2. Übungsblatt (mit Lösungen)

3.0 VU Formale Modellierung

Lara Spendier, Gernot Salzer

29. April 2012

Aufgabe 1 (0.3 Punkte)

Vereinfachen bzw. berechnen Sie die folgenden Ausdrücke.

(a) $\{\mathtt{1}, \mathtt{10}, \mathtt{11}\} \cap \{\}$

(b) $\{\mathtt{1}, \mathtt{10}, \mathtt{11}\} \cdot \{\}$

(c) $\{\mathtt{a}\}^* \cdot (\{\mathtt{a}\} \cup \{\varepsilon\})$

(d) $(\{\mathtt{a}\}^* \cdot \{\mathtt{a}\}^+) \cdot \{\mathtt{a}\}^*$

(e) $\{\}^*$

(f) $(\{\mathtt{0}\} \cup \{\varepsilon\}) \cdot (\{\varepsilon\} \cup \{\mathtt{1}\})$

Lösung

(a) $(\{\mathtt{1}, \mathtt{10}, \mathtt{11}\} \cap \{\} = \{\}$

(b) $\{\mathtt{1}, \mathtt{10}, \mathtt{11}\} \cdot \{\} = \{\}$

(c) $\{\mathtt{a}\}^* \cdot (\{\mathtt{a}\} \cup \{\varepsilon\}) = \{\mathtt{a}\}^* \cdot \{\mathtt{a}\} \cup \{\mathtt{a}\}^* \cdot \{\varepsilon\} = \{\mathtt{a}\}^+ \cup \{\mathtt{a}\}^* = \{\mathtt{a}\}^*$

(d) $(\{\mathtt{a}\}^* \cdot \{\mathtt{a}\}^+) \cdot \{\mathtt{a}\}^* = \{\mathtt{a}\}^+ \cdot \{\mathtt{a}\}^* = \{\mathtt{a}\}^+$

(e) $\{\}^* = \{\varepsilon\}$

(f) $(\{\mathtt{0}\} \cup \{\varepsilon\}) \cdot (\{\varepsilon\} \cup \{\mathtt{1}\}) = \{\varepsilon, 0\} \cdot \{\varepsilon, 1\} = \{\varepsilon, \mathtt{0}, \mathtt{1}, \mathtt{01}\}$

Aufgabe 2 (0.3 Punkte)

Welche der folgenden Gleichungen sind für alle Sprachen L_1, L_2, L_3 gültig? Begründen Sie Ihre Antwort bzw. geben Sie ein Gegenbeispiel, wenn die Gleichung nicht gültig ist.

(a) $L_1 \cdot L_2 = L_2 \cdot L_1$

(b) $L_1 \cup L_2 = L_2 \cup L_1$

(c) $L_1 \cdot (L_2 \cup L_3) = (L_1 \cdot L_2) \cup (L_1 \cdot L_3)$

(d) $(L_1 \cup L_2)^+ = L_1^+ \cup L_2^+$

Lösung

(a) $L_1 \cdot L_2 = L_2 \cdot L_1$ ist nicht allgemein gültig, da die Konkatenation von Worten und daher auch von Sprachen nicht kommutativ ist. Für $L_1 = \{\mathtt{a}\}$ und $L_2 = \{\mathtt{b}\}$ etwa erhalten wir $L_1 \cdot L_2 = \{\mathtt{ab}\} \neq \{\mathtt{ba}\} = L_2 \cdot L_1$.

(b) $L_1 \cup L_2 = L_2 \cup L_1$ ist für alle Sprachen L_1, L_2 gültig, da Mengenvereinigung kommutativ ist.

(c) $L_1 \cdot (L_2 \cup L_3) = (L_1 \cdot L_2) \cup (L_1 \cdot L_3)$ gilt für alle Sprachen L_1, L_2, L_3 (Distributivgesetz).

(d) $(L_1 \cup L_2)^+ = L_1^+ \cup L_2^+$ ist nicht allgemein gültig. Für $L_1 = \{\mathtt{0}\}$ und $L_2 = \{\mathtt{1}\}$ etwa erhalten wir $(L_1 \cup L_2)^+ = \{\mathtt{0}, \mathtt{1}\}^+$, was verschieden von $\{\mathtt{0}\}^+ \cup \{\mathtt{1}\}^+$ ist.

Aufgabe 3 (0.4 Punkte)

Sei $\mathcal{A} = \langle Q, \Sigma, \delta, q_0, F \rangle$ ein beliebiger deterministischer Automat und $L = \mathcal{L}(\mathcal{A})$ die von ihm akzeptierte Sprache. Geben Sie ein Verfahren an, um daraus einen Automaten $\mathcal{A}'$ für das Komplement dieser Sprache zu erhalten; es soll also $\mathcal{L}(\mathcal{A}') = \Sigma^* \setminus L$ gelten.[1] Der Automat $\mathcal{A}'$ akzeptiert somit jene Worte, die $\mathcal{A}$ nicht akzeptiert, und akzeptiert jene Worte nicht, die $\mathcal{A}$ akzeptiert. Welche Eigenschaft regulärer Sprachen ergibt sich daraus? Lässt sich das Verfahren auch auf nicht-deterministische Automaten anwenden? Wenden Sie Ihr Verfahren auf einen selbst gewählten Automaten an und konstruieren Sie den Automaten für die Komplementärsprache.

Hinweis: Überlegen Sie sich die Aufgabenstellung zuerst an Hand einfacher konkreter Automaten und verallgemeinern Sie dann Ihre Beobachtungen.

[1] Differenz zweier Mengen A und B: $A \setminus B = \{ x \in A \mid x \notin B \}$. Anstelle des Schrägstrichs wird auch das normale Minuszeichen verwendet.

(a) Automat für die Sprache $\{\texttt{ab}\}^*$ (b) Automat für das Komplement $\Sigma^* \setminus \{\texttt{ab}\}^*$

Abbildung 1: Beispiel für die Komplementbildung bei Automaten (Aufgabe 3)

Lösung

In einem deterministischen Automaten ist die Übergangsfunktion δ total, d.h., es gibt zu jedem Zustand q und jedem Symbol s einen eindeutig definierten Folgezustand $\delta(q, s)$. Daher ist auch die erweiterte Übergangsfunktion δ^* total, und der vom Startzustand q_0 mit einem Wort w erreichbare Zustand $q = \delta^*(q_0, w)$ ist eindeutig festgelegt. Der Automat akzeptiert das Wort w genau dann, wenn q ein Endzustand ist. Es genügt daher, den Status der Zustände zu vertauschen und jeden Nicht-Endzustand zu einem Endzustand zu machen und umgekehrt. Ein Automat für das Komplement der Sprache $\mathcal{L}(\mathcal{A})$ lässt sich somit durch $\mathcal{A}' = \langle Q, \Sigma, \delta, q_0, Q \setminus F\rangle$ definieren. Ein Beispiel für diese Konstruktion ist in Abbildung 1 angegeben.

