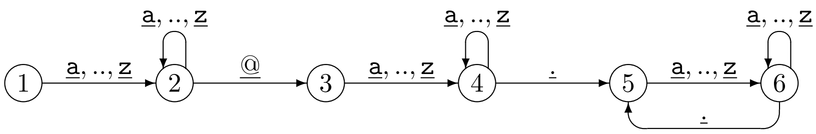
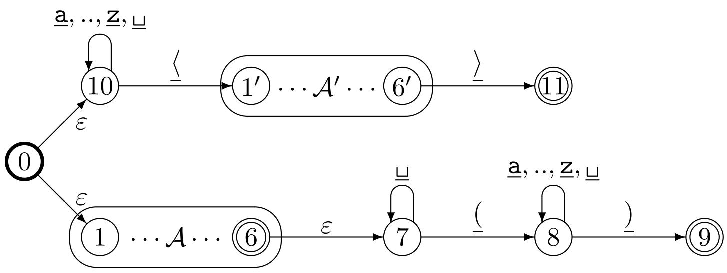
Theoretische Informatik 1 – Übungsblatt 3 (SS2006) Lösungen

Aufgabe 3.1 Geben Sie einen endlichen Automaten an, der die Sprache $\mathcal{E}$ der Email-Adressen akzeptiert, wie sie in Aufgabe 1.5 beschrieben ist. Es genügt die graphische Darstellung des Automaten.

Lösung Wir beschreiben zunächst die eigentlichen Email-Adressen durch den Automaten $\mathcal{A}$:

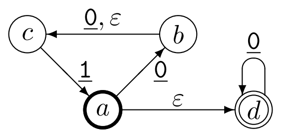


Sei $\mathcal{A}'$ eine Kopie des Automaten $\mathcal{A}$, wobei die Zustände in $1'$, …, $6'$ umbenannt werden. Ein (indeterministischer) Automat für die gesamte Sprache ist dann:



Aufgabe 3.2 Geben Sie einen minimalen deterministischen Automaten an, der die durch den `egrep`-Ausdruck `(00?1)*0*` beschriebene Sprache akzeptiert. Verwenden Sie zum Nachweis der Minimalität eine Unterscheidbarkeitstabelle.

Lösung Wir konstruieren je einen Automaten für die Sprachen `(00?1)*` und `0*` und verbinden sie dann mit einer ε-Kante:



Nun berechnen wir die erweiterte Übergangsfunktion δ^* dieses Automaten sowie die Übergangsfunktion $\hat{\delta}$ des deterministischen Automaten:

δ^*	0	1
a	$\{b,c,d\}$	$\{\}$
b	$\{c\}$	$\{a,d\}$
c	$\{\}$	$\{a,d\}$
d	$\{d\}$	$\{\}$

$\hat{\delta}$	0	1
$\{a\}$	$\{b,c,d\}$	$\{\}$
$\{b,c,d\}$	$\{c,d\}$	$\{a,d\}$
$\{\}$	$\{\}$	$\{\}$
$\{c,d\}$	$\{d\}$	$\{a,d\}$
$\{a,d\}$	$\{b,c,d\}$	$\{\}$
$\{d\}$	$\{d\}$	$\{\}$

Endzustände sind alle Mengen, die den ursprünglichen Endzustand d enthalten – $\{b,c,d\}$, $\{c,d\}$, $\{a,d\}$ und $\{d\}$ – sowie der Startzustand $\{a\}$, da der ursprüngliche Automat auch das Leerwort akzeptiert. Ehe wir nun die Unterscheidbarkeitstabelle erstellen, führen wir einen Vereinfachungsschritt durch. Die Zustände $\{a\}$ und $\{a,d\}$ sind beides Endzustände und daher nicht durch das

Leerwort unterscheidbar; sie können aber auch durch keine anderen Wörter unterschieden werden, da sie sowohl für $\underline{0}$ als auch für $\underline{1}$ identische Folgezustände aufweisen. Daher sind die beiden Zustände ununterscheidbar und können zu einem zusammengefasst werden, den wir $\{a\}$ nennen.

