Homework 7 – Due Thursday, March 26 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.**

1. **(**EVEN_{TM}**)** Consider the following computational problem: Given a Turing machine M (over some alphabet Σ), does there exist a string w of even length such that M accepts input w?

   (a) Formulate this problem as a language EVEN_{TM}.

   (b) Show that EVEN_{TM} is Turing-recognizable. You may give a high-level description of the Turing machine your construct.

   Hint: Think about the TM M_3 from Homework 6, Problem 4. What is it doing, and why did it have to be constructed that way? A similar idea will be helpful for this problem.

2. (**Countable sets**)

   (a) Aliens from the planet Foobar have finite single-strand DNA sequences consisting of the nucleobases A, C, G, and T. For example, ACGTTAG and CGATCGACTGCA are both possible DNA sequences. Let $\mathcal{F}$ be the set of all possible DNA sequences for residents of Foobar. Show that $\mathcal{F}$ is countable.

   Hint: A good way to show that a set is countable is to describe how to exhaustively list the elements in stages $1, 2, 3, \dots$, where only finitely many elements are listed in each stage. The required bijection from $\mathbb{N}$ is obtained by mapping each natural number to the nth distinct item in this list.

   (b) Aliens from the planet Foobarbaz have finite single-strand DNA sequences from a countably infinite set of nucleobases $\{A_1, A_2, A_3, \dots\}$. For example, $A_2A_4A_8A_1A_9$ and $A_3A_{747}A_{9999999}A_4$ are both possible DNA sequences. Let $\mathcal{B}$ be the set of all possible DNA sequences for residents of Foobarbaz. Show that $\mathcal{B}$ is countable.

3. (**Uncountable sets**)

   (a) Aliens from the planet Fubar'd have *(countably) infinite* single-strand DNA sequences from the set of nucleobases $\{\mathsf{A}, \mathsf{C}, \mathsf{G}, \mathsf{T}\}$. Let $\mathcal{D}$ be the set of all possible DNA sequences for residents of Fubar'd, so $\mathcal{D} = \{a_1 a_2 a_3 \ldots \mid a_i \in \{\mathsf{A}, \mathsf{C}, \mathsf{G}, \mathsf{T}\}, i \in \mathbb{N}\}$. Show that $\mathcal{D}$ is **uncountable**.

   (b) A function $f : \mathbb{N} \to \mathbb{N}$ is *rapidly growing* if $f(i+1) \geq 2f(i)$ for every $i \in \mathbb{N}$. So $f(1) = 3, f(2) = 9, f(3) = 27, f(4) = 81 \ldots$ are the first few values of (what appears to be) a rapidly growing function f. But neither the function g where $g(1) = 2, g(2) = 1, g(3) = 7, \ldots$ nor the function h where $h(1) = 4, h(2) = 7, h(3) = 14$ are rapidly growing.

   Show that $\mathcal{R} = \{f : \mathbb{N} \to \mathbb{N} \mid f \text{ is rapidly growing}\}$, the set of all rapidly growing functions, is uncountable.

   Hint: When you construct a function contradicting the diagonal, make sure that it is indeed a member of $\mathcal{R}$, i.e., that it is rapidly growing.

4. (**Unrecognizability**) Consider the explicit undecidable language described in Lecture 13: $UD = \{\langle M \rangle \mid M \text{ is a TM that does not accept on input } \langle M \rangle\}$. Show that this language is not Turing-recognizable.

5. (**Reduction Mad-Libs**) A language $S \subseteq \{0, 1, \ldots, 9\}^*$ is *brainrotten* if every string in S contains "67" as a substring. For example, the empty language, $\{0123456789, 4267\}$, and $\{(67)^n \mid n \geq 1\} = \{67, 6767, 676767, \ldots\}$ are all brainrotten, but $\{6, 7, 67\}$ and $\{(67)^n \mid n \geq 0\} = \{\varepsilon, 67, 6767, \ldots\}$ are not brainrotten. The language $BR_{\text{TM}} = \{\langle M \rangle \mid L(M) \text{ is brainrotten}\}$ corresponds to the following computational problem: Given the encoding of a TM M, does M recognize a brainrotten language? This exercise will walk you through a proof, by reduction, that BR_{TM} is undecidable.

   Assume, for the sake of contradiction, that BR_{TM} is decidable by a TM R. That is, there is a TM R that accepts $\langle M \rangle$ whenever $L(M)$ is brainrotten, and rejects $\langle M \rangle$ whenever $L(M)$ is not brainrotten. We will use R to construct a new TM T that decides the (undecidable) language A_{TM}.

Algorithm 1: $T(\langle M, w \rangle)$

Input : Encoding of a basic TM M over input alphabet $\{0, 1, \ldots, 9\}$, string $w \in \{0, 1, \ldots, 9\}^*$

1. Construct the following TM N:
   $N =$ "On input a string $x \in \{0, 1, \ldots, 9\}^*$:
   If x contains "67" as a substring, *accept*.
   Else, run M on input w. If it accepts, *accept*. If it rejects, *reject*."
2. Run R on input $\langle N \rangle$. If it accepts, ___(i)___. If it rejects, ___(ii)___.

   (a) This proof is by reduction from a language A to a language B. What are the languages A and B? (Make it clear in your solution which one is A and which one is B, since the order matters a lot!)

   (b) Consider the machine N constructed inside algorithm T. If M accepts on input w, what is the language $L(N)$? Is $L(N)$ brainrotten in this case?

   (c) If M does not accept on input w, what is the language $L(N)$? Is $L(N)$ brainrotten in this case?

   (d) Fill in the blanks labeled (i) and (ii) with *accept* or *reject* decisions to guarantee the following conditions: If M accepts input w, then T accepts input $\langle M, w \rangle$, and if M does not accept input w, then T rejects input $\langle M, w \rangle$. Use parts (b) and (c) to explain why these conditions hold for your choices of how to fill in the blanks.

(Your job is done now, but you may want to keep reading to see the exciting conclusion of the proof.) By part (d), the TM T exactly decides the language A_{TM}. But this language is undecidable, which is a contradiction. Hence our assumption that BR_{TM} was decidable is false, so we conclude that BR_{TM} is an undecidable language.