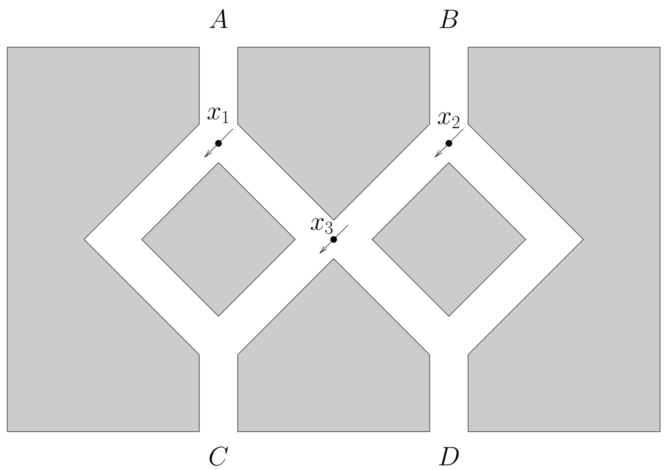
Universität Freiburg
PD Dr. A. Jakoby
Sommer 2007

Übungen zum Repetitorium
Informatik III

Übungsaufgaben zu Formalen Sprachen und Automaten

1. Untersuchen Sie das folgende Spiel:



Eine Murmel wird in A oder B fallen gelassen. Die Hebel x_1, x_2, x_3 lassen die Murmel entweder nach rechts oder nach links fallen, je nachdem in welche Richtung der Pfeil zeigt. Immer wenn eine Murmel auf einen Hebel trifft, veranlasst sie den Hebel, seinen Zustand zu wechseln, so dass die Murmel, die als nächstes auf den Hebel trifft, den alternativen Weg nehmen muss.

   (a) Modellieren Sie dieses Spiel mit Hilfe eines endlichen Automaten. Bezeichnen Sie hierbei eine Murmel, die durch A fällt, durch eine 0-Eingabe und eine Murmel, die durch B fällt, durch eine 1-Eingabe. Ein Eingabewort wird akzeptiert, wenn die letzte Murmel bei D herauskommt.

   (b) Beschreiben Sie die durch diesen Automaten akzeptierte Sprache.

2. Geben Sie deterministische endliche Automaten für die folgenden Sprachen über dem Alphabet $\{0, 1\}$ an.

   (a) Die Menge aller Zeichenketten, die mit 00 enden.

   (b) Die Menge aller Zeichenketten mit drei aufeinander folgenden Nullen.

   (c) Die Menge aller Zeichenketten, in denen jeder Block von fünf aufeinanderfolgenden Zeichen zumindest zwei Nullen enthält.

   (d) Die Menge aller Zeichenketten, deren zehntletztes Symbol eine Eins ist.

3. Für $k \in \mathbb{N}$ definiere

$$N_k \quad := \quad \{ \; w \in \{0,1\}^* \mid \text{sei } v \in \mathbb{N} \text{ die Zahl mit Binärdarstellung } w, \text{ dann ist } v \bmod k = 0 \; \} \; .$$

Zeigen Sie:

(a) Für jedes feste k ist N_k eine reguläre Sprache.

(b) Für jede Teilmenge $I \subset \mathbb{N}$ mit konstanter Größe ist die Sprache

$$\bigcup_{i \in I} \{ \; \mathrm{bin}(i) \# w \mid w \in N_i \; \}$$

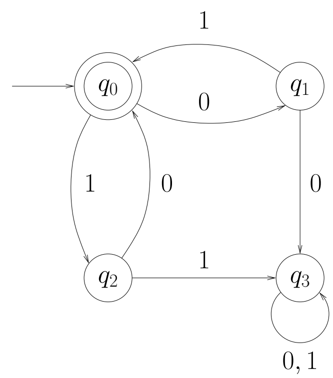
eine reguläre Sprache, wobei $\mathrm{bin}(i)$ die Binärdarstellung von i ist.

(c) Die Sprache

$$\bigcup_{i \in \mathbb{N}} \{ \; \mathrm{bin}(i) \# w \mid w \in N_i \; \}$$

ist nicht regulär.

4. Betrachten Sie den folgenden endlichen Automaten:



(a) Beweisen Sie, dass dieser DFA die Menge aller Zeichenketten über dem Alphabet $\{0,1\}$ akzeptiert, die eine gleiche Anzahl von Nullen und Einsen enthalten, und für die jeder Präfix höchstens eine Null mehr hat als Einsen bzw. eine Eins mehr hat als Nullen.

(b) Untersuchen Sie, ob dieser DFA minimal ist.

(c) Können Sie einen NFA für diese Sprache konstruieren, der mit drei Zuständen auskommt?

