Quiz 2 Solutions

Problem 1 [20 points] Let $G = (\{S, A\}, \{0, 1\}, R, S)$ be the CFG having as rules

\[
S \rightarrow 0A1 \\
A \rightarrow AA \mid 10
\]

Specify a CFG $G'$ in Chomsky Normal Form such that $L(G') = L(G)$ and $G'$ has at most 6 rules.

We can apply the transforms we saw in class or are in the proof of Theorem 2.6 of the textbook. But we don’t need all of them. We introduce a new variable $B$ to stand for $0A$. We also introduce variables $Z, W$ to stand for $0, 1$ respectively. Our grammar is now $G' = (\{S, A, B, Z, W\}, \{0, 1\}, R', S)$ where the rules are

\[
S \rightarrow BW \\
B \rightarrow ZA \\
A \rightarrow AA \mid WZ \\
Z \rightarrow 0 \\
W \rightarrow 1
\]

Problem 2 [20 points] Recall that $|w|$ denotes the length of a string $w$. Let

\[
A = \{ w \in \{0, 1\}^* : |w| \text{ is even} \}.
\]

Specify a CFG $G$ such that $L(G) = A$ and $G$ has at most six rules.

Let $G = (\{S, T\}, \{0, 1\}, R, S)$ be the CFG having the following four rules:

\[
S \rightarrow TTSS \mid \varepsilon \\
T \rightarrow 0 \mid 1
\]

Problem 3 [20 points] Let $A$ be the following language over the alphabet $\{0, 1\}$:

\[
A = \{ 1^n0^{n+m}1^m : n, m \geq 0 \}.
\]

Draw the state diagram of a PDA $M$ such that $M$ recognizes $A$ and $M$ has at most six states.

The PDA is depicted in Figure 1.
Figure 1: PDA for Problem 3. Here “e” stands for \( \varepsilon \). The state with the boldface outline is the final state.

---

**Problem 4 [20 points]** Let

\[
A = \{ 0^{n^2} : n \geq 0 \}
\]

Prove that \( A \) is not a context-free language.

This is simpler than the examples we have seen in class or the homeworks because there is only one case to consider and also the proof is very similar to the proof we did in class that \( \{ 0^{2n} : n \geq 0 \} \) is not regular.

We use the usual template exemplified both in class and in the solutions to Problem Set 2. The proof is by contradiction.

**Assume:** \( A = \{ 0^{n^2} : n \geq 0 \} \) is a CFL.

The assumption means that the pumping lemma (Theorem 2.19, page 115 of the text) applies to \( A \). We imagine ourselves “interacting” with the lemma as follows:
We give it \( A \), and it returns a pumping length \( p \). Now, we choose a string \( s \in A \) of length greater than \( p \), and return it to the lemma. The choice is \( s = 0^{p^2} \). Because \( s \) is in \( A \) and has a length greater than \( p \), the pumping lemma says that \( s \) can be split into \( uvxyz \) which obey the three conditions of the pumping lemma. The lemma returns \( u, v, x, y, z \) to us. We then choose \( i = 2 \) and return it to the lemma. At this point, the lemma guarantees that

1. \( uv^2xy^2z \in A \)
2. \(|vy| > 0\)
3. \(|vxy| \leq p\)

Now our goal is to show that conditions (2) and (3) imply that \( uv^2xy^2z \) cannot be in \( A \), contradicting condition (1) above. This means our assumption that \( A \) was a CFL is false.

We do not know where \( vxy \) sits in the string \( s = uvxyz \), but since \( s \) is all zeros, so is \( vxy \). This means there exist integers \( i, j \) such that \( v = 0^i \) and \( y = 0^j \). Conditions (2) and (3) tell us that \( 1 \leq i + j \leq p \). Now

\[
|uv^2xy^2z| = |uvxyz| + |vy| = |s| + i + j = p^2 + i + j.
\]

For \( uv^2xy^2z \) to be in \( A \), it must be that \( |uv^2xy^2z| = q^2 \) for some integer \( q \). However, since \( 1 \leq i + j \leq p \) we have from the above

\[
p^2 + 1 \leq |uv^2xy^2z| \leq p^2 + p.
\]

But \( p^2 + p < p^2 + 2p + 1 = (p + 1)^2 \), so from the above we get

\[
p^2 < |uv^2xy^2z| < (p + 1)^2.
\]

This shows that there is no integer \( q \) such that \( |uv^2xy^2z| = q^2 \).
**Problem 5 [20 points]** To languages $A, B \subseteq \Sigma^*$ we associate the language

$$A \oplus B = \{ w \in \Sigma^* : \text{w \in A or w \in B but not both} \}.$$ 

Prove that the class of context-free languages is *not* closed under the $\oplus$ operation. In other words, prove that the following statement is false:

**Statement X**: For all languages $A, B$: if $A, B$ are CFLs, then so is $A \oplus B$.

Let us consider:

**Statement C**: For all languages $C$: if $C$ is a CFL then so is $\overline{C}$.

Now we proceed as follows:

**Assume**: Statement X is true

**Claim**: Statement C is true.

If we can prove this claim, it concludes the proof, because we know that statement C is false. So we proceed to show that the claim is true assuming statement X is true.

Let $C$ be a CFL. Assuming statement X, we want to show that $\overline{C}$ is also a CFL. We note that $\overline{C} = C \oplus \Sigma^*$. But $C$ is a CFL by assumption, and $\Sigma^*$ is certainly a CFL, so statement X says $\overline{C}$ is a CFL, as desired.

The template here is the one given in class to show that the class of CFLs is not closed under complement. Recall we did it by showing that if it were, it would also be closed under intersection, thereby contradicting the fact that it is not closed under intersection.