Automaten und Formale Sprachen
12. Vorlesung

Prof. Dr. Dietrich Kuske

FG Automaten und Logik, TU Ilmenau

Wintersemester 2023/24

Von Grammatiken zu PDAs

Konstruktion

Sei $G = (V, \Sigma, S, P)$ eine kontextfreie Grammatik in Greibach-Normalform.
Konstruiere den PDA $M_G = (Z, \Sigma, \Gamma, \iota, \delta, \#)$:

- $Z = \{\iota\}$ . . . es gibt nur einen Zustand
- $\Gamma = V$ . . . die Nichtterminale der Grammatik bilden das Kelleralphabet
- für $A \in V$, $a \in \Sigma$:
  $\delta(\iota, a, A) = \big\{(\iota, B_1 \ldots B_k) \bigm| (A \to aB_1 \ldots B_k) \in P\big\}$
  $\delta(\iota, \varepsilon, A) = \emptyset$ für $A \in V$ . . . es gibt keine ε-Transitionen
- $\# = S$ . . . das Kellerinitialisierungssymbol ist das Startsymbol der Grammatik

Ziel

$L(M_G) = L(G)$

Beispiel: G hat die Regeln

$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

Linksableitung	PDA-Berechnung
$S \quad \Rightarrow_G \quad 0SE$	$(\iota, 001011, S) \quad \vdash \quad (\iota, 01011, SE)$

$(S \to 0SE) \in P$, also $(\iota, SE) \in \delta(\iota, 0, S)$

Beispiel: G hat die Regeln

$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

Linksableitung			PDA-Berechnung		
S	$\Rightarrow_G$	$0SE$	$(\iota, 001011, S)$	$\vdash$	$(\iota, 01011, SE)$
	$\Rightarrow_G$	$00ESE$		$\vdash$	$(\iota, 1011, ESE)$

$(S \to 0ES) \in P$, also $(\iota, ES) \in \delta(\iota, 0, S)$

Beispiel: G hat die Regeln

$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

$$\begin{array}{llllcl}
 & \text{Linksableitung} & & \text{PDA-Berechnung} & & \\
S & \Rightarrow_G & 0SE & (\iota, 001011, S) & \vdash & (\iota, 01011, SE) \\
 & \Rightarrow_G & 00ESE & & \vdash & (\iota, 1011, ESE) \\
 & \Rightarrow_G & 001SE & & \vdash & (\iota, 011, SE)
\end{array}$$

$(E \to 1) \in P$, also $(\iota, \varepsilon) \in \delta(\iota, 1, E)$

Beispiel: G hat die Regeln

$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

Linksableitung			PDA-Berechnung		
S	$\Rightarrow_G$	$0SE$	$(\iota, 001011, S)$	$\vdash$	$(\iota, 01011, SE)$
	$\Rightarrow_G$	$00ESE$		$\vdash$	$(\iota, 1011, ESE)$
	$\Rightarrow_G$	$001SE$		$\vdash$	$(\iota, 011, SE)$
	$\Rightarrow_G$	$0010EE$		$\vdash$	$(\iota, 11, EE)$

$(S \to 0E) \in P$, also $(\iota, E) \in \delta(\iota, 0, S)$

Beispiel: G hat die Regeln

$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

$$\begin{array}{lll|lll}
\multicolumn{3}{c}{\text{Linksableitung}} & \multicolumn{3}{c}{\text{PDA-Berechnung}} \\
S & \Rightarrow_G & 0SE & (\iota, 001011, S) & \vdash & (\iota, 01011, SE) \\
 & \Rightarrow_G & 00ESE & & \vdash & (\iota, 1011, ESE) \\
 & \Rightarrow_G & 001SE & & \vdash & (\iota, 011, SE) \\
 & \Rightarrow_G & 0010EE & & \vdash & (\iota, 11, EE) \\
 & \Rightarrow_G & 00101E & & \vdash & (\iota, 1, E)
\end{array}$$

$(E \to 1) \in P$, also $(\iota, \varepsilon) \in \delta(\iota, 1, E)$

Beispiel: G hat die Regeln

$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

Linksableitung			PDA-Berechnung		
S	$\Rightarrow_G$	$0SE$	$(\iota, 001011, S)$	$\vdash$	$(\iota, 01011, SE)$
	$\Rightarrow_G$	$00ESE$		$\vdash$	$(\iota, 1011, ESE)$
	$\Rightarrow_G$	$001SE$		$\vdash$	$(\iota, 011, SE)$
	$\Rightarrow_G$	$0010EE$		$\vdash$	$(\iota, 11, EE)$
	$\Rightarrow_G$	$00101E$		$\vdash$	$(\iota, 1, E)$
	$\Rightarrow_G$	001011		$\vdash$	$(\iota, \varepsilon, \varepsilon)$

$(E \to 1) \in P$, also $(\iota, \varepsilon) \in \delta(\iota, 1, E)$

Lemma

$L(G) \subseteq L(M_G)$.

