Solutions for Homework 6
CIS 770: Formal Language Theory

Problem 1. [Category: Design] If A and B are languages, define $A \diamond B = \{xy \,|\, x \in A,\ y \in B \text{ and } |x| = |y|\}$. Show that if A and B are regular languages, then $A \diamond B$ is context free. If you construct a CFG or PDA for $A \diamond B$, you need not prove that your construction is correct, but your intuitions behind the construction should be clearly spelt out. **[10 points]**

Solution: Closure Properties Proof: If A and B are regular languages over alphabet Σ, then $L_1 = A \circ \{\#\} \circ B = \{x\#y \,|\, x \in A \text{ and } y \in B\}$ (where $\# \notin \Sigma$) is also regular because regular languages are closed under concatenation. Consider $L_2 = \{x\#y \mid x \in \Sigma^*,\ y \in \Sigma^* \text{ and } |x| = |y|\}$ is context-free because L_2 is generated by the grammar $G = (\{S\}, \Sigma \cup \{\#\}, R, S)$ where

$$R = \{S \to aSb \mid a, b \in \Sigma\} \cup \{S \to \#\}$$

Since L_1 is regular and L_2 is context-free, the language $L_3 = L_1 \cap L_2$ is context-free. Observe that

$$L_3 = L_1 \cap L_2 = \{x\#y \mid x \in A,\ y \in B, \text{ and } |x| = |y|\}$$

Finally, consider the homomorphism $h : (\Sigma \cup \{\#\})^* \to \Sigma^*$ where $h(a) = a$ for $a \in \Sigma$ and $h(\#) = \epsilon$. Then,

$$h(L_3) = \{xy \mid x \in A,\ y \in B \text{ and } |x| = |y|\} = A \diamond B$$

Thus, $A \diamond B$ is context-free.

Proof by Construction: If A and B are regular then there are DFAs $M_A = (Q_A, \Sigma, \delta_A, q_A, F_A)$ and $M_B = (Q_B, \Sigma, \delta_B, q_B, F_B)$ such that $\mathbf{L}(M_A) = A$ and $\mathbf{L}(M_B) = B$. Using M_A and M_B, we will construct a PDA P to recognize $A \diamond B$. The PDA will work as follows. It will read the input and simulate M_A on the string. As it reads the symbols, it will count the length of the string read by pushing symbols onto the stack. At some point, P will nondeterministically decide that it has read exactly half the input. At this point, it will check that M_A is indeed in an accept state (as this means that the input it has read so far is in A), and then move to the initial state of M_B to start simulating M_B on the remaining input. As it reads the remaining input it will pop symbols from the stack to ensure that it does indeed read a string of length equal to the part on which M_A was simulated. When it is finished reading the entire input, it will check that M_B is in an accept state and that the stack is empty.

The formal construction is as follows. $P = (Q, \Sigma, \Gamma, \delta, q_0, F)$ where

- $Q = Q_A \cup Q_B \cup \{q_0, q_F\}$ where $q_0, q_F \notin Q_A \cup Q_B$
- $\Gamma = \{\$, 0\}$; \$ will be used as a bottom of stack symbol and 0 will be used to count the first half.
- $F = \{q_F\}$
- δ is given as follows

$$\delta(q, a, x) = \begin{cases} \{(q_A, \$)\} & \text{if } q = q_0,\ a = x = \epsilon \\ \{(q', 0)\} & \text{if } q \in Q_A,\ a \in \Sigma,\ q' = \delta_A(q, a), \text{ and } x = \epsilon \\ \{(q_B, \epsilon)\} & \text{if } q \in F_A,\ a = x = \epsilon \\ \{(q', \epsilon)\} & \text{if } q \in Q_B,\ a \in \Sigma,\ q' = \delta_B(q, a), \text{ and } x = 0 \\ \{(q_F, \epsilon)\} & \text{if } q \in F_B,\ a = \epsilon, \text{ and } x = \$ \\ \emptyset & \text{otherwise} \end{cases}$$

$\blacksquare$

Problem 2. [Category: Proof] Let B be the language of all palindromes over $\{0,1\}$ containing an equal number of 0s and 1s. Prove that B is not context-free. **[10 points]**

Solution: Pumping Lemma Proof: We will show that B does not satisfy the pumping lemma. Let p be any pumping length. Consider the string $z = 0^p1^p1^p0^p$; since z is a palindrome and has equal number of 0s and 1s, $z \in B$. Let u, v, w, x, y be any division of z such that (a) $z = uvwxy$, (b) $|vx| > 0$, and (c) $|vwx| \leq p$.

