Technische Universität München
Fakultät für Informatik
Prof. Tobias Nipkow, Ph.D.
Sascha Böhme, Lars Noschinski

Sommersemester 2011
Lösungsblatt 4
20. Juni 2011

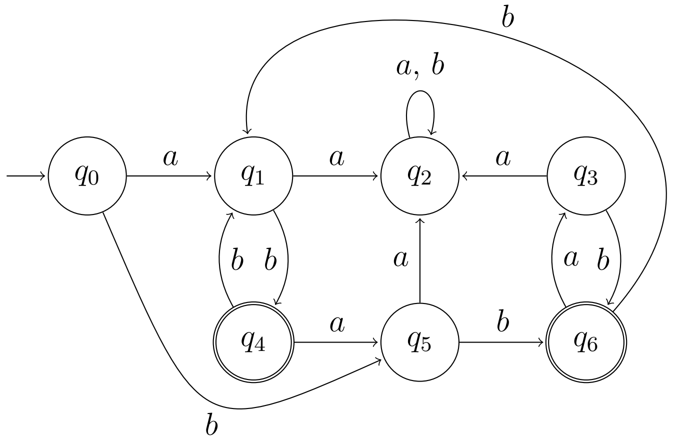
Einführung in die Theoretische Informatik

Hinweis: *Bitte beachten Sie unbedingt die Hinweise zum Übungsablauf und zu den Aufgabentypen auf der THEO-Website (`http://theo.in.tum.de/`).*

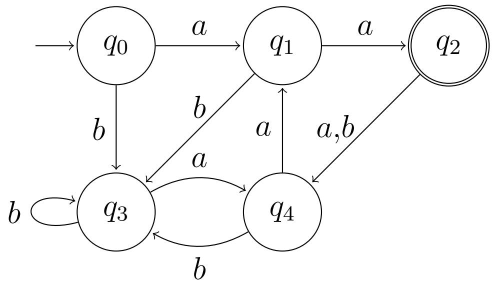
Hausaufgabe 1 (6 Punkte)

Wir betrachten die beiden folgenden deterministischen Automaten.

1.



2.



Verwenden Sie das in der Vorstellung vorgestellte Verfahren, um diese Automaten zu minimieren. Gehen Sie dabei wie folgt vor:

1. Stellen Sie die Tabelle aus der Vorlesung auf und geben Sie zu jedem unterscheidbaren Paar von Zuständen an, mit welchem Zeichen (oder ϵ) sie unterschieden werden können.

2. Verwenden Sie die aufgestellte Tabelle, um den Quotientenautomat zu konstruieren.

Lösungsvorschlag

1. Die Tabellen sehen folgendermaßen aus. Dabei ist eine Zelle mit i/c beschriftet, wenn die beiden Zustände mit dem Buchstaben c durch ein Zustandspaar, dass mit einer Zahl $j < i$ beschriftet ist, unterschieden werden können. Mit X sind diejenigen Zellen beschriftet, die wegen der Endzustände unterschieden werden können

1)

	0	1	2	3	4	5
1	$1/b$					
2	$2/a$	$1/b$				
3	$1/b$		$1/b$			
4	X	X	X	X		
5	$1/b$		$1/b$		X	
6	X	X	X	X		X

2)

	0	1	2	3
1	$1/a$			
2	X	X		
3	$2/a$	$1/a$	X	
4		$1/a$	X	$2/a$

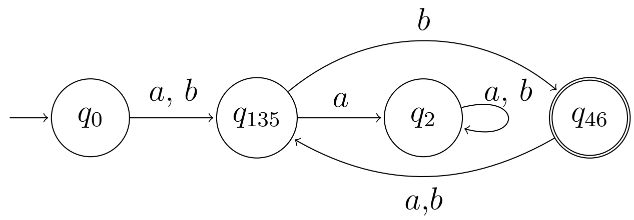
1)

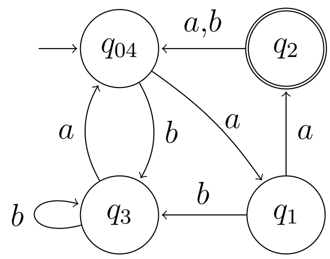
	0	1	2	3	4	5
1	$1/b$					
2	$2/a$	$1/b$				
3	$1/b$		$1/b$			
4	X	X	X	X		
5	$1/b$		$1/b$		X	
6	X	X	X	X		X

2)