Diese Überlegung zeigt, dass das Komplement jeder durch einen Automaten akzeptierten Sprache wieder durch einen Automaten beschrieben werden kann. Die Familie der Sprachen, die von endlichen Automaten akzeptiert werden, ist daher abgeschlossen gegenüber Komplementbildung. In der Vorlesung haben wir gezeigt, dass diese Spachfamilie genau die regulären Sprachen sind. Daher sind auch die regulären Sprachen abgeschlossen gegenüber Komplementbildung: Das Komplement einer regulären Sprache ist wieder regulär.

Das angegebene Verfahren lässt sich *nicht* auf indeterministische Automaten anwenden. In der Regel sind mit einem bestimmten Wort vom Startzustand aus mehrere Zustände erreichbar. Wenn mindestens einer dieser Zustände ein Endzustand ist, wird das Wort akzeptiert. Nehmen wir nun an, dass sich unter den erreichbaren Zuständen zusätzlich ein Nicht-Endzustand befindet. Dreht man wie oben beschrieben den Status aller Zustände um, wird aus dem vormaligen Nicht-Endzustand ein Endzustand. Somit wird das Wort immer noch akzeptiert, obwohl es nicht im Komplement der Sprache liegt. Um einen Automat für das Komplement zu finden, muss man zuerst den indeterministischen Automaten determinisieren (das entsprechende Verfahren wurde nicht in der Vorlesung besprochen, existiert aber) und dann erst die obige Methode anwenden.

Aufgabe 4 (0.4 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 dem Alphabet Σ und $L_1 = \mathcal{L}(\mathcal{A}_1)$ bzw. $L_2 = \mathcal{L}(\mathcal{A}_2)$ die von ihnen akzeptierten Sprachen. Geben Sie ein Verfahren an, um daraus einen Automaten $\mathcal{A}$ für den Durchschnitt dieser Sprachen zu erhalten; es soll also $\mathcal{L}(\mathcal{A}) = L_1 \cap L_2$ gelten. Der Automat $\mathcal{A}$ akzeptiert somit genau jene Worte, die sowohl von $\mathcal{A}_1$ als auch von $\mathcal{A}_2$ akzeptiert werden. Welche Eigenschaft regulärer Sprachen ergibt sich daraus? Geben Sie als Beispiel zwei konkrete Automaten und den daraus mit Ihrem Verfahren konstruierten Automaten für die Schnittsprache an.

Hinweis: Überlegen Sie sich die Aufgabenstellung zuerst an Hand einfacher konkreter Automaten und verallgemeinern Sie dann Ihre Beobachtungen.

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

$$\delta((q_1, q_2), s) = (\delta_1(q_1, s), \delta_2(q_2, s))$$

für alle $(q_1, q_2) \in Q_1 \times Q_2$ und alle $s \in \Sigma$.

Als Beispiel betrachten wir die Automaten

$$\begin{aligned} \mathcal{A}_1 &= \langle \{1, 2, 3\}, \{\mathtt{a}, \mathtt{b}\}, \delta_1, 1, \{2\} \rangle \\ \mathcal{A}_2 &= \langle \{x, y\}, \{\mathtt{a}, \mathtt{b}\}, \delta_2, x, \{y\} \rangle \end{aligned}$$

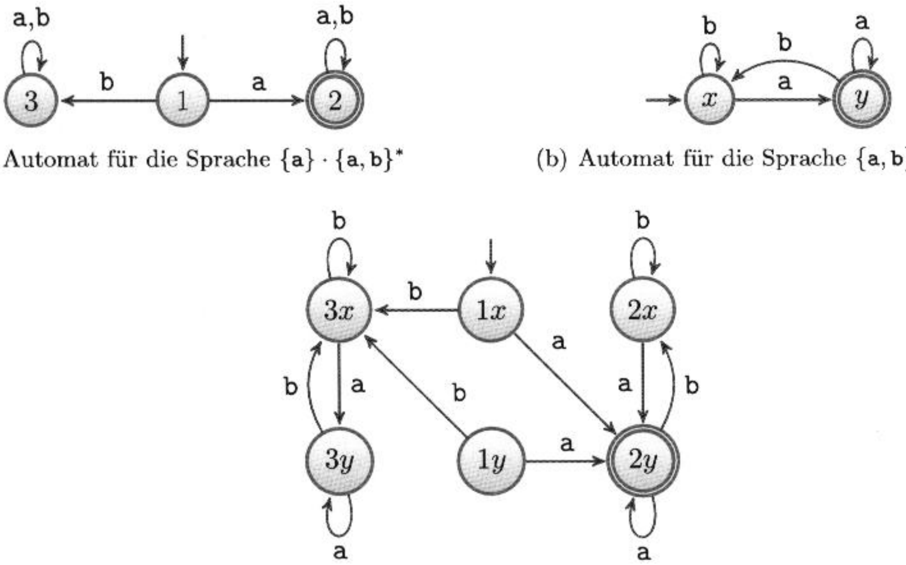
mit den Übergangsfunktionen

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 Worte 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

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



(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 2: Beispiel für die Schnittbildung bei Automaten (Aufgabe 4)

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 Zustandsbezeichnugnen (q_1, q_2) auf $q_1 q_2$ verkürzt wurden.[2] 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 2 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.

[2] 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.

Aufgabe 5 (0.3 Punkte)

PKWs, die in den Niederlanden seit 2008 neu angemeldet werden, bekommen ein Kennzeichen der Form `NL␣99-XXX-9`, wobei gilt:

- Das Kennzeichen beginnt mit der Buchstabenkombination `NL`, gefolgt von einem Leerzeichen.
- `9` steht für eine beliebige Ziffer zwischen 0 und 9.
- `X` steht für einen beliebigen Großbuchstaben des Alphabets ausgenommen Vokale (`A`, `E`, `I`, `O`, `U` und `Y`) und die Konsonanten `C` und `Q`.[3]

(a) Geben Sie einen regulären Ausdruck in algebraischer Notation an, der die Menge aller solchen Autokennzeichen beschreibt.

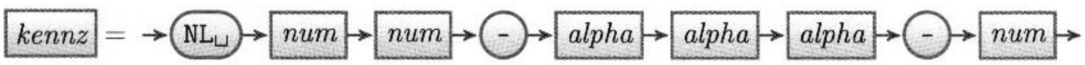
(b) Geben Sie einen POSIX Extended Regular Expression an, der alle Zeilen beschreibt, die ausschließlich ein Kennzeichen enthalten.