5. Betrachten Sie die folgenden beiden NFAs:

(a) $(\{p,q,r,s\}, \{0,1\}, \delta_1, p, \{s\})$ mit

	0	1
p	p,q	p
q	r	r
r	s	-
s	s	s

(b) $(\{p,q,r,s\}, \{0,1\}, \delta_2, p, \{q,s\})$ mit

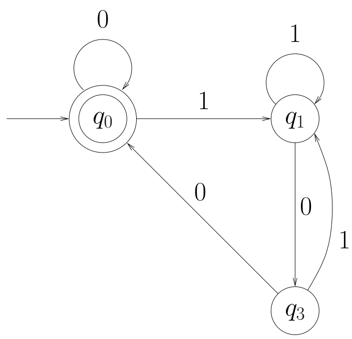
	0	1
p	p,q	q
q	r	q,r
r	s	p
s	-	p

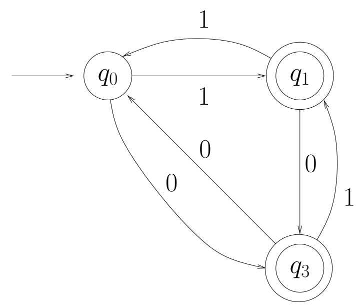
Geben Sie zu jedem dieser Automaten einen äquivalenten DFA an.

6. Konstruieren Sie einen endlichen Automaten, die zu den folgenden regulären Ausdrücken äquivalent sind:

   (a) $10 \cup (0 \cup 11)0^*1$
   (b) $01(((10)^* \cup 111)^* \cup 0)^*1$
   (c) $((0 \cup 1)(0 \cup 1))^* \cup ((0 \cup 1)(0 \cup 1)(0 \cup 1))^*$

7. Konstruieren Sie reguläre Ausdrücke, die den Automaten der beiden folgenden Zustandsdiagrammen entsprechen.





8. Beweisen oder widerlegen Sie die folgenden Gleichungen für reguläre Ausdrücke r, s, t:

   (a) $(\varepsilon \cup r)^* \;=\; r^*$
   (b) $(rs \cup r)^*r \;=\; r(sr \cup r)^*$
   (c) $s(rs \cup s)^*r \;=\; rr^*s(rr^*s)^*$
   (d) $(r \cup s)^* \;=\; r^* \cup s^*$

9. Welche der folgenden Sprachen sind regulär und welche sind kontextfrei? Beweisen Sie Ihre Antwort.

   (a) $\{\ 0^{2^n} \mid n \geq 1\ \}$
   (b) $\{\ 0^m 1^n 0^{m+n} \mid m, n \geq 1\ \}$
   (c) $\{\ 0^p \mid p$ ist eine Primzahl $\}$
   (d) Die Menge aller Zeichenketten über dem Alphabet $\{0, 1\}$, die keine drei aufeinanderfolgenden Nullen haben.
   (e) Die Menge aller Primzahlen kleiner $2^{2^{100}}$.
   (f) Die Menge aller Zeichenketten über dem Alphabet $\{0, 1, 2\}$, die die gleiche Anzahl von Nullen und Einsen haben.
   (g) $\{\ x \in \{0,1\}^* \mid x$ in $L((0 \cup 1)^*)$ und $x = x^{\mathrm{rev}}\ \}$ wobei x^{rev} die rückwärts gelesene Zeichenkette x bezeichnet.
   (h) $\{\ x \in \{0,1\}^* \mid x$ in $L(0 \cup 1)$ und $x = x^{\mathrm{rev}}\ \}$
   (i) $\{xwx^{\mathrm{rev}} \mid x, w \in \{0,1\}^* \setminus \{\varepsilon\}\ \}$

10. Sei $L \subseteq \{0,1\}^*$ eine reguläre Sprache. Welche der folgenden Sprachen ist dann auch regulär? Begründen Sie Ihre Antwort.

   (a) $\{\ a_1 a_3 \ldots a_{2n-1} \mid a_1 a_2 \ldots a_{2n} \in L$ und $a_1, a_2, \ldots, a_{2n} \in \{0,1\}\ \}$
   (b) $\{\ x \in L \mid \forall y \in \{0,1\}^* \setminus \{\varepsilon\} :\ xy \notin L\ \}$
   (c) $\{\ x \in L \mid$ für alle $y \in L$ ist y kein Präfix von $x\ \}$
   (d) $\{\ x^{\mathrm{rev}} \mid x \in L\ \}$

11. Betrachten Sie die Sprache

$$L \;=\; \{\ 0^n 1^n \mid n \in \mathbb{N}\ \}\ .$$

   (a) Wie lauten die Äquivalenzklassen im Satz von Myhill-Nerode für L.
   (b) Verwenden Sie die Antwort aus dem ersten Teil dieser Aufgabe um zu zeigen, dass L nicht regulär ist.

12. Geben Sie für die folgenden Mengen jeweils eine erzeugende kontextfreie Grammatik an.
   (a) Die Menge aller korrekten Klammerausdrücke.
   (b) Die Menge aller Zeichenketten über dem Alphabet $\{a, b\}$ mit genau doppelt so vielen as wie bs.
   (c) Die Menge aller Zeichenketten über dem Alphabet $\{a, b, \cup, ^*, (,), \epsilon, \emptyset\}$, die wohlgeformte reguläre Ausdrücke über dem Alphabet $\{a, b\}$ sind. Beachten Sie, dass wir hierbei zwischen ε als leere Zeichenkette der Grammatik und dem ϵ als leere Zeichenkette in einem regulären Ausdruck (d.h. in einem Wort der Sprache) unterscheiden müssen.
   (d) Die Menge aller Zeichenketten über dem Alphabet $\{a, b\}$, die nicht von der Form ww sind.
   (e) $\{ \ a^i b^j c^k \mid i \neq j \text{ oder } j \neq k\}$.

