Tutorium
Automaten und Formale Sprachen
Das Pumping-Lemma

Christopher Blöcker, B. Sc.
inf9900@fh-wedel.de

SS 2012

Frage

Wie lässt sich feststellen, ob

$$\mathcal{L} \in \mathcal{L}_3?$$

Antwort

Eine Sprache ist vom Typ-3, wenn sie von einem endlichen Automaten akzeptiert wird.

Da Endliche Automaten, Reguläre Ausdrücke und Typ-3-Grammatiken äquivalent sind (sie lassen sich ineinander überführen) ist für den Nachweis $\mathcal{L} \in \mathcal{L}_3$ ein Endlicher Automat $\mathcal{A}$, ein Regulärer Ausdruck $\mathcal{R}$ oder eine Typ-3-Grammatik $\mathcal{G}$ mit

$$\mathcal{L}_{\mathcal{A}} = \mathcal{L} \text{ bzw. } \mathcal{L}_{\mathcal{R}} = \mathcal{L} \text{ bzw. } \mathcal{L}_{\mathcal{G}} = \mathcal{L}$$

anzugeben.

Frage

Wie lässt sich feststellen, ob

$$\mathcal{L} \notin \mathcal{L}_3?$$

Antwort

Mit Hilfe des Pumping-Lemmas.

Das Pumping-Lemma gilt für jede reguläre Sprache.

$$\mathcal{L} \in \mathcal{L}_3 \rightarrow \mathrm{PL}\left(\mathcal{L}\right).$$

Wenn die Bedingungen des Pumping-Lemmas nicht erfüllt sind, so kann die betroffene Sprache nicht regulär sein.

Durch Kontraposition erhält man

$$\overline{\mathrm{PL}(\mathcal{L})} \rightarrow \mathcal{L} \notin \mathcal{L}_3.$$

Das Pumping-Lemma

Sei $\mathcal{L} \subseteq \Sigma^*$ eine reguläre Sprache.
Dann gibt es eine natürliche Zahl $p \in \mathbb{N}$ derart, dass sich jedes Wort $x \in \mathcal{L}$ mit $|x| \geq p$ zerlegen lässt in drei Teilworte

$$x = uvw$$

mit

1. $|v| > 0$
2. $|uv| \leq p$
3. $\bigwedge\limits_{i \in \mathbb{N}} uv^i w \in \mathcal{L}$

Vorsicht!

Es gilt

$$\mathcal{L} \in \mathcal{L}_3 \to \mathrm{PL}\left(\mathcal{L}\right).$$

Die Umkehrung gilt jedoch **nicht**! Die folgende Aussage ist also **falsch**:

$$\lightning\ \mathrm{PL}\left(\mathcal{L}\right) \to \mathcal{L} \in \mathcal{L}_3.\ \lightning$$

Aussage des Pumping-Lemmas

Sei $\mathcal{L} \in \mathcal{L}_3$ eine reguläre Sprache. Dann gilt das Pumping-Lemma für $\mathcal{L}$, also:

$$\bigvee_{p \in \mathbb{N}} \quad \bigwedge_{\substack{x \in \mathcal{L} \\ |x| \geq p}} \quad \bigvee_{\substack{x = uvw \\ |uv| \leq p \\ |v| > 0}} \quad \bigwedge_{i \in \mathbb{N}} uv^i w \in \mathcal{L}.$$

In Worten

Wenn $\mathcal{L}$ eine reguläre Sprache ist, dann

- **gibt es** eine Zahl p
  (die Pumping-Länge, die wir i.d.R. **nicht konkret** kennen)
- und es gilt **für jedes** Wort $x \in \mathcal{L}$ mit der **Mindestlänge** p,
- dass es eine **zulässige** Zerteilung von x in drei Teilworte **gibt**,
- sodass sich x für **jede** Zahl $i \in \mathbb{N}$ „aufpumpen“ lässt und das entstehende Wort x' ebenfalls in $\mathcal{L}$ liegt.