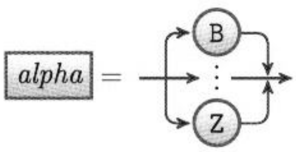
(c) Zeichnen Sie das Syntaxdiagramm, das Ihrem regulären Ausdruck aus Teil a entspricht.

Lösung

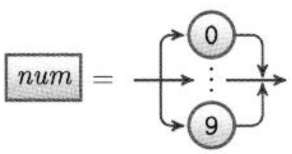
(a) *kennz* = `NL␣`*num num*`-`*alpha alpha alpha*`-`*num*
wobei *alpha* und *num* Abkürzungen sind:
$alpha = \mathtt{B} + \mathtt{D} + \mathtt{F} + \mathtt{G} + \mathtt{H} + \mathtt{J} + \mathtt{K} + \mathtt{L} + \mathtt{M} + \mathtt{N} + \mathtt{P} + \mathtt{R} + \mathtt{S} + \mathtt{T} + \mathtt{V} + \mathtt{W} + \mathtt{X} + \mathtt{Z}$
$num = \mathtt{0} + \cdots + \mathtt{9}$

(b) `^NL␣[0-9]{2}-[BDFGHJKLMNPRSTVWXZ]{3}-[0-9]$`

(c)




Die Punkte stehen für die Buchstaben `DFGHJKLMNPRSTVWX`.



Die Punkte stehen für die Ziffern `12345678`.

[3] Vokale werden ausgeschlossen, um Wortbildungen zu vermeiden. Bei C und Q besteht Verwechslungsgefahr mit der Ziffer 0. Genau genommen werden auch noch jene Buchstabenkombinationen vermieden, die eine negative Bedeutung besitzen, etwa Abkürzungen für Schimpfworte oder für Bezeichnungen aus dem zweiten Weltkrieg.

Aufgabe 6 (0.3 Punkte)

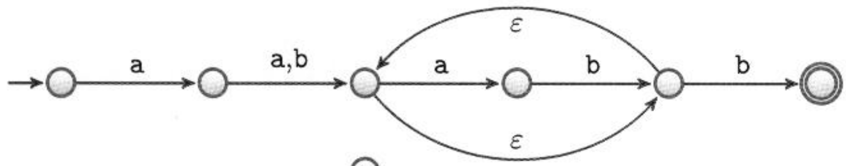
Konstruieren Sie endliche Automaten, die dieselbe Sprache beschreiben wie die folgenden regulären Ausdrücke.

(a) $\mathtt{a}(\mathtt{a}+\mathtt{b})(\mathtt{ab})^*\mathtt{b}$

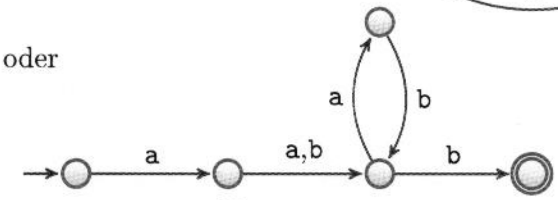
(b) $(\mathtt{ab}+(\mathtt{ac})^*)\mathtt{ca}^+$

Lösung

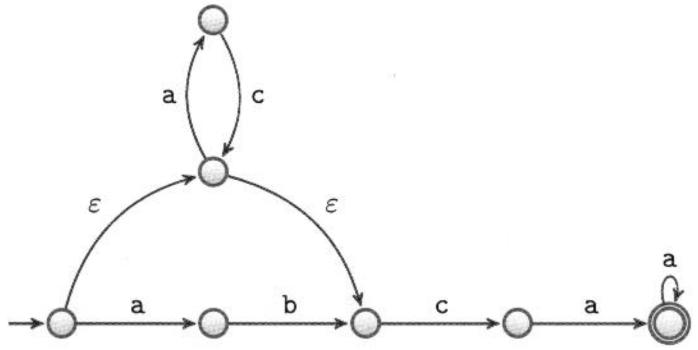
(a)



oder

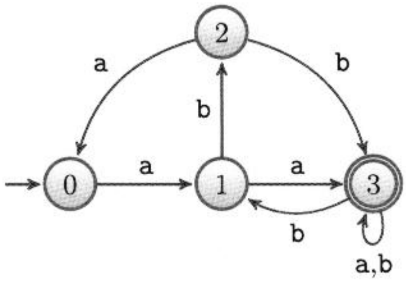


(b)



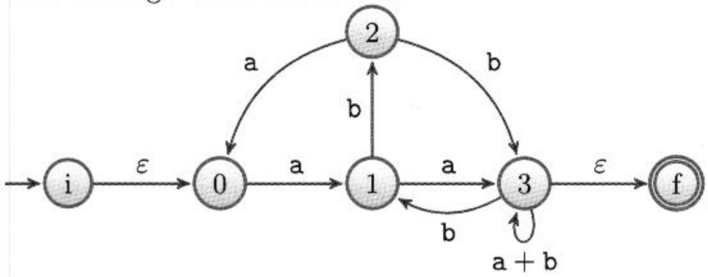
Aufgabe 7 (0.4 Punkte)

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!



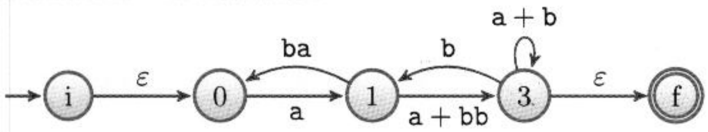
Lösung

Neuer Anfangs- und Endzustand:

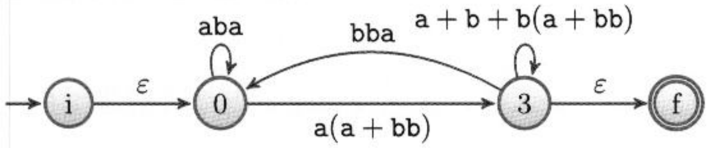


Wir eliminieren die Zustände in der Reihenfolge 2, 1, 0 und 3; andere Reihenfolgen sind ebenfalls möglich.

