Name ______________________ Student No. ______________________

No aids allowed. Answer all questions on test paper. Use backs of sheets for scratch work.

Total Marks: 50 [each question 10 marks]

1. For any string $w = w_1 w_2 \dots w_n$, the *reverse* of w, denoted w^R, is the string w in reverse order, that is $w_n \dots w_2 w_1$. If L is a language, then $L^R = \{w^R | w \in L\}$. Show that if L is regular, then so is L^R.

   Solution: If L is regular, there exists a regular expression E such that $L = L(E)$. We show how to transform E into E^R so that $L^R = L(E^R)$. The transformation is by structural induction on E. If $E = \varepsilon, \emptyset, a$ (where $a \in \Sigma$), then $E^R = E$. If $E = E_1 \cup E_2, E_1 E_2, E_1^*$, then $E^R = E_1^R \cup E_2^R, E_2^R E_1^R, (E_1^R)^*$, respectively.

2. Show 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.

   Solution: If $w \in L(G)$, and $|w| \geq 1$, then a derivation of w can be re-ordered, so that in the first part we derive all the variables, and in the second part we derive all the terminals. Thus, the derivation looks as follows:

$$S \Rightarrow \ldots \Rightarrow x \Rightarrow \ldots \Rightarrow w$$

   where $x \in V^*$, $|x| = |w|$, and after obtaining x, only rules of the form $A \longrightarrow a$ are used. Thus, the second part of the derivation has $|w|$ many steps (each variable A is replaced by a terminal a). To obtain x, we need to obtain a string of $|x| = |w|$ many variables. This requires $|w| - 1$ steps, since each step is an application of the rule $A \longrightarrow BC$.

3. **Rice's Theorem:** Let P be any problem about Turing machines that satisfies the following two properties:

   (a) If M_1, M_2 are two TMs such that $L(M_1) = L(M_2)$, then $\langle M_1 \rangle \in P$ iff $\langle M_2 \rangle \in P$.

   (b) There exist M_1, M_2 such that $\langle M_1 \rangle \in P$, but $\langle M_2 \rangle \notin P$.

   Show that P is undecidable.

   Solution: To show that P is undecidable, we show how to reduce A_{TM} to it. Note that properties (a) and (b) also hold for $\overline{P}$. We know that either P or $\overline{P}$ contain a TM M_{empty}, i.e., $L(M_{\text{empty}}) = \emptyset$. Since showing that P is not decidable implies that $\overline{P}$ is not decidable, we may assume that $\langle M_{\text{empty}} \rangle \notin P$.

   Suppose that P is decidable by some TM M_P.

   Let $M_{\langle M,w \rangle}$ be designed as follows: on input x, first simulate M on w. If M accepts w, then simulate M_1 on x. If M rejects w, then reject x.

   Note that there are two possibilities: M accepts w, in which case $L(M_{\langle M,w \rangle}) = L(M_1)$, and so by property (a) $\langle M_{\langle M,w \rangle} \rangle \in P$, or M does not accept w (by rejecting, or by not halting), then $L(M_{\langle M,w \rangle}) = \emptyset$, so since $\langle M_{\text{empty}} \rangle \notin P$, by (a) $\langle M_{\langle M,w \rangle} \rangle \notin P$.

   Therefore, $\langle M, w \rangle \in A_{\mathrm{TM}}$ iff $\langle M_{\langle M,w \rangle} \rangle \in P$. Hence P cannot be decidable.

4. Show that the Kolmogorov function $K(x)$ is not computable.

Hint: Suppose that $K(x)$ is computable. Let $f(n)$ be the lexicographically first string x such that $K(x) \geq n$. (It was shown in class that for all n, $f(n)$ exists, so you can assume that f is well defined.) Now derive the contradiction.

Solution: Since $K(x)$ is computable, so is f. To see this, let $x_1, x_2, x_3, \ldots$ be the lexicographic ordering of all strings (over $\{0,1\}$). To find $f(n)$, compute $K(x_i)$ for $i = 1, 2, 3, \ldots$, until you find the first x_k such that $K(x_k) \geq n$. We know that such a k exists for all n, so f is computed by this algorithm (which always halts).

Let M be the Turing machine that implements the algorithm computing f. It follows that for each n, there exists an x such that $K(x) \geq n$ and $\langle M, \langle n \rangle \rangle$ represents x. Here $\langle n \rangle$ is the binary representation of the number n. Therefore,

$$|\langle M, \langle n \rangle \rangle| = 2|\langle M \rangle| + 2 + |\langle n \rangle| = c + \log_2(n)$$

For a sufficiently large n, $c + \log_2(n) < n$, contradiction that $K(x) \geq n$.

5. Let L be the language of encodings of Turing machines which on the empty input visit all their states. That is,

$$L = \{\langle M\rangle |\text{ on input } \varepsilon,\ M \text{ visits all its states }\}$$

Show that L is undecidable

Solution: We show a reduction from A_{TM} to show that L is undecidable:

Given a pair $\langle M, w\rangle$, construct a Turing machine M', which on any input (and hence on ε in particular), ignores its input, and simulates M on w. If M accepts w, M' writes a special symbol (e.g., $\heartsuit$) on the tape, enters a new state $q_{\heartsuit}$, and then cycles through all its states and halts.

To see that $\langle M, w\rangle \in A_{\mathrm{TM}} \iff \langle M'\rangle \in L$, note that if M accepts w, then M' will cycle through all its states. If M does not accept w, then M' will never enter state $q_{\heartsuit}$.