3.2.10 Chomského normální tvar. Jedním z důležitých typů CF gramatik, je gramatika v Chomského normálním tvaru. Je to CF gramatika, kde pravá strana pravidla je buď terminální písmeno nebo dvojice neterminálů. Gramatiky v Chomského normálním tvaru budeme využívat jednak pro důkaz Pumping lemmatu pro bezkontextové jazyky, jednak při polynomiálním algoritmu, který pro danou gramatiku a dané terminální slovo rozhodne, zda je slovo generováno.

Definice. Je dána bezkontextová gramatika $\mathcal{G} = (N, \Sigma, S, P)$. Řekneme, že gramatika $\mathcal{G}$ je v *Chomského normálním tvaru*, jestliže má pouze pravidla tvaru

$$A \to BC,\ A \to a \quad \text{kde} \quad A, B, C \in N,\ a \in \Sigma.$$

□

Poznamenejme, že derivační strom terminálního slova v CF gramatice v Chomského normálním tvaru je vždy binární strom a to takový, že předchůdce listu má vždy jen jednoho následníka.

3.2.11 Věta. Pro každou bezkontextovou gramatiku $\mathcal{G}$ existuje bezkontextová gramatika $\mathcal{G}'$ v Chomského normálním tvaru taková, že

$$L(\mathcal{G}') = L(\mathcal{G}) \setminus \{\varepsilon\}.$$

□

Větu dokážeme postupem, který pro danou CF gramatiky $\mathcal{G}$ sestrojí CF gramatiku $\mathcal{G}'$ v Chomského normálním tvaru generující stejný jazyk.

Postup nalezení ekvivalentní gramatiky v Chomského normálním tvaru.

1. Pro gramatiku $\mathcal{G}$ nejprve sestrojíme nevypouštěcí gramatiku $\mathcal{G}_1$, která generuje všechna neprázdná slova z jazyka $L(\mathcal{G})$ (viz 3.1.8).
2. Následně v gramatice $\mathcal{G}_1$ odstraníme pravidla typu $A \to B$, kde $A, B \in N$, takto: Pro každou dvojici neterminálů A, C takovou, že
$$A \Rightarrow X_1 \Rightarrow X_2 \Rightarrow \ldots \Rightarrow C,$$
a pro každé pravidlo $C \to \alpha$ do gramatiky přidáme pravidlo $A \to \alpha$.
3. Dále:
   a) pro každý terminál $a \in \Sigma$ zavedeme nový neterminál X_a a přidáme pravidlo $X_a \to a$,
   b) pro každé pravidlo $A \to \alpha$, kde $|\alpha| \geq 2$, nahradíme v pravé straně pravidla každý terminál b neterminálem X_b.
4. Nyní stačí nahradit pravidla typu
$$A \to B_1 B_2 \ldots B_k,\ k \geq 3$$
pravidly s novými neterminály $C_1, C_2, \ldots, C_{k-2}$:
$$A \to B_1 C_1,\ C_1 \to B_2 C_2,\ \ldots, C_{k-2} \to B_{k-1} B_k.$$

Tím dostaneme hledanou gramatiku $\mathcal{G}'$ v Chomského normálním tvaru.

3.2.12 Následující tvrzení využijeme při zdůvodnění pumping lemmatu pro bezkontextové jazyky.

Tvrzení. Je dána bezkontextová gramatika $\mathcal{G} = (N, \Sigma, S, P)$ v Chomského normálním tvaru a dále derivační strom s výsledkem $w \in \Sigma^\star$.

Jestliže nejdelší orientovaná cesta v tomto derivačním stromu má délku n, pak pro délku slova w platí

$$|w| \leq 2^{n-1}.$$

□

Nástin důkazu: Pro $n = 1$ se jedná o použití jediného pravidla a proto je w rovno jedinému písmenu $w = b$, $b \in \Sigma$.

Předpokládejme, že $n > 1$. Nejdelší výsledek derivačního stromu nastane, když v derivačním stromě mají všechny vrcholy kromě předposlední hladiny vždy dva následníky. Pak ale předposlední hladina má 2^{n-1} vrcholů a tudíž slovo w je délky 2^{n-1}. (Dokažte si indukcí.)