Informatik III - Tutorium IX & X (SR -107)

Tut Nr. 12 – Üb12, Informationstheorie

David Münch

Universität Karlsruhe (TH)
Institut für Informatik
IAKS Beth

4. Februar 2009

Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

Inhaltsverzeichnis

Inhaltsverzeichnis

Inhaltsverzeichnis

Inhaltsverzeichnis

Inhaltsverzeichnis

Organisatorisches

Email: muenchdavid@gmail.com

https://www.stud.uni-karlsruhe.de/~uhbro/

Tutorium 09: Mittwochs 8:00 Uhr - Raum -107
Tutorium 10: Mittwochs 9:45 Uhr - Raum -107

Letzte Übungsblattabgabe heute.

Was wollen wir heute erreichen?

Was wollen wir heute erreichen?

- Übungsblatt 12 besprechen

Was wollen wir heute erreichen?

- Übungsblatt 12 besprechen
- Einführen von Grundbegriffen der Informationstheorie

Was wollen wir heute erreichen?

- Übungsblatt 12 besprechen
- Einführen von Grundbegriffen der Informationstheorie
- Überblick über Huffmancodes geben

Was wollen wir heute erreichen?

- Übungsblatt 12 besprechen
- Einführen von Grundbegriffen der Informationstheorie
- Überblick über Huffmancodes geben
- Kanäle verstehen

Was wollen wir heute erreichen?

- Übungsblatt 12 besprechen
- Einführen von Grundbegriffen der Informationstheorie
- Überblick über Huffmancodes geben
- Kanäle verstehen
- One-Time-Pad

Aufgabe 2

Alice möchte verschlüsselt mit Bob kommunizieren, die beiden entscheiden sich für das Rabin-Kryptosystem. Dazu erzeugt sich Alice ein Schlüsselpaar mit den Parametern $p = 7$ und $q = 23$, ihr geheimer Schlüssel ist $(p, q) = (7, 23)$.

1. Was ist der zum geheimen Schlüssel von Alice passsende öffentliche Schlüssel?

Aufgabe 2

Alice möchte verschlüsselt mit Bob kommunizieren, die beiden entscheiden sich für das Rabin-Kryptosystem. Dazu erzeugt sich Alice ein Schlüsselpaar mit den Parametern $p = 7$ und $q = 23$, ihr geheimer Schlüssel ist $(p, q) = (7, 23)$.

1. Was ist der zum geheimen Schlüssel von Alice passsende öffentliche Schlüssel?
2. Verschlüsseln Sie die Klartexte $m_1 = 3$ und $m_2 = 29$ mit Alice's öffentlichem Schlüssel.

Aufgabe 2

Alice möchte verschlüsselt mit Bob kommunizieren, die beiden entscheiden sich für das Rabin-Kryptosystem. Dazu erzeugt sich Alice ein Schlüsselpaar mit den Parametern $p = 7$ und $q = 23$, ihr geheimer Schlüssel ist $(p, q) = (7, 23)$.

1. Was ist der zum geheimen Schlüssel von Alice passsende öffentliche Schlüssel?
2. Verschlüsseln Sie die Klartexte $m_1 = 3$ und $m_2 = 29$ mit Alice's öffentlichem Schlüssel.
3. Entschlüsseln Sie das Chiffrat $c = 8$ mit Alice's geheimem Schlüssel. Welche Nachrichten sind als Klartext möglich?

Lösung:

1) Alice's öffentlicher Schlüssel ist $n := pq = 7 * 23 = 161$.

Lösung:

2) **Verschlüsselung von** m_1**:** Chiffrat $c_1 := m_1^2 \mod n = 3^2 \mod 161 = 9$.

Bemerkung: *An diesem Beispiel sieht man, dass das Verschlüsseln nicht sicher ist, wenn die Nachricht zu kurz ist. Hört ein Angreifer die Nachricht* c_1 *ab, so kann er daraus ganz einfach* m_1 *berechnen, in dem er die Quadratwurzel aus* 9 *in* $\mathbb{R}$ *zieht, Wurzelziehen in* $\mathbb{Z}_n$ *ist hier nicht nötig.*
Dieses Problem kann man beheben, indem man (zu kurze) Nachrichten grundsätzlich immer zu einer bestimmten Länge mit Bits auffüllt.

Verschlüsselung von m_2**:** Chiffrat $c_2 = m_2^2 \mod n = 29^2 \mod 161 = 841 \mod 161 = 36$.

Da $m_2^2 > n$ *ist, besteht das Problem von oben hier nicht.*

Lösung:

3) Berechnen der Quadratwurzeln von c modulo p und q:

$m_p = \sqrt{c} \mod p = c^{\frac{p+1}{4}} \mod p = 8^2 \mod 7 = 1.$

$m_q = \sqrt{c} \mod q = c^{\frac{q+1}{4}} \mod q = 8^6 \mod 23 = 13.$

Berechnen der y_p, y_q mit $y_p * p + y_q * q = 1$ mit dem erweiterten euklidschen Algorithmus:

Schritt	q	p	r_0	r_1	s_0	s_1	a_0 div a_1
Init.	23	7	1	0	0	1	23 div 7 = 3
1	7	2	0	1	1	−3	7 div 2 = 3
2	2	1	1	−3	−3	10	2 div 1 = 2
3	$1 = ggT$	0	$-3 =: y_q$	7	$10 =: y_p$	−23	-

Wir erhalten $y_p = 10$, $y_q = -3$. Test:

$y_p * p + y_q * q = 7 * 10 + (-3) * 23 = 1.$

Lösung:

Berechnen der vier möglichen Quadratwurzeln von $c = 8$ modulo n:

$r := (y_p * p * m_q + y_q * q * m_p)$

$\mod n = (10 * 7 * 13 + (-3) * 23 * 1) \mod 161 = 36$

$-r := n - r = 161 - 36 = 125$

$s := (y_p * p * m_q - y_q * q * m_p)$

$\mod n = (10 * 7 * 13 - (-3) * 23 * 1) \mod 161 = 13$

$-s := n - s = 161 - 13 = 148$

Als Klartexte sind die vier Nachrichten $36, 125, 13$ und 148 möglich.

