ETH Zürich
Dept. Informatik

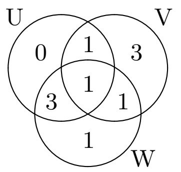
Prof. Markus Gross
WS 05/06

Informationstheorie

Lösung 4

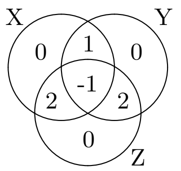
4.1 Entropiediagramme

a)



H(UVW) = 10	I(U;V|W) = 1
	I(U;W|V) = 3
H(U) = 5	I(V;W|U) = 1
H(V) = 6	
H(W) = 6	H(U|VW) = 0
	H(V|UW) = 3
I(U;V) = 2	H(W|UV) = 1
I(U;W) = 4	
I(V;W) = 2	R(U;V;W) = 1

b)



H(XYZ) = 4	I(X;Y|Z) = 1
	I(X;Z|Y) = 2
H(X) = 2	I(Y;Z|X) = 2
H(Y) = 2	
H(Z) = 3	H(X|YZ) = 0
	H(Y|XZ) = 0
I(X;Y) = 0	H(Z|XY) = 0
I(X;Z) = 1	
I(Y;Z) = 1	R(X;Y;Z) = -1

c)

(i) z.B.: $X = [X_1, X_2]$, $Y = [X_3, X_4, X_5]$ und $Z = [Y, X_2, X_4, X_5]$, wobei $X_1, \dots X_5$ unabhängige, faire Münzwürfe sind und $Y = X_1 \otimes X_3$ ist.

(ii) Dieses Diagramm ist nicht realisierbar, weil $I(X;Y) = -1 < 0$.

(iii) z.B.: Seien X_1, X_2, X_3, X_4 unabhängige, *unfaire* Münzwürfe mit $\forall i : \mathrm{P}(X_i = 1) \approx 0.11 \Rightarrow \mathrm{h}(X_i) = \frac{1}{2}$. $X := [X_1, X_2, X_4]; Y := [X_1, X_3, X_4]; Z := [X_2, X_3, X_4]$.

(iv) z.B.: Seien X und Y unabhängige, faire Münzwürfe. $Z := X \otimes Y$.

4.2 Verteilung für eine gegebene Entropie

a) Eine Wahrscheinlichkeitsverteilung mit den folgenden Wahrscheinlichkeiten besitzt die verlangten Eigentschaften: $\frac{1}{2}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}$ und $\frac{1}{8}$.

b) Eine solche Wahrscheinlichkeitsverteilung ist nicht möglich, da für jede Zufallsvariable X mit Wertebereich $\mathcal{X}$ gilt: $H(X) \leq \log_2(|\mathcal{X}|)$, und $3 > \log_2(5) \approx 2.585$.

4.3 Berechnung und Entropie

a) Wahr, denn $H(X) = H(X\ f(X))) \geq H(f(X))$. Die erste Gleichung gilt, weil $X\ f(X)$ dieselbe Wahrscheinlichkeitsverteilung hat wie X.

b) Wahr. $H(f(X)\ g(Y)) = H(l(XY)) \leq H(XY)$, wobei $l(XY) := f(X)\ g(Y)$.

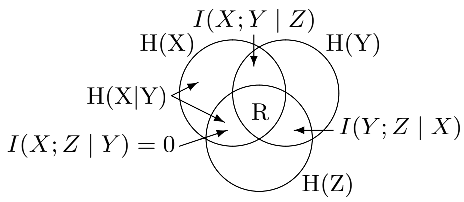
c) Nicht allgemeingültig. Gegenbeispiel: X sei der Ausfall einer fairen Münze. Definiere $Y := X, f(z) = z, g(z) = 0$. Dann gilt: $H(f(X) \mid g(Y)) = H(X \mid 0) = H(X) = 1 > 0 = H(X \mid Y)$.

d) Wahr, denn $I(f(X); Z) = H(Z) - H(Z \mid f(X)) \leq H(Z) - H(Z \mid X) = I(X; Z)$ und somit $I(f(X); g(Y)) \leq I(X; g(Y)) = I(g(Y); X) \leq I(Y; X) = I(X; Y)$.

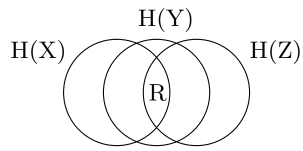
4.4 Markov-Ketten & Entropiediagramme

a)

Man erhält folgendes Diagramm:



... was sich, da $I(X; Z \mid Y) = 0$, zur Verdeutlichung auch so darstellen lässt:



Es unterscheidet sich von einem Diagramm dreier beliebig verteilter Zufallsvariablen durch die Tatsache, dass $I(X; Z \mid Y) = 0$.

b) Es gilt $I(X; Z) = R(X; Y; Z) + I(X; Z \mid Y)$ und $I(Y; Z) = R(X; Y; Z) + I(Y; Z \mid X)$.
Da $I(X; Z \mid Y) = 0$ und $I(Y; Z \mid X) \geq 0$, folgt sofort dass $I(X; Z) \leq I(Y; Z)$.
Der Beweis für die zweite Ungleichung geht analog.

