Computer Science 452 - Homework Assignment #5

Hari Amoor, NetID: hra25

February 18, 2020

Problem 2.11: Supply a PDA for a given CFG G_4.

We supply this equivalent CFG G in Greibach-Normal Form:

$$\begin{aligned} S &\to E \\ E &\to K \mid T \\ T &\to L \mid M \mid Y \mid Z \\ K &\to E \mid T \\ Y &\to (E) \\ Z &\to a \end{aligned}$$

Now, we generate a PDA P for G. P has only one state q, and the start symbol of P is the start symbol S of G. Furthermore, the input symbols in P are the terminal symbols $a, (,)$, in G and the stack symbols will be the non-terminal symbols iN G. The transition rules are as follows:

$$\begin{aligned} \delta(q, \epsilon, S) &= (q, E) \\ \delta(q, \epsilon, E) &= \{(q, K), (q, T)\} \\ \delta(q,), E) &= (q, \epsilon) \\ \delta(q, \epsilon, T) &= \{(q, K), (q, T), (q, Y), (q, Z)\} \\ \delta(q, \epsilon, K) &= \{(q, E), (q, T)\} \\ \delta(q, (, Y) &= (q, E) \\ \delta(q, a, Z) &= (q, a) \end{aligned}$$

Problem 2.20, Claim: If A is context-free and B is regular, then $A \setminus B$ as defined is context-free.

Proof. Let X be a PDA for A and Y a finite automaton for B. We provide a PDA Z that recognizes $A \setminus B$.

We initialize Z with $Q_Z = Q_X \times Q_Y$. First, Z reads the input word w and advances in X, ignoring Y. When and if the Z reaches the end of w, i.e. by coming to a success state in X, Z guesses a word x that will continue to advance in X and start to advance in Y; in order for Z to accept w, x has to reach an acceptance state in both automata simultaneously.

If there is an accepting run of Z, then $w \in A \setminus B$ since it guessed a witness x. Similarly, if there exists an x satisfying $wx \in A$, then it will be discovered by Z. Thus, since Z exists, it must be the case that $A \setminus B$ is context-free. □

Problem 2.25, Claim: The set of CFLs is closed under the SUFFIX operator, as defined.

Proof. Let A be a CFL and H be a CFG that generates A. We define a CFG L for SUFFIX(A) as follows:

Suppose w.l.o.g. that H is in Chomsky-Normal Form. First, for each variable X in H, include variables X and X' in L. Now:

1. For every rule in H of the form $X \to YZ$, include the following rules into L:

$$X \to YZ$$
$$X' \to Y'Z \mid Z'$$

2. For every rule in H of the form $X \to \sigma$, include the following rules into L:

$$X \to \sigma$$
$$X' \to \sigma \mid \epsilon$$

The CFL generated by L is exactly equal to SUFFIX(A), as required. □

Problem 2.30(a, d): Use the pumping lemma to show that the following languages are not context-free.

1. **The given language A.**

   Proof. Suppose for contradiction that A is context-free. The string $s = 0^k 1^k 0^k 1^k$ is in A, where k is the pumping length of A. Since s is sufficiently large, the decomposition $s = uvxyz$ must exist as per the pumping lemma.

   Now, the following cases are exhaustive:

   Case 1: vxy contains all 0s and is contained within the first or second string of 0s.

   Here, since $|vy| > 0$, either v or y must contain at least one 0. Now, observe that $uv^0xy^0z \notin A$; this contradicts the pumping lemma. **Case 2:** vxy contains all 1s and is contained within the first or second string of 1s.

   By the same reasoning in Case 1, it is clear that a contradiction can be derived.

   Case 3: vxy contains both 0s and 1s.

   Suppose w.l.o.g. that vxy contains 0s followed by 1s. Then, vxy straddles either the first division between 0s and 1s or the second division. Now, because $|vxy| \leq k$, it holds that pumping up or down will only affect the substrings immediately adjacent to the division being straddled; the other two substrings would be unaffected. Thus, pumping s will result in a string not in the language, which contradicts the pumping lemma.

   Consequently, the claim holds. □

2. **The given language L.**

Proof. Suppose for contradiction that L is context-free, and let p denote its pumping length. Consider $s = 0^p1^p\#0^p1^p \in L$. The following cases are exhaustive:

Case 1: vxy does not contain the # symbol.

Here, it must be the case that vxy is entirely on one side of the # symbol; then, pumping v and y will result in a string not in L, which is a contradiction.

Case 2: Either v or y contains the symbol # (assume v w.l.o.g.).

Here, $uv^0xy^0z \notin L$ because it does not contain the # symbol. This contradicts the pumping lemma.
Case 3: x contains the symbol #.

Here, v is a substring of the first string of 1s, and y is a substring of the second string of 0s. Thus, $uv^0xy^0z \notin L$ because there is either a smaller number of 1s in t_1 or a larger number of 0s in t_2. This is, therefore, a contradiction.

Thus, the claim holds. □

Problem 2.32, Claim: The language C, as given, is not context-free.

Proof. Suppose for contradiction that C is context-free with pumping length p. Let $s = 1^p3^p2^p4^p$. Clearly, the decomposition $s = uvxyz$ must exist.

W.l.o.g., suppose vxy includes the symbol 1. Then, it cannot include 2. With this, $uv^2xy^2z \notin C$, because the number of 1s in the same is not equal to the number of 2s.

This directly contradicts the pumping lemma. Thus, the claim holds. □

Problem 2.46: Show that the grammar G, as given is ambiguous, and provide an unambiguous grammar for the same.

Let L be the language of strings of the form a^nb^n. $L(G)$ is the language $\{x_1x_2 \ldots x_k \mid x_i \in L\}$.

Consider the string $s = x_1x_2x_3$ where $x_i = ab$. Here, s can be parsed either as two separate subexpressions (x_1x_2, x_3) or as two separate subexpressions (x_1, x_2x_3). Thus, since s has two separate parse trees, the claim holds that this is ambiguous.

An unambiguous CFG H is provided as follows:

$$S \to TS \mid T$$
$$T \to aTb \mid ab$$

In order to prove that H is unambigious, we let $s \in \Sigma^\star$ be arbitrary, and show that s has only one parse tree under H.

Problem 2.48: Prove the following.

(a) **Claim: C_1, as given, is a CFL.**

Proof. The following CFG generates C_1:

$$S \to T1T$$
$$T \to 0T \mid 1T \mid 0 \mid 1$$

Thus, the claim holds that C_1 is context-free. □

(b) **Claim: C_2 is not context-free.**

Proof. Suppose for contradiction that C_2 is a CFL, and let p be its pumping length. Let $s = w10^pw$, where $w = 0^{p+2}$. Then, the decomposition $s = uvxyz$ is well-defined w.r.t. the pumping lemma.

The following cases are exhaustive:

Case 1: vxy contains one, but not both, of the 1s in s.

Here, if the 1 is in x, we are done; with v and y pumped to a sufficiently large length, we can generate a string not in C_2. So, suppose w.l.o.g. that the 1 is in v. Then, clearly $uv^2xy^2z \notin C_2$, because you can pump at most two of the strings of 0s. This contradicts the pumping lemma.

Case 2: vxy does not contain either of the 1s in s.

Pumping v and y here will result in arbitrarily many 0s; thus, when done to a sufficiently large length, one can generate a string outside of C_2, which is a contradiction.

Thus, the claim holds. □