Aufgabe 1

Sei $L \subseteq \{0,1\}^*$ eine Sprache aus $\mathcal{BPP}$, das heißt es gibt einen Polynomialzeitalgorithmus A, der bei Eingabe eines Wortes $x \in \{0,1\}^*$ und „Zufalls“ $r \in \{0,1\}^{p(|x|)}$ (In der Kryptographie Coin oder Cointoss genannt), wobei $p(n)$ ein nur von A abhängiges Polynom ist, folgendes Ausgabeverhalten zeigt ($\mathcal{U}(M)$ bezeichnet hierbei die Gleichverteilung auf einer Menge M).

- Ist $x \in L$, so ist $P(A(x,r) = 1 | r \in \mathcal{U}(\{0,1\}^{(p(|x|))})) > \frac{2}{3}$ also die Wahrschleinlichkeit dass x von A akzeptiert wird für ein zufällig gleichverteiltes $r \in \{0,1\}^{(p(|y|)}$ größer als $2/3$.
- Ist $x \notin L$, so ist $P(A(x,r) = 0 | r \in \mathcal{U}(\{0,1\}^{(p(|x|))})) > \frac{2}{3}$ also die Wahrschleinlichkeit dass x von A nicht akzeptiert wird für ein zufällig gleichverteiltes $r \in \{0,1\}^{(p(|y|)}$ größer als $2/3$.

Es gilt also $P(A(x,r) = \chi_L(x) | r \in \mathcal{U}(\{0,1\}^{(p(|x|))})) > \frac{2}{3}$, wobei $\chi_L : \{0,1\}^* \to \{0,1\}$ die charakteristische Funktion von L ist.

Fortsetzung Aufgabe 1

Zeigen sie nun, dass es zu jedem $L \in \mathcal{BPP}$ einen Polynomialzeitalgorithmus A' gibt, für welchen gilt

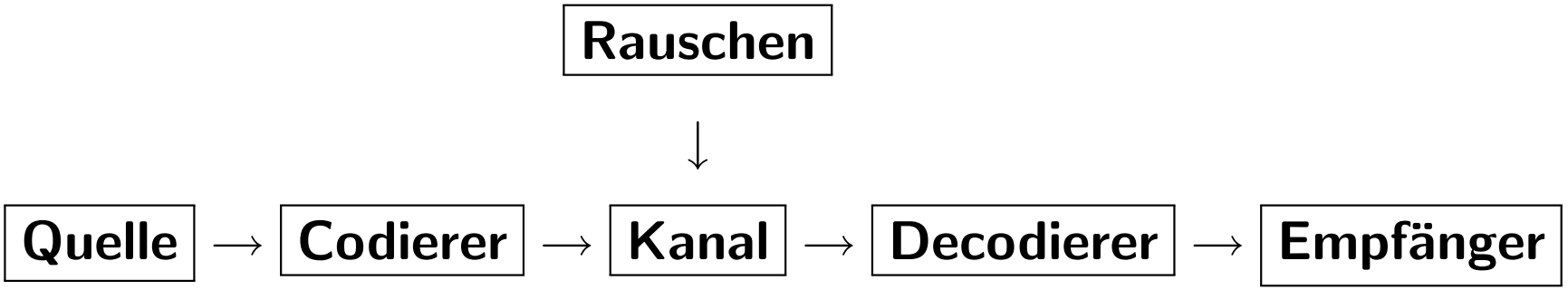
$$P(A'(x,r) = \chi_L(x) | r \in \mathcal{U}(\{0,1\}^{(p'(|x|))})) > 1 - \frac{1}{2^{|x|}}$$

wobei $p'(n)$ ein nur von A' abhängiges Polynom ist. Das heißt also, dass es zu jedem $\mathcal{BPP}$-Algorithmus A einen weiteren $\mathcal{BPP}$-Algorithmus A' gibt,welcher die selbe Sprache erkennt, dessen Erfolgswahrscheinlichkeit asymptotisch exponentiell nahe bei 1 liegt.

Hinweis: Erinnern sie sich an die Definition eines Bernoulli-Experiments und konstruieren sie einen Mehrheitsentscheider. Schätzen sie die Terme in den auftretenden Binomialverteilungen geeignet ab.

Grundbegriffe der Informationstheorie

Kommunikationssystem



Die Informationstheorie versucht nun die einzelnen Blöcke durch ein mathematisches Modell zu bescheiben.

Zentrale Fragen der Informationstheorie

- Wieviel Information enthält eine Nachricht?

Zentrale Fragen der Informationstheorie

- Wieviel Information enthält eine Nachricht?
- Wieviel von der Nachricht ist überflüssig?

Zentrale Fragen der Informationstheorie

- Wieviel Information enthält eine Nachricht?
- Wieviel von der Nachricht ist überflüssig?
- Kann die Nachricht noch gelesen werden wenn Störungen auftreten?