	0	1	2	3
1	$1/a$			
2	X	X		
3	$2/a$	$1/a$	X	
4		$1/a$	X	$2/a$

2. Die dazugehörigen Quotientenautomaten sehen folgendermaßen aus:





Hausaufgabe 2 (4 Punkte)

Konstruieren Sie durch Lösen entsprechender Gleichungen einen regulären Ausdruck für die Sprache $L(M)$ des DFA $M = (\{q_0, q_1, q_2, q_3, q_4\}, \{a, b\}, \delta, q_0, \{q_2, q_4\})$ mit folgender Übergangstabelle:

q_i	$\delta(q_i, a)$	$\delta(q_i, b)$
q_0	q_1	q_2
q_1	q_3	q_4
q_2	q_2	q_1
q_3	q_3	q_3
q_4	q_4	q_2

Lösungsvorschlag

Die entsprechenden Gleichungen sind:

$$X_0 \equiv aX_1 \mid bX_2 \tag{1}$$
$$X_1 \equiv aX_3 \mid bX_4 \tag{2}$$
$$X_2 \equiv aX_2 \mid bX_1 \mid \boldsymbol{\epsilon} \tag{3}$$
$$X_3 \equiv aX_3 \mid bX_3 \tag{4}$$
$$X_4 \equiv aX_4 \mid bX_2 \mid \boldsymbol{\epsilon} \tag{5}$$

Gleichung (4) läßt sich leicht lösen:

$$X_3 \equiv (a \mid b)X_3 \equiv (a \mid b)^*\boldsymbol{\emptyset} \equiv \boldsymbol{\emptyset} \tag{6}$$

Damit läßt sich X_3 aus (2) eliminieren:

$$X_1 \equiv a\boldsymbol{\emptyset} \mid bX_4 \equiv bX_4 \tag{7}$$

Dann folgt:

$$X_0 \equiv a(bX_4) \mid bX_2 \equiv abX_4 \mid bX_2 \tag{8}$$
$$X_2 \equiv aX_2 \mid b(bX_4) \mid \boldsymbol{\epsilon} \equiv aX_2 \mid bbX_4 \mid \boldsymbol{\epsilon} \tag{9}$$

Mit Arden's Lemma erhält man aus (9):

$$X_2 \equiv a^*(bbX_4 \mid \boldsymbol{\epsilon}) \equiv a^*bbX_4 \mid a^* \tag{10}$$

Einsetzen in (5) und Anwenden von Arden's Lemma ergibt:

$$\begin{aligned} X_4 &\equiv aX_4 \mid b(a^*bbX_4 \mid a^*) \mid \boldsymbol{\epsilon} \\ &\equiv (a \mid ba^*bb)X_4 \mid ba^* \mid \boldsymbol{\epsilon} \\ &\equiv (a \mid ba^*bb)^*(ba^* \mid \boldsymbol{\epsilon}) \end{aligned}$$

Im Folgenden kürzen wir den regulären Ausdruck ba^*bb mit γ ab.
Durch Einsetzen in (8) und unter Verwendung von (10) erhalten wir den gesuchten regulären Ausdruck:

$$\begin{aligned} X_0 &\equiv abX_4 \mid b(a^*bbX_4 \mid a^*) \\ &\equiv (ab \mid \gamma)X_4 \mid ba^* \\ &\equiv (ab \mid \gamma)(a \mid \gamma)^*(ba^* \mid \boldsymbol{\epsilon}) \mid ba^* \end{aligned}$$

Hausaufgabe 3 (5 Punkte)

Sei $\Sigma = \{0, 1, \#\}$. Ein Wort $w \in \Sigma^*$ kann als Binärdarstellung einer natürlichen Zahl aufgefasst werden, die wir mit $\|w\|$ bezeichnen. Wir betrachten zwei verschiedene Additionssprachen.

1. Zeigen Sie: Die Sprache

$$\begin{aligned} L_1 = \{u_1v_1w_1 \cdots u_nv_nw_n \mid n \in \mathbb{N} \wedge \forall i.\ u_i, v_i, w_i \in \{0,1\} \\ \wedge \|u_1 \cdots u_n\| = \|v_1 \cdots v_n\| + \|w_1 \cdots w_n\|\} \end{aligned}$$

   ist regulär.

2. Verwenden Sie das Pumping-Lemma, um zu zeigen, dass die Sprache

$$L_2 = \{u\#v\#w \mid u, v, w \in \{0,1\}^* \wedge \|u\| = \|v\| + \|w\|\}$$

   nicht regulär ist.

