Automata and Grammars

SS 2018

Assignment 2: Solutions to Selected Problems

Seminary: Thursday, March 8, 2018.

Problem 2.1. [Normal Form for Regular Grammars]

Determine a grammar in *right normal form* that generates the same language as the following grammar $G = (\{S, A, B, C, D\}, \{a, b\}, S, P)$, where P contains the following productions:

$$\begin{array}{lllllllll} S & \to & aaA, & S & \to & bbB, & S & \to & C, \\ A & \to & aaA, & A & \to & B, & & & \\ B & \to & bbB, & B & \to & A, & B & \to & C, \\ C & \to & aaA, & C & \to & bbB, & C & \to & D, \\ D & \to & aaA, & D & \to & bbB, & D & \to & \varepsilon. \end{array}$$

Hint: Use the construction from the proof of Theorem 2.5.

Solution. The construction proceeds in three steps:

1. Eliminate the ε-productions.
2. Eliminate the chain productions.
3. Split long productions into short ones.

1. $N_1 = \{ X \in \{S, A, B, C, D\} \mid X \to_P^* \varepsilon \} = \{S, A, B, C, D\}$.

   Now we remove all ε-rules and introduce replacement rules, which results in the set of productions P_1:

$$\begin{array}{lllllllllllllll} S & \to & aaA, & S & \to & bbB, & S & \to & C, & S & \to & aa, & S & \to & bb, \\ A & \to & aaA, & A & \to & B, & A & \to & aa, & & & & & & \\ B & \to & bbB, & B & \to & A, & B & \to & C, & B & \to & bb, & & & \\ C & \to & aaA, & C & \to & bbB, & C & \to & D, & C & \to & aa, & C & \to & bb, \\ D & \to & aaA, & D & \to & bbB, & D & \to & aa, & D & \to & bb. & & & \end{array}$$

   As $S \in N_1$, we also add the production $S \to \varepsilon$. Observe that S does not occur on the right-hand side of any production.

2. P_1 contains the following chain rules:

$$S \to C, \quad A \to B, \quad B \to A, \quad B \to C, \quad C \to D.$$

   Hence, $A \leftrightarrow B < C$, and $S < C < D$. First we replace all occurrences of B by A, deleting chain rules of the form $X \to X$ and multiple occurrences of productions, which yields P_2:

$$\begin{array}{llllllllllllllllll} S & \to & aaA, & S & \to & bbA, & S & \to & C, & S & \to & aa, & S & \to & bb, & S & \to & \varepsilon, \\ A & \to & aaA, & & & & A & \to & aa, & & & & & & & & & \\ A & \to & bbA, & & & & A & \to & C, & A & \to & bb, & & & & & & \\ C & \to & aaA, & C & \to & bbA, & C & \to & D, & C & \to & aa, & C & \to & bb, & & & \\ D & \to & aaA, & D & \to & bbA, & D & \to & aa, & D & \to & bb. & & & & & & \end{array}$$

Finally, we remove all chain rules and introduce corresponding replacement rules, proceeding from D down to S, which gives P_3:

$$\begin{array}{llllllllllllll}
S & \to & aaA, & S & \to & bbA, & S & \to & aa, & S & \to & bb, & S \to \varepsilon, \\
A & \to & aaA, & A & \to & aa, & A & \to & bbA, & A & \to & bb, & \\
C & \to & aaA, & C & \to & bbA, & C & \to & aa, & C & \to & bb, & \\
D & \to & aaA, & D & \to & bbA, & D & \to & aa, & D & \to & bb. &
\end{array}$$

3. As C and D are not reachable from S, we can delete all C- and D-productions from P_3. The remaining S- and A-productions are split into shorter productions, where $S_1, S_2, S_3, S_4, A_1, A_2, A_3, A_4$ are new nonterminals, which yields the final set of productions P_4 in right normal form:

$$\begin{array}{llllllllllllll}
S & \to & aS_1, & S & \to & bS_2, & S & \to & aS_3, & S & \to & bS_4, & S \to \varepsilon, \\
S_1 & \to & aA, & S_2 & \to & bA, & S_3 & \to & a, & S_4 & \to & b, & \\
A & \to & aA_1, & A & \to & aA_2, & A & \to & bA_3, & A & \to & bA_4, & \\
A_1 & \to & aA, & A_2 & \to & a, & A_3 & \to & bA, & A_4 & \to & b. &
\end{array}$$

□

Problem 2.2 [Closure Properties for Regular Grammars]

Let $G_1 = (N_1, \Sigma, S_1, P_1)$ and $G_2 = (N_2, \Sigma, S_2, P_2)$ be two right regular grammars such that N_1 and N_2 are disjoint.

