Chomsky Normal Form

Exercise 1

Given the following grammar \( G = (V_N, V_T, P, S) \) where \( V_N, V_T \) are sets of nonterminals and terminals, \( S \) is the start symbol and \( P \) is a set of production rules as follows:

\[
\begin{align*}
S & \rightarrow OS \mid BD \mid EO \\
B & \rightarrow C \mid E \mid 1 \\
D & \rightarrow 0 \mid AO \mid S \mid E \\
E & \rightarrow BE \mid SE \\
C & \rightarrow O \\
F & \rightarrow C
\end{align*}
\]

a) Simplify the grammar.
b) Transform the simplified grammar into Chomsky Normal Form.

Exercise 2: Given the following grammar \( G = (V_N, V_T, P, S) \) where \( V_N=\{S,A,B\}, V_T=\{0,1\}, S \) is the start symbol, and \( P \) is as follows:

\[
\begin{align*}
S & \rightarrow A \mid B \\
A & \rightarrow OA \mid B \mid 00 \\
B & \rightarrow 1B0 \mid 10
\end{align*}
\]

Transform the grammar into Chomsky normal form.

Exercise 3: Given the grammar \( G = (V_N, V_T, P, S) \) where \( V_N=\{S,A,B\}, V_T=\{0,1\}, S \) is the start symbol, \( P \) is as follows:

\[
\begin{align*}
S & \rightarrow A \mid B \\
A & \rightarrow OS0 \mid B \mid 00 \\
B & \rightarrow 1B0 \mid 10
\end{align*}
\]

Transform the grammar into Chomsky normal form.
1) a) Elimination of E-production:

Nullable Symbols: It. 0: \{B, D\}
It. 1: \{B, D, S\}
It. 2: \{B, D, S\}

we get:

\[ S \rightarrow OS \mid BD \mid E0 \mid O \mid B \mid D \mid E \]
\[ B \rightarrow C \mid E \mid 1 \]
\[ D \rightarrow 0 \mid AO \mid S \mid E \]
\[ E \rightarrow BE \mid SE \mid E \]
\[ C \rightarrow 0 \]
\[ F \rightarrow C \]

Since \( S \) is nullable the resulting grammar is:

\[ G_1 = (V_N, V_T, \rho_1, S') \] where \( V_N = V_N \cup \{S'\} \)

\[ V_T = V_T \] \( \rho_1 \) is as follows:

\[ S' \rightarrow S \mid E \]
\[ S \rightarrow OS \mid BD \mid E0 \mid O \mid B \mid D \]
\[ B \rightarrow C \mid 1 \]
\[ D \rightarrow 0 \mid AO \mid S \]
\[ E \rightarrow BE \mid SE \mid E \]
\[ C \rightarrow 0 \]
\[ F \rightarrow C \]

Elimination of Unit Production.

Graph of unit productions:

\[ S \rightarrow S \]
\[ S \rightarrow B \rightarrow D \rightarrow F \rightarrow E \]

Reachability:

\[ S' \rightarrow S, S' \rightarrow B, S' \rightarrow D, S' \rightarrow C \]
\[ S \rightarrow *B, S \rightarrow *D, S \rightarrow *C \]
\[ B \rightarrow *C \]
\[ D \rightarrow *S, D \rightarrow *B, D \rightarrow *C \]
\[ F \rightarrow *C \]
we get: $G_2 = (V_{N_2}, V_{T_2}, P_2, S')$ where

$V_{N_2} = V_{N_1}$, $V_{T_2} = V_{T_1}$, $P_2$ is as follows:

$$S' \rightarrow e | OS | 0 | BD | 1 | AO | EO$$

$$S \rightarrow OS | 0 | BD | 1 | AO | EO$$

$$B \rightarrow 0 | 1$$

$$D \rightarrow 0 | AO | OS | BD | 1 | EO$$

$$E \rightarrow BE | SE$$

$$C \rightarrow 0$$

$$F \rightarrow 0$$

Elimination of useless symbols (non-generating)

Generating Symbols: it a: $\{0, 1\}$

it 1: $\{0, 1, B, C, D, F, S, S'\}$

it 2: $\{0, 1, B, C, D, F, S, S'\}$

non-generating symbols: $\{E, F\}$

we get $G_3 = (V_{N_3}, V_{T_3}, P_3, S')$ where $V_{N_3} = \{B, C, D, F, S, S'\}$

$V_{T_3} = \{0, 1\}$, $P_3$ is as follows:

$$S' \rightarrow e | OS | 0 | BD | 1$$

$$S \rightarrow OS | 0 | BD | 1$$

$$B \rightarrow 0 | 1$$

$$D \rightarrow 0 | OS | BD | 1$$

$$C \rightarrow 0$$

$$F \rightarrow 0$$
Elimination of useless symbols - (unreachable)

Reachable symbols:  

It. 0: \{S\'}
It. 1: \{S, S, 0, B, D, 1\}
It. 2: \{S, S, 0, B, D, 1\}

Unreachable Symbols: \{C, S\'}

we get \(G_4 = (V_{N_4}, V_{T_4}, P_4, S')\) where \(V_{N_4} = \{S', S, B, D\}\)  
\(V_{T_4} = \{0, 1\}\), and \(P_4\) is defined as follows:

\[
\begin{align*}
S' &\rightarrow e | 0S | 0 | BD | 1 | 0 \\
S &\rightarrow 0S | 0 | BD | 1 | 0 \\
B &\rightarrow 0 | 1 \\
D &\rightarrow 0 | 0S | BD | 1 \\
\end{align*}
\]

1b) Since we do not have long Productions, we only need to remove "mixed bodies" (such as 0S) by introducing some new Productions

we get \(G_5 = (V_{N_5}, V_{T_5}, P_5, S')\) where \(V_{N_5} = V_{N_4} \cup \{N_0\}\) and \(V_{T_5} = \{0, 1\}\) and \(P_5\) is as follows:

\[
\begin{align*}
S' &\rightarrow e | N_0 S | 0 | BD | 1 \\
S &\rightarrow N_0 S | 0 | BD | 1 \\
B &\rightarrow 0 | 1 \\
D &\rightarrow 0 | N_0 S | BD | 1 \\
N_0 &\rightarrow 0 \\
\end{align*}
\]
2) Elimination of $E$-productions:

Since there aren't any nullable symbol we skip this phase.

Elimination of unit productions:

Graph of unit productions:

![Graph of unit productions]

Reachability:

$s \Rightarrow^* A, s \Rightarrow^* B$

$a \Rightarrow^* b$

We get $G_1 = (V_M, V_F, P_1, S)$ where $V_M = V_F = V_F$

$P_1$ is as follows:

$s \rightarrow OA0 \mid 00 \mid 180 \mid 10$

$a \rightarrow OA0 \mid 180 \mid 10 \mid 00$

$b \rightarrow 180 \mid 10$

Elimination of useless symbols (non-generating)

Generating Symbols: It 0: \{1, 0\}

It 1: \{1, 0, A, B, S\}

It has all the symbols

Since non-generating symbols is an empty set, we also skip this phase.

Elimination of useless symbols (un-reachable)

Reachable Symbols: It 0: \{S\}

It 1: \{S, 0, 1, A, B\}

It has all the symbols

Since unreachable symbols is an empty set, we also skip this phase.
2) Continue with removing "mixed bodies"

we get $G_2 = (V_{T_2}, V_{N_2}, P_2, S)$ where

$V_{T_2} = V_{T_1} \cup \{N_0, N_1\}$, $V_{N_2} = V_N = \{0, 1\}$, $P_2$ is as follows:

- $S \rightarrow N_0 A N_0 | N_0 N_0 | N_1 B N_0 | N_1 N_0$
- $A \rightarrow N_0 A N_0 | N_1 B N_0 | N_1 N_0 | N_0 N_0$
- $B \rightarrow N_1 B N_0 | N_1 N_0$
- $N_1 \rightarrow 1$
- $N_0 \rightarrow 0$

"Factor" long productions,

we get $G_3 = (V_{T_3}, V_{N_3}, P_3, S)$ where

$V_{T_3} = V_{T_2} \cup \{B_1, B_2\}$, $V_{N_3} = \{0, 1\}$, $P_3$ is as follows:

- $S \rightarrow N_0 B_1 | N_0 N_0 | N_1 B_2 | N_1 N_0$
- $A \rightarrow N_0 B_1 | N_1 B_2 | N_1 N_0 | N_0 N_0$
- $B \rightarrow N_1 B_2 | N_1 N_0$
- $N_1 \rightarrow 1$
- $N_0 \rightarrow 0$
- $B_1 \rightarrow A N_0$
- $B_2 \rightarrow B N_0$
5) we only provide the solution.

Note that we have eliminated the useless symbol $A$.

The resulting grammar is $G = (V_N, V_T, P, S)$

where $V_N = \{ S, B, N_0, N_1, B_1, B_2 \}$, $V_T = \{ 0, 1 \}$

and $P$ is as follows:

$S \rightarrow N_0 B_1 \mid N_0 N_0 \mid N_1 B_2 \mid N_1 N_0$

$B \rightarrow N_1 B_2 \mid N_1 N_0$

$N_0 \rightarrow 0$

$N_1 \rightarrow 1$

$B_1 \rightarrow SN_0$

$B_2 \rightarrow BN_0$