LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Formale Sprachen und Komplexität
Sommersemester 2019

Zentralübung 16.05.2019: Reguläre Ausdrücke, Anwenden des Pumping-Lemmas

Prof. Dr. David Sabel

LFE Theoretische Informatik

Letzte Änderung der Folien: 16. Mai 2019

Aufgabe

Geben Sie reguläre Ausdrücke für jede der folgenden formalen Sprachen an, die genau die jeweilige formale Sprache erzeugen:

a) $\{u \in \{a,b\}^* \mid |u| = 4\}$ $\quad (a|b)(a|b)(a|b)(a|b)$

b) $\{u \in \{a,b\}^* \mid |u| \leq 4\}$ $\quad (\varepsilon|a|b)(\varepsilon|a|b)(\varepsilon|a|b)(\varepsilon|a|b)$
oder
$(\varepsilon \mid (a|b) \mid (a|b)(a|b) \mid (a|b)(a|b)(a|b) \mid (a|b)(a|b)(a|b)(a|b))$

c) $\{u \in \{a,b\}^* \mid |u| \geq 4\}$ $\quad (a|b)(a|b)(a|b)(a|b)(a|b)^*$

d) $\{u \in \{a,b\}^* \mid u$ enthält nicht $bbbb$ als Teilwort$\}$
$(\varepsilon|b|bb|bbb)(a|ab|abb|abbb)^*$
oder $(a|ba|bba|bbba)^*(\varepsilon|b|bb|bbb)$

Daten

Geben Sie einen regulären Ausdruck an, der die Sprache aller gültigen Daten im Format TT-MM-JJJJ erzeugt, außer den 29. Februar für Schaltjahre.

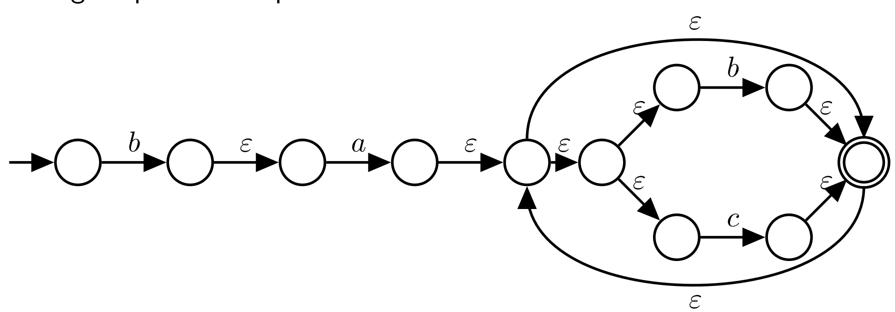
$$
\begin{array}{l}
(\quad (((0(1|...|9))\mid 1(0|...|9)\mid 2(0|...|8))(0(1|...|9))|1(0|1|2))\\
\quad |29(01|03|04|05|06|07|08|09|10|11|12)\\
\quad |30(01|03|04|05|06|07|08|09|10|11|12)\\
\quad |31(01|03|05|07|08|10|12)\\
)\\
(0|...|9)(0|...|9)(0|...|9)(0|...|9)
\end{array}
$$

Aufgabe: Regulärer Ausdruck $\to$ Endlicher Automat

Konstruieren Sie mit dem Verfahren aus der Vorlesung einen NFA mit ε-Übergängen und eindeutigen Start- und Endzuständen, der als Sprache genau die durch den regulären Ausdruck

$$ba(b|c)^*$$

erzeugte Sprache akzeptiert.

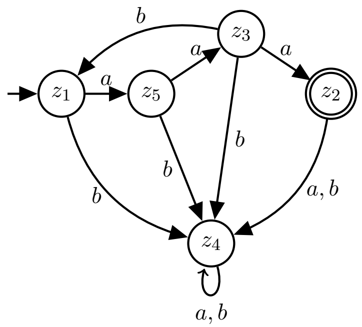


DFA $\to$ regulärer Ausdruck

Wesentliche Ideen: Seien $\{z_1, \ldots, z_n\}$ die Zustände des DFA, konstruiere reguläre Ausdrücke $\alpha_{i,j}^k$:

- $\alpha_{i,j}^k$ repräsentiert alle Worte, mit denen man im DFA von z_i zu z_j kommt, ohne als Zwischenzustand ein z_r zu besuchen, mit $r > k$.
- $\alpha_{i,i}^0$ ist $(\varepsilon|a_1|\ldots|a_m)$ für alle $a_x \in \Sigma$ mit $\delta(z_i, a_x) = z_i$
- $\alpha_{i,j}^0$ ist entweder $(a_1|\ldots|a_m)$ für alle $a_x \in \Sigma$ mit $\delta(z_i, a_x) = z_j$ oder $\emptyset$ (falls es kein solches a_x gibt)
- $\alpha_{i,j}^{k+1} = \alpha_{i,j}^k \mid \alpha_{i,k+1}^k(\alpha_{k+1,k+1}^k)^*\alpha_{k+1,j}^k$
- Sei z_1 der Startzustand und $z_l, z_{l+1}, \ldots, z_q$ die Endzustände. Dann ist $\alpha_{1,l}^n|\ldots|\alpha_{1,q}^n$ der zum DFA passende reguläre Ausdruck.

Beispiel



Durch Hinschauen?
z.B. $aa(baa)^*a$

Durch Kleene's Konstruktion (angedeutet und mit Hinschauen)
Wir brauchen $\alpha_{1,2}^5$:

$$\begin{aligned}
\alpha_{1,2}^5 &= \alpha_{1,2}^4|\alpha_{1,5}^4(\alpha_{5,5}^4)^*\alpha_{5,2}^4 \\
&= \emptyset|\alpha_{1,5}^4(\alpha_{5,5}^4)^*\alpha_{5,2}^4 \\
&= \alpha_{1,5}^4(\alpha_{5,5}^4)^*\alpha_{5,2}^4 \\
&= a(\varepsilon|aba)^*aa \\
&= a(aba)^*aa \text{ (Vereinfachung)} \\
\alpha_{1,2}^4 &= \emptyset \text{ (durch Hinschauen, jeder Weg von } z_1 \text{ zu } z_2 \text{ muss } z_5 \text{ benutzen)} \\
\alpha_{1,5}^4 &= a \text{ (durch Hinschauen, Schleife läuft über } z_5\text{: geht nicht)} \\
\alpha_{5,5}^4 &= \varepsilon|aba \text{ (weitere Schleifen laufen durch } z_5\text{: geht nicht)} \\
\alpha_{5,2}^4 &= aa \text{ (durch Hinschauen, Schleife läuft über } z_5\text{: geht nicht)}
\end{aligned}$$

Anwendung des Pumping-Lemmas

- Pumping-Lemma:
  Sprache regulär $\Longrightarrow$ Sprache erfüllt die Pumping-Eigenschaft
- Zeige, dass eine Sprache nicht regulär ist, durch Kontraposition:

Sprache erfüllt **nicht** die Pumping-Eigenschaft
$\Longrightarrow$ Sprache ist **nicht** regulär

Erfüllt / erfüllt nicht die Pumping-Eigenschaft

Formale Sprache L erfüllt die Pumping-Eigenschaft

Es gibt eine Zahl $n \in \mathbb{N}_{>0}$, sodass jedes Wort $z \in L$, welches Mindestlänge n hat (d.h. $|z| \geq n$), als $z = uvw$ geschrieben werden kann, so dass gilt:

- $|uv| \leq n$
- $|v| \geq 1$
- für alle $i \geq 0$: $uv^iw \in L$.

Formale Sprache L erfüllt nicht die Pumping-Eigenschaft:

Für jede Zahl $n \in \mathbb{N}_{>0}$ gibt es ein Wort $z \in L$ mit $|z| \geq n$, sodass für jede Zerlegung $z = uvw$ mit

- $|uv| \leq n$ und
- $|v| \geq 1$

ein $i \geq 0$ existiert mit $uv^iw \notin L$.

Anwendung des Pumping-Lemmas

- Pumping-Lemma:
  Sprache regulär $\implies$ Sprache erfüllt die Pumping-Eigenschaft
- Zeige, dass eine Sprache nicht regulär ist, durch Kontraposition:

Sprache erfüllt **nicht** die Pumping-Eigenschaft
$\implies$ Sprache ist **nicht** regulär

Formale Sprache L erfüllt nicht die Pumping-Eigenschaft:

Für jede Zahl $n \in \mathbb{N}_{>0}$ gibt es ein Wort $z \in L$ mit $|z| \geq n$, sodass für jede Zerlegung $z = uvw$ mit

- $|uv| \leq n$ und
- $|v| \geq 1$

ein $i \geq 0$ existiert mit $uv^iw \notin L$.

Beispiel

Aufgabe

Zeigen Sie: $L = \{a^j b^j \mid j \in \mathbb{N}\}$ ist nicht regulär.

Korrekter Beweis:

1. Die Zahl $n \in \mathbb{N}_{>0}$ sei beliebig gewählt („vom Gegner")
   Sei $n \in \mathbb{N}_{>0}$ vom Gegner gewählt
2. Wähle ein Wort $z \in L$ mit $|z| \geq n$ (wir suchen aus)
   Wir wählen $z = a^n b^n \in L$.
3. Sei $z = uvw$ beliebige Zerlegung mit $|uv| \leq n, |v| \geq 1$ („vom Gegner")
   Sei z zerlegt in $z = uvw$ mit $|uv| \leq n$ und $|v| \geq 1$
4. Für jede solche Zerlegung gebe ein $i \in \mathbb{N}$ an mit $uv^iw \notin L$ (wir suchen aus)
   Dann ist $u = a^j$ $v = a^k$ und $w = a^l b^n$ mit $k > 0$, $j + k + l = n$ und damit für $i = 0$: $uv^iw = a^{n-1}b^n \notin L$

Beispiel: Finde den Fehler

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

Behauptung

L ist nicht regulär.

~~Beweis~~ mit dem Pumping-Lemma:

- Sei $n \in \mathbb{N}_{>0}$ beliebig.
- Wir wählen $z = a^n a^n \in L$.
- Sei $z = uvw$ mit $|uv| \leq n$ und $v \leq 1$ eine Zerlegung von z.
- Dann ist $u = a^d$, $v = a^e$, und $w = a^{n+n-d-e}$ und $e > 1$, $d + e \leq n$. Dann ist $\cancel{uv^0w = a^{n-e}a^n \notin L}$.
  ~~D.h. das Pumping-Lemma zeigt, dass L nicht regulär ist.~~

Beachte: L ist regulär, z.B. wird L durch den regulären Ausdruck $(aa)^*$ erzeugt!!!

Beispiel: L erfüllt die Pumping-Eigenschaft

Satz

$L = \{a^n a^n \mid n \in \mathbb{N}\}$ erfüllt die Pumping-Eigenschaft

Beweis:

- Sei $n = 2$.
- Sei $z \in L$ mit $|z| \geq n$.
- Wir zerlegen $z = uvw$ mit $u = \varepsilon$, $v = z[1]z[2]$, w der Suffix von z ohne die ersten beiden Buchstaben.
- Da $z \in L$, ist $z = a^j a^j$ und dann gilt: $v = aa$, $w = a^{j-1}a^{j-1}$. Daher gilt auch: $uv^iw = a^i a^i a^{j-1} a^{j-1} = a^{i+j-1}a^{i+j-1} \in L$ für alle $i \in \mathbb{N}$.

Aufgabe

Zeigen Sie mit dem Pumping-Lemma, dass die folgenden Sprachen nicht regulär sind:

- $L_1 = \{w \in \{a,b\}^* \mid \#_a(w) = \#_b(w)\}$
- $L_2 = \{a^n bbc^m \mid n,m \in \mathbb{N}, m > n\}$
- $L_3 = \{a^n \$ a^n \mid n \in \mathbb{N}\}$
- $L_4 = \{w \in \{a,b\}^* \mid |w| = n^2 \text{ mit } n \in \mathbb{N}\}$

Satz

$L_1 = \{w \in \{a,b\}^* \mid \#_a(w) = \#_b(w)\}$ ist nicht regulär.

Beweis:

1. Die Zahl $n \in \mathbb{N}_{>0}$ sei beliebig gewählt („vom Gegner“).

   Sei $n \in \mathbb{N}_{>0}$ beliebig.
2. Wähle ein Wort $z \in L_1$ mit $|z| \geq n$ (wir suchen aus)

   Sei $z = a^n b^n$ (dann gilt $z \in L_1$ und $|z| \geq n$)
3. Sei $z = uvw$ beliebige Zerlegung mit $|uv| \leq n, |v| \geq 1$ („vom Gegner“)

   Sei $z = uvw$ mit $|uv| \leq n, |v| \geq 1$ eine Zerlegung von z.
4. Für jede solche Zerlegung gebe ein $i \in \mathbb{N}$ an mit $uv^iw \notin L_1$ (wir suchen aus)

   Da $|uv| \leq n$ und $|v| > 0$ gilt $v = a^k$ mit $k > 0$ und daher $uv^0w = a^{n-k}b^n \notin L_1$ (d.h. für $i = 0$ gilt $uv^iw \notin L_1$)

Satz

$L_2 = \{a^n bbc^m \mid n,m \in \mathbb{N}, m > n\}$ ist nicht regulär.

Beweis:

1. Die Zahl $n \in \mathbb{N}_{>0}$ sei beliebig gewählt („vom Gegner“).

   Sei $n \in \mathbb{N}_{>0}$ beliebig.
2. Wähle ein Wort $z \in L_2$ mit $|z| \geq n$ (wir suchen aus)

   Sei $z = a^n bbc^{n+1}$ (dann gilt $z \in L_2$ und $|z| \geq n$)
3. Sei $z = uvw$ beliebige Zerlegung mit $|uv| \leq n, |v| \geq 1$ („vom Gegner“)

   Sei $z = uvw$ mit $|uv| \leq n, |v| \geq 1$ eine Zerlegung von z.
4. Für jede solche Zerlegung gebe ein $i \in \mathbb{N}$ an mit $uv^iw \notin L_2$ (wir suchen aus)

   Da $|uv| \leq n$ und $|v| > 0$ gilt $v = a^k$ mit $k > 0$ und daher $uv^3w = a^{n+2k}bbc^{n+1} \notin L_2$ (d.h. für $i = 3$ gilt $uv^iw \notin L_2$)

Satz

$L_3 = \{a^n \$ a^n \mid n \in \mathbb{N}\}$ ist nicht regulär.

Beweis:

1. Die Zahl $n \in \mathbb{N}_{>0}$ sei beliebig gewählt („vom Gegner“).

   Sei $n \in \mathbb{N}_{>0}$ beliebig.
2. Wähle ein Wort $z \in L_3$ mit $|z| \geq n$ (wir suchen aus)

   Sei $z = a^n \$ a^n$ (dann gilt $z \in L_3$ und $|z| \geq n$)
3. Sei $z = uvw$ beliebige Zerlegung mit $|uv| \leq n, |v| \geq 1$ („vom Gegner“)

   Sei $z = uvw$ mit $|uv| \leq n, |v| \geq 1$ eine Zerlegung von z.
4. Für jede solche Zerlegung gebe ein $i \in \mathbb{N}$ an mit $uv^iw \notin L_3$ (wir suchen aus)

   Da $|uv| \leq n$ und $|v| > 0$ gilt $v = a^k$ mit $k > 0$ und daher $uv^0w = a^{n-k}\$a^n \notin L_3$ (d.h. für $i = 0$ gilt $uv^iw \notin L_3$)

Satz

$L_4 = \{w \in \{a,b\}^* \mid |w| = n^2 \text{ mit } n \in \mathbb{N}\}$ ist nicht regulär.

Beweis:

1. Die Zahl $n \in \mathbb{N}_{>0}$ sei beliebig gewählt („vom Gegner“).

   Sei $n \in \mathbb{N}_{>0}$ beliebig.

2. Wähle ein Wort $z \in L_4$ mit $|z| \geq n$ (wir suchen aus)

   Sei $z = a^{n^2}$ (dann gilt $z \in L_4$ und $|z| \geq n$)

3. Sei $z = uvw$ beliebige Zerlegung mit $|uv| \leq n, |v| \geq 1$ („vom Gegner“)

   Sei $z = uvw$ mit $|uv| \leq n, |v| \geq 1$ eine Zerlegung von z.

4. Für jede solche Zerlegung gebe ein $i \in \mathbb{N}$ an mit $uv^iw \notin L_4$ (wir suchen aus)

   Dann ist $u = a^l, v = a^k$ und $w = a^{n^2-k-l}$, wobei $k + l \leq n$ und $k > 0$.
   Daher ist $uv^0w = a^{l+n^2-k-l} = a^{n^2-k}$, wobei $|a^{n^2-k}| = n^2 - k$. $a^{n^2-k} \notin L_4$,
   da $n^2 - k$ keine Quadratzahl sein kann:
   - *$n^2 - k < n^2$ (da $k > 0$)*
   - *$n^2 - k \geq n^2 - n$ (da $k + l \leq n$) und $n^2 - n = n(n-1) > (n-1)^2$*
     daher $n^2 - k > (n-1)^2$

   D.h. für $i = 0$ gilt $uv^iw \notin L_4$