AUFGABEN ZUR INFORMATIONSTHEORIE

Aufgabe 1

Wir betrachten das folgende Zufallsexperiment:

Eine faire Münze wird so lange geworfen, bis erstmals Kopf erscheint. Die Zufallsvariable X bezeichne die Anzahl der dazu notwendigen Versuche. Wird beispielsweise schon im ersten Versuch Kopf geworfen, so gilt X = 1, werden zwei Würfe benötigt, X = 2, usw.

a) Ergänzen Sie die untenstehende Tabelle mit den entsprechenden Wahrscheinlichkeiten P(X).

X	P(X)
1	
2	
3	
4	
5	
⋮	

b) Leiten Sie einen allgemeinen Ausdruck für P(X=x) her.

c) Bestimmen Sie die Entropie H(X) des Zufallsexperiments in bit.

d) Finden Sie eine effiziente Folge von Ja-Nein Fragen zur Ermittlung des Werts der Zufallsvariable X. Wieviele Fragen müssen im Mittel gestellt werden. Vergleichen Sie diese Grösse mit der Entropie H(X).

Die folgenden Beziehungen könnten sich als nützlich erweisen:

$$\sum_{n=1}^{\infty} r^n = \frac{r}{1-r} \qquad \sum_{n=1}^{\infty} n \cdot r^n = \frac{r}{(1-r)^2}$$

Aufgabe 2

Eine Nachrichtenquelle generiert die vier Symbole A, B, C und D mit den Wahrscheinlichkeiten $P(A) = 0.8$, $P(B) = 0.1$, $P(C) = 0.08$ und $P(D) = 0.02$.

a) Wie gross ist die Entropie der Quelle? Wie gross wäre die Entropie der Quelle, falls alle Zeichen mit gleicher Wahrscheinlichkeit auftreten würden?

b) Entwerfen Sie für diese Quelle einen Huffman-Code. Wie gross ist die mittlere Codewortlänge?

c) Zwei aufeinanderfolgende Symbole werden in einer Symbolkette zusammengefasst und gemeinsam codiert. Verifizieren Sie, dass die Entropie der Symbolkette gerade dopppelt so gross ist wie diejenige eines Einzelsymbols. Entwerfen Sie einen Huffmann-Code für die 16 möglichen Symbolpaare. Wie gross ist in diesem Fall die mittlere Codewortlänge pro Symbol?

Aufgabe 3

Eine zweiwertige Symbolquelle mit den Wahrscheinlichkeiten $P_1 = 0.9$ und $P_2 = 0.1$ soll so codiert werden, dass im Mittel maximal 0.81 bit/Symbol verwendet werden müssen.

a) Wieviele Symbole müssen gleichzeitig codiert werden, damit dies garantiert werden kann?

b) Entwerfen Sie einen Code, der die Anforderungen erfüllt. Wie gross ist dessen durchschnittliche Codewortlänge pro Symbol? Vergleichen Sei dies mit der Entropie der Quelle.

Aufgabe 4

Information soll mit der Rate $R \le C$ über einen Gauss'schen Kanal (Bandbreite B, Rauschleistungsdichte η) übertragen werden. Die dazu pro bit benötigte Energie bezeichnen wir mit E_b.

a) Wie ist der Zusammenhang zwischen der Nutzsignalleistung P_S und der Energie pro bit E_b?

b) Welche Energie pro bit wird minimal benötigt, damit eine fehlerfreie Übertragung möglich ist? Wie gross ist zu diesem Zweck die Rate R zu wählen?

c) Das Verhältnis E_b/η wird als Signal- zu Rauschverhältnis pro bit bezeichnet. Wie gross muss dieses bei einem nicht bandbegrenzten Kanal ($B \to \infty$) mindestens sein, damit eine fehlerfreie Übertragung möglich ist?

Die thermische Rauschleistung P_N, die ein beliebiger Widerstand abgeben kann, ist gegeben durch

$$P_N = \eta \cdot B = k \cdot T \cdot B \, .$$

Darin sind $k = 1.38 \cdot 10^{-23}$ J/K die Boltzmannkonstante, T die absolute Temperatur und B die Bandbreite.

d) Welche Energie benötigen Sie minimal, um bei Zimmertemperatur ($T = 293°K$) die Informationsmenge 1 bit über einen nicht bandbegrenzten, Gauss'schen Kanal zu übertragen?

