

# Převod BKG do Chomského normální formy (CNF)

## Definice CNF

Bezkontextová gramatika  $G = (N, \Sigma, P, S)$  je v *Chomského normálním tvaru*, pokud každé pravidlo z množiny  $P$  má tvar:

$$A \rightarrow BC \quad \text{nebo} \quad A \rightarrow a,$$

kde  $A, B, C \in N$ ,  $a \in \Sigma$ . Jedinou výjimkou je pravidlo  $S \rightarrow \lambda$ , pokud se  $S$  nevyskytuje na pravé straně žádného pravidla.

## Postup převodu do CNF

### 1. Základní zjednodušení

Nejprve gramatiku převedeme na ekvivalentní tvar bez nepotřebných konstrukcí:

1. odstraníme zbytečné symboly (nedosažitelné a negenerující),
2. odstraníme  $\lambda$ -pravidla (s výjimkou  $S \rightarrow \lambda$ ),
3. odstraníme jednoduchá pravidla typu  $A \rightarrow B$ .

Tím získáme gramatiku, která má pouze netriviální produkce.

### 2. Úprava pravidel do CNF

Následně projdeme všechna pravidla a upravíme je tak, aby splňovala formu CNF:

1. Pokud je pravidlo již tvaru  $A \rightarrow a$  nebo  $A \rightarrow BC$ , ponecháme jej.
2. Pokud se terminál vyskytuje na pravé straně s dalšími symboly, nahradíme terminál novým neterminálem:

$$A \rightarrow aBC \Rightarrow A \rightarrow XBC, \quad X \rightarrow a.$$

3. Pokud má pravá strana více než dva neterminály, rozdělíme ji pomocí nových neterminálů na dvojice:

$$A \rightarrow ABCD \Rightarrow A \rightarrow A_1D, \quad A_1 \rightarrow ABC.$$

nebo obecně postupným rozkladem:

$$A \rightarrow BCD \Rightarrow A \rightarrow XD, \quad X \rightarrow BC.$$

Po těchto úpravách mají všechny produkce formu CNF.

## Výsledná gramatika

Výsledná gramatika je ekvivalentní s původní gramatikou, zachovává jazyk  $L(G)$  a obsahuje pouze pravidla dvou tvarů:

$$A \rightarrow BC, \quad A \rightarrow a.$$

Tento tvar se používá pro algoritmy CYK, formální důkazy vlastností CFL a praktickou analýzu gramatik.