

Informationsgehalt (Entscheidungsgehalt einer Nachricht)

$$I_r(a_i) = \log_r \frac{1}{P(a_i)}$$

$$r=2: I_2(a_i) = \text{ld} \frac{1}{P(a_i)}$$

Entropie (= mittlerer Informationsgehalt der Nachricht einer Quelle)

Gedächtnislose Quelle

$$H_r(A) = \sum_{i=1}^q P(a_i) \cdot I_r(a_i) = \sum_{i=1}^q P(a_i) \cdot \log_r \frac{1}{P(a_i)}$$

$$A = \{a_1, \dots, a_q\}, \sum_{i=1}^q P(a_i) = 1$$

$$r=2: H_2(A) = \sum_{i=1}^q P(a_i) \cdot \text{ld} \frac{1}{P(a_i)} \quad [\text{bit}]$$

$$\text{binäre Entropiefunktion: } H_r(p) = p \cdot \log_r \frac{1}{p} + (1-p) \cdot \log_r \frac{1}{(1-p)}$$

Die Entropie einer Nachrichtenquelle ist maximal, wenn alle Nachrichten mit der gleichen Wahrscheinlichkeit $P(a_i) = \frac{1}{q}$ $\forall i = 1, \dots, q$ gesendet werden.

n-te Quellenerweiterung

(n-te Erweiterung einer gedächtnislosen Quelle)

$$P(\vec{a}_n) = P(a_{n1}, a_{n2}, \dots, a_{nn}) = P(a_{n1}) \cdot P(a_{n2}) \cdot \dots \cdot P(a_{nn})$$

$$H(A^n) = n \cdot H(A)$$

Markoff-Quelle (Quelle mit Gedächtnis)

Bei einer q -nären Markoff-Quelle M mit der Einflusslänge $L+1$ (Rückgrifttiefe = Gedächtnistiefe L) hängt die Wahrscheinlichkeit der Nachricht zum Zeitpunkt $i, a(i) \in A$ von den letzten L gesendeten Nachrichten ab:

$$P(a(i) | \underbrace{(a(i-1), \dots, a(i-L))}_{\vec{z}(i)}) = P(a(i) | \vec{z}(i))$$

Informationsgehalt für einen Zustand $\vec{z}_j$:

$$I(a_u | \vec{z}_j) = \log \frac{1}{P(a_u | \vec{z}_j)}$$

Mittlerer Informationsgehalt für einen Zustand $\vec{z}_j$:

$$H(A | \vec{z}_j) = \sum_{u=1}^q P(a_u | \vec{z}_j) \cdot \log \frac{1}{P(a_u | \vec{z}_j)}$$

Mittlerer Informationsgehalt für alle Zustände:

$$H(A) = \sum_{j=1}^{q^L} P(\vec{z}_j) \cdot \sum_{u=1}^q P(a_u | \vec{z}_j) \cdot \log \frac{1}{P(a_u | \vec{z}_j)}$$

$$[M] = \begin{bmatrix} P(\vec{z}_1 | \vec{z}_1) & \dots & P(\vec{z}_q | \vec{z}_1) \\ \vdots & & \vdots \\ P(\vec{z}_1 | \vec{z}_q) & \dots & P(\vec{z}_q | \vec{z}_q) \end{bmatrix} \Rightarrow \vec{P}(i) = \vec{P}(i-1) \cdot [M] \\ = \vec{P}(0) \cdot [M]^i$$

$$\text{Mittlere Codewortlänge: } L(C) = \sum_{i=1}^q p_i \cdot \ell_i(C)$$

ℓ_i = Anzahl der Komponenten des CW
= Länge des CW

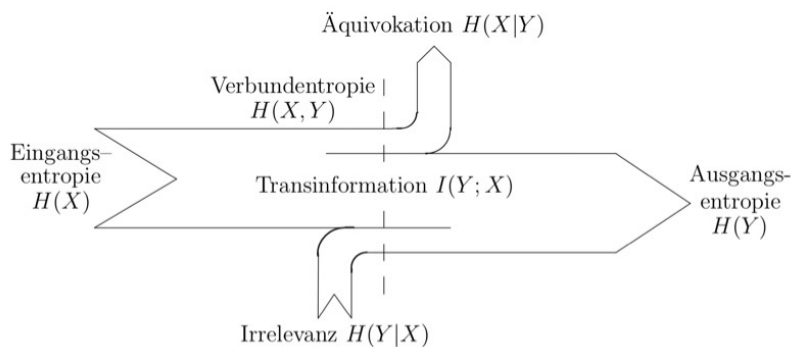
Effizienz eines Codes C :

$$\eta(C) = \frac{H_r(A)}{L(C)} = \frac{\text{Entropie}}{\text{mittlere CW-Länge}}$$