Die folgende Beziehung könnte sich als nützlich erweisen:

$$\lim_{x\to 0}\frac{2^x-1}{x}=\ln(2)$$

Aufgabe 5

a) Zeigen Sie, dass für alle $x > 0$ gilt: $\log_2(x) \le (x - 1)/\ln(2)$.
Tipp: Zeichnen Sie die Kurven graphisch auf und denken Sie daran, dass $\log_2(x) = \ln(x)/\ln(2)$.

b) Verwenden Sie die obige Beziehung, um zu zeigen, dass für die Entropie einer Quelle mit L möglichen Symbolen immer gilt:

$$H \le \log_2(L)$$

Tipp: Zeigen Sie, dass

$$\begin{aligned} H-\log_2(L) &= \sum_{i=1}^{L} P(u_i)\cdot\log_2\left(\frac{1}{P(u_i)}\right)-\log_2(L) \\ &= \sum_{i=1}^{L} P(u_i)\cdot\left(\log_2\left(\frac{1}{P(u_i)}\right)-\log_2(L)\right) \end{aligned}$$

kleiner oder gleich Null ist.

c) Unter welchen Voraussetzung gilt $H = \log_2(L)$?

Aufgabe 6

Wir betrachten eine Quelle, die aus einem Symbolvorrat von jeweils 9 Symbolen auswählt, deren Auftrittswahrscheinlichkeiten $P(u_i)$ in der folgenden Tabelle angeben sind:

i	$P(u_i)$	$Q_i=\sum_{j=1}^{i-1}P(u_j)$	$L(u_i)=\left\lceil \log_2\left(\frac{1}{P(u_i)}\right)\right\rceil$	Die ersten $L(u_i)$ Nachkommastellen einer Binärdarstellung von Q_i
1	0.22	0.00	3	000
2	0.19	0.22	3	001
3	0.15	0.41		
4	0.12			
5	0.08			
6	0.07			
7	0.07			
8	0.06			
9	0.04			

Die in der obigen Tabelle angegebenen Werte Q_i sind die aufsummierten Auftrittswahrscheinlichkeiten

$$Q_i = \sum_{j=1}^{i-1} P(u_j)\,.$$

Berechnen Sie zunächst die Codewortlängen $L(u_i) = \lceil \log_2(1/P(u_i)) \rceil$ und tragen Sie diese in die Tabelle ein. Danach ermitteln Sie die Binärdarstellung der aufsummierten Wahrscheinlichkeit Q_i, wobei die Binärdarstellung jeweils nach $L(u_i)$ Ziffern abgebrochen wird. Beispiel:

$$Q_2 = 0.22 = 0 \cdot 2^{-1} + 0 \cdot 2^{-2} + 1 \cdot 2^{-3} + 1 \cdot 2^{-4} + 1 \cdot 2^{-5} + \ldots$$

hat die Binärdarstellung

$$0.00111\ldots$$

die nach der 3. Stelle abgebrochen und in die Tabelle eingetragen wird.

a) Vervollständigen Sie die Tabelle.

b) Die Nachkommastellen der Binärdarstellung werden als Code betrachtet. Erfüllt der derart konstruierte Code die Präfixbedingung?

c) Zeichnen Sie den Binärbaum des Codes auf und bestimmen Sie die mittlere Codewortlänge. Wie vergleicht sich diese mit der Entropie der Quelle?

d) Konstruieren Sie für die gleiche Quelle einen Huffman Code.

Aufgabe 7

Eine Nachrichtenquelle hat einen Symbolvorrat von 3 Symbolen (A,B,C) mit den folgenden Auftrittswahrscheinlichkeiten:

$$P(A) = 0.7,\ P(B) = 0.2,\ P(C) = 0.1$$

a) Einzelne Ausgangssymbole werden optimal in binäre, präfixfreie Zeichenfolgen codiert. Welche Aussage können Sie über die durchschnittliche Anzahl binärer Ziffern pro Symbol machen?

b) Nun werden jeweils 5 Symbole in eine binäre, präfixfreie Zeichenfolge codiert. Wie lautet in diesem Fall die Aussage über die durchschnittliche Anzahl binärer Ziffern pro Symbol?

c) Konstruieren Sie einen optimalen Code, der jeweils zwei Symbole der Nachrichtenquelle in eine binäre, präfixfreie Zeichenfolge codiert. Wie gross ist die durchschnittliche Codewortlänge pro Symbol?

d) Ist der in c) konstruierte Code eindeutig?