Beweis:

Sei $w \in L(G)$. Dann hat w eine Linksableitung

$$S = \alpha_0 \Rightarrow_G \alpha_1 \Rightarrow_G \cdots \Rightarrow_G \alpha_t = w\,.$$

Da G in Greibach-Normalform ist, haben alle Regeln die Form $A \to aB_1 \cdots B_k$. Also existieren $a_1, a_2, \ldots, a_t \in \Sigma$ und $\gamma_0, \ldots, \gamma_t \in V^*$ mit

$$\alpha_i = a_1 a_2 \cdots a_i \, \gamma_i \text{ für alle } 0 \le i \le t\,.$$

Insbesondere gelten also $S = \alpha_0 = \varepsilon\,\gamma_0$ und $a_1 \cdots a_t \, \gamma_t = \alpha_t = w \in \Sigma^*$, also $\gamma_0 = S$, $w = a_1 \cdots a_t$ und $\gamma_t = \varepsilon$.

<u>Behauptung:</u> Für alle $0 \leq i < t$ gilt

$$(\iota, a_{i+1} \cdots a_t, \gamma_i) \vdash (\iota, a_{i+2} \cdots a_t, \gamma_{i+1}) \,.$$

<u>Beweis der Behauptung:</u>

$$a_1 \cdots a_i \, \gamma_i = \alpha_i \Longrightarrow_G \alpha_{i+1} = a_1 \cdots a_{i+1} \, \gamma_{i+1}$$

ist Schritt in einer Linksableitung. Also existieren $(A \to a_{i+1} B_1 \cdots B_k) \in P$ und $\eta_i \in V^*$ mit $\gamma_i = A\eta_i$ und $\gamma_{i+1} = B_1 \cdots B_k \eta_i$. Also gilt $(\iota, B_1 \cdots B_k) \in \delta(\iota, a_{i+1}, A)$ und damit

$$\begin{aligned} (\iota, a_{i+1} \cdots a_t, \gamma_i) &= (\iota, a_{i+1} \cdots a_t, A\eta_i) \\ &\vdash (\iota, a_{i+2} \cdots a_t, B_1 \cdots B_k \eta_i) \\ &= (\iota, a_{i+2} \cdots a_t, \gamma_{i+1}) \,. \end{aligned}$$

q.e.d.

Da die Behauptung für alle $0 \leq i < t$ gilt, erhalten wir

$$
\begin{aligned}
(\iota, w, S) &= (\iota, a_1 \dots a_t, \gamma_0) \text{ nach Folie 12.4 unten} \\
&\vdash (\iota, a_2 \dots a_t, \gamma_1) \\
&\vdash (\iota, a_3 \dots a_t, \gamma_2) \\
&\vdots \\
&\vdash (\iota, \varepsilon, \gamma_t) \\
&= (\iota, \varepsilon, \varepsilon) \text{ nach Folie 12.4 unten}
\end{aligned}
$$

und damit $w \in L(M_G)$. $\square$

Lemma

$L(M_G) \subseteq L(G)$.

Beweis:

Sei $w \in L(M_G)$. Dann existieren $t \in \mathbb{N}$ und Konfigurationen (ι, w_i, γ_i) für $0 \leq i \leq t$, so daß

- $(\iota, w_0, \gamma_0) = (\iota, w, S)$,
- $(\iota, w_i, \gamma_i) \vdash (\iota, w_{i+1}, \gamma_{i+1})$ für alle $0 \leq i < t$ und
- $(\iota, w_t, \gamma_t) = (\iota, \varepsilon, \varepsilon)$.

Da der PDA M_G keine ε-Transitionen hat, existieren Buchstaben $a_1, \ldots, a_t \in \Sigma$ mit $w_i = a_{i+1} w_{i+1}$, also insbesondere $w = w_0 = a_1 a_2 \cdots a_t$.

