

Parsing

Beispiel: Transformation in CNF

Kurt Eberle

k.eberle@lingenio.de

03. August 2018

Outline

Chomsky-Normal-Form

Beispiel-Grammatik

Transformation

Outline

Chomsky-Normal-Form

Beispiel-Grammatik

Transformation

Chomsky-Normal-Form

Grammatik $G = (V, \Sigma, P, S)$ ist in CNF gdw G hat die Form

- ▶ $A \rightarrow BC$
- ▶ $A \rightarrow a$
- ▶ ($S \rightarrow \epsilon$; in diesem Falle erscheint S nicht auf der rechten Seite einer Regel)

für $A, B, C \in V, a \in \Sigma$.

Outline

Chomsky-Normal-Form

Beispiel-Grammatik

Transformation

Grammatik

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

- ▶ $V = \{S, A, B, C, D, E, F, G, H, I\}$
- ▶ $\Sigma = \{c, d, e, f, g, h, i\}$
- ▶ und P :

$S \rightarrow ABC$	$D \rightarrow \varepsilon$
$S \rightarrow \varepsilon$	$E \rightarrow deE$
$A \rightarrow DEF$	$E \rightarrow \varepsilon$
$B \rightarrow GHI$	$F \rightarrow efF$
$B \rightarrow \varepsilon$	$F \rightarrow \varepsilon$
$C \rightarrow cC$	$G \rightarrow g$
$C \rightarrow \varepsilon$	$H \rightarrow gh$
$D \rightarrow cdD$	$I \rightarrow hi$

Outline

Chomsky-Normal-Form

Beispiel-Grammatik

Transformation

1. Neue einstellige Regeln

- ▶ $\Sigma = \{c, d, e, f, g, h, i\}$

- ▶ einstellige Regeln

$$X_c \rightarrow c \quad X_g \rightarrow g$$

$$X_d \rightarrow d \quad X_h \rightarrow h$$

$$X_e \rightarrow e \quad X_i \rightarrow i$$

$$X_f \rightarrow f$$

- ▶ Ersetzung

$$S \rightarrow ABC$$

$$D \rightarrow \varepsilon$$

$$S \rightarrow \varepsilon$$

$$E \rightarrow X_d X_e E$$

$$A \rightarrow DEF$$

$$E \rightarrow \varepsilon$$

$$B \rightarrow GHI$$

$$F \rightarrow X_e X_f F$$

$$B \rightarrow \varepsilon$$

$$F \rightarrow \varepsilon$$

$$C \rightarrow X_c C$$

$$G \rightarrow X_g$$

$$C \rightarrow \varepsilon$$

$$H \rightarrow X_g X_h$$

$$D \rightarrow X_c X_d D$$

$$I \rightarrow X_h X_i$$

2. Neue Symbolpaar-Regeln

► Neue Paar-Regeln

$$X_{AB} \rightarrow AB \quad X_{cd} \rightarrow X_c X_d$$

$$X_{DE} \rightarrow DE \quad X_{de} \rightarrow X_d X_e$$

$$X_{GH} \rightarrow GH \quad X_{ef} \rightarrow X_e X_f$$

► Ersetzung

$$S \rightarrow X_{ABC} \quad D \rightarrow \varepsilon$$

$$S \rightarrow \varepsilon \quad E \rightarrow X_{de} E$$

$$A \rightarrow X_{DEF} \quad E \rightarrow \varepsilon$$

$$B \rightarrow X_{GHI} \quad F \rightarrow X_{ef} F$$

$$B \rightarrow \varepsilon \quad F \rightarrow \varepsilon$$

$$C \rightarrow X_c C \quad G \rightarrow X_g$$

$$C \rightarrow \varepsilon \quad H \rightarrow X_g X_h$$

$$D \rightarrow X_{cd} D \quad I \rightarrow X_h X_i$$

3. S- ϵ -Regel

- ▶ Neue Regeln

$$S' \rightarrow S \quad S' \rightarrow \epsilon$$

- ▶ Ergebnis: (+ Neue Regeln aus Schritt 1 und 2)

$$S' \rightarrow S \quad D \rightarrow X_{cd}D$$

$$S' \rightarrow \epsilon \quad D \rightarrow \epsilon$$

$$S \rightarrow X_{ABC} \quad E \rightarrow X_{de}E$$

$$S \rightarrow \epsilon \quad E \rightarrow \epsilon$$

$$A \rightarrow X_{DEF} \quad F \rightarrow X_{ef}F$$

$$B \rightarrow X_{GHI} \quad F \rightarrow \epsilon$$

$$B \rightarrow \epsilon \quad G \rightarrow X_g$$

$$C \rightarrow X_c C \quad H \rightarrow X_g X_h$$

$$C \rightarrow \epsilon \quad I \rightarrow X_h X_i$$

4. Andere (nicht S') ϵ -Regeln (erster Durchgang)

- Tilge Regeln $S \rightarrow \epsilon$ $D \rightarrow \epsilon$
 $B \rightarrow \epsilon$ $E \rightarrow \epsilon$
 $C \rightarrow \epsilon$ $F \rightarrow \epsilon$