	$\hat{\delta}'$	$\underline{0}$	$\underline{1}$
EZ	$\{a\}$	$\{b,c,d\}$	$\{\}$
EZ	$\{b,c,d\}$	$\{c,d\}$	$\{a\}$
	$\{\}$	$\{\}$	$\{\}$
EZ	$\{c,d\}$	$\{d\}$	$\{a\}$
EZ	$\{d\}$	$\{d\}$	$\{\}$

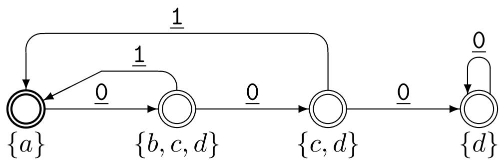
EZ … Endzustand

{b,c,d}	$\underline{1}$			
{}	ε	ε		
{c,d}	$\underline{1}$	$\underline{0}\,\underline{1}$	ε	
{d}	$\underline{0}\,\underline{1}$	$\underline{1}$	ε	$\underline{1}$
	{a}	{b,c,d}	{}	{c,d}

Alle Felder in der Unterscheidbarkeitstabelle sind markiert, es können also keine weiteren Zustände zusammengefasst werden. Wir erhalten den minimalen Automaten

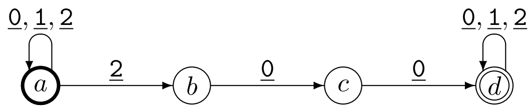
$$\langle\{\{a\},\ \{b,c,d\},\ \{\},\ \{c,d\},\ \{d\}\},\{\underline{0},\underline{1}\},\hat{\delta}',\{a\},\{\{a\},\ \{b,c,d\},\ \{c,d\},\ \{d\}\}\rangle$$

bzw. in graphischer Darstellung



Aufgabe 3.3 Sei L die Menge aller Wörter über dem Alphabet $\{\underline{0},\underline{1},\underline{2}\}$, die das Wort $\underline{2}\underline{0}\underline{0}$ *nicht* enthalten. Geben Sie einen endlichen Automaten für L an. Es genügt die graphische Darstellung des Automaten.
Hinweis: Ein Lösungsweg besteht darin, zuerst einen deterministischen Automaten für das Komplement der Sprache L zu konstruieren und diesen dann entsprechend anzupassen.

Lösung Wir konstruieren zunächst einen Automaten für jene Sprache, die aus allen Wörtern besteht, die das Wort $\underline{2}\underline{0}\underline{0}$ enthalten:



und determinisieren ihn dann:

δ^*	$\underline{0}$	$\underline{1}$	$\underline{2}$
a	$\{a\}$	$\{a\}$	$\{a,b\}$
b	$\{c\}$	$\{\}$	$\{\}$
c	$\{d\}$	$\{\}$	$\{\}$
d	$\{d\}$	$\{d\}$	$\{d\}$

	$\hat{\delta}$	$\underline{0}$	$\underline{1}$	$\underline{2}$
	$\{a\}$	$\{a\}$	$\{a\}$	$\{a,b\}$
	$\{a,b\}$	$\{a,c\}$	$\{a\}$	$\{a,b\}$
	$\{a,c\}$	$\{a,d\}$	$\{a\}$	$\{a,b\}$
EZ	$\{a,d\}$	$\{a,d\}$	$\{a,d\}$	$\{a,b,d\}$
EZ	$\{a,b,d\}$	$\{a,c,d\}$	$\{a,d\}$	$\{a,b,d\}$
EZ	$\{a,c,d\}$	$\{a,d\}$	$\{a,d\}$	$\{a,b,d\}$

EZ … Endzustand

Die drei Endzustände haben die Eigenschaft, dass man mit jedem beliebigen Wort wieder einen Endzustand erreicht; sie sind somit nicht unterscheidbar und können zusammengefasst werden:

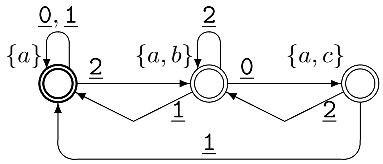
	$\hat{\delta}'$	$\underline{0}$	$\underline{1}$	$\underline{2}$
	$\{a\}$	$\{a\}$	$\{a\}$	$\{a,b\}$
	$\{a,b\}$	$\{a,c\}$	$\{a\}$	$\{a,b\}$
	$\{a,c\}$	$\{a,d\}$	$\{a\}$	$\{a,b\}$
EZ	$\{a,d\}$	$\{a,d\}$	$\{a,d\}$	$\{a,d\}$

Um nun den gesuchten Automat für die komplementäre Sprache zu erhalten, machen wir Endzustände zu Nichtendzuständen und umgekehrt. Laut Unterscheidbarkeitstabelle können keine Zustände zusammengefasst werden.

