

- ▶ A, B endliche Alphabete, A : Quellalphabet, B : Zielalphabet
- ▶ Ein (stationärer, gedächtnisfreier) Kanal gegeben durch Matrix

$$\mathbf{p} = [p_{a,b}]_{(a,b) \in A \times B}$$

- ▶ $p_{a,b}$ ist die (bedingte) Wahrscheinlichkeit, dass $b \in B$ empfangen wird, falls $a \in A$ gesendet wurde
- ▶ Die Kanalmatrix \mathbf{p} ist eine *stochastische Matrix*, d.h. für jedes $a \in A$ ist $\mathbf{p}_a = (p_{a,b})_{b \in B}$ eine WV auf B , d.h. $\sum_{b \in B} p_{a,b} = 1$, m.a.W.

$$\mathcal{R}_a = (B, \mathbf{p}_a) \text{ ist eine Quelle auf } B$$

Man schreibt auch $p(b|a)$ für $p_{a,b}$ im Sinne von: "bedingte Wahrscheinlichkeit"



- ▶ Stationarität und Gedächtnisfreiheit des Kanals äussern sich in der Aussage für die Übertragung von Wörtern:

$$A^n \ni (a_1, a_2, \dots, a_n) \mapsto (b_1, b_2, \dots, b_n) \in B^n$$

mit Wahrscheinlichkeit $p_{a_1, b_1} \cdot p_{a_2, b_2} \cdots p_{a_n, b_n}$

- ▶ Es genügt hier, die Übertragung von einzelnen Symbolen zu untersuchen.



1. Der störungsfreie Kanal: $A = B$ und

$$\mathbf{p} = E^{(m)}, \text{ d.h. } p_{a,b} = \delta_{a,b} = \begin{cases} 1 & \text{für } a = b \\ 0 & \text{für } a \neq b \end{cases}$$

2. Der total gestörte Kanal: $\#A = m, \#B = n$ und

$$\mathbf{p} = (p_{a,b})_{a \in A, b \in B} \text{ mit } p_{a,b} = \frac{1}{n} \quad (a \in A, b \in B)$$

3. Der binäre symmetrische Kanal BSC $_\rho$: $A = B = \mathbb{B} = \{0, 1\}$ mit

$$\mathbf{p} = \begin{matrix} & 0 & 1 \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \end{matrix}$$

4. Der binäre Auslöschungskanal BEC $_\rho$: $A = \mathbb{B} = \{0, 1\}$, $B = A \cup \{*\}$ mit

$$\mathbf{p} = \begin{matrix} & 0 & * & 1 \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{pmatrix} 1-p & p & 0 \\ 0 & p & 1-p \end{pmatrix} \end{matrix}$$



- ▶ Eine Quelle $\mathcal{R} = (A, \mathbf{r})$ ist gegeben durch eine WV $\mathbf{r} = (r_a)_{a \in A}$ auf A
 - ▶ r_a ist die Wahrscheinlichkeit, dass $a \in A$ gesendet wird.

- ▶ Der Vektor

$$\mathbf{c} = (c_b)_{b \in B} = \mathbf{r} \cdot \mathbf{p}, \text{ d.h. } c_b = r_a \cdot p_{a,b} \quad (b \in B)$$

ist eine WV auf dem Zielalphabet B , d.h.

$$\mathcal{C} = (B, \mathbf{c}) \text{ ist eine Quelle auf } B$$

- ▶ c_b ist die Wahrscheinlichkeit für das Empfangen von $b \in B$, falls die $a \in A$ mit Wahrscheinlichkeit p_a gesendet werden.



- ▶ Eine WV \mathbf{q} auf $A \times B$ wird definiert durch

$$q_{a,b} = r_a \cdot p_{a,b} \quad (a \in A, b \in B)$$

$\mathcal{Q} = (A \times B, \mathbf{q})$ bezeichnet die entsprechende Quelle.