Aufgabe 8

Gegeben sei eine Informationsquelle, welche die sechs Symbole A, B, C, D, E und F mit den Wahrscheinlichkeiten

$P(A) = 0.3$,	$P(B) = 0.24$
$P(C) = 0.2$,	$P(D) = 0.15$
$P(E) = 0.07$,	$P(F) = 0.04$

erzeugt.

a) Berechnen Sie die Entropie der Informationsquelle.

b) Konstruieren Sie für die gegebene Quelle einen binären präfixfreien Code mit optimaler mittlerer Codewortlänge. Geben Sie diesen in Tabellenform an. Wie gross ist dessen mittlere Codewortlänge?

c) Die Symbole der Informationsquelle sollen mit Hilfe des unter b) konstruierten Codes über einen Telefoniekanal mit einer Bandbreite von 3.1 kHz und einem Signal-zu-Rauschverhältnis von 30 dB übertragen werden. Welche Symbolrate (in Symbolen pro Sekunde) kann bei fehlerfreier Übertragung auch mit beliebig hohem Aufwand nicht überschritten werden?
Tip: Falls Sie sie Teilaufgabe b) nicht lösen konnten, nehmen Sie einen mittlere Codewortlänge von $E[W] = 2.45$ an.

d) Jeweils drei Symbole der Informationsquelle werden nun mit einem optimalen, präfixfreien Code in ein binäres Codewort codiert. Welche Aussage können Sie über die mittlere Codewortlänge pro Symbol machen?

e) Die Symbole der Informationsquelle sollen mit einer Symbolrate von $r = 20'000$ Symbolen/s über einen nicht bandbegrenzten AWGN-Kanal mit einer Rauschleistungsdichte von $\eta = 30\mu W/Hz$ übertragen werden. Zu diesem Zweck werden jeweils sehr viele Symbole in ein einziges binäres Codewort codiert. Welche minimale Signalleistung wird dazu benötigt?

Aufgabe 9

Ein Gerät zur automatischen Bildauswertung unterscheidet nur die Zustände Schwarz und Weiss. Auf den verwendeten Bildern treten die Zustände mit den Wahrscheinlichkeiten $P(\text{Schwarz}) = 0.2$ und $P(\text{Weiss}) = 0.8$ auf.

a) Führen Sie eine Huffman-Codierung unter der Voraussetzung durch, dass jeweils 2 Bildpunkte zu einem Symbol zusammengefasst werden! Zeichnen Sie den entsprechenden Codebaum und geben Sie die Codetabelle an.

b) Wie gross ist die mittlere Codewortlänge pro Symbol?

Aufgabe 10

Sie wollen Daten über einen bandbegrenzten AWGN-Kanal übertragen. Die Bandbreite beträgt $B = 3.1$ kHz, das Signal-zu-Rauschverhältnis ist 40 dB. Die zu übertragenden Symbole stammen aus einer Nachrichtenquelle mit einer Entropie von $H = 1.76$ bit/Symbol.

a) Die spektrale Leistungsdichte des Rauschens beträgt $\eta = 3$ nW/Hz. Wie gross ist die Signalleistung?

b) Berechnen Sie die Kapazität des gegebenen Kanals.

c) Die einzelnen Symbole werden mit einem optimalen, binären, präfixfreien Code codiert. Wieviele Symbole können Sie (bei beliebig hohem Aufwand) pro Sekunde minimal über den gegebenen Kanal übertragen? Anders gefragt: Welche Symbolrate können Sie mit den vorhandenen Angaben garantieren?

d) Wie kann die Symbolrate verbessert werden und welche maximale Symbolrate kann damit erreicht werden?

e) Um welchen Faktor ändert sich die Kanalkapazität, wenn die Bandbreite verdoppelt wird, die Signalleistung jedoch gleich bleibt?

Aufgabe 11

Bestimmen Sie den Wahrheitsgehalt der nachfolgenden Aussagen (**Falsch gesetzte Kreuze führen zu einem Punkteabzug!**) Voraussetzung: Die Signalleistung soll endlich bleiben!

	Wahr	Falsch
Über einen Kanal mit der Bandbreite 3.4 kHz können grundsätzlich maximal 6.8 kbits/s übertragen werden.	❑	❑
Über einen verrauschten Kanal mit unendlicher Bandbreite können beliebig viele bits/s übertragen werden.	❑	❑
Über einen rauschfreien Kanal mit endlicher Bandbreite können beliebig viele bits/s übertragen werden.	❑	❑

Aufgabe 12

Tipp: Die Teilaufgaben d) und e) können ohne die Resultate der vorhergehenden Teilaufgaben gelöst werden.

Sie sollen einen binären präfixfreien Code für 5 Symbole entwerfen. Die Codewortlängen für die ersten vier Symbole sind mit $w_1 = 1$, $w_2 = 2$, $w_3 = 4$ und $w_4 = 4$ vorgegeben.

a) Wie gross wählen Sie die Länge w_5 des fünften Codeworts, damit die durchschnittliche Codewortlänge möglichst klein wird (inkl. Begründung)?

b) Geben Sie einen binären präfixfreien Code mit den vorgegebenen Codewortlängen an.

