Klausur zur Vorlesung „Formale Sprachen und Automaten“

SS 2022 / 22. August 2022

Vorname: ______________________________

Nachname: ______________________________

Matrikelnummer: ______________________________

Aufgabe	**Punktzahl**	**Erreicht**
1	14	
2	9	
3	6	
4	6	
5	6	
6	6	
7	6	
8	6	
9	6	
10	0	
Σ	65	

Generelle Hinweise:

- Prüfungsdauer: **120 Minuten**
- Hilfsmittel: Ein beidseitig beschriebenes DIN-A4-Blatt.
- Benutzen Sie ein **dokumentenechtes Schreibgerät**.
- Überprüfen Sie die Ihnen ausgehändigte Klausur auf Vollständigkeit (**10 Aufgaben** auf 11 Seiten).
- Tragen Sie **auf jedes Blatt** Ihren **Namen** und Ihre **Matr.-Nr.** in die entsprechenden Felder ein.
- Schreiben Sie Ihre Lösungen in die dafür vorgesehenen Felder. Reicht der Platz in einem Feld nicht aus, so benutzen Sie die Rückseite des entsprechenden Blattes und vermerken Sie dies auf der Vorderseite. Reicht der Platz dennoch nicht aus, können Sie die Aufsicht nach zusätzlichen Blättern fragen.
- Schreiben Sie bitte **deutlich**. Unleserliche Lösungen sind ungültig.
- Ein **Täuschungsversuch** führt umgehend zum Ausschluss und **Nichtbestehen**. Es erfolgt keine Vorwarnung.
- Alle mitgeführten **elektronischen Geräte** sind vor der Klausur bzw. spätestens jetzt auszuschalten.

Inhaltliche Hinweise:

- Endliche Automaten können wahlweise grafisch oder tabellarisch angegeben werden.

Aufgabe 1. (14 Punkte) Beantworten Sie die folgenden sieben Fragen. Für jede Teilaufgabe gibt es zwei Punkte, wenn Sie die Aufgabe vollständig und korrekt beantwortet haben.

(1) Seien K, L, M Sprachen mit $K \subseteq L \subseteq M$ und sei $\varepsilon \notin L$, aber $\varepsilon \in M$. Welche der folgenden 3 Aussagen sind richtig? Begründen Sie kurz Ihre Antworten!
(a) $K^+ = K^* \backslash \{\varepsilon\}$ (b) $L^* \cup M^* = (L \backslash K)^* \cup M^*$ (c) $M^+ = M^* \backslash \{\varepsilon\}$

> **Lösung:**
> (a) ist richtig, da $\varepsilon \notin K$ ist.
> (b) ist richtig, da beide Seiten gleich M^* sind.
> (c) ist falsch, da $\varepsilon \in M$ ist.

(2) Gilt $L(((a|b)^*|c)^*) = L((a|(b|c)^*)^*)$? Begründen Sie kurz Ihre Antwort.

> **Lösung:**
> Ja, beide reguläre Ausdrücke erzeugen beliebige Wörter über $\Sigma = \{a, b, c\}$.

(3) Wie viele Myhill-Nerode-Äquivalenzklassen hat eine kontextfreie Sprache mindestens? Geben Sie auch ein Beispiel für eine Sprache mit dieser Anzahl an Äquivalenzklassen an.

> **Lösung:**
> 1, denn jede reguläre Sprache ist auch kontextfrei.
> Beispiel: Σ^* für ein endliches Alphabet Σ.

(4) Gibt es eine Grammatik $G = (\{S, A\}, \Sigma, P, S)$ in Chomsky-Normalform, sodass die Sprache $L(G)$
(a) von einem DFA (b) von einer Turingmaschine akzeptiert wird?
Begründen Sie Ihre Antworten!

> **Lösung:** Ja: Seien die Produktionen $S \rightarrow AA$, $A \rightarrow a$, dann ist $L(G) = \{a^2\}$ und die Sprache wird sowohl von einem DFA als auch von einer Turingmaschine akzeptiert (und die Grammatik ist in Chomsky-Normalform).

(5) Sei $L = \{awb \mid w \in \{a,b\}^*\} \subseteq \{a,b\}^*$. Geben Sie kontextfreie, aber nicht reguläre Sprachen L_1 und L_2 an mit $L_1 \subseteq L \subseteq L_2$.