Optimale CW-Länge l_j für Einzelwahrscheinlichkeiten P_j :

$$\lceil \log \frac{1}{P_j} \rceil \leq l_j \leq \lceil \log \frac{1}{P_j} \rceil + 1, \quad l_j \in \mathbb{N}$$

Informationstransport über Kanäle



Irrelevanz $\hat{=}$ durch Kanal vorgeküscht Information

$$H(Y|X) = \sum_{x \in X} \sum_{y \in Y} \underbrace{P(y|x) \cdot P(x)}_{P(x,y)} \cdot \log \frac{1}{P(y|x)}$$

Äquivokation $\hat{=}$ bei der Übertragung verlorene Information

$$H(X|Y) = \sum_{x \in X} \sum_{y \in Y} \underbrace{P(x|y) \cdot P(y)}_{P(y,x) = P(x,y)} \cdot \log \frac{1}{P(x|y)}$$

Verbund-Entropie:

$$H(X|Y) + H(Y) = H(Y, X) = H(X, Y) = H(Y|X) + H(X)$$

Transinformation (mittlere wechselseitige Information)

$\hat{=}$ Ausgabeentropie, die ausschließlich von der Kanaleingabe herrührt

$$\begin{aligned}
 \text{Transinformationsgehalt } I(x, y) &= \log \frac{P(x|y)}{P(x)} \\
 &= \log \frac{P(x, y)}{P(x) \cdot P(y)} \\
 &= \log \frac{P(y|x)}{P(y)}
 \end{aligned}$$

Kanalkapazität

Die Kanalkapazität ist ein Maß für den maximal möglichen mittleren Informationsfluss $\hat{=}$ maximale Transinformation durch einen Kanal mit der Kanalmatrix $[P_u]$

$$\begin{aligned}
 C &= \max_{\vec{P}_x} I(X; Y) \quad \text{mit } \vec{P}_x: 0 \leq P(x) \leq 1 \quad \forall x \in X \\
 &\quad \text{und } \sum_{x \in X} P(x) = 1 \\
 &= \max_{\vec{P}_x} H(Y) - H(Y|X)
 \end{aligned}$$

Gedächtnislose Quelle: $H(X^n) = n \cdot H(X)$ n -te Erweiterung

Gedächtnisloser Kanal: $C^{(n)} = n \cdot C$

Decoder - Strategie bei gestörten Kanal

$$P_e = 1 - P_{\text{richtig}} \quad \min(P_e) \rightarrow \max(P_{\text{richtig}}) \rightarrow \max_{f(y)} (P(f(y) = x | y)) \quad \forall y \in Y$$

$$\{\hat{x}\} = \arg \max_{\tilde{x} \in X} P(\tilde{x} = \hat{x} | y) \xrightarrow{\text{Bayes}} \hat{x} = \arg \max_{\tilde{x} \in X} \frac{P(y | \tilde{x}) \cdot P(\tilde{x})}{P(y)} = \arg \max_{\tilde{x} \in X} P(y | \tilde{x}) \cdot P(\tilde{x})$$

irrelevant grösse: fällt weg, da arg max

$$\{\hat{x}\} = \arg \max_{\tilde{x} \in X} P(y | \tilde{x})$$

Maximum-Likelihood
 $P(y | \hat{x}) \geq P(y | \tilde{x}) \quad \forall \tilde{x} \in X$

Maximum A Posteriori
 $P(y | \hat{x}) \cdot P(\hat{x}) \geq P(y | \tilde{x}) \cdot P(\tilde{x}) \quad \forall \tilde{x} \in X$

ML: Bsp: $j=2, k=3, X = \{-, +\}, Y = \{-1, 0, +1\}$

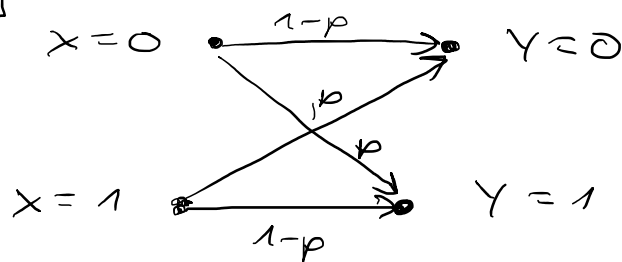
$$[P_k] = \begin{bmatrix} 0,5 & 0,4 & 0,1 \\ 0,2 & 0,3 & 0,5 \end{bmatrix} \begin{matrix} - \\ 0 \\ 1 \end{matrix}$$

