Theorie der Informatik

G. Röger | Universität Basel
Frühjahrssemester 2020 | Fachbereich Informatik

Übungsblatt 6 — Lösungen

Aufgabe 6.1 (Reguläre Ausdrücke und Pumping-Lemma für reguläre Sprachen; 3 Punkte)
Sind die folgenden Sprachen über $\Sigma = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$ regulär? Falls ja, beweisen Sie es, indem Sie einen regulären Ausdruck angeben, der die Sprache beschreibt. Falls nein, beweisen Sie es mit dem Pumping-Lemma.

(a) $L_1 = \{\mathtt{a}^n\mathtt{b}^m\mathtt{c}^{n+m} \mid n, m \in \mathbb{N}_0\}$

Lösung:

Angenommen L_1 ist regulär. Dann sei p eine Pumpingzahl von L_1. Das Wort $w = \mathtt{a}^p\mathtt{b}\mathtt{c}^{p+1}$ ist in L_1 und erfüllt $|w| \geq p$. Vom Pumping-Lemma wissen wir, dass es Wörter x, y und z gibt mit $w = xyz$, $|xy| \leq p$, $|y| \geq 1$ und für alle $i \geq 0 : xy^iz \in L_1$.

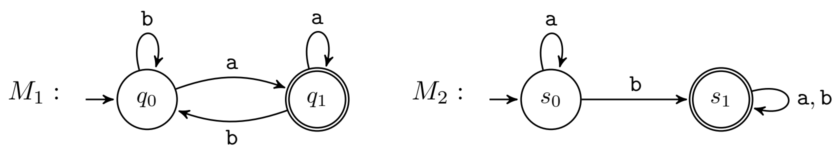
Aus $|xy| \leq p$ können wir schliessen, dass xy nur aus $\mathtt{a}$s besteht. Wenn wir w kleiner pumpen, d.h. wir wählen $i = 0$, dann erhalten wir das Wort $w_0 := xy^0z = \mathtt{a}^{p-|y|}\mathtt{b}\mathtt{c}^{p+1}$. Wegen $|y| \geq 1$ wissen wir, dass $p - |y| + 1 < p + 1$ und sehen, dass $w_0 \notin L_1$ (da die Anzahl von $\mathtt{a}$s addiert mit der Anzahl von $\mathtt{b}$s nicht die Anzahl von $\mathtt{c}$s ergibt und die Eigenschaften der Sprache daher nicht erfüllt sind). Dies ist ein Widerspruch zum Pumping-Lemma und deshalb kann L_1 nicht regulär sein.

(b) $L_2 = \{\mathtt{a}^2\mathtt{b}^n\mathtt{a}^2\mathtt{c}^m \mid n, m \in \mathbb{N}_0\}$

Lösung:

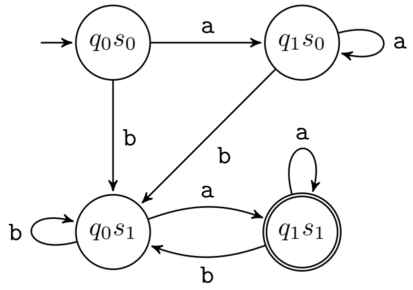
L_2 ist regulär, da sie von dem regulären Ausdruck $\mathtt{aab}^*\mathtt{aac}^*$ beschrieben wird.

Aufgabe 6.2 (Kreuzproduktautomat; 2 Punkte)
Gegeben sind die folgenden beiden DFAs M_1 und M_2.



Geben Sie den Kreuzproduktautomaten an, der $\mathcal{L}(M_1) \cap \mathcal{L}(M_2)$ akzeptiert.

Lösung:
$\mathcal{L}(M_1) = \{w \mid w \text{ endet mit } \mathtt{a}\}$, $\mathcal{L}(M_2) = \{w \mid w \text{ enthält mindestens ein } \mathtt{b}\}$
$\mathcal{L}(M_1) \cap \mathcal{L}(M_2) = \{w \mid w \text{ endet mit } \mathtt{a} \text{ und enthält mindestens ein } \mathtt{b}\}$
Der Kreuzproduktautomat sieht folgendermassen aus:



Aufgabe 6.3 (Chomsky-Normalform, 3 Punkte)

Geben Sie eine Grammatik G' in Chomsky-Normalform an, die dieselbe Sprache erzeugt wie die kontextfreie Grammatik $G = \langle \Sigma, V, P, \mathrm{S} \rangle$ mit $\Sigma = \{\mathtt{a}, \mathtt{b}\}$, $V = \{\mathrm{S}, \mathrm{W}, \mathrm{X}, \mathrm{Y}, \mathrm{Z}\}$ und den folgenden Regeln in P:

$\mathrm{S} \to \varepsilon$	$\mathrm{S} \to \mathrm{XW}$	$\mathrm{S} \to \mathrm{Z}$	$\mathrm{W} \to \mathrm{X}$	$\mathrm{X} \to \mathtt{a}\mathrm{Z}\mathtt{b}$
$\mathrm{Y} \to \mathrm{W}$	$\mathrm{Y} \to \mathtt{b}\mathrm{Y}$	$\mathrm{Z} \to \mathtt{bb}$	$\mathrm{Z} \to \mathrm{Z}\mathtt{a}$	$\mathrm{X} \to \mathrm{Y}$

Geben Sie genügend Zwischenschritte an, damit Ihre Konstruktion nachvollziehbar ist.

Lösung:

Schritt 1: Eliminiere Regeln der Form $A \to B$ (wobei $A, B \in V$).

Zuerst eliminieren wir den Zyklus aus W, X und Y.

$\mathrm{S} \to \varepsilon$	$\mathrm{S} \to \mathbf{RR}$	$\mathrm{S} \to \mathrm{Z}$		$\mathbf{R} \to \mathtt{a}\mathrm{Z}\mathtt{b}$
	$\mathbf{R} \to \mathtt{b}\mathbf{R}$	$\mathrm{Z} \to \mathtt{bb}$	$\mathrm{Z} \to \mathrm{Z}\mathtt{a}$	

Dann benennen wir Variablen in Regeln $A \to B$ so um, dass für alle Regeln $X_i \to X_j$ folgt, dass $i < j$. Hier gibt es nur eine Regel dieser Form: $\mathrm{S} \to \mathrm{Z}$.

$\mathbf{X}_1 \to \varepsilon$	$\mathbf{X}_1 \to \mathrm{RR}$	$\mathbf{X}_1 \to \mathbf{X}_2$		$\mathrm{R} \to \mathtt{a}\mathbf{X}_2\mathtt{b}$
	$\mathrm{R} \to \mathtt{b}\mathrm{R}$	$\mathbf{X}_2 \to \mathtt{bb}$	$\mathbf{X}_2 \to \mathbf{X}_2\mathtt{a}$	

Eliminiere alle Regeln der Form $X_k \to X_{k'}$. In diesem Fall ist das nur $\mathrm{X}_1 \to \mathrm{X}_2$.

$\mathrm{X}_1 \to \varepsilon$	$\mathrm{X}_1 \to \mathrm{RR}$	$\mathbf{X}_1 \to \mathtt{bb}$	$\mathbf{X}_1 \to \mathbf{X}_2\mathtt{a}$	$\mathrm{R} \to \mathtt{a}\mathrm{X}_2\mathtt{b}$
	$\mathrm{R} \to \mathtt{b}\mathrm{R}$	$\mathrm{X}_2 \to \mathtt{bb}$	$\mathrm{X}_2 \to \mathrm{X}_2\mathtt{a}$	

Schritt 2: Eliminiere Regeln mit Terminalzeichen, die nicht die Form $A \to a$ haben.

$\mathrm{X}_1 \to \varepsilon$	$\mathrm{X}_1 \to \mathrm{RR}$	$\mathrm{X}_1 \to \mathbf{BB}$	$\mathrm{X}_1 \to \mathrm{X}_2\mathbf{A}$	$\mathrm{R} \to \mathbf{A}\mathrm{X}_2\mathbf{B}$
	$\mathrm{R} \to \mathbf{B}\mathrm{R}$	$\mathrm{X}_2 \to \mathbf{BB}$	$\mathrm{X}_2 \to \mathrm{X}_2\mathbf{A}$	
$\mathbf{A} \to \mathtt{a}$	$\mathbf{B} \to \mathtt{b}$			

Schritt 3: Eliminiere Regeln der Form $A \to B_1B_2 \dots B_k$ mit $k > 2$. Wir haben eine solche Regel: $\mathrm{R} \to \mathrm{AX_2B}$.

Dies resultiert in der Grammatik $G' = \langle \Sigma, V, P, \mathrm{X}_1 \rangle$ mit dem Terminalalphabet $\Sigma = \{\mathtt{a}, \mathtt{b}\}$, den Variablen $V = \{\mathrm{R}, \mathrm{X}_1, \mathrm{X}_2, \mathrm{A}, \mathrm{B}, \mathrm{K}_1\}$ und den folgenden Regeln in P:

$\mathrm{X}_1 \to \varepsilon$	$\mathrm{X}_1 \to \mathrm{RR}$	$\mathrm{X}_1 \to \mathrm{BB}$	$\mathrm{X}_1 \to \mathrm{X}_2\mathrm{A}$	$\mathbf{R} \to \mathbf{AK}_1$
	$\mathrm{R} \to \mathrm{BR}$	$\mathrm{X}_2 \to \mathrm{BB}$	$\mathrm{X}_2 \to \mathrm{X}_2\mathrm{A}$	$\mathbf{K}_1 \to \mathbf{X}_2\mathbf{B}$
$\mathrm{A} \to \mathtt{a}$	$\mathrm{B} \to \mathtt{b}$			

Aufgabe 6.4 (Länge von Ableitungen bei Chomsky-Normalform; 2 Punkte)

Sei G eine Grammatik in Chomsky-Normalform und $w \in \mathcal{L}(G)$ ein nicht-leeres Wort ($w \neq \varepsilon$), das von G erzeugt wird. Zeigen Sie, dass jede Ableitung von w aus der Startvariable von G genau $2|w| - 1$ Ableitungsschritte hat.

Lösung:

Eine Grammatik ist in Chomsky-Normalform, wenn jede Regel die Form (a) $A \to BC$ oder (b) $A \to a$ oder (c) $S \to \varepsilon$ hat, wobei A, B, C Nichtterminale (Variablen) sind, S die Startvariable ist, B und C nicht die Startvariable sind und a ein Terminalzeichen ist.

Sei w ein Wort der Länge $n > 0$ aus der Sprache, die von der Grammatik generiert wird. Da Regeln der Sorte (c) nur verwendet werden können, um das leere Wort abzuleiten, sind sie hier nicht relevant. Um die n Terminale von w einzuführen, muss jede Ableitung genau n mal Regeln der

Sorte (b) anwenden. Jede dieser Regelanwendungen entfernt ein Nichtterminal, was bedeutet, dass für jede dieser Regelanwendungen vorher ein Nichtterminal existieren muss. Diese Nichtterminale müssen durch die Anwendung von Regeln vom Typ (a) eingeführt worden sein, wobei jede Anwendung die Zahl der Nichtterminale um 1 erhöht. Da die Startvariable bereits ein Nichtterminal ist, muss also *mindestens* $n-1$ mal eine solche Regel angewendet werden. Da Nichtterminale nur von Regeln vom Typ (b) eliminiert werden können, können auch *höchstens* $n-1$ Regelanwendungen vom Typ (a) in der Ableitung vorkommen. Daher besteht jede Ableitung von w aus exakt $n+n-1=2n-1$ Ableitungsschritten.