

Abschlußeigenschaften kontextfreier Sprachen

Satz: Die kontextfreien Sprachen sind abgeschlossen gegen

(i) **Vereinigung**

Beweisidee: G_1 und G_2 seien Grammatiken mit den Startsymbolen S_1, S_2 . Dann bilde die Vereinigungsgrammatik G mit dem Startsymbol S durch Hinzufügen der Regeln $S \rightarrow S_1$ und $S \rightarrow S_2$.

(ii) **Konkatenation**

Beweisidee: G_1 und G_2 seien Grammatiken mit den Startsymbolen S_1, S_2 . Dann bilde die Grammatik G mit dem Startsymbol S durch Hinzufügen der Regel $S \rightarrow S_1 S_2$.

(iii) **Sternbildung**

Beweisidee: G sei eine Grammatik mit dem Startsymbol S . Dann bilde die Grammatik G mit dem Startsymbol S' durch Hinzufügen der Regeln $S' \rightarrow \epsilon$ und $S' \rightarrow SS'$.

(iv) **Homomorphismen**

(v) **Durchschnitt regulärer Mengen**

Beweisidee: Parallele Simulation einer regulären Menge A und einer kontextfreien Grammatik G durch Übergang zum kartesischen Produkt

Beweis:

Sei L eine kontextfreie Sprache und $G = (N, T, P, \bar{\delta})$ eine kontextfrei Grammatik in Chomsky-Normalform mit $L = L(G)$.

Sei R eine reguläre Menge und $A = (X, S, \delta, s_0, F)$ ein Akzeptor mit $R = L(A)$.

Bilde eine Grammatik $\bar{G} = (\bar{N}, T \cup X, \bar{P}, \bar{\delta})$ mit:

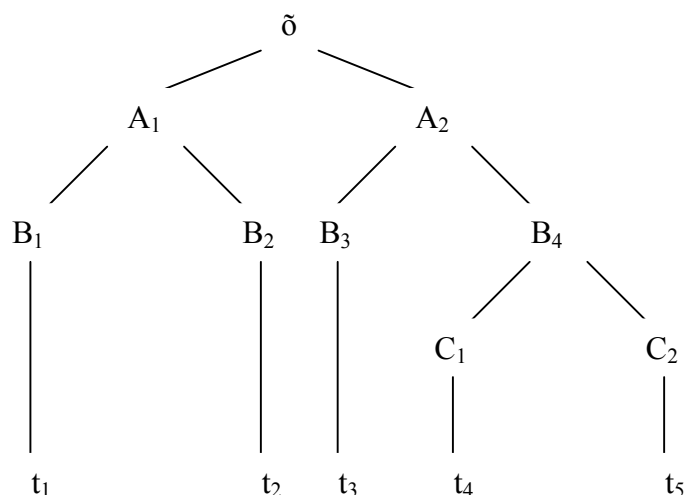
- $\bar{N} = \{\bar{\delta}\} \cup (S \times N \times S)$

- $\bar{P} = \{\bar{\delta} \rightarrow (s_0, \bar{\delta}, s) \mid s \in F\} \cup \{(s_1, A, s_2) \rightarrow (s_1, B_1, s_3) (s_3, B_2, s_2) \mid \forall s_1, s_2, s_3 \in S \text{ und } \forall \text{ Regeln } A \rightarrow B_1 B_2 \in P, A, B_1, B_2 \in N\} \cup \{(s_1, A, s_2) \rightarrow t \mid \forall A \rightarrow t \in P \text{ und } \forall s_1, s_2 \in S \text{ mit } \delta(s_1, t) = s_2\}$

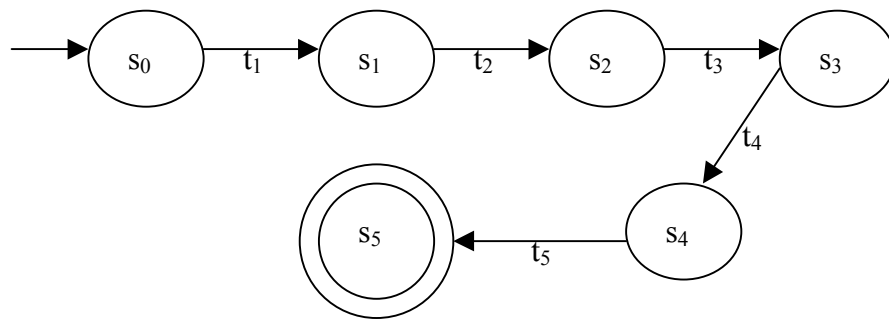
Falls $\epsilon \in R \cap L$, dann füge $\bar{\delta} \rightarrow \epsilon$ hinzu.

