

Abschlusseigenschaften kontextfreier Sprachen

Satz J :

Die kontextfreien Sprachen sind abgeschlossen gegen:

(i) Vereinigung

Beweisidee: G_1 und G_2 seien Grammatiken mit den Startsymbolen S_1 und S_2 ;
bilde nun die Vereinigungsgrammatik mit dem Startsymbol S , das diesen
Regeln folgt: $S \rightarrow S_1, S \rightarrow S_2$

(ii) Konkatenation

Beweisidee: G_1 und G_2 seien Grammatiken mit den Startsymbolen S_1 und S_2 ;
dann bilde die neue Regel $S \rightarrow S_1 S_2$

(iii) Sternbildung

Beweisidee: G sei eine Grammatik mit dem Startsymbol S ; bilde nun die Grammatik G'
mit dem Startsymbol S' und folgenden Regeln: $S' \rightarrow SS'$ und $S' \rightarrow \varepsilon$

(iv) Homomorphismen

(v) Durchschnitt und reguläre Mengen

Beweisidee: Parallele Simulation eines Akzeptors A mit G , wobei als Beweistechnik der
Übergang zum kartesischen Produkt der beiden Beschreibungsmechanismen
genutzt werden soll

Beweis: Sei L eine kontextfreie Sprache und $G = (N, T, P, \tilde{\sigma})$ eine kontextfreie
Grammatik mit $L = L(G)$.

O.B.d.A. sei G in Chomsky-Normalform $[A \rightarrow BB', A \rightarrow t]$.

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

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

$$\bar{N} = \{\bar{\sigma}\} \cup (S \times N \times S) \text{ und}$$

$$\bar{P} = \{\bar{\sigma} \rightarrow (s_0, \sigma, s) \mid s \in F\} \cup$$

$$\{(s_1, B_1, s_2) \rightarrow (s_1, B_1, s_2)(s_3, B_2, s_2) \mid \text{für alle } s_1, s_2 \in S \text{ und für alle}$$

$$\text{Regeln } A \rightarrow B_1 B_2 \in P, A, B_1, B_2 \in N\} \cup$$

$$\{(s_1, A, s_2) \rightarrow t \mid \text{für alle } A \rightarrow t \in P \text{ und für alle } s_1, s_2 \in S \text{ mit}$$

$$\delta(s_1, t) = s_2\}.$$

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

Es folgt die Herleitung $L(\bar{G}) = L(G) \cap L(A)$:

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

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

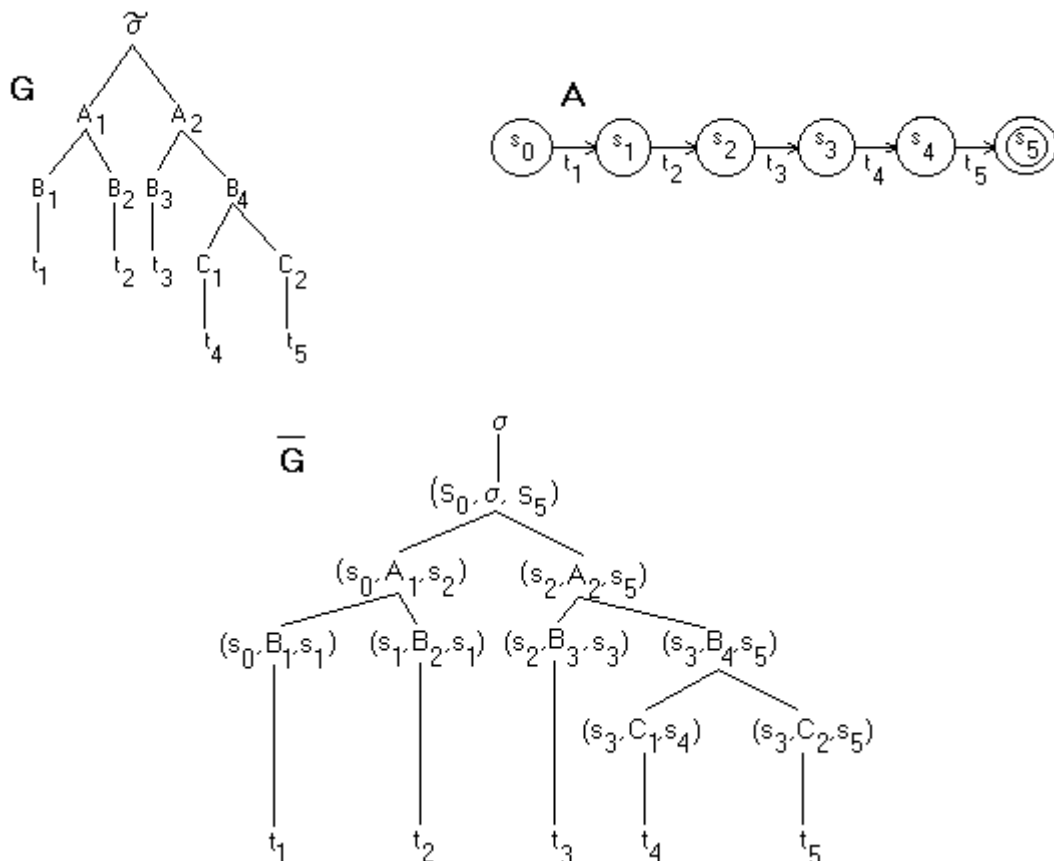
Für das abgeleitete Wort $t_1 t_2 \dots t_r \in L(\bar{G})$ muss gelten:

$$\delta(s_{i-1}, t_i) = s_i \text{ mit Anfangszustand } s_0 \text{ und } s_r \in F, i = 1, \dots, r, \text{ also } L(\bar{G}) \subseteq L(A).$$

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

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

Beispiel: $\tilde{\sigma} \rightarrow' t_1 t_2 t_3 t_4 t_5 \in L(A) \cap L(G)$



Bemerkung:

Die Klasse der kontextfreien Sprachen CF ist nicht abgeschlossen gegen Durchschnitt und Komplement.

