Lösungen zu den Übungsaufgaben im Buch *Automaten und Sprachen: Theoretische Informatik für die Praxis* von Andreas Müller, ISBN 978-3-662-70145-4 (Softcover), ISBN 978-3-662-70146-1 (eBook), https://link.springer.com/book/10.1007/978-3-662-70146-1. Website zum Buch: https://autospr.ch

Kapitel 2: Nicht reguläre Sprachen

2.1. Zeigen Sie: Das Komplement einer nicht regulären Sprache ist nicht regulär.

Lösung. Die Sprache L sei nicht regulär. Dann kann $\bar{L}$ nur entweder regulär sein, oder nicht regulär. Wenn $\bar{L}$ regulär wäre, dann müsste das das Komplement davon auch regulär sein, also $\bar{\bar{L}} = L$ müsste regulär sein. Dies kann also nicht sein, es bleibt also nur noch die Alternative, dass $\bar{L}$ nicht regulär ist. ○

2.2. Welche der folgenden Aussagen ist wahr bzw. falsch:

a) Jede Teilmenge einer regulären Sprache ist regulär.

b) Jede Obermenge einer regulären Sprache ist regulär.

c) Die Vereinigung zweier nicht regulärer Sprachen ist nicht regulär.

d) Die Schnittmenge zweier nicht regulärer Sprachen ist nicht regulär.

Lösung.

a) Σ^* ist regulär, wäre a) wahr, müsste jede Sprache regulär sein, aber wir kennen nicht reguläre Sprachen, zum Beispiel $\{0^n1^n \mid n \in \mathbb{N}\}$.

b) $\emptyset$ ist regulär, da aber $\emptyset$ in jeder anderen Teilmenge von Σ^* enthalten ist, müssten alle Sprachen regulär sein, im Widerspruch zur Tatsache, dass wir nicht reguläre Sprachen kennen.

c) Wenn L nicht regulär ist, dann ist auch $\bar{L}$ nicht regulär. Wäre nämlich $\bar{L}$ regulär, müsste auch $\bar{\bar{L}} = L$ regulär sein. Es ist aber auch $L \cup \bar{L} = \Sigma^*$ die Vereinigung zweier nicht regulärer Sprachen, trotzdem ist Σ^* natürlich regulär.

d) Das Komplement einer nicht regulären Sprache ist ebenfalls nicht regulär. Wäre die Vereinigung zweier nicht regulärer Sprachen L_1 und L_2 nicht regulär müsste auch $\bar{L}_1 \cup \bar{L}_2$ nicht regulär sein, und damit auch $\overline{\bar{L}_1 \cup \bar{L}_2} = L_1 \cap L_2$. Dass die Schnittmenge nicht regulärer Sprachen aber trotzdem regulär sein kann haben wir bereits in b) eingesehen. ○

2.3. Betrachten Sie über dem Alphabet $\Sigma = \{\texttt{0}, \texttt{1}, \text{␣}\}$ die Sprache L, deren Wörter aus zwei durch genau ein Leerzeichen getrennten Binärzahlen bestehen, wobei die zweite größer sein muss als die erste. Zur Sprache gehören also zum Beispiel

$$\texttt{10000␣10001},\quad \texttt{0␣10001}$$

aber nicht

$$\texttt{10000␣10000},\quad \texttt{10000␣␣␣10001},\quad \texttt{10001␣0}.$$

Ist L regulär?

Lösung. Nein, wie man mit dem Pumping Lemma nachweisen kann. Wäre L regulär, müsste es eine Pumping Length N geben, so dass Wörter mit mindestens Länge N aufgepumpt werden können. Wir wählen das Wort

$$w = \mathtt{1}^N \textvisiblespace \mathtt{1}\mathtt{0}^N \qquad\qquad w = \underbrace{\mathtt{1111}\dots\mathtt{1}}_{N\text{ Einsen}}\textvisiblespace\mathtt{1}\underbrace{\mathtt{0000}\dots\mathtt{0}}_{N\text{ Nullen}}$$

es erfüllt die Bedingung $w \in L$. Da $|w| = 2N + 2 > N$ muss es eine Unterteilung $w = xyz$ geben so, dass $|xy| \le N$ ist, also x und y müssen vollständig aus Einsen bestehen. Es ist also $y = \mathtt{1}^n$ mit $n > 0$. Pumpt man diese Wort auf, erhöht sich nur die Zahl der Einsen im Teil vor dem Leerzeichen:

$$xy^2z = \mathtt{1}^{N+n} \textvisiblespace \mathtt{1}\mathtt{0}^N \qquad\qquad xy^2z = \underbrace{\mathtt{1111}\dots\mathtt{1}}_{N+n}\textvisiblespace\mathtt{1}\underbrace{\mathtt{0000}\dots\mathtt{0}}_{N}$$

Die Zahl links vom Blank ist $2^{N+n} - 1$, die Zahl rechts ist 2^N, aber $2^{N+n} - 1 > 2^N$, das aufgepumpte Wort kann also nicht mehr in L sein.