Lösungsvorschlag

1. Da reguläre Sprachen abgeschlossen sind unter Spiegelungen zeigen wir stattdessen, dass die Sprache

$$\begin{aligned} L_1^R = \{w_nv_nu_n \cdots w_1v_1u_1 \mid n \in \mathbb{N} \wedge \forall i.\ u_i, v_i, w_i \in \Sigma \\ \wedge \|u_1 \cdots u_n\| = \|v_1 \cdots v_n\| + \|w_1 \cdots w_n\|\} \end{aligned}$$

   regulär ist.

   Dazu geben wir einen Automaten auf der Zustandsmenge $Q = \{a_0, a_1, b_0, b_1, b_2, c_0, c_1, c_2, c_3\}$ an, der die Eingabe in Dreier-Blöcken verarbeitet. Den Zuständen wird dabei folgende Bedeutung zugeordnet:

   - a_i: Der Automat liest als nächstes ein Zeichen von w ein, und es gibt einen Übertrag i.
   - b_i: Der Automat liest als nächstes ein Zeichen von v ein. Die Summe aus Übertrag und dem letzten Zeichen aus w beträgt i.
   - c_i: Der Automat liest als nächstes ein Zeichen von u ein. Die Summe aus Übertrag und den letzten Zeichen aus w und v beträgt i.

Startzustand und einziger Endzustand ist dann a_0. Die Übergangsrelation sieht wie folgt aus:

q_i	$\delta(q_i, 0)$	$\delta(q_i, 1)$
a_0	$\{b_0\}$	$\{b_1\}$
a_1	$\{b_1\}$	$\{b_2\}$
b_0	$\{c_0\}$	$\{c_1\}$
b_1	$\{c_1\}$	$\{c_2\}$
b_2	$\{c_2\}$	$\{c_3\}$
c_0	$\{a_0\}$	$\varnothing$
c_1	$\varnothing$	$\{a_0\}$
c_2	$\{a_1\}$	$\varnothing$
c_3	$\varnothing$	$\{a_1\}$

In den Übergängen von a nach b bzw. b nach c werden also die Ziffern von w, v mit dem Übertrag aufsummiert; in den Übergängen von c nach a wird verglichen, ob die Summe mit der aktuellen Ziffer von u übereinstimmt und ein eventueller Übertrag für die nächste Runde übernommen.

Wir bezeichnen den Automat $(Q, \Sigma, \delta, a_0, \{a_0, a_1\})$ mit A.

Formaler Korrektheitsbeweis (wurde in der Übung nicht erwartet): Betrachtet man die Zustandsübergangsrelation δ, so sieht man, dass für alle $i, j \in 0, 1$ ein Übergang von a_i zu a_j (ohne dass daziwschen ein weiteres a_k, $k \in 0, 1$ vorkommt) immer in genau drei Schritten erfolgt. Da a_0 der Startzustand ist und a_0, a_1 die einzigen Endzustände sind, akzeptiert der Automat A nur Wörter, deren Länge ein vielfaches von 3 ist.

Wir setzen jetzt: $R_0 = L(A)$ und $R_1 = L((Q, \Sigma, \delta, a_1, \{a_0, a_1\}))$. Wir zeigen jetzt per Induktion über n: $x = w_1 v_1 u_1 \cdots w_n v_n u_n \in R_0 \Leftrightarrow \|u\| + \|v\| = \|w\|$ und $x \in R_1 \Leftrightarrow 1 + \|u\| + \|v\| = \|w\|$.

2. Angenommen, L_2 sei regulär. Dann sei n eine Pumping-Lemma-Zahl, und wir wählen das Wort $z = 1^n\#0\#1^n \in L_2$. Für jede Zerlegung $z = uvw$ mit $|uv| \leq n$ und $v \neq \epsilon$ gibt es $k \geq 0$, $l > 0$ so dass $u = 1^k$, $v = 1^l$ und $w = 1^{(n-k-l)}\#0\#1^n$. Aber dann ist $uv^0w = 1^{(n-l)}\#0\#1^n$ nicht in L, denn die Summe aus $\|1^{(n-l)}\|$ und $\|0\|$ ist nicht $\|1^n\|$.

   Damit erhalten wir einen Widerspruch zur Annahme, und somit ist L_2 nicht regulär.

Hausaufgabe 4 (5 Punkte)

Für zwei Sprachen L_1 und L_2 über einem Alphabet Σ definieren wir die Quotientensprache $L_1/L_2 = \{u \mid uv \in L_1 \wedge v \in L_2\}$. Zeigen Sie: Ist L_1 regulär, so ist auch L_1/L_2 regulär.

