Homework 9 – Due Thursday, April 18 at 11:59 PM

Reminder Collaboration is permitted, but you must write the solutions *by yourself without assistance*, and be ready to explain them orally to the course staff if asked. You must also identify your collaborators and write "Collaborators: none" if you worked by yourself. Getting solutions from outside sources such as the Web or students not enrolled in the class is strictly forbidden. Collaboration is not allowed on problems marked "INDIVIDUAL."

Note You may use various generalizations of the Turing machine model we have seen in class, such as TMs with two-way infinite tapes, stay-put, or multiple tapes. If you choose to use such a generalization, state clearly and precisely what model you are using. **You may describe Turing machines at a high-level on this assignment.**

Problems There are 4 required problems and 1 bonus problem.

1. (**Time and Space Warmup**) Let $A = \{x\#^m x^R \mid x \in \{0,1\}^m \text{ for some } m \geq 0\}$. Show that $A \in \mathsf{TIME}(n^2)$ and $A \in \mathsf{SPACE}(n)$ by i) giving an **implementation-level** description of a basic, single-tape Turing machine M that decides A, ii) briefly explain why your TM correctly decides A, and iii) analyzing the running time and space usage of M.

2. (**Hierarchy Theorems**) You may assume without saying so that any reasonable-looking function (logarithms, polynomials, exponentials, and combinations thereof) is time-constructible.

   The time complexity class $\mathsf{P} = \bigcup_{k=1}^{\infty} \mathsf{TIME}(n^k)$ consists of all languages decidable in polynomial time. The time complexity class $\mathsf{EXP} = \bigcup_{k=1}^{\infty} \mathsf{TIME}(2^{n^k})$ consists of all languages decidable in exponential time (i.e., in time that is an exponential of a polynomial).

   (a) Show that $\mathsf{P} \subseteq \mathsf{TIME}(n^{\log n})$.

   (b) Use the time hierarchy theorem to show that $\mathsf{EXP} \not\subseteq \mathsf{TIME}(n^{\log n})$.

   (c) Combine parts (a) and (b) to conclude that $\mathsf{P} \neq \mathsf{EXP}$.

3. (**Encodings and Runtime**) The runtime of an algorithm (Turing machine) is always measured *as a function if its input length*. The goal of this problem is to help you understand what that means and why it can be subtle.

   (a) Prove that there is **no** polynomial-time algorithm that takes as input a natural number k (written in binary) and outputs (i.e., writes to its tape) the number $k!$ (again, written in binary).

   Hint: Show that the expected output for this problem is so long that it is impossible for a poly-time algorithm to even write it down.

   (b) Give a high-level description of a polynomial-time algorithm that takes as input a natural number k (written in **unary**, i.e., the string $\underbrace{11\ldots11}_{k \text{ times}}$) and outputs the number $k!$ (written in **binary**). Explain why your algorithm is correct and why it runs in polynomial time.

   Hint: You can use without proof the fact that Turing machines can perform basic arithmetic operations on binary numbers, like addition and multiplication, in polynomial time.

4. (**Closure Properties**)

   (a) Show that P is closed under the union operation, i.e., show that for all languages $L_1, L_2 \in \mathsf{P}$, we have $L_1 \cup L_2 \in \mathsf{P}$.

   (b) This question will walk you through a proof that P is closed under the star operation. Let L be a language in P, and let M be a TM deciding L in time $O(n^c)$ for some constant c. Consider the following algorithm S_1 that decides L^*. Explain why S_1 does *not* run in polynomial time.

Algorithm: $S_1(w)$

Input : String $w = w_1 w_2 \dots w_n$ of length n

1. For each $k \leq n$ and each way to break w into a concatentation of *strings* $s_1 \circ s_2 \circ \cdots \circ s_k$:
2. For each $i = 1, \dots, k$:
3. Run M on input s_i.
4. If all runs have accepted, *accept*.
5. *Reject*.

   (c) The following recursive algorithm uses a slightly different idea. Its correctness relies on the fact that a string $w_1 w_2 \dots w_n \in L^*$ if and only if there exists an index $i \in \{0, \dots, n-1\}$ such that $w_1 w_2 \dots w_i \in L^*$ and $w_{i+1} \dots w_n \in L$. Explain why S_2 does *not* run in polynomial time.

Algorithm: $S_2(w)$

Input : String $w = w_1 w_2 \dots w_n$ of length n

1. If $n = 0$: *Accept*.
2. For each $i = 0, 1, 2, \dots, n-1$:
3. (Recursively) run S_2 on input $w_1 w_2 \dots w_i$ and run M on input $w_{i+1} \dots w_n$.
4. If both runs accept, *accept*.
5. *Reject*.

   (d) The main issue with the recursive algorithm in part (c) is that it for each prefix $w_1 w_2 \dots w_i$ of w, it keeps repeating the work of checking whether that prefix is in L^*. Wouldn't it be great if for each i, you only had to check whether $w_1 w_2 \dots w_i$ is in L^* once?

   It turns out you can, and this forms the basis of an actual polynomial time algorithm. Think of filling out an array $T[0, 1, \dots, n]$ where each cell $T[i]$ contains the answer to the question, "is $w_1 w_2 \dots w_i \in L^*$?" Design an algorithm S_3 that systematically fills in this array and uses it to determine whether $w \in L^*$. Briefly explain why your algorithm runs in polynomial time, and how it allows you to conclude that P is closed under the star operation.

5. (**Bonus problem**) Show that your algorithm from problem 1(a) is optimal: There is no basic, single-tape TM algorithm deciding the language $A = \{x\#^m x^R \mid x \in \{0,1\}^m \text{ for some } m \geq 0\}$ in time $o(n^2)$.