Behauptung: Für alle $0 \leq i < t$ gilt

$$a_1 a_2 \cdots a_i \gamma_i \Rightarrow_G a_1 a_2 \cdots a_{i+1} \gamma_{i+1} \, .$$

Beweis der Behauptung:
Es gilt

$$(\iota, a_{i+1} w_{i+1}, \gamma_i) = (\iota, w_i, \gamma_i) \vdash (\iota, w_{i+1}, \gamma_{i+1}) \, .$$

Also existieren $\eta_i \in V^*$ und $A, B_1, \ldots, B_k \in V$ mit

$$\gamma_i = A\eta_i \, , \ (\iota, B_1 \cdots B_k) \in \delta(\iota, a_{i+1}, A) \text{ und } \gamma_{i+1} = B_1 \cdots B_k \eta_i \, .$$

Damit haben wir aber $(A \to a_{i+1} B_1 \cdots B_k) \in P$. Hieraus folgt

$$\begin{array}{rcl} a_1 \cdots a_i \gamma_i & = & a_1 \cdots a_i \, A\eta_i \\ & \Rightarrow_G & a_1 \cdots a_i \, a_{i+1} B_1 \cdots B_k \eta_i \\ & = & a_1 \cdots a_{i+1} \gamma_{i+1} \end{array}$$

q.e.d.

Da die Behauptung für alle $0 \leq i < t$ gilt, erhalten wir

$$\begin{array}{rcl} S & = & \gamma_0 \text{ nach Folie 12.7} \\ & \Rightarrow_G & a_1\gamma_1 \\ & \Rightarrow_G & a_1 a_2\gamma_2 \\ & \vdots & \\ & \Rightarrow_G & a_1 \ldots a_t\gamma_t \\ & = & w \text{ nach Folie 12.7} \end{array}$$

und damit $w \in L(G)$ $\square$

Proposition

Jede kontextfreie Sprache L ist Sprache eines PDA M mit nur einem Zustand. Gilt $\varepsilon \notin L$, so werden keine ε-Transitionen benötigt.

Beweis:
Sei L kontextfrei. Dann existiert eine Grammatik $G = (V, \Sigma, S, P)$ in Greibach-Normalform mit $L(G) = L \setminus \{\varepsilon\}$. Nach den beiden gezeigten Lemmata existiert ein PDA M mit einem Zustand, ohne ε-Transitionen und mit $L(M) = L(G) = L \setminus \{\varepsilon\}$. Gilt $\varepsilon \notin L$, so ist Proposition bewiesen.

Angenommen, $\varepsilon \in L$.
Wir können o.E. annehmen, daß S auf keiner rechten Seite einer Regel aus G vorkommt. Dann wird der Kellerautomat M_G das Kellerinitialisierungssymbol S im ersten Schritt auf dem Keller ersetzen und niemals wieder auf den Keller schreiben. Fügt man die Transition $\delta(\iota, \varepsilon, S) = \big\{(\iota, \varepsilon)\big\}$ zu M_G hinzu, so erhält man einen PDA M'_G mit $L(M'_G) = L(M_G) \cup \{\varepsilon\} = L(G) \cup \{\varepsilon\} = L \setminus \{\varepsilon\} \cup \{\varepsilon\} = L.$ $\square$

Beispiel: G hat die Regeln

$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

PDA-Berechnung	Ableitungsbaum
$(\iota, 001011, S)$	S

Beispiel: G hat die Regeln

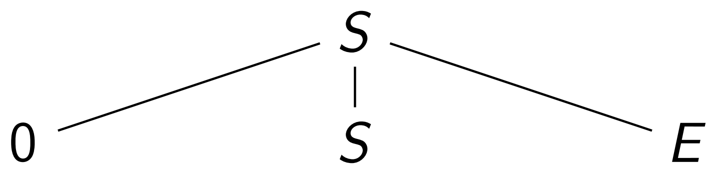
$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

<u>PDA-Berechnung</u>

$(\iota, 001011, S)$

$\vdash (\iota, 01011, SE)$

<u>Ableitungsbaum</u>



$(S \to 0SE) \in P$

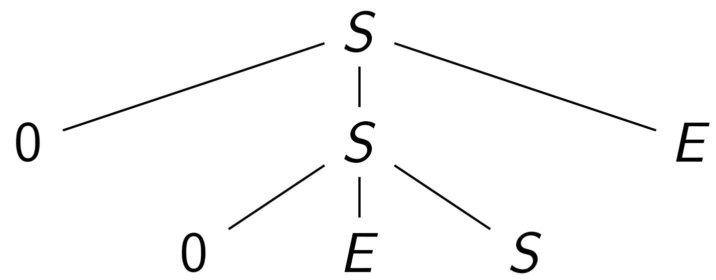
Beispiel: G hat die Regeln

$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

<u>PDA-Berechnung</u>

$$(\iota, 001011, S)$$
$$\vdash (\iota, 01011, SE)$$
$$\vdash (\iota, 1011, ESE)$$

<u>Ableitungsbaum</u>



$(S \to 0ES) \in P$

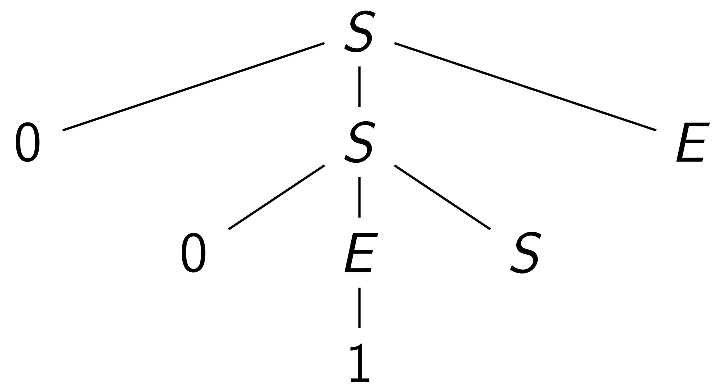
Beispiel: G hat die Regeln

$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

PDA-Berechnung

$$\begin{aligned}
&(\iota, 001011, S)\\
&\quad\vdash (\iota, 01011, SE)\\
&\quad\vdash (\iota, 1011, ESE)\\
&\quad\vdash (\iota, 011, SE)
\end{aligned}$$

Ableitungsbaum



$(E \to 1) \in P$

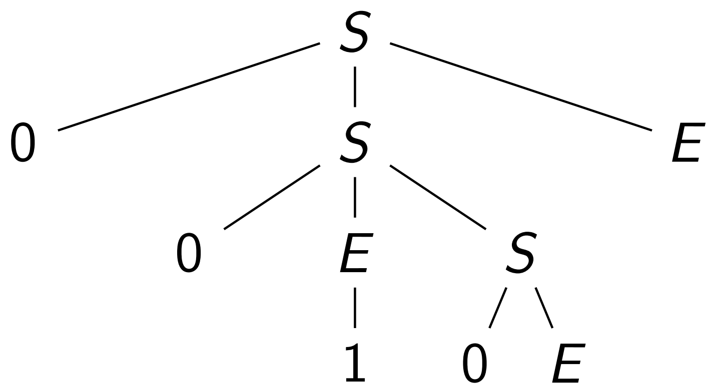
Beispiel: G hat die Regeln

$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

PDA-Berechnung

$$\begin{aligned}
&(\iota, 001011, S) \\
&\quad \vdash (\iota, 01011, SE) \\
&\quad \vdash (\iota, 1011, ESE) \\
&\quad \vdash (\iota, 011, SE) \\
&\quad \vdash (\iota, 11, EE)
\end{aligned}$$

Ableitungsbaum



$(S \to 0E) \in P$

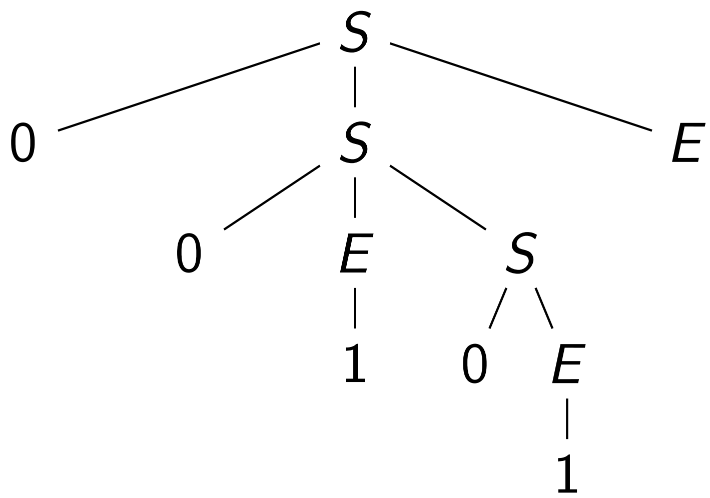
Beispiel: G hat die Regeln

$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

PDA-Berechnung

$$
\begin{aligned}
&(\iota, 001011, S)\\
&\quad\vdash (\iota, 01011, SE)\\
&\quad\vdash (\iota, 1011, ESE)\\
&\quad\vdash (\iota, 011, SE)\\
&\quad\vdash (\iota, 11, EE)\\
&\quad\vdash (\iota, 1, E)
\end{aligned}
$$

Ableitungsbaum



$(E \to 1) \in P$

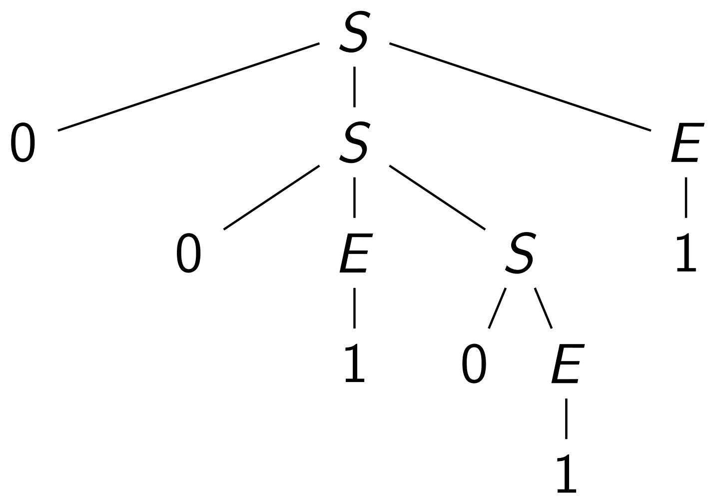
Beispiel: G hat die Regeln

$$S \to 0SES \mid 0ES \mid 0SE \mid 0E \text{ und } E \to 1$$

<u>PDA-Berechnung</u>

$$\begin{aligned} &(\iota, 001011, S) \\ &\quad \vdash (\iota, 01011, SE) \\ &\quad \vdash (\iota, 1011, ESE) \\ &\quad \vdash (\iota, 011, SE) \\ &\quad \vdash (\iota, 11, EE) \\ &\quad \vdash (\iota, 1, E) \\ &\quad \vdash (\iota, \varepsilon, \varepsilon) \end{aligned}$$

<u>Ableitungsbaum</u>



$(E \to 1) \in P$

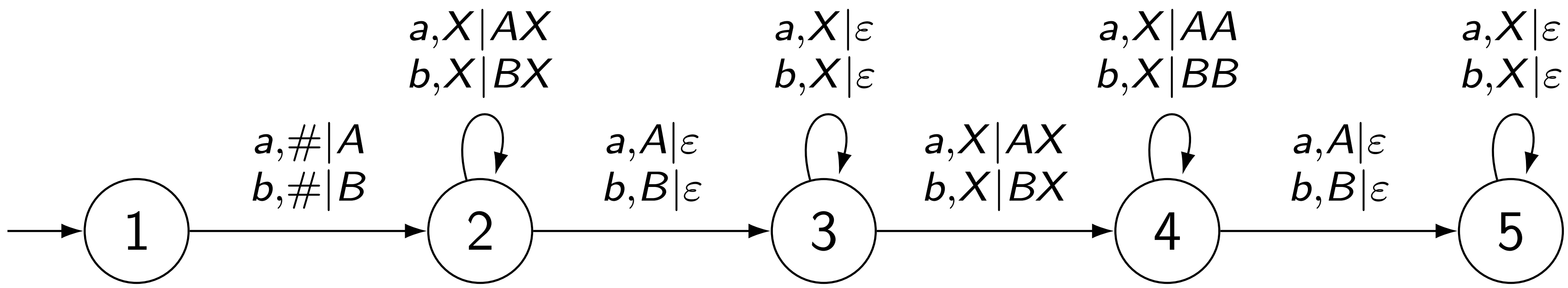
Bemerkungen:

- vorgestellte Methode heißt LL-Parsing: Berechnungen des PDA entsprechen Linksableitungen und Wort wird von links nach rechts gelesen
  oder Top-Down-Parsing: Ableitungsbaum wird an der Wurzel beginnend erzeugt.
- Aus Grammatik-Regeln $A \to aU \mid aV$ entsteht der Nichtdeterminismus $(\iota, U), (\iota, V) \in \delta(\iota, a, A)$. LL(1)-Grammatiken enthalten keine solche Regeln, so daß M_G deterministisch wird - und erlauben, die Frage „Wird das Wort w von der Grammatik G erzeugt?“ in linearer Zeit zu beantworten.
- allgemeiner: in LL(k)-Grammatiken legen die nächsten k zu lesenden Buchstaben fest, welche Regel angewandt wird - auch dann kann man deterministischen PDA konstruieren.
- LR-Parsing oder Bottom-Up-Parsing sind alternative Methoden, einen PDA aus einer kontextfreien Grammatik zu konstruieren, auch LR(k)-Grammatiken führen zu deterministischen PDAs.

Von PDAs zu Grammatiken

Beispiel:

PDA M: (hierbei steht X für ein beliebiges Kellersymbol aus $\{A, B\}$)



eine akzeptierende Berechnung:

$$\begin{array}{llll}
(1, aabbabbbaa, \#) & \vdash (2, abbabbbaa, A) & \vdash (2, bbabbbaa, AA) \\
 & \vdash (2, babbbaa, BAA) & \vdash (3, abbbaa, AA) \\
 & \vdash (4, bbbaa, AAA) & \vdash (4, bbaa, BBAA) \\
 & \vdash (5, baa, BAA) & \vdash (5, aa, AA) \\
 & \vdash (5, a, A) & \vdash (5, \varepsilon, \varepsilon)
\end{array}$$

die akzeptierende Berechnung nochmal anders hingeschrieben:

$a \quad a \quad b \quad b \quad a \quad b \quad b \quad b \quad a \quad a$

						$_4B_5$				
			$_2B_3$		$_4A_5$	$_5B_5$	$_5B_5$			
		$_2A_5$	$_3A_5$	$_3A_5$	$_5A_5$	$_5A_5$	$_5A_5$	$_5A_5$		
$_1\#_5$	$_2A_5$	$_5A_5$	$_5A_5$	$_5A_5$	$_5A_5$	$_5A_5$	$_5A_5$	$_5A_5$	$_5A_5$	
1	2	2	2	3	4	4	5	5	5	5

Bedeutung von $\boxed{{}_iX_j}$:

- X ... Kellersymbol
- i ... dieses Symbol wird im Zustand i ersetzt werden
- j ... sobald der Keller das erste Mal echt kürzer sein wird, befindet der PDA sich im Zustand j

Ableitung der Grammatik:

$$
\begin{aligned}
S \quad &\Rightarrow \boxed{{}_1\#_5} \\
&\Rightarrow a\,\boxed{{}_2A_5} \\
&\Rightarrow aa\,\boxed{{}_2A_5}\,\boxed{{}_5A_5} \\
&\Rightarrow aab\,\boxed{{}_2B_3}\,\boxed{{}_3A_5}\,\boxed{{}_5A_5} \\
&\Rightarrow aabb\,\boxed{{}_3A_5}\,\boxed{{}_5A_5} \\
&\Rightarrow aabba\,\boxed{{}_4A_5}\,\boxed{{}_5A_5}\,\boxed{{}_5A_5} \\
&\Rightarrow aabbab\,\boxed{{}_4B_5}\,\boxed{{}_5B_5}\,\boxed{{}_5A_5}\,\boxed{{}_5A_5} \\
&\Rightarrow aabbabb\,\boxed{{}_5B_5}\,\boxed{{}_5A_5}\,\boxed{{}_5A_5} \\
&\Rightarrow aabbabbb\,\boxed{{}_5A_5}\,\boxed{{}_5A_5} \\
&\Rightarrow aabbabbba\,\boxed{{}_5A_5} \\
&\Rightarrow aabbabbbaa
\end{aligned}
$$

Ziel / Idee

kontextfreie Grammatik G, so daß für alle $w \in \Sigma^*$:

$$(i, w, A) \vdash^* (j, \varepsilon, \varepsilon) \text{ gdw. } \boxed{{}_iA_j} \Rightarrow_G^* w$$

Schreibweise

Im Folgenden schreiben wir (i, A, j) an Stelle von $\boxed{{}_iA_j}$ (um doppelte Indizes zu vermeiden).

Konstruktion („Tripel-Konstruktion“)

Sei $M = (Z, \Sigma, \Gamma, \delta, \iota, \#)$ ein PDA. Konstruiere die kontextfreie Grammatik $G_M = (V, \Sigma, P, S)$:

- $V = \{S\} \cup (Z \times \Gamma \times Z)$
- Wir haben die folgenden Produktionen:
    - $S \to (\iota, \#, z)$ für alle $z \in Z$
    - $(z, A, z_k) \to a\,(z_0, B_1, z_1)\,(z_1, B_2, z_2)\,\ldots\,(z_{k-1}, B_k, z_k)$ für alle $z, z_0, z_1, \ldots, z_k \in Z$, $A \in \Gamma$ und $a \in \Sigma \cup \{\varepsilon\}$ mit $(z_0, B_1 B_2 \ldots B_k) \in \delta(z, a, A)$

Insbesondere wird aus der Transition $(z_0, \varepsilon) \in \delta(z, a, A)$ wegen $k = 0$ und damit $z_k = z_0$ die Produktion

$$(z, A, z_0) \to a$$

Für jede Transition $(z_0, B_1 B_2 \cdots B_k) \in \delta(z, a, A)$ und alle Zustände $z_1, \ldots, z_{k-1}, z_k$ bilde die folgende Regel der Grammatik:

$$(z, A, z_k) \rightarrow a\,(z_0, B_1, z_1)\,(z_1, B_2, z_2)\,(z_2, B_3, z_3)\,\cdots\,(z_{k-1}, B_k, z_k)$$

1. Schreibe die Transition als Muster der Regel auf.
2. Schließe die Lücken vor den Nichtterminalen B_2 bis B_k durch beliebige Zustände z_1 bis z_{k-1} und die Lücke hinter A durch beliebigen Zustand z_k.
3. Schließe die Lücken hinter B_1 bis B_{k-1} durch die folgenden Zustände z_1 bis z_{k-1} und die Lücke hinter B_k durch den Zustand z_k.

Lemma

Für alle $z_0, z_1, z_2, z_3, \ldots, z_\ell \in Z$, $B_1, B_2, \ldots, B_\ell \in \Gamma$ und $w \in \Sigma^*$ mit

$$(z_0, B_1, z_1)(z_1, B_2, z_2) \cdots (z_{\ell-1}, B_\ell, z_\ell) \Longrightarrow^*_{G_M} w$$

gilt

$$(z_0, w, B_1 B_2 \cdots B_\ell) \vdash^* (z_\ell, \varepsilon, \varepsilon)\,.$$

Beweis: per Induktion über die Länge n der Ableitung.

IA $n = 0$: Dann gilt

$$(z_0, B_1, z_1)(z_1, B_2, z_2) \cdots (z_{\ell-1}, B_\ell, z_\ell) = w\,,$$

also $\ell = 0$ und $w = \varepsilon$ und damit

$$(z_0, w, B_1 \cdots B_\ell) = (z_0, \varepsilon, \varepsilon) = (z_\ell, \varepsilon, \varepsilon)\,.$$

IS Gelte $(z_0, B_1, z_1)(z_1, B_2, z_2)\cdots(z_{\ell-1}, B_\ell, z_\ell) \Longrightarrow_{G_M}^{n+1} w$.
Wir betrachten eine Linksableitung der Länge $n+1$. Aufgrund der Gestalt der Produktionen der Grammatik G_M hat diese die Form

$$
\begin{array}{ll}
 & (z_0, B_1, z_1)(z_1, B_2, z_2)\cdots(z_{\ell-1}, B_\ell, z_\ell) \\
\Rightarrow_{G_M} & a(z_0', C_1, z_1')(z_1', C_2, z_2')\cdots(z_{k-1}', C_k, z_1)\,(z_1, B_2, z_2)\cdots(z_{\ell-1}, B_\ell, z_\ell) \\
\Rightarrow_{G_M}^{n} & w\,.
\end{array}
$$

Insbesondere ist $(z_0, B_1, z_1) \to a(z_0', C_1, z_1')(z_1', C_2, z_2')\cdots(z_{k-1}', C_k, z_1)$ eine Produktion, $a \in \Sigma \cup \{\varepsilon\}$ und $w = aw'$ für ein Wort $w' \in \Sigma^*$.
Nach Konstruktion der Grammatik G_M erhalten wir

$$(z_0', C_1\,C_2\cdots C_k) \in \delta(z_0, a, B_1)\,,$$

also

$$(z_0, aw', B_1\cdots B_\ell) \vdash (z_0', w', C_1\cdots C_k\,B_2\cdots B_\ell)\,.$$

Wir haben auch

$$(z_0', C_1, z_1') \cdots (z_{k-1}', C_k, z_1)\,(z_1, B_2, z_2) \cdots (z_{\ell-1}, B_\ell, z_\ell) \Longrightarrow_{G_M}^{n} w'\,.$$

Nach der IV folgt

$$(z_0', w', C_1 C_2 \cdots C_k\, B_2 B_3 \cdots B_\ell) \vdash^* (z_\ell, \varepsilon, \varepsilon)\,.$$

Damit erhalten wir

$$\begin{aligned} (z_0, w, B_1 \cdots B_\ell) &= (z_0, aw', B_1 \cdots B_\ell) \\ &\vdash (z_0', w', C_1 \cdots C_k\, B_2 \cdots B_\ell) \\ &\vdash^* (z_\ell, \varepsilon, \varepsilon)\,. \end{aligned}$$

$\square$

Lemma

Für alle $z_0, z_\ell \in Z$, $B_1, B_2, \ldots, B_\ell \in \Gamma$ und $w \in \Sigma^*$ mit

$$(z_0, w, B_1 B_2 \cdots B_\ell) \vdash^* (z_\ell, \varepsilon, \varepsilon)$$

existieren Zustände $z_1, z_2, \ldots, z_{\ell-1}$ mit

$$(z_0, B_1, z_1)(z_1, B_2, z_2) \cdots (z_{\ell-1}, B_\ell, z_\ell) \Longrightarrow^*_{G_M} w$$

Beweis: per Induktion über die Länge n der Berechnung.

IA $n = 0$: Dann gelten $(z_0, w, B_1 \cdots B_\ell) = (z_\ell, \varepsilon, \varepsilon)$, also

$$(z_0, B_1, z_1)(z_1, B_2, z_2) \cdots (z_{\ell-1}, B_\ell, z_\ell) = \varepsilon \Longrightarrow^*_{G_M} \varepsilon = w$$

IS Gelte $(z_0, w, B_1 B_2 \cdots B_\ell) \vdash^{n+1} (z_\ell, \varepsilon, \varepsilon)$.
Dann existiert eine Konfiguration $(z_0', w', C_1 C_2 \cdots C_k)$ mit

$$(z_0, w, B_1 B_2 \cdots B_\ell) \vdash (z_0', w', C_1 C_2 \cdots C_k\, B_2 B_3 \cdots B_\ell) \vdash^n (z_\ell, \varepsilon, \varepsilon)\,.$$

Nach IV existieren Zustände $z'_1, z'_2, \ldots, z'_{k-1}, z_1, z_2, \ldots, z_{\ell-1}$ mit

$(z'_0, C_1, z'_1)(z'_1, C_2, z'_2)\cdots(z'_{\ell-1}, C_\ell, z_1)\,(z_1, B_2, z_2)\cdots(z_{\ell-1}, B_\ell, z_\ell) \;\Rightarrow^*_{G_M} w'$

Wegen

$$(z_0, w, B_1 B_2 \cdots B_\ell) \vdash (z'_0, w', C_1 C_2 \cdots C_k\, B_2 B_3 \cdots B_\ell)$$

existiert $a \in \Sigma \cup \{\varepsilon\}$ mit $(z'_0, C_1 \cdots C_k) \in \delta(z_0, a, B_1)$ und $w = aw'$.
Nach Konstruktion der Grammatik G_M erhalten wir die Produktion

$$(z_0, B_1, z_1) \to a(z'_0, C_1, z'_1)(z'_1, C_2, z'_2)\cdots(z'_{\ell-1}, C_\ell, z_1)\,.$$

Damit gilt

$$\begin{aligned}
& (z_0, B_1, z_1)(z_1, B_2, z_2)\cdots(z_{\ell-1}, B_\ell, z_\ell) \\
\Rightarrow_{G_M}\; & a(z'_0, C_1, z'_1)(z'_1, C_2, z'_2)\cdots(z'_{\ell-1}, C_\ell, z_1)(z_1, B_2, z_2)\cdots(z_{\ell-1}, B_\ell, z_\ell) \\
\Rightarrow^*_{G_M}\; & aw' = w\,.
\end{aligned}$$

$\square$

Proposition

Ist $M = (Z, \Sigma, \Gamma, \delta, \iota, \#)$ ein PDA, so ist $L(M)$ kontextfrei.

Beweis: Wir zeigen zunächst $L(M) \subseteq L(G_M)$: Sei $w \in L(M)$. Dann existiert $z' \in Z$ mit $(\iota, w, \#) \vdash^* (z', \varepsilon, \varepsilon)$. Wir erhalten

$$\begin{aligned} S &\Longrightarrow_{G_M} (\iota, \#, z') && \text{(nach Konstruktion von } G_M\text{)} \\ &\Longrightarrow^*_{G_M} w && \text{(nach Lemma auf Folie 12.19).} \end{aligned}$$

Also gilt $w \in L(G_M)$, d.h. $L(M) \subseteq L(G_M)$.

Wir zeigen jetzt $L(M) \supseteq L(G_M)$:
Sei $w \in L(G_M)$.
Dann existiert $z' \in Z$ mit $S \implies_{G_M} (\iota, \#, z') \implies^*_{G_M} w$.
Aus dem Lemma auf Folie 12.22 folgt $(\iota, w, \#) \vdash^* (z', \varepsilon, \varepsilon)$ und damit $w \in L(M)$. $\square$

Satz

Sei L eine Sprache. Dann sind äquivalent:

1. L ist kontextfrei.
2. Es gibt einen PDA M mit $L(M) = L$.
3. Es gibt einen PDA M mit nur einem Zustand und $L(M) = L$.

Gilt $\varepsilon \notin L$, so sind diese Aussagen äquivalent zu

4. Es gibt einen PDA M mit nur einem Zustand und ohne ε-Transitionen, so daß $L(M) = L$ gilt.

Beweis:

(1)$\Rightarrow$(3) bzw. (4): Proposition auf Folie 12.10
(4)$\Rightarrow$(3)$\Rightarrow$(2): trivial
(2)$\Rightarrow$(1): Proposition auf Folie 12.24 $\square$

Zusammenfassung 12. Vorlesung

in dieser Vorlesung neu

- Kellerautomaten akzeptieren genau die kontextfreien Sprachen
- dabei ist ein einzelner Zustand ausreichend (und oft braucht man auch keine ε-Transitionen)

kommende Vorlesung

- Abschlußeigenschaften der Klasse der kontextfreien Sprachen
- Nicht-Abschlußeigenschaften der Klasse der kontextfreien Sprachen hierfür: Pumping-Lemma (vgl. Vorlesung 6)