Dieser Widerspruch zeigt, dass L nicht regulär sein kann. ○

2.4. Sei $\Sigma = \{\mathtt{a}, \mathtt{b}, \dots, \mathtt{z}\}$ das Alphabet bestehend aus allen Kleinbuchstaben. Die Sprache

$$L = \{w \in \Sigma^* \mid \text{es gibt zwei verschiedene Buchstaben } a, b \in \Sigma \text{ mit } |w|_a = |w|_b > 1\}$$

besteht aus Wörtern, die mindestens zwei Buchstaben mehr als einmal und in gleicher Anzahl enthalten. Die Wörter

`essen`, `rapperswil`, `seenachtfest`

sind in L, nicht aber

`trinken`, `pfaeffikon`, `montag`.

Ist die Sprache L regulär?

Lösung. Nein, wie man mit dem Pumping-Lemma für reguläre Sprachen bewesein kann. Wir nehmen dazu an, dass L regulär sei, das Pumping Lemma für reguläre Sprachen besagt dann, dass es eine Zahl N, die Pumping Length, gibt, so dass Wörter größerer Länge die Pumpeigenschaft haben.

Wir konstruieren daher ein Wort in L, welches diese Eigenschaft hat:

$$w = \mathtt{a}^N\mathtt{b}^N$$

Das Wort ist in L, wenn $N > 1$ ist, was wir im Folgenden annehmen. Das Pumping-Lemma besagt jetzt, dass es eine Zerlegung in drei Teile $w = xyz$ gibt, wobei $|xy| \le N$ und $|y| > 1$ ist. Insbesondere bestehen x und y ausschließlich aus Buchstaben $\mathtt{a}$. Ein aufgepumptes Wort hat die Form

$$w_k = xy^kz = \mathtt{a}^{N+|y|(k-1)}\mathtt{b}^N,$$

d. h. für $k > 1$ ist die Zahl der einzigen beiden vorkommenden Buchstaben $\mathtt{a}$ und $\mathtt{b}$ nicht mehr gleich, also $w_k \notin L$, im Widerspruch zur Behauptung des Pumping-Lemmas. Dieser Widerspruch zeigt, dass L nicht regulär sein kann. ○

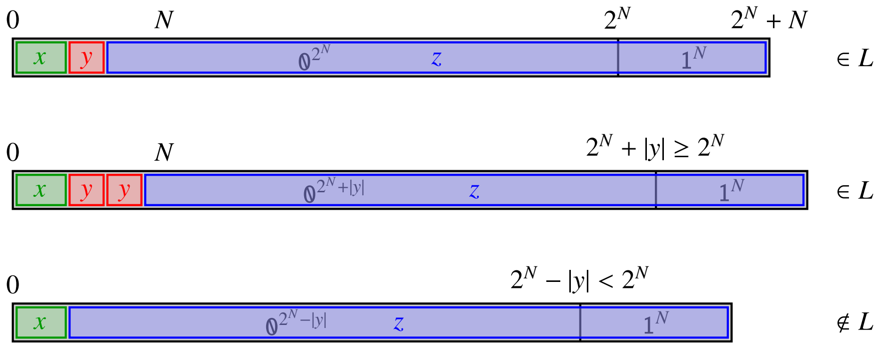
2.5. Ist die Sprache

$$L = \{w \in \Sigma^* \mid |w|_{\mathtt{0}} \ge 2^{|w|_{\mathtt{1}}}\}$$

über dem Alphabet $\Sigma = \{\mathtt{0}, \mathtt{1}\}$ regulär?

Lösung. Die Sprache ist nicht regulär, wie man mit dem Pumping Lemma für reguläre Sprachen zeigen kann.

1. Annahme ist regulär.
2. Nach dem Pumping Lemma gibt es die Pumping Length N.
3. Wähle das Wort $w = \mathtt{0}^{2^N}1^N \in L$
4. Es gibt eine Aufteilung des Wortes $w = xyz$ mit $|xy| \le N$ und $|y| > 0$. Der Teil y besteht ausschließlich aus Nullen.



5. Beim Aufpumpen nimmt die Zahl der Nullen zu, dies steht aber nicht im Widerspruch zu der Sprachdefinition. Beim Abpumpen nimmt jedoch die Zahl der Nullen ab, es ist dann $|xz|_{\mathtt{w}} = 2^N - |y|$. Die Zahl der Einsen ist immer noch N, damit haben wir
$$|xy|_{\mathtt{0}} = 2^N - |y| < 2^N = |xy|_1,$$
im Widerspruch zur Aussage des Pumping Lemmas
6. Dieser Widerspruch zeigt, dass die Sprache L nicht regulär ist. ○

Es gibt natürlich viele weitere Möglichkeiten, den gesuchten Widerspruch herzustellen. Das oben gewählte Wort ist möglicherweise dasjenigen, welches die meisten Studenten als erstes wählen würden, aber das Wort $w = 1^N\mathtt{0}^{2^N}$ ist noch etwas einfacher. Beim Aufpumpen nimmt die Anzahl der Einsen zu, die Anzahl der Nullen aber nicht. Nach Sprachdefinition müsste aber auch die Anzahl der Nullen zunehmen, da ja $|w|_{\mathtt{0}} \ge 2^{|w|_1}$ sein muss. Der Widerspruch zeigt wieder, dass L nicht regulär sein kann.

2.6. In dieser Aufgabe geht es um die Frage, ob man binäre Ganzzahlen mit Hilfe eines endlichen Automaten auf Gleichheit testen kann. Sei $\Sigma = \{\mathtt{0}, 1, \mathtt{=}\}$. Ist die Sprache
$$L = \{w\mathtt{=}w \mid w \in \{\mathtt{0}, 1\}^*\} \tag{1}$$
regulär?

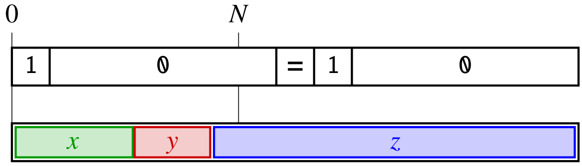
Hinweis. Bitte beachten Sie die beiden unterschiedlichen Bedeutungen von = bzw. $\texttt{=}$ in Formel (1). Das eine ist das mathematische Gleichheitszeichen, das andere ein Symbol im Alphabet Σ. Die beiden Zeichen sind zwar in zwei verschiedenen Schriftarten gesetzt, die normalerweise Zeichen von Formelsymbolen zu unterscheiden gestatten, aber im Falle des Gleichheitszeichens sind die Unterschiede sehr klein.

Die Gleichung (1) besagt, dass die Sprache L aus Zeichenketten von Nullen, Einsen und Gleichheitszeichen besteht, die wie folgt strukturiert sind. Sie beginnen mit einer Folge von Nullen und Einsen, dann kommt ein Gleichheitszeichen. Anschließend wird die Folge von Nullen und Einsen nochmals wiederholt.

Lösung. Die Sprache ist nicht regulär, wie man mithilfe des Pumping Lemmas beweisen kann. ❶ Dazu nehmen wir an, die Sprache sei regulär. ❷ Das Pumping Lemma garantiert, dass es eine Zahl N gibt, so dass sich Wörter länger als diese Zahl in einer speziellen Art zerlegen und "aufpumpen" lassen. ❸ Wir wählen das Wort

$$w = \texttt{10}^N\texttt{=}\texttt{10}^N,$$

welches offenbar zur Sprache gehört und auch länger ist als N. ❹ Das Pumping Lemma garantiert, dass $w = xyz$ geschrieben werden kann mit $|xy| \le N$ und $|y| > 0$:



Aus der Konstruktion sieht man, dass y eine der beiden Formen

$$\texttt{0}^k \quad \text{oder} \quad \texttt{10}^k$$

haben muss (die zweite nur falls $|x| = 0$). ❺ Welche Form auch immer y hat, nach dem Aufpumpen wird auf der linken Seite des Gleichheitszeichens eine Binärzahl mit einer führenden $\texttt{1}$ und mit mehr Stellen stehen, die Gleichheit ist also nicht mehr erfüllt, das aufgepumpte Wort ist nicht in der Sprache, $xy^lz \notin L$. ❻ Dieser Widerspruch zur Aussage des Pumping Lemmas zeigt, dass die Annahme, L sei regulär, falsch gewesen sein muss. ○

2.7. Ist die Sprache

$$L = \{w \in \{\texttt{1}\}^* \mid |w| \text{ ist prim.}\}$$

regulär?

Lösung. Nein, wie wir mit dem Pumping Lemma zeigen können. Dazu nehmen wir an, die Sprache L sei regulär. Nach dem Pumping Lemma gibt es daher die Pumping Length N, Wörter mit Länge mindestens N haben die Pump-Eigenschaft. Sei also

$$w = \texttt{1}^p \qquad \text{mit einer Primzahl } p \ge N,$$

so eine Primzahl p gibt es, weil es unendlich viele Primzahlen gibt. Nach dem Pumping Lemma gibt es jetzt eine Unterteilung $w = xyz$, wobei $|y| > 0$, mit der Eigenschaft, dass auch alle aufgepumpten

Wörter xy^kz in L sind, d. h. prime Länge haben. Wenn man aber genau mit $k = p + 1$ pumpt, erhält man als Länge

$$|xy^kz| = |x| + k|y| + |z| = |x| + (p + 1)|y| + |z| = p|y| + \underbrace{|x| + |y| + |z|}_{= p} = p(|y| + 1)$$

Die Länge von $xy^{(p+1)}z$ ist daher durch p und $|y| + 1$ teilbar, kann also keine Primzahl sein, $xy^{(p+1)}z \notin L$. Dieser Widerspruch zeigt, dass die Annahme, L sei regulär, nicht aufrecht erhalten werden kann. ○

2.8. Sei $\Sigma = \{0, 1\}$ und

$$L = \{w \in \Sigma^* \mid w \text{ ist die Binärdarstellung einer Primzahl.}\}$$

Ist L regulär?

Lösung. Wir lösen die Aufgabe für beliebige Basis b des Zahlensystems, nicht nur wie in der Aufgabe gefordert für $b = 2$.

Im Folgenden sollen Eigenschaften der Darstellung von Primzahlen zur Basis b untersucht werden. Jede natürliche Zahl kann als Zeichenkette aus Zeichen aus der Menge $\Sigma = \{\mathtt{0}, \mathtt{1}, \mathtt{2}, \dots, b - 1\}$ der b-adischen Ziffern geschrieben werden. Die Aufgabe stellt daher die Frage, ob die Menge

$$L = \{w \in \Sigma^* \mid w \text{ ist } b\text{-adische Darstellung einer Primzahl}\}$$

eine reguläre Sprache ist. Die Antwort ist:

Satz 2.1. *Die b-adischen Darstellungen der Primzahlen bilden keine reguläre Sprache.*

Der Satz 2.1 verbindet b-adische Darstellungen einer Zahl, also Zeichenketten, mit Werten. Wir definieren für diesen Übergang die Funktion

$$v_b : \Sigma^* \to \mathbb{N} : w \mapsto v_b(w),$$

wobei $v_b(w)$ der Zahlenwert der als b-adische Zahl aufgefassten Zeichenkette sei.

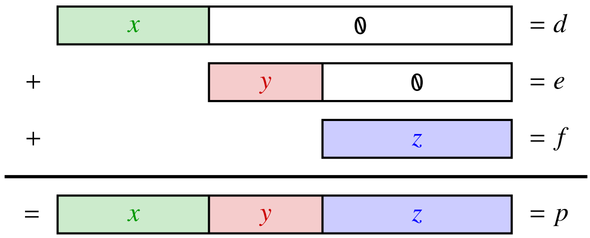
Beweis. Für den Beweis des Satzes verwenden wir das Pumping Lemma. Wir nehmen also an, dass L regulär ist. Wörter mindestens so lange wie die pumping length N angibt haben die Pumpeigenschaft. Wir wählen daher eine Zeichenkette w ausreichender Längen, $|w| \geq N$, die die b-adische Darstellung einer Primzahl ist, wir bezeichnen die Primzahl mit $p = v_b(w)$.

Nach dem Pumping Lemma gibt es eine Unterteilung der Zeichenkette w in drei Teile $x, y, z \in \Sigma^*$ mit $w = xyz$. Von den vom Pumping Lemma garantierten Eigenschaften brauchen wir nur $|y| > 0$. Wir setzen $w_k = xy^kz$, insbesondere ist $w = w_1$. Das Pumping Lemma behauptet, dass alle Wörter w_k wieder b-adische Darstellungen von Primzahlen sind. Es genügt also ein k zu finden, so dass $v_b(w_k)$ nicht prim ist.

Den Zeichenketten x, y und z entsprechen die drei Zahlen

$$\begin{aligned} d &= v_b(x)b^{|y|+|z|}, \\ e &= v_b(y)b^{|z|} \quad \text{und} \\ f &= v_b(z), \end{aligned}$$

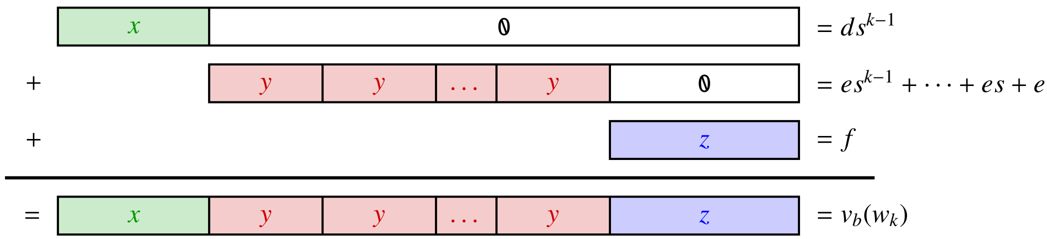
die zusammen die Summe p haben, $d + e + f = p$. Dies lässt sich auch graphisch verdeutlichen:



Die Pumpoperation kann jetzt algebraisch ausgedrückt werden. Sie verschiebt Zeichen um $|y|$ Stellen, was in der Basis b der Multiplikation mit $s = b^{|y|}$ entspricht. Wegen $|y| > 0$ ist $s > 1$. Es gilt

$$v_b(w_k) = ds^{k-1} + e(s^{k-1} + s^{k-2} + \cdots + s + 1) + f. \tag{2}$$

wie man auch aus der graphischen Darstellung



ablesen kann. Wir möchten dies für den Fall $k = p$ berechnen, allerdings nur modulo p. Die Summenformel für die geometrische Reihe liefert für den Klammerausdruck in (2)

$$s^{k-1} + s^{k-2} + \cdots + s + 1 = \frac{1 - s^k}{1 - s}. \tag{3}$$

Da $1 < s < p$ ist, ist $1 - s$ invertierbar modulo p, und (3) ist auch richtig modulo p.

Wir verwenden jetzt $k = p$ und wenden den kleinen Satz von Fermat an, der besagt, dass

$$\begin{aligned} a^p &\equiv a \mod p \\ \text{oder} \qquad a^{p-1} &\equiv 1 \mod p \end{aligned}$$

ist für jedes $a \neq 0$, also auch für $a = s$. Setzt man dies in (3) und (2) ein, folgt

$$\begin{aligned} s^{p-1} + s^{p-2} + \cdots + s + 1 \qquad &= \frac{1 - s^p}{1 - s} \equiv \frac{1 - s}{1 - s} = 1 \mod p, \\ v_b(w_p) = ds^{p-1} + e(s^{p-1} + s^{p-2} + \cdots + s + 1) + f &\equiv d + e + f = p \equiv 0 \mod p \end{aligned}$$

Also ist $v_b(w_p)$ durch p teilbar, und damit stellt w_p keine Primzahl dar.

Dieser Widerspruch zeigt, dass die Annahme, L sei regulär, nicht zutreffend ist. Damit ist der Satz bewiesen. □

Die Aufgabe ist der Spezialfall $b = 2$, und mit Satz 2.1 auch beantwortet. ○

2.9. Ist die Sprache

$$L = \{\mathtt{1}^n\mathtt{+}\mathtt{1}^m\mathtt{=}\mathtt{1}^{n+m} \mid n, m \geq 0\}$$

über dem Alphabet $\Sigma = \{\mathtt{1}\}$ regulär?

Lösung. Die Sprache L ist nicht regulär, wie man mit dem Pumping-Lemma für reguläre Sprachen zeigen kann.

1. Wir nehmen an, L sei regulär.
2. Nach dem Pumping-Lemma gibt es die Pumping-Length N.
3. Wir konstruieren das Wort $w = \texttt{1}^N\texttt{+1}^N\texttt{=1}^{2N}$.
4. Nach dem Pumping-Lemma gibt es eine Unterteilung $w = xyz$, wobei $|xy| \le N$ ist und $|y| = b > 0$. Der Teil y besteht also ausschließlich aus $\texttt{1}$, $y = \texttt{1}^b$.
5. Beim Pumpen wird aus w das Wort
$$xy^kz = \texttt{1}^{N+(k-1)b}\texttt{+1}^N\texttt{=1}^{2N}.$$
Für $k \ne 1$ ist $N + (k-1)b + N = 2N + (k-1)b \ne 2N$, das gepumpte Wort ist also nicht in L.
6. Aus diesem Widerspruch zur Aussage des Pumping-Lemmas folgt, dass L nicht regulär sein kann.

○