Fachbereich Mathematik
Prof. Dr. Ulrich Kohlenbach
Dr. Benno van den Berg

TECHNISCHE UNIVERSITÄT DARMSTADT

Februar 2009

Extra Aufgaben Formale Grundlagen der Informatik I
WS 2008/2009

(E8.1) [Induktion]

Für jedes Wort $w \in \Sigma^*$, definieren wir die Umkehrung w^{-1} formell durch:

$$\begin{array}{rcl} \varepsilon^{-1} & = & \varepsilon \\ (w \cdot a)^{-1} & = & a \cdot w^{-1}, \text{ wobei } w \in \Sigma^*, a \in \Sigma. \end{array}$$

Beweisen Sie, mittels struktureller Induktion, dass

$$(v \cdot w)^{-1} = w^{-1} \cdot v^{-1}$$

für beliebige Wörter $v, w \in \Sigma^*$.

Musterlösung.

Wir beweisen die Aussage

$$(v \cdot w)^{-1} = w^{-1} \cdot v^{-1}$$

mittels struktureller Induktion über w (und lassen $v \in \Sigma^*$ fest).

Induktionsanfang: $w = \varepsilon$.

$$\begin{array}{rcll} (v \cdot \varepsilon)^{-1} & = & v^{-1} & \text{Definition von } \cdot \\ & = & \varepsilon \cdot v^{-1} & \text{Definition von } \cdot \\ & = & \varepsilon^{-1} \cdot v^{-1} & \text{Definition von } (-)^{-1} \end{array}$$

Induktionsschritt: $w = z \cdot a$, wobei $a \in \Sigma$ und die Aussage schon bewiesen worden ist für $z \in \Sigma^*$ (Induktionshypothese).

$$\begin{array}{rcll} (v \cdot (z \cdot a))^{-1} & = & ((v \cdot z) \cdot a)^{-1} & \text{Associativität von } \cdot \\ & = & a \cdot (v \cdot z)^{-1} & \text{Definition von } (-)^{-1} \\ & = & a \cdot (z^{-1} \cdot v^{-1}) & \text{Induktionshypothese} \\ & = & (a \cdot z^{-1}) \cdot v^{-1} & \text{Associativität von } \cdot \\ & = & (z \cdot a)^{-1} \cdot v^{-1} & \text{Definition von } (-)^{-1} \end{array}$$

(E8.2) [Reguläre Sprachen]

Unter welchen der folgenden Operationen ist die Menge der regulären Sprachen abgeschlossen?

(i) In jedem Wort werden alle Buchstaben a durch b ersetzt und alle b durch a.

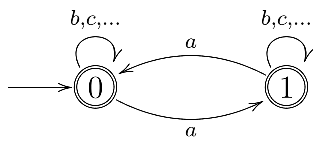
(ii) Jedes zweite Vorkommen des Buchstaben a wird durch das Wort aba ersetzt.

(iii) Die Buchstaben in jedem Wort dürfen beliebig umsortiert werden, d.h. ist etwa das Wort $aaba$ in der Sprache, so fügen wir auch die Wörter $aaab$, $abaa$ und $baaa$ hinzu.

Musterlösung.

(i) Die Menge der regulären Sprachen ist under dieser Operation abgeschlossen. Hat man einen regulären Ausdruck für eine Sprache, so kann man daraus einen Ausdruck für die neue Sprache konstruieren, indem man jedes a durch b und jedes b durch a ersetzt.

(ii) Die Menge der regulären Sprachen ist unter dieser Operation abgeschlossen. Aus einem DFA für die alte Sprache können wir wie folgt einen Automaten für die neue Sprache konstruieren. Zuerst bilden wir das Produkt mit dem Automaten



um die Anzahl der a zu zählen. Im resultierenden Automaten können wir jede a-Transition

$$(q, 1) \longrightarrow (p, 0),$$

welche bei einem Zustand mit zweiter Komponente 1 beginnt, ersetzen durch eine Folge van Transitionen

$$(q, 1) \xrightarrow{a} \bullet \xrightarrow{b} \bullet \xrightarrow{a} (p, 0).$$

(Hierzu müssen wir für jede solche Transition zwei neue Zwischenzustände einführen.)

