Theoretische Informatik 1

WiSe 23/24

Prof. Dr. Sebastian Siebertz
AG Theoretische Informatik
MZH, Raum 3160
siebertz@uni-bremen.de

Universität Bremen

Diese VL: Nachweis der Nichtregularität einer Sprache

- Nachweis der Regularität
  - Konstruiere einen endlichen Automaten DEA/NEA für die Sprache.
- Nachweis der *Nicht*regularität
  - *Beweise*, dass es *keinen* DEA/NEA für die Sprache geben kann.

    $\rightarrow$ viel schwieriger.
- Idee:
  - Verwende, dass jeder DEA/NEA nur endlich viele Zustände hat.
  - Finde darauf basierende Eigenschaft, die jede reguläre Sprache erfüllt.
  - Weise dann nach, dass die betreffende Sprache die Eigenschaft verletzt.

Beispiel

Beispiel

$L = \{a^n b^n \mid n \geq 0\}$ ist *nicht* regulär.

Beweis.

- Angenommen es gibt NEA $\mathcal{A} = (Q, \Sigma, q_s, \Delta, F)$ mit $L(\mathcal{A}) = L$.
- Sei $n_0 := |Q|$ und betrachte das Wort $w = a^{n_0} b^{n_0} \in L$.
- Nach Annahme existiert ein Lauf

$$p_0 \xrightarrow{a}_{\mathcal{A}} p_1 \xrightarrow{a}_{\mathcal{A}} \cdots \xrightarrow{a}_{\mathcal{A}} p_{n_0} \xrightarrow{b}_{\mathcal{A}} \cdots \xrightarrow{b}_{\mathcal{A}} p_{2n_0}$$

mit $p_0 = q_s$ und $p_{2n_0} \in F$.

Beispiel fortgesetzt

$$p_0 \xrightarrow{a}_{\mathcal{A}} p_1 \xrightarrow{a}_{\mathcal{A}} \cdots \xrightarrow{a}_{\mathcal{A}} p_{n_0} \xrightarrow{b}_{\mathcal{A}} \cdots \xrightarrow{b}_{\mathcal{A}} p_{2n_0}$$

- Wegen $n_0 = |Q|$ können die $n_0 + 1$ Zustände $p_0, \ldots, p_{n_0}$ nicht alle verschieden sein.
- Es gibt also i, j mit $0 \leq i < j \leq n_0$ und $p_i = p_j$ und wir können Lauf auch schreiben als
  $$p_0 \xRightarrow{x}_{\mathcal{A}} p_i \xRightarrow{y}_{\mathcal{A}} p_i = p_j \xRightarrow{z}_{\mathcal{A}} p_{2n_0},$$
  für $x, y, z \in \Sigma^*$ mit $w = xyz$, $|xy| \leq n_0$ und $|y| \neq 0$.
- Dann ist auch
  $$p_0 \xRightarrow{x}_{\mathcal{A}} p_i \xRightarrow{y}_{\mathcal{A}} p_i \xRightarrow{y}_{\mathcal{A}} p_i \xRightarrow{z}_{\mathcal{A}} p_{2n_0}$$
  auf dem Wort $xyyz$ ein gültiger Lauf.

Beispiel fortgesetzt

$$p_0 \xRightarrow{x}_{\mathcal{A}} p_i \xRightarrow{y}_{\mathcal{A}} p_i \xRightarrow{y}_{\mathcal{A}} p_i \xRightarrow{z}_{\mathcal{A}} p_{2n_0}$$

- Sei $k_1 = |x|$, $k_2 = |y| > 0$, $k_3 = |z| - n_0$.
- Also enthält $L(\mathcal{A})$ das Wort $xyyz$, welches die Form
$$a^{k_1}\, a^{2k_2}\, a^{k_3}\, b^{n_0}$$
mit $k_1 + 2k_2 + k_3 > n_0$ hat.
- Also hat $xyyz$ mehr a's als b's und kann damit nicht zu L gehören, obwohl es von $\mathcal{A}$ akzeptiert wird.
- Dies ist ein Widerspruch;
also war die Annahme falsch, dass L regulär ist!

Das Pumping Lemma für reguläre Sprachen

- Wenn L eine reguläre Sprache ist, dann gilt:

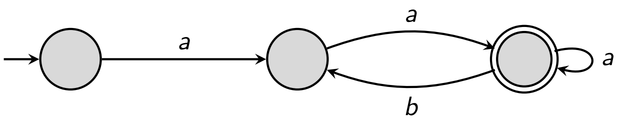
  es existiert $n_0 \geq 1$, so dass

  für alle $w \in L$ mit $|w| \geq n_0$ gilt:

  es existiert eine Zerlegung $w = xyz$ mit $y \neq \varepsilon$ und $|xy| \leq n_0$, so dass

  für alle $k \geq 0$ gilt $xy^kz \in L$.

- Beispiel



  - Wähle $n_0 = 3$.
  - Z.B. das Wort $aabaa \in L$ lässt sich zerlegen als $x = a, y = ab, z = aa$.
  - Dann $xy^0z = aaa \in L$, $xy^1z = aabaa \in L$, $xy^2z = aababaa \in L$, ...

Beweis Pumping Lemma

- Sei $\mathcal{A} = (Q, \Sigma, q_s, \Delta, F)$ ein NEA mit $L(\mathcal{A}) = L$.
- Wir wählen $n_0 := |Q|$.
- Sei nun $w = a_1 \cdots a_m \in L$ ein Wort mit $m \geq n_0$.
- Dann existiert ein Lauf

$$p_0 \xrightarrow{a_1}_{\mathcal{A}} p_1 \xrightarrow{a_2}_{\mathcal{A}} \cdots \xrightarrow{a_{n_0}}_{\mathcal{A}} p_{n_0} \xrightarrow{a_{n_0+1}}_{\mathcal{A}} \cdots \xrightarrow{a_m}_{\mathcal{A}} p_m$$

  mit $p_0 = q_s$ und $p_m \in F$.
- Wegen $n_0 = |Q|$ können die $n_0 + 1$ Zustände $p_0, \ldots, p_{n_0}$ nicht alle verschieden sein.

Beweis Pumping Lemma

$$p_0 \xrightarrow{a_1}_{\mathcal{A}} p_1 \xrightarrow{a_2}_{\mathcal{A}} \cdots \xrightarrow{a_{n_0}}_{\mathcal{A}} p_{n_0} \xrightarrow{a_{n_0+1}}_{\mathcal{A}} \cdots \xrightarrow{a_m}_{\mathcal{A}} p_m$$

- Es gibt also i, j mit $0 \leq i < j \leq n_0$ und $p_i = p_j$.
- Wir wählen

$$x := a_1 \cdots a_i, \quad y := a_{i+1} \cdots a_j, \quad z := a_{j+1} \cdots a_m.$$

- Es gilt $y \neq \varepsilon$ (da $i < j$)
- und $|xy| \leq n_0$ (da $j \leq n_0$)
- sowie

$$q_s = p_0 \xRightarrow{x}_{\mathcal{A}} p_i \xRightarrow{y}_{\mathcal{A}} p_i = p_j \xRightarrow{z}_{\mathcal{A}} p_m \in F.$$

Beweis Pumping Lemma

$$q_s = p_0 \stackrel{x}{\Longrightarrow}_{\mathcal{A}} p_i \stackrel{y}{\Longrightarrow}_{\mathcal{A}} p_i = p_j \stackrel{z}{\Longrightarrow}_{\mathcal{A}} p_m \in F.$$

- Folglich gilt für alle $k \geq 0$ auch $p_i \stackrel{y^k}{\Longrightarrow}_{\mathcal{A}} p_i$,also

$$q_0 = p_0 \stackrel{x}{\Longrightarrow}_{\mathcal{A}} p_i \stackrel{y^k}{\Longrightarrow}_{\mathcal{A}} p_i = p_j \stackrel{z}{\Longrightarrow}_{\mathcal{A}} p_m \in F,$$

was $xy^k z \in L$ zeigt.

Zurück zum Beispiel

Beispiel

$L = \{a^n b^n \mid n \geq 0\}$ ist *nicht* regulär.

Beweis mit Hilfe des Pumping Lemmas.

- Angenommen L ist regulär.
- Dann **existiert** $n_0 \geq 1$, so dass
  für alle $w \in L$ mit $|w| \geq n_0$ gilt:
  es existiert eine Zerl. $w = xyz$ mit $y \neq \varepsilon$ und $|xy| \leq n_0$, so dass
  für alle $k \geq 0$ gilt $xy^k z \in L$.
- Fixiere ein solches n_0.

Beispiel fortgesetzt

- Betrachte das Wort $w = a^{n_0} b^{n_0} \in L$.
- Da $|w| \geq n_0$, gibt es eine Zerlegung $a^{n_0} b^{n_0} = xyz$ mit $|y| \geq 1$ und $|xy| \leq n_0$, so dass $xy^k z \in L$ für alle $k \geq 0$.
- Da $|xy| \leq n_0$ muss y ganz in a^{n_0} liegen.
- Also ist
$$x = a^{k_1}, \quad y = a^{k_2}, \quad z = a^{k_3} b^{n_0}$$
mit $k_2 > 0$ und $n_0 = k_1 + k_2 + k_3$.
- Damit ist aber
$$xy^0 z = xz = a^{k_1 + k_3} b^{n_0} \notin L,$$
da $k_1 + k_3 < n_0$.
- Widerspruch. Deshalb muss die Annahme „L ist regulär“ falsch sein.

Zurück zum Pumping Lemma

- Pumping-Lemma für reguläre Sprachen:

  Wenn L eine reguläre Sprache ist, dann gilt:

  es existiert $n_0 \geq 1$, so dass

  für alle $w \in L$ mit $|w| \geq n_0$ gilt:

  es existiert eine Zerlegung $w = xyz$ mit $y \neq \varepsilon$ und $|xy| \leq n_0$, so dass

  für alle $k \geq 0$ gilt $xy^k z \in L$.

- Pumping-Lemma als Kontraposition:

  Wenn X, dann Y $\Leftrightarrow$ wenn nicht Y, dann nicht X.

  Sei L eine Sprache. Wenn

  für alle $n_0 \geq 1$

  existiert ein $w \in L$ mit $|w| \geq n_0$, so dass

  für alle Zerlegungen $w = xyz$ mit $y \neq \varepsilon$ und $|xy| \leq n_0$ gilt:

  es existiert $k \geq 0$ mit $xy^k z \notin L$.

  Dann ist L **nicht** regulär.

Spieltheoretische Sicht auf das Pumping Lemma

- Pumping-Lemma als Kontraposition:

  Sei L eine Sprache. Wenn

  für alle $n_0 \geq 1$

  existiert ein $w \in L$ mit $|w| \geq n_0$, so dass

  für alle Zerlegungen $w = xyz$ mit $y \neq \varepsilon$ und $|xy| \leq n_0$ gilt:

  es existiert $k \geq 0$ mit $xy^k z \notin L$.

  Dann ist L **nicht** regulär.

  - Wir spielen in Runden gegen eine Gegnerin.
  - Wir möchten zeigen, dass L nicht regulär ist, die Gegnerin möchte zeigen, dass L regulär ist.
  - In den Zeilen, die mit „es existiert“ beginnen, sind wir am Zug.
  - In den Zeilen, die mit „für alle“ beginnen, ist die Gegnerin am Zug.
  - Wenn wir das Spiel stets gewinnen können – und zwar unabhängig davon, was die Gegnerin tut – dann ist L nicht regulär.

Beispiel

Beispiel

$L = \{a^n \mid n \text{ ist Quadratzahl}\}$ ist nicht regulär.

- Sei $n_0 \geq 1$ von der Gegnerin gewählt.
- Wir wählen $w = a^m$ mit $m = (n_0 + 1)^2$.
  (es gilt $w \in L$ und $|w| > n_0$)
- Sei $xyz = w$ eine Zerlegung von der Gegnerin gewählt mit den Eigenschaften
  $$y \neq \varepsilon \quad \text{und} \quad |xy| \leq n_0.$$
- Wir müssen nun ein $k \geq 0$ finden, so dass $xy^k z \notin L$.
  - Beobachtung: Eine Zahl der Form $s^2 + t$ mit $0 < t < s$ ist keine Quadratzahl, denn
    $$s^2 \quad < \quad s^2 + t \quad < \quad s^2 + 2s + 1 \quad = \quad (s+1)^2.$$

Beispiel fortgesetzt

- Wir suchen $k \geq 0$ so dass $|xy^k z| = s^2 + t$ für ein s und $0 < t < s$.
- Betrachte $k = 4 \cdot |y| \cdot |w|^2$. Es gilt:

$$\begin{aligned} |xy^k z| &= |y| \cdot k + |x| + |z| \\ &= 4 \cdot |y|^2 \cdot |w|^2 + |x| + |z| \\ &= (2 \cdot |y| \cdot |w|)^2 + |x| + |z|. \end{aligned}$$

- Mit $s = 2 \cdot |y| \cdot |w|$ und $t = |x| + |z|$ gilt nun wie gewünscht $|xy^k z| = s^2 + t$ und $0 < t < s$ ($|z| > 0$, da $|w| > n_0$ und $|xy| \leq n_0$).
- Also ist $xy^k z \notin L$, was zu zeigen war (d. h. wir gewinnen das Spiel und haben bewiesen, dass L nicht regulär ist).

Noch ein Beispiel

Beispiel

$L = \{a^n b^m \mid n \neq m\}$ ist *nicht* regulär.

- Sei $n_0 \geq 1$ von der Gegnerin gewählt.
- Wir wählen $w = a^{n_0} b^{n_0! + n_0}$,
  wobei $n_0! = 1 \cdot 2 \cdot \ldots \cdot (n_0 - 1) \cdot n_0$ (Fakultätsoperation).
- Sei $xyz = w$ eine Zerlegung von der Gegnerin gewählt mit den Eigenschaften

$$y \neq \varepsilon \quad \text{und} \quad |xy| \leq n_0.$$

- Wir müssen nun ein $k \geq 0$ finden, so dass $xy^k z \notin L$.

Noch ein Beispiel fortgesetzt

- Erinnerung: $w = a^{n_0} b^{n_0! + n_0}$
- Da $|xy| \leq n_0$ enthält y nur a's, also

$$x = a^{k_1}, \quad y = a^{k_2}, \quad z = a^{k_3} b^{n_0! + n_0}$$

  für geeignete k_1, k_2, k_3 mit $k_1 + k_2 + k_3 = n_0$ und $k_2 > 0$ (da $y \neq \varepsilon$).
- Da $0 < k_2 \leq n_0$ existiert eine Zahl ℓ mit $\ell \cdot k_2 = n_0!$
- Wir wählen $k = \ell + 1$, dann ist die Anzahl a's im Wort $xy^{\ell+1}z$:

$$\begin{aligned} k_1 + (\ell + 1)k_2 + k_3 &= \underbrace{k_1 + k_2 + k_3}_{=n_0} + \underbrace{\ell \cdot k_2}_{=n_0!} \\ &= n_0 + n_0! \end{aligned}$$

- Also ist $xy^{\ell+1}z = a^{n_0 + n_0!} b^{n_0 + n_0!} \notin L$, was zu zeigen war (d. h. wir gewinnen das Spiel und haben bewiesen, dass L nicht regulär ist).

Grenzen des Pumping-Lemmas

- Das Pumping-Lemma ist nützlich, um Nichterregularität nachzuweisen.
- Es gibt aber Sprachen, die *nicht regulär* sind, aber *die Eigenschaften des Lemmas trotzdem erfüllen*.
- Die formulierte Eigenschaft ist also *notwendig* für Regularität, aber *nicht hinreichend*.

Beispiel

- Beispiel

  $L = \{a^m b^n c^n \mid m, n \geq 1\} \cup \{b^m c^n \mid m, n \geq 0\}$ ist *nicht regulär*, erfüllt aber die Eigenschaften des Pumping Lemmas.

- Wir zeigen, dass

  es existiert $n_0 \geq 1$, so dass

  für alle $w \in L$ mit $|w| \geq n_0$ gilt:

  es existiert eine Zerlegung $w = xyz$ mit $y \neq \varepsilon$ und $|xy| \leq n_0$, so dass

  für alle $k \geq 0$ gilt $xy^k z \in L$.

Beispiel fortgesetzt

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

- Wähle $n_0 = 3$.
- Sei $w \in L$ mit $|w| \geq 3$ beliebig.
- Es tritt einer der folgenden drei Fälle ein.
  - $w = a^m b^n c^n$ mit $m, n \geq 1$ (w ist aus dem „1. Teil" von L)
  - $w = b^m c^n$ mit $m \geq 1$ und $n \geq 0$ (w ist aus dem „2. Teil" von L und beginnt mit b)
  - $w = c^n$ mit $n \geq 1$ (w ist aus dem „3. Teil" von L und beginnt mit c, da $|w| \geq 3$)
- Wir zeigen: in jedem dieser drei Fälle lässt sich w zerlegen in $w = xyz$ mit $y \neq \varepsilon$ und $|xy| \leq n_0$ sowie $xy^k z \in L$ für alle $k \geq 0$.

Beispiel fortgesetzt

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

- 1. Fall: Wenn $w = a^m b^n c^n$ mit $m, n \geq 1$,dann betrachte die Zerlegung

$$x = \varepsilon, \quad y = a, \quad z = a^{m-1} b^n c^n.$$

Dann ist für *jedes* $k \geq 0$ das Wort

$$xy^k z = a^{k+(m-1)} b^n c^n$$

in L, denn die Anzahlen der a's und b's sind im „1. Teil“ unabhängig (und falls $(m-1) + k = 0$, ist $xy^k z$ im „2. Teil“ von L).

Beispiel fortgesetzt

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

- 2. Fall: Wenn $w = b^m c^n$ mit $m \geq 1$ und $n \geq 0$,dann betrachte die Zerlegung

$$x = \varepsilon, \quad y = b, \quad z = b^{m-1} c^n.$$

Dann ist für *jedes* $k \geq 0$ das Wort

$$xy^k z = b^{k+(m-1)} c^n$$

in L, denn die Anzahlen der b's und c's sind im „2. Teil" unabhängig.

Beispiel fortgesetzt

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

- 3. Fall: Wenn $w = c^n$ mit $n \geq 1$,dann betrachte die Zerlegung

  $$x = \varepsilon, \quad y = c, \quad z = c^{n-1}.$$

  Dann ist für *jedes* $k \geq 0$ das Wort

  $$xy^k z = c^{k+(n-1)}$$

  in L (wiederum im „2. Teil").
- Also gilt in jedem Fall $xy^k z \in L$ und wir können mit Hilfe des Pumping Lemmas nicht nachweisen, dass L nicht regulär ist.

Nachweis der Nichtregularität über Abschlusseigenschaften

- Manchmal können wir auch einfacher über die Abschlusseigenschaften von regulären Sprachen beweisen, dass eine Sprache nicht regulär ist.
- Beispiel:

  $L := \{a^n b^m \mid n \neq m\}$ ist *nicht* regulär.
- Statt das Pumping-Lemma zu benutzen, können wir auch verwenden, dass bereits bekannt ist, dass $L' = \{a^n b^n \mid n \geq 0\}$ *nicht* regulär ist.
- Wäre nämlich L regulär, so auch

$$L' = \overline{L} \cap (\{a\}^* \cdot \{b\}^*).$$

- Da wir schon wissen, dass L' nicht regulär ist, kann auch L nicht regulär sein.

Zusammenfassung

- Das Pumping-Lemma ist ein wichtiges Werkzeug zum Nachweis der Nichtregularität.
- Es kann *nicht* verwendet werden um zu zeigen, dass eine Sprache *regulär* ist!
- Bequem ist die Betrachtungsweise als Kontraposition und Spiel.
- Die Methode ist jedoch kein Automatismus, *erfordert Kreativität*.
- Es gibt stärkere Versionen des Pumping-Lemmas, die die regulären Sprachen charakterisieren (diese behandeln wir aber nicht in der Vorlesung).
- Wichtige Intuition: Sprachen, deren Erkennen *unbeschränktes Zählen* erfordert, sind (in der Regel) *nicht regulär*.