# Einführung in die Computerlinguistik
# – Pumpinglemma für reguläre Sprachen

Dozentin: Wiebke Petersen

6. Foliensatz

# Vorbemerkungen

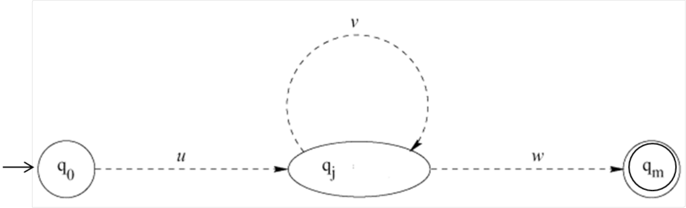
- Ein Lemma ist ein Hilfssatz.
- Lemmata werden häufig immer wieder in ähnlichen Beweisen angewandt.
- Das Pumpinglemma ermöglicht es zu zeigen, dass eine Sprache nicht regulär ist.

# Pumping-lemma für reguläre Sprachen

**Lemma (Pumping-Lemma)**

*Sei $L$ eine unendliche reguläre Sprache, dann gilt für jedes genügend lange Wort $z \in L$, daß es so in Teilworte $z = uvw$ $(u, w \in \Sigma^*,\ v \in \Sigma^+)$ zerlegt werden kann, dass jedes der Worte $uv^iw \in L$ $(i \geq 0)$ ein Wort der Sprache $L$ ist.*

**Beweisidee:**




Wenn ein Wort länger ist, als der Automat Zustände hat, dann muß bei der Verarbeitung des Wortes ein Zustand zweimal besucht werden. Es gibt somit eine Schleife, die beliebig oft durchlaufen werden kann; das Wort kann in dem Schleifenbereich aufgepumpt werden.

# Über die Aussagekraft des Pumpinglemmas

- Das Pumpinglemma sagt, daß wenn eine Sprache unendlich und regulär ist, dann muß sie auch pumpbar sein.
  (reg $\rightarrow$ pump)
- Vorsicht: Aus der Pumpbarkeit einer Sprache kann nicht auf ihre Regularität geschlossen werden. (reg $\rightarrow$ pump $\not\Leftrightarrow$ pump $\rightarrow$ reg)
- Aber aus der Nichtpumpbarkeit einer Sprache kann geschlossen werden, dass sie nicht regulär ist. (reg $\rightarrow$ pump $\Leftrightarrow$ $\neg$reg $\vee$ pump $\Leftrightarrow$ pump $\vee$ $\neg$reg $\Leftrightarrow$ $\neg$pump $\rightarrow$ $\neg$reg)

# $L = \{a^n b^n : n \geq 0\}$ ist nicht regulär

- $L = \{a^n b^n : n \geq 0\}$:
  $L$ ist unendlich. Wäre $L$ regulär, dann müßte es für genügend lange Worte die geforderte pumpbare Zerlegung geben: aber

# $L = \{a^n b^n \ : \ n \geq 0\}$ ist nicht regulär

- $L = \{a^n b^n \ : \ n \geq 0\}$:
  $L$ ist unendlich. Wäre $L$ regulär, dann müßte es für genügend lange Worte die geforderte pumpbare Zerlegung geben: aber
  1. das pumpbare Teilwort kann nicht nur aus $a$'s bestehen, sonst würden beim Pumpen zuviele $a$'s entstehen
     $(aa(aa)^2bbbb = aaaaaabbbb)$ .
  2. das pumpbare Teilwort kann nicht nur aus $b$'s bestehen, sonst würden beim Pumpen zuviele $b$'s entstehen.
     $(aaaab(bb)^2b = aaaabbbbbb)$
  3. das pumpbare Teilwort kann nicht aus $a$'s und $b$'s bestehen, da beim Pumpen die Sortierung der $a$'s und $b$'s verloren ginge.
     $(aaa(ab)^2bbb = aaaababbbb)$

# $L = \{a^n b a^n : n \geq 0\}$ ist nicht regulär

- $L = \{a^n b a^n : n \geq 0\}$:
  $L$ ist unendlich. Wäre $L$ regulär, dann müßte es für genügend lange Worte die geforderte pumpbare Zerlegung geben: aber

# $L = \{a^n b a^n \;:\; n \geq 0\}$ ist nicht regulär

- $L = \{a^n b a^n \;:\; n \geq 0\}$:
  $L$ ist unendlich. Wäre $L$ regulär, dann müßte es für genügend lange Worte die geforderte pumpbare Zerlegung geben: aber
  1. das pumpbare Teilwort kann nicht nur aus $a$'s bestehen, sonst würden beim Pumpen auf einer Seite des $b$'s zuviele $a$'s entstehen $(aa(aa)^2 baaaa = aaaaaabaaaa)$ .
  2. das pumpbare Teilwort darf nicht $b$ beinhalten, sonst würden durch das Pumpen Wörter mit mehr als einem $b$ entstehen. $(aaa(aba)^2 aaa = aaaabaabaaaa)$

# Die Sprache $L_{\text{pal}}$ der Palindrome über dem Alphabet {a,b} ist nicht regulär

- Wäre die Sprache $L_{\text{pal}}$ der Palindrome über dem Alpahbet $\{a, b\}$ regulär, dann müßte die Schnittmenge dieser Palindromsprache mit einer regulären Sprache ebenfalls eine reguläre Sprache sein (Satz von Kleene, vgl. Foliensatz 5).
- $L(a^*ba^*)$ ist eine reguläre Sprache.
- $L_{\text{pal}} \cap L(a^*ba^*) = L(a^nba^n)$.
- $L(a^nba^n)$ ist keine reguläre Sprache (siehe vorherige Folie).
- Die Sprache $L_{\text{pal}}$ der Palindrome über dem Alpahbet $\{a, b\}$ kann keine reguläre Sprache sein.