Beispiel: $\bar{\delta} \rightarrow^* t_1 t_2 t_3 t_4 t_5 \in (L(A) \cap L(G))$

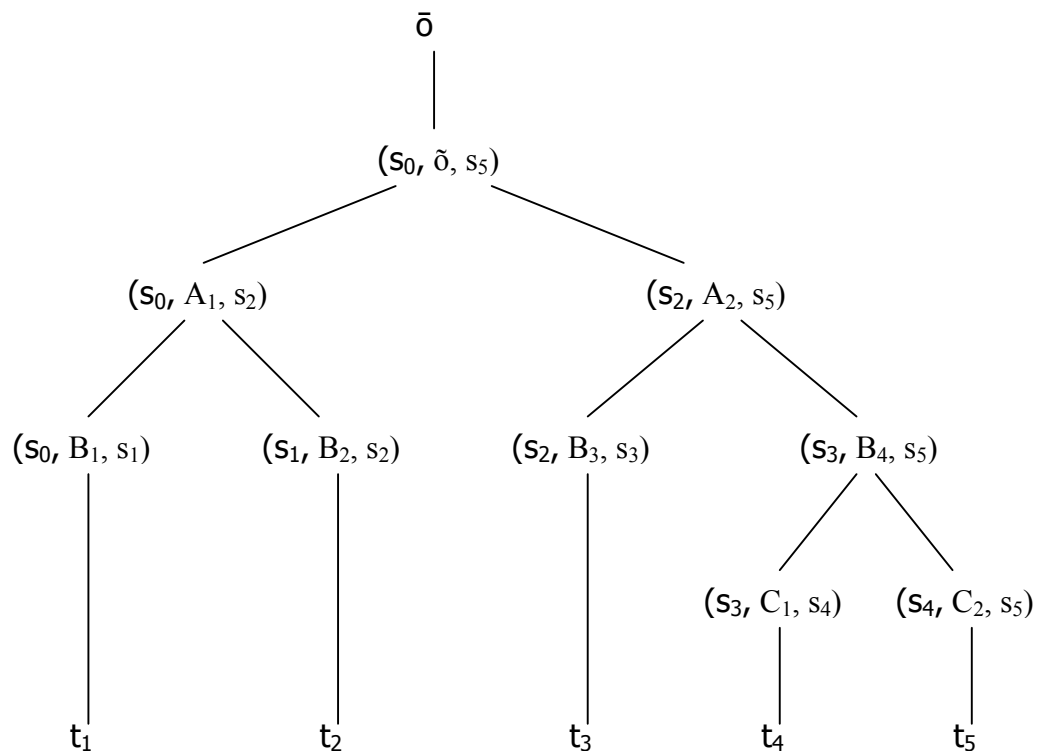
- in G gibt es eine Ableitung:



- in A gibt es einen Weg:



- in $\bar{G}$ gibt es dann eine Ableitung der Form:



$L(\bar{G}) \subseteq L(G) \cap L(A)$, weil:

Betrachtet man in einer Ableitung für $\bar{G}$ nur die Mittelkomponenten, so erhält man eine Ableitung für G , also $L(\bar{G}) \subseteq L(G)$.

Für das abgeleitete Wort $t_1 t_2 \dots t_r \in L(G)$ muß gelten $\delta(s_{i-1}, t_i) = s_i$, $i = 1, \dots, r$ mit Anfangszustand s_0 und $s_r \in F$, also $L(\bar{G}) \subseteq L(A)$.

$L(\bar{G}) \supseteq L(G) \cap L(A)$, weil:

Wenn eine Ableitung $u \in L(G)$ vorliegt und eine Überföhrungsfunktion $\delta(s_{i-1}, t_i) = s_i$, $i = 1, \dots, r$, $u = t_1 \dots t_r$, dann erhält man eine Ableitung in $\bar{G}$, indem man die Zustände geeignet an die Nonterminals links und rechts "anklebt". $\square$

Bemerkung

CF ist nicht abgeschlossen gegen Durchschnitt und Komplement.

Begründung:

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

Dann gilt $L \cap L^c = \{a^n b^n c^n \mid n \geq 0\}$.

Mit Hilfe des Pumping-Lemmas kann man zeigen, daß $L \cap L^c$ nicht kontextfrei ist (aber kontextsensitiv).

Nach den deMorganschen Regeln gilt $L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$. Da CF gegen Vereinigung abgeschlossen ist, würde aus dem Abschluß gegen Komplement der Abschluß gegen Durchschnitt folgen.

Folgerung K

(Weitere Vereinfachung des Satzes von Chomsky-Schützenberger durch Reduzierung der Anzahl der Klammerpaare auf 2)

$L \subseteq L^*$ ist kontextfrei $\Leftrightarrow$ Es existiert ein Alphabet T^c , eine reguläre Menge $R \subseteq T^{c*}$ sowie zwei Homomorphismen $h: T^{c*} \rightarrow T^*$ und $g: T^{c*} \rightarrow T_2^*$ mit $L = h(R \cap g^{-1}(D_2))$.
(∇D_n)