Lösungsvorschlag

Sei $M = (Q, \Sigma, \delta, q_0, F)$ ein DFA mit $L(M) = L_1$. Wir behaupten, der Automat $M' = (Q, \Sigma, \delta, q_0, F')$ mit

$$F' = \{q \in Q \mid \exists u, v \in \Sigma^*.\ \delta(q_0, u) = q \wedge v \in L_2 \wedge uv \in L_1\}$$

erkennt L_1/L_2.

Sei $w \in L(M')$. Dann gibt es ein $q \in F'$, so dass $\hat{\delta}(q_0, w) = q$ ist. Also gibt es $u, v \in \Sigma^*$ mit $\hat{\delta}(q_0, u) = q$, $v \in L_2$ und $uv \in L_1$. Weiterhin gilt

$$\hat{\delta}(q_0, uv) = \hat{\delta}(\hat{\delta}(q_0, u), v) = \hat{\delta}(\hat{\delta}(q_0, w), v) = \hat{\delta}(q_0, wv)$$

und damit $wv \in L_1$ und damit $w \in L_1/L_2$.
Sei jetzt umgekehrt $w \in L_1/L_2$ und $\hat{\delta}(q_0, w) = q$. Dann gibt es ein $v \in L_2$, so dass $wv \in L_1$. Also gilt nach Definition von F' auch $q \in F'$ und damit $w \in L(M')$.
Damit haben wir gezeigt, dass die Sprache L_1/L_2 von einem DFA akzeptiert wird. Also ist L_1/L_2 regulär.

Quiz 1

Beantworten Sie kurz die folgenden Fragen:

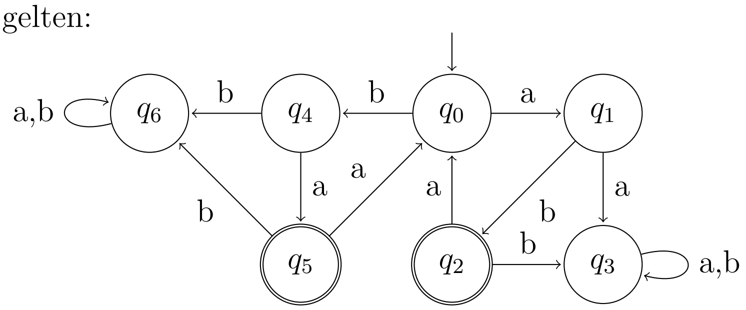
1. Gibt es endliche, nicht kontextfreie Sprachen?
2. Ist die Sprache $\{a^m b^n \mid m < n\}$ kontextfrei?
3. Welche Sprache beschreibt die Grammatik mit den Produktionen $S \to aS \mid bB$ und $B \to bBb$?

Lösungsvorschlag

1. Nein, denn jede endliche Sprache ist regulär und damit kontextfrei.
2. Ja, sie wird von der Grammatik $G = (\{S\}, \{a, b\}, P, S)$ mit $P = \{S \to aSb \mid Sb \mid b\}$ erzeugt.
3. Die leere Sprache, denn jede endliche Ableitung endet mit einem Wort, dass ein Nichtterminal enthält.

Tutoraufgabe 1

1. Sei $\Sigma = \{a, b\}$. Zeigen Sie, dass im untenstehenden DFA $q_3 \not\equiv_M q_4$ und $q_3 \equiv_M q_6$ gelten:



2. Sei $L = L(a^*b^*)$. Zeigen oder widerlegen Sie die folgenden Äquivalenzbeziehungen:

$$aa \equiv_L \epsilon\,, \qquad ab \equiv_L aa\,, \qquad aba \equiv_L abba\,, \qquad aba \equiv_L \epsilon\,.$$

3. Sei nun $\Sigma_n = \{a_1, \ldots, a_n\}$ und

$$L_n = \{w \in \Sigma_n^* \mid \forall a \in \Sigma_n.\ |w|_a = 1\}\,.$$

Zeigen Sie, dass jeder DFA, der L_n akzeptiert, mindestens $2^n + 1$ Zustände haben muss. Betrachten Sie dazu die Äquivalenzklassen von $\equiv_{L_n}$.

Lösungsvorschlag

1. $\hat{\delta}(q_3, a) = q_3 \notin F$, aber $\hat{\delta}(q_4, a) = q_5 \in F$. Damit ist die Definition von Äquivalenz nicht erfüllt.