Lösung:
$L_1 = \{a^n b^n \mid n \geq 1\}$
$L_2 = L \cup \{b^n a^n \mid n \geq 1\}$

(6) Geben Sie zwei Möglichkeiten an, zu zeigen, dass eine Sprache nicht regulär ist.

Lösung:

- Pumping-Lemma
- Zeige, dass die Sprache unendlich viele Myhill-Nerode ÄK induziert

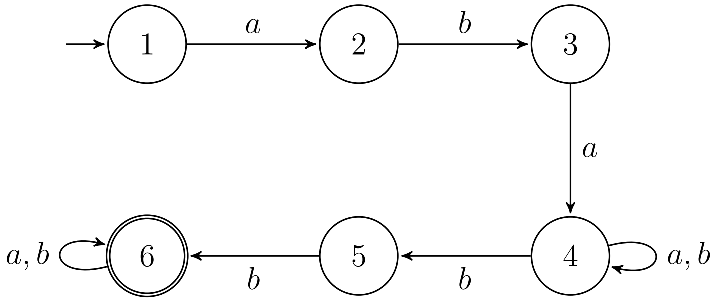
(7) Welche der folgenden Sprachklassen ist abgeschlossen unter Schnittbildung? Reguläre Sprachen (Typ 3), kontextfreie Sprachen (Typ 2), kontextsensitive Sprachen (Typ 1), rekursiv aufzählbare Sprachen (Typ 0), unäre Sprachen, endliche Sprachen.

Lösung: Alle außer Typ 2.

Aufgabe 2. (9 Punkte) Sei $\Sigma = \{a, b\}$. Geben Sie für jede der folgenden Sprachen einen endlichen Automaten und einen regulären Ausdruck an.

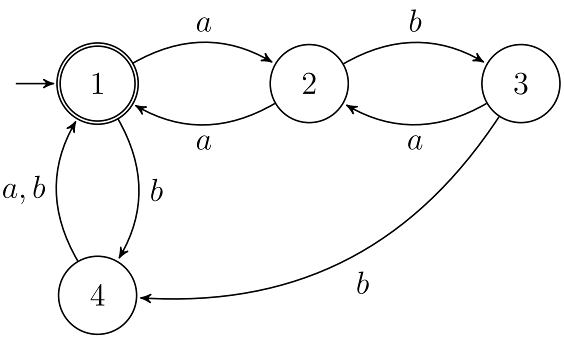
(a) $L_1 = \{w \in \Sigma^* \mid w$ beginnt mit aba und enthält das Teilwort $bb\}$

Lösung: $aba(a|b)^*bb(a|b)^*$



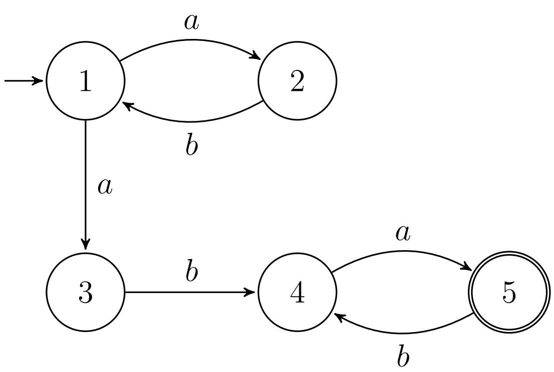
(b) $L_2 = \{w \in \Sigma^* \mid |w|$ ist gerade und w endet nicht mit $ab\}$

Lösung: $\varepsilon|(ab|aa|bb|ba)^*(aa|bb|ba)$



(c) $L_3 = \{(ab)^n a(ba)^m \mid n \geq 0$ und $m \geq 1\}$

Lösung: $(ab)^*aba(ba)^*$

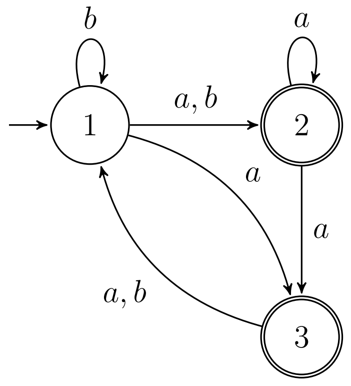


Aufgabe 3. (6 Punkte) Gegeben sei der NFA $M = (\{1, 2, 3\}, \{a, b\}, \delta, \{1\}, \{2, 3\})$, wobei δ gegeben ist durch:

δ	a	b
1	$\{2, 3\}$	$\{1, 2\}$
2	$\{2, 3\}$	$\emptyset$
3	$\{1\}$	$\{1\}$

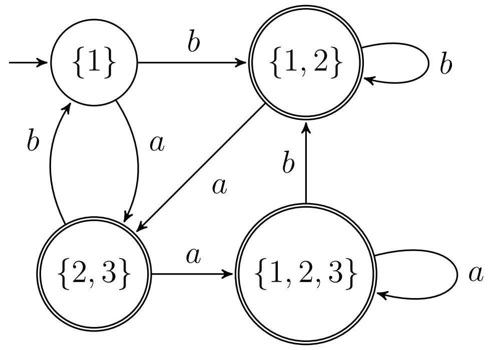
(a) Zeichnen Sie das zu M gehörige Automatendiagramm.

Lösung:

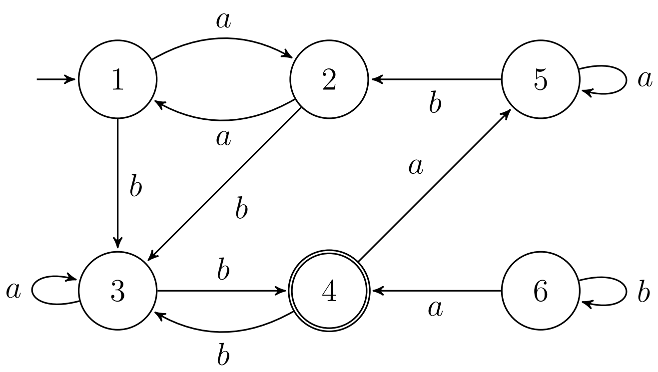


(b) Geben Sie mittels Potenzmengenkonstruktion einen zu M äquivalenten DFA an. Es genügt, den vom Startzustand erreichbaren Teil anzugeben.

Lösung:



Aufgabe 4. (6 Punkte) Minimieren Sie den folgenden DFA mit dem Algorithmus aus der Vorlesung. **Geben Sie an, welche Zustandspaare in welcher Reihenfolge markiert werden, und zeichnen Sie den erhaltenen minimalen DFA**.

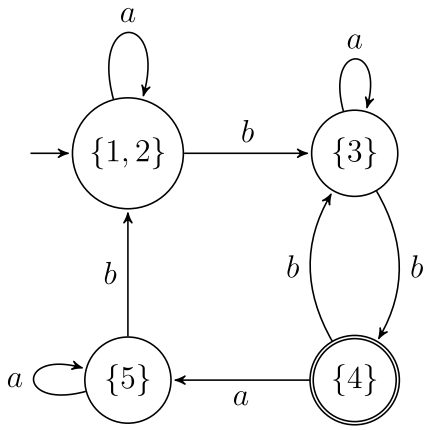


Lösung: Zustand 6 wird entfernt, weil nicht erreichbar.

2		-	-	-
3	2	3	-	-
4	1	1	1	-
5	6	5	4	1
	1	2	3	4

1. Trenne End- und Nichtendzustände.
2. Markiere $(3,1)$ wegen $\delta(3,b) = 4$ und $\delta(1,b) = 3$ und $(4,3)$ markiert.
3. Markiere $(3,2)$ wegen $\delta(3,a) = 3$ und $\delta(2,a) = 1$ und $(3,1)$ markiert.
4. Markiere $(5,3)$ wegen $\delta(5,b) = 2$ und $\delta(3,b) = 4$ und $(4,2)$ markiert.
5. Markiere $(5,2)$ wegen $\delta(5,b) = 2$ und $\delta(2,b) = 3$ und $(3,2)$ markiert.
6. Markiere $(5,1)$ wegen $\delta(5,b) = 2$ und $\delta(1,b) = 3$ und $(3,2)$ markiert.

1 und 2 sind erkennungsäquivalent.



Aufgabe 5. (6 Punkte) Gegeben sei folgende Sprache $L \subseteq \{a, b\}^*$ mit

$$L = \{a^n b^m \mid m \geq 0, n \geq m \text{ und } |n - m| \leq 7\}.$$

Zeigen Sie, dass L nicht regulär ist.