We will argue that $uv^2wx^2y \notin B$. We have to consider a few possibilites based on the form of v and x.

1. Both v and x contain only 0s. Then, since $|vx| > 0$, uv^2wx^2y will have more 0s than 1s, and so will not be a member of B.
2. Both v and x contain only 1s. Then, analogous to the previous case, $uv^2wx^2y \notin B$ because it has more 1s than 0s.
3. Suppose one among v and x has both 0 and 1 symbols, or v and x have symbols of opposite kinds. In that case, uv^2wx^2y is not a palindrome.
4. The last case, when both v and x contain both 0 and 1 symbols is not possible because we have $|vwx| \leq p$.

Closure Properties Proof: Let $L_1 = B \cap \mathbf{L}(0^*1^*0^*)$. Observe that any string in L_1 must begin and end with the same number of 0s (as strings in B are palindromes) and the number of 1s must be equal to the number of 0s (as this is a requirement for strings in B). Thus, $L_1 = \{0^n1^{2n}0^n \mid n \geq 0\}$. Consider a homomorphism $h : \{a, b, c\}^* \to \{0, 1\}^*$ defined as: $h(a) = 0$, $h(b) = 11$, $h(c) = 0$. Then

$$L_2 = h^{-1}(L_1) \cap \mathbf{L}(a^*b^*c^*) = \{a^nb^nc^n \mid n \geq 0\}$$

Since L_2 is not a CFL and is constructed from B using operations that preserve context-freeness, we can conclude that B is not context-free. $\blacksquare$

Problem 3. [Category: Proof] Let $A = \{wtw^R \mid w, t \in \{0,1\}^* \text{ and } |w| = |t|\}$. Prove that A is not context-free. **[10 points]**

Solution: Pumping Lemma Proof: We will show that A does not satisfy the pumping lemma. Let p be any pumping length. Consider the string $z = 0^{2p}0^p1^p0^{2p}$; $|z| > p$ and $z \in A$. Consider any division of z into u, v, w, x, y such that (a) $u = uvwxy$, (b) $|vwx| \leq p$ and (c) $|vx| > 0$. We will consider various cases based on the form of vwx.

1. Consider the case when vwx is contained in the first two-thirds of z. Then, in uv^2wx^2y (which is of length $> 6p$ because $|vx| > 0$), position $2p + 1$ from the end (which is now in the last third) is a 1 but the position $2p + 1$ from the begining is a 0. Thus, $uv^2wx^2y \notin A$.
2. Consider the case when vwx has a non-empty intersection with the last third. Observe that, since $|vwx| \leq p$, in this case, vwx must be completely contained within the second half of z. Here are a few subcases to consider.
    - Suppose x has both 0s and 1s, i.e., it spans the boundary of 1^p and 0^{2p}. Then, in uv^2wx^2y (of length $> 6p$), position $2p + 1$ from the begining is a 0 while position $2p + 1$ from the end is a 1. Thus, $uv^2wx^2y \notin A$.

- Suppose $x = \epsilon$ and v contains both 0s and 1s. Then uv^2wx^2y (of length $> 6p$) again has 0 in position $2p+1$ from the begining but 1 at position $2p+1$ from the end and so $uv^2wx^2y \notin A$.
- Suppose $x \neq \epsilon$ but contains only 0s, or vwx is completely contained in the last third of z. Then consider $z_0 = uv^{3p+4}wx^{3p+4}y$. Now, $|z_0| > 9p+4$. Moreover, the symbol at position $3p+1$ from the begining is a 1, while the symbol at position $3p+1$ from the end is a 0. Thus, $z_0 \notin A$.

■