2. Übungsblatt (mit Lösungen)

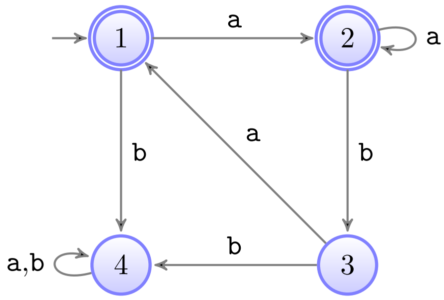
3.0 VU Formale Modellierung

Gernot Salzer, Marion Scholz

3. Dezember 2017

Aufgabe 1 (5 Punkte)

Sei $\mathcal{A}$ der folgende endliche Automat.



(a) Welche der Wörter ε, `aa`, `ab` bzw. `aba` akzeptiert der Automat?

(b) Geben Sie fünf weitere Wörter an, die von $\mathcal{A}$ akzeptiert werden.

(c) Berechnen Sie schrittweise $\delta^*(1, \texttt{ababa})$.

(d) Beschreiben Sie $\mathcal{L}(\mathcal{A})$, die von $\mathcal{A}$ akzeptierte Sprache.

(e) Spezifizieren Sie $\mathcal{A}$ in tabellarischer Form. Handelt es sich bei $\mathcal{A}$ um einen determinstischen oder indeterministischen Automaten?

Lösung

(a) $\mathcal{A}$ akzeptiert ε, `aa` und `aba`, aber nicht `ab`.

(b) $\mathcal{A}$ akzeptiert zum Beispiel `a`, `aaa`, `aaaa`, `aaba` und `abaa`.

(c) $\delta^*(1, \texttt{ababa}) = \delta^*(\delta(1, \texttt{a}), \texttt{baba})$
$= \delta^*(2, \texttt{baba})$
$= \delta^*(\delta(2, \texttt{b}), \texttt{aba})$
$= \delta^*(3, \texttt{aba})$
$= \delta^*(\delta(3, \texttt{a}), \texttt{ba})$
$= \delta^*(1, \texttt{ba})$
$= \delta^*(\delta(1, \texttt{b}), \texttt{a})$
$= \delta^*(4, \texttt{a})$
$= \delta^*(\delta(4, \texttt{a}), \varepsilon)$
$= \delta^*(4, \varepsilon)$
$= 4$

(d) Die Wörter bestehen aus Blöcken der Form $\texttt{a}^+\texttt{ba}$, denen $\texttt{a}^*$ folgt. Es gilt also

$$\mathcal{L}(\mathcal{A}) = (\{\texttt{a}\}^+\{\texttt{ba}\})^*\{\texttt{a}\}^* \ .$$

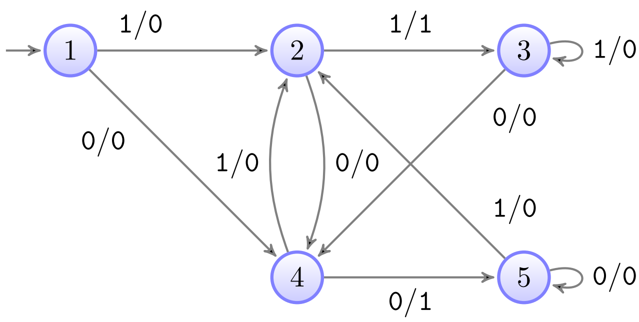
(e) $\mathcal{A} = \langle\{1, 2, 3, 4\}, \{\texttt{a}, \texttt{b}\}, \delta, 1, \{1, 2\}\rangle$, wobei die Übergangsfunktion δ durch folgende Tabelle definiert ist:

δ	a	b
1	2	4
2	2	3
3	1	4
4	4	4

$\mathcal{A}$ ist ein deterministischer Automat, da der momentane Zustand und die nächste Eingabe eindeutig den Folgezustand bestimmen. Das äußert sich in der Tabelle dadurch, dass jeder Eintrag einen einzigen Zustand enthält.

Aufgabe 2 **(3 Punkte)**

Sei $\mathcal{A}$ der folgende Mealy-Automat.



(a) Geben Sie die Ausgabe zur Eingabe 101100111010 an.

(b) Berechnen Sie schrittweise $\delta^*(1, \texttt{00100})$ und $\gamma^*(1, \texttt{00100})$.

(c) Beschreiben Sie die Übersetzungsfunktion $[\mathcal{A}]$.