ML $\Rightarrow$

$$\begin{aligned} f(-1) &= - \\ f(0) &= - \\ f(+1) &= + \end{aligned}$$

$$[P_k] = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

BSC = Binärer Symmetrischer Kanal



Rate $R = \frac{k}{n} = \frac{\text{Info-Länge}}{\text{Block-Länge}}$

Shannons Kanalcodierungssatz für BSC

gegeben sei ein BSC mit Fehlerwahrscheinlichkeit p und der Kanalkapazität $C = 1 - H(p)$ sowie ein Blockcode der Länge n . Sei $\epsilon > 0$ und $M = 2^{n(C-\epsilon)}$. Dann gibt es bei einer hinreichend großen Blocklänge n unter den 2^n möglichen Blöcken eine Teilmenge von M Blöcken derart, dass ihre Decodierfehlerwahrscheinlichkeit eine von ϵ abhängige Schranke unterschreitet.

$\Rightarrow$ Die M Blöcke können $\lfloor n(C-\epsilon) \rfloor$ Bit übertragen

$$\Rightarrow \text{Rate } R = \frac{k}{n} = \frac{\lfloor n(C-\epsilon) \rfloor}{n} \approx C - \epsilon \quad \Rightarrow \quad \boxed{R < C \text{ (bit)}}$$

Zur fehlerfreien Übertragung darf die Kanalkapazität nie kleiner sein als die Coderate.

Verallgemeinerung:

Sei X eine gedächtnislose q -näre Quelle mit der Entropie $H(X)$ und einer Quellsymbolrate von $\frac{1}{T_x}$. Der gedächtnislose diskrete q -näre Kanal mit der Kanalkapazität C habe eine Sendesymbolrate von $\frac{1}{T_c}$. Wenn gilt:

$$\frac{H(x)}{T_x} < \frac{C}{T_c}$$

dann existieren für hinreichend große Blocklängen Codes, die eine beliebig kleine Restfehlerwahrscheinlichkeit ermöglichen.

Umkehrung:

Wenn $C < \frac{T_c}{T_x} \cdot H(x)$, dann gibt es keine Codierung, die eine beliebig kleine Restfehlerwahrscheinlichkeit ermöglicht.

Kanalkapazität wertkontinuierlicher Kanäle

Mittelwert $\bar{x} = E[x]$

Varianz $\sigma_x^2 = E[(x - \bar{x})^2]$ = mittlere Wechselleistung

Entropie $H(x) \rightarrow \infty$, da beliebig viele Information darstellbar

differenzieller Informationsgehalt: $i(x) = \log \frac{1}{p_x(x)}$ [bit]

differenzielle Entropie: $h(x) [\text{bit}] = \int_{-\infty}^{\infty} p_x(\xi) \log \frac{1}{p_x(\xi)} d\xi$

...

Kanal kapazität pro Zeiteinheit

Kanal kapazität C in $[bit/s]$

Symbol dauer $\tau \Rightarrow$ Nyquist rate $\frac{1}{\tau}$, Kanalbandbreite $B = \frac{1}{2\tau} [Hz]$

Anzahl der Symbole je Zeiteinheit T : $n = \frac{T}{\tau} = 2 \cdot B \cdot T$

Kanal kapazität

$$C = \frac{n \cdot C_s}{T} = \frac{2 \cdot B \cdot T}{2 \cdot T} \log \left(1 + \frac{P}{N_0 \cdot B} \right) \cdot [bit]$$

$$= \frac{B}{[Hz]} \cdot \log \left(1 + \frac{P}{N_0 \cdot B} \right) \cdot \left[\frac{bit}{s} \right]$$

Spektrale Effizienz:

$$\frac{C}{B} = \log \left(1 + \frac{P}{N_0 \cdot B} \right) \left[\frac{bit}{s \cdot Hz} \right]$$

Satz von Shannon:

Ein auf die Bandbreite B begrenzter gedächtnisloser Kanal, der durch additives weißes Gaußrauschen der Leistung $N_0 \cdot B$ gestört wird, hat pro Zeiteinheit die Kanal kapazität:

$$C = B \cdot \log \left(1 + \frac{P}{N_0 \cdot B} \right) [bit]$$

wobei für die maximale Transformation das übertragene Signal mit der endlichen mittleren Leistung P eine gaußsche Wahrscheinlichkeit aufzuweisen hat.