Midterm

Note

This is a closed-book exam. Each problem accounts for 10 points, unless otherwise marked.

Problems

1. Let $L = \{w \in \{0,1\}^* \mid w \text{ does not contain 011 or 10 as a substring}\}$.
   (a) Draw the state diagram of a DFA, with as few states as possible, that recognizes L. The fewer states your DFA has, the more points you will be credited for this problem.
   (b) Translate the DFA in (a) systematically to an equivalent context-free grammar (using the procedure discussed in class).

2. Let $L = \{w \in \{0,1\}^* \mid w \text{ contains 110 as a substring or ends with 1}\}$.
   (a) Draw the state diagram of an NFA, with as few states as possible, that recognizes L. The fewer states your NFA has, the more points you will be credited for this problem.
   (b) Convert the preceding NFA systematically into an equivalent DFA (using the procedure discussed in class). Do not attempt to optimize the number of states, though you may omit the unreachable states.

3. Let the *rotational closure* of language A be $RC(A) = \{yx \mid xy \in A\}$.
   (a) Show that, for any language A, we have $RC(A) = RC(RC(A))$ (i.e., rotational closure, as an operation/function, is idempotent).
   (b) Show that the class of regular languages is closed under rotational closure.

4. We define the *avoids* operation for languages A and B to be

   A *avoids* $B = \{w \mid w \in A \text{ and } w \text{ doesn't contain any string in } B \text{ as a substring}\}$.

   Prove that the class of regular languages is closed under the *avoids* operation.

5. Consider the following CFG discussed in class, where for convenience the variables have been renamed with single letters.

$$\begin{array}{lcl} E & \rightarrow & E + T \mid T \\ T & \rightarrow & T \times F \mid F \\ F & \rightarrow & (E) \mid a \end{array}$$

   (a) Give the (leftmost) derivation and parse tree for the string $(a + a \times a)$.

(b) Convert the grammar into an equivalent PDA (that recognize the same language).

6. Draw the state diagram of a PDA that recognizes the following language: $\{w \in \{0,1\}^* \mid w \text{ has twice as many 1s as 0s}\}$. Please make the PDA as simple and deterministic as possible and explain the intuition behind the PDA.

7. For two given languages A and B, define $A \diamond B = \{xy \mid x \in A \text{ and } y \in B \text{ and } |x| = |y|\}$. Prove that, if A and B are regular, then $A \diamond B$ is context-free. (Hint: construct a PDA where the stack is used to ensure that x and y are of equal length.)

8. Prove *by induction* that, if G is a CFG in Chomsky normal form, then for any string $w \in L(G)$ of length $n \geq 1$, exactly $2n - 1$ steps are required for any derivation of w.

9. Let A be the language of all palindromes over $\{0, 1\}$ with equal numbers of 0s and 1s. Prove, using the pumping lemma, that A is not context free. (Note: a *palindrome* is a string that reads the same forward and backward.)

10. Let $A = \{wtw^R \mid w, t \in \{0,1\}^* \text{ and } |w| = |t|\}$, where w^R is the reverse of w. Prove that A is not context free.

Appendix

- A context-free grammar is in **Chomsky normal form** if every rule is of the form

$$\begin{array}{lll} A & \to & BC \quad \text{or} \\ A & \to & a \end{array}$$

  where a is any terminal and A, B, and C are any variables—except that B and C may not be the start variable. In addition,

$$S \to \varepsilon$$

  is permitted if S is the start variable.

- (Pumping Lemma for Context-Free Languages)

  If A is a context-free language, then there is a number p such that, if s is a string in A and $|s| \geq p$, then s may be divided into five pieces, $s = uvxyz$, satisfying the conditions:

  1. for each $i \geq 0$, $uv^ixy^iz \in A$,
  2. $|vy| > 0$, and
  3. $|vxy| \leq p$.