Lösung

(a) $\gamma^*(1, \texttt{101100111010}) = \texttt{000101010000}$

(b)

$$\begin{aligned}
\delta^*(1, \texttt{00100}) &= \delta^*(\delta(1, \texttt{0}), \texttt{0100}) \\
&= \delta^*(4, \texttt{0100}) \\
&= \delta^*(\delta(4, \texttt{0}), \texttt{100}) \\
&= \delta^*(5, \texttt{100}) \\
&= \delta^*(\delta(5, \texttt{1}), \texttt{00}) \\
&= \delta^*(2, \texttt{00}) \\
&= \delta^*(\delta(2, \texttt{0}), \texttt{0}) \\
&= \delta^*(4, \texttt{0}) \\
&= \delta^*(\delta(4, \texttt{0}), \varepsilon) \\
&= \delta^*(5, \varepsilon) \\
&= 5
\end{aligned}$$

$$\begin{aligned}
\gamma^*(1, \texttt{00100}) &= \gamma(1, \texttt{0}) \cdot \gamma^*(\delta(1, \texttt{0}), \texttt{0100}) \\
&= \texttt{0} \cdot \gamma^*(4, \texttt{0100}) \\
&= \texttt{0} \cdot \gamma(4, \texttt{0}) \cdot \gamma^*(\delta(4, \texttt{0}), \texttt{100}) \\
&= \texttt{0} \cdot \texttt{1} \cdot \gamma^*(5, \texttt{100}) \\
&= \texttt{01} \cdot \gamma(5, \texttt{1}) \cdot \gamma^*(\delta(5, \texttt{1}), \texttt{00}) \\
&= \texttt{01} \cdot \texttt{0} \cdot \gamma^*(2, \texttt{00}) \\
&= \texttt{010} \cdot \gamma(2, \texttt{0}) \cdot \gamma^*(\delta(2, \texttt{0}), \texttt{0}) \\
&= \texttt{010} \cdot \texttt{0} \cdot \gamma^*(4, \texttt{0}) \\
&= \texttt{0100} \cdot \gamma(4, \texttt{0}) \cdot \gamma^*(\delta(4, \texttt{0}), \varepsilon) \\
&= \texttt{0100} \cdot \texttt{1} \cdot \gamma^*(5, \varepsilon) \\
&= \texttt{01001} \cdot \varepsilon = \texttt{01001}
\end{aligned}$$

(c) Der Mealy-Automat erkennt eine Folge von mehr als einem 0er oder 1er. Immer wenn 0 oder 1 mehrmals hintereinander vorkommt, wird einmal das Symbol 1 ausgegeben, sonst 0.

Aufgabe 3 (3 Punkte)

Sei L die Sprache

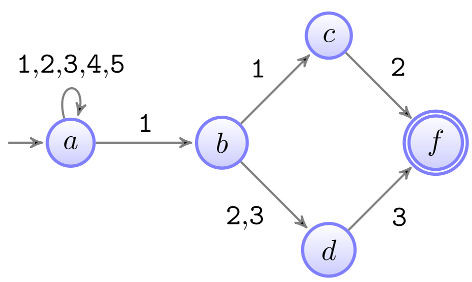
$$\{\, w \in \{\texttt{1}, \texttt{2}, \texttt{3}, \texttt{4}, \texttt{5}\}^* \mid w \text{ endet mit der Ziffernfolge } \texttt{112}, \texttt{123} \text{ oder } \texttt{133} \,\} \enspace .$$

Geben Sie einen *deterministischen* Automaten für L an. Gehen Sie folgendermaßen vor:

1. Konstruieren Sie einen indeterministischen Automaten für diese Sprache.
2. Wandeln Sie den indeterministischen Automaten mit Hilfe des in der Vorlesung besprochenen Verfahrens in einen deterministischen um.

Lösung

Wir konstruieren zunächst auf möglichst direktem Weg einen beliebigen Automaten für die gesuchte Sprache. Dieser ist im Zustand a indeterministisch: Für das Symbol 1 gibt es zwei mögliche Folgezustände. Zusätzlich zur graphischen Darstellung geben wir die Übergangsfunktion auch als Tabelle an, da diese bei der Determinisierung hilft.



