IB102 – úkol 9 – řešení **Odevzdání: 29. 11. 2010**

Vypracoval: James Bond UČO: 007
Skupina: MI6

1. [2 body] Mějme následující jazyk:

$$L = \{w \in \{a,b\}^* \mid w = w^R\} \cup \{w \in \{a,b\}^* \mid \#_a(w) \bmod 2 = \#_b(w) \bmod 2\}$$

Sestrojte *jednoznačnou* bezkontextovou gramatiku generující jazyk L. Stručně vysvětlete, proč je Vaše gramatika jednoznačná.

Řešení: Zřejmě umíme sestrojit jednoznačnou gramatiku pro $\{w \in \{a,b\}^* \mid w = w^R\}$ (jazyk všech palindromů) i pro $\{w \in \{a,b\}^* \mid \#_a(w) \bmod 2 = \#_b(w) \bmod 2\}$ (jazyk všech slov se stejnou paritou symbolů a a b). Sjednocení těchto jazyků ale není disjunktní. Budeme se tedy nejprve snažit napsat L jako disjunktní sjednocení dvou jazyků, pro které umíme sestrojit jednoznačné gramatiky.

Použijeme následující úvahu: Všechny palindromy sudé délky splňují podmínku stejné parity a a b, zatímco žádný palindrom liché délky tuto podmínku nesplňuje. Jazyk L můžeme tedy ekvivalentně napsat takto:

$$L = \{w \in \{a,b\}^* \mid w = w^R, |w| \text{ je lichá}\} \cup \{w \in \{a,b\}^* \mid \#_a(w) \bmod 2 = \#_b(w) \bmod 2\}$$

Hledaná gramatika je tedy např. $G = (\{S, X, Y_{00}, Y_{01}, Y_{10}, Y_{11}\}, \{a,b\}, P, S)$, kde

$$\begin{aligned} P = \{ \quad S &\rightarrow X \mid Y_{00}, \\ X &\rightarrow a \mid b \mid aXa \mid bXb, \\ Y_{00} &\rightarrow aY_{10} \mid bY_{01} \mid \varepsilon, \\ Y_{01} &\rightarrow aY_{11} \mid bY_{00}, \\ Y_{10} &\rightarrow aY_{00} \mid bY_{11}, \\ Y_{11} &\rightarrow aY_{01} \mid bY_{10} \mid \varepsilon \quad \}. \end{aligned}$$

(Zde neterminál X generuje právě všechny palindromy liché délky a neterminál Y_{00} generuje právě všechna slova se stejnou paritou a a b.)

IB102 – úkol 9 – řešení **Odevzdání: 29. 11. 2010**

Vypracoval: James Bond UČO: 007
Skupina: MI6

2. [**3 body**] Mějme gramatiku $G = (\{S, A, B, C, D, E, F\}, \{a, b, c\}, P, S)$, kde

$$
\begin{array}{rrcl}
P = \{ & S & \rightarrow & AB \mid cC, \\
& A & \rightarrow & \varepsilon \mid aA \mid SEb, \\
& B & \rightarrow & \varepsilon \mid bB \mid AEc, \\
& C & \rightarrow & D \mid AcB, \\
& D & \rightarrow & Bb \mid Abc \mid C, \\
& E & \rightarrow & abE \mid Ec \mid EF, \\
& F & \rightarrow & Ea \mid ba \ \ \}.
\end{array}
$$

Ke gramatice G sestrojte (použitím algoritmů z přednášky) ekvivalentní gramatiku v Chomského normální formě.

Poznámka: Gramatika v CNF musí být vždy redukovaná.

Řešení: Nejprve z gramatiky odstraníme nepoužitelné symboly (tento krok bychom nemuseli v tuto chvíli dělat, je ale výhodné jej provést, protože se tím zbavíme zbytečných částí gramatiky a dále budeme pracovat s menší gramatikou).

