

1 Informationstheorie

1.1 Typen von Datenquellen

1.1.1 Discrete Memoryless Source (DMS)

- Discrete heisst, dass die Quelle (zeitlich) einzelne Ereignisse liefert.
- Memoryless bedeutet, die Quelle erinnert sich beim Produzieren eines Ereignisses nicht an die Vorgeschichte. → Die Ereignisse sind (statistisch) unabhängig voneinander

1.1.2 Binary Memoryless Source (BMS)

- Bei dieser Quelle handelt es sich um eine DMS, die aber nur zwei verschiedene Ereignisse erzeugt.
- Ausgabe ist eine Folge von 0 und 1

1.2 Zweier-Logarithmus

$$x = \log_2(K) = \frac{\log_{10}(K)}{\log_{10}(2)}$$

1.3 Gleiche Wahrscheinlichkeit

- Je mehr Fälle es gibt, desto seltener tritt ein bestimmtes Ereignis ein.
- Je seltener ein Ereignis ist, desto höher ist sein Informationsgehalt.
- N sei wieder die Anzahl der möglichen Ereignisse. Wenn alle Ereigniswerte x_n die Gleiche Auftretungswahrscheinlichkeit $P(x_n)$ haben, gilt:

$$P(x_n) = \frac{1}{N} \rightarrow N = \frac{1}{P(x_n)}$$

1.4 Informationsgehalt von Ereignissen

- Je seltener ein Ereignis eintritt, desto grösser ist der Informationsgehalt (Überraschungseffekt)
- Die folgende Formel gilt allgemein:

$$I(x_n) = \log_2\left(\frac{1}{P(x_n)}\right)$$

1.5 Entropie

Den mittleren Informationsgehalt von Quellen nennt man Entropie:

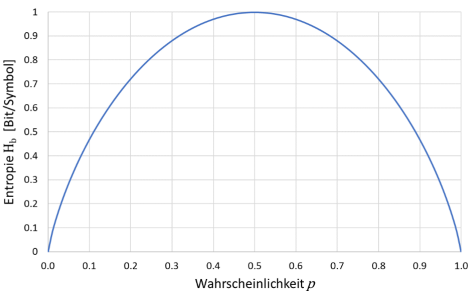
$$H(X) = \sum_{n=0}^{N-1} P(x_n) \cdot I(x_n) = \sum_{n=0}^{N-1} P(x_n) \cdot \log_2(P(x_n))$$

Die Masseinheit der Entropie ist *Bit/Symbol*.

1.5.1 Entropie Binary Memoryless Source

Eine BMS kennt nur zwei Symbole. Ist p die Auftretungswahrscheinlichkeit des eines Symbols, folgt dass $(1 - p)$ jene des anderen Symbols ist.

$$H_b = p \cdot \log_2\left(\frac{1}{p}\right) + (1 - p) \cdot \log_2\left(\frac{1}{1 - p}\right)$$



2 Quellencodierung

2.1 Redundanz

2.1.1 Codewortlänge

Symbol	Code	Codewortlänge
x_0	$c_0 = (10)$	$\ell_0 = 2Bit$
x_1	$c_1 = (110)$	$\ell_1 = 3Bit$
x_2	$c_2 = (1110)$	$\ell_2 = 4Bit$