δ^*	1	2	3	4	5
a	$\{a,b\}$	$\{a\}$	$\{a\}$	$\{a\}$	$\{a\}$
b	$\{c\}$	$\{d\}$	$\{d\}$	$\{\}$	$\{\}$
c	$\{\}$	$\{f\}$	$\{\}$	$\{\}$	$\{\}$
d	$\{\}$	$\{\}$	$\{f\}$	$\{\}$	$\{\}$
f	$\{\}$	$\{\}$	$\{\}$	$\{\}$	$\{\}$

Einen deterministischen Automaten erhalten wir, indem wir den indeterminischen Automaten simulieren. Ein Zustand des deterministischen Automaten repräsentiert dabei jene Zustände des indeterministischen, in denen sich dieser zu diesem Zeitpunkt befinden kann. Der Startzustand wird mit $\{a\}$ bezeichnet, da sich der indeterministische Automat zu Beginn im Zustand a (und nur in diesem) befindet. Von diesem Zustand ausgehend erstellen wir zeilenweise die Tabelle für die Übergangsfunktion des deterministischen Automaten.

δ^*	1	2	3	4	5
$\{a\}$	$\{a,b\}$	$\{a\}$	$\{a\}$	$\{a\}$	$\{a\}$
$\{a,b\}$	$\{a,b,c\}$	$\{a,d\}$	$\{a,d\}$	$\{a\}$	$\{a\}$
$\{a,d\}$	$\{a,b\}$	$\{a\}$	$\{a,f\}$	$\{a\}$	$\{a\}$
$\{a,f\}$	$\{a,b\}$	$\{a\}$	$\{a\}$	$\{a\}$	$\{a\}$
$\{a,b,c\}$	$\{a,b,c\}$	$\{a,d,f\}$	$\{a,d\}$	$\{a\}$	$\{a\}$
$\{a,d,f\}$	$\{a,b\}$	$\{a\}$	$\{a,f\}$	$\{a\}$	$\{a\}$

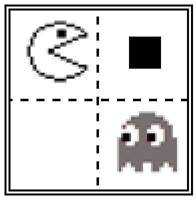
Jene Zustände, die einer Situation entsprechen, in der der indeterministische Automat einen Endzustand erreicht hat, sind die Endzustände des deterministischen Automaten; in diesem Beispiel sind das alle Zustände, deren Bezeichnung f enthält. Dieser wird somit durch das Tupel $\langle \widehat{Q}, \Sigma, \hat{\delta}, \{a\}, \widehat{F} \rangle$ beschrieben, wobei

$$\begin{aligned}
\Sigma &= \{1,2,3,4,5\} \\
\widehat{F} &= \{\{a,f\},\{a,d,f\}\} \\
\widehat{Q} &= \{\{a\},\{a,b\},\{a,d\},\{a,b,c\}\} \cup \widehat{F}
\end{aligned}$$

Aufgabe 4 (4 Punkte)

Bei Pac-Man, einem Videospiel aus den 80ern, muss die Spielfigur Pac-Man Punkte in einem Labyrinth fressen, während sie von Geistern verfolgt wird. Pac-Man und jeder Geist kann pro Zug ein Feld nach oben, unten, links oder rechts bewegt werden, vorausgesetzt, es ist keine Mauer im Weg. Bewegt Pac-Man sich auf ein Feld mit einem Punkt, so frisst er diesen Punkt. Sind alle Punkte gefressen, gilt das Level als gewonnen. Befinden sich Pac-Man und ein Geist auf dem gleichen Feld, so wird Pac-Man vom Geist gefressen und das Spiel ist beendet. Befinden sich der letzte Punkt, Pac-Man und ein Geist auf demselben Feld, hat der Geist Vorrang, d.h., Pac-Man verliert auch in diesem Fall.

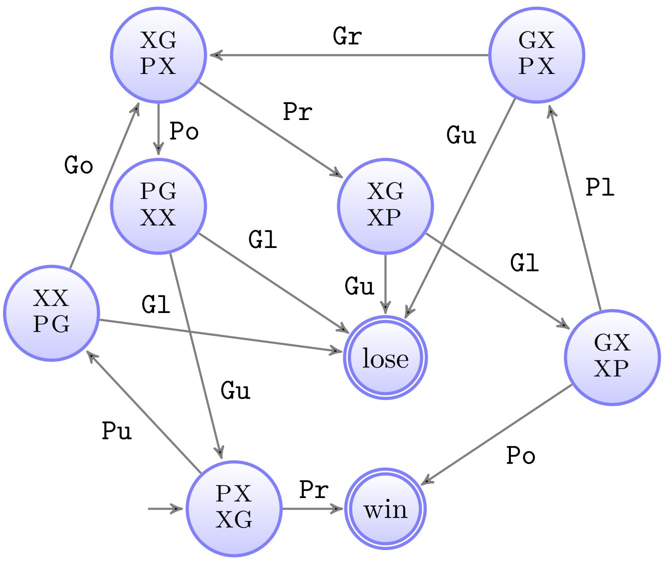
Nehmen Sie an, dass das aktuelle Level, das Pac-Man bewältigen muss, so aussieht wie rechts skizziert. Zu Beginn befindet sich Pac-Man im linken oberen Feld, im rechten unteren Feld ist ein Geist und im rechten oberen Feld ist ein Punkt (schwarzes Quadrat). Rundherum ist eine Mauer (doppelte Linie). Pac-Man gewinnt das Level, wenn er den einen Punkt frisst. Pac-Man verliert, falls er von dem Geist gefressen wird.



- Überlegen Sie, welche Informationen notwendig sind, um den Zustand des Systems zu beschreiben.
- Legen Sie die möglichen Aktionen fest, die zu einem Zustandswechsel führen können.
- Geben Sie einen endlichen Automaten an, der das Systemverhalten vollständig beschreibt. Gehen Sie dabei davon aus, dass Pac-Man und der Geist abwechselnd je einen Zug machen, wobei Pac-Man beginnt.

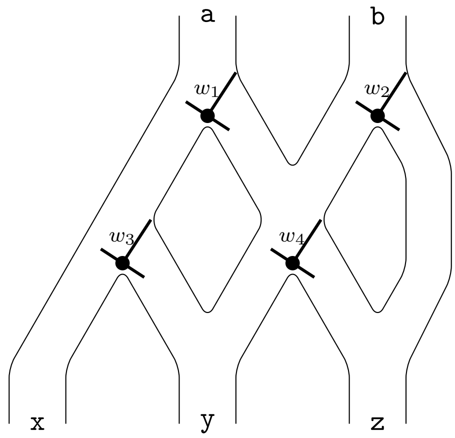
Lösung

Der Systemzustand ist hinreichend festgelegt, wenn bekannt ist, wo Pac-Man und der Geist sich gerade befinden. Wir verwenden P bzw. G um anzugeben, ob sich auf dem jeweiligen Feld gerade Pac-Man (P) oder der Geist (G) befindet. Die Aktionen sind mit jeweils zwei Buchstaben gekennzeichnet, die angeben, wer sich ein Feld in welche Richtung bewegt. Dabei stehen `r`, `l`, `o` und `u` für rechts, links, oben und unten. Bewegt Pac-Man sich auf das Feld rechts oben, also dort wo der einzige Punkt eingezeichnet ist, gewinnt er – vorausgesetzt dort befindet sich zu diesem Zeitpunkt nicht der Geist.



Aufgabe 5 (4 Punkte)

Max erhält zum Geburtstag ein Knobelspiel, das aus einem Würfel und einer Stahlkugel besteht. Der Würfel besitzt oben zwei Löcher (bezeichnet mit `a` und `b`) und unten drei (bezeichnet mit `x`, `y` und `z`). Wirft man die Kugel bei einem der beiden oberen Löcher hinein, kommt sie bei einem der unteren drei wieder heraus. Die Aufgabe besteht nun darin das Loch vorherzusagen, bei dem die Kugel erscheinen wird. Um die Sache schwieriger zu gestalten, sind die Eingänge nicht fest mit den Ausgängen verbunden, sondern werden durch Weichen umgeleitet, die sich mit jeder vorbeikommenden Kugel verstellen. Nach einiger Zeit verliert Max die Geduld und zerlegt den Würfel. Er findet vier Weichen (w_1 bis w_4) vor, die folgendermaßen angeordnet sind:



Bei der momentanen Weichenstellung werden die Kugeln nach links geleitet, sodass die Kugel vom Eingang `a` zum Ausgang `x` bzw. vom Eingang `b` zum Ausgang `y` rollen würde. Dabei würden die Weichen w_1 und w_3 bzw. w_2 und w_4 umgeschaltet werden. Wirft man etwa die Kugel viermal in den Eingang `a`, kommt sie zuerst bei Ausgang `x`, dann zweimal bei Ausgang `y` und schließlich bei Ausgang `z` zum Vorschein; die Weichen befinden sich danach wieder in der Ausgangsstellung.
Modellieren Sie das Verhalten dieses Spiels mit Hilfe eines Moore- oder Mealy-Automaten. Bei Eingabe eines Wortes über $\{\mathtt{a}, \mathtt{b}\}$ soll der Automat jene Ausgänge liefern, bei denen die Kugel erscheinen würde. Sie können die Übergangs- und Ausgabefunktion graphisch oder als Tabelle darstellen.
Berechnen Sie $\gamma^*(q_0, \mathtt{aabaabb})$, wobei γ die Ausgabefunktion und q_0 den Anfangszustand Ihres Automaten bezeichnet.

