Theorem

Es sei $G = (N, T, P, S)$ eine CFG ohne ϵ-Produktionen. Dann gibt es einen PDA M mit $N(M) = L(G)$.

Beweis.

$M = (\{q\}, T, N \cup T, \delta, q, S)$ wobei

- $\delta(q, a, a) = \{(q, \epsilon)\}$ für alle $a \in T$
- $\delta(q, \epsilon, A) = \{\, (q, \alpha) \mid A \to \alpha \in P \,\}$ für alle $A \in N$.

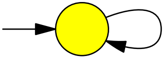
Behauptung:

$(q, w, S) \overset{*}{\vdash} (q, \epsilon, \epsilon)$ gdw. $S \overset{*}{\Rightarrow} w$ $\square$

Beispiel

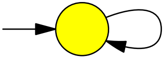
$$S \to aSa \mid bSb \mid a \mid b$$

Zugehöriger PDA:



$$\begin{array}{r|l} \epsilon, S & aSa \\ \epsilon, S & bSb \\ \epsilon, S & a \\ \epsilon, S & b \\ a, a & \epsilon \\ b, b & \epsilon \end{array}$$

bzw.



$$\begin{array}{r|l} a, S & Sa \\ b, S & Sb \\ a, S & \epsilon \\ b, S & \epsilon \\ a, a & \epsilon \\ b, b & \epsilon \end{array}$$

Das Kellerbodensymbol ist S.

Theorem

Es sei $M = (Q, \Sigma, \Gamma, \delta, q_0, \Gamma_0)$ *ein PDA.*

Dann gibt es eine CFG G *mit* $L(G) = N(M)$*.*

Beweis.

Für jedes $p \in Q$ erzeuge eine Grammatik
$G_p = (N, \Sigma, P, [q_0, \Gamma_0, p])$ wobei $N = \{[q, Z, r] | q, r \in Q, Z \in \Gamma\}$.
P enthält folgende Regeln:

$$[q, Z, q_k] \to a[r, \gamma_1, q_1][q_1, \gamma_2, q_2] \cdots [q_{k-1}, \gamma_k, q_k]$$

falls

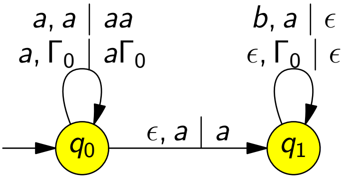
- $(r, \gamma) \in \delta(q, a, Z)$ mit $a \in \Sigma \cup \{\epsilon\}$
- $\gamma = \gamma_1 \cdots \gamma_k$ mit $\gamma_i \in \Gamma$
- $q_i \in Q$ (alle Möglichkeiten!)

Spezialfall: $[q, Z, r] \to a$ falls $\gamma = \epsilon$.

Es gilt $(q_0, w, \Gamma_0) \overset{*}{\vdash} (p, \epsilon, \epsilon)$ gdw. $[q_0, \Gamma_0, p] \overset{*}{\Rightarrow} w$ d.h.

$$N(M) = \bigcup_{p \in Q} L(G_p).$$

Beispiel



$$\begin{aligned}
[q_0, \Gamma_0, q_1] &\to a[q_0, a, q_1][q_1, \Gamma_0, q_1] \\
[q_0, a, q_1] &\to [q_1, a, q_1] \\
[q_1, a, q_1] &\to b \\
[q_1, \Gamma_0, q_1] &\to \epsilon
\end{aligned}$$

$$[q_0, \Gamma_0, q_1] \Rightarrow a[q_0, a, q_1][q_1, \Gamma_0, q_1] \Rightarrow ab[q_1, \Gamma_0, q_1] \Rightarrow ab$$

Definition

Ein PDA $M = (Q, \Sigma, \Gamma, \delta, q_0, \Gamma_0, F)$ ist ein DPDA, falls

1. $|\delta(q, a, X)| \leq 1$ für alle $q \in Q, a \in \Sigma \cup \{\epsilon\}, X \in \Gamma$
2. Falls $\delta(q, a, X) \neq \emptyset$ für ein $q \in Q, X \in \Gamma, a \in \Sigma$, dann ist $\delta(q, \epsilon, X) = \emptyset$.

Eine Sprache L ist eine deterministische CFL, falls es einen DPDA M gibt mit $L(M) = L$. Sie heißen DCFL.

Theorem

DCFL ist unter Komplement abgeschlossen.

D.h. falls $L \in DCFL$, dann auch $\Sigma^* \backslash L \in DCFL$.

Frage: Warum ist der Beweis nicht trivial?

Lemma

Es sei $L \in DCFL$.

Dann gibt es einen DPDA M, der folgende Eigenschaften hat:

1. $L(M) = L$.
2. *Jede erreichbare Konfiguration (q, w, γ) hat eine Nachfolgekonfiguration, falls $w \neq \epsilon$.*
3. *Es gibt keine unendlichen Folgen $(q, \epsilon, \gamma) \vdash (q', \epsilon, \gamma') \vdash \cdots$*

Beweis.

