Grammatik alternativer Sprachspezifikationsmechanismus

$G = (\Sigma, V, S, P)$

- Σ Terminalalphabet
- V Variablen- (Nicht-terminal) –alphabet
- $S \in V$ Startvariable
- $P \subset F V F \times F$ “Produktionen” ($F = (\Sigma \cup V)^*$)

Σ, V, S, P müssen endlich sein
$(\alpha,\beta) \in P$ wird geschrieben als $\alpha \to \beta$

Grammatik alternativer Sprachspezifikationsmechanismus

$G = (\Sigma, V, S, P)$

- Σ Terminalalphabet
- V Variablen- (Nicht-terminal) –alphabet
- $S \in V$ Startvariable
- $P \subset F V F \times F$ “Produktionen” ($F = (\Sigma \cup V)^*$)

Σ, V, S, P müssen endlich sein
$(\alpha,\beta) \in P$ wird geschrieben als $\alpha \to \beta$

G induziert Ableitungsschritt-Relation (Derivationsschritt-Relation)
$\Rightarrow_G$ auf F durch $\gamma\alpha\gamma' \Rightarrow_G \gamma\beta\gamma'$ wenn $\alpha \to \beta$ Produktion in P

(also, Teilstring α kann durch β ersetzt werden)

$\Rightarrow_G^*$ reflexive, transitive Hülle von $\Rightarrow_G$: Ableitungsrelation (Derivationsrelation) auf F

G induziert Ableitungsschritt-Relation (Derivationsschritt-Relation)
$\Rightarrow_G$ auf F durch $\gamma\alpha\gamma' \Rightarrow_G \gamma\beta\gamma'$ wenn $\alpha \to \beta$ Produktion in P

$\Rightarrow_G^*$ reflexive, transitive Hülle von $\Rightarrow_G$: Ableitungsrelation (Derivationsrelation) auf F

Die von G generierte Sprache

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

Beispiel: $G = (\Sigma, V, S, P)$ mit $\Sigma = \{a,b\}$, $V = \{S,A,B,C\}$ und

$P = \{ S \to SABC,\ S \to \varepsilon,$
$SA \to a,\ BA \to AB,\ CA \to AC,\ CB \to BC,$
$aA \to aa,\ aB \to ab,\ bB \to bb,\ bC \to bc,\ cC \to cc \}$

Beispielableitung:

$\underline{S} \Rightarrow_G \underline{S}ABC \Rightarrow_G \underline{SA}BCABC \Rightarrow_G aB\underline{CA}BC \Rightarrow_G a\underline{BA}CBC \Rightarrow_G$

$\underline{a}\underline{A}BCBC \Rightarrow_G aaB\underline{CB}C \Rightarrow_G a\underline{aB}BCC \Rightarrow_G$

$a\underline{ab}\underline{B}CC \Rightarrow_G aa\underline{b}\underline{C}C \Rightarrow_G aab\underline{c}\underline{C} \Rightarrow_G aabbcc$

Beispiel: $G = (\Sigma, V, S, P)$ mit $\Sigma=\{a,b\}$, $V=\{S,A,B,C\}$ und

$$P = \{\ S\to SABC\ ,\ S\to\varepsilon\ ,\ SA\to a\ ,\ BA\to AB\ ,\ CA\to AC\ ,\ CB\to BC\ ,\ aA\to aa\ ,\ aB\to ab\ ,\ bB\to bb\ ,\ bC\to bc\ ,\ cC\to cc\ \}$$

Behauptung: $L(G) = \{\ a^n b^n c^n \mid n\in\mathbb{N}\ \}$

Beweis: "$\subseteq$"

Zeige dass nach jedem Ableitungsschritt folgende drei Aussagen für den abgeleiteten String u gelten:

(i) $\#_a(u)+\#_A(u) = \#_b(u)+\#_B(u) = \#_c(u)+\#_C(u)$

(ii) Die Kleinbuchstaben in u bilden einen Präfix von u

(iii) Kein b kommt vor einem a, kein c kommt vor einem b

Wenn nur mehr Kleinbuchstaben übrig sind, hat u die gewünschte Form $a^n b^n c^n$

Beweis: "$\supseteq$"

Induktion

Beispiel: $G = (\Sigma, V, S, P)$ mit $\Sigma=\{a,b\}$, $V=\{S,L,R\}$ und

$$P = \{\ S\to aLbc\ ,\ S\to\varepsilon\ ,\ aL\to aabR\ ,\ Rb\to bR\ ,\ Rc\to Lcc\ ,\ bL\to Lb\ ,\ L\to\varepsilon\ \}$$

Beispielableitung:

$\underline{S} \Rightarrow_G a\underline{L}bc \Rightarrow_G aab\underline{R}\underline{b}c \Rightarrow_G aabb\underline{R}\underline{c} \Rightarrow_G aab\underline{b}\underline{L}cc \Rightarrow_G$

$aa\underline{b}\underline{L}bcc \Rightarrow_G a\underline{a}\underline{L}bbcc \Rightarrow_G aaab\underline{R}\underline{b}bcc \Rightarrow_G$

$aaabb\underline{R}\underline{b}cc \Rightarrow_G aaabbb\underline{R}\underline{c}c \Rightarrow_G aaabbb\underline{L}ccc \Rightarrow_G aaabbbccc$

Beispiel: $G = (\Sigma, V, S, P)$ mit $\Sigma=\{a,b\}$, $V=\{S,L,R\}$ und

$$P = \{\ S\to aLbc\ ,\ S\to\varepsilon\ ,\ aL\to aabR\ ,\ Rb\to bR\ ,\ Rc\to Lcc\ ,\ bL\to Lb\ ,\ L\to\varepsilon\ \}$$

Behauptung: $L(G) = \{\ a^n b^n c^n \mid n\in\mathbb{N}\ \}$

Beweis: "$\subseteq$"

Jeder aus S abgeleitete String hat einer der folgenden 4 Formen für irgendein $n>0$ und irgendein $0\le i\le n$:

ε , $a^n b^i R b^{n-i} c^{n-1}$, $a^n b^i L b^{n-i} c^n$, $a^n b^n c^n$

(Anwendung von Regeln aus P auf einen String aus diesen 4 Klassen ergibt immer einen String aus diesen 4 Klassen.)

Beweis: "$\supseteq$"

Zeige durch Induktion, dass aus S jeder String der Form $a^n L b^n c^n$ abgeleitet werden kann (und daher auch $a^n b^n c^n$).

Chomsky Hierarchie für Grammatiken und Sprachen

Typ 0 (unbeschränkt)

Typ 1 (kontextsensitiv)
nur Regeln $\alpha\to\beta$ mit $|\alpha|\le|\beta|$
(außer $S\to\varepsilon$, aber dann S auf keiner rechten Seite einer Regel)

Typ 2 (kontextfrei)
nur Regeln $A\to\alpha$ mit $A\in V$

Typ 3 (rechtslinear)
nur Regeln $A\to uB$, $A\to u$, $A\to\varepsilon$ mit $A,B\in V$ und $u\in\Sigma$

Chomsky Hierarchie für Grammatiken und Sprachen

Typ 0 (unbeschränkt)

Typ 1 (kontextsensitiv)
nur Regeln $\alpha \rightarrow \beta$ mit $|\alpha| \leq |\beta|$
(außer $S \rightarrow \varepsilon$, aber dann S auf keiner rechten Seite einer Regel)

Typ 2 (kontextfrei)
nur Regeln $A \rightarrow \alpha$ mit $A \in V$

Typ 3 (rechtslinear)
nur Regeln $A \rightarrow uB$, $A \rightarrow u$, $A \rightarrow \varepsilon$ mit $A,B \in V$ und $u \in \Sigma$

Anmerkung: Bei kontextfreien und rechtslinearen Grammatiken können Regeln der Form $A \rightarrow \varepsilon$ zugelassen werden, weil sie leicht entfernt werden können (abgesehen von $S \rightarrow \varepsilon$)

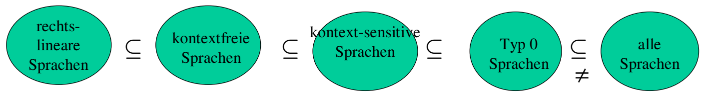
Chomsky Hierarchie für Grammatiken und Sprachen

Typ 0 (unbeschränkt)

Typ 1 (kontextsensitiv)
nur Regeln $\alpha \rightarrow \beta$ mit $|\alpha| \leq |\beta|$
(außer $S \rightarrow \varepsilon$, aber dann S auf keiner rechten Seite einer Regel)

Typ 2 (kontextfrei)
nur Regeln $A \rightarrow \alpha$ mit $A \in V$

Typ 3 (rechtslinear)
nur Regeln $A \rightarrow uB$, $A \rightarrow u$, $A \rightarrow \varepsilon$ mit $A,B \in V$ und $u \in \Sigma$



Satz: Die rechtslinearen Sprachen sind genau die regulären Sprachen.

