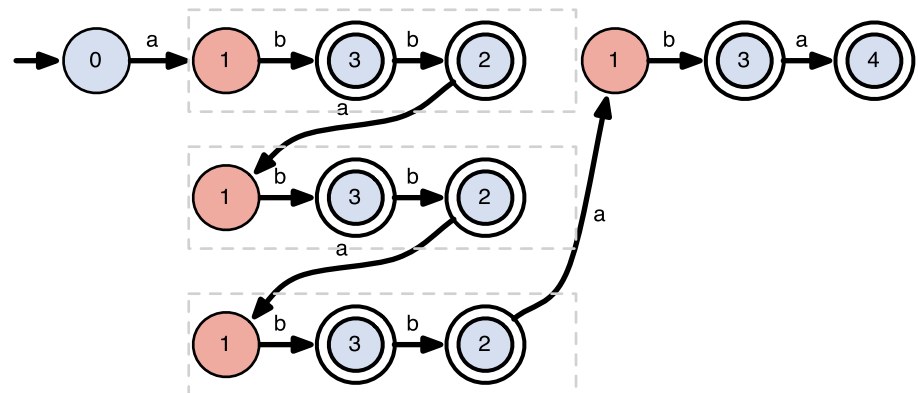


Berechenbarkeitstheorie

7. Vorlesung



Dr. Franziska Jahnke

Institut für Mathematische Logik und Grundlagenforschung

WWU Münster

Das Pumpinglemma

- Strukturelles Lemma über eine Eigenschaft von regulären Sprachen
- Hat eine Sprache diese Eigenschaft **nicht**, kann Sie nicht regulär sein (Anwendung: Beweis Nichtregularität)

Definition

Eine Sprache heißt **pumpbar** gdw., es gibt ein $k \in \mathbb{N}$, sodass für alle Wörter $w \in L$ mit $|w| \geq k$ es eine Zerlegung von $w = xyz$ gibt, für die gilt:

1. $|xy| \leq k$,
2. $\forall i \geq 0$: $xy^iz \in L$,
3. $|y| > 0$.

Pumpinglemma

$$L \in \text{REG} \Rightarrow L \text{ ist pumpbar}$$

Beispiel

$$L = \{a^k \mid k \text{ ungerade}\} \in \text{REG}$$

Zur Erinnerung: pumbar heißt

$$\exists k \in \mathbf{N} \forall w \in L: (|w| \geq k) \exists xyz = w \text{ mit } |xy| \leq k, |y| > 0, \text{ und } \forall i \geq 0: xy^i z \in L$$