4.5 Kumulierter Gewinn oder Verlust

Sei Y_i die Zufallsvariable, die dem i-ten Münzwurf entspricht. Codieren wir Kopf mit 1 und Zahl mit -1, so lässt sich X_i schreiben als $X_i = \sum_{j=1}^{i} Y_j$.

a) Für jedes i ist (wegen der Linearität des Erwartungswertes)

$$E(X_i) = E\left(\sum_{j=1}^{i} Y_j\right) = \sum_{j=1}^{i} E(Y_j) = 0.$$

Die gemeinsame Verteilung ist:

$P_{X_2X_3}$		X_3 -3	-1	1	3
	-2	1/8	1/8	0	0
X_2	0	0	1/4	1/4	0
	2	0	0	1/8	1/8

b) Es ist $P_{X_3}(-3) = P_{X_3}(3) = \frac{1}{8}$ und $P_{X_3}(-1) = P_{X_3}(1) = \frac{3}{8}$.
Also erhält man $H(X_3) = -(2 \cdot \frac{1}{8} \log \frac{1}{8} + 2 \cdot \frac{3}{8} \log \frac{3}{8}) \approx 1.811$.

$$\begin{aligned}
H(X_0 \cdots X_{10}) &= H(X_0 \cdots X_9) + H(X_{10}|X_0 \cdots X_9) \\
&= H(X_0 \cdots X_8) + H(X_9|X_0 \cdots X_8) + H(X_{10}|X_0 \cdots X_9) \\
&= H(X_0) + \sum_{i=1}^{10} H(X_i|X_0 \cdots X_{i-1}) \\
&= H(X_0) + \sum_{i=1}^{10} H(X_i|X_{i-1}) \\
&= H(X_0) + 10 = 10.
\end{aligned}$$

Im letzten Schritt haben wir ausgenutzt, dass $H(X_i|X_{i-1}) = 1$, da die Unsicherheit über X_i gegeben X_{i-1} genau der Unsicherheit über den letzten Münzwurf entspricht. Dies kann auch bewiesen werden:

$$\begin{aligned}
H(X_{i+1}|X_i) &= H(X_iX_{i+1}) - H(X_i) && (1) \\
&= H(X_iY_{i+1}) - H(X_i) && (2) \\
&= H(X_i) + H(Y_{i+1}) - H(X_i) && (3) \\
&= H(Y_{i+1}) = 1
\end{aligned}$$

In (1) wurde die Definition der bedingten Entropie eingesetzt. Schritt (2) gilt, da aus (X_i, X_{i+1}) das Paar (X_i, Y_{i+1}) berechnet werden kann und umgekehrt. Gleichheit (3) gilt, da X_i und Y_{i+1} unabhängig sind.

c) Wir bemerken zunächst, dass $H(X_{i+2}|X_i) = 1.5$ unabhängig von i ist. Die Unsicherheit über X_{i+2} wenn man X_i kennt, entpricht genau der Unsicherheit über die Summe der letzten beiden Münzwürfe. Diese Summe nimmt die Werte 2,0 oder -2 mit den Wahrscheinlichkeiten 1/4, 1/2 und 1/4 an, unabhängig davon, wieviele Münzen bereits geworfen wurden.

Ähnlich wie in (b) lässt sich auch dies formal beweisen. Sei $S_i := Y_{i+1} + Y_{i+2}$, dann gilt

$$\begin{aligned}
H(X_{i+2}|X_i) &= H(X_iX_{i+2}) - H(X_i) \\
&= H(X_iS_i) - H(X_i) \\
&= H(X_i) + H(S_i) - H(X_i) \\
&= H(S_i) = H([\frac{1}{4}, \frac{1}{2}, \frac{1}{4}]) = 1.5
\end{aligned}$$

Somit ergibt sich $I(X_1; X_3) = H(X_3) - H(X_3|X_1) = H(X_3) - 1.5 = -(2 \cdot \frac{1}{8} \log \frac{1}{8} + 2 \cdot \frac{3}{8} \log \frac{3}{8}) - 1.5 \approx 0.3113$.

Analog berechnet man $I(X_5; X_7) = H(X_7) - H(X_7|X_5) = H(X_7) - 1.5$. Um die Verteilung von X_7 zu finden, überlegen wir uns als Beispiel $P(X_7 = 3)$. Eine Summe von 3 kann man nur erreichen, wenn von den 7 Münzwürfen 5 Kopf und 2 Zahl zeigen. Jede Sequenz von 7 Münzwürfen hat Wahrscheinlichkeit 2^{-7} und es gibt $\binom{7}{5}$ Sequenzen mit 5 mal Kopf. Somit ist $P(X_7 = 3) = \binom{7}{5} 2^{-7}$. Allgemein ist $P(X_7 = 2i - 7) = \binom{7}{i} 2^{-7}$ für $i = 0, \ldots, 7$. Andere Werte kann X_7 nicht annehmen.Damit ist $H(X_7) = H([\binom{7}{0} p, \binom{7}{1} p, \ldots, \binom{7}{7} p]) \approx 2.45$, wobei $p = 2^{-7}$. Schliesslich ist $I(X_5; X_7) \approx 0.947$.