1. Construct a right regular grammar G from G_1 and G_2 such that $L(G) = L(G_1) \cup L(G_2)$.
2. Construct a right regular grammar G from G_1 and G_2 such that $L(G) = L(G_1) \cap L(G_2)$.
3. Construct a right regular grammar G from G_1 and G_2 such that $L(G) = L(G_1) \cdot L(G_2)$.
4. Construct a right regular grammar G from G_1 such that $L(G) = (L(G_1))^*$.

Hint: You may assume that the grammars G_1 and G_2 are in right normal form.

Solution. Let G_1 and G_2 be in right normal form, and let S be a new nonterminal.

1. Let $G = (\{S\} \cup N_1 \cup N_2, \Sigma, S, P)$, where P is defined as follows:

$$P = P_1 \cup P_2 \cup \{ S \to \alpha_1 \mid (S_1 \to \alpha_1) \in P_1 \} \cup \{ S \to \alpha_2 \mid (S_2 \to \alpha_2) \in P_2 \}.$$

   Then it is easily seen that $L(G) = L(G_1) \cup L(G_2)$.

2. Let $G = (N, \Sigma, [S_1, S_2], P)$, where $N = \{ [A_1, A_2] \mid A_1 \in N_1, A_2 \in N_2 \}$ and P is defined as follows:

$$\begin{array}{lll}
P & = & \{ [A_1, A_2] \to a[B_1, B_2] \mid a \in \Sigma, (A_1 \to aB_1) \in P_1, \text{ and } (A_2 \to aB_2) \in P_2 \} \cup \\
 & & \{ [A_1, A_2] \to a \mid a \in \Sigma, (A_1 \to a) \in P_1, \text{ and } (A_2 \to a) \in P_2 \} \cup \\
 & & \{ [S_1, S_2] \to \varepsilon \mid (S_1 \to \varepsilon) \in P_1 \text{ and } (S_2 \to \varepsilon) \in P_2 \}.
\end{array}$$

   Then $L(G) = L(G_1) \cap L(G_2)$.

3. Let $G = (\{S\} \cup N_1 \cup N_2, S, P)$, where P is defined as follows:

$$\begin{array}{lll}
P & = & \{ S \to aA_1 \mid (S_1 \to aA_1) \in P_1 \} \cup \{ S \to aS_2 \mid (S_1 \to a) \in P_1 \} \cup \\
 & & \{ A_1 \to aB_1 \mid (A_1 \to aB_1) \in P_1 \} \cup \{ A_1 \to aS_2 \mid (A_1 \to a) \in P_1 \} \cup \\
 & & P_2 \cup \{ S \to S_2 \mid (S_1 \to \varepsilon) \in P_1 \}.
\end{array}$$

   Then $L(G) = L(G_1) \cdot L(G_2)$.

4. Let $G = (\{S\} \cup N_1, S, P)$, where P is defined as follows:

$$P = \{S \to \varepsilon, S \to S_1\} \cup P_1 \cup \{ A \to aS_1 \mid (A \to a) \in P_1 \}$$

Then $L(G) = (L(G_1))^*$. □

Problem 2.4 [DFAs]

Determine the languages that are accepted by the following deterministic finite-state automata $(\{q_0, q_1, q_2\}, \{0, 1\}, \delta, q_0, F)$, where δ and F are defined as follows:

(a) $F = \{q_1, q_2\}$ and δ :

	0	1
q_0	q_1	q_0
q_1	q_2	q_1
q_2	q_0	q_2

,

(b) $F = \{q_2\}$ and δ :

	0	1
q_0	q_0	q_1
q_1	q_0	q_2
q_2	q_0	q_2

,

(c) $F = \{q_1, q_2\}$ and δ :

	0	1
q_0	q_0	q_1
q_1	q_2	q_1
q_2	q_0	q_1

.

Solution. The above DFAs accept the following languages:

(a) $L_a = \{ w \in \{0,1\}^+ \mid |w|_0 \equiv 1 \mod 3 \text{ or } |w|_0 \equiv 2 \mod 3 \}$,
(b) $L_b = \{ w \in \{0,1\}^+ \mid w = u11 \text{ for some } u \in \{0,1\}^* \}$,
(c) $L_c = \{ w \in \{0,1\}^+ \mid w = u1 \text{ or } w = u10 \text{ for some } u \in \{0,1\}^* \}$.

□