Beweis:
1) L rechtslinear $\Rightarrow$ L regulär
Idee: zeige, dass L nur endlich viele Fortsetzungssprachen hat

2) L regulär $\Rightarrow$ L rechtslinear

L regulär $\Rightarrow$ $L=L(M)$ für DEA $M=(\Sigma,Q,s,F,\Delta)$

Betrachte Grammatik $G = (\Sigma,Q,s,P)$ mit

$P = \{ p \rightarrow uq \mid (p,u,q) \in \Delta \} \cup \{ p \rightarrow \varepsilon \mid p \in F \}$

Kontextfreie Grammatiken und Sprachen

Bsp: kfG für geklammerte arithmetische Ausdrücke über Binärzahlen und Variablennamen über $\{a,b\}^*$

$G = (\Sigma,V,E,P)$ mit $\Sigma = \{a,b,0,1,(,),*,+\}$
$V=\{E,K,W\}$
und Produktionen P: $E \rightarrow E+E \mid E*E \mid (E) \mid K \mid W$
$W \rightarrow aW \mid bW \mid a \mid b$
$K \rightarrow 0K \mid 1K \mid 0 \mid 1$

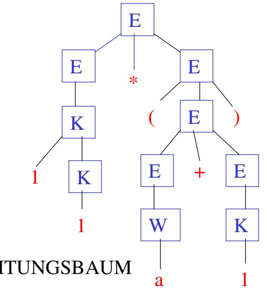
$11*(a+1) \in L(G)$ $\quad$ $(a+b)(a+1) \notin L(G)$

$G = (\Sigma, V, E, P)$ mit $\Sigma = \{a,b,0,1,(,),*,+\}$
$V = \{E,K,W\}$
und Produktionen P: $E \rightarrow E+E \mid E*E \mid (E) \mid K \mid W$
$W \rightarrow aW \mid bW \mid a \mid b$
$K \rightarrow 0K \mid 1K \mid 0 \mid 1$

$\underline{E} \Rightarrow E*\underline{E} \Rightarrow E*(\underline{E}) \Rightarrow \underline{E}*(E+E)$
$\Rightarrow K*(\underline{E}+E) \Rightarrow K*(W+\underline{E})$
$\Rightarrow \underline{K}*(W+K) \Rightarrow 1\underline{K}*(W+K)$
$\Rightarrow 11*(W+\underline{K}) \Rightarrow 11*(\underline{W}+1)$
$\Rightarrow 11*(a+1)$

$G = (\Sigma, V, E, P)$ mit $\Sigma = \{a,b,0,1,(,),*,+\}$
$V = \{E,K,W\}$
und Produktionen P: $E \rightarrow E+E \mid E*E \mid (E) \mid K \mid W$
$W \rightarrow aW \mid bW \mid a \mid b$
$K \rightarrow 0K \mid 1K \mid 0 \mid 1$

$\underline{E} \Rightarrow E*\underline{E} \Rightarrow E*(\underline{E}) \Rightarrow \underline{E}*(E+E)$
$\Rightarrow K*(\underline{E}+E) \Rightarrow K*(W+\underline{E})$
$\Rightarrow \underline{K}*(W+K) \Rightarrow 1\underline{K}*(W+K)$
$\Rightarrow 11*(W+\underline{K}) \Rightarrow 11*(\underline{W}+1)$
$\Rightarrow 11*(a+1)$



ABLEITUNGSBAUM

Ableitungsbaum:

geordneter Baum mit Knotenbeschriftung aus $\Sigma \cup V \cup \{\varepsilon\}$

Knoten N beschriftet mit $x \in \Sigma \cup \{\varepsilon\}$ hat keine Kinder

Knoten N beschriftet mit $y \in V$ und mit Kindern $N_1,\cdots,N_k$
beschriftet mit $x_1,\cdots,x_k$
nur möglich, wenn $y \rightarrow x_1 x_2 \cdots x_k$ eine Produktionsregel

Wurzelbeschriftung A und Blätterbeschriftung $\alpha_1\alpha_2\cdots\alpha_m$ genau dann möglich wenn $A \Rightarrow^* \alpha_1\alpha_2\cdots\alpha_m$