13. Betrachten Sie die Grammatik G mit den Produktionen

$$S \longrightarrow aS \mid aSbS \mid \varepsilon \ .$$

   Beweisen Sie, dass

$$L(G) \ = \ \{ \ x \in \{a, b\}^* \mid \text{jedes Präfix von } x \text{ hat mindestens soviel } a\text{s wie } b\text{s} \ \} \ .$$

14. Die Grammatik

$$E \longrightarrow E + E \mid E \cdot E \mid (E) \mid \mathbf{id}$$

   erzeugt die Menge der arithmetischen Ausdrücke mit $+$, $\cdot$, Klammern und **id**. Diese Grammatik ist mehrdeutig, da $\mathbf{id} + \mathbf{id} \cdot \mathbf{id}$ durch zwei verschiedene Ableitungen erzeugt werden kann, d.h. durch zwei verschiedene Ableitungen mit unterschiedlichen Ableitungsbäumen.
   (a) Konstruieren Sie eine äquivalente nicht mehrdeutige Grammatik.
   (b) Konstruieren Sie eine nicht mehrdeutige Grammatik für alle arithmetischen Ausdrücke ohne redundante Klammern. Eine Menge von Klammern nennen wir redundant, wenn ihr Entfernen den Ausdruck (bzw. dessen Wert) nicht verändert.

15. G sei eine kontextfreie Grammatik, die wohlgeformte aussagenlogische Formeln mit den Prädikaten p und q erzeugt. Für die Produktionen erhalten wir

$$S \longrightarrow \neg S \mid [S \rightarrow S] \mid p \mid q \ .$$

   Die Terminale sind $p, q, \neg, [,]$ und $\rightarrow$.
   (a) Bestimmen Sie eine Grammatik für $L(G)$ in Chomsky-Normalform.
   (b) Zeigen Sie unter Anwendung des CYK-Algorithmus, dass $\neg\neg[\neg q \rightarrow p]$ in der Sprache $L(G)$ ist.
   (c) Untersuchen Sie ob $\neg\neg[q\neg \rightarrow p]$ in $L(G)$ ist. Beweisen Sie Ihre Antwort.

16. Konstruieren Sie eine Kellerautomaten, der die Sprache der Grammatik mit dem Startsymbol S und den folgenden Produktionen akzeptiert:

$$\begin{aligned} S &\longrightarrow aAA \\ A &\longrightarrow aS \mid bS \mid a \ . \end{aligned}$$

17. Geben Sie eine kontextfreie Grammatik für die Sprache $L(A)$ für den PDA

$$A \ = \ (\{q_0, q_1, q_2, q_3, q_4\}, \{0, 1\}, \{Z_0, X\}, \delta, q_0, \{q_4\})$$

   mit

$$\begin{aligned} \delta(q_0, \varepsilon, \varepsilon) &= \{(q_1, Z_0)\} \\ \delta(q_1, \varepsilon, Z_0) &= \{(q_4, Z_0)\} \\ \delta(q_1, 1, Z_0) &= \{(q_2, Z_0)\} \\ \delta(q_1, 1, X) &= \{(q_2, X)\} \\ \delta(q_1, 0, X) &= \{(q_3, X)\} \\ \delta(q_2, \varepsilon, \varepsilon) &= \{(q_1, X)\} \\ \delta(q_3, 1, X) &= \{(q_3, \varepsilon)\} \\ \delta(q_3, 0, Z_0) &= \{(q_1, Z_0)\} \end{aligned}$$

   an.

18. Untersuchen Sie, ob die folgenden Mengen keine kontextfreien Sprachen sind.

    (a) $\{ \ a^i b^j c^k \mid i < j < k\}$
    (b) $\{ \ a^i b^j \mid j = i^2 \ \}$
    (c) Die Menge der Zeichenketten aus a, b, c, die von allen jeweils die gleiche Anzahl enthalten.
    (d) $\{ \ a^i b^i c^j \mid i \leq j \leq 2i\}$
    (e) $\{ \ a^i b^j \mid i \neq j$ und $i \neq 2j \ \}$
    (f) $\{ \ b_i \# b_{i+1} \mid b_i$ ist die Binärdarstellung von i und $i \geq 1 \ \}$
    (g) $\{ \ ww^{\text{rev}}w \mid w \in \{0,1\}^* \ \}$
    (h) $\{ \ wxw \mid w, x \in \{0,1\}^* \ \}$
    (i) $\{ \ ww \mid w, x \in \{0,1\}^* \ \}$

Ein großer Teil dieser Aufgaben ist dem Buch von J. E. Hopcroft, J. D. Ullman, *Einführung in die Automatentheorie, Formale Sprachen und Komplexitätstheorie*, Addison-Wesley (2. Auflage), entnommen.