Elementare Wahrscheinlichkeit
Lösung der Prüfung vom 18. 12. 2013

1. Entropie und Kodierung

1.1 Die Abbildungen φ_i sind wie folgt definiert

	1	2	3	4	5	6	7	8
φ_1	101	11	0001	1100	1101	0010	01	0011
φ_2	0	100	101	110	11100	11101	11110	11111
φ_3	10	01	000	001	110	1110	1101	1111

a) Nur φ_2 ist ein Präfixkode: Es gilt $\varphi_1(4) = \varphi_1(2)00$ und $\varphi_3(7) = \varphi_3(5)1$. Dass φ_2 ein Präfixkode ist, sieht man wenn man die zugehörigen Knoten im binären Baum markiert. Keiner dieser Knoten liegt auf dem Weg von der Wurzel zu einem anderen solche Knoten. (je 0.5 Punkte für jede richtige Antwort).

b) Gemäss Lemma 1 existiert ein Präfixkode mit Längen $\ell(i)$ genau dann, wenn

$$\sum_{i=1}^{8} 2^{-\ell(i)} \leq 1.$$

Für φ_1 ist die Summe links gleich

$$\frac{1}{8} + \frac{1}{4} + \frac{4}{16} + \frac{1}{4} + \frac{1}{16} = \frac{15}{16},$$

also existiert ein Präfix-Kode φ' mit gleichen Kodelängen wie φ_i. Markiert man die zu φ_1 gehörigen Knoten im binären Baum, dann sieht man dass es genügt $\varphi'(4) = 1000$ und $\varphi'(5) = 1001$ zu setzen und $\varphi'(i) = \varphi_1(i)$ für die anderen i's. Für φ_3 ist die Summe links gleich

$$\frac{2}{4} + \frac{3}{8} + \frac{3}{16} = \frac{17}{16}.$$

Also kann es keinen solchen Präfix-Kode φ' geben. (Idee mit Kraft-Ungleichung: 0.5 Punkte. Umsetzung je 0.5 Punkte).

1.2 Die Entropie einer Wahrscheinlichkeit $\mathbb{P}$ ist definiert als

$$H(\mathbb{P}) = \sum_{i=1}^{n} p(i) \cdot - \log_2 p(i).$$

Die erwartete Länge eines beliebigen Präfixkodes unter der Verteilung $\mathbb{P}$ ist mindestens gleich $H(\mathbb{P})$. Die erwartete Länge des optimalen Kodes ist um weniger als 1 grösser als die Entropie. (Definition Entropie 1 Pkt., jede Ungleichung 0.5 Punkte).

1.3 Der optimale Kode ist der von Huffmann. Dieser Kode geht aus vom optimalen Kode von $n-1$ Symbolen für die Verteilung $\mathbb{P}'$ und unterscheidet dann $n-1$ und n durch ein zuätzliches Bit. Also ist

$$\mathbb{E}(\ell) = \mathbb{E}'(\ell') + p(n-1) + p(n).$$

Nach der Definition der Entropie ist

$$\begin{aligned} H(\mathbb{P}) - H(\mathbb{P}') &= -p(n-1)\log_2 p(n-1) - p(n)\log_2 p(n) + (p(n-1)+p(n))\log_2(p(n)+p(n-1)) \\ &= -(p(n)+p(n-1))\left(\frac{p(n-1)}{p(n)+p(n-1)}\log_2\frac{p(n-1)}{p(n)+p(n-1)} + \right. \\ &\qquad \left. \frac{p(n)}{p(n)+p(n-1)}\log_2\frac{p(n)}{p(n)+p(n-1)}\right). \end{aligned}$$

Also ist die Behauptung

$$\mathbb{E}(\ell) - \mathbb{E}'(\ell') \geq H(\mathbb{P}) - H(\mathbb{P}')$$

äquivalent dazu, dass für alle $x \in (0,1)$

$$-x\log_2(x) - (1-x)\log_2(1-x) \leq 1.$$

Das gilt gemäss Lemma 2, mit Gleichheit genau dann, wenn $x = p(n-1)/(p(n)+p(n-1)) = 1/2$. Statt Lemma 2 zu bemühen, kann man das auch leicht selber nachprüfen. Alternativ kann man die Formel

$$H(\mathbb{P}) - H(\mathbb{P}') = p(n-1)\log_2\left(\frac{p(n-1)+p(n)}{p(n-1)}\right) + p(n)\log_2\left(\frac{p(n-1)+p(n)}{p(n)}\right)$$

zusammen mit der Ungleichung $\log(x) \leq \log(2) + \frac{1}{2}(x-2)$ verwenden. (Je 1 Punkt für $\mathbb{E}(\ell) - \mathbb{E}'(\ell')$ und $H(\mathbb{P}) - H(\mathbb{P}')$, 1.5 Punkte für eine richtige Ungleichung, 0.5 Punkte für $p(n) = p(n-1)$ als Bedingung für Gleichheit.)

Der Shannon Kode mit $\ell(i) = \lceil -\log_2\ p(i)\rceil$ hilft hier nicht weiter, weil er im Allgemeinen nicht optimal ist und weil man nur

$$\mathbb{E}(\ell) - \mathbb{E}'(\ell') > H(\mathbb{P}) - H(\mathbb{P}') - 1$$

erhält. Er liefert nur die hinreichende, aber nicht notwendige Bedingung $-\log_2 p(i) \in \mathbb{N}$ für Gleichheit. Dafür gibt es 0.5 Punkte.

2. Mischen von Karten

2.1 Weil es $(n-1)!$ Permutationen gibt mit $\omega(1) = 1$, gilt für solche Permutationen

$$q(\omega) = \frac{1}{(n-1)!} > p(\omega) = \frac{1}{n!}.$$

Für alle anderen Permutationen ist $q(\omega) = 0$. Also folgt

$$\|P - Q\| = \sum_{\omega; q(\omega) > p(\omega)} (q(\omega) - p(\omega)) = \sum_{\omega;\omega(1)=1}\left(\frac{1}{(n-1)!} - \frac{1}{n!}\right) = (n-1)!\frac{n-1}{n!} = 1 - \frac{1}{n}.$$

(0.5 Punkte für Definition des Abstandes, 0.5 Punkte für die richtige Angabe von1 $q(\omega)$ inklusive der ω's mit $q(\omega) = 0$, 1 Punkt für die Berechnung, bzw 0.5 Punkte falls ein Rechenfehler).

2.2 a) ω_1 hat 4 runs, nämlich 35, 278, 6 und 14. ω_2 hat 2 runs, nämlich 347 und 12568, und ω_3 hat 6 runs, nämlich 4, 13, 28, 7, 6 und 5. (1 Punkt falls alle 3 Antworten richtig, 0.5 Punkte falls 1 Fehler, sonst 0 Punkte).

b) Die Antwort folgt aus Satz 4:

$$p_1(\omega_2) = 2^{-8},\ p_1(\omega_1) = p_1(\omega_3) = 0$$

und

$$p_2(\omega_1) = 2^{-16},\ p_2(\omega_2) = 2^{-16}\binom{10}{8},\ p_2(\omega_3) = 0.$$

(0.5 Punkte für die Fälle mit Wahrscheinlichkeit Null, je 0.5 Punkte für die 3 anderen Fälle).

2.3 Ein riffle shuffle wird beschrieben durch $k =$ Anzahl Karten im linken Teilstapel und die Teilmenge L der k Plätze, den die Karten des linken Teilstapels nachher einnehmen. Jede Kombination (k, L) hat Wahrscheinlichkeit 2^{-n}. Die oberste Karte bleibt zuoberst genau dann, wenn entweder $k = 0$ oder wenn $k > 0$ und $1 \in L$. Da es $\binom{n-1}{k-1}$ Mengen mit $k \geq 1$ Elementen gibt, die 1 enthalten, ist die gesuchte Wahrscheinlichkeit gleich

$$2^{-n}\left(1 + \sum_{k=1}^{n}\binom{n-1}{k-1}\right) = 2^{-n}(1 + 2^{n-1}) = \frac{1}{2} + 2^{-n}.$$

(1 Punkt, wenn das Ereignis mit den Grössen k und L ausgedrückt wird, 1 Punkt für die Durchführung, 0.5 Punkte Abzug bei Rechenfehler. Wenn das Resultat $\frac{1}{2}$ ohne richtige Begründung gegeben wird, 0.5 Punkte.).

3. Markovketten

3.1 Auf Grund der Definition von Π ist sicher $\Pi(i,j) > 0$ für alle $i \neq j$ und $\sum_{j=1}^{n} \Pi(i,j) = 1$ für alle i. Π ist also eine stochastische Matrix, falls $\Pi(i,i) \geq 0$ für alle i. Dies gilt, weil für $i \neq j$ $\Pi(i,j) \leq \Lambda(i,j)$ und daher

$$\sum_{j;j\neq i} \Pi(i,j) \leq \sum_{j;j\neq i} \Lambda(i,j) \leq 1,$$

denn nach Voraussetzung ist Λ eine stochastische Matrix. (je 0.5 Punkte für die offensichtlichen Eigenschaften, 1 Punkt für $\Pi(i,i) \geq 0$).

Die Bedingung für die Reversibilität lautet $\gamma(i)\Pi(i,j) = \gamma(j)\Pi(j,i)$, also

$$\gamma(i)\Lambda(i,j)\min\left(1, \frac{\gamma(j)\Lambda(j,i)}{\gamma(i)\Lambda(i,j)}\right) = \gamma(j)\Lambda(j,i)\min\left(1, \frac{\gamma(i)\Lambda(i,j)}{\gamma(j)\Lambda(j,i)}\right).$$

Wir nehmen o.B.d.A. an, dass $\gamma(j)\Lambda(j,i) \leq \gamma(i)\Lambda(i,j)$. Dann ist

$$\min\left(1, \frac{\gamma(j)\Lambda(j,i)}{\gamma(i)\Lambda(i,j)}\right) = \frac{\gamma(j)\Lambda(j,i)}{\gamma(i)\Lambda(i,j)},\ \min\left(1, \frac{\gamma(i)\Lambda(i,j)}{\gamma(j)\Lambda(j,i)}\right) = 1,$$

und damit ist die Bedingung für Reversibilität erfüllt. (1 Punkt für die Fallunterscheidung, 1 Punkt für die Durchführung).

3.2

a) Man erhält

$$\Pi = \begin{pmatrix}
0 & 0.5 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.5 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0.5 & 0 \\
0 & 0 & 0 & 0 & 0.5 & 0 & 0.5 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0.5 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0.5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0.5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.$$

(1 Punkt)

b) Die Zustände 10 und 11 werden nicht mehr verlassen, also sind sie rekurrent, und mit keinen anderen Zuständen verbunden. Von einem Zustand $i < 10$ kann man nur in Zustände $j > i$ gelangen, also sind diese Zustände alle transient und ebenfalls nicht verbunden mit anderen Zuständen. (1 Punkt für die Äquivalenzklassen, je 0.5 Punkte für transient, bzw. rekurrent).

c) Sei u_i die Wahrscheinlichkeit, von i aus in den Zustand 10 zu gelangen, und v_i die Wahrscheinlichkeit, von i aus in den Zustand 11 zu gelangen. Gemäss Satz 5 lautet das Gleichungssystem für u_i und v_i

$$\begin{pmatrix} 1 & -0.5 & 0 & -0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -0.5 & 0 & -0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -0.5 & 0 & -0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -0.5 & 0 & -0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \\ u_3 & v_3 \\ u_4 & v_4 \\ u_5 & v_5 \\ u_6 & v_6 \\ u_7 & v_7 \\ u_8 & v_8 \\ u_9 & v_9 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0.5 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0.5 & 0 \\ 0 & 0.5 \\ 0 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$$

(Gleichung $(I - Q)U = R$ 1 Punkt, richtige Angabe von Q, U und R 1 Punkt).