Exercises for FLL, Fall 2018, sheet 7 – Solutions

Return Thursday Nov 1, in class

Exercise 1. (requires a little work!) Convert the following grammar $G = (V, T, P, S)$ into CNF, by (i) eliminating ε-productions, (ii) eliminating unit productions, (iii) eliminating useless symbols, (iv) putting the resulting grammar in CNF.

$S \to 0A0 \mid 1B1 \mid AB$

$A \to C$

$B \to S \mid A$

$C \to S \mid \varepsilon$

Solution: (i) **a.** Finding nullable variables: NULL(1) = $\{C\}$, NULL(2) = $\{C, A\}$, NULL(3) = $\{C, A, B\}$, NULL(4) = NULL(5) = $\{C, A, B, S\}$. **b.** For $S \to 0A0$ add $\{S \to 0A0, S \to 00\}$ to P', for $S \to 1B1$ add $\{S \to 1B1, S \to 11\}$ to P', for $S \to AB$ add $\{S \to AB, S \to B, S \to A\}$ to P', for $A \to C$ add $\{A \to C\}$ to P', for the remaining rules add $\{B \to S, B \to A, C \to S\}$ to P'. This gives a new set P'

$S \to 0A0 \mid 00 \mid 1B1 \mid 11 \mid AB \mid B \mid A$

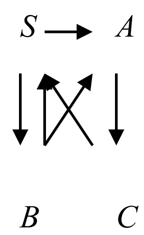
$A \to C$

$B \to S \mid A$

$C \to S$

(ii) **a.** Finding unit pairs: PAIRS(1) = $\{(A, A), (B, B), (C, C), (S, S)\}$, PAIRS(2) = $\{(A, A), (B, B), (C, C), (S, S), (S, B), (S, A), (A, C), (B, S), (B, A), (C, S)\}$, PAIRS(3) = $\{(A, A), (B, B), (C, C), (S, S), (S, B), (S, A), (A, C), (B, S), (B, A), (C, S), (S, C), (A, S), (B, C), (C, A), (C, B)\}$,

PAIRS(4) = PAIRS(5) = $\{(A, A), (B, B), (C, C), (S, S), (S, B), (A, C), (B, S), (B, A), (C, S), (S, A), (A, S), (B, C), (C, B), (S, C), (A, B), (C, A)\}$. An easier way to see that here *all* pairs are unit pairs is to check the following directed graph created by the unit transitions from P' and see that it is cyclic, that is, every node is transitively reachable from every other node:



$S \to 0A0 \mid 00 \mid 1B1 \mid 11 \mid AB \mid B \mid A$

$A \to C$

$B \to S \mid A$

$C \to S$

b. Stripping from P' all unit productions and then adding all productions of the form $A \to \alpha$, where $B \to \alpha$ is a non-unit production in P' and (A, B) is a unit pair, yields $P'' =$

$S \to 0A0 \mid 00 \mid 1B1 \mid 11 \mid AB$

$A \to 0A0 \mid 00 \mid 1B1 \mid 11 \mid AB$

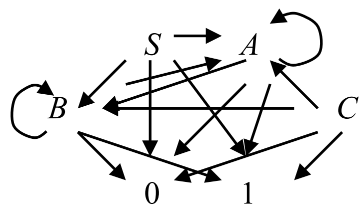
$B \to 0A0 \mid 00 \mid 1B1 \mid 11 \mid AB$

$C \to 0A0 \mid 00 \mid 1B1 \mid 11 \mid AB$

(iii) **a.** We first detect all generating symbols. GEN(1) = {0,1}, GEN(2) = GEN(3) = {0, 1, A, B, C, S}.

b. Deleting from G all nongenerating symbols and productions in which such symbols occur, yields $G_2 = (V, T, P'', S)$, because there are no non-generating symbols or productions.

c. Next we find all reachable symbols of G_2. The graph described in the lecture notes is



From this we see that the reachable symbols are $\{0, 1, S, A, B\}$.

d. Finally we eliminate from G_2 all non-reachable symbols and productions in which such symbols occur, to obtain $G_1 = (\{S, A, B\}, \{0, 1\}, P''', S)$, where $P''' =$

$$S \to 0A0 \mid 00 \mid 1B1 \mid 11 \mid AB$$

$$A \to 0A0 \mid 00 \mid 1B1 \mid 11 \mid AB$$

$$B \to 0A0 \mid 00 \mid 1B1 \mid 11 \mid AB$$

(iv) In the last step, we obtain a CNF grammar by carrying out the two steps given in the proof of theorem 4.10 in the lecture notes.

a. Arrange that all bodies of lenght 2 or more consists only of variables. This gives us productions $P'''' =$

$$S \to A_0AA_0 \mid A_0A_0 \mid A_1BA_1 \mid A_1A_1 \mid AB$$

$$A \to A_0AA_0 \mid A_0A_0 \mid A_1BA_1 \mid A_1A_1 \mid AB$$

$$B \to A_0AA_0 \mid A_0A_0 \mid A_1BA_1 \mid A_1A_1 \mid AB$$

$$A_0 \to 0$$

$$A_1 \to 1$$

b. Break productions with all-variable bodies of length 3 or more into sequences of productions of the form $A \to BC$. This gives us the final rule set P_{CNF} =

$$S \to A_0A' \mid A_0A_0 \mid A_1B' \mid A_1A_1 \mid AB$$

$$A \to A_0A' \mid A_0A_0 \mid A_1B' \mid A_1A_1 \mid AB$$

$$B \to A_0A' \mid A_0A_0 \mid A_1B' \mid A_1A_1 \mid AB$$

$$A' \to AA_0$$
$$B' \to BA_1$$
$$A_0 \to 0$$

$$A_1 \to 1$$