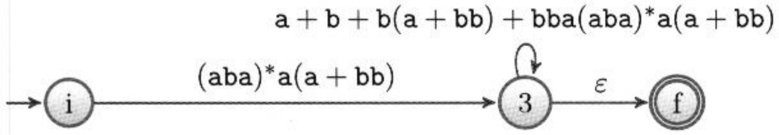
Elimination von Zustand 2:



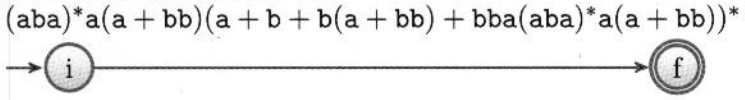
Elimination von Zustand 1:



Elimination von Zustand 0:



Elimination von Zustand 3:



Da $(\texttt{a}+\texttt{b})^*$ bereits alle möglichen Wörter über $\{\texttt{a},\texttt{b}\}$ enthält, gilt $(\texttt{a}+\texttt{b}+\cdots)^* = (\texttt{a}+\texttt{b})^*$, der reguläre Ausdruck vereinfacht sich damit zu

$$(\texttt{aba})^*(\texttt{aa} + \texttt{abb})(\texttt{a} + \texttt{b})^* \ .$$

Aufgabe 8 (0.3 Punkte)

Gegeben sei die Grammatik $G = \langle N, T, P, S\rangle$ mit den Nonterminalen $N = \{S, H, T, X, W\}$, den Terminalen

$$T = \{\texttt{0}, \ldots, \texttt{9}, \texttt{a}, \ldots, \texttt{z}, \texttt{A}, \ldots, \texttt{Z}, \texttt{?}, \texttt{.}, \texttt{!}, \texttt{<}, \texttt{>}, \texttt{/}, \sqcup\}$$

und folgender Menge P von Ersetzungsregeln:

$S \rightarrow$ <note> $H\,T$ </note>

$H \rightarrow$ <to> W </to>

$T \rightarrow$ <text> X </text>

$X \rightarrow W \mid$ <i> X </i> $\mid$ X $\mid XX$

$W \rightarrow \varepsilon \mid WW \mid$ 0$W \mid \cdots \mid$ 9$W \mid$ a$W \mid \cdots \mid$ z$W \mid$ A$W \mid \cdots \mid$ Z$W \mid$?$W \mid$.$W \mid$!$W \mid$ ␣W

Überprüfen Sie für die nachfolgenden Wörter, ob sie in der von der Grammatik G spezifizierten Sprache $\mathcal{L}(G)$ liegen. Falls ja, geben Sie eine Parallelableitung an. Falls nein, argumentieren Sie, warum nicht.

(a) `<note><to>Lucy</to><text>Hi!<b>Wann</b>treffen␣wir␣uns?</text></note>`

(b) `<note><to>Myself</to><text><i><b>Morgen</i></b>␣FMOD␣lernen!</text></note>`

Geben Sie eine zu G äquivalente Grammatik an, deren Produktionen mit Hilfe von EBNF-Notationen und den in der Vorlesung besprochenen Richtlinien möglichst verständlich gestaltet sind.

Lösung

(a) Ja, das Wort liegt in der Sprache $\mathcal{L}(G)$:

$S \Rightarrow_P$ <note> $H\ T$ </note>

$\Rightarrow_P$ <note><to> W </to><text> X </text></note>

$\Rightarrow_P$ <note><to> WW </to><text> XX </text></note>

$\Rightarrow_P$ <note><to>L W c W </to><text> WXX </text></note>

$\Rightarrow_P$ <note><to>Lu W cy W </to><text>H W X W </text></note>

$\Rightarrow_P$ <note><to>Lucy</to><text>Hi W W WW </text></note>

$\Rightarrow_P \cdots$ <text>Hi! W WW $WWWWW$ </text></note>

$\Rightarrow_P \cdots$ <text>Hi!W W n W $WWWWWWWWW$ </text></note>

$\Rightarrow_P \cdots$ <text>Hi!Wa W nn W t W e W f W n W w W r W u W s W </text></note>

$\Rightarrow_P \cdots$ <text>Hi!Wanntr W ef W fe W n␣ W wi W r␣ W un W s? W </text></note>

$\Rightarrow_P$ <note><to>Lucy</to><text>Hi!Wanntreffen␣wir␣uns?</text></note>

(b) Nein, dieses Wort ist nicht ableitbar und liegt daher nicht in $\mathcal{L}(G)$: Die Zeichenfolgen <i> und </i> bzw. und können durch die Produktionen

$X \rightarrow \cdots \mid$ <i> X </i> $\mid$ X $\mid XX$

immer nur paarweise und symmetrisch verschachtelt erzeugt werden. Eine Verschränkung <i> $\cdots$ </i> wie im angegebenen Wort ist dadurch nicht möglich.

Andere Argumentation: Die Zeichenfolge <i> des gegebenen Wortes kann nur durch eine Ableitung der Form

$$\cdots X \cdots \Rightarrow \cdots \texttt{<i>}\, X\, \texttt{</i>} \Rightarrow \cdots \Rightarrow \cdots \texttt{<i>}\, X\, \texttt{</i>} \Rightarrow \cdots$$

entstehen. Dadurch entsteht allerdings auch die Zeichenfolge </i>, die aber im gegebenen Wort nicht auftritt. Daher ist das Wort nicht ableitbar.

Die angegebene Grammatik besitzt hinsichtlich Lesbarkeit einige Schwächen: schlechte Strukturierung, keine sprechenden Variablenbezeichnungen, vermeidbare Rekursionen sowie schlechte Unterscheidbarkeit von Terminal- und Nonterminalsymbolen. Eine verständlichere Grammatik ist

$$\langle \{\mathit{Note}, \mathit{To}, \mathit{Text}, \ldots, \mathit{Punctuation}\}, T, P, \mathit{Note} \rangle \ ,$$

wobei T wie oben definiert ist und P aus folgenden Produktionen besteht:

Note	→ "<note>" *To Text* "</note>"
To	→ "<to>" *Plain* "</to>"
Text	→ "<text>" *Formatted* "</text>"
Formatted	→ { *Plain* | "<i>" *Formatted* "</i>" | "" *Formatted* "" }
Plain	→ { *Letter* | *Digit* | *Punctuation* }
Letter	→ "a" | ··· | "z" | "A" | ··· | "Z"
Digit	→ "0" | ··· | "9"
Punctuation	→ "␣" | "." | "?" | "!"

EBNF

Aufgabe 9 (0.4 Punkte)

Ein *Dokument* der Sprache HTML beginnt mit der Zeichenfolge <html>, dann folgen Kopf und Hauptteil des Dokumentes sowie die Zeichenfolge </html>. Der *Kopf* eines Dokumentes besteht aus dem Dokumenttitel, dem <head> vorangeht und </head> folgt. Der *Dokumenttitel* besteht aus einem Text eingeschlossen zwischen <title> und </title>. Der *Hauptteil* des Dokumentes beginnt mit <body> und endet mit </body>; dazwischen liegen Texte, geordnete und ungeordnete Listen in beliebiger Reihenfolge und Anzahl (auch gar nichts ist erlaubt). Ein *Text* ist eine nicht-leere Folge von Buchstaben, Ziffern, Leerzeichen und den Sonderzeichen ,, ;, :, ., ! und ?. Eine *geordnete Liste* besteht aus <ol>, einer nicht-leeren Folge von Listeneinträgen und der Zeichenfolge . Ein *Listeneintrag* besteht aus einer möglicherweise leere Folge von Texte, geordneten und ungeordneten Listen, die zwischen <li> und eingeschlossen ist. Eine *ungeordnete Liste* sieht ebenso aus wie eine geordnete, nur werden <ul> und an Stelle von <ol> und verwendet. Beispiel eines solchen Dokuments:

```
<html><head><title>Beispiel</title></head>
      <body>Ich bin ein Text. Ungeordnete Liste:
            <ul><li>Listeneintrag</li>
                <li>Geordnete Liste in einem Listeneintrag:
```

```
                  <ol><li>Schon wieder ein Eintrag.</li></ol>
              </li>
          </ul>
    </body>
</html>
```

(a) Geben Sie für diese Sprache der HTML-Dokumente eine kontextfreie Grammatik in EBNF an. Verwenden Sie so weit wie möglich EBNF-Notationen, um die Grammatik übersichtlich zu halten.