(iii) Die Menge der regulären Sprache ist unter dieser Operation nicht abgeschlossen. Wenn wir diese Operation auf die sprache $L_0 = L((ab)^*)$ anwenden, erhalten wir die Sprache $L_1 = \{w \in \{a, b\}^* \,:\, |w|_a = |w|_b\}$. Nach dem Schnitt mit $L(a^*b^*)$ erhalten wir

$$L_2 = L_1 \cap L(a^*b^*) = \{a^n b^n \,:\, n \in \mathbb{N}\}.$$

Wäre die Menge der regulären Sprachen unter der Operation abgeschlossen, dann wären auch L_1 und L_2 regulär. Aber wir wissen, dass L_2 nicht regulär ist.

(E8.3) [Automaten]

Sei $\Sigma = \{a, b, c\}$ und

$$L = L((c + (ab + b)^*ba^*)^*).$$

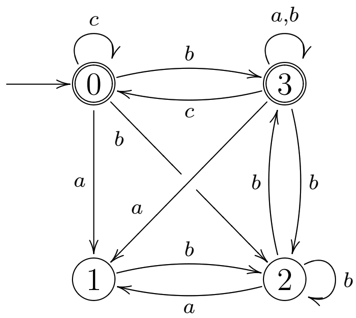
(i) Geben Sie einen NFA für L an.

(ii) Geben Sie einen DFA für L an.

(iii) Konstruieren Sie den minimalen DFA für L.

Musterlösung.

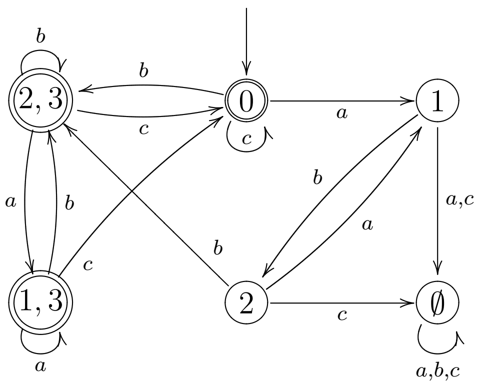
(i) Ein NFA, der die Sprache L erkennt, ist:



(ii)

δ	a	b	c
$\{0\}$	$\{1\}$	$\{2,3\}$	$\{0\}$
$\{1\}$	$\emptyset$	$\{2\}$	$\emptyset$
$\{2,3\}$	$\{1,3\}$	$\{2,3\}$	$\{0\}$
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$\{2\}$	$\{1\}$	$\{2,3\}$	$\emptyset$
$\{1,3\}$	$\{1,3\}$	$\{2,3\}$	$\{0\}$

Akzeptierende Zustände sind $\{0\}, \{2,3\}, \{1,3\}$.



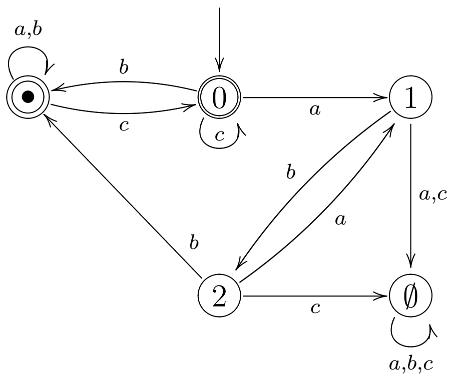
(iii) Wir bestimmen die Relationen $\not\sim_i$.

$\not\sim_0$	$\{0\}$	$\{1\}$	$\{2,3\}$	$\emptyset$	$\{2\}$	$\{1,3\}$
$\{0\}$		$\times$		$\times$	$\times$	
$\{1\}$	$\times$		$\times$			$\times$
$\{2,3\}$		$\times$		$\times$	$\times$	
$\emptyset$	$\times$		$\times$			$\times$
$\{2\}$	$\times$		$\times$			$\times$
$\{1,3\}$		$\times$		$\times$	$\times$	

$\not\sim_1$	$\{0\}$	$\{1\}$	$\{2,3\}$	$\emptyset$	$\{2\}$	$\{1,3\}$
$\{0\}$		$\times$	$\times$	$\times$	$\times$	$\times$
$\{1\}$	$\times$		$\times$		$\times$	$\times$
$\{2,3\}$	$\times$	$\times$		$\times$	$\times$	
$\emptyset$	$\times$		$\times$		$\times$	$\times$
$\{2\}$	$\times$	$\times$	$\times$	$\times$		$\times$
$\{1,3\}$	$\times$	$\times$		$\times$	$\times$	

