Linearer Kongruenzgenerator: $X_{i+1} = (a x_i + b) \bmod n$

Satz: Die maximale Periodenlänge n wird genau dann erreicht, wenn: (→ Knuth, Vol(2))

(1) $ggT(b, n) = 1$

(2) $a \equiv 1 \pmod{p}$ für alle Primteiler p von n

(3) $4 \mid n \implies a \equiv 1 \pmod 4$

Korollar 1: Falls $n = 2^k$ $(k \geq 2)$, dann wird die maximale Periode n genau dann erreicht, falls

b ungerade und $a \equiv 1 \pmod 4$

Korollar 2: Falls n Primzahl, dann wird die maximale Periode n genau dann erreicht, falls

$\underbrace{ggT(b, n) = 1}_{b \neq 0,\ b \neq n}$ und $a = 1$

Beispiel zu Korollar 1:

Sei $n = 16$, $a = 5$, $b = 3$, $X_0 = 1$

ergibt die Folge:

1, 8, 11, 10, 5, 12, 15, 14, 9, 0, 3, 2, 13, 4, 7, 6, 1

Beispiel zu Korollar 2:

Sei $n = 17$, $b = 8$, $a = 1$, $X_0 = 1$

ergibt:

1, 9, 0, 8, 16, 7, 15, 6, 14, 5, 13, 4, 12, 3, 11, 2, 10, 1

Vergleiche: Double Hashing (Vorlesung Algorithmen)
Man wählt Hashtabellengröße als Primzahl n,
und $b \in \{1, \dots, n-1\}$ ist ~~die~~ der Wert der 2. Hashfunktion
bei einer Kollision.

Anders ausgedrückt: Die Struktur $(\mathbb{Z}_n, \cdot_{\bmod n})$
ist eine zyklische Gruppe, wobei jedes Element
$b \in \mathbb{Z}_n \setminus \{0\}$ ein Erzeuger ist.

Entropiebegriff nach Shannon

Sei $P = (p_1, p_2, \dots, p_n)$ eine Wahrscheinlichkeitsverteilung, also $\sum_{i=1}^{n} p_i = 1$. Dann ist

$$H(P) := -\sum_{i=1}^{n} p_i \cdot \log_2 p_i \qquad \text{(wobei } 0 \cdot \log 0 = 0\text{)}$$

die Entropie von P.

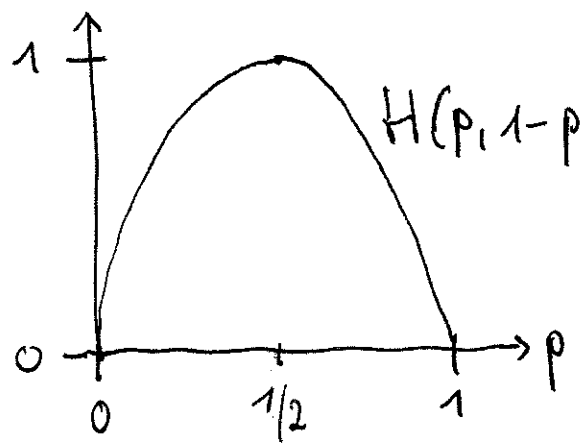
Es gelten folgende Eckwerte:

$$0 \leq H(P) \leq \log_2 n$$

↑ Gleichheit falls P deterministische Verteilung ist: $P = (0, \dots, 0, 1, 0, \dots, 0)$

↖ Gleichheit falls P die Gleichverteilung ist: $P = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$

Spezialfall $n = 2$: $P = (p, 1-p)$



(Ähnelt dem Verlauf von $-\log_2 IC(p, 1-p)$)

Interpretation von $H(P)$:

P ist die Wahrscheinlichkeitsverteilung, die einem Zufallsexperiment zugrunde liegt.

$H(P)$ kann man interpretieren als Maßzahl für die Unsicherheit über den Ausgang des Experiments, bevor man das Ergebnis erfährt.

Man kann $H(P)$ auch interpretieren als den Informationsgehalt, den man erhält, wenn man das Ergebnis des Experiments erfährt.

Beispiele: $H(\text{Münzwurf}) = 1 \quad [\text{bit}]$

$H(\underbrace{\tfrac{1}{26}, \tfrac{1}{26}, \ldots, \tfrac{1}{26}}) = \log_2 26 = 4{,}7$