- ▶ Für $(a, b) \in A \times B$ ist $q_{a,b}$ die Wahrscheinlichkeit für das Auftreten des Sende-Empfangspaares (a, b) , falls die $a \in A$ mit Wahrscheinlichkeit p_a gesendet werden.
- ▶ Wegen

$$\sum_{b \in B} q_{a,b} = r_a \quad (a \in A) \quad \text{und} \quad \sum_{a \in A} q_{a,b} = c_b \quad (b \in B)$$

sind \mathbf{r} und \mathbf{c} die Marginalverteilungen von \mathbf{q} .

- ▶ Für jedes $b \in B$ ist

$$\mathbf{c}_b = \left(\frac{q_{a,b}}{c_b} \right)_{a \in A} =$$

eine WV auf A und $\mathcal{C}_b = (A, \mathbf{c}_b)$ eine Quelle auf dem Alphabet A .

- ▶ $\frac{q_{a,b}}{c_b}$ ist die (bedingte) Wahrscheinlichkeit, dass $a \in A$ gesendet worden ist, falls $b \in B$ empfangen wurde, auch als $p(a|b)$ geschrieben.

- ▶ Jede der Quellen \mathcal{R}_a ($a \in A$) hat die Entropie

$$H(\mathcal{R}_a) = - \sum_{b \in B} p_{a,b} \cdot \log p_{a,b}.$$

Die erwartete Entropie (gewichtete Entropie bezüglich der Quellverteilung \mathbf{r} auf A) ist dann

$$H(\mathcal{C}|\mathcal{R}) = \sum_{a \in A} r_a \cdot H(\mathcal{R}_a).$$

Man spricht von der *bedingten Entropie* von \mathcal{C} bezüglich \mathcal{R} .

- ▶ Jede der Quellen \mathcal{C}_b ($b \in B$) hat die Entropie

$$H(\mathcal{C}_b) = - \sum_{a \in A} (q_{a,b}/c_b) \cdot \log(q_{a,b}/c_b).$$

Die erwartete Entropie (gewichtete Entropie bezüglich der Quellverteilung \mathbf{c} auf B) ist dann

$$H(\mathcal{R}|\mathcal{C}) = \sum_{b \in B} c_b \cdot H(\mathcal{C}_b).$$

Man spricht von der *bedingten Entropie* von \mathcal{R} bezüglich \mathcal{C} .

- Die Entropie der Quelle $\mathcal{Q} = (A \times B, \mathbf{q})$ ist

$$H(\mathcal{Q}) = H(\mathbf{q}) = - \sum_{(a,b) \in A \times B} q_{a,b} \cdot \log q_{a,b}$$

- Es gilt

$$H(\mathcal{Q}) = H(\mathcal{R}) + H(\mathcal{C}|\mathcal{R}) = H(\mathcal{C}) + H(\mathcal{R}|\mathcal{C})$$

Beweis: Einsetzen und ausrechnen!

- Es gilt

$$H(\mathcal{Q}) \leq H(\mathcal{R}) + H(\mathcal{C})$$

Beweis: $\mathbf{r}^t \cdot \mathbf{c} = (r_a \cdot c_b)_{(a,b) \in A \times B}$ ist eine WV auf $A \times B$.
Aus dem Fundamentallema von GIBBS folgt:

$$\begin{aligned} H(\mathcal{Q}) &= - \sum_{(a,b) \in A \times B} q_{a,b} \cdot \log q_{a,b} \leq - \sum_{(a,b) \in A \times B} q_{a,b} \cdot \log(r_a \cdot c_b) \\ &= - \sum_{(a,b) \in A \times B} q_{a,b} \cdot \log r_a - \sum_{(a,b) \in A \times B} q_{a,b} \cdot \log c_b \\ &= - \sum_{a \in A} r_a \underbrace{\left(\sum_{b \in B} p_{a,b} \right)}_{r_a} \cdot \log r_a - \sum_{b \in B} \underbrace{\left(\sum_{a \in A} q_{a,b} \right)}_{c_b} \cdot \log c_b \\ &= H(\mathcal{R}) + H(\mathcal{C}), \end{aligned}$$

wobei Gleichheit genau dann gilt, wenn $q_{a,b} = r_a \cdot c_b$ ist für alle $(a, b) \in A \times B$, d.h. wenn $p_{a,b} = c_b$ unabhängig von $a \in A$ ist.

