Exam #1 (Automata)

Due: February 27, 2024

Academic Honesty & Collaboration. I expect you to work alone, and for your work to represent wholly your own efforts. You are, of course, allowed to ask your classmates (and me) general questions about the material. Please include the following information at the top of your submission, along with your name.

- Written sources used: (Include textbook(s), complete citations for web or other written sources. Write none if no sources used)
- Help obtained: (Include names of anyone other than the instructor.)

Extensions. I am happy to offer extensions, but I require (for this exam) that you ask *in advance* and at least prior to *Friday, Feb 23*. Further, I expect you to be able to show either (i) proof of effort and some progress on the exam, or (ii) some minor proof of conflict in schedule. (Exceptions will be made, of course, in cases of sickness, injury, or emergency.) Finally, please talk to me *as soon as possible* if you know you have a conflict in schedule (e.g., another take-home exam in a separate course) and you need some help meeting this deadline.

The latest extension I can/will give is to Thursday, Feb 29.

Extra Credit. Extra credit questions will be graded (more harshly) on an E/O/N scale:

- **E**xcellent. No errors or only a few tiny errors. Argument is easy to follow. Would be accepted by the broader mathematical community as a valid proof of the claim. Full credit.
- **Ok**ay. Has most of the right answer, but some minor errors. Argument is easy to follow. Would be accepted by the broader mathematical community after some revision. Maximum half credit.
- **N**o credit. Major errors and/or too many minor errors. Would not be accepted by the broader mathematical community.

This is to say, I am looking for correct or nearly correct answers, and will not be offering extra credit to anything less.

N.b. For each question that calls for a proof of regularity, I will allow you to exhibit not only the regular models we discussed in class (NFAs, DFAs, and regular expressions) but also any model we showed to be equivalently regular, e.g., Pointed DFAs and All-NFAs, etc. ***Many of the problems below are made dramatically easier by choosing the right model to exhibit.***

#1 Sandwich (50 pts)

Suppose A is a language with alphabet Σ and s is some character not in Σ. Define the sandwich of s by language A, written $\|A\|_s$, as follows:

$$\|A\|_s = \{xsy \mid x, y \in A\}$$

Prove the following claim.

Claim 1. *If A is a regular language with alphabet Σ and $s \notin \Sigma$, then $\|A\|_s$ is regular.*

#2 Intersection (50 pts)

Define the intersection of sets A and B as follows:

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}$$

Prove or disprove the following claim.

Claim 2. *If languages A and B are regular then $A \cap B$ is regular.*

#3 (Ir)regularity (50 pts)

Let $\Sigma = \{0, 1\}$ and let

$$A = \{w \mid w \text{ has an equal number of zeros and ones}\}$$
$$B = \{w \mid w \text{ contains an equal number of occurrences of the substrings 01 and 10}\}$$

For example, $101 \in B$ because 101 contains a single 01 and a single 10, but $1010 \notin B$ because 1010 contains two 10s and one 01.

Exactly one of A and B is regular.

1. (25 pts). Which language is regular? Give a proof that it is regular.
2. (25 pts). Which language is irregular? Use the pumping lemma to prove it is irregular.

#4 Regular Grammars (50 pts)

Consider the following definition, which connects your study of context-free grammars (CFGs) with regular languages.

Definition 1 (Right-Regular Grammars)**.** *A right-regular grammar (RRG) is a tuple (N, Σ, P, S) where*

1. *N is a finite set of nonterminal symbols,*

2. *Σ is an alphabet of terminal symbols disjoint from N,*

3. *$P \subseteq N \times \Sigma_\epsilon N_\epsilon$ is a finite set of production rules, and*

4. *$S \in N$ is the start symbol.*

Observe that the type of production rules P is different than that of CFGs. To be clear, the production rules of a right-regular grammar each have one of the following forms:

- $A \to a$,
- $A \to aB$, or
- $A \to B$.

That is, the right hand side of a right-regular grammar production rule is *only* ever: a terminal (alone), a terminal before a nonterminal, or just a nonterminal.

Prove the following claim by structural induction.

Claim 3. *Every regular expression R can be translated into a right-regular grammar G such that $L(R) = L(G)$.*

#5 Distinguishability, again (40 pts extra credit)

Recall the definition of distinguishability that was first described in Lab 1A.

Definition 2 (Distinguishability)**.** *Let L be a set of strings and x and y be strings (not necessarily in L).*

- *We say that x and y are distinguishable by L if there exists a string z (also not necessarily from L itself) such that exactly one of xz and yz are in L.*
- *We say that x and y are indistinguishable by L if they are not distinguishable.*

Furthermore, if x and y are strings indistinguishable by L, we write $x \equiv_L y$.

Distinguishability can generalize to n many strings that are each mutually distinguishable. We call this pairwise distinguishability.

Definition 3 (Pairwise distinguishability & Index)**.** *Let L be a language and let X be a set of strings. We say that X is pairwise distinguishable by L if every two distinct strings in X are distinghuishable by L. Define the index of L to be the cardinality of the greatest such X that is pairwise distinguishable by L. The index of L may be finite or infinite.*

For full credit, give a correct proof of the following claim.

Claim 4. *The language A is regular if and only if it has finite index.*

Recall that $\equiv_A$ is an equivalence relation on A. In particular, the index of A is equal to the number of partitions induced on A by the relation $\equiv_A$. As a corollary, the following claim is equivalent.

Claim 5. *The language A is regular iff the number of equivalence classes of $\equiv_L$ is finite.*

(You may thus prove either.)