$\not\sim_2$	$\{0\}$	$\{1\}$	$\{2,3\}$	$\emptyset$	$\{2\}$	$\{1,3\}$
$\{0\}$		$\times$	$\times$	$\times$	$\times$	$\times$
$\{1\}$	$\times$		$\times$	$\times$	$\times$	$\times$
$\{2,3\}$	$\times$	$\times$		$\times$	$\times$	
$\emptyset$	$\times$	$\times$	$\times$		$\times$	$\times$
$\{2\}$	$\times$	$\times$	$\times$	$\times$		$\times$
$\{1,3\}$	$\times$	$\times$		$\times$	$\times$	

Da $\sim_2 = \sim_3$ können wir die Zustände $\{2,3\}$ und $\{1,3\}$ identifizieren. Ein minimaler DFA für L ist deshalb:



(E8.4) [Kellerautomaten]

Konstruieren Sie einen PDA für die Sprache L der Palindrome über dem Alphabet $\Sigma = \{a, b\}$.

Musterlösung.

Ein PDA für L ist $\mathcal{P} = (\Sigma, Q, q_0, \Delta, A, \Gamma, \#)$ mit $Q = \{q_0, q_1\}$, $\Gamma = \{\#, A, B\}$, $A = \{q_1\}$ und Transitionen

$$
\begin{array}{llllll}
\{ & (q_0, & \#, & \varepsilon, & \varepsilon, & q_1) \\
 & (q_0, & \#, & a, & \varepsilon, & q_1) \\
 & (q_0, & \#, & b, & \varepsilon, & q_1) \\
 & (q_0, & \#, & a, & A, & q_0) \\
 & (q_0, & \#, & b, & B, & q_0) \\
 & (q_0, & A, & a, & AA, & q_0) \\
 & (q_0, & B, & a, & AB, & q_0) \\
 & (q_0, & A, & b, & BA, & q_0) \\
 & (q_0, & B, & b, & BB, & q_0) \\
 & (q_0, & A, & a, & A, & q_1) \\
 & (q_0, & B, & a, & B, & q_1) \\
 & (q_0, & A, & b, & A, & q_1) \\
 & (q_0, & B, & b, & B, & q_1) \\
 & (q_0, & A, & A, & \varepsilon, & q_1) \\
 & (q_0, & B, & b, & \varepsilon, & q_1) \\
 & (q_1, & A, & a, & \varepsilon, & q_1) \\
 & (q_1, & B, & b, & \varepsilon, & q_1) \ \}.
\end{array}
$$

(E8.5) [Kontextfreie Sprachen]

Welche von den folgenden Sprachen sind kontextfrei?

(i) $L_1 = \{a^n b^m \in \{a, b\}^* \, : \, n \geqslant m\}$

(ii) $L_2 = \{a^{n!} \in \{a\}^* \, : \, n \geqslant 0\}$

(iii) $L_3 = \{a^p \in \{a\}^* \, : \, p \text{ prim}\}$

Begründen Sie Ihre Antwort!

Musterlösung.

Die Sprache L_1 ist kontextfrei, die Sprachen L_2 und L_3 sind das nicht.

(i) Um zu zeigen, dass L_1 kontextfrei ist, können wir entweder eine kontextfreie Grammatik konstruieren, oder einen Kellerautomat angeben, der L_1 erkennt.

Eine kontextfreie Grammatik wäre:

$$X_0 \quad \rightarrow \quad aX_0 \,|\, aX_0b \,|\, \varepsilon$$

Sei $\mathcal{P} = (\Sigma, Q, q_a, \Delta, A, \Gamma, \#)$ der Kellerautomat mit Eingabealphabet $\Sigma = \{a, b\}$, Zustandsmenge $Q = \{q_a, q_b\}$, $A = \{q_a, q_b\}$ als Menge der akzeptierenden Zustände, Kelleralphabet $\Gamma = \{\#, |\}$ und Übergangsrelation Δ gegeben durch

$$\begin{array}{ll} \{ & (q_a, \#, \varepsilon, \varepsilon, q_a) \\ & (q_a, \#, a, |, q_a) \\ & (q_a, |, a, ||, q_a) \\ & (q_a, |, \varepsilon, \varepsilon, q_a) \\ & (q_a, |, b, \varepsilon, q_b) \\ & (q_b, |, b, \varepsilon, q_b) \quad \}. \end{array}$$

Dann erkennt $\mathcal{P}$ die Sprache L_1.

