Definition Information $I(p)$

Definition $I(p)$

Die Information $I(p)$ eines Symbols mit Quellws $p > 0$ beträgt

$$I(p) = \log \frac{1}{p}.$$

Die Einheit der Information bezeichnet man als Bit.

Beispiele für Information

- $Q = \{0, 1\}$ mit $p_1 = p_2 = \frac{1}{2}$. Dann ist $I(\frac{1}{2}) = 1$, d.h. für jedes gesendete Symbol erhält der Empfänger 1 Bit an Information.
- $Q = \{0, 1\}$ mit $p_1 = 1, p_2 = 0$. Dann ist $I(1) = 0$, d.h. der Empfänger enthält 0 Bit an Information pro gesendetem Zeichen.
- Beamer-Bild SXGA: Auflösung $1280 * 1024$, 256 Farben
  - $2^{1280*1024*8}$ mögliche Folien. Annahme: Jede gleich wahrscheinlich.
  - Information in Bit: $I(2^{-1280*1024*8}) = 1280 * 1024 * 8 = 10.485.760$
- Meine Erklärung dieser Folie:
  $\leq$ 1000 Worte, $\leq$ 10.000 Worte Vokabular
  - Information meiner Erklärung: $I(10.000^{-1000}) < 13.288$
  - Beweis für “Ein Bild sagt mehr als 1000 Worte!”

Entropie einer Quelle

Definition Entropie einer Quelle

Sei Q eine Quelle mit Quellws $P = \{p_1, \ldots, p_n\}$. Wir bezeichnen mit

$$H(Q) = \sum_{i=1}^{n} p_i I(p_i) = \sum_{i=1}^{n} p_i \log \frac{1}{p_i} = -\sum_{i=1}^{n} p_i \log p_i$$

die Entropie von Q.

- Für $p_i = 0$ definieren wir $p_i \log \frac{1}{p_i} = 0$.
- Entropie ist die durchschnittliche Information pro Quellsymbol.
- $P = \{\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\} : H(Q) = \sum_{i=1}^{n} \frac{1}{n} \log n = \log n$
- $P = \{\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}, 0\} : H(Q) = \sum_{i=1}^{n} \frac{1}{n} \log n = \log n$
- $P = \{1, 0, 0, \ldots, 0\} : H(Q) = 1 * \log 1 = 0$

Wollen zeigen: $0 \leq H(Q) \leq \log n$.

Wechsel zu anderer Ws-Verteilung

Lemma Wechsel Ws-Verteilung

Seien $P = \{p_1, \ldots, p_n\}$ eine Ws-Verteilung und $Q = \{q_1, \ldots, q_n\}$ mit $\sum_{i=1}^{n} q_i \leq 1$. Dann gilt

$$\sum_{i=1}^{n} p_i I(p_i) \leq \sum_{i=1}^{n} p_i I(q_i).$$

Gleichheit gilt genau dann, wenn $p_i = q_i$ für alle $i = 1, \ldots, n$.

Nützliche Ungleichung für das Rechnen mit logs:

$$x - 1 \geq \ln x = \log x \cdot \ln 2 \qquad \text{für alle } x > 0$$

Gleichheit gilt gdw $x = 1$.

Beweis des Lemmas

$$
\begin{aligned}
\sum_{i=1}^{n} p_i I(p_i) - \sum_{i=1}^{n} p_i I(q_i) &= \sum_{i=1}^{n} p_i \left(\log \frac{1}{p_i} - \log \frac{1}{q_i} \right) \\
&= \sum_{i=1}^{n} p_i \log \frac{q_i}{p_i} \\
&\leq \frac{1}{\ln 2} \sum_{i=1}^{n} p_i \left(\frac{q_i}{p_i} - 1 \right) \\
&= \frac{1}{\ln 2} \left(\sum_{i=1}^{n} q_i - \sum_{i=1}^{n} p_i \right) \\
&= \frac{1}{\ln 2} \left(\sum_{i=1}^{n} q_i - 1 \right) \leq 0.
\end{aligned}
$$

Gleichheit gilt gdw $\frac{q_i}{p_i} = 1$ für alle $i = 1, \ldots, n$.

Untere und obere Schranken für $H(P)$

Satz Schranken für $H(P)$

Sei Q eine Quelle mit Ws-Verteilung $P = \{p_1, \ldots, p_n\}$. Dann gilt

$$0 \leq H(Q) \leq \log n.$$

Weiterhin gilt $H(Q) = \log n$ gdw alle $p_i = \frac{1}{n}$ für $i = 1, \ldots, n$ und $H(Q) = 0$ gdw $p_i = 1$ für ein $i \in [n]$.