Lösung: Pumping Lemma mit $w = a^n b^n$:

Sei $n \in \mathbb{N}$ und sei $w = a^n b^n$. Dann ist $w \in L$ und $|w| \geq n$. Sei $w = xyz$ eine Zerlegung von w mit $|xy| \leq n$ und $|y| \geq 1$. Also gibt es $k, \ell \in \mathbb{N}$ mit $x = a^k$, $y = a^\ell$ und $z = a^{n-k-\ell} b^n$. Mit Pumpfaktor 0 erhalten wir

$$xy^0 z = a^k a^{n-k-\ell} b^n = a^{n-\ell} b^n.$$

Da $\ell \geq 1$ gilt nun $n - \ell < n$, also $xy^0 z \notin L$. Somit ist L nach dem Pumping Lemma nicht regulär.

Aufgabe 6. (6 Punkte) Gegeben sei der folgende nichtdeterministische Kellerautomat $M = (Z, \Sigma, \Gamma, \delta, z_0, \#)$, wobei $Z = \{z_0, z_1, z_2\}$, $\Sigma = \{a, b, c\}$, $\Gamma = \{\#, A\}$ und

(1) $\delta(z_0, a, \#) = \{(z_0, A)\}$
(2) $\delta(z_0, a, A) = \{(z_0, AA)\}$
(3) $\delta(z_0, b, A) = \{(z_1, \varepsilon)\}$
(4) $\delta(z_1, b, A) = \{(z_1, \varepsilon)\}$
(5) $\delta(z_1, c, A) = \{(z_2, \varepsilon)\}$
(6) $\delta(z_2, c, A) = \{(z_2, \varepsilon)\}$.

Die Modifikationen aus Teilaufgabe (b) wird für (c) nicht beibehalten.

(a) Welche Sprache wird von dem Kellerautomaten erkannt?

Lösung: $L(M) = \{a^n b^m c^{n-m} \mid n \geq m \geq 1\}$

(b) Anstelle von Transition (5) nehmen wir (5*) $\delta(z_1, c, A) = \{(z_2, A\#)\}$. Welche Sprache erkennt der Automat nun?

Lösung: $L(M) = \{a^n b^n \mid n \geq 1\}$

(c) Ersetze Transition (3) durch
(3*) $\delta(z_0, b, A) = \{(z_1, AA)\}$
und ersetze Transition (4) durch
(4*) $\delta(z_1, b, A) = \{(z_1, AA)\}$.
Welche Sprache erkennt der Kellerautomat nach dieser Modifikation?

Lösung: $L(M) = \{a^n b^m c^{n+m} \mid n, m \geq 1\}$

Aufgabe 7. (6 Punkte) Gegeben sei die Grammatik $G = (\{S, A, B, C\}, \{a, b, c\}, S, P)$ mit

$$\begin{aligned} P : S &\to BC \mid AC \mid BA \\ A &\to AA \mid BB \mid a \\ B &\to BA \mid b \\ C &\to AC \mid c. \end{aligned}$$

Testen Sie mit dem CYK-Algorithmus, ob das Wort $babbbc$ in $L(G)$ enthalten ist.

Lösung: Der CYK-Algorithmus liefert folgende Tabelle:

Länge	b	a	b	b	b	c
1	B	A	B	B	B	C
2	S, B	$\emptyset$	A	A	S	-
3	A	A	S, B	S, C	-	-
4	S, B	$\emptyset$	S	-	-	-
5	A	$\emptyset$	-	-	-	-
6	S, C	-	-	-	-	-

Also ist $babbbc \in L(G)$.

Aufgabe 8. (6 Punkte) Geben Sie für

$$L = \{w \in \{a,b,c,d\}^* \mid w = a^n b^{n+m} c^{m+k} d^k, n,m,k \geq 1\}$$

eine kontextfreie Grammatik an.

Lösung: Es gilt $L(G) = L$ mit $G = (\{S,N,M,K\},\{a,b,c,d\},P,S)$, wobei

$$\begin{aligned} P: \quad & S \to NMK \\ & N \to aNb \mid ab \\ & M \to bMc \mid bc \\ & K \to cKd \mid cd. \end{aligned}$$

Aufgabe 9. (6 Punkte) Sei $M = (\{z_0, z_1, z_f\}, \{a, b\}, \{a, b, \Box\}, \delta, z_0, \Box, \{z_f\})$ eine deterministische Turingmaschine mit

$$\begin{aligned}
\delta(z_0, a) &= (z_1, a, R)\\
\delta(z_0, b) &= (z_0, b, R)\\
\delta(z_1, a) &= (z_0, a, R)\\
\delta(z_1, b) &= (z_1, b, R)\\
\delta(z_0, \Box) &= (z_0, \Box, N)\\
\delta(z_1, \Box) &= (z_f, \Box, N)
\end{aligned}$$

(a) Akzeptiert M das Wort $w_1 = abbba$? ◯ ja $\surd$ **nein**

(b) Akzeptiert M das Wort $w_2 = abbbaab$? $\surd$ **ja** ◯ nein

(c) Welche Sprache akzeptiert M?

Lösung: $L(M) = \{w \in \Sigma^* \mid \text{Anzahl a's in } w \text{ ist ungerade}\}$

Aufgabe 10. (5 Punkte (Bonus)) Sei $\Sigma = \{a, b\}$ und sei $\Sigma_\# = \Sigma \cup \{\#\}$. Weiterhin sei $\Sigma_\#^2$ die Menge aller 2-Tupel über $\Sigma_\#$, d.h.

$$\Sigma_\#^2 = \{(a,a), (a,b), (b,a), (b,b), (a,\#), (\#,a), (b,\#), (\#,b), (\#,\#)\}.$$

Seien $w_1 = x_1 \dots x_\ell$ und $w_2 = y_1 \dots y_k$ Wörter in Σ^*, also $x_i, y_j \in \Sigma$ für $1 \leq i \leq \ell$ und $1 \leq j \leq k$ sowie $\ell, k \geq 0$. Die *Konvolution* $w_1 \otimes w_2 \in (\Sigma_\#^2)^*$ der Wörter w_1 und w_2 ist definiert als

$$w_1 \otimes w_2 = \begin{cases} (x_1, y_1) \cdots (x_k, y_k)(x_{k+1}, \#) \cdots (x_\ell, \#) & \text{falls } \ell > k, \\ (x_1, y_1) \cdots (x_\ell, y_\ell)(\#, y_{\ell+1}) \cdots (\#, y_k) & \text{falls } \ell < k, \\ (x_1, y_1) \cdots (x_\ell, y_\ell) & \text{falls } \ell = k. \end{cases}$$

Die Konvolution ist also ein Wort über dem Alphabet $\Sigma_\#^2$. Ein 2-Tupel entspricht nun *einem* Zeichen eines solchen Worts. Zum Beispiel erhalten wir

$$\begin{aligned} aa \otimes bbbb &= (a,b)(a,b)(\#,b)(\#,b) \\ abab \otimes a &= (a,a)(b,\#)(a,\#)(b,\#) \\ bb \otimes bb &= (b,b)(b,b). \end{aligned}$$

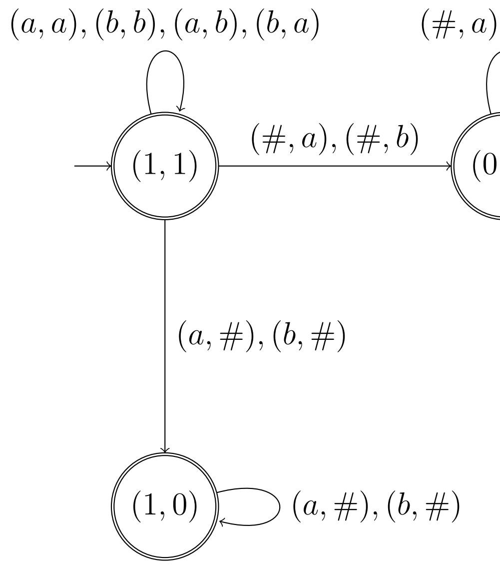
(a) Berechnen Sie die Konvolution $aab \otimes bababb$.

Lösung: $aab \otimes bababb = (a,b)(a,a)(b,b)(\#,a)(\#,b)(\#,b)$

(b) Die Sprache

$$L = \{w \otimes v \mid w, v \in \Sigma^*\}$$

aller Konvolutionen zweier Wörter aus Σ^* ist regulär. Geben Sie einen endlichen Automaten für L an. Beachten Sie, dass Wörter wie $(a,a)(\#,\#)$ oder $(a,b)(\#,b)(b,\#)$ *nicht* zu L gehören.



Lösung: