COMP2805: Solution to Assignment # 3

November 19, 2010

Question 1

Let $\Sigma = 0, 1$. Write CFGs that generate the following languages:

- $\{w | w$ contains at least three 1s$\}$.

  Let $G = (V, \Sigma, R, S)$ be a context-free grammar where:
  $V = \{S, A\}$
  $\Sigma = \{0, 1\}$
  $R =$
  $\{$
  $S \rightarrow A1A1A1A$
  $A \rightarrow AA|1|0|\epsilon$
  $\}$

- $\{w | w$ starts and ends with the same symbol$\}$.

  Let $G = (V, \Sigma, R, S)$ be a context-free grammar where:
  $V = \{S, A\}$
  $\Sigma = \{0, 1\}$
  $R =$
  $\{$
  $S \rightarrow 1A1|0A0|0|1$
  $A \rightarrow AA|0|1|\epsilon$
  $\}$

- $\{w |$ the length of w is odd$\}$.

  Let $G = (V, \Sigma, R, S)$ be a context-free grammar where:
  $V = \{S\}$
  $\Sigma = \{0, 1\}$
  $R =$
  $\{$
  $S \rightarrow 0|1|0S0|0S1|1S0|1S1$
  $\}$

- $\{w |$ the length of w is odd and its middle symbol is 0$\}$.

Let $G = (V, \Sigma, R, S)$ be a context-free grammar where:
$V = \{S\}$
$\Sigma = \{0, 1\}$
$R =$
$\{$
$S \rightarrow 0|0S0|0S1|1S0|1S1$
$\}$
(*Same as the previous solution, but we removed* $S \rightarrow 1$ *because of the constraint on the middle symbol*)

- $\{w|w$ contains more 1s than 0s$\}$.

Let $G = (V, \Sigma, R, S)$ be a context-free grammar where:
$V = \{S, A\}$
$\Sigma = \{0, 1\}$
$R =$
$\{$
$S \rightarrow 0S1|1S0|S01|S10|01S|10S|A$
$A \rightarrow 1A|1$
$\}$
(*rule* S *make sure to have a* 1 *for each* 0 *than rule* A *is there to add at least one more* 1)

- $\{w|w = w^R$; that is, w is a palindrome of **even** length $\}$.

Let $G = (V, \Sigma, R, S)$ be a context-free grammar where:
$V = \{S\}$
$\Sigma = \{0, 1\}$
$R =$
$\{$
$S \rightarrow 0S0|1S1|\epsilon$
$\}$

Question 2

Using the pumping lemma for Regular Languages, prove that the following language, with the alphabet $\Sigma = \{\langle, \rangle\}$, is not Regular:
$L = \{\langle^N\rangle^M, N < M\}$

NOTE: We change (for $\langle$ and) for $\rangle$ for the readability.

We will prove by contradiction that L is not Regular. Assume L is regular. Let $p \geq 1$ be the pumping length, as given by the pumping lemma. Consider the string $s = \langle^p \rangle \rangle^p$. Clearly, this string is in L and its length exceeds the pumping length. Therefore, it can be rewritten as $s = xyz$, where $y \neq \epsilon$, $|xy| \leq p$, and $xy^iz \in L, \forall i \geq 0$.

Observe that, since $|xy| \leq p$, the string y will contain only $\langle$'s. Then, consider $s = xy^2z$. The number of $\langle$'s in s is equal to or greater than the number of $\rangle$'s in s, therefore $xy^2z \notin S$. However by the pumping lemma it is in L. Therefore we have a contradiction, and L is not Regular.

Question 3

Let A and B be context-free languages over the same alphabet Σ. Prove that:

1. **Prove that the union A U B of A and B is also context-free.**
2. **Prove that the concatenation AB of A and B is also context-free.**
3. **Prove that the star A* of A is also context-free.**

Suppose that the language A has the Context Free Grammars $G_1 = (V_1, \Sigma, R_1, S_1)$ and B has the CFG $G_2 = (V_2, \Sigma, R_2, S_2)$, where V_1 and V_2 are the set of variables, Σ is the set of terminals, S_1 and S_2 are the start variables, R_1 and R_2 are the set of rules.

The proof is by construction. So, we construct CFGs that can recognize $A \cup B$, $A \cdot B$, and A^*.

1. $A \cup B$
   $G_3 = (V_3, \Sigma, R_3, S_3)$. Here $V_3 = V_1 \cup V_2 \cup \{S_3\}$, and $R_3 = R_1 \cup R_2 \cup \{S_3 \to S_1, S_3 \to S_2\}$.
2. $A \cdot B$
   $G_3 = (V_3, \Sigma, R_3, S_3)$. Here $V_3 = V_1 \cup V_2 \cup S_3$, and $R_3 = R_1 \cup R_2 \cup \{S_3 \to S_1 S_2\}$.
3. A^*
   $G = (V, \Sigma, R, S)$. Here $V = V_1 \cup S$, and $R = R_1 \cup \{S \to S_1 S, S \to \epsilon\}$.

Since CFGs that can recognize $A \cup B$, $A \circ B$, and A^* exist, CFL is closed under union, concatenation, and kleene-star.

Question 4

Define the following two languages $A = \{a^m b^n c^n : m \geq 0, n \geq 0\}$ **and** $B = \{a^m b^m c^n : m \geq 0, n \geq 0\}$**.**

- Prove that both A and B are context-free, by constructing two grammars, one that generates A and one that generates B.

  Let $G_A = (V, \Sigma, R, S)$ be a context-free that generates A:
  $V = \{S, A, B\}$
  $\Sigma = \{a, b, c\}$
  $R =$
  $\{$
  $S \to BA$
  $A \to bAc | \epsilon$
  $B \to aB | \epsilon$
  $\}$

  Let $G_B = (V, \Sigma, R, S)$ be a context-free that generates B:
  $V = \{S, A, B\}$
  $\Sigma = \{a, b, c\}$
  $R =$
  $\{$
  $S \to AB$
  $A \to aAb | \epsilon$
  $B \to cB | \epsilon$
  $\}$

- We have seen that the language $\{a^n b^n c^n : n \geq 0\}$ is not context-free. Explain why this implies that the intersection of two context-free languages is not necessarily context-free.

$A \cap B = \{a^m b^n c^n : m \geq 0, n \geq 0\} \cap \{a^m b^m c^n : m \geq 0, n \geq 0\}$
$= \{a^m b^k c^n : k = n = m, m \geq 0, n \geq 0\}$
$= \{a^n b^n c^n : n \geq 0\}$

We know A and B are CFG because we defined their grammars in part 1. We know $A \cap B$ is $\{a^n b^n c^n : n \geq 0\}$ and is not context-free (from Section 3.8.2). Therefore there exists an intersection of two CFG that is not context free.

- Use De Morgans Law to conclude that the complement of a context-free language is not necessarily context-free.

  Recall that the complement of some context free grammars are context free (*ex.: $\overline{A}$ and $\overline{B}$ are CFGs*) and the union of two context free grammars results in a context free grammar (*therefore $\overline{A} \cup \overline{B}$ is CFG*).

  $\overline{(\overline{A} \cup \overline{B})} = (A \cap B) = \{a^n b^n c^n : n \geq 0\}$ is not context-free

  Therefore there exists a complement of a context-free language is not context-free.

Question 5

Convert the following Context Free Grammars to the Chomsky Normal Form.

1. $S \longrightarrow SS$; $S \longrightarrow (S)$; $S \longrightarrow \epsilon$, where$\Sigma = \{(,)\}$

   Let $G_1 = (V, \Sigma, R, S')$ be a context-free grammar where:
   $V = \{S', S, A, L, R\}$
   $\Sigma = \{(,)\}$
   $R =$
   $\{$
   $S' \rightarrow S$
   $S \rightarrow SS|LA|LR$
   $A \rightarrow SR$
   $L \rightarrow ($
   $R \rightarrow)$
   $\}$

2. $S \longrightarrow aSb$; $S \longrightarrow bSa$; $S \longrightarrow \epsilon$, where $\Sigma = \{a, b\}$

   Let $G_2 = (V, \Sigma, R, S')$ be a context-free grammar where:

$V = \{S', S", S, A, B, \}$
$\Sigma = \{a, b\}$
$R =$
$\{$

$S' \rightarrow S$
$S \rightarrow AB$
$S \rightarrow BA$
$S \rightarrow AS"$
$S" \rightarrow SB$
$S \rightarrow BS"$
$S" \rightarrow SA$
$B \rightarrow b$
$A \rightarrow a$

$\}$

Question 6

Find a CFG generating $L = \{0^m 1^n 0^m 1^k | m, n, k \geq 0\}$. Give straightforward arguments to show that your answer is right.

$G(L) = \{V, \Sigma, R, S\}$

$V = \{S, A, B\}$
$\Sigma = \{0, 1\}$
$R =$
$\{$

$S \rightarrow A|BA$
$A \rightarrow 1A|\epsilon$
$B \rightarrow 0B0|A$

$\}$

We can observe that the number of zeros is maintained by variable B. Both part with ones are completely independant from each other and controlled by variable A. Finally, m, n and k can all be equal to zero because of the rule $A \rightarrow \epsilon$.

Question 7

If a Context Free Grammar G is in the Chomsky Normal Form, show that for any string w in $L(G)$, exactly $2|w| - 1$ steps are required for any derivation of w.

All rules in the Chomsky Normal Form are of the form:

- $X \rightarrow AB$, where $X, A, B \in V$ and $S \notin \{A, B\}$
- $X \rightarrow a$, where $X \in V$ and $a \in \Sigma$
- $S \rightarrow \epsilon$, where S is the starting rule

In order to generate the string w (*assuming that* $w \neq \epsilon$), we need to apply the first rule $|w| - 1$ times. The string w now contains a sequence of concatenated variables, where each variable corresponds to the second rule. Apply the second rule to each variable. Since there are $|w|$ variables, we need to apply the second rule $|w|$ times. In total, it took $|w| - 1 + |w| = 2|w| - 1$ steps to generate a string w.