Gleichverteilung auf 26 Buchstaben

$H(\text{Verteilung der (Einzel-) Buchstaben in deutscher Sprache}) = 4{,}1$

$H(\text{Verteilung von Bigrammen in Deutsch}) = 2 \cdot 3{,}5$

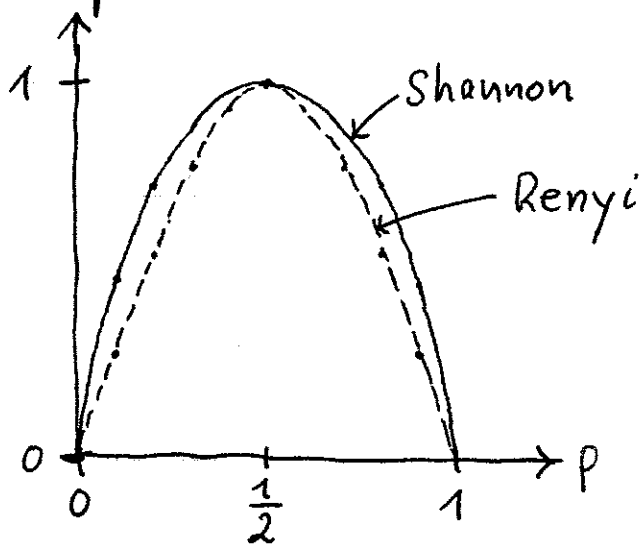
$H(\text{Verteilung von Trigrammen in Deutsch}) = 3 \cdot 3{,}2$

für großes n: $H(\text{Verteilung von } n\text{-Grammen in Deutsch}) = n \cdot 1{,}5$

Shannon-Entropie versus Renyi-Entropie

$$-\sum_{i=1}^{n} p_i \log_2 p_i \qquad -\log_2\left(\sum_{i=1}^{n} p_i^2\right)$$

für den Fall $n=2$, also $P=(p, 1-p)$



Frage: Aus einem Skatspiel mit 32 (verschiedenen) Karten wird zufällig eine Karte gezogen. Welchen Informationsgehalt (= Shannon-Entropie) hat dieses Zufallsexperiment? Welchen Wert hat der Koinzidenzindex bzw. Renyi-Entropie?

Antwort: Shannon: $-\sum_{i=1}^{32} \frac{1}{32} \cdot \log_2 \frac{1}{32} = \log_2(32) = 5$ [bit]

Koinzidenzindex: $\sum_{i=1}^{32} \left(\frac{1}{32}\right)^2 = \frac{1}{32} \approx 0{,}031$

Renyi: $-\log_2\left(\frac{1}{32}\right) = 5$

Wenn X eine Zufallsvariable ist, die ihre Werte mit Wahrscheinlichkeitsverteilung P annimmt, so schreibt man auch $H(X)$ statt $H(P)$.

Varianten der Definition:

gemeinsame Entropie zweier Zufallsvariablen

$$H(X,Y) := -\sum_{i,j} P(X=x_i, Y=y_j) \cdot \log_2 P(X=x_i, Y=y_j)$$

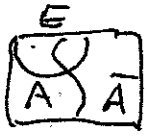
Satz: $H(X,Y) \leq H(X) + H(Y)$

↑ Gleichheit, falls X und Y stochastisch unabhängig

bedingte Entropie

$$H(X|E) := -\sum_i P(X=x_i|E) \cdot \log_2 P(X=x_i|E)$$

bedingte W'heit: $P(A|E) = \frac{P(A \cap E)}{P(E)}$



$$H(X|Y) := \sum_j H(X|Y=y_j) \cdot \log_2 P(X|Y=y_j)$$

$$\sum_j P(Y=y_j) \cdot H(X|Y=y_j)$$

Eigenschaften:

$H(X|Y) = 0$, falls X durch Y eindeutig bestimmt wird, also $X = f(Y)$.

$H(X|Y) \leq H(X)$, falls X und Y unabhängig sind
↑ Gleichheit

(dann gilt auch: $H(Y|X) = H(Y)$)

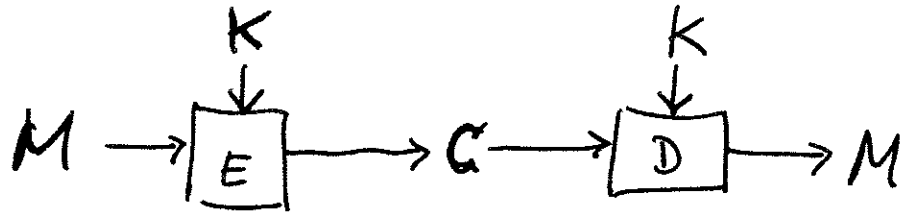
$$H(X,Y) = H(X) + H(Y|X) \qquad (*)$$
$$= H(Y) + H(X|Y)$$

Man definiert:

$$I(X,Y) := H(X) + H(Y) - H(X,Y)$$
$$= H(X) - H(X|Y) \qquad \text{wegen } (*)$$
$$= H(Y) - H(Y|X)$$

Transinformation, Synentropie, mutual information

Shannon's Modell eines symmetrischen Kryptosystems:



M, K, C werden als Zufallsvariablen betrachtet

Bsp: $P(M=m)$ ... W'heit, dass konkrete Nachricht gleich m ist.