Konstruktion:

Wir erreichen

1. durch einen Fangzustand und ein zusätzliches Kellerbodensymbol, das nie entfernt wird,
2. durch: Falls es eine solche unendliche Folge gibt, die mit (q, ϵ, Z) beginnt, dann setze
   - $\delta(q, \epsilon, Z) := \{(q_\emptyset\ Z)\}$, wobei $q_\emptyset$ der Fangzustand ist, falls ab q kein Endzustand durchlaufen wird,
   - $\delta(q, \epsilon, Z) := \{(q_f, Z)\}$ (neuer Endzustand) und $\delta(q_f, \epsilon, Z) := \{(q_\emptyset, Z)\}$.

Wir haben jetzt einen PDA, der die ganze Eingabe liest.

(Er blockiert nie.) □

Beweis (des Theorems)

Ersetze Q durch $Q' = Q \times \{1,2,3\}$ und F durch $F' = Q \times \{3\}$ und δ durch δ' mit:
Falls $\delta(q, \epsilon, Z) = \{(p, \gamma)\}$ setze

$$\delta'((q,1),\epsilon,Z) := \begin{cases} \{((p,1),\gamma)\} & \text{falls } p \notin F \\ \{((p,2),\gamma)\} & \text{falls } p \in F \end{cases}$$

somit $\delta'((q,2),\epsilon,Z) := \{((p,2),\gamma)\}$
Falls $\delta(q,a,Z) = \{(p,\gamma)\}$ für $a \in \Sigma$ setze
$\delta'((q,1),\epsilon,Z) := \{((q,3),Z)\}$ und

$$\delta'((q,2),a,Z) = \delta'((q,3),a,Z) := \begin{cases} \{(p,1),\gamma\} & \text{falls } p \notin F \\ \{(p,2),\gamma\} & \text{falls } p \in F \end{cases}$$

Neuer Startzustand: $(q_0, 1)$ falls $q_0 \in F$, $(q_0, 2)$ sonst.
„1“: Nach dem letzten Zeichens keinen Endzustand durchlaufen
„2“: Nach dem letzten Zeichens einen Endzustand durchlaufen

Theorem

Es seien L und L' kontextfreie und R eine reguläre Sprache. Dann sind $L \cap R$, $L \setminus R$, $L \cup L'$ und LL' wieder kontextfreie Sprachen.

Beweis.

Schnitt und Differenz: Produktautomat.

Vereinigung: $G_1 = (N, T, P_1, S_1)$ und $G_2 = (N, T, P_2, S_2)$.
$G = (N, T, P_1 \cup P_2 \cup \{S \to S_1 \mid S_2\}, S)$, $L(G) = L(G_1) \cup L(G_2)$.

Konkatenation: $G_1 = (N, T, P_1, S_1)$ und $G_2 = (N, T, P_2, S_2)$.
$G = (N, T, P_1 \cup P_2 \cup \{S \to S_1 S_2\}, S)$, $L(G) = L(G_1)L(G_2)$.

Kleene'sche Hülle: $G_1 = (N, T, P_1, S_1)$.
$G = (N, T, P_1 \cup \{S \to \epsilon \mid SS \mid S_1\}, S)$, $L(G) = L(G_1)^*$. □

Theorem

CFL ist nicht unter Komplement abgeschlossen und damit auch nicht unter Schritt.

Beweis.

Es sei $L = \{\, a^i b^j c^k \mid i \geq j \vee k \geq j \vee i + k \leq j \,\}$

Offensichtlich ist L kontextfrei.

Es sei $\bar{L} = \{a, b, c\}^* \setminus L$ und $L' = \bar{L} \cap a^* b^* c^*$.

Falls $\bar{L}$ kontextfrei wäre, dann auch L'.

$L' = \{\, a^i b^j c^k \mid i < j \wedge k < j \wedge i + k > j \,\}$ ist aber nicht kontextfrei (Pumping-Lemma).

Also ist auch $\bar{L}$ nicht kontextfrei. $\square$

Theorem

DCFL ist nicht unter Vereinigung und daher auch nicht unter Schnitt abgeschlossen.

Beweis.

Wir definieren drei DCFLs:

- $L_1 = \{\, a^i b^j c^k \mid i \geq j \,\}$,
- $L_2 = \{\, a^i b^j c^k \mid k \geq j \,\}$,
- $L_3 = \{\, a^i b^j c^k \mid i + k \leq j \,\}$.

$L_1 \cup L_2 \cup L_3$ ist eine CFL, ihr Komplement ist aber keine CFL.

Wäre $L_1 \cup L_2 \cup L_3$ eine DCFL, wäre auch das Komplement eine CFL. $\square$

Die Chomsky-Hierarchie besteht aus vier Stufen:

- Chomsky-0: Rekursiv aufzählbare Sprachen
  Unbeschränkte Grammatiken: $abAc \rightarrow Bcb$
- Chomsky-1: Kontextsensitive Sprachen
  Kontextsensitive Grammatiken: $abAc \rightarrow abBCc$, $aBCD \rightarrow abCaAbCD$, kein $A \rightarrow \epsilon$
- Chomsky-2: Kontextfreie Sprachen
  Kontextfreie Grammatiken: $A \rightarrow BC$, $B \rightarrow bCaAb$
- Chomsky-3: Reguläre Sprachen
  Linkslineare Grammatiken: $A \rightarrow a$, $B \rightarrow bC$