Lösung zu Aufgabe 2

Die Informationsquelle aus Übung 1 liefert statistisch unabhängige Zeichen. Die Wahrscheinlichkeit für die Sequenz von zwei Symbolen x_i, x_j ist daher

$$p(x_i, x_j) = p(x_i) \cdot p(x_j) \ . \tag{1}$$

In diesem Fall ist die Auftrittswahrscheinlichkeit eines Zeichens x_i unter der Bedingung, dass das letzte Zeichen x_j bekannt ist, einfach die Auftrittswahrscheinlichkeit für x_i

$$p(x_i|x_j) = p(x_i) \ . \tag{2}$$

Nun wird die Quelle realistischer beschrieben. Sie liefert statistisch abhängige Symbole, d.h., die Wahrscheinlichkeit eines aktuellen Symbols x_i ist abhängig von den vorherigen Zuständen $\{x_j,\, x_k,\, x_l,\, \ldots\}$ und deren Auftrittswahrscheinlichkeiten. Allgemein gilt

$$p(x_i) = \underbrace{p\big(\, x_i \,\big|\, \{x_j,\, x_k,\, x_l,\, \ldots\}\big)}_{\text{Übergangswahrscheinlichkeit}} \cdot p\big(\, \{x_j,\, x_k,\, x_l,\, \ldots\}\,\big) \ . \tag{3}$$

In der Praxis (sowie in diesem Beispiel) können Quellen häufig durch eine Markoff-Quelle erster Ordnung modelliert werden. Das bedeutet, dass nur das Wissen des letzten Zeichens notwendig ist, um die Übergangswahrscheinlichkeit bestimmen zu können. Demnach hängt die Wahrscheinlichkeit des aktuellen Symbols x_i nur vom letzten Symbol x_j und dessen Auftrittswahrscheinlichkeit ab. Es gilt:

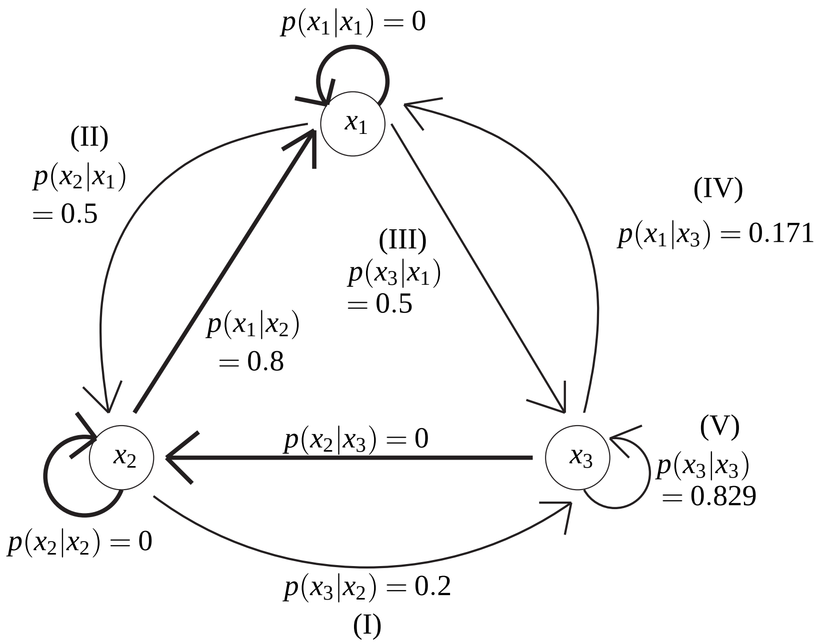
$$p\big(\, x_i \,\big|\, \{x_j,\, x_k,\, x_l,\, \ldots\}\big) = p(x_i|x_j) \ .$$

Folgende allgemeine Zusammenhänge müssen bekannt sein, um die Aufgabe lösen zu können:

$$p(x_i) = \sum_{j=1}^{N} p(x_i, x_j) = \sum_{j=1}^{N} p(x_i|x_j) \cdot p(x_j) \tag{4}$$

$$\sum_{i=1}^{N} p(x_i|x_j) = 1 \ . \tag{5}$$

2.1 Das Markoff-Diagramm ist unten dargestellt. Die gegebenen Wahrscheinlichkeiten sind fettgedruckt.



Die anderen Wahrscheinlichkeiten wurden folgendermaßen bestimmt:

$$\begin{aligned}
p(x_1|x_2) + p(x_2|x_2) + p(x_3|x_2) &= 1 \\
\Longrightarrow p(x_3|x_2) &= 1 - p(x_1|x_2) = 0.2 && \text{(I)}
\end{aligned}$$

$$\begin{aligned}
p(x_2) &= p(x_2|x_1)\cdot p(x_1) + p(x_2|x_2)\cdot p(x_2) + p(x_2|x_3)\cdot p(x_3) \\
&= p(x_2|x_1)\cdot p(x_1) \\
\Longrightarrow p(x_2|x_1) &= \frac{p(x_2)}{p(x_1)} = \frac{0.1}{0.2} = 0.5 && \text{(II)}
\end{aligned}$$

$$\begin{aligned}
\Longrightarrow p(x_3|x_1) &= 1 - p(x_2|x_1) - p(x_1|x_1) \\
&= 1 - p(x_2|x_1) = 0.5 && \text{(III)}
\end{aligned}$$

$$\begin{aligned}
p(x_1) &= p(x_1|x_2)\cdot p(x_2) + p(x_1|x_3)\cdot p(x_3) \\
\Longrightarrow p(x_1|x_3) &= \frac{p(x_1) - p(x_1|x_2)\cdot p(x_2)}{p(x_3)} = \frac{0.12}{0.7} = \frac{12}{70} \\
&= 0.171 && \text{(IV)}
\end{aligned}$$

$$\begin{aligned}
\Longrightarrow p(x_3|x_3) &= 1 - p(x_1|x_3) - p(x_2|x_3) \\
&= 1 - p(x_1|x_3) = \frac{58}{70} = 0.829 && \text{(V)}
\end{aligned}$$

2.2 Zwei Symbole X_1, X_2 werden zu einem neuen Zeichen Z zusammengefasst, der Ausgang der Quelle ist z.B.:

$$\Big(\underbrace{x_3,x_1,}_{z_n} \quad \underbrace{x_3,x_2,}_{z_{n-1}} \quad \underbrace{x_3,x_2,}_{z_{n-2}} \quad \cdots\Big)$$

$$\begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ \text{aktuelle} & \text{letzte} & \text{vorletzte} \\ \text{Zeichen} & \text{Zeichen} & \text{Zeichen} \end{array}$$

$$H(X,Y) = \sum_{i=1}^{3}\sum_{j=1}^{3} p(x_i,x_j)\cdot \mathrm{ld}\left(\frac{1}{p(x_i,x_j)}\right)$$

with $p(x_i,x_j) = p(x_i|x_j)\cdot p(x_j) \Longrightarrow$

$$\begin{aligned}
p(x_1,x_1) &= 0\cdot 0.2 = 0 \\
p(x_1.x_2) &= 0.8\cdot 0.1 = 0.08 \\
p(x_1,x_3) &= 0.171\cdot 0.7 = 0.12 \\
p(x_2,x_1) &= 0.5\cdot 0.2 = 0.1 \\
p(x_2,x_2) &= 0\cdot 0.1 = 0 \\
p(x_2,x_3) &= 0\cdot 0.7 = 0 \\
p(x_3,x_1) &= 0.5\cdot 0.2 = 0.1 \\
p(x_3,x_2) &= 0.2\cdot 0.1 = 0.02 \\
p(x_3,x_3) &= 0.829\cdot 0.7 = 0.58
\end{aligned}$$

$$\Longrightarrow H(X,Y) = 1.892 \quad < \quad 2\cdot H(X) = 2.314$$

$$\hat{=} H(X,Y)\Big|_{X,Y \text{ statistisch unabhängig}}$$

2.3 Codieren von statistisch abhängigen Symbolen
$\Longrightarrow$ Codieren unter Ausnutzung der Übergangswahrscheinlichkeiten

Definition:

$$\begin{aligned}
x_i &\;\hat{=}\; \text{aktuelle Zeichen der Sequenz} \\
x_j &\;\hat{=}\; \text{letzte Zeichen der Sequenz} \\
x_k &\;\hat{=}\; \text{vorletzte Zeichen der Sequenz} \\
x_l &\;\hat{=}\; \text{Zeichen vor } x_k
\end{aligned}$$

$\Longrightarrow$ Beispiel für die Zeichensequenz $(\underbrace{x_i\,,\,x_j}_{z_n}\,,\underbrace{x_k\,,\,,\,x_l}_{z_{n-1}}\ldots)$

$$\begin{aligned}
\Longrightarrow \quad z_n &= \{x_i,x_j\} \quad \hat{=} \quad \text{aktuelle Zeichenpaar} \\
\Longrightarrow \quad z_{n-1} &= \{x_k,x_l\} \quad \hat{=} \quad \text{vorheriges Zeichenpaar}
\end{aligned}$$

Codieren unter Ausnutzung von:

$$
\begin{aligned}
p(z_n|z_{n-1}) &= p(\{x_i,x_j\}|\{x_k,x_l\}) \\
&= p(\{x_i,x_j\}|x_k) \\
&= \frac{p(x_i,x_j,x_k)}{p(x_k)} \\
&= \frac{p(x_i|\{x_j,x_k\})\cdot p(x_j,x_k)}{p(x_k)} \\
&= p(x_i|\{x_j,x_k\})\cdot p(x_j|x_k) \\
&= p(x_i|x_j)\cdot p(x_j|x_k)
\end{aligned}
$$

Das Codewort für das Zeichenpaar $z_n = \{x_i,x_j\}$ hängt vom Symbol x_k, das vor $\{x_i,x_j\}$ generiert wurde, ab.

x_k	x_j	x_i	$p(\{x_i,x_j\}\|x_k)$	Code	$L(\{x_i,x_j\}\|x_k)/\frac{\text{bit}}{\text{Zeichenpaar}}$
x_1	x_2	x_1	$0.5\cdot 0.8 = 0.4$	00	0.8
x_1	x_2	x_3	$0.5\cdot 0.2 = 0.1$	010	0.3
x_1	x_3	x_1	$0.5\cdot 0.171 = 0.086$	011	0.258
x_1	x_3	x_3	$0.5\cdot 0.829 = 0.415$	1	0.415
					$L_{x_1} = 1.77$
x_2	x_1	x_2	$0.8\cdot 0.5 = 0.4$	1	0.4
x_2	x_1	x_3	$0.8\cdot 0.5 = 0.4$	00	0.8
x_2	x_3	x_1	$0.2\cdot 0.171 = 0.034$	011	0.102
x_2	x_3	x_3	$0.2\cdot 0.829 = 0.166$	010	0.498
					$L_{x_2} = 1.8$
x_3	x_1	x_2	$0.171\cdot 0.5 = 0.086$	101	0.258
x_3	x_1	x_3	$0.171\cdot 0.5 = 0.086$	100	0.258
x_3	x_3	x_1	$0.829\cdot 0.171 = 0.142$	11	0.284
x_3	x_3	x_3	$0.829\cdot 0.829 = 0.687$	0	0.687
					$L_{x_3} = 1.487$

$$
\begin{aligned}
\overline{L}\{x_i,x_j\} &= p(x_1)\cdot L_{x_1} + p(x_2)\cdot L_{x_2} + p(x_3)\cdot L_{x2} \\
&= 0.354 + 0.18 + 1.041 \\
&= 1,575\ \frac{\text{bit}}{\text{Zeichenpaar}} \\
\Longrightarrow \overline{L} &= \frac{\overline{L}\{x_i,x_j\}}{2} = 0.788\ \frac{\text{bit}}{\text{Zeichen}}
\end{aligned}
$$