

Exercise 5.1

Step 1: ϵ Rules

$S \rightarrow \epsilon$ is the only occurrence of ϵ and S never occurs on the right-hand side. We don't have to do anything here.

Step 2: Chain Rules

Since the CMS (WHAT DO YOU MEAN BY THAT??? \rightarrow CNF?) can only have a terminal or **two** Variables on the right-hand side, we have to eliminate every Rule that has only a single Variable on the right.

Chains in G :

$$S \rightarrow Z \rightarrow bb \quad S \rightarrow Z \rightarrow Za \quad X \rightarrow Y \rightarrow bY \quad Y \rightarrow X \rightarrow aZb$$

We eliminate the chains by linking the end result of the chain directly with the first Variable. Since $X \rightarrow Y$ and $Y \rightarrow X$, we can just write them together as X that points to the results of X and Y .

$$S \rightarrow bb \mid Za \mid XX \quad Z \rightarrow bb \mid Za \quad X \rightarrow aZb \mid bX$$

Step 3: Order

Now we have to eliminate all the rules that have terminals **and** Variables on the right-hand side. We do so by adding a new Variable for each terminal and having each new Variable point to exactly one terminal.

So we create the new rules:

$$A \rightarrow a \quad \text{and} \quad B \rightarrow b$$

and can now rewrite the Rules above to:

$$S \rightarrow BB \mid ZA \mid XX \quad Z \rightarrow BB \mid ZA \quad X \rightarrow AZB \mid BX$$

Step 4: Shorten

To get our CNF we have to shorten the right-hand side to exactly two variables per rule. Since $X \rightarrow AZB$ is too long, we add $V \rightarrow ZB$ and get our final Grammar $G' = \langle \{S, V, X, Z\}, \{a, b\}, P, S \rangle$ with the following rules in P :

$$\begin{array}{llll} S \rightarrow BB \mid ZA \mid XX & Z \rightarrow BB \mid ZA & X \rightarrow AV \mid BX & V \rightarrow ZB \\ A \rightarrow a & B \rightarrow b & & \end{array}$$

Exercise 5.2

Proof. Length of Derivations in Chomsky Normal Form:

In the case $|w| = 1$ we can only have one terminal Rule $S \rightarrow a$ because $S \rightarrow AB$ would be replaced by at least two terminal rules and therefore be $|w| \geq 2$.

With $S \rightarrow AB$ we can generate any word w with $|w| \geq 2$ in the language G by replacing the A and (NOT SURE IF OR WAS CORRECT!) B .

$$S \rightarrow AB \xrightarrow{(B \rightarrow CD)} ACD \xrightarrow{(D \rightarrow EF)} ACEF \rightarrow \dots \rightarrow X_1 \dots X_n$$

By the definition of the Chomsky Normal Form, the Grammar has either a terminal or two Variables on the right-hand side. Therefore with every Step, exactly one Variable can be added. So to generate $X_1 \dots X_n$ exactly $n - 1$ steps are needed.

Now we have to replace every Variable X with a terminal symbol. By definition we can only replace one X at a time and therefore need n more steps to generate our final word w .

In total we need exactly $2n - 1$ steps and since $n = |w|$ the statement is true. □

Exercise 5.3

Σ read input	step	Stack status
ϵ	$q_0 \rightarrow q_1$	#
a	$q_1 \rightarrow q_1$	#X
ϵ	$q_1 \rightarrow q_2$	#X
a	$q_2 \rightarrow q_2$	#XY
a	$q_2 \rightarrow q_2$	#XYY
d	$q_2 \rightarrow q_3$	#XYY
(a) b	$q_3 \rightarrow q_4$	#XYY
a	$q_4 \rightarrow q_3$	#XY
b	$q_3 \rightarrow q_4$	#XY
a	$q_4 \rightarrow q_3$	#X
ϵ	$q_3 \rightarrow q_5$	#X
c	$q_5 \rightarrow q_6$	#X
c	$q_6 \rightarrow q_5$	#
ϵ	$q_5 \rightarrow q_7$	ϵ

(b)

This PDA accepts the following language: $w = \{ a^x d (ba)^y (cc)^z \mid a = y + z, \quad x, y, z \in \mathbb{N}_0 \}$
The self-loops q_1 and q_2 stack a 's. Then follows a single d and the loop $q_3 \rightarrow q_4 \rightarrow q_3$ (ab) which

removes anything stacked by q_1 followed by the loop $q_5 \rightarrow q_6 \rightarrow q_5$ (cc) which removes anything stacked by q_2 . So in the end the visits of $q_1 + q_2$ have to equal the visits of $q_3 + q_5$.

Exercise 5.4