(ii) Um zu beweisen, dass L_2 und L_3 nicht kontextfrei sind, zeigen wir, dass in beiden Fällen das Pumping Lemma verletzt ist (ganz analog zu Aufgabe (E5.3)). Das heißt:

> Für jedes $i \in \mathbb{N}$ gibt es ein $x \in L$ mit $|x| \geqslant i$, so dass für alle Zeichenreihen y, u, v, w, z, mit $x = y \cdot u \cdot v \cdot w \cdot z$, $|uw| > 0$ und $|uvw| \leqslant i$, es ein $j \in \mathbb{N}$ gibt mit
> $$y \cdot u^j \cdot v \cdot w^j \cdot z \notin L.$$

Wir zeigen erst, dass L_2 nicht kontextfrei ist. Sei $i \in \mathbb{N}$ beliebig und betrachte

$$x = a^{(i+1)!}.$$

Offensichtlich ist $x \in L_2$ und $|x| \geqslant i$. Seien also y, u, v, w, z, mit $x = y \cdot u \cdot v \cdot w \cdot z$, $|uw| > 0$ und $|uvw| \leqslant i$. Für $i = 0$ haben wir dann einen Widerspruch ($0 < |uw| \leqslant |uvw| \leqslant 0$) und für $i = 1$ können wir $j = 2$ wählen (für $i = 1$ und $j = 2$ ist $x = aa$, $uw = a$ und $y \cdot u^j \cdot v \cdot w^j \cdot z = a^3 \notin L_2$). Wir nehmen deshalb an, dass $i > 1$ und verwenden, dass dann $(i+1)! - i > i!$. Jetzt wählen wir $j = 0$ und behaupten

$$x' = y \cdot u^j \cdot v \cdot w^j \cdot z = yvz \notin L_2.$$

Da $i \geqslant |uw| > 0$, kann die Länge von x' unmöglich die Form $n!$ haben:

$$(i+1)! = |x| > |x| - |uw| = |x'| \geqslant (i+1)! - i > i!.$$

(iii) Wir zeigen, dass L_3 nicht kontextfrei ist. Sei $i \in \mathbb{N}$ beliebig und betrachte

$$x = a^l,$$

wobei $l > i + 1$ eine Primzahl ist. Offensichtlich ist $x \in L_3$ und $|x| \geqslant i$. Seien also y, u, v, w, z, mit $x = y \cdot u \cdot v \cdot w \cdot z$, $|uw| > 0$ und $|uvw| \leqslant i$. Aus $|uvw| \leqslant i$ folgt, dass $|yz| > 1$ und $|yvz| > 1$. Wählen wir $j = |yvz|$ und $x' = y \cdot u^j \cdot v \cdot w^j \cdot z$. Wir bestimmen die Länge von x':

$$|x'| = |y \cdot u^j \cdot v \cdot w^j \cdot z| = |yvz| + j|uw| = j(|uw| + 1).$$

Weil $j > 1$ und $|uw| + 1 > 1$, ist $|x'|$ nicht prim und $x' \notin L_3$.

(E8.6) [Chomsky Hierarchie]

Geben Sie an, ob die folgenden Aussagen richtig oder falsch sind und begründen Sie Ihre Antwort.

(a) Jede kontextfreie Sprache hat ein aufzählbares Komplement.

(b) Sind L_1 und L_2 kontextfreie Sprachen, dann ist auch $L_1 \backslash L_2$ kontextfrei.

(c) Jede Sprache mit endlichem Komplement ist regulär.

(d) Ist L_1 regulär und L_2 beliebig, dann ist

$$L = \{x \in \Sigma^* \, : \text{ es existiert ein } y \in L_2, \text{ so dass } xy \in L_1\}$$

regulär.

Musterlösung.

(a) Richtig: jede kontextfreie Sprache ist entscheidbar, entscheidbare Sprachen sind unter Komplement abgeschlossen und entscheidbare Sprachen sind aufzählbar.

(b) Falsch: nehmen wir $L_1 = \Sigma^*$, dann würde (b) bedeuten, dass die kontextfreien Sprachen unter Komplement abgeschlossen sind: das sind sie aber nicht.

(c) Richtig: endliche Sprachen sind regulär und reguläre Sprachen sind unter Komplement abgeschlossen.

(d) Richtig: sei $\mathcal{A}$ ein DFA, der L_1 erkennt und sei A die Menge der akzeptierenden Zustände in $\mathcal{A}$. Durch Änderung der Menge der akzeptierenden Zustände können wir aus $\mathcal{A}$ einen Automaten $\mathcal{B}$ bilden, der L erkennt: ein Zustand p ist akzeptierend in $\mathcal{B}$, wenn ein Element $w \in L_2$ existiert mit $\hat{\delta}(p, w) \in A$.