	$\hat{\delta}''$	$\underline{0}$	$\underline{1}$	$\underline{2}$
EZ	$\{a\}$	$\{a\}$	$\{a\}$	$\{a,b\}$
EZ	$\{a,b\}$	$\{a,c\}$	$\{a\}$	$\{a,b\}$
EZ	$\{a,c\}$	$\{a,d\}$	$\{a\}$	$\{a,b\}$
	$\{a,d\}$	$\{a,d\}$	$\{a,d\}$	$\{a,d\}$

{a,b}	$\underline{0}\,\underline{0}$		
{a,c}	$\underline{0}$	$\underline{0}$	
{a,d}	ε	ε	ε
	{a}	{a,b}	{a,c}

Der Zustand $\{a,d\}$ fungiert als „Falle“, die wir in der graphischen Darstellung weglassen.



Aufgabe 3.4 Sei v^R die Spiegelung von v, also z.B. $(\underline{1}\,\underline{0}\,\underline{0})^R = \underline{0}\,\underline{0}\,\underline{1}$ oder $(\underline{1}\,\underline{0}\,\underline{1}\,\underline{0}\,\underline{1}\,\underline{1}\,\underline{1})^R = \underline{1}\,\underline{1}\,\underline{1}\,\underline{0}\,\underline{1}\,\underline{0}\,\underline{1}$.

a) Sei L die Sprache $\{vv^R \mid v \in \{\underline{0},\underline{1}\}^+\}$. Ist L regulär? Begründen Sie Ihre Antwort.

Lösung L ist nicht regulär, wie wir mit Hilfe des Pumping Lemmas in der Spielvariante zeigen.

1) Der Opponent wählt eine Schranke m.
2) Der Proponent wählt das Wort $w = \underline{0}^m\underline{1}\,\underline{1}\,\underline{0}^m$. Es liegt in der Sprache, da es die Form vv^R besitzt, wobei $v = \underline{0}^m\underline{1}$ gilt. Weiters erfüllt das Wort die Längenbedingung: $|w| = 2m + 2 \geq m$.
3) Der Opponent zerlegt w nun in drei Teile xyz mit der Einschränkung, dass $|xy| \leq m$ und $|y| \geq 1$ gelten muss. Da der erste Nuller-Block die Länge m besitzt, können x und y nur aus Nullern bestehen:

$$\begin{array}{rcll} x & = & \underline{0}^j & 0 \leq j < m \\ y & = & \underline{0}^k & 0 < k \leq m \\ z & = & \underline{0}^{m-j-k}\underline{1}\,\underline{1}\,\underline{0}^m & \end{array}$$

4) Der Proponent muss nun ein $i \geq 0$ finden, sodass w_i nicht in L liegt. Es gilt:

$$w_i = xy^iz = \underline{0}^j(\underline{0}^k)^i\underline{0}^{m-j-k}\underline{1}\,\underline{1}\,\underline{0}^m = \underline{0}^{m+k(i-1)}\underline{1}\,\underline{1}\,\underline{0}^m$$

Der Proponent kann z.B. $i = 0$ wählen: $w_0 = \underline{0}^{m-k}\underline{1}\,\underline{1}\,\underline{0}^m$. Dieses Wort besitzt nicht die Form vv^R, es liegt also nicht in L.

Der Proponent kann daher das Spiel immer gewinnen, er besitzt eine Gewinnstrategie. Somit ist die Sprache nicht regulär.

b) Sei L die Sprache $\{uvv^Rw \mid u, v, w \in \{\underline{0},\underline{1}\}^+\}$. Ist L regulär? Begründen Sie Ihre Antwort.

Lösung L ist regulär. Jedes Wort aus L hat die Eigenschaft, dass es an der Stelle, wo v und v^R aufeinandertreffen, die Zeichenfolge $\underline{0}\,\underline{0}$ oder $\underline{1}\,\underline{1}$ enthält (letztes Zeichen von v und erstes Zeichen von v^R). Umgekehrt kann jedes Wort, das die Zeichenfolge $\underline{0}\,\underline{0}$ oder $\underline{1}\,\underline{1}$ enthält, als Wort aus L aufgefasst werden: die Zeichenfolge entspricht vv^R, alles davor dem Wort u, alles danach dem Wort w. Somit ist L genau die Menge aller Wörter, die $\underline{0}\,\underline{0}$ oder $\underline{1}\,\underline{1}$ enthalten, d.h., $L = \{\underline{0},\underline{1}\}^+ \cdot \{\underline{0}\,\underline{0}, \underline{1}\,\underline{1}\} \cdot \{\underline{0},\underline{1}\}^+$.

