Theory of Computation, CSCI 438 spring 2022
Chomsky Normal Form, pg. 108-111, Feb. 18

1. Convert the following grammar to Chomsky Normal Form.

$S \rightarrow aSb \mid \varepsilon$

Step 1. Add a new start symbol

$S_0 \rightarrow S$
$S \rightarrow aSb \mid \varepsilon$

Step 2. Remove ε-productions. There is only one ε-productions: $S \rightarrow \varepsilon$.

Removing $S \rightarrow \varepsilon$ gives:
$S_0 \rightarrow S \mid \varepsilon$
$S \rightarrow aSb \mid ab$

The ε-production $S_0 \rightarrow \varepsilon$ can't be removed. It indicates that ε is in the language.

Step 3. Remove all unit productions. There is only one unit production: $S_0 \rightarrow S$.

Remove $S_0 \rightarrow S$:
$S_0 \rightarrow aSb \mid ab \mid \varepsilon$
$S \rightarrow aSb \mid ab$

Step 4. Add new variables as needed to get the grammar to adhere to rule of Chomsky normal form.

$S_0 \rightarrow AV \mid AB \mid \varepsilon$
$S \rightarrow AV \mid AB$
$A \rightarrow a$
$V \rightarrow SB$
$B \rightarrow b$

2. Exercise 2.14 Convert the following CFG into an equivalent CFG in Chomsky normal form, using the procedure given in Theorem 2.9.

A→ BAB | B | ε
B→ 00 | ε

Step 1. Add a new start symbol

S→ A
A→ BAB | B | ε
B → 00 | ε

Step 2. Remove ε-productions. There are two ε-productions: A→ ε and B→ ε. Removing these in either order gives the same result. Removing B→ ε first is more efficient. I'll remove A→ ε first.

Removing A→ ε gives:
S→ A | ε
A→ BAB | BB | B
B → 00 | ε

Removing B→ ε gives:
S→ A | ε
A→ BAB | AB | BA | BB | B | ε
B → 00

This created A→ ε again. Removing A→ ε gives:
S→ A | ε
A→ BAB | BB | AB | B | BA | ε
B → 00

Removing A→ ε again gives:
S→ A | ε
A→ BAB | BB | AB | B | BA
B → 00

Step 3. Remove all unit productions. There are two unit productions: S→ A and A→ B. Removing these in either order gives the same result. Removing A→ B first is more efficient. I'll remove S→ A first.

Remove S→ A:
S→ BAB | BB | AB | B | BA | ε
A→ BAB | BB | AB | B | BA
B → 00

Remove S→ B:
S→ BAB | BB | AB | 00 | BA | ε
A→ BAB | BB | AB | B | BA

B → 00

Remove A→ B:
S→ BAB | BB | AB | 00 | BA | ε
A→ BAB | BB | AB | 00 | BA
B → 00

Step 4. Add new variables as needed to get the grammar to adhere to rule of Chomsky normal form.

S→ BD | BB | AB | CC | BA | ε
A→ BD | BB | AB | CC | BA
B → CC
C → 0
D → AB