Aufgabe 1 (AGS 10.11)

Bei einem Spiel werfen Sie zwei unabhängige Münzen und erhalten einen Gewinn, wenn nach dem Wurf beide Münzen *auf der gleichen Seite landen*. Sie können bei diesem Spiel nur beobachten, ob Sie gewonnen oder verloren haben. Nehmen Sie an, dass die erste Münze sehr dick ist und daher beim Werfen auch auf dem Rand R landen kann. Die Menge der möglichen Ergebnisse ist daher $X = \{K, Z, R\} \times \{K, Z\}$. yield (R, Z) = keinGewinn

(a) Geben Sie den Analysator A für dieses Szenario an.

$A(\text{Gewinn}) = \{(K,K), (Z,Z)\}$

$A(\text{keinGewinn}) = \{(Z,K), (K,Z), (R,K), (R,Z)\}$

(b) Sie spielen das Spiel 24 Mal und gewinnen 6 Mal. Geben Sie den Korpus h mit unvollständigen Daten an.

$h(\text{Gewinn}) = 6$ $\quad$ $h(\text{keinGewinn}) = 18$

(c) Gegeben ist die initiale Wahrscheinlichkeitsverteilung $q_0 = q_0^1 \times q_0^2$ über den vollständigen Daten, mit $q_0^1(K) = 2/5$, $q_0^1(R) = 1/5$ und $q_0^2(K) = 1/3$. Dabei ist q_0^1 die Wahrscheinlichkeitsverteilung der ersten und q_0^2 die der zweiten Münze. Führen Sie den E-Schritt des EM-Algorithmus aus, erweitern Sie also den Korpus h zum Korpus h_1 über den vollständigen Daten.

$q_0^1(K) = \frac{2}{5}$ $\quad$ $q_0^2(K) = \frac{1}{3}$

$q_0^1(Z) = \frac{2}{5}$ $\quad$ $q_0^2(Z) = \frac{2}{3}$

$q_0^1(R) = \frac{1}{5}$

$q_0(K,K) = q_0^1(K) \cdot q_0^2(K) = \frac{2}{5} \cdot \frac{1}{3} = \frac{2}{15}$

$q_0(Z,K) = q_0^1(Z) \cdot q_0^2(K) = \frac{2}{5} \cdot \frac{1}{3} = \frac{2}{15}$

$q_0(R,K) = q_0^1(R) \cdot q_0^2(K) = \frac{1}{5} \cdot \frac{1}{3} = \frac{1}{15}$

$q_0(K,Z) = q_0^1(K) \cdot q_0^2(Z) = \frac{2}{5} \cdot \frac{2}{3} = \frac{4}{15}$

$q_0(Z,Z) = q_0^1(Z) \cdot q_0^2(Z) = \frac{2}{5} \cdot \frac{2}{3} = \frac{4}{15}$

$q_0(R,Z) = q_0^1(R) \cdot q_0^2(Z) = \frac{1}{5} \cdot \frac{2}{3} = \frac{2}{15}$

E-Schritt: $h_1(x) = h(yield(x)) \cdot \dfrac{q_0(x)}{\sum\limits_{x' \in A(yield(x))} q_0(x')}$

$$h_1(K,K) = h(Gewinn) \cdot \frac{q_0(K,K)}{\sum\limits_{x' \in \{(K,K),(Z,Z)\}} q_0(x')}$$

$$= h(Gewinn) \cdot \frac{q_0(K,K)}{q_0(K,K) + q_0(Z,Z)}$$

$$= 6 \cdot \frac{\frac{2}{15}}{\frac{2}{15} + \frac{4}{15}}$$

$$= 6 \cdot \frac{2}{6}$$

$$= 2$$

$$h_1(Z,Z) = h(Gewinn) \cdot \frac{q_0(Z,Z)}{q_0(K,K) + q_0(Z,Z)}$$

$$= 6 \cdot \frac{\frac{4}{15}}{\frac{6}{15}} = 4$$

$$h_1(Z,K) = h(kein Gewinn) \cdot \frac{q_0(Z,K)}{q_0(Z,K) + q_0(K,Z) + q_0(R,Z) + q_0(R,K)}$$

$$= 4$$

$h_1(K,Z) = 8$

$h_1(R,K) = 2$

$h_1(R,Z) = 4$

(d) Führen Sie nun den M-Schritt aus. Bestimmen Sie dafür zunächst die Teilkorpora h_1^1 und h_1^2 für die erste bzw. zweite Münze.