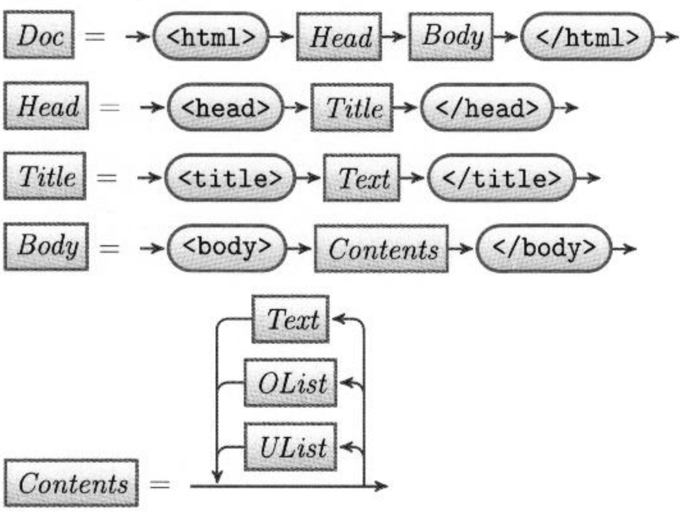
(b) Geben Sie für Ihre Grammatik ein Syntaxdiagramm an.

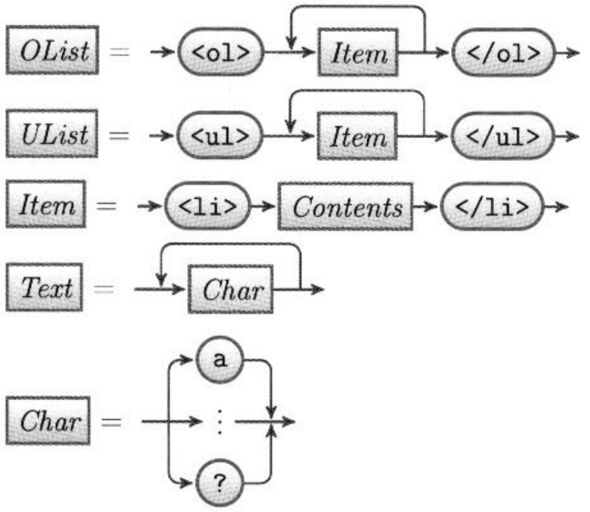
Lösung

(a) $\langle N, T, P, \mathit{Doc}\rangle$, wobei

$$\begin{aligned}
N &= \{\mathit{Doc}, \mathit{Head}, \mathit{Title}, \mathit{Body}, \mathit{OList}, \mathit{UList}, \mathit{Item}, \mathit{Contents}, \mathit{Text}, \mathit{Char}\},\\
T &= \{\cdots \textit{alle Zeichen in den Produktionen zwischen Anführungszeichen} \cdots\},\\
P &= \{ \quad \mathit{Doc} \to \texttt{"<html>"}\ \mathit{Head}\ \mathit{Body}\ \texttt{"</html>"},\\
&\qquad \mathit{Head} \to \texttt{"<head>"}\ \mathit{Title}\ \texttt{"</head>"},\\
&\qquad \mathit{Title} \to \texttt{"<title>"}\ \mathit{Text}\ \texttt{"</title>"},\\
&\qquad \mathit{Body} \to \texttt{"<body>"}\ \mathit{Contents}\ \texttt{"</body>"},\\
&\qquad \mathit{Contents} \to \texttt{\{}\ \mathit{Text} \mid \mathit{OList} \mid \mathit{UList}\ \texttt{\}},\\
&\qquad \mathit{OList} \to \texttt{"
"}\ \mathit{Item}\ \texttt{\{}\ \mathit{Item}\ \texttt{\}}\ \texttt{"
"},\\
&\qquad \mathit{UList} \to \texttt{"
"}\ \mathit{Item}\ \texttt{\{}\ \mathit{Item}\ \texttt{\}}\ \texttt{"
"},\\
&\qquad \mathit{Item} \to \texttt{"
"}\ \mathit{Contents}\ \texttt{"
"},\\
&\qquad \mathit{Text} \to \mathit{Char}\ \texttt{\{}\ \mathit{Char}\ \texttt{\}},\\
&\qquad \mathit{Char} \to \texttt{"a"} \mid \cdots \mid \texttt{"z"} \mid \texttt{"A"} \mid \cdots \mid \texttt{"Z"} \mid \texttt{"0"} \mid \cdots \mid \texttt{"9"} \mid\\
&\qquad\qquad \texttt{"\textvisiblespace"} \mid \texttt{","} \mid \texttt{";"} \mid \texttt{":"} \mid \texttt{"."} \mid \texttt{"!"} \mid \texttt{"?"}\ \}.
\end{aligned}$$

(b) Syntaxdiagramme:





Aufgabe 10 (0.4 Punkte)

Sei $\mathcal{A} = \langle Q, \Sigma, \delta, q_0, F\rangle$ ein beliebiger deterministischer Automat. Geben Sie ein Verfahren zur Konstruktion einer kontextfreien Grammatik G an, die dieselbe Sprache beschreibt, d.h., es soll $\mathcal{L}(G) = \mathcal{L}(\mathcal{A})$ gelten. Welche Eigenschaft regulärer und kontextfreier Sprachen ergibt sich daraus? Lässt sich das Verfahren auch auf nicht-deterministische Automaten anwenden? Geben Sie als Beispiel einen konkreten Automaten und die daraus mit Ihrem Verfahren konstruierte Grammatik an.

Hinweis: Überlegen Sie sich die Aufgabenstellung zuerst an Hand einfacher konkreter Automaten und verallgemeinern Sie dann Ihre Beobachtungen.