In diesem Fall: ■ $k = 3$

■ $\forall w \in L$ zerteile w wie folgt

$$w = \underbrace{a}_{x} \underbrace{aaa}_{y} \underbrace{\dots a}_{z}$$

■ $|xy| \leq k$ und $|y| > 0$ sind erfüllt

■ $\forall i \geq 0$ gilt: w und $xy^i z$ unterscheiden sich um $2i$ as

→ $xy^i z$ besitzt eine ungerade Anzahl von as

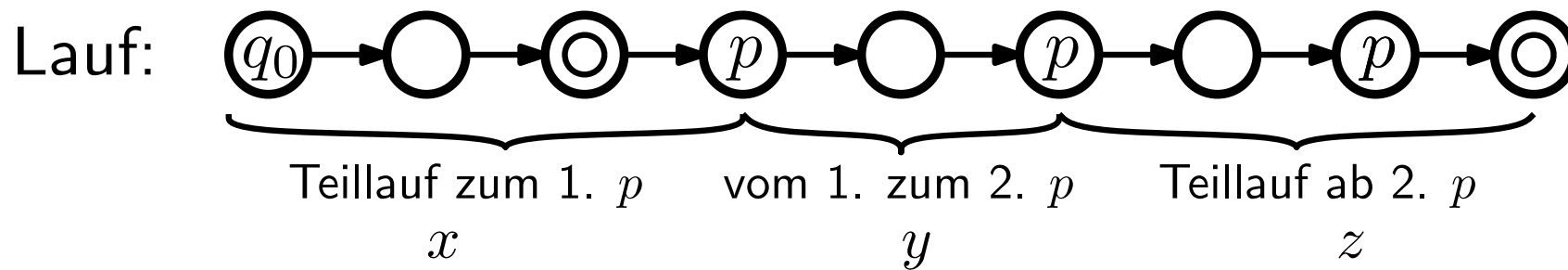
→ $xy^i z \in L$

■ L ist pumpbar

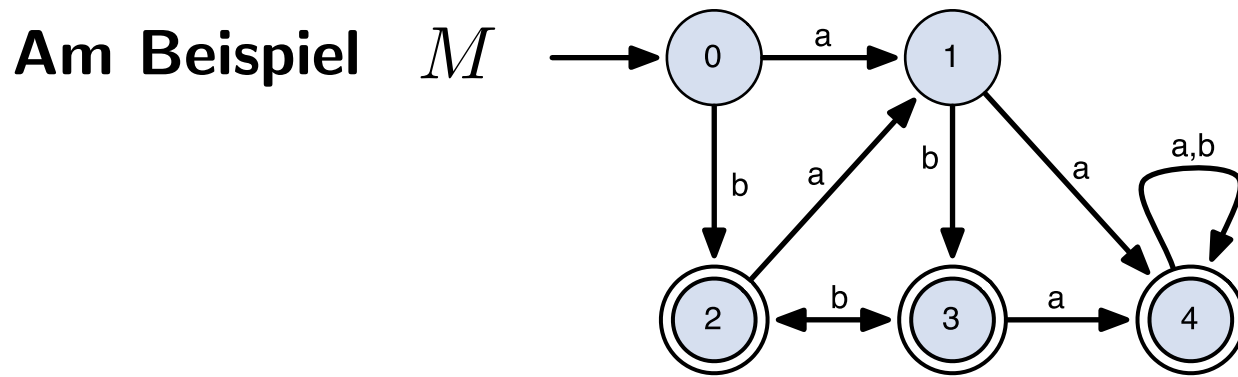
Beweis des Pumpinglemmas

Zu zeigen: L ist regulär, dann ist L pumpbar

- sei L regulär, und $M = (Q, \Sigma, \delta, q_0, F)$ ein DEA der L erkennt
- wir setzen $k = |Q| + 1$
- jeder w -Lauf ($|w| \geq k$) besucht mindestens $k = |Q| + 1$ Zustände
 $\rightarrow \geq 1$ Zustand doppelt
- sei p der 1. Zustand der mehrfach besucht wurde



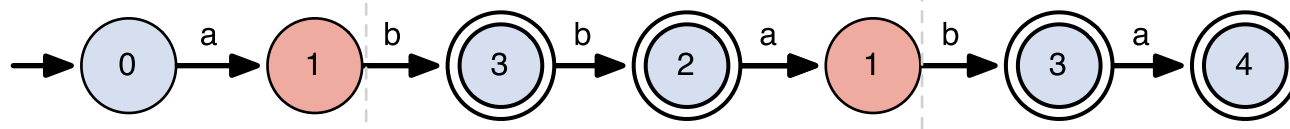
- Eigenschaft 1: $|y| > 0$
- Eigenschaft 2: $|xy| \leq k$
- Eigenschaft 3: $\forall i \geq 0: xy^i z \in L$



$$|Q| = 5$$

$$k = 5$$

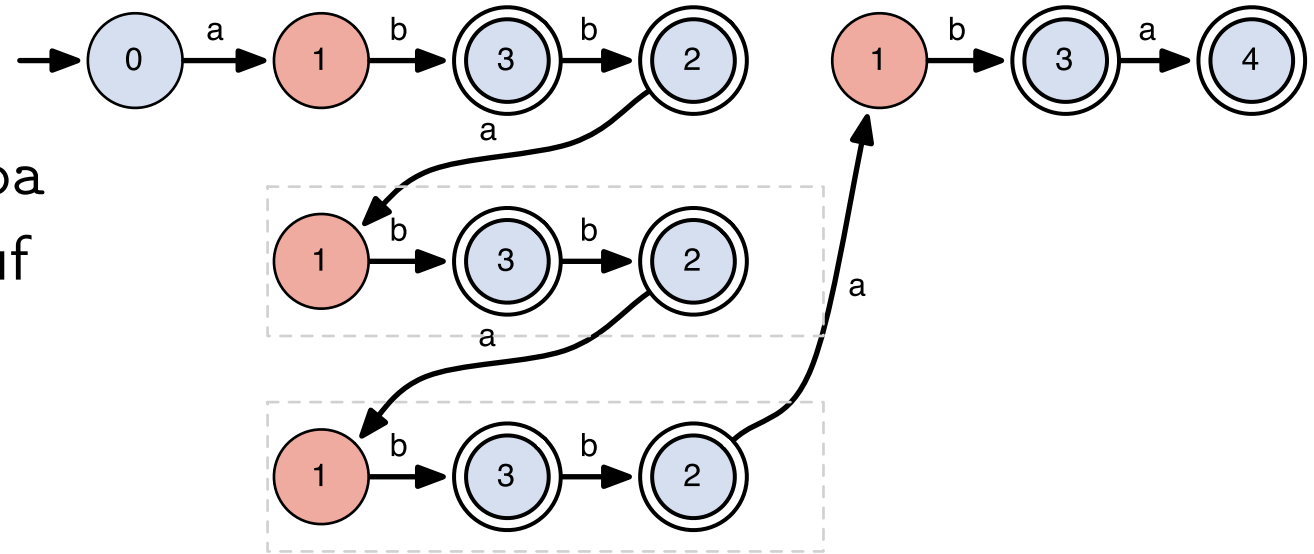
$w = \text{abbaba}$ hat folgenden Lauf



erster doppelter Zustand ist 1
 $x = a, y = \text{bba}, z = \text{ba}$

Wiederholung $i = 3$

$xyyyz = \text{abbabbaba}$
 hat akzeptierenden Lauf



Anwendung des Pumpinglemmas

6

- PL: $L \in \text{REG} \Rightarrow L$ ist pumpbar
- Umkehrung PL: L nicht pumpbar $\Rightarrow L \notin \text{REG}$
- **nicht pumpbar heißt:**
 - $\forall k \in \mathbb{N}$
 - $\exists w \in L: (|w| \geq k)$
 - $\forall xyz = w$ mit $|xy| \leq k, |y| > 0$ gilt
 - $\exists i \geq 0: xy^i z \notin L$
- Interpretation als 2-Spieler Spiel:
 - Gegenspieler wählt bei den \forall Quantoren.
 - Wir wählen bei den \exists Quantoren.
 - Wahl in Abhängigkeit der vorigen Auswahlen.
 - Haben wir eine Gewinnstrategie ist die Sprache nicht pumpbar und somit nicht regulär.

1. Beispiel $L = \{a^n b^n \mid n \geq 0\}$

Zu zeigen:

$$\forall k \in \mathbf{N} \exists w \in L: (|w| \geq k) \forall xyz = w \text{ mit } |xy| \leq k, |y| > 0 \text{ gilt} \\ \exists i \geq 0: xy^i z \notin L$$

- Gegenspieler bestimmt k
- **wir wählen** $w \in L$ in Abhängigkeit von k , in unserem Fall $w = a^k b^k$
- Gegenspieler wählt Zerteilung von w
- **wir zeigen:** für jede seiner Wahlen finden wir ein i , sodass $xy^i z \notin L$
- in unserem Fall wissen wir dass $y \in L(a^*)$ und $y \neq \varepsilon$
- d.h. für $i = 2$ ist $xy^i z = a^{k+|y|} b^k \notin L$
- also ist L nicht pumpbar und somit nicht regulär

2. Beispiel $L = \{uu \mid u \in \{a, b\}^*\}$

8

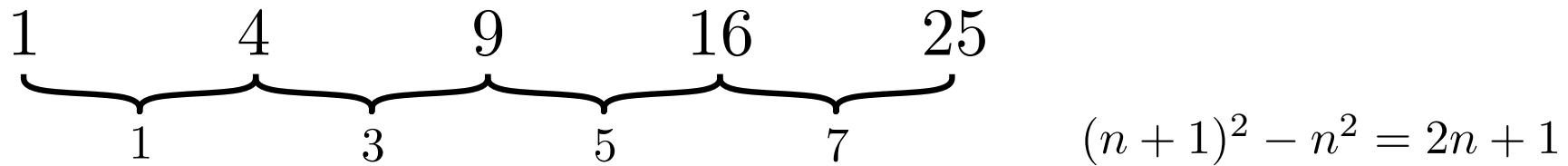
Zur Erinnerung

$\forall k \in \mathbf{N} \exists w \in L: (|w| \geq k) \forall xyz = w \text{ mit } |xy| \leq k, |y| > 0 \text{ gilt}$
 $\exists i \geq 0: xy^i z \notin L$

- **wir wählen** $w \in L$ in Abhängigkeit von k , in unserem Fall $w = a^k b a^k b$
- **wir zeigen** für jede Zerteilung gibt es ein i , so dass $xy^i z \notin L$
 - y besteht nur aus a s
 - $xy^i z$ enthält genau zweimal das Zeichen b
 - b s legen Endpositionen von u fest
 - $xy^0 z = a^{k-|y|} b a^k b \notin L$
 - L ist nicht pumpbar, also ist L nicht regulär

3. Beispiel $L = \{a^{n^2} \mid n \geq 0\}$

- Abstände zwischen Quadratzahlen wachsen linear



- für Pumplänge k wählen wir $w = a^{k^2} \in L$
- da $\Sigma = \{a\}$ spielt bei der Zerteilung von w nur die Länge von y eine Rolle
- wir wählen $i = 2$, d.h. $|xy^iz| \leq k^2 + k$
- die zu k^2 nächste Quadratzahl ist $(k+1)^2 = k^2 + 2k + 1$
- $k^2 < |xy^2z| \leq k^2 + k < k^2 + 2k + 1 = (k+1)^2$
- $|xy^2z|$ liegt zwischen k^2 und $(k+1)^2$ und kann demnach keine Quadratzahl sein

$\rightarrow L$ nicht pumpbar $\rightarrow L \notin \text{REG}$

2. Kapitel

Kontextfreie Sprachen

Kontextfreie Grammatiken

- Grammatiken = Formalismus zum Beschreiben von Sprachen mit Ableitungsregeln
- Ziel: $\{a^n b^n \mid n \geq 0\}$ soll erkannt werden
- Ziel: Modell soll *beherrschbar* bleiben

Kontextfreie Grammatik

Eine Kontextfreie Grammatik G ist ein 4 Tupel

(V, Σ, R, S) , bestehend aus

- V endliche Menge von **Variablen**,
- $\Sigma \neq \emptyset$ endliche Menge von **Terminalsymbolen**,
- $R \subseteq V \times (V \cup \Sigma)^*$ einer endliche Menge von **Regeln**,
- $S \in V$ einem **Startsymbol**.

■ Schreibweise:

- $(x, w) \in R$ wird geschrieben als $x \rightarrow w$

linke Seite \nearrow \nwarrow rechte Seite

- Zusammenfassen aller Regeln mit der gleichen linken Seite

$(x, w_1), (x, w_2), (x, w_3) \in R$ geschrieben als $x \rightarrow w_1 \mid w_2 \mid w_3$

- Variablen = Großbuchstaben
- Terminalsymbole = Kleinbuchstaben, Ziffern, Sonderzeichen
- wenn nicht anders angegeben, Startsymbol immer S
- es genügt i.A. die Regeln der Grammatik anzugeben

- Beispiel:
- $$\begin{array}{lcl} S & \rightarrow & aAA \mid \varepsilon \\ A & \rightarrow & aX_1 \mid a \\ X_1 & \rightarrow & baA \end{array}$$

- wenn $u, v, w \in (\Sigma \cup V)^*$ und $A \rightarrow w$ eine Regel, dann leitet uAv zu uww ab
Schreibweise: $uAv \Rightarrow uww$
- wiederholtes Ableiten: $uAv \Rightarrow \dots \Rightarrow \dots \Rightarrow t$
Schreibweise: $uAv \Rightarrow^* t$

Definition

Für eine kf Grammatik G ist die Sprache der Grammatik $L(G)$ definiert als

$$L(G) := \{w \in \Sigma^* \mid S \Rightarrow^* w\}.$$

- also alle Wörter **über** Σ die ich aus dem Startsymbol ableiten kann

Beispiel 1

$$S \rightarrow aSb \mid \varepsilon$$

$$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow \dots \Rightarrow a^k S b^k \Rightarrow a^k b^k$$

$$L(G) = \{a^n b^n \mid n \geq 0\}$$

Beispiel 2

$$S \rightarrow aSa \mid bSb \mid \varepsilon$$

$$S \Rightarrow aSa \Rightarrow abSba \Rightarrow \dots \Rightarrow uS\bar{u} \Rightarrow u\bar{u}$$

$$L(G) = \{u\bar{u} \mid u \in \{a, b\}^*\}$$

Beispiel 3

$$S \rightarrow (S) \mid SS \mid \varepsilon$$

$$L(G) = \{w \in \{(,)\}^* \mid w \text{ korrekt geklammert}\}$$

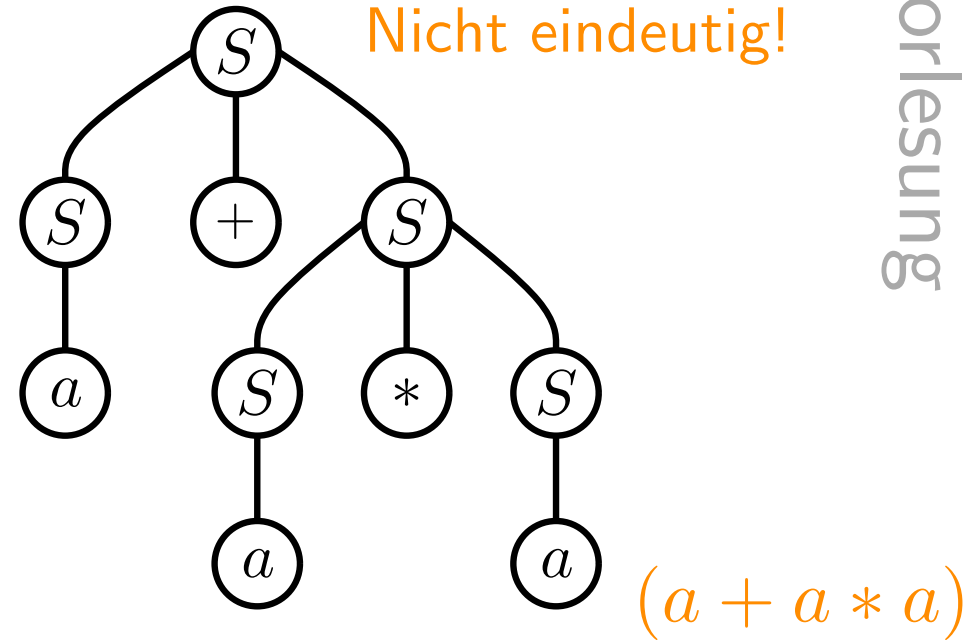
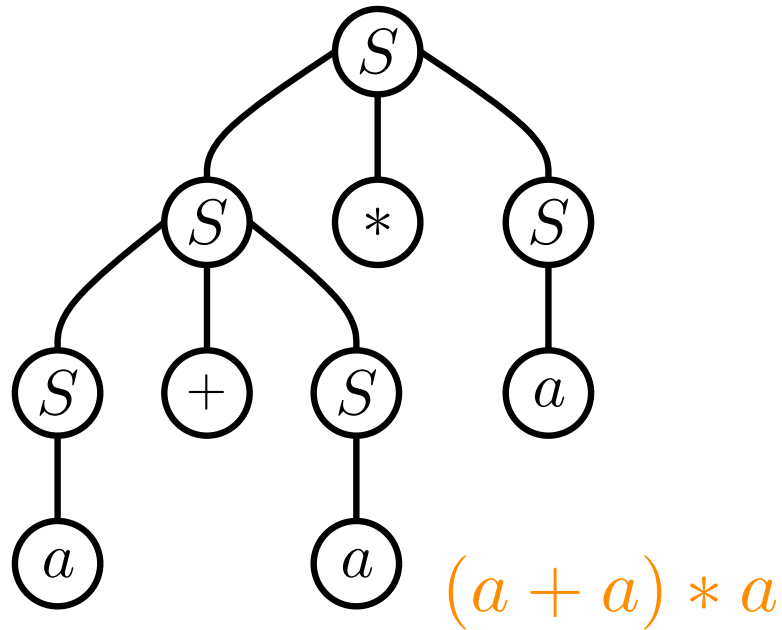
Ableitungsbäume

- beim DEA/NEA *bezeugte* der Lauf eines Wortes die Berechnung
- lineare Ableitungssequenz $S \Rightarrow \dots \Rightarrow \dots \Rightarrow w$ erfasst die *Struktur* der Berechnung unzureichend
- **Alternative:** Ableitungsbaum eines Wortes w
 - Knotenbeschrifteter Baum mit Elementen aus $\Sigma \cup V$
 - ausgezeichnete Wurzel, mit S beschriftet
 - Blätter enthalten Terminalsymbole
 - wenn Knoten Bezeichnung $X \in V$ hat und die Kinder (von links) u_1, u_2, u_3, \dots , dann muss es die Regel geben $X \rightarrow u_1 u_2 u_3$ geben
 - w = Terminale des Baumes in Post-Order

Beispiel $S \rightarrow (S + S) \mid S * S \mid a$

16

$$w = a + a * a$$



Nicht eindeutig!

Definition

Gibt es ein Wort in einer kf. Grammatik G mit zwei unterschiedlichen Ableitungsbäumen so heißt G

mehrdeutig.

Mehrdeutigkeit gilt es zu vermeiden!

Definition

Die Sprache $L(G)$ einer kf. Grammatik G heißt **kontextfreie Sprache**.

$$\text{CFL} := \{L \mid L \text{ ist kontextfrei}\}$$

Satz 10

$$L \in \text{REG} \Rightarrow L \in \text{CFL}$$

- Aus Satz 10 folgt, dass $\text{REG} \subsetneq \text{CFL} \dots$
... denn $\{a^n b^n \mid n \geq 0\}$ ist kontextfrei aber nicht regulär