Die Auftretenswahrscheinlichkeiten für die fünf Symbole sind nun wie folgt gegeben:

$$P(u_1) = 0.3,\ P(u_2) = 0.3,\ P(u_3) = 0.1,\ P(u_4) = 0.1,\ P(u_5) = 0.2.$$

c) Wie gross ist die durchschnittliche Codewortlänge des unter b) konstruierten Codes?

d) Konstruieren Sie nun für die Symbolquelle einen Huffmann-Code. Wie gross ist nun die durchschnittliche Codewortlänge?

e) Um wie viele Prozent kann die durchschnittliche Codewortlänge pro Symbol noch verbessert werden und wie müsste man dazu vorgehen?

Aufgabe 13

a) Welche der nachfolgenden Codes sind sicher keine Huffman-Codes (ganz gleich wie die Auftretenswahrscheinlichkeiten verteilt sind)? Mit Begründung!

(i) {0, 10, 11}　　(ii) {00, 01, 10, 110}　　(iii) {01, 10}

Aufgabe 14

Wir betrachten eine Informationsquelle, deren 5 mögliche Symbole die folgenden Auftretenswahrscheinlichkeiten $P(u_i)$ aufweisen:

u_i	A	B	C	D	E
$P(u_i)$	1/2	1/6	1/6	1/8	1/24

Beim so genannten Shannon-Code werden die Codewortlängen gemäss der folgenden Beziehung gewählt:

$$w_i = \left\lceil \log_2\left(\frac{1}{P(u_i)}\right)\right\rceil$$

a) Zeigen Sie, dass mit diesen Codewortlängen ein binärer, präfixfreier Code für die gegebene Quelle gebildet werden kann.

b) Berechnen Sie die durchschnittliche Codewortlänge dieses Codes.

c) Konstruieren Sie für die gegebene Quelle zwei binäre präfixfreie Codes mit optimaler mittlerer Codewortlänge. Die beiden Codes sollen unterschiedliche Mengen von Codewortlängen aufweisen.

d) Welches sind die durchschnittlichen Codewortlängen der beiden unter c) konstruierten Codes?

e) Überprüfen Sie den Wahrheitsgehalt der folgenden Aussage: „Die Codewortlängen eines optimalen binären präfixfreien Codes sind für alle Symbole immer kleiner gleich der Codewortlängen eines Shannon-Codes". Begründen Sie Ihre Antwort.

Unter Umständen ist man nicht daran interessiert, einen Code mit möglichst kleiner mittlerer Codewortlänge zu konstruieren. Es ist beispielsweise denkbar, dass das Symbol „E" (obwohl es nur relativ selten auftritt) eine hohe Priorität hat und deshalb einem kurzen Codewort zugeordnet werden sollte. Um dies zu berücksichtigen werden für jedes Symbol u_i unterschiedliche Kosten pro Bit c_i definiert. Die Kosten für die Übertragung eines Symbols ergeben sich aus dem Produkt der Codewortlänge w_i mit den Kosten pro Bit c_i. Es soll nun ein Code mit möglichst kleinen durchschnittlichen Kosten

$$E[c_i \cdot w_i] = \sum_{i=1}^{5} P(u_i) \cdot c_i \cdot w_i$$

konstruiert werden.

u_i	A	B	C	D	E
$P(u_i)$	1/2	1/6	1/6	1/8	1/24
c_i	1	1	2	2	6

f) Konstruieren Sie für die gegebene Quelle einen binären präfixfreien Code mit minimalen durchschnittlichen Kosten. Wie gross sind die durchschnittlichen Kosten für die Übertragung eines Symbols?

Aufgabe 15

Bei einem fairen Würfel ist eine Fläche mit einem roten Punkt (rt), die anderen fünf Flächen sind mit schwarzen Punkten (sz) gekennzeichnet. Der Würfel wird so lange geworfen, bis die rote Fläche erscheint, jedoch höchstens n = 4 mal. Damit ergeben sich die in der nachfolgenden Tabelle angegebenen Ereignisse.

U	Farbfolge	$P(u_i)$
u_1	rt	
u_2	sz, rt	
u_3	sz, sz, rt	
u_4	sz,sz,sz,rt	
u_5	sz, sz, sz, sz	

a) Berechnen Sie zu jedem der fünf möglichen Ereignisse die dazugehörige Auftretenswahrscheinlichkeit und tragen Sie diese in die Tabelle ein.

b) Berechnen Sie die Entropie dieses Zufallsexperiments.

c) Konstruieren Sie nun einen binären, präfixfreien Code zur Darstellung des Ereignis mit optimaler durchschnittlicher Codewortlänge und geben Sie ihn in Tabellenform an.

d) Wie gross ist die durchschnittliche Codewortlänge des unter c) konstruierten Codes?

e) Die Ereignisse sollen mit dem unter c) konstruierten Code über einen bandbegrenzten AWGN-Kanal mit der Bandbreite B = 3100 Hz übertragen werden. Wie gross muss das Signal-zu-Rauschverhältnis in Dezibel mindestens sein, damit $R = 10^4$ Ereignisse pro Sekunde mit beliebig kleiner Fehlerwahrscheinlichkeit übertragen werden können?

f) Durch welche Massnahme kann die durchschnittliche Codewortlänge pro Ereignis noch verbessert werden? Um wie viele Dezibel kann dadurch die Anforderung an das Signal-zu-Rauschverhältnis gemildert werden?

Das Experiment wird nun derart modifiziert, dass immer das Auftreten der roten Fläche abgewartet wird, also $n \to \infty$.

g) Geben Sie einen allgemeinen Ausdruck für die Wahrscheinlichkeit $P(u_i)$ an, dass die rote Fläche im i-ten Wurf erscheint.

h) Berechnen Sie nun die Entropie des Zufallsexperiments.

Tipps:

$$\sum_{i=1}^{\infty} \alpha^{i-1} = \frac{1}{1-\alpha}$$

$$\sum_{i=1}^{\infty} (i-1) \cdot \alpha^{i-1} = \frac{\alpha}{(1-\alpha)^2}$$

Aufgabe 16

Eine Quelle gibt zufällig zwei Symbole X und Y aus. Das Symbol X stammt aus der Menge {a, b, c, d}, das Symbol Y aus der Menge {α, β, γ}. Die Wahrscheinlichkeit P(X = x, Y = y) für ein beliebiges Ereignis (X = x, Y = y) ist durch die nachfolgende Tabelle gegeben:

	a	b	c	d
α	0.16	0.11	0.01	0.10
β	0.13	0.01	0.00	0.10
γ	0.06	0.04	0.17	0.11

Die Wahrscheinlichkeit für das Ereignis (X = a, Y = α) beträgt also beispielsweise 16%.

a) Vorerst betrachten wir jede Kombination (X = x, Y = y) der beiden Symbole als ein Ergebnis. Das Zufallsexperiment kann also 12 mögliche Ergebnisse liefern (wobei eines, nämlich (c, β), die Wahrscheinlichkeit 0 besitzt). Berechnen Sie nun die Entropie

$$H(X,Y)$$

dieser Quelle.

b) Konstruieren Sie für diese Quelle einen Huffman-Code. Welche durchschnittliche Codewortlänge weist dieser Code auf?

Tipp: Beachten Sie, dass für das Ereignis (c, β) kein Codewort notwendig ist, da es nie auftritt.

c) Wenn wir nur das Symbol Y betrachten, mit welcher Wahrscheinlichkeit $P(Y = \alpha)$ tritt dann der Wert $Y = \alpha$ auf?
Berechnen Sie ebenfalls die Wahrscheinlichkeiten $P(Y = \beta)$ und $P(Y = \gamma)$.

d) Berechnen Sie die Entropie H(Y) der Zufallsvariablen Y.

e) Die Beobachtung des Symbols Y habe die Erkenntnis $Y = \alpha$ geliefert. In diesem Fall ist nur noch das Symbol X zufällig und kann vier mögliche Werte annehmen. Welche Wahrscheinlichkeit $P(X = a \mid Y = \alpha)$ besitzt das Ereignis X = a, falls $Y = \alpha$ gegeben ist?
Berechnen Sie ebenfalls $P(X = b \mid Y = \alpha)$, $P(X = c \mid Y = \alpha)$ und $P(X = b \mid Y = \alpha)$.

f) Unter der Voraussetzung, dass $Y = \alpha$ bekannt ist, ist nur noch das Symbol X ungewiss. Das Zufallsexperiment kann also vier mögliche Ergebnisse liefern, deren Wahrscheinlichkeiten in e) berechnet wurden. Berechnen Sie in diesem Fall die Ungewissheit (Entropie) $H(X \mid Y = \alpha)$ über den Wert des Symbols X.

Aufgabe 17

In der deutschen Sprache kommen die Buchstaben E, N, I, S, R, A und T mit den folgenden Wahrscheinlichkeiten vor:

Buchstabe	Wahrscheinlichkeit
E	0.174
N	0.098
I	0.076
S	0.073
R	0.070
A	0.065
T	0.062

Alle anderen Buchstaben sollen mit dem Symbol Λ dargestellt werden.

a) Mit welcher Wahrscheinlichkeit tritt das Symbol Λ auf?

b) Berechnen Sie die Entropie H dieser Informationsquelle.

c) Bestimmen Sie für diese Quelle einen binären, präfixfreien Code mit minimaler durchschnittlicher Codewortlänge. Geben Sie den Code in Tabellenform an.

d) Wie viele Bits werden durchschnittlich zur Codierung eines Symbols benötigt?

Aufgabe 18

Die Beobachtung einer Nachrichtenquelle hat ergeben, dass diese die drei Symbole A, B und C mit den Wahrscheinlichkeiten P(A) = 4/9, P(B) = 3/9 und P(C) = 2/9 ausgibt.

a) Berechnen Sie die Entropie dieser Quelle.

b) Erstellen Sie für diese Quelle einen Huffman-Code, wenn jeweils zwei aufeinanderfolgende Symbole in ein binäres Codewort abgebildet werden.

c) Wie gross ist die durchschnittliche Anzahl binärer Ziffern pro Symbol?

Die genauere Beobachtung der gleichen Quelle ergibt, dass die Wahrscheinlichkeit für das Symbol X_n vom vorhergehenden Symbol X_{n-1} abhängt. (Die Quelle verfügt also über ein Gedächtnis von einem Symbol). Die entsprechenden Übergangswahrscheinlichkeiten sind in der nachfolgenden Tabelle angegeben:

$P(X_n = A \mid X_{n-1} = A) = 1/2$	$P(X_n = B \mid X_{n-1} = A) = 1/4$	$P(X_n = C \mid X_{n-1} = A) = 1/4$
$P(X_n = A \mid X_{n-1} = B) = 1/3$	$P(X_n = B \mid X_{n-1} = B) = 1/3$	$P(X_n = C \mid X_{n-1} = B) = 1/3$
$P(X_n = A \mid X_{n-1} = C) = 1/2$	$P(X_n = B \mid X_{n-1} = C) = 1/2$	$P(X_n = C \mid X_{n-1} = C) = 0$

d) Geben Sie für die neun möglichen Symbolpaare (AA, AB, ..., CC) jeweils deren Auftretenswahrscheinlichkeit an.

e) Berechnen Sie die Entropie $H(X_n, X_{n-1})$, d.h. die Ungewissheit über zwei aufeinanderfolgende Symbole.

f) Erstellen Sie für diese Quelle mit Gedächtnis einen Huffman-Code, wenn jeweils zwei aufeinanderfolgende Symbole in ein binäres Codewort abgebildet werden.

g) Wie gross ist die durchschnittliche Anzahl binärer Ziffern pro Symbol?

h) Berechnen Sie die Entropie $H(X_n \mid X_{n-1} = A)$ des Symbols X_n, falls bekannt ist, dass das vorhergehende Symbol $X_{n-1} = A$ war.

i) Berechnen Sie die Entropie $H(X_n \mid X_{n-1} = B)$.

j) Berechnen Sie die Entropie $H(X_n \mid X_{n-1} = C)$.

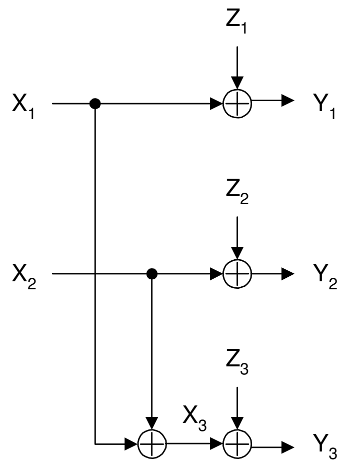
k) Berechnen Sie die sogenannte Entropierate $H(X_n \mid X_{n-1})$.

Tipp: $P_{X,Y}(x, y) = P_X(x) \cdot P_{Y|X}(y \mid x)$

Aufgabe 19

X_1 und X_2 sind voneinander unabhängige, binäre Zufallswerte, welche die Werte 0 und 1 mit der Wahrscheinlichkeit ½ annehmen.

Z_1, Z_2 und Z_3 sind voneinander unabhängige, binäre Zufallswerte, welche den Wert 1 mit der Wahrscheinlichkeit p annehmen.



a) Geben Sie die Wahrscheinlichkeiten für die in der nachfolgenden Tabelle angegebenen Ereignisse an.

Ereignis	**Wahrscheinlichkeit**
$Y_1 = 0$	
$Y_1 = 1$	
$(X_1, X_2) = (0, 0)$	
$(X_1, X_2) = (0, 1)$	
$(X_1, X_2) = (1, 0)$	
$(X_1, X_2) = (1, 1)$	
$Y_3 = 0$	
$Y_3 = 1$	

b) Berechnen Sie die Entropien $H(Y_1)$, $H(Y_3)$, $H(X_1, X_2)$, $H(X_1, X_2, X_3)$ und $H(Y_1, Y_2)$.

c) Berechnen Sie die bedingte Entropie $H(Y_1|X_1)$. Geben Sie das Resultat in Funktion der binären Entropiefunktion h(p) an.

d) Berechnen Sie die bedingte Entropie $H(Y_1, Y_2| X_1, X_2)$. Geben Sie das Resultat in Funktion der binären Entropiefunktion h(p) an.

e) Berechnen Sie die Transinformation $I(X_1,X_2;Y_1,Y_2)$.

