IB102 – úkol 7, příklad 1 – řešení **Odevzdání: 18. 11. 2013**

Vypracoval(a): UČO:
Skupina:

1. [2 body] Uvažme následující bezkontextovou gramatiku:

$$
\begin{aligned}
G &= (\{S, A, B, C, D, E\}, \{x, y, z\}, P, S)\\
P &= \{S \;\rightarrow\; \varepsilon \mid Bx,\\
&\quad\;\; A \;\rightarrow\; Sx \mid DS,\\
&\quad\;\; B \;\rightarrow\; Ay \mid AB \mid Ez,\\
&\quad\;\; C \;\rightarrow\; xxz,\\
&\quad\;\; D \;\rightarrow\; \varepsilon \mid y \mid ExC \mid ExS,\\
&\quad\;\; E \;\rightarrow\; Ex \mid EA\}
\end{aligned}
$$

Převeďte gramatiku G na ekvivalentní vlastní gramatiku neobsahující levou rekurzi. Do řešení uveďte celý postup převodu, zejména následující mezivýsledky:

- gramatiku G_1 (ekvivalentní G), která neobsahuje nepoužitelné symboly;
- gramatiku G_2 (ekvivalentní G_1), která neobsahuje nepoužitelné symboly ani ε-pravidla (nezapomeňte uvést množinu N_ε obsahující všechny neterminály, které se dají přepsat na ε);
- gramatiku G_3 (ekvivalentní G_2), která neobsahuje nepoužitelné symboly, ε-pravidla ani jednoduchá pravidla;
- výslednou gramatiku G_4 (ekvivalentní G_3), která neobsahuje nepoužitelné symboly, ε-pravidla, jednoduchá pravidla, nepřímou ani přímou levou rekurzi (uveďte, jaké uspořádání neterminálů jste zvolili při odstraňování nepřímé levé rekurze).

Prvním krokem je eliminace nepoužitelných symbolů. Postupujeme podle algoritmu z přednášky (6. přednáška, slide 9): nejdříve odstraníme nenormované symboly (slide 3), poté odstraníme nedosažitelné symboly (slide 7). Výsledná gramatika G_1 vypadá následovně:

$$
\begin{aligned}
G_1 &= (\{S, A, B, D\}, \{x, y\}, P_1, S)\\
P_1 &= \{S \;\rightarrow\; \varepsilon \mid Bx,\\
&\quad\;\; A \;\rightarrow\; Sx \mid DS,\\
&\quad\;\; B \;\rightarrow\; Ay \mid AB,\\
&\quad\;\; D \;\rightarrow\; \varepsilon \mid y\}
\end{aligned}
$$

Druhým krokem je odstranění ε-pravidel. Postupujeme podle algoritmu z přednášky (6. přednáška, slide 13). Množina N_ε neterminálů, které se dají přepsat na ε, je $N_\varepsilon = \{S, D, A\}$. Jelikož $S \in N_\varepsilon$, musíme zavést nový iniciální neterminál S', který se bude moct přepsat na původní iniciální neterminál S nebo na ε. Výsledná gramatika G_2 vypadá následovně:

$$
\begin{aligned}
G_2 &= (\{S', S, A, B, D\}, \{x, y\}, P_2, S') \\
P_2 &= \{S' \ \rightarrow \ S \ \mid \varepsilon, \\
&\quad S \ \rightarrow \ Bx, \\
&\quad A \ \rightarrow \ Sx \mid x \mid D \mid S \ \mid DS, \\
&\quad B \ \rightarrow \ Ay \mid y \mid AB \mid B, \\
&\quad D \ \rightarrow \ y\}
\end{aligned}
$$

Dalším krokem je eliminace jednoduchých pravidel. Opět použijeme algoritmus z přednášky (6. přednáška, slide 17). Výsledkem algoritmu bude gramatika G_3:

$$
\begin{aligned}
G_3 &= (\{S', S, A, B, D\}, \{x, y\}, P_3, S') \\
P_3 &= \{S' \ \rightarrow \ Bx \mid \varepsilon, \\
&\quad S \ \rightarrow \ Bx, \\
&\quad A \ \rightarrow \ Sx \mid x \mid y \mid Bx \mid DS, \\
&\quad B \ \rightarrow \ Ay \mid y \mid AB, \\
&\quad D \ \rightarrow \ y\}
\end{aligned}
$$

Zbývá nám už jenom odstranit levou rekurzi (nepřímou a přímou), opět se inspirujeme přednáškou (7. přednáška, slidy 3–5). Neterminální symboly uspořádáme například následovně: $S' < S < A < B < D$.

- **neterminál S':**
  Hledáme všechna pravidla tvaru $S' \rightarrow X\alpha$ kde X je neterminál nacházející se v uspořádání před S' a $\alpha \in \{S', S, A, B, D, x, y\}^*$. Jelikož takové neterminály nejsou, nic se nezmění. S' nemá ani přímou levou rekurzi, množinu pravidel tedy neměníme.

- **neterminál S:**
  Hledáme všechna pravidla tvaru $S \rightarrow X\alpha$ kde X je neterminál nacházející se v uspořádání před S a $\alpha \in \{S', S, A, B, D, x, y\}^*$. Máme jedinou možnost: $X = S'$. Jelikož žádné pravidlo takového tvaru nemáme, nic se nezmění. S nemá ani přímou levou rekurzi, množinu pravidel tedy neměníme.

- **neterminál A:**
  Hledáme všechna pravidla tvaru $A \rightarrow X\alpha$ kde X je neterminál nacházející se v uspořádání před A, přičemž tyto neterminály uvažujeme v pořadí daném uspořádáním, tedy nejprve $X = S'$, potom $X = S$ a $\alpha \in \{S', S, A, B, D, x, y\}^*$. Pro S' žádné takové pravidlo neexistuje, pro S takové pravidlo existuje jedno ($A \rightarrow Sx$), nahradíme jej tedy pravidlem $A \rightarrow Bxx$ (za S jsme dosadili pravou stranu všech pravidel tvaru $S \rightarrow \beta$, $\beta \in \{S', S, A, B, D, x, y\}^*$). Původní pravidlo $A \rightarrow Sx$ vypustíme.

  Neterminál A přímou levou rekurzi nemá, další změny tedy neprovádíme.

A pravidla tedy nyní vypadají takto:

$$A \ \rightarrow \ Bxx \mid x \mid y \mid Bx \mid DS$$

- **neterminál B:**
Hledáme všechna pravidla tvaru $B \rightarrow X\alpha$ kde X je neterminál nacházející se v uspořádání před B, tedy postupně $X = S'$, $X = S$, $X = A$ a $\alpha \in \{S', S, A, B, D, x, y\}^*$. Pro S' a S taková pravidla nemáme, pro A taková pravidla máme dvě ($B \rightarrow Ay$ a $B \rightarrow AB$). Každé z nich nahradíme sadou pravidel vzniklých substitucí prvního neterminálu A za pravé strany pravidel $A \rightarrow \beta$, $\beta \in \{S', S, A, B, D, x, y\}^*$. Pozor, pravidla pro přepis neterminálu A jsme v předešlém kroku měnili, musíme použít tuto novou množinu pravidel! Po substituci dostáváme:

$$B \ \rightarrow Bxxy \mid xy \mid yy \mid Bxy \mid DSy \mid y \mid BxxB \mid xB \mid yB \mid BxB \mid DSB$$

Nyní musíme odstranit přímou levou rekurzi na neterminálu B. Zavedeme nový neterminál B' a podle předpisu (7. přednáška, slide 3) vytvoříme pro B a B' následující množiny pravidel:

$$\begin{aligned} B \ &\rightarrow xy \mid yy \mid DSy \mid y \mid xB \mid yB \mid DSB \mid \\ &\quad\ xyB' \mid yyB' \mid DSyB' \mid yB' \mid xBB' \mid yBB' \mid DSBB' \\ B' \ &\rightarrow xxy \mid xy \mid xxB \mid xB \mid xxyB' \mid xyB' \mid xxBB' \mid xBB' \end{aligned}$$

- **neterminál D:**
Hledáme všechna pravidla tvaru $D \rightarrow X\alpha$ kde X je neterminál nacházející se v uspořádání před D, tedy postupně $X = S'$, $X = S$, $X = A$, $X = B$ a $\alpha \in \{S', S, A, B, D, x, y\}^*$. Jelikož žádné pravidlo takového tvaru nemáme, nic se nezmění. D nemá ani přímou levou rekurzi, množinu pravidel tedy neměníme.

Výsledná gramatika G_4 tedy vypadá následovně:

$$\begin{aligned} G_4 &= (\{S', S, A, B, B', D\}, \{x, y\}, P_4, S') \\ P_4 &= \{S' \ \rightarrow Bx \mid \varepsilon, \\ &\qquad S \ \rightarrow Bx, \\ &\qquad A \ \rightarrow Bxx \mid x \mid y \mid Bx \mid DS, \\ &\qquad B \ \rightarrow xy \mid yy \mid DSy \mid y \mid xB \mid yB \mid DSB \mid \\ &\qquad\qquad xyB' \mid yyB' \mid DSyB' \mid yB' \mid xBB' \mid yBB' \mid DSBB', \\ &\qquad B' \ \rightarrow xxy \mid xy \mid xxB \mid xB \mid xxyB' \mid xyB' \mid xxBB' \mid xBB', \\ &\qquad D \ \rightarrow y\} \end{aligned}$$