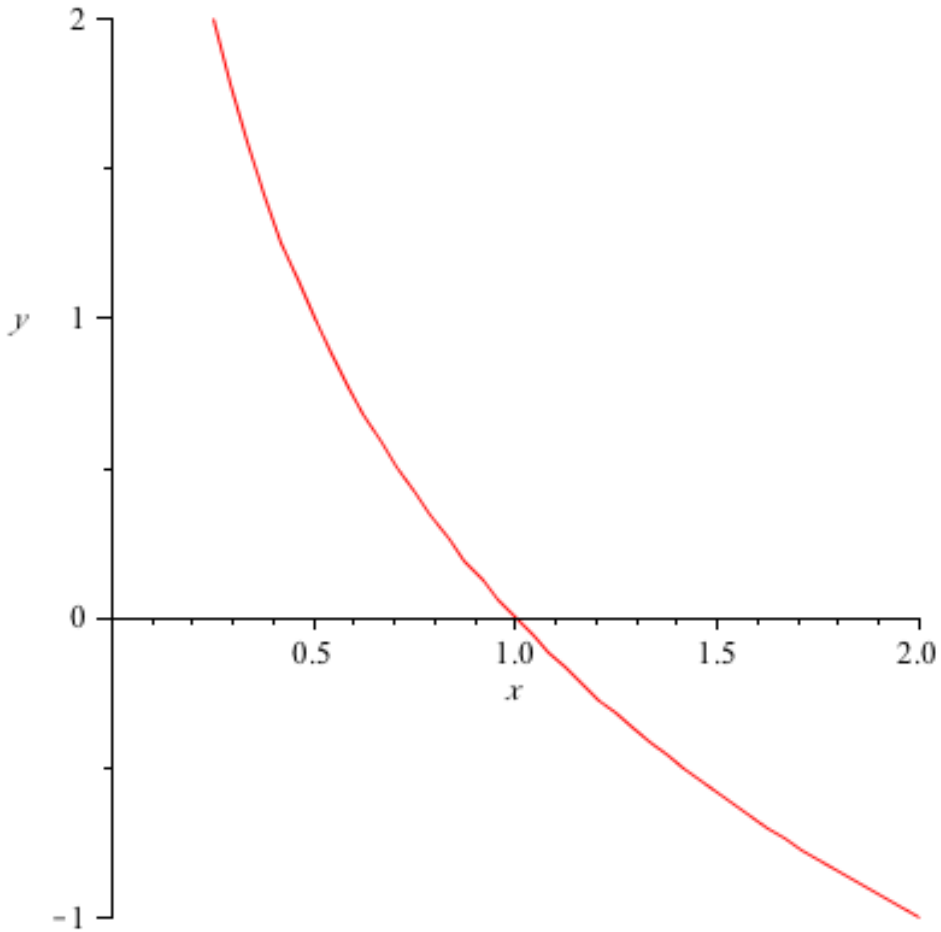
Beispiel:
Sendet man statt einer Eins n Einsen, dann kann man für $n \to \infty$ die Fehlerwahrscheinlichkeit beliebig klein machen auf Kosten der Übertragungsgeschwindigkeit.

Entropie

Definition:

Sei X eine Zufallsvariable mit den unabhängigen Ereignissen $x_1, ..., x_m$ und sei $p_i = p(x_i) = P(X = x_i)$ die Wahrscheinlichkeit eines Ereignisses x_i, das mit der Unsicherheit (Information) $h(x_i)$ behaftet ist.
Eine Ereignissfolge $x_i x_j$ tritt mit der Wahrscheinlichkeit $p_{ij} = p(x_i x_j) = p_i \cdot p_j$ auf und hat die Unsicherheit $h(p_{ij}) = h(p_i) + h(p_j)$.
U.a. daraus folgt, dass die Unsicherheitsfunktion (oder Information eines Zeichens) von der Gestalt $h(p) = -\log_b p$ sein muss.

Entropie



Entropie

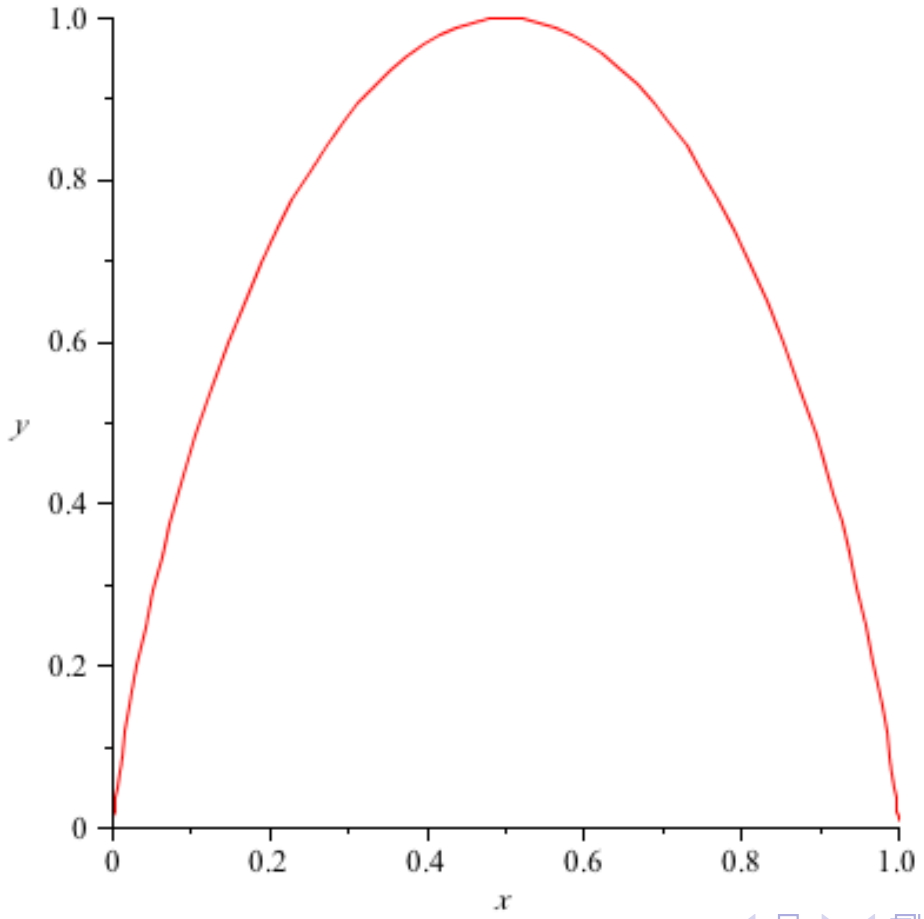
Definition: Entropie

Die Entropie (Informationsgehalt) einer Zufallsvariablen X ist $H(X) = H(p_1, ..., p_m) = -\sum_{i=1}^{m} p_i \log p_i$

Die Entropie ist die Information pro Zeichen, die wir erwarten.
Oder: Die Entropie ist die durchschnittliche Anzahl von Entscheidungen (bits), die benötigt werden, um ein Zeichen aus einer Zeichenmenge zu identifizieren oder zu isolieren.