Aufgabe 20

Am Ausgang einer Informationsquelle wird das Wort „MISSISSIPPISHIP“ beobachtet. Wir gehen davon aus, dass diese Beobachtung für die Quelle typisch ist, d.h. die relative Häufigkeit der Buchstaben entspricht der Wahrscheinlichkeit.

a) Geben Sie an, mit welcher Wahrscheinlichkeit (eigentlich relativer Häufigkeit) jeder Buchstabe auftritt.

b) Konstruieren Sie einen binären, präfixfreien Code, mit dem das Wort „MISSISSIPPISHIP“ möglichst effizient codiert werden kann.

c) Wie viele binäre Ziffern werden im Durchschnitt gebraucht, um einen Buchstaben der Quelle zu codieren?

d) Wie gross ist die Entropie der Quelle?

Aufgabe 21

In einem Haufen von $n = 6$ Münzen sind $n - 1 = 5$ Münzen fair. Eine Münze ist so gezinkt, dass sie immer Kopf zeigt. Das Zufallsexperiment läuft wie folgt ab:

1. Eine Münze wird zufällig ausgewählt.
2. Diese Münze wird dann $k = 3$ mal hintereinander geworfen. Dabei wird die Anzahl Würfe, die Kopf zeigen, notiert.

Wir definieren die folgenden Zufallsvariablen:

$$X = \begin{cases} 0 & \text{Die gewählte Münze ist fair} \\ 1 & \text{Die gewählte Münze ist gezinkt} \end{cases}$$

Y: Anzahl Würfe, die Kopf zeigen.

a) Wie gross sind die Wahrscheinlichkeiten $P(X = 0)$ und $P(X = 1)$?

b) Wie gross ist die Ungewissheit $H(X)$ darüber, ob die gewählte Münze fair ist?

c) Berechnen Sie für $y = 0, 1, 2, 3$ die Wahrscheinlichkeiten $P(Y = y | X = 0)$ und $P(Y = y | X = 1)$.

d) Wie gross sind die Wahrscheinlichkeiten $P(Y = 3)$ und $P(X = 0 | Y = 3)$?

e) Wie gross ist die Entropie $H(X | Y)$?

f) Wie gross ist die Transinformation $I(X;Y)$?

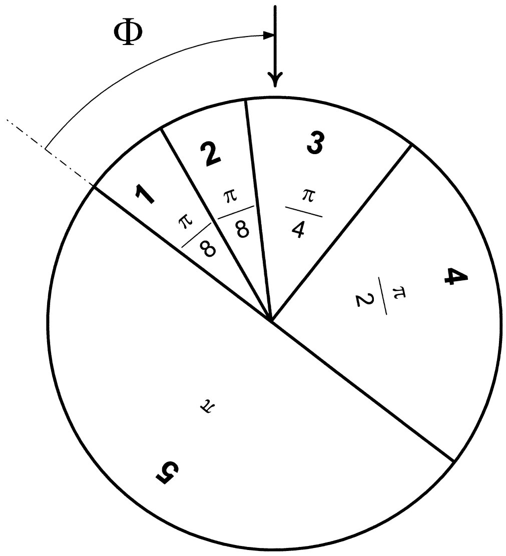
Tipps:

$$P(Y = y) = P(X = 0) \cdot P(Y = y \mid X = 0) + P(X = 1) \cdot P(Y = y \mid X = 1)$$

$$P(X = x \mid Y = y) = \frac{P(X = x) \cdot P(Y = y \mid X = x)}{P(Y = y)}$$

Aufgabe 22

Das folgende Glücksrad wird verwendet, um zufällig die Zahlen 1, 2, 3, 4, und 5 zu wählen. Der Winkel Φ gibt an, in welcher Position das Glücksrad stehen geblieben ist.



Es darf zunächst angenommen werden, dass das Glücksrad ideal arbeitet, d.h. jede Stellung (jedes Φ) ist gleich wahrscheinlich.

a) Konstruieren Sie einen Huffman-Code um das Ergebnis des Glücksrades zu codieren und notieren Sie für jedes Feld das Codewort.

b) Wie gross ist die durchschnittliche Codewortlänge?

