NOTE: For questions 2 and 3, you will receive 2 marks for a BLANK answer to part (a), 1 mark for a BLANK answer to part (b) and 0.5 marks for a BLANK answer to part (c).

Name: ID:

---

1. (10 MARKS)

   (a) If L_1 is regular and L_2 is any language such that $L_2 \subseteq L_1$, then L_2 is regular. T F

   (b) If L_1 is regular and L_2 is context-free, then $L_1 \cap L_2$ is regular. T F

   (c) If M is a NFA with n states, then any DFA accepting $L(M)$ must have at least 2^n states. T F

   (d) Every CNF grammar is unambiguous T F

   (e) Which ONE of the following is an *unambiguous* grammar for strings of balance parentheses

      i. $S \rightarrow (S) \mid SS$
      ii. $S \rightarrow \epsilon \mid (S) \mid SS$
      iii. $S \rightarrow \epsilon \mid (S)S$
      iv. $S \rightarrow () \mid \epsilon \mid (S) \mid SS$
      v. None of the above

2. (a) (5 Marks) Use the pumping lemma to prove that

$$L = \{w \in \{(,)\}^* \mid w \text{ is balanced}\}$$

is not regular

(b) (3 Marks) For a string $w = w_1 \dots w_k$, $k \geq 0$, define $w^R = w_k w_{k-1} \dots w_1$. For any language L, define $L^R = \{w^R \mid w \in L\}$. Suppose L is regular. Is L^R always regular? If your answer is "yes", give a construction that proves this is the case. If your answer is "no", give a regular language L and prove that L^R is not regular.

(c) (2 Marks) Let

$$L = \{w \in \{0,1\}^* \mid w \text{ contains the same number of occurrences of the substrings 01 and 10}\}$$

For example $101 \in L$, but $1010 \notin L$. Prove that L is regular by giving a FA or regular expression that defines it, or use the pumping lemma or closure properties of regular languages to prove that it is not regular.

3. (a) (5 Marks) Convert the grammar $S \to \epsilon \mid (S)S$ to CNF. Give each step of the CNF construction. If the step does not apply to this grammar, explain why.

(b) (3 Marks) Give a grammar for the language

$$\{w \in \{0,1\}^* \mid w \text{ contains the same number of 0's and 1's.}\}$$

. (For full marks, give an unambiguous grammar.)

(c) (2 Marks) Let G be a CNF grammar, and $w = w_1 \ldots w_k \in L(G)$. How many steps must there be in the derivation of w? Give a justification for your answer.