Verwendung der negierten Aussage

Wenn das Pumping-Lemma für eine Sprache $\mathcal{L}$ nicht gilt, so kann die betroffene Sprache nicht regulär sein, also wenn

$$\bigvee_{\substack{x \in \mathcal{L} \\ |x| \geq p}} \quad \bigwedge_{\substack{x = uvw \\ |uv| \leq p \\ |v| > 0}} \quad \bigvee_{i \in \mathbb{N}} uv^i w \notin \mathcal{L}.$$

In Worten

Eine Sprache $\mathcal{L}$ ist **nicht** regulär, wenn

- es ein Wort $x \in \mathcal{L}$ mit der **Mindestlänge** p **gibt**,
- bei dem zu **jeder zulässigen** Zerlegung in drei Teilworte
- eine Zahl $i \in \mathbb{N}$ **existiert**, sodass das „aufgepumpte" Wort x' **nicht** in $\mathcal{L}$ liegt.

Hinweis

Wir können nicht nur „aufpumpen", sondern auch „Luft herauslassen".

Problem

Es gibt Sprachen mit $\mathcal{L} \notin \mathcal{L}_3$, für die aber das Pumping-Lemma erfüllt ist.

Beispiel

$$L = \{a^m b^n c^n \mid m, n \geq 1\} \cup \{b^m c^n \mid m, n \geq 0\}$$

NERODE-Relation

Seien $x, y \in \Sigma^*, \mathcal{L} \subseteq \Sigma^*, \mathcal{R} \subseteq \Sigma^* \times \Sigma^*$.

$$\mathcal{R}_{\mathcal{N}} : x \sim y \Leftrightarrow \bigwedge_{z \in \Sigma^*} (xz \in \mathcal{L}) \Leftrightarrow (yz \in \mathcal{L})$$

Satz von MYHILL und NERODE

Wenn der Index der NERODE-Relation endlich ist, also wenn es nur endlich viele Äquivalenzklassen bezüglich $\mathcal{R}_{\mathcal{N}}$ gibt, so gilt $\mathcal{L} \in \mathcal{L}_3$.

Aufgaben

Zeigen Sie, dass die folgenden Sprachen nicht regulär sind.

1. $L_1 = \{a^n b^n \mid n \in \mathbb{N}\}$
2. $L_2 = \{w \in \{0,1\}^* \mid w \text{ ist kein Palindrom}\}$
3. $L_3 = \{a^p \mid p \text{ ist eine Primzahl}\}$
4. $L_4 = \left\{a^{n^2} \mid n \in \mathbb{N}\right\}$
5. $L_5 = \left\{a^{2^n} \mid n \in \mathbb{N}\right\}$

Herangehensweise

1. Annahme, L sei regulär
2. Wenn L regulär ist, so muss das Pumping-Lemma gelten
3. Finden eines Wortes, für das es keine zum Pumpen geeignete Unterteilung gibt (Wie? Kreativität! Es gibt keinen Algorithmus.)
4. Gegenbeispiel gefunden
   $\rightarrow$ Annahme muss falsch sein
   $\rightarrow$ L kann nicht regulär sein

$L_1 = \{a^n b^n \mid n \in \mathbb{N}\}$

$$\text{Sei } x = a^p b^p \text{ mit } |x| = 2p,$$

$$\text{und die Unterteilung } x = uvw = \underbrace{a^j}_{u} \cdot \underbrace{a^{p-j-k}}_{v} \cdot \underbrace{a^k b^p}_{w}$$

Damit sind alle möglichen Unterteilungen erfasst.

Unabhängig von j und k wählen wir $i = 0$ zum „aufpumpen“.

Da wegen Bedingung 2. des Pumping-Lemmas gelten muss $j + k < p$, entstehen dabei Wörter der Form $x' = a^q b^p$ mit $q < p$.

Die entstehenden Wörter x' liegen nicht in L_1, somit gilt

$$L_1 \notin \mathcal{L}_3$$

□

$L_2 = \{w \in \{0,1\}^* \mid w \text{ ist kein Palindrom}\}$

Aufgrund der Abschlusseigenschaften regulärer Sprachen gilt

$$L_2 \in \mathcal{L}_3 \Leftrightarrow \overline{L_2} \in \mathcal{L}_3$$

$$\overline{L_2} = \{w \in \{0,1\}^* \mid w \text{ ist ein Palindrom}\}$$

$$\text{Sei } x = 1^p 0 1^p \text{ mit } |x| = 2p+1$$

$$\text{und die Unterteilung } x = uvw = \underbrace{1^j}_{u} \cdot \underbrace{1^{p-j-k}}_{v} \cdot \underbrace{1^k 0 1^p}_{w}$$

mit $j + k < p$. Damit sind alle möglichen Unterteilungen erfasst. Beim Pumpen entstehen für jede Unterteilung Wörter, die kein Palindrom sind.

Es folgt

$$\overline{L_2} \notin \mathcal{L}_3 \rightarrow L_2 \notin \mathcal{L}_3$$

$\square$

$L_3 = \{a^p \mid p \text{ ist eine Primzahl}\}$

$$\text{Sei } x = a^p \text{ mit } |x| = p$$

$$\text{und die Unterteilung } x = uvw = \underbrace{a^j}_{u} \cdot \underbrace{a^{p-j-k}}_{v} \cdot \underbrace{a^k}_{w}$$

Die Länge des entstehenden Wortes x' ist dann

$$|x'| = (j+k) + i(p-j-k)$$

Wir wählen zum Pumpen für i einen Wert, sodass $|x'|$ nicht prim sein kann.
Sei $i = j + k$, dann ergibt sich als Wortlänge

$$|x'| = (j+k) + (j+k)(p-j-k) = (j+k)(1+p-j-k)$$

Die Länge lässt sich in zwei Faktoren zerteilen und kann somit nicht prim sein.
Für den Sonderfall $j + k = 1$ wählen wir $i = 0$. □

$L_4 = \left\{ a^{n^2} \mid n \in \mathbb{N} \right\}$

Beobachtung: Zu L_4 gehören Wörter der Länge $0, 1, 4, 9, 16, 25, 36, 49, \ldots$

$$\text{Sei } x = a^{p^2} \text{ mit } |x| = p^2$$

$$\text{und die Unterteilung } x = uvw = \underbrace{a^j}_{u} \cdot \underbrace{a^{p^2-j-k}}_{v} \cdot \underbrace{a^k}_{w}$$

Zum Pumpem wählen wir $i = 2$.
Wegen Bedingung 1. des Pumping-Lemma, $|v| > 0$, gilt

$$|x'| = |uv^2w| > p^2.$$

Und wegen Bedingung 2. des Pumping-Lemma, $|uv| \leq p \rightarrow |v| \leq p$:

$$\underbrace{|x'| = |uv^2w| \leq p^2 + p}_{\text{aufgepumptes Wort}} < \underbrace{p^2 + 2p + 1 = (p+1)^2}_{\text{nächste gültige Länge}}$$

Also: $x' \notin L_4 \rightarrow L_4 \notin \mathcal{L}_3$. $\square$

$L_5 = \left\{ a^{2^n} \mid n \in \mathbb{N} \right\}$

Beobachtung: Zu L_5 gehören Wörter der Länge $1, 2, 4, 8, 16, 32, 64, \ldots$

$$\text{Sei } x = a^{2^p} \text{ mit } |x| = 2^p$$
$$\text{und die Unterteilung } x = uvw = \underbrace{a^j}_{u} \cdot \underbrace{a^{2^p - j - k}}_{v} \cdot \underbrace{a^k}_{w}$$

Zum Pumpem wählen wir $i = 2$.
Wegen Bedingung 1. des Pumping-Lemma, $|v| > 0$, gilt

$$|x'| = |uv^2w| > 2^p.$$

Und wegen Bedingung 2. des Pumping-Lemma, $|uv| \leq p \rightarrow |v| \leq p$:

$$\text{mit } p < 2^p\text{: } \underbrace{|x'| = |uv^2w| \leq 2^p + p}_{\text{aufgepumptes Wort}} < \underbrace{2 \cdot 2^p = 2^{p+1}}_{\text{nächste gültige Länge}}$$

Also: $x' \notin L_5 \rightarrow L_5 \notin \mathcal{L}_3$. $\square$