M ist gemäß deutscher Sprache verteilt

Modellannahmen:

M und K sind stochast. unabhängig

Somit gelten folgende Gleichungen:

$$H(M,K) = H(M) + H(K)$$

$$H(M|K) = H(M)$$

$$H(K|M) = H(K) \leq \log_2 \underbrace{|K|}$$

Mächtigkeit der Grundmenge als Anzahl möglicher Schlüssel

Es gilt $C = E(M,K)$. Somit gilt:

$$H(C \mid M,K) = 0$$

Es gilt $M = D(C,K)$. Somit gilt:

$$H(M \mid C,K) = 0$$

$$H(M,K,C) = H(M) + \underbrace{H(K \mid M)}_{=H(K)} + \underbrace{H(C \mid M,K)}_{=0}$$

$$= H(M) + H(K)$$

Ebenso:

$$H(M,K,C) = H(K) + H(C \mid K) + \underbrace{H(M \mid C,K)}_{=0}$$

$$= H(K) + H(C \mid K) = H(K,C)$$

Eingesetzt ergibt sich:

$$\begin{aligned} H(K|C) &= H(K,C) - H(C) \\ &= H(M,K,C) - H(C) \\ &= H(M) + H(K) - H(C) . \end{aligned}$$

Def: Kryptosystem $(\mathcal{M}, \mathcal{K}, \mathcal{C})$ mit Verschlüsselungsfunktionen D, E heißt <u>absolut sicher</u> (oder: informationstheoretisch sicher), falls

$$H(M) = H(M|C).$$

Das heißt inhaltlich: Keinerlei Information über M wird bekannt, wenn man C kennt.

Heißt nichts anderes als dass die Zufallsvariablen M und C unabhängig sind.

<u>Beispiel:</u> Bei Vigenère-Chiffre, sofern Schlüssellänge $\ll$ Klartextlänge, lässt sich aus Chiffre in nahezu eindeutiger Weise der Klartext ermitteln. In diesem Fall gilt also: $H(M) =$ Klartextlänge $\cdot 1{,}5$

$$\text{aber } H(M|C) \approx 0$$

Vernam-chiffre

Def: Ein **one-time pad** Kryptosystem hat folgende Eigenschaften: $|M| = |K| = |C|$

(Zum Beispiel: $M = K = C = \Sigma^n$ für ein Alphabet Σ und Blocklänge n.)

Die Schlüssel $k \in K$ werden unter Gleichverteilung gewählt: $P(K = k) = \frac{1}{|K|}$.

Zu gegebenem m und c gibt es eindeutigen Schlüssel $k = k_{m,c}$ mit $E(m,k) = c$.

Bei festem m und Variation aller k-Werte sollte $c = E(m,k)$ alle Werte in C durchlaufen.

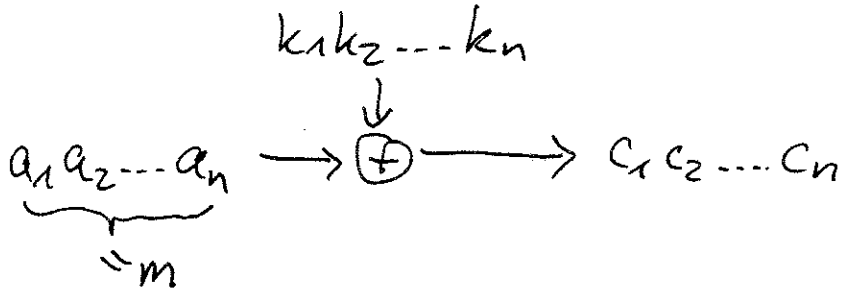
Beispiel für one-time pad (wobei $\Sigma = \{0, 1, \ldots, \ell-1\}$)

$$E(a_1 a_2 \ldots a_n, k_1 k_2 \ldots k_n) = c_1 c_2 \ldots c_n$$

wobei $c_i = (a_i + k_i) \bmod |\Sigma|$

Spezialfall: $\Sigma = \{0,1\}$, dann ist

$$c_i = a_i \text{ XOR } k_i$$



Satz: Ein one-time pad Kryptosystem ist absolut sicher.

Beweis: Zu zeigen: Zufallsvariablen M und C sind unabhängig.

Zu gegebenem m,c gibt es genau einen Schlüssel $k = k_{m,c}$ mit $E(m,k) = c$ bzw. $D(c,k) = m$.

Somit gilt für alle m,c:

$$P(M=m, C=c) = \sum_{\substack{k: \\ E(m,k)=c}} P(M=m) \cdot P(K=k)$$

↖ Unabhängigkeit von M und K

$$= P(M=m) \cdot P(K=k_{m,c})$$

$$= P(M=m) \cdot \frac{1}{|K|}$$

$$P(C=c) = \sum_{m,k:} P(M=m) \cdot P(K=k)$$

$$= \sum_m P(M=m) \cdot \underbrace{P(K=k_{m,c})}_{=\frac{1}{|K|}}$$

$$= \frac{1}{|K|} \cdot \underbrace{\sum_m P(M=m)}_{=1} = \frac{1}{|K|}$$

Zusammenfasst:

$$\underline{P(M=m, C=c)} = P(M=m) \cdot \frac{1}{|K|} = \underline{P(M=m) \cdot P(C=c)}$$

Somit sind M und C unabhängig. □