Lösung

Der Zustand des Würfels wird durch die Stellung der vier Weichen festgelegt. Wir bezeichnen die Zustände mit $w_1w_2w_3w_4$, wobei $w_i = 0$ bedeutet, dass die Weiche rechts steht (wie in der Skizze eingezeichnet), und $w_i = 1$, dass sich die Weiche in der linken Stellung befindet. Das Spiel wird durch den Mealy-Automaten

$$\langle \{0000, 0001, \ldots, 1111\}, \{\mathtt{a}, \mathtt{b}\}, \{\mathtt{x}, \mathtt{y}, \mathtt{z}\}, \delta, \gamma, 0000 \rangle$$

beschrieben, wobei die Übergangsfunktion δ und die Ausgabefunktion γ durch folgende Tabelle definiert sind.

	δ		γ	
	`a`	`b`	`a`	`b`
0000	1010	0101	`x`	`y`
0001	1011	0100	`x`	`z`
0010	1000	0111	`y`	`y`
0011	1001	0110	`y`	`z`
0100	1110	0000	`x`	`z`
0101	1111	0001	`x`	`z`
0110	1100	0010	`y`	`z`
0111	1101	0011	`y`	`z`
1000	0001	1101	`y`	`y`
1001	0000	1100	`z`	`z`
1010	0011	1111	`y`	`y`
1011	0010	1110	`z`	`z`
1100	0101	1000	`y`	`z`
1101	0100	1001	`z`	`z`
1110	0111	1010	`y`	`z`
1111	0110	1011	`z`	`z`

Für die Eingabe `aabaabb` erhalten wir $\gamma^*(0000, \texttt{aabaabb}) = \texttt{xyzyyzz}$.

Aufgabe 6 (3 Punkte)

Seien $\mathcal{A}_1 = \langle Q_1, \Sigma, \delta_1, i_1, F_1\rangle$ und $\mathcal{A}_2 = \langle Q_2, \Sigma, \delta_2, i_2, F_2\rangle$ zwei beliebige deterministische Automaten über demselben Alphabet Σ. Geben Sie eine Methode an, um daraus einen Automaten $\mathcal{A}$ für den Durchschnitt der beiden zugehörigen Sprachen zu konstruieren, d.h., es soll $\mathcal{L}(\mathcal{A}) = \mathcal{L}(\mathcal{A}_1) \cap \mathcal{L}(\mathcal{A}_2)$ gelten.

Hinweis: Überlegen Sie sich die Aufgabenstellung zuerst an Hand einfacher konkreter Automaten und verallgemeinern Sie dann Ihre Beobachtungen. Wählen Sie für den neuen Automaten $\mathcal{A}$ Zustände mit der Bezeichnung q_1q_2, die signalisieren, dass sich der Automat $\mathcal{A}_1$ momentan im Zustand q_1 und der Automat $\mathcal{A}_2$ im Zustand q_2 befindet.

Lösung

Wir konstruieren einen Automaten $\mathcal{A}$, der die beiden Automaten $\mathcal{A}_1$ und $\mathcal{A}_2$ parallel ausführt. Als Zustände für $\mathcal{A}$ verwenden wir Paare (q_1, q_2), wobei $q_1 \in Q_1$ ein Zustand des ersten Automaten und $q_2 \in Q_2$ ein Zustand des zweiten Automaten ist. Der neue Automat befindet sich bei Eingabe eines Wortes w im Zustand (q_1, q_2), wenn sich der erste Automat bei diesem Wort im Zustand q_1 und der zweite im Zustand q_2 befinden würde. Der Startzustand (i_1, i_2) entspricht der Situation, in der sich die beiden ursprünglichen Automaten im Startzustand befinden. Ein Übergang mit dem Symbol s von (q_1, q_2) nach (q_1', q_2') existiert genau dann, wenn wenn man mit diesem Symbol in $\mathcal{A}_1$

