COT 4420 - Formal Languages and Automata Theory
HW #4
Instructor: Dimitris A. Pados
Teaching Assistant: Andrew Steinberg

Problem 1 - The Empty String
Consider an arbitrary grammar $G = \{V, T, S, P\}$ for which $\lambda \notin L(G)$. Explain how you can formally modify G to include λ in $L(G)$.

Problem 2 - Grammar Simplifications
Consider the grammar with alphabet $T = \{a, b\}$ and production rules
$S \to aC \mid A \mid C$
$A \to a$
$B \to aa$
$C \to aCb$.
Simpilfy the grammar eliminating all useless variables and productions, if any. What is the language of this grammar?

Problem 3 - General CF to Chomsky Grammars
Consider the grammar with alphabet $T = \{a, b\}$ and production rules
$S \to AS \mid ASS$
$A \to SA \mid aa \mid b$.
Convert this grammar to an equivalent (in language) Chomsky grammar.

Problem 4 - The Chomsky Membership Algorithm
Consider the grammar with alphabet $T = \{a, b\}$ and production rules
$S \to AB$
$A \to BA \mid a$
$B \to CC \mid b$
$C \to AB \mid a$.
(i) Is this grammar regular? Why? (ii) Is this grammar linear? Why? (iii) Is this grammar context-free? Why? (iv) Is this grammar Chomsky? Why? (v) Execute on paper the Chomsky Membership Algorithm in complete detail to test the membership of the string $w = baab$ in the language of this grammar.

Problem 5 - Non-deterministic Pushdown Automata
Design a non-deterministic pushdown automaton for the language $L = \{w \in \{a, b\}^* : n_a(w) + 2 = n_b(w)\}$, where $n_a(w)$, $n_b(w)$ denote the number of a's and b's in w, correspondingly.

1) You can modify G to include λ in $L(G)$ by adding the production $S_0 \to S \mid \lambda$ and changing the start variable to S_0. This new grammar is $\hat{G} = (\hat{V}, T, S_0, \hat{P})$ where $\hat{V} = V \cup \{S_0\}$, $\hat{P} = P \cup \{S_0 \to S \mid \lambda\}$. Then, $L(\hat{G}) = L(G) \cup \{\lambda\}$.

2)

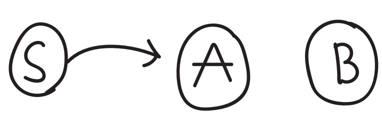
Step 1: Identify all potentially useful variables

S, A, B

Step 2: Remove all non-potentially useful variables

$S \to A$
$A \to a$
$B \to aa$

Step 3: Make a dependency graph



Step 4: Remove potentially useful variables that can not be reached from the start variable

$\hat{V} = \{S, A\}$ $\quad \hat{G} = \{\hat{V}, \hat{T}, S, \hat{P}\}$

$\hat{P}$: $S \to A$ $\quad \hat{T} = \{a\}$ $\quad L(G) = L(\hat{G}) = \{a\}$

$A \to a$

3)

Step 1: Make all productions "variables only" on the right side.

$S \to AS \mid ASS$

$A \to SA \mid N_a N_a \mid b$

$N_a \to a$

Step 2: "Decompose" the variable-only rules so there is only two variables on the right side.

$S \to AS \mid AW_1$

$A \to SA \mid N_a N_a \mid b$

$N_a \to a$

$W_1 \to SS$

Final Chomsky grammar:

$\hat{P}$:

$S \to AS \mid AW_1$

$A \to SA \mid N_a N_a \mid b$

$N_a \to a$

$W_1 \to SS$

$\hat{G} = \{\hat{V}, T, S, \hat{P}\}$ $\quad L(G) = L(\hat{G})$

$\hat{V} = \{S, A, N_a, W_1\}$

4) The grammar is non-regular because it is neither a right-linear or left-linear grammar. This is because the production rules have more than one variable on the right side. The grammar is also not linear because it has more than one variable on the right side. The grammar is context free because it contains a combination of variables and symbols in any order on the right and one variable on the left. The grammar is chomsky because it either has two variables on the right or one symbol on the right and it always has one variable on the left.

$W = baab$

V_{11} V_{22} V_{33} V_{44}

$\{B\}$ $\{A,C\}$ $\{A,C\}$ $\{B\}$

V_{12} V_{23} V_{34}

$\{A\}$ $\{B\}$ $\{S\}$

V_{13} V_{24}

$\emptyset$ $\emptyset$

$S \notin V_{1n}$

V_{14}

$\emptyset$

$\therefore W \notin L(G)$

5) $L = \{ w \in \{a,b\}^* : n_a(w) + 2 = n_b(w) \}$