# Vorlesung 12b

# Relative Entropie

## Teil1:

## Definition und Interpretation

## Zur Wiederholung:

Sei $S$ eine endliche oder abzählbare Menge
und $\rho$ eine Verteilung auf $S$.

Die **binäre Entropie** von $\rho$ ist

$$\mathbf{H}_2[\rho] := -\sum_{a \in S} \rho(a) \log_2 \rho(a)$$

(die bis auf maximal ein Bit kleinstmögliche erwartete Länge
eines binären Präfixcodes unter der Verteilung $\rho$)

Anstatt an binäre Präfixcodes kann man auch an ternäre, oder allgemeiner (für $b \in \mathbb{R}_+$) an $b$-äre Präfixcodes denken.

Die Entropie zur Basis $b$ der Verteilung $\rho$ ist

$$\mathbf{H}_b[\rho] := -\sum_{a \in S} \rho(a) \log_b \rho(a).$$

Die prominentesten Wahlen sind $b = 2$ und $b = e$. Im Folgenden denken wir uns ein $b > 0$ fest gewählt und lassen das Subskript $b$ weg.

**Definition der relativen Entropie:**

Seien $\rho$ und $\pi$ Wahrscheinlichkeitsverteilungen mit Gewichten $\rho(a)$ und $\pi(a)$, $a \in S$. Dann ist die *relative Entropie* von $\rho$ bzgl. $\pi$ definiert als

$$D(\rho \| \pi) := \sum_{a \in S} \rho(a) \log \frac{\rho(a)}{\pi(a)}$$

$$= -\sum_{a \in S} \rho(a) \log \pi(a) - \mathbf{H}[\rho] \ ,$$

wobei die Summanden mit $\rho(a) = 0$ gleich 0 gesetzt werden.

**Interpretation der relativen Entropie:**

Man denke sich einen zufälligen Buchstaben mit Verteilung $\rho$

mit einem Shannon-Code codiert, der nicht der Verteilung $\rho$,

sondern der Verteilung $\pi$ angepasst ist,

also mit Codewortlängen

$$-\log \pi(a) \le \ell(a) < -\log \pi(a) + 1.$$

Dann ändert sich die erwarteten Codelänge

im Vergleich zu dem an $\rho$ angepassten Shannon-Code

(bis auf höchstens 1) um

$$-\sum_a \rho(a) \log \pi(a) - \Big( -\sum_a \rho(a) \log \rho(a) \Big) = D(\rho \| \pi).$$

$D(\rho\|\pi)$ ist also (bis auf höchstens 1)

der erwartete Verlust an Bits, den man macht,

wenn man einen gemäß $\rho$ verteilten zufälligen Buchstaben

mit einem an $\pi$ (statt an $\rho$) angepassten

Shannon-Code codiert.

Im Teil 2 werden wir beweisen, dass immer $D(\rho\|\pi) \geq 0$ gilt.