von q_1 nach $q_1' = \delta_1(q_1, s)$ und in $\mathcal{A}_2$ von q_2 nach $q_2' = \delta_2(q_2, s)$ gelangt. Ein Wort wird von beiden Automaten akzeptiert (und liegt daher im Durchschnitt der Sprachen), wenn der neue Automat einen Zustand (q_1, q_2) erreicht, bei dem beide Zustände Endzustände im jeweiligen Automaten sind, wenn also $q_1 \in F_1$ und $q_2 \in F_2$ gilt. Ein Automat für den Durchschnitt der Sprachen $\mathcal{L}(\mathcal{A}_1)$ und $\mathcal{L}(\mathcal{A}_2)$ lässt sich somit durch

$$\mathcal{A} = \langle Q_1 \times Q_2, \Sigma, \delta, (i_1, i_2), F_1 \times F_2 \rangle$$

definieren, wobei die Übergangsfunktion festgelegt ist durch

$$\delta((q_1, q_2), s) = (\delta_1(q_1, s), \delta_2(q_2, s)) \quad \text{für alle } (q_1, q_2) \in Q_1 \times Q_2 \text{ und alle } s \in \Sigma.$$

Als Beispiel betrachten wir die Automaten

$$\mathcal{A}_1 = \langle \{1, 2, 3\}, \{\mathtt{a}, \mathtt{b}\}, \delta_1, 1, \{2\} \rangle$$

$$\mathcal{A}_2 = \langle \{x, y\}, \{\mathtt{a}, \mathtt{b}\}, \delta_2, x, \{y\} \rangle$$

δ_1	a	b
1	2	3
2	2	2
3	3	3

δ_2	a	b
x	y	x
y	y	x

Automat $\mathcal{A}_1$ akzeptiert alle Wörter über $\Sigma = \{\mathtt{a}, \mathtt{b}\}$, die mit a beginnen, und Automat $\mathcal{A}_2$ jene, die mit a aufhören, d.h., $\mathcal{L}(\mathcal{A}_1) = \{\mathtt{a}\} \cdot \{\mathtt{a}, \mathtt{b}\}^*$ und $\mathcal{L}(\mathcal{A}_2) = \{\mathtt{a}, \mathtt{b}\}^* \cdot \{\mathtt{a}\}$. Das oben beschriebene Verfahren liefert den Automaten

$$\mathcal{A} = \langle \{1x, 1y, 2x, 2y, 3x, 3y\}, \{\mathtt{a}, \mathtt{b}\}, \delta, 1x, \{2y\} \rangle$$

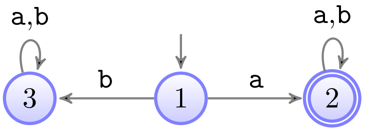
mit der Übergangsfunktion

δ	a	b
$1x$	$2y$	$3x$
$1y$	$2y$	$3x$
$2x$	$2y$	$2x$
$2y$	$2y$	$2x$
$3x$	$3y$	$3x$
$3y$	$3y$	$3x$

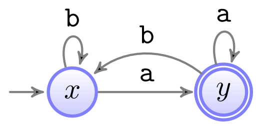
wobei die Zustandsbezeichnungen (q_1, q_2) auf q_1q_2 verkürzt wurden.[1] Für den Automaten $\mathcal{A}$ gilt $\mathcal{L}(\mathcal{A}) = \mathcal{L}(\mathcal{A}_1) \cap \mathcal{L}(\mathcal{A}_2) = \{\mathtt{a}\} \cdot \{\mathtt{a}, \mathtt{b}\}^* \cdot \{\mathtt{a}\}$, d.h., der Automat akzeptiert alle Wörter, die sowohl mit a beginnen als auch aufhören. Abbildung 1 stellt die drei Automaten graphisch dar.

Diese Überlegung zeigt, dass der Schnitt von Sprachen, die von Automaten akzeptiert werden, wieder durch einen Automaten beschrieben werden kann. Die Familie der Sprachen, die von endlichen Automaten akzeptiert werden, ist daher abgeschlossen gegenüber Durchschnitt. In der Vorlesung haben wir gezeigt, dass diese Spachfamilie genau die regulären Sprachen sind. Daher sind auch die regulären Sprachen abgeschlossen gegenüber Durschnitt: Der Durchschnitt zweier regulärer Sprachen ist wieder regulär.

