frac{1}{M P_i} \quad x$$

$$\leq \frac{1}{\ln 2} \sum_{i=1}^{M} P_i \left(1 - \frac{1}{M P_i}\right)$$

$$= \frac{1}{\ln 2} \left[1 - 1\right] = 0$$

$$\Longrightarrow H(x) \leq \log_2 M$$

1.4: <u>Quellencodierung einer diskreten gedächtnis-freien Quelle</u>

Quellencodierung: Erstellung einer Abbildung (Zuordnung), welche die Symbole (einfach, paarweise zusammengefasst, oder mehrfach zusammengefasst) auf Sequenzen von neuen Symbolen abbildet. Das Ziel der Abbildung ist das Erreichen der neuen Quelle ihre maximale Entropie.

Die Wahrscheinlichkeit der neuen Symbole ist dann mehr oder weniger homogen verteilt.

Beispiele:

1-) binäre Codierung: Abbildung der Symbole auf Sequenzen von $C = \{c_1, c_2\}$

2-) Trinäre Codierung: Abbildung der Symbole auf Sequenzen von $C = \{c_1, c_2, c_3\}$

3.) M-äre Codierung: Abbildung der Symbole auf Sequenzen von

$C = \{ c_1, c_2, \ldots, c_M \}$

Ziel $\quad P\{c_i\} = \dfrac{1}{\text{Anzahl der C-Symbole}}$

$\Longrightarrow \quad H(C) = \text{Max.}$

Fluß von Symbolen $x_i$ (Rate: $r$)

$X = \{x_1, x_2, \ldots, x_M\} \longrightarrow$ Quellencodierer mit $M_c$ Symbolen $\longrightarrow C = \{c_1, c_2, \ldots, c_{M_c}\}$

Fluß von Symbolen $c_j$ (Rate: $r_c$)

| | |
|---|---|
| Entropie $= H(X)$ | Entropie $= H(C)$ |
| Informationsrate $= R$ | Informationsrate $= R_c$ |
| $R = r\, H(X)$ | $R_c = r_c \cdot H(C)$ |

Angenommen, $X_i$ wurde auf eine Sequenz von C-Symbolen abgebildet, deren Länge $N_i$ ist.

$$X_i \longrightarrow \underbrace{\boxed{c_i \mid c_j \mid c_k \mid \cdots \mid c_u}}_{N_i\text{-Stellen}}$$

Hier definiert man die Mittlere Codlänge

$$\bar{N} = \sum_{i=1}^{M} P_i N_i$$

Dies bedeutet, daß der Codierer durchschnittlich $\bar{N}$, C-Symbole produziert für jedes X-Symbol. $\Longrightarrow$ $\boxed{r_c = \bar{N} \cdot r}$

Für den Fall, daß die Informationsrate erhalten wird, gilt $\boxed{R = R_c}$

$$\Longrightarrow \quad r\, H(X) = r_c H(c) = \bar{N} \cdot r\, H(c)$$

$$\Longrightarrow \quad \boxed{\bar{N} = \frac{H(X)}{H(c)}}$$

Das Maximieren von $H(c)$ führt dann zum Minimieren von $\bar{N}$

<u>Beispiel</u> : Binäre Codierung

$$C = \{c_1, c_2\} \Longrightarrow H(c) = \Omega(p)$$

wobei $p = P\{c_1\}$

$$r_c = r_b \Longrightarrow \bar{N}_b = \frac{r_b}{r} = \frac{H(x)}{\Omega(p)}$$

Binäre mittlere Codlänge

Da $\Omega(p) \leq 1 \Longrightarrow \bar{N}_b \geq H(x)$

Die Entropie einer Quelle ist die minimale mittlere Codlänge einer binären Codierung.