- Füge Regeln hinzu (für B, C, D, E, F)
- | | |
|------------------------|-------------------------------|
| $X_{AB} \rightarrow A$ | $X_{DE} \rightarrow D$ |
| $C \rightarrow X_c$ | $X_{DE} \rightarrow \epsilon$ |
| $S \rightarrow X_{AB}$ | $E \rightarrow X_{de}$ |
| $X_{DE} \rightarrow E$ | $F \rightarrow X_{ef}$ |
| $D \rightarrow X_{cd}$ | $A \rightarrow X_{DE}$ |

- Ergebnis: (+ Terminal-Regeln)

$X_{AB} \rightarrow AB$	$S' \rightarrow S$	$C \rightarrow X_c C$	
$X_{AB} \rightarrow A$	$S' \rightarrow \epsilon$	$C \rightarrow X_c$	$F \rightarrow X_{ef} F$
$X_{DE} \rightarrow DE$	$S \rightarrow X_{AB} C$	$D \rightarrow X_{cd} D$	$F \rightarrow X_{ef}$
$X_{DE} \rightarrow D$	$S \rightarrow X_{AB}$	$D \rightarrow X_{cd}$	$G \rightarrow X_g$
$X_{DE} \rightarrow E$	$A \rightarrow X_{DE} F$	$E \rightarrow X_{de} E$	$H \rightarrow X_g X_h$
$X_{DE} \rightarrow \epsilon$	$A \rightarrow X_{DE}$	$E \rightarrow X_{de}$	$I \rightarrow X_h X_i$
$X_{GH} \rightarrow GH$	$B \rightarrow X_{GH} I$		

4. Nicht-S'- ϵ -Regeln (zweiter Durchgang)

- ▶ Tilge Regeln $X_{DE} \rightarrow \epsilon$
- ▶ Füge Regeln hinzu (für X_{DE}) $A \rightarrow F$ $A \rightarrow \epsilon$
- ▶ Ergebnis: (+ Terminal-Regeln)

$X_{AB} \rightarrow AB$	$S' \rightarrow S$	$B \rightarrow X_{GH}I$	
$X_{AB} \rightarrow A$	$S' \rightarrow \epsilon$	$C \rightarrow X_c C$	$F \rightarrow X_{ef} F$
$X_{DE} \rightarrow DE$	$S \rightarrow X_{AB} C$	$C \rightarrow X_c$	$F \rightarrow X_{ef}$
$X_{DE} \rightarrow D$	$S \rightarrow X_{AB}$	$D \rightarrow X_{cd} D$	$G \rightarrow X_g$
$X_{DE} \rightarrow E$	$A \rightarrow X_{DE} F$	$D \rightarrow X_{cd}$	$H \rightarrow X_g X_h$
$X_{GH} \rightarrow GH$	$A \rightarrow F$	$E \rightarrow X_{de} E$	$I \rightarrow X_h X_i$
	$A \rightarrow X_{DE}$	$E \rightarrow X_{de}$	
	$A \rightarrow \epsilon$		

4. Nicht-S'- ϵ -Regeln (dritter Durchgang)

- ▶ Tilge Regeln $A \rightarrow \epsilon$
- ▶ Füge Regeln hinzu (für A) $X_{AB} \rightarrow B$ $X_{AB} \rightarrow \epsilon$
- ▶ Ergebnis: (+ Terminal-Regeln)

$$X_{AB} \rightarrow AB$$

$$X_{AB} \rightarrow A$$

$$X_{AB} \rightarrow B$$

$$X_{AB} \rightarrow \epsilon$$

$$X_{DE} \rightarrow DE$$

$$X_{DE} \rightarrow D$$

$$X_{DE} \rightarrow E$$

$$X_{GH} \rightarrow GH$$

$$S' \rightarrow S$$

$$S' \rightarrow \epsilon$$

$$S \rightarrow X_{AB}C$$

$$S \rightarrow X_{AB}$$

$$A \rightarrow X_{DEF}$$

$$A \rightarrow F$$

$$A \rightarrow X_{DE}$$

$$B \rightarrow X_{GHI}$$

$$C \rightarrow X_c C$$

$$C \rightarrow X_c$$

$$D \rightarrow X_{cd}D$$

$$D \rightarrow X_{cd}$$

$$E \rightarrow X_{de}E$$

$$E \rightarrow X_{de}$$

$$F \rightarrow X_{ef}F$$

$$F \rightarrow X_{ef}$$

$$G \rightarrow X_g$$

$$H \rightarrow X_g X_h$$

$$I \rightarrow X_h X_i$$

4. Nicht-S'- ϵ -Regeln (vierter Durchgang)

- ▶ Tilge Regeln $X_{AB} \rightarrow \epsilon$
- ▶ Füge Regeln hinzu (für X_{AB}) $S \rightarrow C \quad S \rightarrow \epsilon$
- ▶ Ergebnis: (+ Terminal-Regeln)

	$S' \rightarrow S$		
$X_{AB} \rightarrow AB$	$S' \rightarrow \epsilon$	$B \rightarrow X_{GH}I$	
$X_{AB} \rightarrow A$	$S \rightarrow X_{AB}C$	$C \rightarrow X_c C$	$F \rightarrow X_{ef}F$
$X_{AB} \rightarrow B$	$S \rightarrow C$	$C \rightarrow X_c$	$F \rightarrow X_{ef}$
$X_{DE} \rightarrow DE$	$S \rightarrow X_{AB}$	$D \rightarrow X_{cd}D$	$G \rightarrow X_g$
$X_{DE} \rightarrow D$	$S \rightarrow \epsilon$	$D \rightarrow X_{cd}$	$H \rightarrow X_g X_h$
$X_{DE} \rightarrow E$	$A \rightarrow X_{DE}F$	$E \rightarrow X_{de}E$	$I \rightarrow X_h X_i$
$X_{GH} \rightarrow GH$	$A \rightarrow F$	$E \rightarrow X_{de}$	
	$A \rightarrow X_{DE}$		

4. Nicht-S'- ϵ -Regeln (fünfter Durchgang)

- ▶ Tilge Regeln $S \rightarrow \epsilon$
- ▶ Füge Regeln hinzu (für S) \emptyset
- ▶ Ergebnis: (+ Terminal-Regeln)

$X_{AB} \rightarrow AB$	$S' \rightarrow S$	$B \rightarrow X_{GH}I$	
$X_{AB} \rightarrow A$	$S' \rightarrow \epsilon$	$C \rightarrow X_c C$	$F \rightarrow X_{ef} F$
$X_{AB} \rightarrow B$	$S \rightarrow X_{AB} C$	$C \rightarrow X_c$	$F \rightarrow X_{ef}$
$X_{DE} \rightarrow DE$	$S \rightarrow C$	$D \rightarrow X_{cd} D$	$G \rightarrow X_g$
$X_{DE} \rightarrow D$	$S \rightarrow X_{AB}$	$D \rightarrow X_{cd}$	$H \rightarrow X_g X_h$
$X_{DE} \rightarrow E$	$A \rightarrow X_{DE} F$	$E \rightarrow X_{de} E$	$I \rightarrow X_h X_i$
$X_{GH} \rightarrow GH$	$A \rightarrow F$	$E \rightarrow X_{de}$	
	$A \rightarrow X_{DE}$		

5. Kettenregeln

► Ersetze Regeln

$X_{AB} \rightarrow A$ $X_{AB} \rightarrow X_{DE}F$, $X_{AB} \rightarrow X_eX_f$, $X_{AB} \rightarrow X_{ef}F$, $X_{AB} \rightarrow DE$

$X_{AB} \rightarrow B$ $X_{AB} \rightarrow X_{GH}I$

$X_{DE} \rightarrow D$ $X_{DE} \rightarrow X_cX_d$, $X_{DE} \rightarrow X_{cd}D$

$X_{DE} \rightarrow E$ $X_{DE} \rightarrow X_dX_e$, $X_{DE} \rightarrow X_{de}E$

$S' \rightarrow S$ $S' \rightarrow X_{AB}C$, $S' \rightarrow X_cC$, $S' \rightarrow c$, $S' \rightarrow AB$

$S \rightarrow C$ $S \rightarrow X_cC$, $S \rightarrow c$

$S \rightarrow X_{AB}$ $S \rightarrow AB$

$A \rightarrow F$ $A \rightarrow X_eX_f$, $A \rightarrow X_{ef}F$

$A \rightarrow X_{DE}$ $A \rightarrow DE$

$C \rightarrow X_c$ $C \rightarrow c$

$D \rightarrow X_{cd}$ $D \rightarrow X_cX_d$

$E \rightarrow X_{de}$ $E \rightarrow X_dX_e$

$F \rightarrow X_{ef}$ $F \rightarrow X_eX_f$

$G \rightarrow X_g$ $G \rightarrow g$

Ergebnis

$$X_{AB} \rightarrow AB$$

$$X_{AB} \rightarrow X_{DE}F | X_e X_f | X_{ef} F | DE$$

$$X_{AB} \rightarrow X_{GH}I$$

$$X_{DE} \rightarrow DE$$

$$X_{DE} \rightarrow X_c X_d | X_{cd} D$$

$$X_{DE} \rightarrow X_d X_e | X_{de} E$$

$$X_{GH} \rightarrow GH$$

$$S' \rightarrow X_{AB}C | X_c C | c | AB$$

$$S' \rightarrow \varepsilon$$

$$A \rightarrow X_{DE}F$$

$$A \rightarrow X_e X_f | X_{ef} F$$

$$A \rightarrow DE$$

$$B \rightarrow X_{GH}I$$

$$C \rightarrow X_c C$$

$$C \rightarrow c$$

$$D \rightarrow X_{cd} D$$

$$D \rightarrow X_c X_d$$

$$E \rightarrow X_{de} E$$

$$E \rightarrow X_d X_e$$

$$F \rightarrow X_{ef} F$$

$$F \rightarrow X_e X_f$$

$$G \rightarrow g$$

$$H \rightarrow X_g X_h$$

$$I \rightarrow X_h X_i$$