- Als *wechselseitige Information* eines Paares $(a, b) \in A \times B$ bezeichnet man

$$I(a, b) = \log \frac{q_{a,b}}{r_a \cdot c_b} = \log q_{a,b} - \log r_a - \log c_b$$

- $I(a, b)$ kann positiv oder negativ sein. Insbesondere

$$I(a, b) = 0 \Leftrightarrow q_{a,b} = r_a \cdot c_b \Leftrightarrow p_{a,b} = c_b$$

1. Der störungsfreie Kanal: es ist

$$q_{a,b} = \begin{cases} r_a & \text{für } a = b \\ 0 & \text{für } a \neq b \end{cases}$$

und somit

$$I(a, b) = \begin{cases} -\log c_b = -\log r_a & \text{für } a = b \\ \text{nicht def.} & \text{für } a \neq b \end{cases}$$

2. Der total gestörte Kanal: wegen $p_{a,b} = \frac{1}{n}$ ist

$$q_{a,b} = r_a \cdot p_{a,b} = \frac{r_a}{n} \quad (a \in A, b \in B)$$

$$c_b = \sum_{a \in A} q_{a,b} = \frac{1}{n} \quad (b \in B)$$

$$I(a, b) = \log \frac{q_{a,b}}{r_a \cdot c_b} = \log 1 = 0 \quad (a \in A, b \in B)$$

3. Der BSC_p mit $p = \frac{1}{4}$, $\mathbf{r} = (\frac{1}{2}, \frac{1}{2})$:

$$\mathbf{c} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\mathbf{q} = \begin{pmatrix} \frac{3}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{3}{8} \end{pmatrix}$$

$$\begin{pmatrix} q_{a,b} \\ r_a \cdot c_b \end{pmatrix}_{a,b \in \mathbb{B}} = \begin{pmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} \end{pmatrix}$$

$$(I(a,b))_{a,b \in \mathbb{B}} = \begin{pmatrix} 0.58\dots & -1 \\ -1 & 0.58\dots \end{pmatrix}$$

Navigation icons

4. Der BEC_p mit $p = \frac{1}{4}$ und $\mathbf{r} = (\frac{1}{2}, \frac{1}{2})$

$$\mathbf{c} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ 0 & \frac{1}{4} & \frac{3}{4} \end{pmatrix} = \begin{pmatrix} \frac{3}{8} & \frac{1}{4} & \frac{3}{8} \end{pmatrix}$$

$$\mathbf{q} = \begin{pmatrix} \frac{3}{8} & \frac{1}{8} & 0 \\ 0 & \frac{1}{8} & \frac{3}{8} \end{pmatrix}$$

$$\begin{pmatrix} q_{a,b} \\ r_a \cdot c_b \end{pmatrix}_{a,b \in \mathbb{B}} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$$

$$(I(a,b))_{a,b \in \mathbb{B}} = \begin{pmatrix} 1 & 0 & * \\ * & 0 & 1 \end{pmatrix}$$

Navigation icons

- Als *wechselseitige Information* zwischen Senderverteilung \mathcal{R} und Empfängervertelung \mathcal{C} für den durch die Kanalmatrix \mathbf{p} gegebenem Kanal bezeichnet man den Erwartungswert der wechselseitigen Information:

$$I_{\mathbf{p}}(\mathcal{R}, \mathcal{C}) = \sum_{a \in A, b \in B} q_{a,b} \cdot I(a,b)$$

$$= H(\mathcal{R}) + H(\mathcal{C}) - H(\mathcal{Q}) \geq 0$$

Offensichtlich gilt auch

$$I_{\mathbf{p}}(\mathcal{R}, \mathcal{C}) = H(\mathcal{R}) - H(\mathcal{R}|\mathcal{C}) = H(\mathcal{C}) - H(\mathcal{C}|\mathcal{R})$$

Die Gleichheit der beiden Differenzen folgt aus den obigen Darstellungen von $H(\mathcal{Q})$.