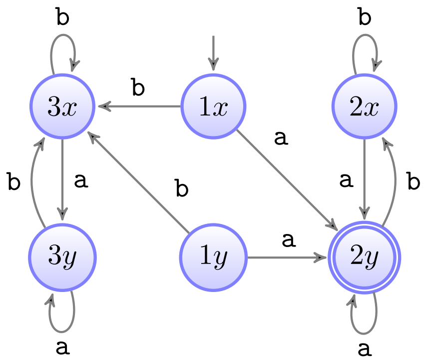
[1] Die konkreten Bezeichnungen der Zustände haben keinen Einfluss auf die akzeptierte Sprache des Automaten. Sie können daher so gewählt werden, dass sie einen Anhaltspunkt für die Rolle bieten, die die Zutände im Automaten spielen.



(a) Automat für die Sprache $\{\mathtt{a}\} \cdot \{\mathtt{a}, \mathtt{b}\}^*$



(b) Automat für die Sprache $\{\mathtt{a}, \mathtt{b}\}^* \cdot \{\mathtt{a}\}$



(c) Automat für die Sprache $\{\mathtt{a}\} \cdot \{\mathtt{a}, \mathtt{b}\}^* \cdot \{\mathtt{a}\}$

Abbildung 1: Beispiel für die Schnittbildung bei Automaten (Aufgabe 6)

Aufgabe 7 (2 Punkte)

Geben Sie endliche Automaten an, die dieselbe Sprache beschreiben wie die folgenden regulären Ausdrücke in algebraischer Notation.

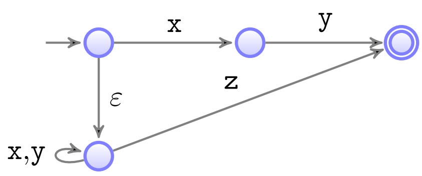
(a) $(\mathtt{x} + \mathtt{y})^*\mathtt{z} + \mathtt{xy}$

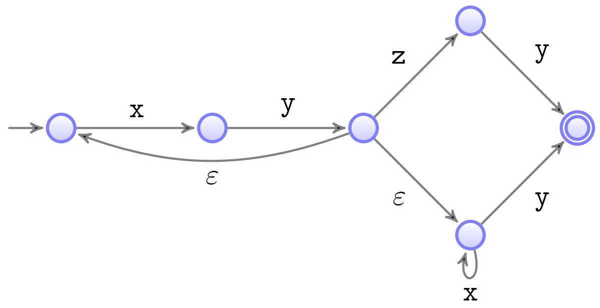
(b) $(\mathtt{xy})^+(\mathtt{z} + \mathtt{x}^*)\mathtt{y}$

Lösung

Die gesuchten Automaten können mit dem allgemeinen Verfahren konstruiert werden, enthalten dann aber in der Regel viel mehr Zustände und ε-Kanten als notwendig. Die folgenden Automaten wurden bereits vereinfacht.

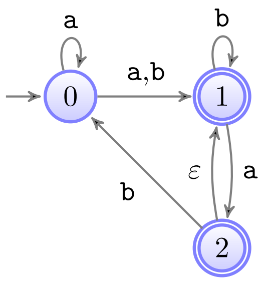
(a)



(b) 

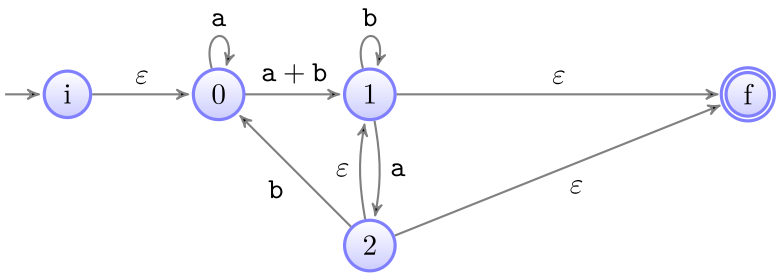
Aufgabe 8 (3 Punkte)

Konstruieren Sie zu folgendem endlichen Automaten einen regulären Ausdruck. Orientieren Sie sich am Algorithmus, der in der Vorlesung besprochen wurde und geben Sie den Automaten nach jeder Zustandselimination an!



