Homework 5

Aleck Zhao

March 29, 2018

1. In class we showed that the class of context free languages is closed under union. This question studies intersection and complement of context free languages. Consider the languages $A = \{a^m b^n c^n : m, n \geq 0\}$ and $B = \{a^n b^n c^m : m, n \geq 0\}$.

   (a) Give a context-free grammar for each of A and B. Then use A and B to show that the class of context free languages is not closed under intersection.

   Solution. For A, we have the rules

$$\begin{aligned} S &\to aS \mid ST \mid \varepsilon \\ T &\to bTc \mid \varepsilon \end{aligned}$$

   and for B, we have the rules

$$\begin{aligned} S &\to aSb \mid ST \mid \varepsilon \\ T &\to cT \mid \varepsilon \end{aligned}$$

   Then $A \cap B = \left\{a^k b^k c^k : k \geq 0\right\}$, which is not a context-free language, so the class of context free languages is not closed under intersection. $\square$

   (b) Use (a) and DeMorgan's Law to show that the class of context-free languages is not closed under complementation.

   Proof. Suppose the class of CFLs was closed under complement. Let C and D be CFLs. Then by assumption C^c and D^c are both CFLs, and thus $C^c \cup D^c$ is a CFL since CFLs are closed under union. By DeMorgan's Law, this is $(C \cap D)^c$ and thus its complement $C \cap D$ is also a CFL. However, if we take $C = A$ and $D = B$ from above, we have a contradiction since $A \cap B$ is not a CFL. Thus, the class of CFLs is not closed under complement. $\square$

2. Let $D = \{xy : x, y \in \{0,1\}^* \text{ and } |x| = |y| \text{ but } x \neq y^{\mathcal{R}}\}$. Give a context-free grammar for D, and formally prove that your grammar generates the given language.

Solution. D is all even-length non-palindromes. We claim that the following rules generate D :

$$S \to 0S0 \mid 1S1 \mid 1T0 \mid 0T1$$
$$T \to 00T \mid 01T \mid 10T \mid 11T \mid \varepsilon$$

Claim 1: every string generated by this grammar is in D. Proceed by induction on the number of steps n in a derivation. For the base case, $n = 2$, since we cannot terminate in 1 step. For $n = 2$, the possible steps taken are $S \to 1T0 \to 10$ or $S \to 0T1 \to 01$, which are both in D. Now, for the inductive hypothesis, suppose that all strings s derivable from the grammar in at most N steps are in D, where $N \geq 2$. Let s' be a string that is derived in $N + 1$ steps.

Now, if the first step to derive s' was either $S \to 0S0$ or $S \to 1S1$, then for the remaining N steps, S is mapped to an even length non-palindrome by hypothesis, and thus $0S0$ and $1S1$ are both even length non-palindromes, and thus in D. If the first step was $S \to 1T0$ or $S \to 0T1$, then s' is not a palindrome since its first and last symbols are different. s' also has even length because each rule for T always adds 2 or 0 symbols. Thus, $s' \in D$, so the claim is proved by induction.

Claim 2: every string in D can be generated by this grammar. Proceed by induction of the length ℓ of the string, which can only be even. The base case is $\ell = 2$ since for $\ell = 0$, the empty string is palindromic. For $\ell = 2$, the strings are 01 and 10, which can be generated as $S \to 0T1 \to 01$ and $S \to 1T0 \to 10$, respectively. Now, for the inductive hypothesis, suppose that all strings in D of length at most $2N$ can be generated by this grammar, where $N \geq 1$. Let $s' \in D$ be a string of length $2N + 2$.

If s' starts and ends with 0, then it can be written as $0s0$, where s is an even length non-palindrome since $s' \in D$, and in particular $|s| = 2N$, so s can be generated by the grammar, so s' can be generated by this grammar with the first step $S \to 0S0$. The case with s' starts and ends with 1 is identical.

If s' starts with 0 and ends with 1, then it can be written as $0s1$, where $|s| = 2N$. The first step must have been $S \to 0T1$, and now the rules for T generate every even length string (this is obvious). Thus, s' can be generated by this grammar. The case when s' starts with 1 and ends with 0 is proved similarly, so the claim is proved by induction. □

3. Prove that the language $L = \left\{0^{2^n} : n \geq 0\right\}$ is not context-free.

Proof. Suppose L was context-free. Then by the pumping lemma, there exists an integer p that is the pumping length. Consider the string $s = 0^{2^{2p}} = uvxyz$. Then since $|vxy| \leq p$, it follows that $0 < |vy| \leq |vxy| \leq p$. Now let

$$s' = uv^0xy^0z = uxz$$

where

$$|s'| = 2^{2p} - |vy| \implies 2^{2p} > |s'| \geq 2^{2p} - p$$

Now, in order to have $s' \in L$, we must have $|s'| = 2^{2p-k}$ for some $k \geq 1$. That is,

$$2^{2p-k} \geq 2^{2p} - p \implies p \geq 2^{2p} - 2^{2p-k} \geq 2^{2p} - 2^{2p-1} = 2^{2p-1}$$

However, for all $p \geq 0$, this is not true (easy to see, or by induction), thus there is no satisfactory pumping length, so L is not context-free. $\square$

4. Let B be the language of all palindromes over $\{0,1\}$ containing an equal number of 0s and 1s. Show that B is not context-free.

Proof. Suppose B was context-free. Then by the pumping lemma, there exists an integer p that is the pumping length. Consider the string $s = 0^p 1^{2p} 0^p = uvxyz$. There are 2 cases.

Case 1: v and y both contain only 1s. Here, $s' = uv^2xy^2z$ will have more 1s than 0s, so $s' \notin B$.

Case 2: either v or y contains a positive number of 0s. Any 0s contained in vxy must come from the same block of 0s, since $|vxy| \leq p$ so the middle block of 1s can't be "crossed." Then in $s' = uv^2xy^2z$, the number of 0s in the beginning and ending blocks won't be the equal anymore, so s' will not be a palindrome.

Thus, B is not context-free. $\square$