Aufgabe 3.5 Zeigen Sie: die Sprache $L = \{(\underline{\mathtt{a}}\,\underline{\mathtt{b}})^{n_1}(\underline{\mathtt{a}}\,\underline{\mathtt{c}})^{n_2} \mid |n_1 - n_2| \leq 5\}$ ist nicht regulär. ($|n|$ bezeichnet den Absolutbetrag der Zahl n.)

Lösung Wir vereinfachen die Sprache zunächst, indem wir einen Homomorphismus anwenden. Reguläre Sprachen sind abgeschlossen unter Homomorphismen, d.h.: Ist h ein Homomorphismus und L eine reguläre Sprache, dann ist auch $h(L)$ eine reguläre Sprache. Der Umkehrschluss ergibt: Ist $h(L)$ keine reguläre Sprache, ist auch L keine reguläre Sprache.
Wir verwenden den Homomorphismus $h(\underline{\mathtt{a}}) = \varepsilon$, $h(\underline{\mathtt{b}}) = \underline{\mathtt{b}}$ und $h(\underline{\mathtt{c}}) = \underline{\mathtt{c}}$. Wir erhalten:

$$\begin{aligned}
h(L) &= h(\{(\underline{\mathtt{a}}\,\underline{\mathtt{b}})^{n_1}(\underline{\mathtt{a}}\,\underline{\mathtt{c}})^{n_2} \mid |n_1 - n_2| \leq 5\}) \\
&= \{h((\underline{\mathtt{a}}\,\underline{\mathtt{b}})^{n_1}(\underline{\mathtt{a}}\,\underline{\mathtt{c}})^{n_2}) \mid |n_1 - n_2| \leq 5\} \\
&= \{(h(\underline{\mathtt{a}})h(\underline{\mathtt{b}}))^{n_1}(h(\underline{\mathtt{a}})h(\underline{\mathtt{c}}))^{n_2}) \mid |n_1 - n_2| \leq 5\} \\
&= \{\underline{\mathtt{b}}^{n_1}\underline{\mathtt{c}}^{n_2} \mid |n_1 - n_2| \leq 5\} \\
&= L'
\end{aligned}$$

Wir zeigen die Nichtregularität von L' mittels des Pumping Lemmas.

1. Der Opponent wählt eine Schranke m.

2. Der Proponent wählt das Wort $w = \underline{\mathtt{b}}^m\underline{\mathtt{c}}^m$. Es liegt in der Sprache, da $|m - m| = 0 \leq 5$ gilt. Weiters erfüllt das Wort die Längenbedingung: $|w| = 2m \geq m$.

3. Der Opponent zerlegt w nun in drei Teile xyz mit der Einschränkung, dass $|xy| \leq m$ und $|y| \geq 1$ gelten muss. Da der erste $\underline{\mathtt{b}}$-Block die Länge m besitzt, können x und y nur aus $\underline{\mathtt{b}}$s bestehen:
$$\begin{aligned}
x &= \underline{\mathtt{b}}^j && 0 \leq j < m \\
y &= \underline{\mathtt{b}}^k && 0 < k \leq m \\
z &= \underline{\mathtt{b}}^{m-j-k}\underline{\mathtt{c}}^m &&
\end{aligned}$$

4. Der Proponent muss nun ein $i \geq 0$ finden, sodass w_i nicht in L liegt. Es gilt:
$$w_i = xy^iz = \underline{\mathtt{b}}^j(\underline{\mathtt{b}}^k)^i\underline{\mathtt{b}}^{m-j-k}\underline{\mathtt{c}}^m = \underline{\mathtt{b}}^{m+k(i-1)}\underline{\mathtt{c}}^m$$
Die Zahl i muss so gewählt werden, dass $|m + k(i-1) - m| \not\leq 5$ gilt, dass also $|k(i-1)| > 5$ gilt. Wir wählen $i = 7$: Wegen $k \geq 1$ gilt $k(7-1) \geq 1 \cdot (7-1) = 6 > 5$, d.h., $w_7 \notin L'$.

Der Proponent kann das Spiel immer gewinnen, er besitzt eine Gewinnstrategie. Somit ist die Sprache $L' = h(L)$ nicht regulär, daher ist auch L nicht regulär.