Parsing Beyond CFG
Homework 6: LCFRS

Laura Kallmeyer, Tatiana Bladier

Sommersemester 2018

Question 1

Consider the following LCFRS in the sRCG format:

$$G = \langle \{A, B, S\}, \{a, b\}, \{V, W, X, Y, Z\}, P, S \rangle$$

where

$$P = \{ \quad \begin{array}{l} S(VWXYZ) \rightarrow A(V, X, Z)B(W, Y), \\ A(aX, aY, aZ) \rightarrow A(X, Y, Z), \\ B(bX, bY) \rightarrow B(X, Y), \\ A(a, a, a) \rightarrow \epsilon, \\ B(b, b) \rightarrow \epsilon \end{array} \quad \}$$

1. What do the yields of A and B look like, given this grammar?
2. What is the string language generated by this LCFRS?
3. Give the same LCFRS in MCFG notation, i.e., with separate composition functions F that describe for each rule how to compute the yield of the righthand side non-terminal from the yields of the lefthand side non-terminals.

Solution:

1. yield(B) $= \{\langle b^n, b^n \rangle \mid n \geq 1\}$; yield(A) $= \{\langle a^n, a^n, a^n \rangle \mid n \geq 1\}$
2. $L = \{a^n b^m a^n b^m a^n \mid n, m \geq 1\}$
3. $G' = \langle \{A, B, S\}, \{a, b, c\}, F, P_1, S \rangle$, where P_1 and F have the following entries:

$$P_1 = \{ \quad \begin{array}{l} S \rightarrow f_1[A, B], \\ A \rightarrow f_2[A], \\ B \rightarrow f_3[B], \\ A \rightarrow f_4[\,], \\ B \rightarrow f_5[\,] \end{array} \quad \} \qquad F = \{ \quad \begin{array}{l} f_1[\langle V, X, Z \rangle, \langle W, Y \rangle] = \langle VWXYZ \rangle, \\ f_2[X, Y, Z] = \langle aX, aY, aZ \rangle, \\ f_3[X, Y] = \langle bX, bY \rangle, \\ f_4[\,] = \langle a, a, a \rangle, \\ f_5[\,] = \langle b, b \rangle \end{array} \quad \}$$

Question 2

Give an LCFRS in the sRCG format (the format from the previous exercise) for the language $L = \{a^n b^n (cd)^n e^n \mid n \geq 0\}$.

Solution:

$$G = \langle \{A, S\}, \{a, b, c, d, e\}, \{W, X, Y, Z\}, P, S \rangle$$

where

$$P = \{ \quad \begin{array}{l} S(\epsilon) \rightarrow \epsilon, \\ S(WXYZ) \rightarrow A(W, X, Y, Z), \\ A(a, b, cd, e) \rightarrow \epsilon \\ A(aW, bX, cdY, eZ) \rightarrow A(W, X, Y, Z) \end{array}$$