Universität Siegen
Lehrstuhl Theoretische Informatik
Markus Lohrey

Musterlösung zu Übungsblatt 5

Aufgabe 1. Seien $A, B, C \subseteq \Sigma^*$. Sind die folgenden Aussagen wahr oder falsch?

(a) Sei $A \cap B = C$. Wenn A und C regulär sind, dann ist B regulär.

(b) Sei $A \cup B = C$. Wenn A und C regulär sind, dann ist B regulär.

(c) Sei $A \cdot B = C$. Wenn A und C regulär sind, dann ist B regulär.

(d) Es gibt eine nicht-reguläre Sprache L so, dass $L \cdot L$ regulär ist.

Lösung zu Aufgabe 1.

(a) falsch
Sei z.B. $\Sigma = \{a, b\}$ mit $A = \{ab\}, B = \{a^n b^n \mid n \in \mathbb{N}\}$.

Dann gilt $C = A \cap B = \{ab\}$.

A und C sind regulär, B aber nicht (siehe Vorlesung).

(b) falsch
Sei z.B. $\Sigma = \{a, b\}$ mit $A = \Sigma^*, B = \{a^n b^n \mid n \in \mathbb{N}\}$.

Dann gilt $C = A \cup B = \Sigma^*$.

A und C sind regular, B aber nicht.

(c) falsch
Sei z.B. $\Sigma = \{a, b\}$ mit $A = \Sigma^*, B = \{a^n b^n \mid n \in \mathbb{N}\}$.

Da das leere Wort ε in B enthalten ist, gilt $C = A \cdot B = \Sigma^*$.

Somit sind A und C regulär, B aber nicht.

(d) wahr
Sei z.B. $L = \{a^{2n} \mid n \in \mathbb{N}\} \cup \{a^k \mid k \text{ ist eine Quadratzahl}\}$.

L ist die Sprache aller Wörter, die nur aus a's bestehen und deren Länge gerade $(0, 2, 4, \ldots)$ oder eine Quadratzahl $(1, 4, 9, 16, \ldots)$ ist.

L ist nicht regulär, da wird die Eigenschaft "k ist eine Quadratzahl" mit einer regulären Sprache nicht ausdrücken können.

Zusätzliche Übung: Erweitern Sie den Beweis aus Aufgabe 3(c) um auch formal zu zeigen, dass L nicht regulär ist.

Betrachte nun $L \cdot L$.

Da $\varepsilon \in L$ und L alle Wörter mit gerader Länge enthält, enthält auch $L \cdot L$ alle Wörter mit gerader Länge.

Zudem gilt $a \in L$, da 1 eine Quadratzahl ist.

Durch Konkatenation der Wörter mit gerader Länge und a erhalten wir alle Wörter mit ungerader Länge.

Somit enthält $L \cdot L$ alle Wörter gerader Länge ($\{a^{2n} \mid n \in \mathbb{N}\}$) und alle Wörter ungerader Länge ($\{a^{2n+1} \mid n \in \mathbb{N}\}$), also $L \cdot L = L(a^*)$.

Somit ist $L \cdot L$ regulär, obwohl L nicht regulär ist.

Aufgabe 2. Betrachten Sie die Sprache $L = \{a^n b^n \mid n \geq 0\}$. Wie sehen die Myhill-Nerode-Äquivalenzklassen bezüglich L aus?

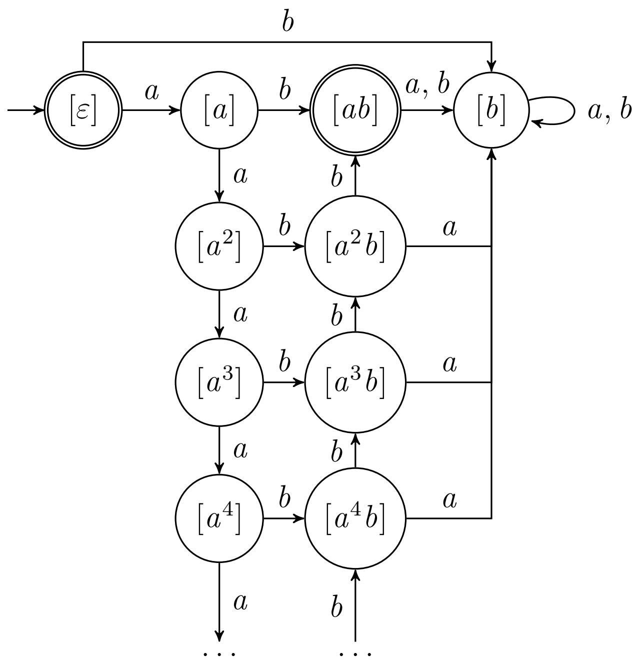
Lösung zu Aufgabe 2. Myhill-Nerode Äquivalenzklassen

- $[b] = \Sigma^* \setminus \{a^n b^m \mid n \geq m \geq 0\}$
  Diese Äquivalenzklasse enthält alle Wörter, die sich nicht zu einem Wort in der Sprache L verlängern lassen. D.h. für jedes Wort $x \in [b]$ gilt, dass für alle $w \in \Sigma^*$ auch $xw \notin L$.

- Für jedes $n \geq 0$ gibt es eine Klasse $[a^n] = \{a^n\}$.
  Diese unendlich vielen Äquivalenzklassen (eine für jedes $n \geq 0$) enthalten jeweils nur ein Wort (a^n). Für $x = a^n$ gilt, dass $xw \in L$ genau dann wenn $w \in \{a^m b^{m+n} \mid m \geq 0\}$.

- Für jedes $n \geq 1$ gibt es eine Klasse $[a^n b] = \{a^{n+m} b^{m+1} \mid m \geq 0\}$.
  Diese unendlich vielen Äquivalenzklassen (eine für jedes $n \geq 1$) enthalten jeweils unendlich viele Wörter. Für jedes Wort $x \in [a^n b]$ gilt, dass $xw \in L$ genau dann wenn $w = b^{n-1}$.

Eine etwas intuitivere Vorstellung über die Äquivalenzklassen erhält man möglicherweise durch den Versuch einen minimalen (unendlichen) deterministischen Automaten für L zu kreieren:



Mit Hilfe der oben aufgelisteten Beschreibungen lässt sich leicht zeigen, dass alle aufgezählten Klassen wirklich von der Myhill-Nerode-Äquivalenz getrennt werden. Zum Beispiel nehmen wir eine Klasse $[a^n]$ und eine Klasse $[a^m]$ mit $n \neq m$. Dann ist $a^n b^n \in L$ während $a^m b^n \notin L$, woraus direkt folgt dass a^n und a^m nicht in der gleichen Äquivalenzklasse liegen. Als weiteres Beispiel betrachten wir eine Klasse $[a^n]$ und eine Klasse $[a^m b]$ für beliebige $n \geq 0$ und $m \geq 1$. Wir haben $a^n ab^{n+1} = a^{n+1}b^{n+1} \in L$ während $a^m bab^{n+1} \notin L$ und somit folgt auch hier dass beide Wörter nicht äquivalent sind. Sie können als selbstständige Übung noch zeigen, dass auch die fehlenden Paare von Klassen jeweils getrennt werden können.

Aufgabe 3. Beweisen oder widerlegen Sie, welche der folgenden Sprachen regulär sind. Wenn ja, geben Sie die Myhill-Nerode-Äquivalenzklassen an.

(a) $\{w \in \Sigma^* \mid w = \text{rev}(w)\}$

(b) $\{ww \mid w \in \Sigma^*\}$

(c) $\{a^{n^2} \mid n \geq 0\}$

(d) $\{a^n b^m \mid n \neq m\}$

(e) $\{w \in \{a, b\}^* \mid \#_a(w) \bmod 4 \in \{1, 3\}\}$

(f) $\{a^p \mid p \text{ ist Primzahl}\}$

(g) $\{(ab)^n \mid n \geq 2\}$

(h) $\{a^n b a^m \mid (n + m) \text{ ist gerade}\}$

Lösung zu Aufgabe 3.

(a) $L = \{w \in \Sigma^* \mid w = \text{rev}(w)\}$ ist **nicht regulär**

Pumping Lemma

Wir folgen dem „Kochrezept“ für das Pumping-Lemma:

Wähle $x = a^n b a^n \in L$. Es gilt $|x| = 2n + 1 \geq n$.

Betrachte alle Zerlegungen $x = uvw$ mit $|v| \geq 1$ und $|uv| \leq n$:

Wir haben $u = a^l, v = a^m, w = a^s b a^n$ $(l + m + s = n)$, weil auf Grund von $|uv| \leq n$ sowohl u als auch v vollständig im ersten a^n liegen müssen.

Wir wählen den Pumpfaktor $i = 2$ und betrachten $uv^i w$:

$uv^2 w = a^l a^{2m} a^s b a^n = a^{l+2m+s} b a^n = a^{(l+m+s)+m} b a^n = a^{n+m} b a^n$
(Durch das Pumpen kommt von $x = uvw$ zu $uv^2 w$ ein zusätzliches v hinzu und somit wird aus dem ersten a^n nun a^{n+m})

Da $m \geq 1$ (wegen $|v| \geq 1$) ist $uv^2 w = a^{n+m} b a^n$ vorwärts gelesen nicht das gleiche wie rückwärts gelesen und somit gilt $uv^2 w \notin L$. Folglich ist die Sprache nicht regulär.

Allgemeiner Hinweis

Es ist oft hilfreich das Wort x so zu wählen, dass uv nur ein unterschiedliches Symbol enthält (hier $uv = a^l a^m$) um Fallunterscheidungen zu vermeiden. Überlegen Sie sich zum Beispiel man hätte $x = (aba)^n$ gewählt. Auch hier gilt $x \in L$ und $|x| \geq n$, aber wie die Zerlegung $x = uvw$ hier aussieht ist wesentlich komplizierter.

Alternativer Beweis mit Hilfe der Myhill-Nerode Äquivalenz

Wir zeigen, dass unendlich viele Myhill-Nerode Äquivalenzklassen bzgl. L existieren ($\text{index}(R_L) = \infty$) und L somit nicht regulär ist.

Behauptung: Für alle $n, m \geq 0$ mit $n \neq m$ gilt $\neg(a^n b \; R_L \; a^m b)$ und folglich gilt auch $[a^n b] \neq [a^m b]$.

Die Behauptung gilt, da $a^n ba^n \in L$ während $a^m ba^n \notin L$ (wegen $n \neq m$), d.h. durch Anhängen von $w = a^n$ landet man einmal in der Sprache (für $a^n b$) und einmal außerhalb der Sprache (für $a^m b$) und somit sind beide Worte nicht äquivalent.

Es folgt, dass $[a^n b]$ für jedes $n \geq 0$ eine eigene Äquivalenzklasse ist. Folglich gilt $\text{index}(R_L) = \infty$.

(b) $L = \{ww \mid w \in \Sigma^*\}$ ist **nicht regulär**

Pumping Lemma

Wähle $x = a^n ba^n b \in L$, $|x| = 2n + 2 \geq n$.

Betrachte alle Zerlegungen $x = uvw$ mit $|v| \geq 1$ und $|uv| \leq n$:

Wir haben $u = a^l, v = a^m, w = a^s ba^n b$ $(l + m + s = n)$.

Wir wählen den Pumpfaktor $i = 2$ und betrachten $uv^i w$:

$uv^2 w = a^l a^{2m} a^s ba^n b = a^{n+m} ba^n b$.

Da $m \geq 1$ (wegen $|v| \geq 1$) lässt sich $uv^2 w = a^{n+m} ba^n b$ nicht in zwei gleiche Worte zerlegen und somit $uv^2 w \notin L$. Folglich ist die Sprache L nicht regulär.

Alternativer Beweis mit Hilfe der Myhill-Nerode Äquivalenz

Wir zeigen wieder $\text{index}(R_L) = \infty$.

Behauptung: Für alle $n, m \geq 0$ mit $n \neq m$ gilt $\neg(a^n b \; R_L \; a^m b)$ und folglich gilt auch $[a^n b] \neq [a^m b]$.

Die Behauptung gilt, da $a^n b a^n b \in L$ während $a^n b a^m b \notin L$. Beachten Sie, dass bei der Teilung von $a^n b a^m b$ in zwei Worte nur dann gleiche Worte entstehen können wenn $n = m$ gilt, aber hier gilt $n \neq m$.

Es folgt, dass $[a^n b]$ für jedes $n \geq 0$ eine eigene Äquivalenzklasse ist. Folglich gilt $\text{index}(R_L) = \infty$.

(c) $L = \{a^{n^2} \mid n \geq 0\}$ ist **nicht regulär**

Pumping Lemma

Wähle $x = a^{n^2} \in L$, somit $|x| = n^2 \geq n$.

Betrachte alle Zerlegungen $x = uvw$ mit $|v| \geq 1$und $|uv| \leq n$: mit $u = a^j, v = a^k, w = a^l$ $(j + k + l = n^2)$.

Wir wählen den Pumpfaktor $i = 2$ und betrachten $uv^i w$:

Wir haben $uv^2 w = a^j a^{2k} a^l = a^{n^2+k}$.

Nun müssen wir zeigen, dass $n^2 + k$ keine Quadratzahl ist und somit $uv^2 w \notin L$.

Wir zeigen, dass $n^2 < n^2 + k < (n + 1)^2$, d.h. $n^2 + k$ liegt zwischen der Quadratzahl n^2 und der nachfolgenden Quadratzahl $(n + 1)^2$ und kann damit selbst keine Quadratzahl sein. Es gilt $n^2 < n^2 + k$ wegen $k = |v| \geq 1$. Außerdem gilt

$$
\begin{aligned}
& n^2 + k \\
\leq\; & n^2 + n && \text{wegen } |uv| \leq n \text{ (und somit } |v| = k \leq n) \\
<\; & n^2 + 2n + 1 \\
=\; & (n + 1)^2
\end{aligned}
$$

Also gilt $uv^2 w \notin L$ und die Sprache ist folglich nicht regulär.

(d) $L = \{a^n b^m \mid n \neq m\}$ ist **nicht regulär**. Der Beweis mit Hilfe des Pumping Lemmas ist hier recht kompliziert im Vergleich zu den bisherigen Sprachen, da hier bei der Wahl des Pumpfaktors der klassische Ansatz mit $i = 2$ oder $i = 0$ nicht funktioniert. Hier empfiehlt sich der Beweis mit Hilfe der Myhill-Nerode-Äquivalenz, der weiterhin recht einfach ist.

Pumping Lemma

Wähle $x = a^n b^{n!+n} \in L$, somit $|x| = n! + 2n \geq n$.

Betrachte alle Zerlegungen $x = uvw$ mit $|v| \geq 1$ und $|uv| \leq n$:

Wir haben $u = a^j, v = a^k, w = a^l b^{n!+n}$ $(j + k + l = n)$.

Wähle den Pumpfaktor $i = \frac{n!}{k} + 1$ und betrachte $uv^i w$ (da $|uv| \leq n$ gilt auch $|v| = k \leq n$ und somit ist $n!$ durch k teilbar):

$uv^i w = a^j a^{i \cdot k} a^l b^{n!+n}$

Es gilt

$$
\begin{aligned}
& j + i \cdot k + l \\
=& j + \left(\frac{n!}{k} + 1\right) \cdot k + l \\
=& j + n! + k + l \\
=& j + k + l + n! \\
=& n + n!
\end{aligned}
$$

Somit gilt $uv^i w = a^{n!+n} b^{n!+n} \notin L$. Somit ist die Sprache ist nicht regulär.

Alternativer Beweis mit Hilfe der Myhill-Nerode Äquivalenz

Wir zeigen $\text{index}(R_L) = \infty$.

Behauptung: Für alle $n, m \geq 0$ mit $n \neq m$ gilt $\neg(a^n \; R_L \; a^m)$ und folglich gilt auch $[a^n] \neq [a^m]$.

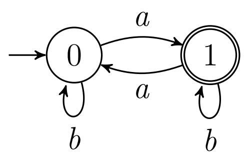
Die Behauptung gilt, da $a^n b^n \notin L$ während $a^m b^n \in L$ (wegen $n \neq m$).

Es folgt, dass $[a^n]$ für jedes $n \geq 0$ eine eigene Äquivalenzklasse ist. Folglich gilt $\text{index}(R_L) = \infty$.

(e) $L = \{w \in \{a,b\}^* \mid \#_a(w) \bmod 4 \in \{1,3\}\}$ ist **regulär**

Es gilt $L = \{w \in \{a,b\}^* \mid \#_a(w) \mod 2 = 1\}$, weil Rest 1 oder 3 bei Teilung durch 4 entspricht genau Rest 1 bei Teilung durch 2.

Es existiert ein DFA für L:



Somit ist L regulär.

Die Myhill-Nerode-Äquivalenzklassen sind:

- $[\varepsilon] = \{\varepsilon, a^2, a^4, aba, \dots\} = \{w \in \{a,b\}^* \mid \#_a(w) \mod 2 = 0\}$ (Zustand 0)
- $[a] = \{a, a^3, a^5, ab, \dots\} = \{w \in \{a,b\}^* \mid \#_a(w) \mod 2 = 1\}$ (Zustand 1)

Allgemeiner Hinweis

Hier sehen wir schön, wie die Myhill-Nerode Äquivalenzklassen den Zuständen eines minimalen Automaten entsprechen.

(f) $L = \{a^p \mid p \text{ ist Primzahl}\}$ ist **nicht regulär**

Pumping Lemma

Wähle $x = a^p$ mit einer Primzahl $p \geq n$, somit $|x| \geq n$.

Betrachte alle Zerelgungen $x = uvw$ mit $|v| \geq 1$ und $|uv| \leq n$:

Wir haben $u = a^k, v = a^l, w = a^m$ $(k + l + m = p)$.

Wähle den Pumpfaktor $i = p + 1$ und betrachte uv^iw:

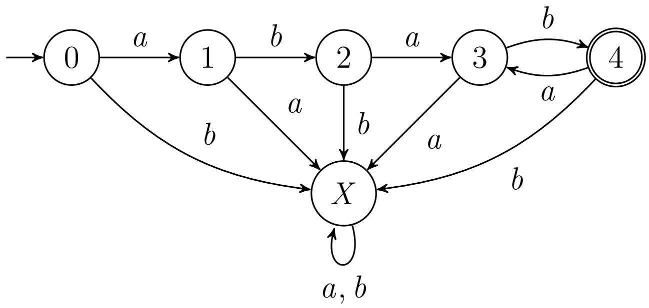
$uv^{p+1}w = a^k a^{l(p+1)} a^m = a^{k+l(p+1)+m} = a^{k+l+m+l\cdot p} = a^{p+l\cdot p} = a^{(l+1)\cdot p}$.

$(l+1)\cdot p$ kann keine Primzahl sein kann, da sie sich in Faktoren $(l+1)$ und p zerlegen lässt, wobei $(l+1) \geq 2$, da nach Voraussetzung $l = |v| \geq 1$.

Daher gilt $uv^iw \notin L$ und somit ist L nicht regulär.

(g) $L = \{(ab)^n \mid n \geq 2\}$ ist **regulär**

Die Sprache ist regulär, da es einen DFA für L gibt:

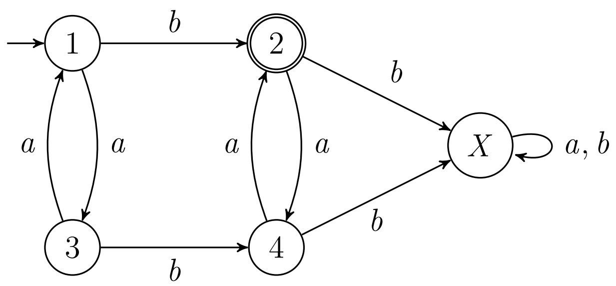


Die Myhill-Nerode-Äquivalenzklassen sind:

- $[\varepsilon] = \{\varepsilon\}$ (Zustand 0)
- $[a] = \{a\}$ (Zustand 1)
- $[ab] = \{ab\}$ (Zustand 2)
- $[aba] = \{aba\}$ (Zustand 3)
- $[abab] = L = \{(ab)^n \mid n \geq 2\}$ (Zustand 4)
- $[b] = \Sigma^* \setminus (L \cup \{\varepsilon, a, ab, aba\})$ (Fangzustand X)

(h) $L = \{a^n ba^m \mid (n + m) \text{ ist gerade}\}$ ist **regulär**

Die Sprache ist regulär, da es einen DFA für L gibt:



Die Myhill-Nerode-Äquivalenzklassen sind:

- $[\varepsilon] = \{a^{2n} \mid n \in \mathbb{N}\}$ (Zustand 1)
- $[a] = \{a^{2n+1} \mid n \in \mathbb{N}\}$ (Zustand 3)
- $[b] = \{a^i ba^j \mid i + j \text{ gerade}\}$ (Zustand 2)
- $[ab] = \{a^i ba^j \mid i + j \text{ ungerade}\}$ (Zustand 4)
- $[bb] = \Sigma^* \setminus L(a^* \mid a^* ba^*) = L(a^* ba^* b(a|b)^*)$ (Fangzustand X)