- Sei $P' = \{\frac{1}{n}, \ldots, \frac{1}{n}\}$ die Gleichverteilung.
- Nach Lemma zum Wechsel von Ws-Verteilungen gilt

$$H(Q) = \sum_{i=1}^{n} p_i \log \frac{1}{p_i} \leq \sum_{i=1}^{n} p_i \log \frac{1}{p'_i} = \log n \sum_{i=1}^{n} p_i = \log n.$$

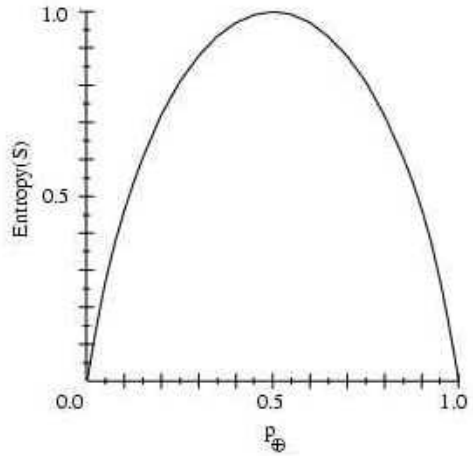
- Gleichheit gilt gdw $p_i = p'_i = \frac{1}{n}$ für alle i.

Untere Schranke für $H(P)$

Verwenden Ungleichung $\log x \geq 0$ für $x \geq 1$. Gleichheit gilt gdw $x = 1$.

$$H(Q) = \sum_{i=1}^{n} p_i \log \frac{1}{p_i} \geq 0,$$

mit Gleichheit gdw $\frac{1}{p_i} = 1$ für ein $i \in [n]$. □



- Binäre Quelle $Q = \{a_1, a_2\}$ mit $P = \{p, 1-p\}$

$$H(Q) = p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}.$$

- $H(Q)$ heißt binäre Entropiefunktion.

Kodieren einer binären Quelle

Szenario: Binäre Quelle Q mit $P = \{\frac{1}{4}, \frac{3}{4}\}$ mit

$$H(Q) = \frac{1}{4} \cdot \log 4 + \frac{3}{4} \cdot \log \frac{4}{3} \approx 0.811.$$

- Huffman-Kodierung von Q:
  $C(a_1) = 0, C(a_2) = 1$ mit $E(C) = 1$.
- **Problem:** Wie können wir a_2 mit kurzem Codewort kodieren?

- **Idee:** Kodieren Zweierblöcke von Quellsymbolen.

Quellerweiterungen von Q

- Betrachten $Q^2 = \{a_1a_1, a_1a_2, a_2a_1, a_2a_2\}$ mit Quellws

$$p_1 = \frac{1}{16}, p_2 = p_3 = \frac{3}{16}, p_4 = \frac{9}{16}.$$

- Huffmann-Kodierung von Q^2 liefert

$$C(a_1a_1) = 000, C(a_1a_2) = 001, C(a_2a_1) = 01, C(a_2a_2) = 1$$

mit $E(C) = 3 \cdot \frac{4}{16} + 2 \cdot \frac{3}{16} + \frac{9}{16} = \frac{27}{16}$.

- Jedes Codewort kodiert zwei Quellsymbole, d.h. die durchschnittliche Codewortlänge pro Quellsymbol ist

$$E(C)/2 = \frac{27}{32} = 0.84375.$$

- *Übung:* Für Q^3 erhält man 0.82292.

k-te Quellerweiterung Q^k

Definition k-te Quellerweiterung

Sei Q eine Quelle mit Alphabet $A = \{a_1, \ldots, a_n\}$ und Ws-Verteilung $P = \{p_1, \ldots, p_n\}$. Die k-te Quellerweiterung Q^k von Q ist definiert über dem Alphabet A^k, wobei $a = a_{i_1} \ldots a_{i_k} \in A^k$ die Quellws $p_{i_1} \cdot p_{i_2} \cdot \ldots \cdot p_{i_k}$ besitzt.

Entropie von Q^k

Sei Q eine Quelle mit k-ter Quellerweiterung Q^k. Dann gilt

$$H(Q^k) = k \cdot H(Q).$$

Beweis für $H(Q^k)$

$$H(Q^k) = \sum_{(i_1,\ldots,i_k)\in[n]^k} p_{i_1}\ldots p_{i_k} \log \frac{1}{p_{i_1}\ldots p_{i_k}}$$

$$= \sum_{(i_1,\ldots,i_k)\in[n]^k} p_{i_1}\ldots p_{i_k} \log \frac{1}{p_{i_1}} + \cdots + \sum_{(i_1,\ldots,i_k)\in[n]^k} p_{i_1}\ldots p_{i_k} \log \frac{1}{p_{i_k}}$$

- Betrachten ersten Summanden

$$\sum_{(i_1,\ldots,i_k)\in[n]^k} p_{i_1}\ldots p_{i_k} \log \frac{1}{p_{i_1}} = \sum_{i_1\in[n]} p_{i_1} \log \frac{1}{p_{i_1}} \cdot \sum_{i_2\in[n]} p_{i_2} \cdot \ldots \cdot \sum_{i_k\in[n]} p_{i_k}$$

$$= \sum_{i_1\in[n]} p_{i_1} \log \frac{1}{p_{i_1}} \cdot 1 \cdot \ldots \cdot 1 = H(Q).$$

- Analog liefern die anderen n-1 Summanden jeweils $H(Q)$.

Kodierungstheorem von Shannon

Kodierungstheorem von Shannon (1948)

Sei Q eine Quelle für $\{a_1, \ldots, a_n\}$ mit Ws-Verteilung $P = \{p_1, \ldots, p_n\}$. Sei C ein kompakter Code für Q. Dann gilt für die erwartete Codewortlänge

$$H(Q) \leq E(C) < H(Q) + 1.$$

Beweis: $H(Q) \leq E(C)$

- Bezeichnen Codewortlängen $\ell_i := |C(a_i)|$ und $q_i := 2^{-\ell_i}$.
- Satz von McMillan: $\sum_{i=1}^{n} q_i = \sum_{i=1}^{n} 2^{-\ell_i} \leq 1$.
- Lemma Wechsel Ws-Verteilung liefert

$$\begin{aligned} H(Q) &= \sum_{i=1}^{n} p_i \log \frac{1}{p_i} \leq \sum_{i=1}^{n} p_i \log \frac{1}{q_i} \\ &= \sum_{i=1}^{n} p_i \log 2^{\ell_i} = \sum_{i=1}^{n} p_i \ell_i = E(C). \end{aligned}$$

$E(C) \leq H(Q) + 1$

- Seien $\ell_1, \ldots, \ell_n$ Codewortlängen mit $\sum_{i=1}^{n} 2^{-\ell_i} \leq 1 = \sum_{i=1}^{n} p_i$.
- Satz von McMillan garantiert Existenz von Code C' für

$$2^{-\ell_i} \leq p_i \quad \Leftrightarrow \quad -\ell_i \leq \log p_i \quad \Leftrightarrow \quad \ell_i \geq \log \frac{1}{p_i}.$$

- Wählen $\ell_i \in \mathbb{N}$ für alle i minimal mit obiger Eigenschaft, d.h.

$$\log \frac{1}{p_i} \leq \ell_i < \log \frac{1}{p_i} + 1.$$

  Ein Code C' mit dieser Eigenschaft heißt *Shannon-Fano Code*.
- Für jeden kompakten Code C gilt

$$\begin{aligned} E(C) \leq E(C') &= \sum_{i=1}^{n} p_i \ell_i < \sum_{i=1}^{n} p_i \left(\log \frac{1}{p_i} + 1 \right) \\ &= \sum_{i=1}^{n} p_i \log \frac{1}{p_i} + \sum_{i=1}^{n} p_i = H(Q) + 1 \end{aligned}$$

Anwendung auf Quellerweiterungen

Korollar zu Shannons Kodierungstheorem

Sei Q eine Quelle mit k-ter Quellerweiterung Q^k. Sei C ein kompakter Code für Q^k. Dann gilt

$$H(Q) \leq \frac{E(C)}{k} < H(Q) + \frac{1}{k}.$$

- Anwendung von Shannon's Kodierungstheorem auf Q^k liefert

$$H(Q^k) \leq E(C) < H(Q^k) + 1.$$

- Anwenden von $H(Q^k) = kH(Q)$ und teilen durch k liefert die Behauptung.

