CS Theory (Spring '24)	Assigned: April 1, 2024
Homework 5	
Instructor: *Josh Alman*	Due: April 10, 2024 at 11:59 pm

Collaboration on homework is allowed. However, you must write up solutions by yourself and understand everything that you hand in. You may also consult other reference materials, but you may not seek out answers from other sources.

Write the solution to each problem on a separate page. This includes writing each part of each problem on a separate page. Do not write your name or any other identifying information on any page of the submission; we are using anonymous grading in GradeScope.

Be sure to answer the final question listing your collaborators and other sources used. See the course webpage for more details.

There are 120 total possible points, and also bonus questions worth an additional 30 points.

1 Mapping Reductions (20 points; each part 10 points)

(a) Prove or disprove: If B is a regular language, and $A \leq_m B$, then A is a regular language.

(b) Prove or disprove: There are languages A, B with $A \neq B$ such that $B \subseteq A$ and $A \leq_m B$.

2 Reverses (40 points; each part 20 points)

Recall that for a string $w \in \{0,1\}^*$, we write w^R to denote its reverse. For a language L, we say L is *reversible* if and only if: for every $w \in L$, we have $w^R \in L$.

(a) Let R_{DFA} be the language over $\Sigma = \{0,1\}$ defined by

$$R_{DFA} := \{\langle D \rangle \mid \; D \text{ is a DFA which recognizes a reversible language}\}.$$

Determine (with proof) whether R_{DFA} is decidable, undecidable but recognizable, or unrecognizable.

(b) Let R_{TM} be the language over $\Sigma = \{0,1\}$ defined by

$$R_{TM} := \{\langle M \rangle \mid \; M \text{ is a Turing Machine which recognizes a reversible language}\}.$$

Determine (with proof) whether R_{TM} is decidable, undecidable but recognizable, or unrecognizable.

3 Binary Representations (30 points; each part 15 points)

Let S be a set. A *binary representation* of S is an injective function

$$f : S \to \{0,1\}^* \setminus \{\varepsilon\}.$$

For example, in class we discussed binary representations of graphs, DFAs, Turing Machines, and other objects. As we discussed in class, when we're deciding whether a subset of S is recognizable or decidable, we typically don't worry about the details of the representation we're using for S. In this problem, we'll worry about the details.

For a set S, a subset $P \subseteq S$, and a binary representation f of S, define the language $L_{P,f}$ by

$$L_{P,f} := \{w \in \{0,1\}^* \mid \text{ there is a } p \in P \text{ such that } f(p) = w\}.$$

(a) Suppose f and g are two binary representations of a set S. Define $c : \{0,1\}^* \to \{0,1\}^*$ by, for any $w \in \{0,1\}^*$,

$$c(w) = \begin{cases} g(s), & \text{if } s \in S \text{ is such that } f(s) = w, \\ \varepsilon, & \text{if there is no } s \in S \text{ with } f(s) = w. \end{cases}$$

Prove that if c is a computable function, and if $P \subseteq S$ is any subset such that $L_{P,g}$ is decidable, then $L_{P,f}$ is also decidable.

(b) Using part a, prove that there is an undecidable language over the alphabet $\Sigma = \{1\}$ (i.e., every string in the language must only have 1s in it.).

(c) (BONUS, 10 additional points) Let G denote the set of all simple graphs (graphs which are undirected, unweighted, and have at most one edge between any pair of nodes), and let $C \subseteq G$ denote the set of all simple graphs that are connected (there is a path between any pair of nodes). Prove or disprove: there is a binary representation f of G such that $L_{C,f}$ is undecidable.

4 The Rambunctious Turing Machine Code Golf Competition (20 points for part a, 8 points for part c)

We say that a Turing Machine M is *rambunctious* if it has all of the following properties:

- It accepts the string 101010101010 and the string 000000111111
- It rejects the string 111111111111 and the string 010101010101
- It loops (never accepts or rejects) on the string 111111000000 and the string 000000000000

Define the language L_{ramb} by

$$L_{ramb} := \{\langle M \rangle \mid \ M \text{ is a rambunctious Turing Machine}\}.$$

(a) Prove that L_{ramb} is undecidable.

(b) (BONUS, up to 20 additional points) Your goal in this problem is to give the code for a rambunctious Turing Machine using the following website: `https://turingmachinesimulator.com/`. Be sure to look at some examples and tutorials on that site to learn about the notation they use.

If your Turing Machine is not rambunctious (or if it doesn't compile on the website) then you will get 0 points. Otherwise, we call your submission "valid", and your score on this problem will be calculated as follows.

Let n be the number of students in the class who make a valid submission. Let k be the number of students in the class who make a valid submission that uses at least as many total lines of code as yours. Then your score will be

$$5 + \left\lceil \frac{15k}{n} \right\rceil.$$

In other words, you want to use as few lines of code as possible, and you will get more points if you use fewer lines of code than other students. Good luck!

(We will measure "total lines of code" as the number of lines in the file you submit. We will post instructions for submitting your code closer to the deadline.)

(c) In light of part a, it may be difficult for the course staff to grade part b. Explain what the apparent issue is, and explain how the course staff might be able to grade it anyway.

5 Collaborators and References and Rules (2 points)

Do all three of the following:

- List who you collaborated with on each problem, including any TAs or students you discussed the problems with in office hours.
- List any reference materials consulted other than the lectures and textbook for our class.
- Follow the formatting rules listed at the beginning of the assignment!