   Alle von q_3 bzw. q_6 ausgehenden Transitionen führen wieder nach q_3 bzw. q_6. Das heißt, für alle $w \in \Sigma^*$ gilt $\hat{\delta}(q_3, w) = q_3 \notin F$ und $\hat{\delta}(q_6, w) = q_6 \notin F$ und damit ist $q_3 \equiv_M q_4$.

2. - $aa \equiv_L \epsilon$:

   $$\begin{aligned} aaw \in L &\iff aaw \in L(a^*b^*) \\ &\iff aaw = a^m b^n && \text{für gewisse } m \geq 2 \text{ und } n \geq 0 \\ &\iff \epsilon w = a^m b^n && \text{für gewisse } m \geq 0 \text{ und } n \geq 0 \\ &\iff \epsilon w \in L(a^*b^*) \iff \epsilon w \in L \end{aligned}$$

   - $ab \not\equiv_L aa$: Für $w = a$ gilt $abw = aba \notin L$ und $aaw = aaa \in L$.
   - $aba \equiv_L abba$: Für alle w gilt $abaw \notin L$ und $abbaw \notin L$.
   - $aba \not\equiv_L \epsilon$: Für $w = \epsilon$ gilt $abaw = aba \notin L$ und $\epsilon w = \epsilon \in L$.

3. Betrachte Äquivalenzklassen von $\equiv_{L_n}$. Für jede Teilmenge $A \subseteq \Sigma$ sei w_A ein beliebiges Wort, welches jedes Zeichen aus A genau einmal enthält.

   Seien nun $A, B \subseteq \Sigma$ beliebig mit $A \neq B$. Dann existiert o.B.d.A. ein $x \in \Sigma$ mit $x \in B$ und $x \notin A$ (sonst vertausche A und B). Nun gilt einerseits $w_A w_{\Sigma \setminus A} \in L_n$ und andererseits $w_B w_{\Sigma \setminus A} \notin L_n$ (denn x kommt sowohl in w_B als auch in $w_{\Sigma \setminus A}$ vor).

   Damit ist $w_A \not\equiv_{L_n} w_B$. Wir erhalten also für jede der 2^n Teilmengen von Σ eine Äquivalenzklasse in $\equiv_{L_n}$. Eine weitere Äquivalenzklasse hat den Repräsentanten aa ($a \in \Sigma$), welches zu keinem der w_A äquivalent ist.

   Da die Äquivalenzklassen die Zustände des kanonischen Minimalautomaten sind, ergibt sich daraus die untere Schranke $2^n + 1$ für die Anzahl der Zustände.

Tutoraufgabe 2

1. Sei $\Sigma = \{0, 1\}$. Geben Sie eine Grammatik an, die alle Wörter beschreibt, bei denen Nullen vor Einsen kommen, die Anzahl der Nullen geringer als die Anzahl der Einsen ist sowie die Länge jedes Wortes ungerade ist.

2. Gegeben sei die CFG $G = (\{S, T, U\}, \{a, b, c\}, P, S)$ mit den Produktionen

   $$S \to aTc \mid bTb \qquad T \to bT \mid Tc \mid aU \qquad U \to a \mid c$$

   Entscheiden Sie, welche der folgenden Wörter in der Sprache $L(G)$ liegen und geben Sie für von der Grammatik erzeugte Wörter eine Kette von Ableitungsschritten an:

   $$aacc \qquad abacbc \qquad bbaccb$$

3. Zeigen Sie formal mit Induktion: $x \to^n y \wedge y \to^m z \Longrightarrow x \to^{n+m} z$.

Lösungsvorschlag

1. $G = (\{S\}, \{0,1\}, P, S)$ mit den Produktionen $S \to 0S1 \mid S11 \mid 1$

2.
   - $S \to aTc \to aaUc \to aacc$
   - $abacbc$ ist nicht in $L(G)$
   - $S \to bTb \to bbTb \to bbTcb \to bbaUcb \to bbaccb$

3. Induktion über m: Wenn $m = 0$, dann ist $y = z$ und es ist nichts zu zeigen. Im Induktionsschritt schließen wir wie folgt:

$$\begin{aligned}
& x \to^n y \quad \wedge \quad y \to^{m+1} z \\
\Longrightarrow\ & x \to^n y \quad \wedge \quad y \to^m w \quad \wedge \quad w \to z && \text{(nach Def.)} \\
\Longrightarrow\ & x \to^{n+m} w \quad \wedge \quad w \to z && \text{(nach Induktionshypothese)} \\
\Longrightarrow\ & x \to^{n+m+1} \to z && \text{(nach Def.)}
\end{aligned}$$

Tutoraufgabe 3

Zeigen Sie: Jede Sprache, die von einer rechtslinearen CFG erzeugt wird, kann auch von einer linkslinearen CFG erzeugt werden.

