

Bezkontextové gramatiky

Definice – bezkontextová gramatika

Bezkontextová gramatika (zkráceně CFG – Context Free Grammar) je čtveřice (N, Σ, P, S) taková, že:

- N je množina neterminálů;
- Σ je množina terminálů, pro kterou platí $N \cap \Sigma = \emptyset$;
- $P \subseteq N \times (N \cup \Sigma)^*$ je množina pravidel;
- S je startovní (počáteční) neterminál.

Poznámka.

Všimněte si dvou věcí:

1. množiny N , Σ mají prázdný průnik, tj. nemůžeme se stát, že by nějaký symbol byl zároveň neterminál i terminál.
2. Množina pravidel P je definována jako podmnožina kartézského součinu, tj. je to relace mezi množinami N a $(N \cup \Sigma)^*$. Znamená to, že levá strana pravidel je vždy jen jeden neterminál a pravá strana může být cokoliv, dokonce i prázdné slovo ε .

Poznámka.

V bodu 2. předchozí poznámky jste možná postřehli rozpor mezi naší aktuální definicí bezkontextové gramatiky a definicí uvedenou v Chomského hierarchii gramatik. Zde se za bezkontextovou gramatiku považuje taková gramatika $G = (N, \Sigma, P, S)$, kde libovolné pravidlo z P je tvaru $A \rightarrow \beta$, kde

- $A \in N$,
- $\beta \in (N \cup \Sigma)^*$,
- $|\beta| \geq 1$

s eventuální výjimkou $S \rightarrow \varepsilon$, pokud se S nevyskytuje na pravé straně žádného jiného pravidla.

Všimněte si, že v definici uvedené na začátku tohoto textu chybí 3. podmínka $|\beta| \geq 1$, tj. můžeme libovolný neterminál odvodit na prázdné slovo. V kapitole 9 si ukážeme, že tento rozdíl není podstatný – každé pravidlo $A \rightarrow \varepsilon$, kde $A \in N$, dokážeme odstranit a nahradit jinými pravidly tak, aby se jazyk generovaný gramatikou nezměnil.

Příklad 1.

Určete jazyk $L = L(G)$, který generuje gramatika $G = (\{S\}, \{a, b\}, P, S)$, kde $P = \{S \rightarrow aSb \mid aSb \mid ab\}$.

Řešení. Gramatika generuje jazyk $L = \{w \in \{a, b\}^* \mid a^n b^n, n \geq 1\}$.

Příklad 2.

Určete jazyk $L = L(G)$, který generuje gramatika $G = (\{S\}, \{a, b\}, P, S)$, kde $P = \{S \rightarrow aSa \mid bSb \mid aa \mid bb\}$.

Řešení. Gramatika generuje jazyk

$$L = \{abba, aa, bb, abbbba, \dots\} = \{w.w^R, w \in \{a, b\}^+\}.$$

Příklad 3.

Navrhněte bezkontextovou gramatiku G , která akceptuje jazyk

$$L = \{w \in \{a, b\}^*, \#_a(w) = \#_b(w)\}.$$

Řešení. $G = (\{S, T, A, B\}, \{a, b\}, P, S)$, kde

$$\begin{aligned} P = \{ & S \rightarrow \varepsilon \mid abS \mid baS \mid aaSBS \mid bbSAS, \\ & A \rightarrow aa, \\ & B \rightarrow bb \} \end{aligned}$$