Lösung

Um einen Automaten in eine äquivalente Grammatik umzuformen, macht man Zustände zu Nonterminalen und Übergänge zu Produktionen. Zusätzlich muss man für jeden Endzustand eine ε-Produktion vorsehen. Das heißt: Wenn der Automat als $\langle Q, \Sigma, \delta, q_0, F\rangle$ gegeben ist, ist $\langle Q, \Sigma, P, q_0\rangle$ mit den Produktionen

$$P = \{\, q \to sq' \mid \delta(q, s) = q' \,\} \cup \{\, q \to \varepsilon \mid q \in F \,\}$$

eine Grammatik, die dieselbe Sprache wie der Automat akzeptiert.

Als Beispiel geben wir eine Grammatik für den Automaten aus Abbildung 2(c) an:

$$G = \langle \{\langle \mathit{1x}\rangle, \langle \mathit{1y}\rangle, \langle \mathit{2x}\rangle, \langle \mathit{2y}\rangle, \langle \mathit{3x}\rangle, \langle \mathit{3y}\rangle\}, \{\mathtt{a}, \mathtt{b}\}, P, \langle \mathit{1x}\rangle\rangle$$

wobei P die Menge folgender Produktionen ist:

$$\begin{aligned}
\langle 1x\rangle &\to \mathsf{a}\,\langle 2y\rangle \mid \mathsf{b}\,\langle 3x\rangle\\
\langle 1y\rangle &\to \mathsf{a}\,\langle 2y\rangle \mid \mathsf{b}\,\langle 3x\rangle\\
\langle 2x\rangle &\to \mathsf{a}\,\langle 2y\rangle \mid \mathsf{b}\,\langle 2x\rangle\\
\langle 2y\rangle &\to \mathsf{a}\,\langle 2y\rangle \mid \mathsf{b}\,\langle 2x\rangle \mid \varepsilon\\
\langle 3x\rangle &\to \mathsf{a}\,\langle 3y\rangle \mid \mathsf{b}\,\langle 3x\rangle\\
\langle 3y\rangle &\to \mathsf{a}\,\langle 3y\rangle \mid \mathsf{b}\,\langle 3x\rangle
\end{aligned}$$

Diese Transformation lässt sich auch auf indeterministische Automaten anwenden; man muss in der Definition der Produktionen lediglich $\delta(q,s)=q'$ durch $(q,s,q')\in\delta$ ersetzen.

Da die von Automaten akzeptierten Sprachen genau die regulären Sprachen sind, folgt daraus, dass die Familie der regulären Sprachen eine Teilmenge der Familie der kontextfreien Sprachen ist. Anders ausgedrückt ist jede reguläre Sprache auch eine kontextfreie Sprache.

Eine andere Methode zur Umwandlung eines Automaten in eine äquivalente Grammatik besteht darin, einen regulären Ausdruck zum Automaten zu konstruieren und diesen mit Hilfe der EBNF-Notation zu beschreiben. Dann reicht eine einzige Produktion, die die Startvariable in diesen regulären Ausdruck übergehen lässt. Soll die Grammatik ohne EBNF-Notationen auskommen, kann man diese wie in der Vorlesung beschrieben durch einfache Produktionen ersetzen.

Aufgabe 11 (0.3 Punkte)

Seien $P/2$, $S/1$ und $K/1$ Prädikatensymbole sowie *lego* und *barbie* Konstantensymbole mit folgender Bedeutung:

$P(x,y)$	... x spielt mit y.
$S(x)$	... x singt.
$K(x)$	... x ist ein Kind.
lego	... Lego
barbie	... Barbie

Drücken Sie die nachfolgenden prädikatenlogischen Formeln als deutsche Sätze aus.

(a) $\exists x\,(K(x)\land P(x,\mathit{lego}))$

(b) $\neg\forall x\,(K(x)\supset(P(x,\mathit{barbie})\land S(x)))$

(c) $\forall x\,(K(x)\supset(P(x,\mathit{barbie})\not\equiv P(x,\mathit{lego})))$

(d) $\exists x\,\forall y\,(K(x)\land(K(y)\supset P(x,y)))$

Lösung

(a) Manche Kinder spielen mit Lego. / Es gibt ein Kind, das mit Lego spielt.

(b) Nicht alle Kinder spielen mit Barbie und singen.

(c) Alle Kinder spielen entweder mit Barbie oder mit Lego (aber nicht mit beiden).

(d) Manche Kinder spielen mit allen Kindern. / Es gibt ein Kind, das mit allen Kindern spielt.

Aufgabe 12 (0.3 Punkte)

Seien *Bestellt*/2, *Assistent*/1 und *Getränk*/1 Prädikatensymbole sowie *tee* und *kaffee* Konstantensymbole mit folgender Bedeutung:

$Bestellt(x,y)$	... x bestellt y
$Assistent(x)$	... x ist ein Assistent
$Getränk(x)$	... x ist ein Getränk
$Heiß(x)$	... x ist heiß
tee	... Tee
kaffee	... Kaffee

Drücken Sie die folgenden Sätze als prädikatenlogische Formeln aus.

(a) Jeder Assistent bestellt entweder Tee oder Kaffee, aber nicht beides.

(b) Manche Assistenten bestellen keinen Kaffee.

(c) Nicht alle Assistenten bestellen ein heißes Getränk.

(d) Jeder Assistent bestellt ein heißes Getränk, aber keinen Tee.

Lösung

(a) $\forall x\,(Assistent(x) \supset (Bestellt(x, tee) \not\equiv Bestellt(x, kaffee)))$

(b) $\exists x\,(Assistent(x) \wedge \neg Bestellt(x, kaffee))$

(c) $\neg\forall x\,\exists y\,(Assistent(x) \supset (Bestellt(x,y) \wedge Getränk(y) \wedge Heiß(y)))$

(d) $\forall x\,\exists y\,(Assistent(x) \supset (Bestellt(x,y) \wedge Getränk(y) \wedge Heiß(y) \wedge \neg Bestellt(x, tee)))$

Aufgabe 13 (0.3 Punkte)

Sue und John veranstalten eine Dinner-Party bei sich zuhause. Sie haben sich bereits ein dreigängiges Menü überlegt, wobei es für Vegetarier und Nicht-Vegetarier unterschiedliche Hauptspeisen gibt. Damit auch während des Essens eine entspannte Stimmung herrscht, wollen sie einen Sitzplan erstellen.

1. Auf jedem Platz sitzt mindestens eine Person.
2. Kein Vegetarier darf einem Nicht-Vegetarier gegenüber sitzen.
3. Neben jedem Mann muss mindestens eine Frau sitzen.