$x_1 \backslash x_2$	K	Z	
K	$h_1(K,K)$	$h_1(K,Z)$	$h_1^1(K)$
Z	$h_1(Z,K)$	$h_1(Z,Z)$	$h_1^1(Z)$
R	$h_1(R,K)$	$h_1(R,Z)$	$h_1^1(R)$
	$h_1^2(K)$	$h_1^2(Z)$	

$\rightsquigarrow$

$x_1 \backslash x_2$	K	Z	
K	2	+ 8	10
Z	4	+ 4	8
R	2	+ 4	6
	8	16	

(e) Schätzen Sie nun die Wahrscheinlichkeitsverteilungen q_1^1 und q_1^2 der beiden Münzen, indem Sie die relative Häufigkeit der Teilkorpora bestimmen.

$$\mathrm{rfe}(h)(x) = \frac{h(x)}{|h(x)|}, \quad |h(x)| = \sum_{x \in X} h(x)$$

$$q_1^1(K) = \mathrm{rfe}(h_1^1)(K) = \frac{10}{24}$$
$$q_1^1(Z) = \mathrm{rfe}(h_1^1)(Z) = \frac{8}{24}$$
$$q_1^1(R) = \mathrm{rfe}(h_1^1)(R) = \frac{6}{24}$$

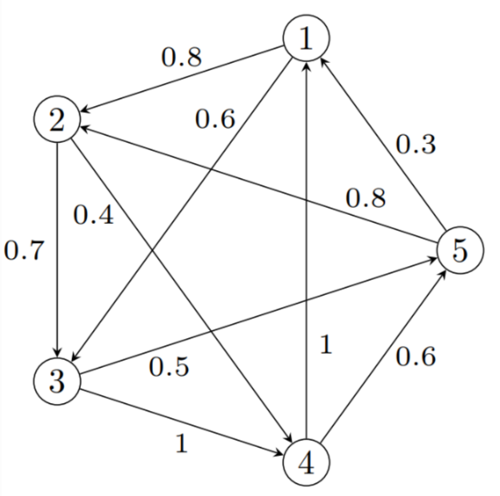
$$q_1^2(K) = \mathrm{rfe}(h_1^2)(K) = \frac{8}{24}$$
$$q_1^2(Z) = \mathrm{rfe}(h_1^2)(Z) = \frac{16}{24}$$

$\max\{0, a\} = a \quad \forall a \in [0,1]$

$a \cdot 1 = a$

Zusatzaufgabe 1 (AGS 9.5.32 ⋆)

Gegeben ist der Viterbi-Semiring $([0,1], \max, \cdot, 0, 1)$ und der gewichtete Graph G über dem Viterbi-Semiring in der nebenstehenden Abbildung.



(a) Geben Sie die modifizierte Adjazenzmatrix von G vollständig an.

$$mA_G = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1 & 0.8 & 0.6 & 0 & 0 \\ 0 & 1 & 0.7 & 0.4 & 0 \\ 0 & 0 & 1 & 1 & 0.5 \\ 1 & 0 & 0 & 1 & 0.6 \\ 0.3 & 0.8 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

(b) Die Matrix $D_G^{(2)}$ ist gegeben. Geben Sie $D_G^{(3)}$ vollständig an!

$$D_G^{(2)} = \begin{pmatrix} 1 & 0.8 & 0.6 & 0.32 & 0 \\ 0 & 1 & 0.7 & 0.4 & 0 \\ 0 & 0 & 1 & 1 & 0.5 \\ 1 & 0.8 & 0.6 & 1 & 0.6 \\ 0.3 & 0.8 & 0.56 & 0.32 & 1 \end{pmatrix}$$

$$D_G^{(3)} = \begin{pmatrix} 1 & 0.8 & 0.6 & 0.6 & 0.3 \\ 0 & 1 & 0.7 & 0.7 & 0.35 \\ 0 & 0 & 1 & 1 & 0.5 \\ 1 & 0.8 & 0.6 & 1 & 0.6 \\ 0.3 & 0.8 & 0.56 & 0.56 & 1 \end{pmatrix}$$

(c) Welche Änderungen ergeben sich in der Berechnung von $D_G^{(5)}$ im Vergleich zu $D_G^{(4)}$?

keine

(d) Gegeben sei das Alphabet $\Sigma = \{a, b, c\}$, der Semiring der formalen Sprachen $(\mathcal{P}(\Sigma^*), \cup, \cdot, \emptyset, \{\varepsilon\})$ über dem Alphabet Σ und die Matrix $D^{(2)}_{G'}$ eines gewichteten Graphen G' über diesem Semiring. Geben Sie die Werte $D^{(3)}_{G'}(u, v)$ für $u \in \{1, 2, 3\}$ und $v \in \{1, 2\}$ an!

$$D^{(2)}_{G'} = \begin{pmatrix} \{\varepsilon\} & \emptyset & \{a\} \\ \emptyset & \{\varepsilon\} & \{d\} \\ \emptyset & \{c\} & \underbrace{\{\varepsilon, b, cd\}} \end{pmatrix}$$

$$D^{(3)}_{G'} = \begin{pmatrix} \{\varepsilon\} & \{a\}\{b,cd\}^*\{c\} & \{a\}\{b,cd\}^* \\ \emptyset & \{\varepsilon\} \cup \{d\}\{b,cd\}^*\{c\} & \{d\}\{b,cd\}^* \\ \emptyset & \{b,cd\}^*\{c\} & \{b,cd\}^* \end{pmatrix}$$

$$\{\varepsilon, b, cd\}^* = \{b, cd\}^*$$

$\varepsilon \in L^*$ für beliebiges L

$$\{\varepsilon, b, cd\}\{\varepsilon, b, cd\}^*\{\varepsilon, b, cd\}$$
$$= \{\varepsilon, b, cd\}^*$$
$$= \{b, cd\}^*$$