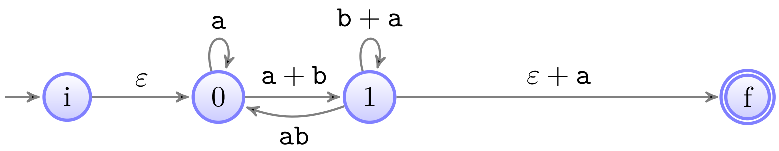
Lösung

Neuer Anfangs- und Endzustand:

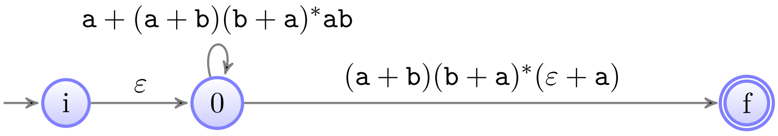


Wir eliminieren die Zustände in der Reihenfolge 2, 1 und 0; die anderen Reihenfolgen sind ebenfalls möglich.

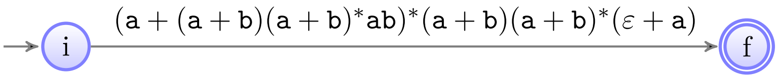
Elimination von Zustand 2:



Elimination von Zustand 1:



Elimination von Zustand 0:



Die Sprache des ursprünglichen Automaten wird also durch $(a+(a+b)^+ab)^*(a+b)^+(\varepsilon+a)$ beschrieben.

Aufgabe 9 **(3 Punkte)**

Auf Milchpackungen in der EU findet man Aufdrucke wie in dieser Abbildung.

Eine kurze wikipedia-Recherche enthüllt das Geheimnis um die Bedeutung der Buchstaben und Ziffern:

> *Das Identitätskennzeichen wird von Lebensmittelunternehmern auf sonstigen tierischen Erzeugnissen bzw. deren Verpackungen angebracht. Die Veterinärnummer ist eine vorgeschriebene Deklaration der Produktionsstätte, die EU-weit auf allen Produkten tierischer Herkunft zu finden ist. Sie besteht aus einer Zahlen-Buchstaben-Kombination, die jedem Betrieb zugeordnet ist* [...] *Vor der Veterinärnummer steht stets das Landeskürzel, z.B. DE für Deutschland oder AT für Österreich. Die abschließenden Zeichen sind stets die landestypische Abkürzung des Begriffs der Europäischen Gemeinschaft, z.B. EG (bzw. in anderen EU-Sprachen: CE, EC, EB, EF, EK, EO, ES, EÜ, EY oder WE). Die Veterinärnummer selbst besteht bei Molkereiprodukten aus einem Kürzel für das Bundesland und einer drei- bis fünfstelligen Zahl für den jeweiligen Betrieb.*

Beschreiben Sie den Aufbau des „Identitätskennzeichens" mit den folgenden Methoden. Beschreiben Sie dabei nur Veterinärnummern von österreichischen und deutschen Produkten, wobei als Kürzel für das Bundesland jegliche Kombinationen aus zwei Großbuchstaben zulässig sind. Treffen Sie sinnvolle Annahmen, wenn Ihnen Informationen fehlen.

(a) Geben Sie einen regulären Ausdruck in algebraischer Notation an.

(b) Geben Sie einen regulären Ausdruck in POSIX-Notation an, der alle Zeilen beschreibt, die *ausschließlich* einen derartigen Dateinamen enthalten.

(c) Zeichnen Sie das Syntaxdiagramm, das Ihrem regulären Ausdruck aus Teil a entspricht.

Lösung

(a) Wir schreiben Symbole des Alphabets unter Anführungszeichen, um sie von algebraischen Symbolen (Metanotation) zu unterscheiden.

$(\texttt{"DE"} + \texttt{"AT"})\,\texttt{"\textvisiblespace"}\,\mathit{Alpha}\,\mathit{Alpha}\,\texttt{"\textvisiblespace"}\,\mathit{num}\,\mathit{num}\,\mathit{num}\,(\mathit{num}\ + \varepsilon)(\mathit{num}\ + \varepsilon)\,\texttt{"\textvisiblespace EG"}$

mit den Abkürzungen

$$\begin{aligned} \mathit{Alpha} &:= (\texttt{"A"} + \cdots + \texttt{"Z"}) \\ \mathit{num} &:= (\texttt{"0"} + \cdots + \texttt{"9"}) \end{aligned}$$

(b) `^(DE|AT)␣[A-Z]{2}␣[0-9]{3,5}␣EG$`

(c) Syntaxdiagramm: