Solutions to Assignment 5

October 27, 2000

Exercise 1 *(40 points) Give context-free grammars generating the following languages.*

1. *The set of strings over the alphabet $\Sigma = \{a, b\}$ with twice as many a's as b's.*
2. *The complement of the language $\{a^n b^n \mid n \geq 0\}$.*
3. $\{w\#x \mid w^R$ *is a substring of* x *for* $w, x \in \{0, 1\}\}$.
4. $\{w \mid w = w^R$ *for* $w \in \{0, 1\}^*\}$.

Solution We present the context-free grammar rules for each case.

1.

$$S \rightarrow SS \mid aaSb \mid abSa \mid baSa \mid \epsilon$$

2.

$$\begin{aligned} S &\rightarrow T \mid aU \mid Vb \\ T &\rightarrow aT \mid Ta \mid bT \mid Tb \mid ba \\ U &\rightarrow aU \mid W \\ V &\rightarrow Vb \mid W \\ W &\rightarrow aWb \mid \epsilon \end{aligned}$$

3.

$$\begin{aligned} S &\rightarrow S0 \mid S1 \\ S &\rightarrow T \\ T &\rightarrow 0T0 \mid 1T1 \\ T &\rightarrow U \\ U &\rightarrow U0 \mid U1 \mid \# \end{aligned}$$

4.

$$S \rightarrow 0S0 \mid 1S1 \mid 0 \mid 1 \mid \epsilon$$

■

Exercise 2 *(15 points) Give a context-free grammar that generates the language.*

$$A = \{a^i b^j c^k \mid i, j, k \geq 0 \text{ and either } i = j \text{ or } j = k\}$$

Is your grammar ambiguous? Why or why not?

Solution Let $G = (V, \Sigma, S, R)$, where $V = \{S, U, V, X, Y\}$, $\Sigma = \{a, b, c\}$ and R consists of the following rules.

$$\begin{array}{rcl} S & \rightarrow & U \mid V \\ U & \rightarrow & aU \mid X \\ V & \rightarrow & Vc \mid Y \\ X & \rightarrow & bXc \mid \epsilon \\ Y & \rightarrow & aYb \mid \epsilon \end{array}$$

The context-free grammar G is ambiguous. For instance, we have $S \Rightarrow U \Rightarrow aU \Rightarrow abUc \Rightarrow abc$ and $S \Rightarrow V \Rightarrow Vc \Rightarrow aYbc \Rightarrow abc$, which are two different derivations of abc. ■

Exercise 4 *(35 points) Use the pumping lemma to show that the following languages are not context-free.*

1. *(15 pts)* $L_1 = \{0^n\#0^{2n}\#0^{3n} \mid n \geq 0\}$
2. *(20 pts)* $L_2 = \{w\#x \mid w$ *is a substring of* x *for* $w, x \in \{0, 1\}\}$.

Solution Assume L_1 is context-free. Let p be its pumping length. and $s = 0^p\#0^{2p}\#0^{3p}$. By the pumping lemma for context-free languages, we have $s = uvxyz$ satisfying the following.

- $uv^ixy^iz \in L_1$ for $i = 0, 1, 2, \ldots$, and
- $|vy| > 0$, and
- $|vxy| \leq p$.

We do a case analysis here.

- Either v or y contains #. Then $uvvxyyz$ contains more than two #'s, which contradicts $uvvxyyz \in L_1$.
- $v = 0^{k_1}$ and $y = 0^{k_2}$ for some $k_1, k_2 \in \mathcal{N}$. We have $k_1 + k_2 > 0$ since $|vy| > 0$.
  - vxy occurs before the second # in s. Then there are $3p + k_1 + k_2$ 0's in $uvvxyyz$ before the second # and only $3p$ $0's$ after, implying $uvvxyyz \notin L_1$.
  - vxy occurs after the first # in s. Then there are $5p + k_1 + k_2$ 0's in $uvvxyyz$ after the first # and only p $0's$ before, implying $uvvxyyz \notin L_1$.

Therefore, L_1 is not context-free.

Assume that L_2 is context-free. Let p be the pumping length of L_2 and $s = 10^p\#10^p$. By the pumping lemma for context-free languages, we have $s = uvxyz$ satisfying the following.

- $uv^ixy^iz \in L_1$ for $i = 0, 1, 2, \ldots$, and
- $|vy| > 0$, and
- $|vxy| \leq p$.

We do a case analysis here.

- Either v or y contains #. Then $uvvxyyz$ contains more than one #'s, which contradicts $uvvxyyz \in L_1$.

- vxy occurs before #. Then $uvvxyyz = w_1\#w_2$ for some w_1, w_2 such that w_1 is longer than w_2. Therefore, w_1 cannot be a substring of w_2, contradicting $uvvxyyz \in L_2$.
- vxy occurs after #. Then $uxz = w_1\#w_2$ for some w_1, w_2 such that w_1 is longer than w_2. Therefore, w_1 cannot be a substring of w_2, contradicting $uxz \in L_2$.
- x straddles #. We have two subcases.
  - y begins with 1. Then $uxz = w_1\#w_2$ for some w_1, w_2 such that w_1 begins with 1 and w_2 consists of 0's. Therefore, w_1 cannot be a substring of w_2, contradicting $uxz \in L_2$.
  - $y = \epsilon$. Then $uvvxyyz = w_1\#w_2$ for some w_1, w_2 such that w_1 is longer than w_2. Therefore, w_1 cannot be a substring of w_2, contradicting $uvvxyyz \in L_2$.

Therefore, L_2 is not context-free. ■