Entropie

Beispiel: Entropie von Münzen: $H(p, 1-p)$



Aufgabe

Ein Empfänger beobachtet eine Quelle über dem Alphabet $X = \{A, B, C, D, R\}$, die folgende Zeichen aussendet:

$$S := ABRACADABRA$$

a) Berechne die Wahrscheinlichkeitsverteilung von S.

b) Welche Information hat das Zeichen A, welche das Zeichen C?

c) Wieviel Bits sind nötig um S zu codieren?

Lösung

a) $|S| = 11$
$\Rightarrow p_A = 5/11, p_B = 2/11, p_C = 1/11, p_D = 1/11, p_R = 2/11$

Lösung

a) $|S| = 11$
$\Rightarrow p_A = 5/11, p_B = 2/11, p_C = 1/11, p_D = 1/11, p_R = 2/11$

b) Mit $I(x_i) = h(p_i) = -\log p_i$ folgt:
$I(A) = -\log 5/11 \approx 1,137 bit$
$I(C) = -\log 1/11 \approx 3,459 bit$

Lösung

a) $|S| = 11$
$\Rightarrow p_A = 5/11, p_B = 2/11, p_C = 1/11, p_D = 1/11, p_R = 2/11$

b) Mit $I(x_i) = h(p_i) = -\log p_i$ folgt:
$I(A) = -\log 5/11 \approx 1,137 bit$
$I(C) = -\log 1/11 \approx 3,459 bit$

c) Berechne Informationsgehalt der kompletten Zeichenkette.
Entropie:
$H(X) = -\sum_{i=1}^{5} p_i \log p_i \approx 2,04 bit/Zeichen$
Multiplizieren von $H(X)$ mit der Länge der Zeichenkette:
$H(X) \cdot |S| = 22,444 bit$
$\Rightarrow$ minimale Anzahl benötigter Bits beträgt 23.

Gemeinsame Entropie

Seien X und Y zwei Zufallsvariablen mit den Werten $x_1, ..., x_m$ bzw. $y_1, ..., y_n$.
Sei $p_{ij} = p(x_i y_j) = P(X = x_i \text{ und } Y = y_j)$.
Wenn x_i und y_i (wie oben) unabhängig sind,
dann gilt: $p_{ij} = p(x_i) \cdot p(y_j)$
sonst $p_{ij} = p(x_i) \cdot p(y_j|x_i) = p(y_j) \cdot p(x_i|y_j)$.

Definition: Gemeinsame Entropie

Die Gemeinsame Entropie von X und Y ist
$H(X, Y) = H(p_{11}, .., p_{mn}) = -\sum_{i,j} p_{ij} \log p_{ij}$

Gemeinsame Entropie

Seien X und Y zwei Zufallsvariablen mit den Werten $x_1, ..., x_m$ bzw. $y_1, ..., y_n$.
Sei $p_{ij} = p(x_i y_j) = P(X = x_i \text{ und } Y = y_j)$.
Wenn x_i und y_i (wie oben) unabhängig sind,
dann gilt: $p_{ij} = p(x_i) \cdot p(y_j)$
sonst $p_{ij} = p(x_i) \cdot p(y_j|x_i) = p(y_j) \cdot p(x_i|y_j)$.

Definition: Gemeinsame Entropie

Die Gemeinsame Entropie von X und Y ist
$H(X, Y) = H(p_{11}, .., p_{mn}) = -\sum_{i,j} p_{ij} \log p_{ij}$

Satz

$H(X_1, ..., X_n) \leq H(X_1) + ... + H(X_n)$, mit Gleichheit $\Leftrightarrow$ X_i unabhängig.

Bedingte Entropie

Definition: Bedingte Entropie

$$H(X|Y) = -\sum_{i,j} p_{ij} \cdot \log p(x_i|y_j)$$

Bedingte Entropie

Definition: Bedingte Entropie

$$H(X|Y) = -\sum_{i,j} p_{ij} \cdot \log p(x_i|y_j)$$

Satz

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

Information

Definition: Information

Information ist die beseitigte Unsicherheit. D.h. die durch Kenntnis von Y über X erzielbare Information ist:
$I(X|Y) = H(X) - H(X|Y) = I(Y|X) \geq 0$

Aufgabe

1. Was ist der Informationsgehalt und die Entropie, wenn eine Quelle mit der Alphabet $\{0, 1\}$ nur lauter 0-en sendet?

Aufgabe

1. Was ist der Informationsgehalt und die Entropie, wenn eine Quelle mit der Alphabet $\{0, 1\}$ nur lauter 0-en sendet?
2. An einer Quelle mit n Zeichen tritt jedes Zeichen gleichverteilt auf. Was ist der Informationsgehalt eines einzelnen Zeichens und die Entropie?

Aufgabe

1. Was ist der Informationsgehalt und die Entropie, wenn eine Quelle mit der Alphabet $\{0, 1\}$ nur lauter 0-en sendet?
2. An einer Quelle mit n Zeichen tritt jedes Zeichen gleichverteilt auf. Was ist der Informationsgehalt eines einzelnen Zeichens und die Entropie?
3. Berechnen Sie die Entropie des Wurfes eines idealen Würfels mit 8 Seiten, dessen Wahrscheinlichkeit für jede Seite $p = \frac{1}{8}$ ist.