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

Dann kann per Pumping-Lemma gezeigt werden, dass $L \cap L' = \{a^n b^n c^n \mid n \geq 0\}$ nicht kontextfrei - $L \cap L'$ also kontextsensitiv ist.

De'Morgansche Regel :

$$L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$$

Da CF abgeschlossen ist gegen Vereinigung, würde aus dem Abschluss gegen Komplement auch der Abschluss gegen Durchschnitt folgen.

Weitere Vereinfachung bei Chomsky-Schützenberger:

a_i, b_i mit $i = 1, \dots, n$; $a_i = [(\dots ($; $b_i =)]^i$

Folgerung K:

$L \subseteq T^*$ kontextfrei $\Leftrightarrow$ Es existiert ein Alphabet T' , eine reguläre Menge $R \subseteq T'^*$ sowie zwei Homomorphismen $h: T'^* \rightarrow T^*$ und $g: T'^* \rightarrow T_2^*$ mit $L = h(R \cap g^{-1}(D_2))$.

Beweis: Für $g: T'^* \rightarrow T_2^*$ und $g(a_i) = a_1 a_2^i$ und $g(b_i) = b_1 b_2^i$ gilt: $D_n = g^{-1}(D_2)$.

„ $\subseteq$ “: klar, da $g(D_n) \subseteq D_2$.

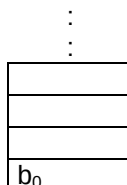
„ $\supseteq$ “: klar, da jedes $u \in g^{-1}(D_2)$ nur auf Klammerungen der Form $\dots a_1 a_2^i b_2^j b_1$ ab gebildet werden kann, also korrespondierende Klammern $\dots a_i b_i \dots$ enthält. Folglich gilt: $u \in D_n$.

□

5.4 PUSHDOWN – AKZEPTOREN

($\hat{=}$ Kellerautomaten, kurz: PDA)

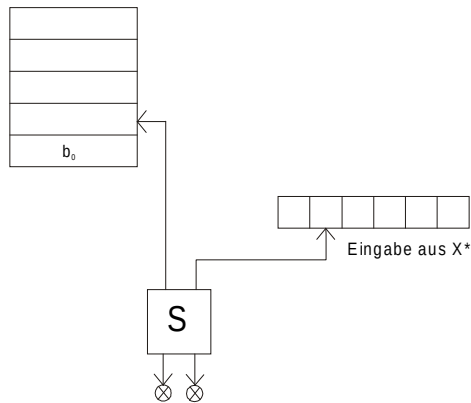
$\Rightarrow$ Ergänzung eines endlichen Akzeptors um einen zusätzlichen Speicher in Stackform



b_0 : bottom-Element
 Es symbolisiert den leeren Keller, wenn es das oberste Element ist.

Stackalphabet: Γ

Eingabe: ϵ ist immer möglich; der PDA muss nicht immer lesen



Definition:

Einen deterministischen pushdown-Akzeptor (DPDA) beschreibt man durch ein 7-Tupel

$\rho = (X, \Gamma, S, \delta, s_0, b_0, F)$ mit

- (i) X als Eingabealphabet,
 Γ als Kelleralphabet (mit $|\Gamma| \geq 2$),
 S als Zustandsmenge
 sind **endliche nichtleere Mengen**.
- (ii) $s_0 \in S$ Anfangszustand,
 $b_0 \in \Gamma$ Bottom-Element,
 $F \subseteq S$ Endzustandsmenge
- (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) $\delta(s, a, c) = (s', u)$ folgt: aus $c \neq b_0$ folgt: b_0 kommt in u' nicht 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 $s \in S$, $a \in (X \cup \{\epsilon\})$, $c \in \Gamma$ 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, dass b_0 nicht in u vorkommt.

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

Arbeitsweise der PDAs in der Sprache der abstrakten Automaten

Definition M:

Sei $p = (X, \Gamma, S, \delta, s_0, b_0, F)$ ein **NPDA**. Dann gilt:

- (i) $K = S \times X^* \times (\{b_0\} \cdot (\Gamma \setminus \{b_0\})^*)$ heißt **Konfigurationsmenge** von p .
- (ii) $\alpha: X^* \rightarrow K$ mit $\alpha(w) = (s_0, w, b_0)$, $\forall w \in X^*$.
- (iii) $\hat{\delta}: X^* \rightarrow 2^K$ 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 p akzeptierte Sprache lautet :**

$$L(p) = \{ w \in X^* \mid \exists k \in N_0, v \in \Gamma^*, s \in F \text{ mit } (s, \epsilon, v) \in \hat{\delta}^k(\alpha(w)) \}$$

[p akzeptiert Wörter, wenn sie vollständig eingelesen sind und ein Endzustand erreicht ist. Der verbleibende Kellerinhalt v interessiert nicht. Später arbeiten wir darauf hin, dass beim Akzeptieren der Keller leer ist.]
- (v) L_{NPDA} (L_{PDA}) ist die Klasse aller Sprachen, die durch nichtdeterministische (deterministische) PDAs erkannt wird.

Beispiel:

