

# 1 Bringen in Chomsky Normalform

Gegeben ist die Grammatik

$$\begin{aligned}G &= (V, \Sigma, P, S) \\V &= \{S, A, B, C, D\} \\ \Sigma &= \{a, b, c\} \\ P &= \{S \rightarrow AB, \\ &\quad A \rightarrow ab|aAbcDa|C, \\ &\quad B \rightarrow c|cB|D, \\ &\quad C \rightarrow A|D|\varepsilon, \\ &\quad D \rightarrow a|c\}\end{aligned}$$

## 1.1 Entfernen von Kettenregeln

Um eine Grammatik in Chomsky Normalform bringen zu können, muss sie als allererstes von Kettenregeln befreit werden.

Hierzu beginnt man damit, alle  $\varepsilon$ -Regeln zu entfernen. Die Menge  $V_1 := \{A \in V \mid A \Rightarrow_G^* \varepsilon\}$  enthält alle Variablen, die sich auf  $\varepsilon$  ableiten lassen können. Bei unserer Grammatik ist das  $V_1 = \{A, C\}$ . Diese  $\varepsilon$ -Regeln, und alle Vorkommen von Variablen die dort hinführen müssen gestrichen werden. Wenn eine Variable allerdings nicht nur zum  $\varepsilon$  führt, müssen neue Regeln eingefügt werden:  $\forall B \rightarrow xAy \wedge A \in V_1 \wedge xy \neq \varepsilon$ : die neue Regel  $B \rightarrow xy$  einfügen. Dies führt zu folgender Grammatik  $G'$

$$\begin{aligned}G' &= (V, \Sigma, P', S) \\V &= \{S, A, B, C, D\} \\ \Sigma &= \{a, b, c\} \\ P' &= \{S \rightarrow AB|B, \\ &\quad A \rightarrow ab|aAbcDa|abcDa|C, \\ &\quad B \rightarrow c|cB|D, \\ &\quad C \rightarrow A|D, \\ &\quad D \rightarrow a|c\}\end{aligned}$$

Nun sind die Kettenregeln an der Reihe. Hierfür muss für jede Variable  $X \in V$  die Äquivalenzklasse bestimmt werden (Variablen die sich ineinander ableiten lassen):

$$\begin{aligned}[A] &= [C] = \{A, C\} \\ [S] &= \{S\} \\ [B] &= \{B\} \\ [D] &= \{D\}\end{aligned}$$

Alle Variablen die in einer gemeinsamen Äquivalenzklasse liegen, haben genau die gleiche Funktion, und können somit zu einer zusammengefasst werden.

$$\begin{aligned}
 G'' &= (V, \Sigma, P'', S) \\
 V &= \{S, A, B, D\} \\
 \Sigma &= \{a, b, c\} \\
 P' &= \{S \rightarrow AB|B, \\
 &\quad A \rightarrow ab|aAbcDa|abcDa|D, \\
 &\quad B \rightarrow c|cB|D, \\
 &\quad D \rightarrow a|c\}
 \end{aligned}$$

Nun müssen können alle Variablen, die nicht sich selbst auf der Rechten Seite enthalten, direkt an den Stellen eingesetzt werden, wo sie auf der Rechten Seite Vorkommen. Und alle Regeln, die auf der Rechten Seite nur eine Variable enthalten, können ebenfalls, direkt ersetzt werden:

$$\begin{aligned}
 G''' &= (V, \Sigma, P''', S) \\
 V &= \{S, A, B, D\} \\
 \Sigma &= \{a, b, c\} \\
 P' &= \{S \rightarrow AB|c|cB|a, \\
 &\quad A \rightarrow ab|aAbcaa|aAcca|abcaa|abcca|a|c, \\
 &\quad B \rightarrow c|cB|a\}
 \end{aligned}$$

## 1.2 Chomsky-Normalisierung

- Zu lange Regeln kürzen
- Terminale durch Var ersetzen