Gegeben sind die folgenden Prädikatensymbole:

$\mathit{Gegenüber}(x, y)$	... x sitzt gegenüber y
$\mathit{Neben}(x, y)$	... x sitzt neben y
$\mathit{Sitzt}(x, y)$	... x sitzt auf Platz y
$\mathit{Mann}(x)$	... x ist ein Mann
$\mathit{Frau}(x)$	... x ist eine Frau
$\mathit{Vegetarier}(x)$	... x ist Vegetarier

Helfen Sie Sue und John, indem Sie die Anforderungen 1–3 mit prädikatenlogischen Formeln modellieren. Benutzen Sie dazu die angegebenen Prädikatensymbole.

(*Graduate Record Examination* (GRE) ist ein Aufnahmetest für Master- und Doktoratsstudien in den USA und anderen englischsprachigen Ländern.)

Lösung

1. $\forall x\, \exists y\, \mathit{Sitzt}(y, x)$

2. $\neg\exists x \exists y\,(\mathit{Vegetarier}(x) \wedge \neg\,\mathit{Vegetarier}(y) \wedge \mathit{Gegenüber}(x,y))$

3. $\forall x\, \exists y\, (\mathit{Mann}(x) \supset (\mathit{Frau}(y) \wedge \mathit{Neben}(x,y)))$

Aufgabe 14 (0.3 Punkte)

Seien *Mag*/2, *Katze*/1 und *Mensch*/1 Prädikatensymbole sowie *odie* ein Konstantensymbol. Die Bedeutung der Symbole ist festgelegt durch folgende Interpretation:

$$\begin{aligned}
\mathcal{U} &= \{\text{Arlene}, \text{Dr.Liz}, \text{Garfield}, \text{Jon}, \text{Nermal}, \text{Odie}, \text{Pooky}\}\\
I(\mathit{Mensch}) &= \{\text{Jon}, \text{Dr.Liz}\}\\
I(\mathit{Katze}) &= \{\text{Arlene}, \text{Garfield}, \text{Nermal}\}\\
I(\mathit{Mag}) &= \{(\text{Jon}, \text{Odie}), (\text{Jon}, \text{Garfield}), (\text{Jon}, \text{Nermal}), (\text{Dr.Liz}, \text{Nermal}),\\
&\qquad (\text{Odie}, \text{Jon}), (\text{Odie}, \text{Dr.Liz}), (\text{Garfield}, \text{Arlene}), (\text{Garfield}, \text{Pooky}),\\
&\qquad (\text{Nermal}, \text{Garfield}), (\text{Arlene}, \text{Garfield})\}\\
I(\mathit{odie}) &= \text{Odie}
\end{aligned}$$

Bestimmen Sie unter Verwendung der Evaluierungsfunktion den Wahrheitswert der folgenden Formeln in der gegebenen Interpretation.

(a) $\forall x\, (\mathit{Mensch}(x) \supset \mathit{Mag}(x, \mathit{odie}))$ ↯

(b) $\forall x\, \exists y\, (\mathit{Mensch}(x) \supset (\mathit{Katze}(y) \wedge \mathit{Mag}(x,y)))$ ✓

(c) $\exists x\, \neg\exists y\, (\mathit{Katze}(x) \wedge \mathit{Katze}(y) \wedge \mathit{Mag}(x,y))$ ✓

Lösung

(a) $\mathsf{val}_{I,\sigma}(\forall x\, \mathit{Mensch}(x) \supset \mathit{Mag}(x, \mathit{odie})) = 1$

$\iff$ $\mathsf{val}_{I,\sigma'}(\mathit{Mensch}(x) \supset \mathit{Mag}(x, \mathit{odie})) = 1$ für alle $\sigma' \overset{x}{\sim} \sigma$

$\iff$ Wenn $\mathsf{val}_{I,\sigma'}(\mathit{Mensch}(x)) = 1$ dann $\mathsf{val}_{I,\sigma'}(\mathit{Mag}(x, \mathit{odie}))) = 1$ für alle $\sigma' \overset{x}{\sim} \sigma$

$\iff$ Wenn $\mathsf{val}_{I,\sigma'}(x) \in I(\mathit{Mensch})$ dann $(\mathsf{val}_{I,\sigma'}(x), \mathsf{val}_{I,\sigma'}(\mathit{odie})) \in I(\mathit{Mag})$ für alle $\sigma' \overset{x}{\sim} \sigma$

$\iff$ Wenn $\sigma'(x) \in \{\text{Jon}, \text{Dr.Liz}\}$ dann $(\sigma'(x), \text{Odie}) \in \{(\text{Jon}, \text{Odie})\}$ für alle $\sigma' \overset{x}{\sim} \sigma$

$\iff$ Für alle $\sigma'(x) \in \{\text{Jon}, \text{Dr.Liz}\}$ gilt $(\sigma'(x), \text{Odie}) \in \{(\text{Jon}, \text{Odie})\}$

Die letzte Aussage ist falsch, da (Dr.Liz, Odie) nicht in $\{(\text{Jon}, \text{Odie})\}$ vorkommt. Somit ist die ursprüngliche Formel nicht wahr in der Interpretation I.

(b) $\mathsf{val}_{I,\sigma}(\forall x\, \exists y\, (\mathit{Mensch}(x) \supset (\mathit{Katze}(y) \wedge \mathit{Mag}(x,y)))) = 1$

$\iff$ $\mathsf{val}_{I,\sigma'}(\exists y\, (\mathit{Mensch}(x) \supset (\mathit{Katze}(y) \wedge \mathit{Mag}(x,y)))) = 1$ für alle $\sigma' \overset{x}{\sim} \sigma$

$\iff$ $\mathsf{val}_{I,\sigma''}(\mathit{Mensch}(x) \supset (\mathit{Katze}(y) \wedge \mathit{Mag}(x,y))) = 1$
für alle $\sigma' \overset{x}{\sim} \sigma$ und ein $\sigma'' \overset{y}{\sim} \sigma'$

$\iff$ Wenn $\mathsf{val}_{I,\sigma''}(\mathit{Mensch}(x)) = 1$ dann $\mathsf{val}_{I,\sigma''}(\mathit{Katze}(y) \wedge \mathit{Mag}(x,y)) = 1$
für alle $\sigma' \overset{x}{\sim} \sigma$ und ein $\sigma'' \overset{y}{\sim} \sigma'$

$\iff$ Wenn $\mathsf{val}_{I,\sigma''}(\mathit{Mensch}(x)) = 1$ dann $\mathsf{val}_{I,\sigma''}(\mathit{Katze}(y)) = 1$ und $\mathsf{val}_{I,\sigma''}(\mathit{Mag}(x,y)) = 1$

für alle $\sigma' \stackrel{x}{\sim} \sigma$ und ein $\sigma'' \stackrel{y}{\sim} \sigma'$

$\iff$ Wenn $\sigma''(x) \in I(\mathit{Mensch})$ dann $\sigma''(y) \in I(\mathit{Katze})$ und $(\sigma''(x), \sigma''(y)) \in I(\mathit{Mag})$
für alle $\sigma' \stackrel{x}{\sim} \sigma$ und ein $\sigma'' \stackrel{y}{\sim} \sigma'$

$\iff$ Für alle $\sigma''(x) \in I(\mathit{Mensch})$ gibt es ein $\sigma''(y) \in I(\mathit{Katze})$
sodass $(\sigma''(x), \sigma''(y)) \in I(\mathit{Mag})$

$\iff$ Für alle $\sigma''(x) \in \{\text{Jon}, \text{Dr.Liz}\}$ gibt es ein $\sigma''(y) \in \{\text{Arlene}, \text{Garfield}, \text{Nermal}\}$
sodass $(\sigma''(x), \sigma''(y)) \in \{\ldots, (\text{Jon}, \text{Garfield}), \ldots, (\text{Dr.Liz}, \text{Nermal}), \ldots\})$

Die letzte Aussage gilt offenbar, daher ist die ursprüngliche Formel wahr in der Interpretation I.

