Übungsblatt 7

Aufgabe 1. Im folgenden sei $L \subseteq \{a, b\}^*$ eine reguläre Sprache und sei $[w]$ die Myhill-Nerode Äquivalenzklasse von $w \in \Sigma^*$ bezüglich der Myhill-Nerode Äquivalenz R_L. Sind die folgenden Aussagen wahr oder falsch?

(a) Wenn $[\varepsilon] \cap L = \varnothing$, dann ist der Startzustand in einem DFA zu L kein Endzustand.

(b) Wenn $[a] \cap L = \varnothing$ und $[b] \cap L = \varnothing$, dann gilt $L = \varnothing$.

(c) Wenn $[\varepsilon] = [b]$ und $[\varepsilon] \subseteq L$, dann gilt $b^n \in L$ für jedes $n \in \mathbb{N}$.

(d) Jede kontextfreie Sprache hat unendlich viele Myhill-Nerode Äquivalenzklassen.

(e) Jede Sprache besitzt mindestens zwei Myhill-Nerode Äquivalenzklassen.

Lösung zu Aufgabe 1. (a) Wahr: Es gilt $\varepsilon \in [\varepsilon]$. Da $[\varepsilon] \cap L = \varnothing$ gilt also $\varepsilon \notin L$. Ein DFA zu L darf also das leere Wort nicht akzeptieren, also ist der Startzustand kein Endzustand.

(b) Falsch: Sei $L = \{w \in \{a, b\}^* \mid |w| \geq 2\}$. Dann ist $[a] = [b] = \{a, b\}$ und es gilt $[a] \cap L = \varnothing$ und $[b] \cap L = \varnothing$. Allerdings ist $L \neq \varnothing$.

(c) Wahr: Wenn $[\varepsilon] = [b]$ gilt insbesondere $[b^n] = [\varepsilon]$ für jedes $n \in \mathbb{N}$. Dies lässt sich z.B. induktiv über die Länge n des Worts b^n zeigen. Für den Induktionsanfang betrachten wir $n = 1$ (der Fall $n = 0$ ist klar). Es gilt $[b] = [\varepsilon]$ nach Voraussetzung. In der Induktionsvoraussetzung gelte $[b^n] = [\varepsilon]$ für ein $n \in \mathbb{N}$. Im Induktionsschritt zeigen wir, dass $[b^{n+1}] = [\varepsilon]$. Es gilt $[b^n] = [\varepsilon]$ nach Induktionsvoraussetzung, also auch $[b^n \cdot b] = [\varepsilon \cdot b] = [b] = [\varepsilon]$ (die letzte Gleichung folgt aus dem Induktionsanfang bzw. der Aufgabenstellung). Somit gilt also $[b^n] = [\varepsilon]$ für jedes $n \in \mathbb{N}$, bzw. $\{b^n \mid n \in \mathbb{N}\} \subseteq [\varepsilon]$. Da nach Voraussetzung $[\varepsilon] \subseteq L$ gilt, folgt, dass $b^n \in L$ für jedes $n \in \mathbb{N}$.

(d) Falsch: Jede reguläre Sprache ist insbesondere auch kontextfrei. Nach dem Satz von Myhill-Nerode erzeugt eine reguläre Sprache jedoch nur endlich viele Myhill-Nerode Äquivalenzklassen.

(e) Falsch: Zum Beispiel die Sprache Σ^* erzeugt nur eine Myhill-Nerode Äquivalenzklasse.

Aufgabe 2. Bestimmen Sie für jede der folgenden Sprachen die Anzahl der Myhill-Nerode-Äquivalenzklassen. Falls die Sprache regulär ist, geben Sie zusätzlich einen minimalen DFA an, der die Sprache erkennt. Falls die Sprache nicht regulär ist, beweisen Sie dies zusätzlich mit Hilfe des Pumping-Lemmas.

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

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

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

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

Lösung zu Aufgabe 2.

(a) **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 $\text{index}(R_L) = \infty$.

$L = \{w \in \Sigma^* \mid w = w^r\}$ ist also **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.

(b) **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 $a^m b a^n b \notin L$. Beachten Sie, dass bei der Teilung von $a^m b a^n b$ 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 $\text{index}(R_L) = \infty$.

$L = \{ww \mid w \in \Sigma^*\}$ ist also **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.

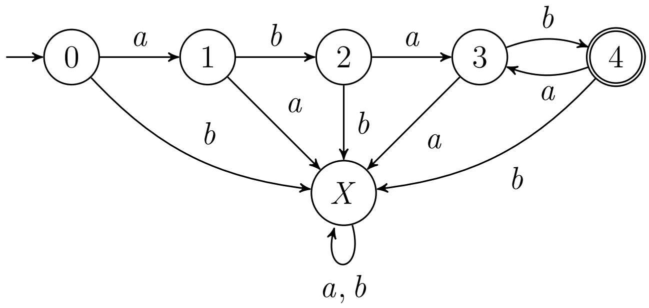
(c) **Myhill-Nerode Äquivalenz**

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)

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

Ein minimaler DFA für L hat 6 Zustände:



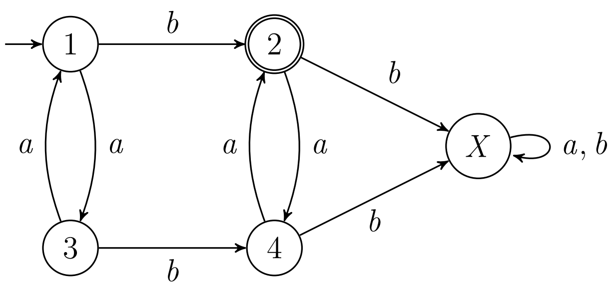
(d) **Myhill-Nerode Äquivalenz**

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)

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

Ein minimaler DFA für L hat 5 Zustände:



Aufgabe 3. Zeigen Sie mit Hilfe des Pumping-Lemmas, dass die folgenden Sprachen nicht regulär sind:

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

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

(c) $\{a^{F(n)} \mid n \geq 0\}$, wobei $(F(n))_{n\geq 0}$ die Folge der *Fibonacci-Zahlen* ist, die wir induktiv definieren durch

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

Lösung zu Aufgabe 3. (a) 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^iw: Wir haben $uv^2w = a^j\, a^{2k}\, a^l = a^{n^2+k}$. Nun müssen wir zeigen, dass $n^2 + k$ keine Quadratzahl ist und somit $uv^2w \notin L$. 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.

(b) 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.

(c) Die ersten Fibonacci-Zahlen (nach der obigen Konvention für die Startwerte) sind also die folgenden:

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

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 $uv^i w$:

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

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.

Aufgabe 4. Bekanntermaßen ist die Sprache

$$L = \{a^{2n} \mid n \in \mathbb{N}\}$$

regulär. Im Folgenden wird versucht zu beweisen, dass sie dies nicht ist. Finden Sie den Fehler.

Beweis: Sei $n \in \mathbb{N}$ beliebig. Wähle $x = a^{2n}$. Dann ist $x \in L$ und $|x| = 2n \geq n$. Betrachten wir alle Zerlegungen $x = uvw$ mit $|uv| \leq n$ und $v \neq \varepsilon$, so ist $uv = a^l$ für ein l mit $1 \leq l \leq n$ und damit $v = a^k$ mit $1 \leq k \leq l$. Damit ist $uv^2w = a^{2n+k} \neq a^{2n}$ und somit $uv^2w \notin L$. Somit ist L nach dem Pumping-Lemma für reguläre Sprachen nicht regulär.

Lösung zu Aufgabe 4. Die Zahl k aus dem Beweis kann eine gerade Zahl sein. In dem Fall gilt $k = 2m$ für ein $m \in \mathbb{N}$ und somit $uv^2w = a^{2n+k} = a^{2n+2m} = a^{2(n+m)} \in L$, da $2(n+m)$ wieder eine gerade Zahl ist.