Midterm

Note

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

Problems

1. Let $A = \{a, b, c, d, e, f\}$ and $R = \{(a, b), (d, c), (d, e)\}$, which is a binary relation on A.
   (a) Give a symmetric and transitive but *not* reflexive binary relation on A that includes R. Please present the relation using a directed graph.
   (b) Find the smallest equivalence relation on A that includes R. Please present the relation using a directed graph.

2. Let L be a language over Σ (i.e., $L \subseteq \Sigma^*$). Two strings x and y in Σ^* are *distinguishable by L* if for some string z in Σ^*, exactly one of xz and yz is in L. When no such z exists, i.e., for every z in Σ^*, either both of xz and yz or neither of them are in L, we say that x and y are *indistinguishable by L*. Is indistinguishability by a language an equivalence relation (over Σ^*)? Please justify your answer.

3. (20 points) Give the state diagrams of DFAs recognizing the following languages. In all parts, the alphabet is $\{\mathtt{0}, \mathtt{1}\}$.
   (a) $\{w \mid w \text{ begins with a } \mathtt{1} \text{ and ends with a } \mathtt{0}\}$.
   (b) $\{w \mid \text{every even position of } w \text{ is a } \mathtt{0}\}$ (Note: see w as $w_1 w_2 \cdots w_n$, where $w_i \in \{\mathtt{0}, \mathtt{1}\}$).

4. Let $L = \{w \in \{0, 1\}^* \mid w \text{ contains } 100 \text{ as a substring or ends with a } 1\}$.
   (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.

5. For languages A and B, let the *shuffle* of A and B be the language $\{w \mid w = a_1 b_1 \cdots a_k b_k, \text{ where } a_1 \cdots a_k \in A \text{ and } b_1 \cdots b_k \in B, \text{ each } a_i, b_i \in \Sigma^*\}$. Show that the class of regular languages is closed under shuffle.

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

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

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

7. Draw the state diagram of a pushdown automaton (PDA) that recognizes the following language: $\{w \in \{a,b,c\}^* \mid$ the number of a's in w equals that of b's or c's$\}$ (no restriction is imposed on the order in which the input symbols may appear). Please make the PDA as simple as possible and explain the intuition behind the PDA.

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. Prove, using the pumping lemma, that $\{x\#wxy \mid w,x,y \in \{a,b\}^*\}$ is not context-free.

Appendix

- Common properties of a binary relation R on A:
  - R is *reflexive* if for every $x \in A$, xRx.
  - R is *symmetric* if for every $x,y \in A$, xRy if and only if yRx.
  - R is *transitive* if for every $x,y,z \in A$, xRy and yRz implies xRz.
- A context-free grammar is in **Chomsky normal form** if every rule is of the form

$$\begin{array}{lcl} A & \rightarrow & BC \quad \text{or} \\ A & \rightarrow & a \end{array}$$

  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,

$$S \rightarrow \varepsilon$$

  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) $|vxy| \leq p$.