Durch einen Konstruktionsfehler ist das Glücksrad nicht genau symmetrisch, weshalb nicht mehr alle Winkel Φ gleich wahrscheinlich sind. Die Wahrscheinlichkeitsdichte ist vielmehr durch die Formel

$$p_\Phi(\varphi) = \begin{cases} \frac{1}{2\cdot\pi}\cdot\left(1+\cos(\varphi)\right) & 0 \le \varphi < 2\cdot\pi \\ 0 & \text{sonst} \end{cases}$$

gegeben.

c) Bestimmen Sie für jedes Feld die Wahrscheinlichkeit, dass es getroffen wird.

d) Konstruieren Sie wiederum einen Huffman Code und notieren Sie die Codeworte.

e) Es werden jeweils 1000 Ziehungen durchgeführt. Das Resultat aller 1000 Ziehungen soll mit möglichst wenigen Bits in einem Codewort codiert werden. Wie gross ist die durchschnittliche Anzahl Bits, die Sie dafür bei optimaler Codierung übertragen müssen, höchstens?

Aufgabe 23

Die 2006 möglichen Symbole einer gedächtnislosen Quelle erscheinen alle mit der gleichen Wahrscheinlichkeit am Ausgang.

a) Gibt es für diese Quelle einen binären, präfixfreien Code mit einer mittleren Codewortlänge von E[W] < 10.5?

b) Gibt es für diese Quelle einen binären, präfixfreien Code mit 42 Codeworten der Länge 10 und 1964 Codeworten der Länge 11?

Aufgabe 24

Für ein einfaches Musikprogramm sollen Noten möglichst platzsparend abgelegt werden. Eine Melodie kann hier aus den Tonhöhen C, D, E, F, G, A, H bestehen, die jeweils als ganze, halbe, viertel oder achtel Noten gespielt werden.

Statistiker haben errechnet, dass jede Tonhöhe/Tondauer eine bestimmte mittlere Häufigkeit in der Musik besitzt:

Tonhöhe						
C	D	E	F	G	A	H
25.9%	14.4	23.3%	11.5%	16.0%	7.1%	1.8%

Tondauer			
1/1	1/2	1/4	1/8
5%	20%	40%	35%

Man kann davon ausgehen, dass Tonhöhe und Tondauer statistisch unabhängig sind!
Musikstücke sollen nach Huffman codiert abgelegt werden. Hierfür gibt es zwei Alternativen:

- Tonhöhe und Tondauer werden getrennt abgelegt. Dafür werden zwei getrennte Huffman-Bäume erzeugt.
- Für jede Kombination aus Tonhöhe und Tondauer (=Note) wird eine Wahrscheinlichkeit bestimmt. Es wird ein Huffman-Baum für die resultierenden 28 Noten erzeugt.

a) Berechnen Sie die Entropie der Tonhöhe, der Tondauer und der Note.

b) Stellen Sie die Huffman-Bäume für die erste Möglichkeit (Tonhöhe und –dauer getrennt codiert) auf. Wie viele binäre Ziffern müssen im Durchschnitt pro Note übertragen werden?

c) Welche Aussage können Sie, ohne einen Code zu konstruieren, über die durchschnittliche Codewortlänge machen, falls die Musikstücke nach der zweiten Methode (Tonhöhe und –dauer gemeinsam codiert) codiert werden?

Kurzfragen

1. Wir betrachten eine Nachrichtenquelle ohne Gedächtnis. Das Ausgangssymbol wird aus einer Menge von endlich vielen Symbolen ausgewählt.
   a) Wovon hängt die Entropie der Nachrichtenquelle ab?
   b) Unter welchen Voraussetzungen ist die Entropie der Nachrichtenquelle gleich Null?
   c) Unter welchen Voraussetzungen ist die Entropie der Nachrichtenquelle maximal?
   d) Was ist der Unterschied zwischen Entropie und Information?
2. Die Ausgangssymbole einer Nachrichtenquelle stammen aus einer Menge mit sieben Elementen, wobei das Symbol x_i mit der Wahrscheinlichkeit $P(x_i)$ ausgegeben wird. Gegeben sind vier Quellencodes.

X	$P(X=x_i)$	Code 1	Code 2	Code 3	Code 4
x_1	0.4	0	0	00	00
x_2	0.2	100	100	01	010
x_3	0.1	101	101	100	011
x_4	0.1	110	110	101	100
x_5	0.1	1110	1110	110	101
x_6	0.05	11110	1101	1110	110
x_7	0.05	11111	1111	1111	111

   a) Welche Codes sind präfixfrei?
   b) Welcher der präfixfreien Codes hat die kürzeste durchschnittliche Codewortlänge?
3. Wieso ist es nicht sinnvoll, ein verschlüsseltes Signal mit einem Quellencode komprimieren zu wollen?