(c) $\mathsf{val}_{I,\sigma}(\exists x\, \neg\exists y\, (\mathit{Katze}(x) \land \mathit{Katze}(y) \land \mathit{Mag}(x,y))) = 1$

$\iff$ $\mathsf{val}_{I,\sigma'}(\neg\exists y\, (\mathit{Katze}(x) \land \mathit{Katze}(y) \land \mathit{Mag}(x,y))) = 1$ für ein $\sigma' \stackrel{x}{\sim} \sigma$

$\iff$ $\mathsf{val}_{I,\sigma'}(\exists y\, (\mathit{Katze}(x) \land \mathit{Katze}(y) \land \mathit{Mag}(x,y))) = 0$ für ein $\sigma' \stackrel{x}{\sim} \sigma$

$\iff$ $\mathsf{val}_{I,\sigma''}(\mathit{Katze}(x) \land \mathit{Katze}(y) \land \mathit{Mag}(x,y)) = 0$ für ein $\sigma' \stackrel{x}{\sim} \sigma$ und alle $\sigma'' \stackrel{y}{\sim} \sigma'$

$\iff$ $\mathsf{val}_{I,\sigma''}(\mathit{Katze}(x)) = 0$ oder $\mathsf{val}_{I,\sigma''}(\mathit{Katze}(y)) = 0$ oder $\mathsf{val}_{I,\sigma''}(\mathit{Mag}(x,y)) = 0$
für ein $\sigma' \stackrel{x}{\sim} \sigma$ und alle $\sigma'' \stackrel{y}{\sim} \sigma'$

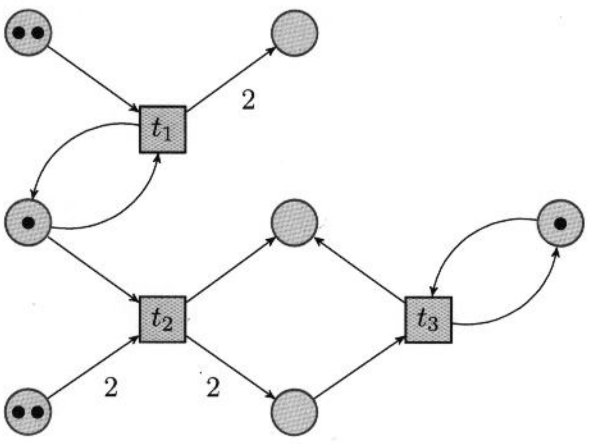
$\iff$ $\sigma''(x) \notin I(\mathit{Katze})$ oder $\sigma''(y) \notin I(\mathit{Katze})$ oder $(\sigma''(x), \sigma''(y)) \notin I(\mathit{Mag})$
für ein $\sigma' \stackrel{x}{\sim} \sigma$ und alle $\sigma'' \stackrel{y}{\sim} \sigma'$

$\iff$ Es gibt ein $\sigma''(x) \in \mathcal{U}$ sodass für alle $\sigma''(y) \in \mathcal{U}$ entweder $\sigma''(x) \notin I(\mathit{Katze})$ oder $\sigma''(y) \notin I(\mathit{Katze})$ oder $(\sigma''(x), \sigma''(y)) \notin I(\mathit{Mag})$ gilt.

Wir wählen $\sigma''(x) = \text{Jon}$. Dann ist die Bedingung $\sigma''(x) \notin I(\mathit{Katze})$ unabhängig von der Wahl für $\sigma''(y)$ immer wahr, die Gesamtaussage gilt. Daher ist die ursprüngliche Formel wahr in der Interpretation I.

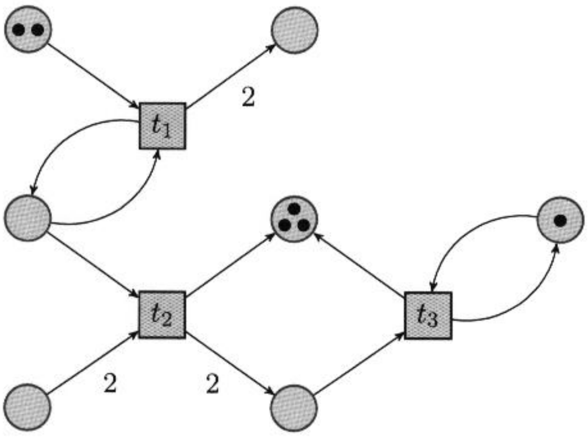
Aufgabe 15 (0.3 Punkte)

Gegeben sei das folgende Petri-Netz mit Anfangsmarkierung. Geben Sie alle möglichen Reihenfolgen an, in denen die Transitionen feuern können. Geben Sie jene erreichbaren Markierungen an, in denen keine Transition aktiviert ist.

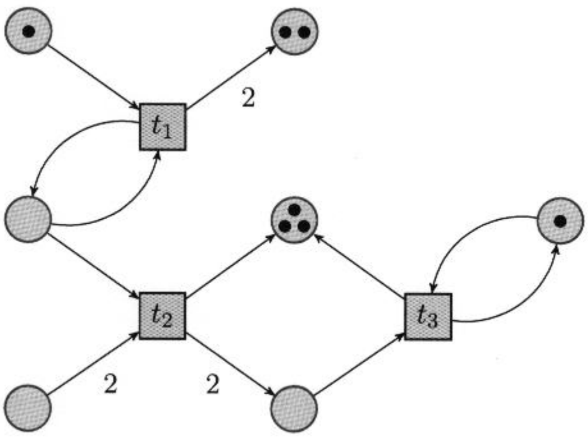


Lösung

Die Transitionsfolge t_2–t_3–t_3 liefert:



Die Transitionsfolge t_1–t_2–t_3–t_3 liefert:



Die Transitionsfolge t_1–t_1–t_2–t_3–t_3 liefert: