Universität Siegen
Lehrstuhl Theoretische Informatik
Markus Lohrey

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 regulär, 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 wir die Eigenschaft "k ist eine Quadratzahl" mit einer regulären Sprache nicht ausdrücken können (in Aufgabe 3 wird gezeigt, dass die Sprache $\{a^k \mid k \text{ ist eine Quadratzahl}\}$ nicht regulär ist;

Der Beweis aus Aufgabe 3 lässt sich erweitern, um zu zeigen, dass L nicht regulär ist).

Behauptung: Es ist $L \cdot L = \{a^k \mid k \geq 0\}$. Die Inklusion $L \cdot L \subseteq \{a^k \mid k \geq 0\}$ ist klar, es bleibt die Inklusion $\{a^k \mid k \geq 0\} \subseteq L \cdot L$ zu zeigen.

Sei $w = a^k$. Wir unterscheiden zwei Fälle:

1. Fall: k ist gerade. Wir können w schreiben als $w = a^k\varepsilon$ (wobei ε das leere Wort bezeichnet), dann ist $a^k \in L$ (da k gerade ist) und $\varepsilon \in L$ und somit $w \in L \cdot L$.

2. Fall: k ist ungerade. Dann ist $k = \ell + 1$, wobei ℓ eine gerade Zahl ist. Wir können w schreiben als $w = a^\ell a$, dann ist $a^\ell \in L$ (da ℓ gerade ist) und $a \in L$, da 1 eine Quadratzahl ist. Daher folgt $w \in L \cdot L$.

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?

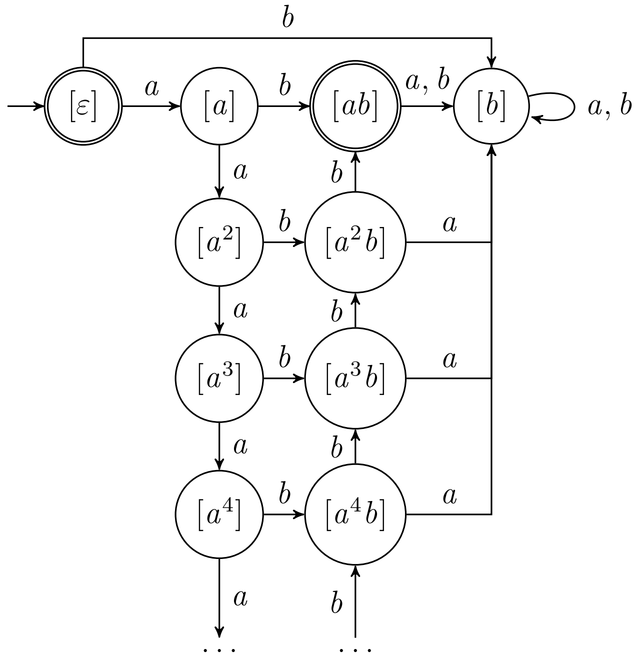
Lösung zu Aufgabe 2. Myhill-Nerode-Äquivalenzklassen

- $[b] = \Sigma^* \setminus \{a^n b^m \mid n \geq m \geq 0\}$
  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. Betrachten wir zum Beispiel 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. Analog lässt sich zeigen, dass die verbleibenden 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) $\{w \in \{a,b\}^* \mid w = w^r\}$

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

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

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

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

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

Lösung zu Aufgabe 3.

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

Pumping Lemma

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

Sei $n \in \mathbb{N}$ beliebig. 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 $u = a^l, v = a^m, w = a^s b a^n$ $(l + m + s = n)$, da $|uv| \leq n$.

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

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

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

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 $[a^n b] \neq [a^m b]$.

Die Behauptung gilt, da $a^n b a^n \in L$ während $a^m b a^n \notin L$ (wegen $n \neq m$) und somit sind beide Worte nicht Myhill-Nerode 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

Sei $n \in \mathbb{N}$ beliebig. Wähle $x = a^n b a^n b \in L$, $|x| = 2n + 2 \geq n$.

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

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

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

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

Da $m \geq 1$ (wegen $|v| \geq 1$) lässt sich $uv^2 w = a^{n+m} b a^n b$ 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

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

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

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

Dieser Beweis lässt sich leicht anpassen, um zu zeigen, dass die Sprache $\{a^{2n} \mid n \geq 0\} \cup \{a^{k^2} \mid k \geq 0\}$ aus Aufgabe 1 nicht regulär ist:

Sei dazu $n \in \mathbb{N}$. Wenn n eine ungerade Zahl ist, wähle $x = a^{n^2}$ und den Pumpfaktor $i = 3$, dann ist $uv^3w = a^{n^2+2k}$ wobei $n^2 + 2k$ keine gerade Zahl ist und es lässt sich wie oben zeigen, dass n^2+2k keine Quadratzahl ist. Wenn n eine gerade Zahl ist, wähle $x = a^{(n+1)^2}$ und den Pumpfaktor $i = 3$, dann ist $uv^3w = a^{(n+1)^2+2k}$, wobei $(n + 1)^2 + 2k$ eine ungerade Zahl ist, und es sich wieder wie oben zeigen lässt, dass $(n+1)^2+2k$ keine Quadratzahl ist.

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

Pumping Lemma

Sei $n \in \mathbb{N}$ beliebig. Wähle $x = a^p$ mit einer Primzahl $p \geq n$, somit $|x| \geq n$.

Betrachte alle Zerlegungen $x = uvw$ 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:

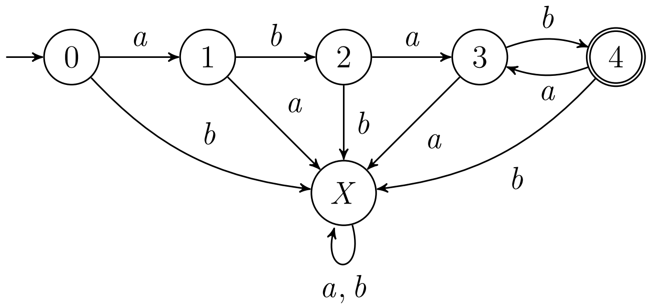
$uv^{p+1}w = a^k a^{l(p+1)} a^m = a^{k+l(p+1)+m} = a^{k+l+m+l\cdot p} = a^{p+l\cdot p} = a^{(l+1)\cdot p}$.

Die Zahl $(l+1)\cdot p$ 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.

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

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

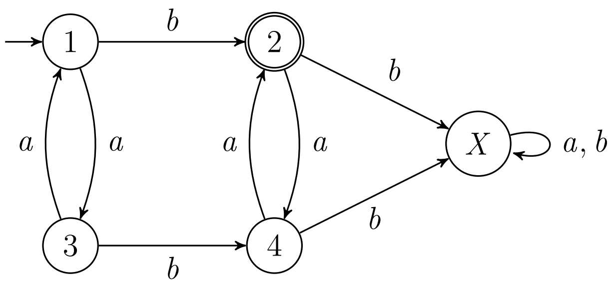


Die Myhill-Nerode-Äquivalenzklassen sind:

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

(f) $L = \{a^n ba^m \mid (n + m) \text{ ist gerade}\}$ 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)
- $[bb] = \Sigma^* \setminus L(a^* \mid a^* ba^*)$ (Fangzustand X)

Aufgabe 4. Die Folge $(F(n))_{n\geq 0}$ der *Fibonacci-Zahlen* ist (mit unserer Konvention) induktiv definiert durch

$$F(0) = 1, \quad F(1) = 2 \quad \text{und } F(n) = F(n-1) + F(n-2) \text{ für } n \geq 2.$$

Beweisen Sie mit Hilfe des Pumping Lemmas, dass die Sprache

$$L = \{a^{F(n)} \mid n \geq 0\}$$

nicht regulär ist.

Lösung zu Aufgabe 4. Die ersten Fibonacci-Zahlen sind also die folgenden:

$$F(0) = 1, F(1) = 2, F(2) = 3, F(3) = 5, F(4) = 8, F(5) = 13, F(6) = 21, \ldots$$

Zunächst zeigen wir induktiv, dass $F(n) > n$ für alle $n \geq 0$: Für den Induktionsanfang betrachten wir $F(0) = 1 > 0$ und $F(1) = 2 > 1$. Als Induktionsvoraussetzung nehmen wir an, dass es ein $n \in \mathbb{N}$ gibt, sodass $F(k) > k$ für alle k mit $0 \leq k \leq n$ und im Induktionsschritt (von n nach $n+1$, für $n \geq 2$) erhalten wir

$$F(n+1) = F(n) + F(n-1) \overset{IV}{\geq} n + 1 + n = 2n + 1 > n + 1.$$

Damit gilt also $F(n) > n$ für alle $n \in \mathbb{N}$.

Wir beweisen nun, dass L nicht regulär ist mit Hilfe des Pumping-Lemmas: Sei nun $n \in \mathbb{N}$ beliebig. Wähle $x = a^{F(n+1)} \in L$, dann gilt $|x| \geq n$. Betrachte alle Zerlegungen $x = uvw$ mit $|v| \geq 1$ und $|uv| \leq n$: Wir haben $u = a^k$, $v = a^l$, $w = a^m$, wobei $k + l + m = F(n+1)$. Wähle den Pumpfaktor $i = 2$ und betrachte uv^iw:

$$uv^2w = a^k a^{2l} a^m = a^{F(n+1)+l}.$$

Zu zeigen ist nun, dass $F(n+1)+l$ keine Fibonacci-Zahl ist, wobei $1 \leq l \leq n$. Wir zeigen, dass $F(n+1) < F(n+1) + l < F(n+2)$: Da $l \geq 1$ gilt $F(n+1) < F(n+1) + l$ und weiterhin gilt

$$F(n+2) = F(n+1) + F(n) > F(n+1) + n \geq F(n+1) + l.$$

Da die Folge der Fibonacci-Zahlen monoton wachsend ist und also $F(n+1) < F(n+1) + l < F(n+2)$ gilt, entspricht $F(n+1) + l$ keiner Fibonacci Zahl. Damit gilt $a^{F(n+1)+l} \notin L$ und somit ist L nicht regulär.