Aufgabe

1. Was ist der Informationsgehalt und die Entropie, wenn eine Quelle mit der Alphabet $\{0, 1\}$ nur lauter 0-en sendet?
2. An einer Quelle mit n Zeichen tritt jedes Zeichen gleichverteilt auf. Was ist der Informationsgehalt eines einzelnen Zeichens und die Entropie?
3. Berechnen Sie die Entropie des Wurfes eines idealen Würfels mit 8 Seiten, dessen Wahrscheinlichkeit für jede Seite $p = \frac{1}{8}$ ist.
4. Was ist der Unterschied zwischen den Folgen, die aus den verschiedenen gedächtnislosen Quellen mit der gleichen Wahrscheinlichkeiten von 0 und 1 gesendet sind, wenn man sie vom Standpunkt Entropie und Ordnung betrachtet.
   a) ...10101010101010101010...
   b) ...01101100110111000010...

Lösung:

1) Sei p die Wahrscheinlichkeitsfunktion. Nach der Aufgabenstellung sendet die Quelle nur Nullen, d.h. $p(x = 0) = 1$ und $p(x = 1) = 0$. Der Informationsgehalt I eines Zeichens, was die Quelle sendet, ist $I(x) = -\log_2 \frac{1}{p(x)}$.

Somit $I(x = 0) = \log_2 \frac{1}{1} = 0$.

Die Entropie, also der mittlere Informationsgehalt ist auch 0:

$\sum_{x \in \{0,1\}} p(x) \cdot I(x) = 1 \cdot 0 = 0.$

Lösung:

Für alle n Zeichen aus der Quelle gilt $p(x) = \frac{1}{n}$. Somit

$$I(x) = \log_2 \frac{1}{p(x)} = -\log_2 p(x) = -\log_2 \frac{1}{n} = \log_2 n.$$

Wenn die Folge aus der n unabhängigen stochastischen Ereignissen besteht, ist der gesamte Informationsgehalt:

$I_{ges}(x) = \sum_{i=1}^{n} \log_2 p(x_i) = n \cdot \log_2 n.$

Lösung:

Sei Z das Alphabet: $|Z| = 8$ und die Wahrscheinlichkeit für jede 8 Seiten des Würfels $p = \frac{1}{8}$

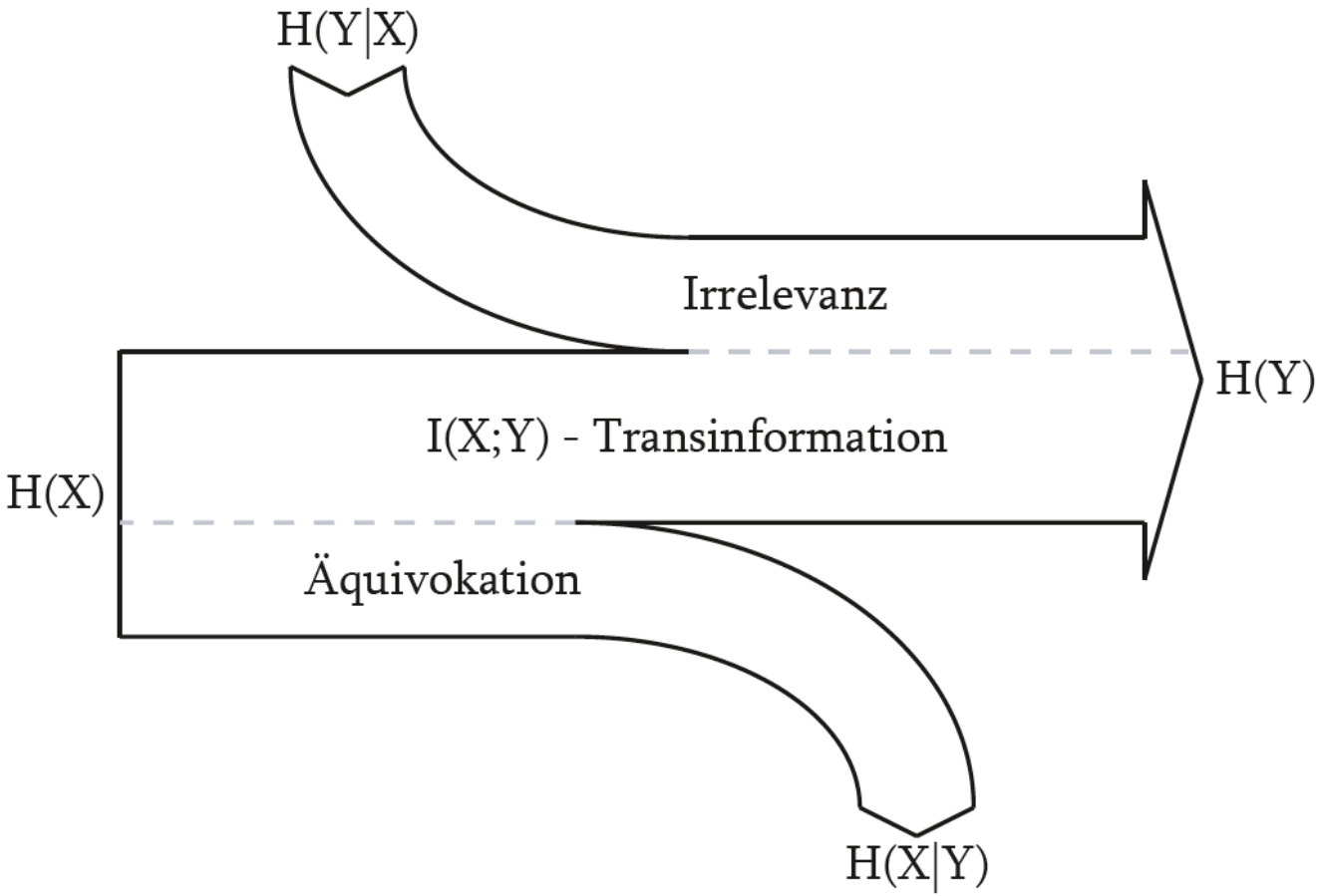
Shannon definierte die Entropie H einer gegebenen Information I über einem Alphabet Z durch
$H(I) = -\sum_{j=1}^{|Z|} p_j \cdot \log_2 p_j$,
wobei unter der Begriff Alphabet hier keine Buchstaben verstanden werden soll, sondern die Menge aller möglichen Ergebnisse des Zufallsprozesses.