Navigation icons

1. Der störungsfreie Kanal

$$I(\mathcal{R}, \mathcal{C}) = - \sum_{a \in A} r_a \log r_a = H(\mathcal{R})$$

2. Der total gestörte Kanal

$$I(\mathcal{R}, \mathcal{C}) = \sum_{a \in A, b \in B} \frac{r_a}{n} \cdot 0 = 0$$

3. Der $BSC_{1/4}$

$$I(\mathcal{R}, \mathcal{C}) = \frac{3}{8} \cdot \log \frac{3}{2} + \frac{1}{8} \cdot \log \frac{1}{2} + \frac{1}{8} \cdot \log \frac{1}{2} + \frac{3}{8} \cdot \log \frac{3}{2} \simeq 0.185\dots$$

4. Der $BEC_{1/4}$

$$I(\mathcal{R}, \mathcal{C}) = \frac{3}{8} \cdot 1 + \frac{1}{8} \cdot 0 + \frac{1}{8} \cdot 0 + \frac{3}{8} \cdot 1 = \frac{6}{8} = 0.75$$

Navigation icons

- ▶ Als *Kapazität* eines Kanals bezeichnet man

$$\max_{\mathbf{r}} I_{\mathbf{p}}(\mathcal{R}, \mathcal{C})$$

wobei über alle WVen \mathbf{r} auf A maximiert wird (optimale Anpassung der Quelle \mathcal{R} an die Übertragungseigenschaften des Kanals)

- ▶ Diese Grösse hängt also nur noch von der Kanalmatrix \mathbf{p} ab, ist also ein Charakteristikum des Kanals.

- ▶ Die Kapazität des binären symmetrischen Kanals mit Fehlerwahrscheinlichkeit p ist $1 - H(p)$.
- ▶ Beweis: Für den BSC_p ist

$$\mathbf{p} = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

und damit

$$\begin{aligned} \mathbf{c} = \mathbf{r} \cdot \mathbf{p} &= [r \quad 1-r] \cdot \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} \\ &= [r+p-2rp \quad 1-r-p+2rp] \end{aligned}$$

Damit ist

$$H(\mathcal{C}) = -(r+p-2rp) \cdot \log(r+p-2rp) - (1-r-p+2rp) \cdot \log(1-r-p+2rp)$$

und

$$H(\mathcal{C}|\mathcal{R}) = r \cdot H(p) + (1-r) \cdot H(p) = H(p).$$

Insbesondere ist hier $H(\mathcal{C}|\mathcal{R})$ nur vom Kanalparameter p und nicht von der Quellverteilung $\mathbf{r} = (r, 1-r)$ abhängig! Dann ist aber

$$\begin{aligned} \max_{\mathbf{r}} I_{\mathbf{p}}(\mathcal{R}, \mathcal{C}) &= \max_{\mathbf{r}} (H(\mathcal{C}) - H(\mathcal{C}|\mathcal{R})) \\ &= \max_{\mathbf{r}} H(\mathcal{C}) - H(p) \\ &= \max_{\mathbf{r}} H(r+p-2rp) - H(p) \end{aligned}$$

Die Entropiefunktion $H(x)$ wird maximal für $x = 1/2$. Das gesuchte Maximum wird also angenommen, wenn $r+p-2rp = 1/2 = 1-r-p+2rp$, also für $r = 1/2$. Das Maximum hat den Wert 1. Also ist $1 - H(p)$ der maximale Wert der wechselseitigen Information.