- Normované neterminály jsou: $N_e = \{A, B, F, S, C, D\}$. (V prvním kroku algoritmu to jsou A, B a F, v druhém se přidají S, C a D, ve třetím kroku se nic nepřidá, proto skončíme.) Odstraníme tedy z gramatiky neterminál E a všechna pravidla jej obsahující.

$$
\begin{array}{rcl}
S & \rightarrow & AB \mid cC \\
A & \rightarrow & \varepsilon \mid aA \\
B & \rightarrow & \varepsilon \mid bB \\
C & \rightarrow & D \mid AcB \\
D & \rightarrow & Bb \mid Abc \mid C \\
F & \rightarrow & ba
\end{array}
$$

- Dosažitelné symboly jsou: $N' = \{S, A, B, C, D\}$ a $\Sigma' = \{c, a, b\}$. (V nultém kroku algoritmu to je S, v prvním kroku se přidají A, B, c a C, ve druhém kroku se přidají a, b a D, ve třetím kroku se nepřidá nic a proto skončíme.) Odstraníme tedy z gramatiky neterminál F.

$$
\begin{array}{rcl}
S & \rightarrow & AB \mid cC \\
A & \rightarrow & \varepsilon \mid aA \\
B & \rightarrow & \varepsilon \mid bB \\
C & \rightarrow & D \mid AcB \\
D & \rightarrow & Bb \mid Abc \mid C
\end{array}
$$

Nyní je třeba odstranit ε-pravidla. Spočítáme si proto nejprve $N_\varepsilon = \{A, B, S\}$ a postupujeme dále podle algoritmu z přednášky (S' je nyní nový počáteční neterminál):

$$\begin{array}{lcl}
S' & \to & S \mid \varepsilon \\
S & \to & AB \mid cC \mid A \mid B \\
A & \to & aA \mid a \\
B & \to & bB \mid b \\
C & \to & D \mid AcB \mid Ac \mid cB \mid c \\
D & \to & Bb \mid Abc \mid C \mid b \mid bc
\end{array}$$

Následně odstraníme jednoduchá pravidla. Nejprve si napočítáme množiny N_X pro všechny neterminály X:

$$\begin{array}{ccc}
N_{S'} = \{S', S, A, B\} & N_A = \{A\} & N_C = \{C, D\} \\
N_S = \{S, A, B\} & N_B = \{B\} & N_D = \{D, C\}
\end{array}$$

A opět pokračujeme dále dle algoritmu z přednášky:

$$\begin{array}{lcl}
S' & \to & \varepsilon \mid AB \mid cC \mid aA \mid a \mid bB \mid b \\
S & \to & AB \mid cC \mid aA \mid a \mid bB \mid b \\
A & \to & aA \mid a \\
B & \to & bB \mid b \\
C & \to & AcB \mid Ac \mid cB \mid c \mid Bb \mid Abc \mid b \mid bc \\
D & \to & AcB \mid Ac \mid cB \mid c \mid Bb \mid Abc \mid b \mid bc
\end{array}$$

Tím nám mohly vzniknout nedosažitelné neterminály (nenormované neterminály odstraněním jednoduchých pravidel vzniknout nemohou). Spočítáme si proto opět množinu dosažitelných neterminálů: $\{S', A, B, C\}$. Odstraníme z gramatiky neterminály S a D.

$$\begin{array}{lcl}
S' & \to & \varepsilon \mid AB \mid cC \mid aA \mid a \mid bB \mid b \\
A & \to & aA \mid a \\
B & \to & bB \mid b \\
C & \to & AcB \mid Ac \mid cB \mid c \mid Bb \mid Abc \mid b \mid bc
\end{array}$$

Nyní můžeme přejít k samotnému převodu na CNF.
Výsledná gramatika je $G' = (\{S', A, B, C, \langle cB\rangle, \langle bc\rangle, a', b', c'\}, \{a, b, c\}, P', S')$, kde

$$\begin{array}{rlcl}
P' = \{ & S' & \to & \varepsilon \mid AB \mid c'C \mid a'A \mid a \mid b'B \mid b, \\
& A & \to & a'A \mid a, \\
& B & \to & b'B \mid b, \\
& C & \to & A\langle cB\rangle \mid Ac' \mid c'B \mid c \mid Bb' \mid A\langle bc\rangle \mid b \mid b'c', \\
& \langle cB\rangle & \to & c'B, \\
& \langle bc\rangle & \to & b'c', \\
& a' & \to & a, \\
& b' & \to & b, \\
& c' & \to & c \ \}.
\end{array}$$