Homework *#7 due December 27,2022 before recitation*

Question 1

A *CFG* is called *right linear* if ***all*** productions are of the form $\boldsymbol{A \rightarrow a\ B}$ or $\boldsymbol{A \rightarrow e}$ and called *left linear* if ***all*** productions are of the form $\boldsymbol{A \rightarrow B\ a}$ or $\boldsymbol{A \rightarrow e}$ where $\boldsymbol{A, B} \in \boldsymbol{V}$ and $\boldsymbol{a} \in \boldsymbol{T}$ and $\boldsymbol{e}$ is the empty string.

Show that both *right linear* and *left linear* grammars generate *regular languages*. Specify finite state machines corresponding respectively to right and left linear grammars.

Main Text: Exercise 7.1.3, 7.1.4, 7.2.1 (b), (c), 7.4.3(b), (c)

HOMEWORK 7

Question I:

Right linear regular grammar:

$S \to X$

or

$S \to XT$

$\Rightarrow X$ = any zero string

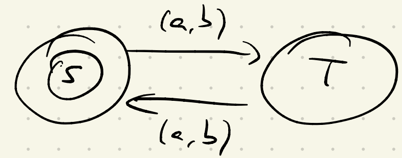
Left linear regular grammar:

$S \to X$

or

$S \to TX$

Stack Machine:



Regular grammar for above stack machine:

$S \to aT \mid \epsilon$

$S \to bT \mid \epsilon$

$T \to aS$

$T \to bS$

A grammar is regular, then it should be either right or left linear.

$G = (\{S\}, \{a,b\}, R1, S)$ R1 as $S \to abS \mid a$ is a right linear, with G we can derive the below string (ababa) as $S \Rightarrow abS \Rightarrow ababS \Rightarrow ababa$

$L(G) = (ab)^* a$

$G1 = (\{S, A, B\}, \{a,b\}, R2, S)$ R2 as the following left linear grammar

$S \to Aab$

$A \to Aab \mid B$

$B \to a$

$\Longrightarrow L(aab(ab)^*)$

Main Text

Exercise 7.1.3:

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

$A \to C$

$B \to S \mid A$

$C \to S \mid \epsilon$

a) Eliminate ϵ-productions

Since C, A, B can produce ϵ, hence S can produce ϵ.

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

b) Eliminating unit productions

We note that A, B, and C will all just produce S again. So we remove them.

$S \to 00 \mid 0A0 \mid 1B1 \mid BB$
$A \to 00 \mid 0A0 \mid 1B1 \mid BB$
$B \to 00 \mid 0A0 \mid 1B1 \mid BB$
$C \to 00 \mid 0A0 \mid 1B1 \mid BB$

c) Eliminating any useless symbols in the resulting grammar.

The variable C has now become unreachable. We also remove A and B, because they are exactly equal to S. Our grammar becomes:

$$S \to 00 \mid 0S0 \mid 1S1 \mid SS$$

d) Chomsky Normal Form

Creating variables $A \to 0$ and $B \to 1$. We then divide the two productions of length 3 using variables C and D.

$$S \to AA \mid AC \mid BD \mid SS$$

$$A \to 0$$

$$B \to 1$$

$$C \to SA$$

$$D \to SB$$

Exercise 7.1.4 :

$S \to AAA \mid B$

$A \to aA \mid B$

$B \to \epsilon$

a) Eliminate ϵ productions

$S \to AAA \mid B \mid \epsilon$

$A \to aA \mid B \mid \epsilon$

$S \to AAA \mid AA \mid A \mid B \mid \epsilon$

$A \to aA \mid a \mid B$

$S \to AAA \mid AA \mid A \mid B$

$A \to aA \mid a \mid B$

b) Eliminate unit productions

$S \to AAA \mid AA \mid aA \mid a \mid B \mid B$

$A \to aA \mid a \mid B$

c) Eliminate useless symbols

$S \to B$ and $A \to B$ are useless symbols

$S \to AAA \mid AA \mid aA \mid a$

$A \to aA \mid a$

D) Convert into CNF.

$C1 \to AA$; $C2 \to a$

$S \to C1A \mid AA \mid C2a \mid a$

$A \to C2A \mid a$

$C1 \to AA$; $C2 \to a$

S, C1, A, C2 are variables

a is terminal symbol

Exercise 7.2.1 (b), (c):

$\boxed{b) \ L = \{a^n b^n c^i \mid i \leq n\}}$

Assume this language is context free. We will apply the PL to reach a contradiction.

Let string be $z = a^p b^p c^p$, we can break z into $uvwxy$, where

→ $|vwx| \leq p$
→ $|vx| > 0$
→ $uv^iwx^iy \in L$ for all $i \geq 0$.

We have 5 different cases:

→ v and x both contain only a's.

→ v contains only a's and x contains only b's

→ v and x both contain only b's.

→ v contains only b's and x contains only c's.

→ v and x both contain only c's.

In every cases strings cannot be in L.

c) $L = \{0^p \mid p \text{ is a prime}\}$

Assume L is context free, and let $z = 0^k$. k is a prime that is at least $n+2$, where n is the pumping length.

$z = uvwxy$, and no matter how we pump v and x, we will get another string in our language. Let $|vx| = m$, and let $z' = uv^{k-m}wx^{k-m}y$. Then $|z'| = m(k-m) + m = (m+1)(k-m)$.

$m+1$ must be at least 2, because $|vx| > 0$. Also, $k-m$ must be at least 2, because we chose k to be at least $n+2$, and $|vx| \leq |vwx| \leq n$. Therefore, $|z'|$ cannot be prime, so L cannot be context free ⚲