$$H(I) = -(8 \cdot p \cdot \log_2 p) = -(8 \cdot 1/8 \cdot \log_2(1/8)) = -(1 \cdot (-3)) = 3$$

Lösung:

Informationsgehalt eines Zeichens sind für die gegebenen Folge nach der Formel gleich, da 0 und 1 von beiden Quellen mit der gleichen Wahrscheinlichkeit gesendet werden. Zu erkennen ist, dass die Zeichen der ersten Quelle durch eine sich wiederholende Struktur geordnet sind. Deshalb würde man intuitiv in der ersten Kette weniger Information als in der zweiten Kette vermuten. Hieraus kann man die Folgen jenachdem entweder als statistisch unabhängiges Ereignis oder jedes Zeichen einzeln betrachten und nicht der eventuelle Zusammenhang mehrerer Zeichen berücksichtigen.
Wenn die Folge aus den voneinander abhängigen Folgenglieder besteht, sind die nächsten Folgenglieder leichter vorherzusagen und damit ist ihre Informationsgehalt, aufgrund der höheren Wahrscheinlichkeit des Zeichens, nahezu Null. Man sieht dies an

Übertragungskanal



Zusammenfassung

Transinformation: $I(X|Y) = H(X) - H(X|Y) = I(Y|X)$

Zusammenfassung

Transinformation: $I(X|Y) = H(X) - H(X|Y) = I(Y|X)$
Äquivokation: $H(X|Y) = H(X) - I(X|Y)$

Zusammenfassung

Transinformation: $I(X|Y) = H(X) - H(X|Y) = I(Y|X)$
Äquivokation: $H(X|Y) = H(X) - I(X|Y)$
Irrelevanz: $H(Y|X) = H(Y) - I(X|Y)$

Zusammenfassung

Transinformation: $I(X|Y) = H(X) - H(X|Y) = I(Y|X)$
Äquivokation: $H(X|Y) = H(X) - I(X|Y)$
Irrelevanz: $H(Y|X) = H(Y) - I(X|Y)$
Totalinformation: $H(Y|X) + I(X|Y) + H(X|Y)$

Kanaleigenschaften

- deterministisch: $H(Y|X) = 0$

Kanaleigenschaften

- deterministisch: $H(Y|X) = 0$
- verlustfrei: $H(X|Y) = 0$

Kanaleigenschaften

- deterministisch: $H(Y|X) = 0$
- verlustfrei: $H(X|Y) = 0$
- störungsfrei: $H(X|Y) = H(Y|X) = 0$

Kanaleigenschaften

- deterministisch: $H(Y|X) = 0$
- verlustfrei: $H(X|Y) = 0$
- störungsfrei: $H(X|Y) = H(Y|X) = 0$
- nutzlos: $I(X|Y) = 0$

Kanaleigenschaften

- deterministisch: $H(Y|X) = 0$
- verlustfrei: $H(X|Y) = 0$
- störungsfrei: $H(X|Y) = H(Y|X) = 0$
- nutzlos: $I(X|Y) = 0$

Die Zufallsvariablen X und Y sind genau dann stochastisch unabhängig, wenn der Kanal nutzlos ist.
$H(X,Y) = H(X) + H(Y)$

Kanalkapazität

Definition: Kanalkapazität

$C = max_{P(X)}\{I(X|Y)\}$
C ist die höchste Informationsmenge, die unter allen möglichen Quellenverteilungen über den Kanal übertragen werden kann.

Aufgabe

Gegeben sei folgender binärer, asymetrischer Kanal über den Alphabeten $\mathcal{X} = \mathcal{Y} = \{0, 1\}$:

$$P(0|0) = 1 - \beta \quad P(1|0) = \beta$$
$$P(0|1) = 0 \quad P(1|1) = 1$$

a) Für welches β hat der Kanal maximale Kapazität?

Aufgabe

Gegeben sei folgender binärer, asymetrischer Kanal über den Alphabeten $\mathcal{X} = \mathcal{Y} = \{0, 1\}$:

$$P(0|0) = 1 - \beta \quad P(1|0) = \beta$$
$$P(0|1) = 0 \quad P(1|1) = 1$$

a) Für welches β hat der Kanal maximale Kapazität?

b) Für welches $H(X)$ wird bei gegebener Wahrscheinlichkeitsverteilung β der Wert von $H(Y)$ maximal?

Aufgabe

Gegeben sei folgender binärer, asymetrischer Kanal über den Alphabeten $\mathcal{X} = \mathcal{Y} = \{0, 1\}$:

$$P(0|0) = 1 - \beta \quad P(1|0) = \beta$$
$$P(0|1) = 0 \quad P(1|1) = 1$$

a) Für welches β hat der Kanal maximale Kapazität?

b) Für welches $H(X)$ wird bei gegebener Wahrscheinlichkeitsverteilung β der Wert von $H(Y)$ maximal?

c) Berechne Irrelevanz, Äquivokation und Transinformation für $\beta = 0.9$ und Gleichverteilung auf X.

Definition: Codierung

Sei $\mathbb{A} = \{a_1, .., a_m\}$ ein Alphabet und A eine Zufallsvariable über $\mathbb{A}$ mit den Wahrscheinlichkeiten $p_i = P(A = a_i)$. Eine Codierung von A über einem Codealphabet $\mathbb{X} = \{x_1, .., x_d\}$ ist eine Abbildung: $C : \mathbb{A} \to \mathbb{X}^+$. Diese wird für ganze Worte erweitert: $C^* : \mathbb{A}^* \to \mathbb{X}^*$, $a_{i_1} ... a_{i_k} = c_{i_1} ... c_{i_k}$

Definition: Codierung

Sei $\mathbb{A} = \{a_1, .., a_m\}$ ein Alphabet und A eine Zufallsvariable über $\mathbb{A}$ mit den Wahrscheinlichkeiten $p_i = P(A = a_i)$. Eine Codierung von A über einem Codealphabet $\mathbb{X} = \{x_1, .., x_d\}$ ist eine Abbildung: $C : \mathbb{A} \to \mathbb{X}^+$. Diese wird für ganze Worte erweitert: $C^* : \mathbb{A}^* \to \mathbb{X}^*$, $a_{i_1} ... a_{i_k} = c_{i_1} ... c_{i_k}$

Definition: Codelänge

Die mittlere Codelänge ist $L(C) = \sum_{i=1}^{m} p_i \cdot l_i$

Codeeigenschaften

- regulär, wenn C injektiv

Codeeigenschaften

- regulär, wenn C injektiv
- dekodierbar, wenn C^* injektiv

Codeeigenschaften

- regulär, wenn C injektiv
- dekodierbar, wenn C^* injektiv
- Präixcode, wenn kein c_i Präfix eines andern c_j ist.

Konstruktion optimaler Codes

Satz

Für jeden Präfix- und dekodierbaren Code gilt die Kraft-Ungleichung:

$$\sum_{i=1}^{m} d^{-l_i} \leq 1$$

Wenn diese Gleichung erfüllt ist, gibt es einen Präfix- bzw. dekodierbaren Code mit diesen Längen l_i

Kodierungstheorem

Theorem

Die Länge eines dekodierbaren Codes C ist mindestens so gross wie die Entropie der Zufallsvariablen A:
$L(C) \geq H_d(p)$
mit Gleichheit $\Leftrightarrow d^{-l_i} = p_i$ für alle i.

Kodierungstheorem

Theorem

Die Länge eines dekodierbaren Codes C ist mindestens so gross wie die Entropie der Zufallsvariablen A:
$L(C) \geq H_d(p)$
mit Gleichheit $\Leftrightarrow d^{-l_i} = p_i$ für alle i.

Ein optimaler Präfixcode L^* hat die Länge:
$H_d(p) \leq L^* < L(C) < H_d(p) + 1$

Huffman-Code

Huffman-Code

Ein Huffman-Code liefert einen beweisbar optimalen Code für gegebene Wahrscheinlichkeiten.

Konstruktionsskizze:

- Erstelle einen Wald mit Bäumen, für jedes Zeichen. Diese Bäume enthalten nur einen Knoten: das Zeichen
- suche die beiden Bäume im Wald, die für die Zeichen mit der kleinsten Wahrscheinlichkeit stehen. Entferne diese Bäume aus dem Wald. Erstelle einen neuen Baum, der die beiden entfernten Bäume als Unterbaum hat. Füge diesen Baum in den Wald ein. Benutze dabei die Summe der Wahrscheinlichkeiten der Unterbäume.
- Wiederhole, bis nur noch ein Baum übrig ist.

Aufgabe

a) Entwickle einen Huffman-Code zu der Zeichenkette $ABRACADABRA$

Aufgabe

a) Entwickle einen Huffman-Code zu der Zeichenkette $ABRACADABRA$

b) Dekodiere ...

Aufgabe

a) Eine Quelle sende Zeichen aus dem Alphabet $X = \{A, B, C, D, E, F\}$ aus.
Die Auftrittswahrscheinlichkeiten seien:
$p_A = 0.25, p_B = 0.06, p_C = 0.04, p_D = 0.1, p_E = 0.35, p_F = 0.2$
Erzeuge einen ternären Huffman-Code C_X mit Symbolen aus $\mathcal{C} = \{1, 2, 3\}$ und berechne die erwartete Codelänge $L(C_X)$.

Aufgabe

a) Eine Quelle sende Zeichen aus dem Alphabet $X = \{A, B, C, D, E, F\}$ aus.
Die Auftrittswahrscheinlichkeiten seien:
$p_A = 0.25, p_B = 0.06, p_C = 0.04, p_D = 0.1, p_E = 0.35, p_F = 0.2$
Erzeuge einen ternären Huffman-Code C_X mit Symbolen aus $\mathcal{C} = \{1, 2, 3\}$ und berechne die erwartete Codelänge $L(C_X)$.

b) Betrachte eine Quelle Y mit Alphabet $\mathcal{Y}, |\mathcal{Y}| = 7$. Für diese Quelle wurde ein optimaler Code C_Y gefunden für den gilt: $L(C_Y) = H_3(Y)$.
Bestimme eine mögliche Wahrscheinlichkeitsverteilung $p(y)$ für die Quelle Y.

Aufgabe

Gegeben sei eine Quelle mit Alphabet $\{A, B, C, D\}$ und mit den Wahrscheinlichkeiten $P(A) = \frac{1}{2}$, $P(A) = \frac{1}{4}$, $P(A) = \frac{1}{8}$, $P(A) = \frac{1}{8}$ auftreten.

1. Berechnen Sie die Entropie der Quelle.

Aufgabe

Gegeben sei eine Quelle mit Alphabet $\{A, B, C, D\}$ und mit den Wahrscheinlichkeiten $P(A) = \frac{1}{2}$, $P(A) = \frac{1}{4}$, $P(A) = \frac{1}{8}$, $P(A) = \frac{1}{8}$ auftreten.

1. Berechnen Sie die Entropie der Quelle.
2. Erstellen Sie einen entsprechende Huffman-Codierung.

Aufgabe

Gegeben sei eine Quelle mit Alphabet $\{A, B, C, D\}$ und mit den Wahrscheinlichkeiten $P(A) = \frac{1}{2}$, $P(A) = \frac{1}{4}$, $P(A) = \frac{1}{8}$, $P(A) = \frac{1}{8}$ auftreten.

1. Berechnen Sie die Entropie der Quelle.
2. Erstellen Sie einen entsprechende Huffman-Codierung.
3. Was ist die mittlere Codewortlänge? Gibt es einen Zusammenhang zur Entropie?

Aufgabe

Gegeben sei eine Quelle mit Alphabet $\{A, B, C, D\}$ und mit den Wahrscheinlichkeiten $P(A) = \frac{1}{2}$, $P(A) = \frac{1}{4}$, $P(A) = \frac{1}{8}$, $P(A) = \frac{1}{8}$ auftreten.

1. Berechnen Sie die Entropie der Quelle.
2. Erstellen Sie einen entsprechende Huffman-Codierung.
3. Was ist die mittlere Codewortlänge? Gibt es einen Zusammenhang zur Entropie?
4. Gegeben sei der folgende Huffman-Baum:
   $A = 00$, $B = 01$ und $C = 1$
   Dekodieren Sie 0110111011100101011. Ist der Huffman-Code geeignet?

Lösung:

1. $H(I) = 1,75$

Lösung:

1. $H(I) = 1,75$
2. Hinweis: Code ist nicht eindeutig.
   $A = 0$, $B = 10$, $C = 110$ und $D = 111$

Lösung:

1. $H(I) = 1,75$
2. Hinweis: Code ist nicht eindeutig.
   $A = 0$, $B = 10$, $C = 110$ und $D = 111$
3. Die mittlere Codewortlänge ergibt sich aus der Summe über alle Codewörter aus dem Produkt ihrer Auftrittswahrscheinlichkeit und ihrer Codewortlänge. Also ist die mittlere Codewortlänge
   $= \frac{1}{2} * 1 + \frac{1}{4} * 2 + \frac{1}{8} * 3 + \frac{1}{8} * 3 = 1,75$. Die mittlere Codewortlänge ist immer größer gleich der Entropie. Hier erreicht sie sogar den minimale Wert, weil die Auftrittswahrscheinlichkeiten 2'er-Potenzen sind.

Lösung:

1. $H(I) = 1,75$
2. Hinweis: Code ist nicht eindeutig.
   $A = 0$, $B = 10$, $C = 110$ und $D = 111$
3. Die mittlere Codewortlänge ergibt sich aus der Summe über alle Codewörter aus dem Produkt ihrer Auftrittswahrscheinlichkeit und ihrer Codewortlänge. Also ist die mittlere Codewortlänge
   $= \frac{1}{2} * 1 + \frac{1}{4} * 2 + \frac{1}{8} * 3 + \frac{1}{8} * 3 = 1,75$. Die mittlere Codewortlänge ist immer größer gleich der Entropie. Hier erreicht sie sogar den minimale Wert, weil die Auftrittswahrscheinlichkeiten 2'er-Potenzen sind.
4. Decodiert ist es BCBCCBCACBBC. Der Huffmancode eignet sichgut, wegen den Aufrittswahrscheinlichkeiten.

One-Time-Pad-Verschlüsselung

Eine buchstabensweise One-Time-Pad-Verschlüsselung:[1] ist die buchstabensweise Addition von Klartext und Schlüssel. Hierzu ordnet man im einfachsten Fall den 26 Großbuchstaben des lateinischen Alphabets Zahlen zu, die ihrer Position im Alphabet entsprechen. Sei m_i die Zahl des i-ten Buchstabens des Klartextes und k_i ist die Zahl des i-ten Buchstabens des Schlüssels. Die $c_i = k_i + m_i \mod 26$ ist die Zahl, die einem Zeichen des Alphabets zugeordnet ist. Zur Entschlüsselung berechnet man dementsprechend die Zahl des i-ten Klartextes $m_i = k_i - c_i \mod 26$. Grundlegende Voraussetzungen zur Wahrung der Sicherheit des One-Time-Pad-Verfahrens sind:

i) der Schlüssel muss geheim bleiben,

ii) der Schlüssel muss unvorhersagbar zufällig sein und

iii) der Schlüssel darf nur einmal verwendet werden!

[1] Normalerweise ist die Verschlüsselungsfunktion XOR über dem Körper $\mathbb{F}_2$

Aufgabe

Gegeben seien ein Chiffretext $c = EAEXMIVKSZQHJKZSVC$ und der dazugehörige Schlüssel $k = WZSLXWMFQUDMPJLYQO$. Geben Sie den Klartext an, der mit dem angegebenen Schlüssel k one-time-pad verschlüsselt ist. Die Addition der Buchstaben sind in der Tabelle zu sehen:

Quellen

Pajor - Informatik 4 Tutorium SS2007

Prautzsch - Skript Informatik 4 SS2008

Ash: Information Theory. Dover 1990. ISBN 0-486-66521-6
http://www.ubka.uni-karlsruhe.de/hylib-bin/suche.cgi?opacdb=UBKA_OPAC&nd=9866904

Reflexion

Was haben wir heute gelernt?

Reflexion

Was haben wir heute gelernt?

- Übungsblatt 7 besprochen

Reflexion

Was haben wir heute gelernt?

- Übungsblatt 7 besprochen
- Grundbegriffe der Informationstheorie kennen gelernt

Reflexion

Was haben wir heute gelernt?

- Übungsblatt 7 besprochen
- Grundbegriffe der Informationstheorie kennen gelernt
- Übertragungskanal behandelt

Reflexion

Was haben wir heute gelernt?

- Übungsblatt 7 besprochen
- Grundbegriffe der Informationstheorie kennen gelernt
- Übertragungskanal behandelt
- Huffmancode

Reflexion

Was haben wir heute gelernt?

- Übungsblatt 7 besprochen
- Grundbegriffe der Informationstheorie kennen gelernt
- Übertragungskanal behandelt
- Huffmancode
- One-Time-Pad Verschlüsselung

Noch Fragen?

Vorschau

Vorschau

- Zero Knowledge

Vorschau

- Zero Knowledge
- evt. Klausurwiederholung, daher bitte konkrete Wünsche per Email!

Bis zum letzten Mal