Lösungsvorschlag

Statt zu einer beliebigen rechtslinearen Grammatik eine äquivalente linkslineare Grammatik zu kontruieren, was sehr aufwendig ist, nehmen wir den Umweg über reguläre Sprachen und Automaten. Wir zeigen dazu, dass (1) die Sprache einer rechtslineare Grammatik regulär ist, (2) jede reguläre Sprache durch eine rechtslineare Grammatik beschrieben werden kann, sowie (3) die geforderte Aussage mithilfe der regulären Abschlußeigenschaft von Spiegelsprachen.

1. Zu einer rechtslinearen kontextfreien Grammatik $G = (V, \Sigma, P, S)$ konstruieren wir einen NFA $N = (V, \Sigma, \delta, S, F)$, der dieselbe Sprache akzeptiert:

$$\delta(A, a) = \{B \in V \mid A \to aB \in P\}$$
$$F = \{A \in V \mid A \to \epsilon \in P\}$$

   Wir beweisen $L(G) = L(N)$ in zwei Teilschritten:

   (a) Sei $A \in V$ beliebig. Dann gilt $w \in L(A) \implies \hat{\delta}(\{A\}, w) \cap F \neq \emptyset$ aufgrund folgender Induktion über die Erzeugung von w:

   - $A \to \epsilon$: Dann muss $A \in F$ sein und somit $\hat{\delta}(A, \epsilon) = \{A\} \cap F \neq \emptyset$.
   - $A \to aB$: Sei $w = au$ mit $u \in L(B)$: Nach Induktionsvoraussetzung gibt es somit ein $C \in \hat{\delta}(B, u) \cap F$. Nun ist

$$\hat{\delta}(\{A\}, au) = \hat{\delta}(\bigcup_{A' \in \{A\}} \delta(A', a), u) = \hat{\delta}(\delta(A, a), u)\ .$$

   Da $B \in \delta(A, a)$, ist auch $C \in \hat{\delta}(\{A\}, au)$.

(b) Für die Gegenrichtung müssen wir für jeden möglichen Ablauf des NFA eine Ableitung konstruieren. Wir zeigen also $\hat{\delta}(\{A\}, w) \cap F \neq \emptyset \implies A \to_G^* w$ mit Induktion über w:

- $w = \epsilon$: Wenn $\hat{\delta}(\{A\}, \epsilon) \cap F \neq \emptyset$, dann ist $A \in F$ und damit $A \to \epsilon \in P$. Damit gilt $A \to_G^* \epsilon$.
- $w = au$: Wenn $\hat{\delta}(\{A\}, au) \cap F \neq \emptyset$, dann gibt es einen Zustand $B \in \delta(A, a)$ so dass $\hat{\delta}(\{B\}, u) \cap F \neq \emptyset$. Mit Induktionshypothese erhalten wir $B \to^* u$ und daraus die Ableitung $A \to aB \to^* au = w$.

2. Sei $M = (Q, \Sigma, \delta, q_0, F)$ ein DFA. Dann definieren wir die rechtslineare Grammatik $G = (Q \cup \Sigma', \Sigma, P, q_0)$ mit

$$P = \{q \to aq' \mid \delta(q, a) = q'\} \cup \{q \to aa', a' \to \epsilon \mid \delta(q, a) \in F\}$$

Offensichtlich ist G rechtslinear. Der Beweis von $L(G) = L(M)$ ist einfach ...

3. Sei G eine rechtslineare Grammatik. Dann ist $L(G)$ nach (1) regulär. Da nach Lemma 2.22 die gespiegelte Sprache $L(G)^R$ ebenfalls regulär ist, gibt es einen DFA M, der $L(G)^R$ akzeptiert. Nach (2) gibt es eine rechtslineare Grammatik G', die $L(M) = L(G)^R$ erzeugt. Die Spiegelung der Produktionen der Grammatik G' liefert dann offensichtlich eine linkslineare Grammatik G'', die $L(M)^R = L(G)$ erzeugt (induktiver Beweis über die Ableitung von Wörtern in $L(G'')$ bzw. von gespiegelten Wörtern w^R in $L(G')$).