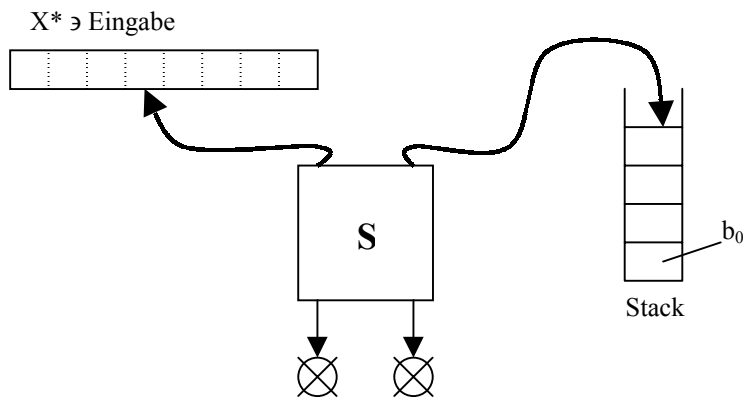
Beweis:

Für $g: T^{c*} \rightarrow T_2^*$ mit $g(a_i) = a_1 a_2^i$, $g(b_i) = b_2^i b_1$ gilt $D_n = g^{-1}(D_2)$, weil:

- $D_n \subseteq g^{-1}(D_2)$, da $g(D_n) \subseteq D_2$
- $D_n \supseteq g^{-1}(D_2)$, da jedes $u \in g^{-1}(D_2)$ nur auf Klammerungen der $a_1 a_2^i b_2^i b_1$ abgebildet werden kann, also korrespondierende Klammern $a_i b_i$ enthält und folglich $u \in D_n$ gilt.

5.4. Pushdown-Akzeptoren (Kellerautomaten, PDA)

Visualisierung:



Definition L

Einen **deterministischen Pushdown-Akzeptor (DPDA)** beschreibt man durch ein 7-Tupel $\rho = (X, \Gamma, S, \delta, s_0, b_0, F)$ mit:

- (i) X (Eingabealphabet), Γ (Kelleralphabet), S (Zustandsmenge) sind endliche nichtleere Mengen und $|\Gamma| \geq 2$
- (ii) $s_0 \in S$ (Anfangszustand), $b_0 \in \Gamma$ (bottom-Element), $F \subseteq S$ (Endzustände)
- (iii) $\delta: S \times (X \cup \{\epsilon\}) \times \Gamma \rightarrow S \times \Gamma^*$ Zustandsübergangsfunktion mit
 - 1) aus $\delta(s, a, b_0) = (s', u)$ folgt $u = b_0 u'$ und b_0 kommt in u' nicht vor
 - 2) aus $\delta(s, a, c) = (s', u)$ und $c \neq b_0$ folgt b_0 kommt nicht in u vor
 - 3) wenn $\delta(s, \epsilon, c)$ definiert ist, dann ist $\delta(s, a, c)$ für alle $a \in X$ undefiniert

Falls $\delta: S \times (X \cup \{\epsilon\}) \times \Gamma \rightarrow 2^{S \times \Gamma^*}$ mit

- 1) $\delta(s, a, c)$ ist für alle $s \in S, a \in (X \cup \{\epsilon\})$ eine endliche Menge
- 2) aus $(s', u) \in \delta(s, a, b_0)$ folgt $u = b_0 u'$ und b_0 kommt in u' nicht vor
- 3) aus $(s', u) \in \delta(s, a, c)$ mit $c \neq b_0$ folgt b_0 kommt in u nicht vor

dann heißt ρ **nichtdeterministischer Pushdown-Akzeptor (NPDA)**.

Definition M

Sei $\rho = (X, \Gamma, S, \delta, s_0, b_0, F)$ ein NPDA. Dann definiert ρ einen abstrakten Automaten durch:

- (i) $K = S \times X^* \times (\{b_0\} \cdot (\Gamma \setminus \{b_0\})^*)$ heißt Konfigurationsmenge von ρ .
- (ii) $\alpha: X^* \rightarrow K$ mit $\alpha(w) = (s_0, w, b_0), \forall w \in X^*$
- (iii) $\hat{\delta}: K \rightarrow 2^K$ sei definiert durch:
 $\forall (s, w, vc) \in K, c \in \Gamma$ gelte: $\hat{\delta}(s, w, vc) = \{(s', w, v\bar{v}) \mid \exists \delta(s, \epsilon, c) = (s', \bar{v})\}$
 $\cup \{(s', w', v\bar{v}) \mid w = aw' \text{ mit } a \in X \text{ und } \exists \delta(s, a, c) = (s', \bar{v})\}$
(Überführungsrelation)

- (iv) die von ρ **akzeptierte Sprache** lautet:

$$L(\rho) = \{w \in X^* \mid \exists k \in \mathbb{N}_0, v \in \Gamma^*, s \in F \text{ mit } (s, \epsilon, v) \in \hat{\delta}^k(\alpha(w))\}.$$

(ρ akzeptiert alle Wörter, wenn sie vollständig gelesen sind und ein Endzustand erreicht ist. Der verbleibende Kellerinhalt wird nicht beachtet. Es gibt jedoch immer einen äquivalenten NPDA, der seinen Keller leert, bevor er terminiert.)

- (v) L_{NPDA} (L_{DPDA}) ist die Klasse aller Sprachen, die durch nichtdeterministische (deterministische) PDAs erkannt werden.