Bedingte Entropie

- Sei X, Y Zufallsvariablen
- Definieren $\mathrm{Ws}(x) = \mathrm{Ws}(X = x)$ und $\mathrm{Ws}(x,y) = \mathrm{Ws}(X = x, Y = y)$.
- X, Y heißen unabhängig $\Leftrightarrow \mathrm{Ws}(x,y) = \mathrm{Ws}(x) \cdot \mathrm{Ws}(y)$

Definition Bedingte Entropie

Wir bezeichnen die Größe $H(Y|X)$

$$
\begin{aligned}
&:= \sum_x \mathrm{Ws}(x) H(Y|X = x) = \sum_x \mathrm{Ws}(x) \left(\sum_y \mathrm{Ws}(y|x) \log \frac{1}{\mathrm{Ws}(y|x)} \right) \\
&= \sum_x \sum_y \mathrm{Ws}(x,y) \log \frac{1}{\mathrm{Ws}(y|x)}
\end{aligned}
$$

als bedingte Entropie von Y gegeben X.

Eigenschaften bedingter Entropie

Rechenregel für die bedingte Entropie

1. Kettenregel:
   $H(X,Y) = \sum_x \sum_y \mathrm{Ws}(x,y) \log \frac{1}{\mathrm{Ws}(x,y)} = H(X) + H(Y|X)$ (Übung)
2. $H(Y|X) \leq H(Y)$. Gleichheit gilt gdw X, Y unabhängig sind. (ohne Beweis)
3. Folgerung aus 1. und 2.: $H(X,Y) \leq H(X) + H(Y)$.

Kryptographische Kodierung

Szenario:

- Drei Klartexte: a, b, c mit Ws $p_1 = 0.5, p_2 = 0.3, p_3 = 0.2$.
- Zwei Schlüssel k_1, k_2 gewählt mit Ws jeweils $\frac{1}{2}$
- Verschlüsselungsfunktionen:
  - $e_{k_1}: a \mapsto d, b \mapsto e, c \mapsto f$
  - $e_{k_2}: a \mapsto d, b \mapsto f, c \mapsto e$

- Seien P,C Zufallsvariablen für den Klar- und Chiffretext.
- Erhalten Chiffretext d, Plaintext muss a sein.
- Erhalten Chiffretext e, Plaintext muss b oder c sein.

$$\mathrm{Ws}(b|e) = \frac{\mathrm{Ws}(b,e)}{\mathrm{Ws}(e)} = \frac{\mathrm{Ws}(b,k_1)}{\mathrm{Ws}(b,k_1) + Ws(c,k_2)} = \frac{0.15}{0.15 + 0.1} = 0.6$$

  Lernen Information über zugrundeliegenden Klartext.
- $H(P) = \sum_i p_i \log \frac{1}{p_i} = 1.485$ und $H(P|C) = 0.485$.
- D.h. für gegebenen Chiffretext sinkt die Unsicherheit.

Perfekte Sicherheit

Definition Perfekte Sicherheit

Ein Kryptosystem ist perfekt sicher, falls $H(P|C) = H(P)$.

One-Time Pad

- Plaintextraum $\mathcal{P}$: $\{0,1\}^n$ mit Ws-Verteilung $p_1, \ldots, p_{2^n}$
- Schlüsselraum $\mathcal{K}$: $\{0,1\}^n$ mit Ws $\frac{1}{2^n}$ für alle Schlüssel
- Verschlüsselung: $c = e_k(x) = x \oplus k$ für $x \in \mathcal{P}, k \in \mathcal{K}$.

Sicherheit des One-Time Pads

Satz One-Time Pad

Das One-Time Pad ist perfekt sicher.

- Wahrscheinlichkeit von Chiffretext c:

$$\mathrm{Ws}(c) = \sum_{x,k:e_k(x)=c} \mathrm{Ws}(x)\mathrm{Ws}(k) = \frac{1}{2^n} \sum_{x,k:e_k(x)=c} \mathrm{Ws}(x) = \frac{1}{2^n}.$$

  Letzte Gleichung: Für jedes x, c existiert genau ein $k = x \oplus c$ mit $e_k(x) = c$.
- $H(K) = H(C) = n$
- Es gilt $H(P,K,C) = H(P,K) = H(P) + H(K)$
- Andererseits
  $H(P,K,C) = H(P,C) = H(P|C) + H(C) = H(P|C) + H(K)$.
- Dies liefert $H(P|C) = H(P)$.