6.045 Midterm

- This exam is open book, open notes.
- The exam is long; manage your time wisely.
- If you can not produce a full proof, clearly state partial results for partial credit.

YOUR NAME: ______________________________

Problem	Points	Score
1	5	
2	5	
3	5	
4	5	
5	5	
Total	25	

Problem 1. Mark each statement (T) for true or (F) for false. All languages are over alphabet $\Sigma = \{0, 1\}$.

_____ 1. If L is regular, then the following language must be regular:

$$\{w \mid w \in L \text{ and } w \text{ ends in } 10101\}$$

Solution. True. This is an intersection of the regular language L and the regular language $(0 \cup 1)^*10101$ and is therefore regular.

_____ 2. If an NFA with n states accepts no string with length less than n, then must accept no strings at all.

Solution. True. We prove the contrapositive. Suppose that the NFA accepts some string. Then there is some path from the start state to an accept state. This implies that there is a simple path (in which every state appears at most once) from the start state to an accept state. (We can obtain a simple path from an arbitrary path by removing loops.) A simple path can contain at most n nodes, and thus at most $n-1$ transitions. Therefore, this simple path defines an accepted string with length at most $n - 1$.

_____ 3. The union of a countable set of regular languages is always regular.

Solution. False. In fact, every language is the union of a countable set of regular languages, each containing a single string. Closure under pairwise union does not imply closure under infinite union.

_____ 4. If L_1 is undecidable and L_2 is decidable, then $L_1 \oplus L_2$ must be undecidable.

Solution. True. Assume for the purpose of contradiction that $L_1 \oplus L_2$ is decidable. Then $(L_1 \oplus L_2) \oplus L_2$ is decidable, because the set of regular languages is closed under symmetric difference. However, $(L_1 \oplus L_2) \oplus L_2 = L_1$, which is undecidable by assumption. This gives a contradiction.

_____ 5. The intersection of a recognizable language and an unrecognizable language is always unrecognizable.

Solution. False. If L_1 is any unrecognizable language and L_2 is the recognizable language ϕ, then $L_1 \cap L_2 = \phi$ is recognizable.

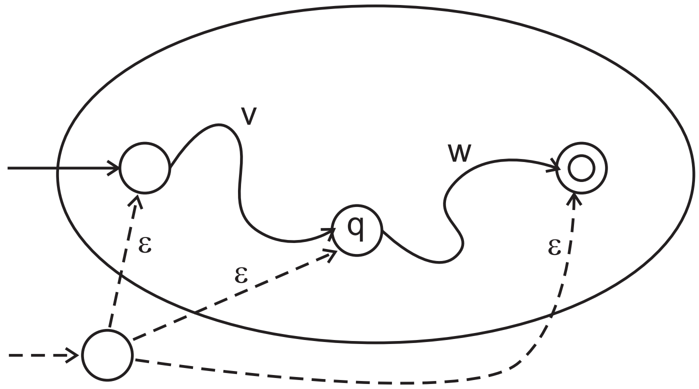
Problem 2. Prove that if L is a regular language over alphabet Σ, then the language of suffixes of L, given by

$$L_S = \{w \mid \text{There exists } v \in \Sigma^* \text{ such that } vw \in L.\}$$

is also regular.

Solution. Let M be a DFA that accepts L. We can construct an NFA N that accepts L_S by adding a new start state with ϵ-transitions to every reachable state of M. This implies that L_S is also regular.

First, we show that N accepts all strings in L_S. Suppose that M accepts string vw. Then M contains a path from the start state to some intermediate node q labelled with symbols in v and a path from q to an accepting state labelled with symbols in w. Then N accepts the suffix w, since there is an arc from the new start state to q labeled epsilon and a path from q to an accepting state labelled with symbols in w as before.



Now we show that N accepts only strings is L_S. Suppose that N accepts string w. Then N contains a path from the start state consisting of an initial epsilon transition to a state q followed by a path from q to an accept state labelled with symbols in w. Since q is a reachable state of M, there exists a path from the start state of M to state q labelled with symbols in some string v. Thus, M accepts a string vw of which w is a suffix.

Problem 3. Prove or disprove the following claim. There is an infinite regular language L such that for every "gap size" $g \geq 0$, there exists a "base length" $b \geq 0$ such that L contains no strings with length between b and $b + g$ inclusive.

Solution. The claim is false.

Assume that such a language L exists. By the pumping lemma, there exists a pumping length p such that every string $w \in L$ with length of at least p is of the form xyz, where:

1. $xy^iz \in L$ for all $i \geq 0$
2. $|y| \geq 1$
3. $|xy| \leq p$

Since L is infinite, there exists some string $w \in L$ with length at least p. Therefore, L contains strings with lengths:

$$|w|, |w| + |y|, |w| + 2|y|, |w| + 3|y|, \ldots$$

This contradicts the claim when the "gap size" g is greater than both $|y|$ and $|w|$.

Problem 4. Are the following statements **True** or **False**? *In each case give a brief explanation. (One sentence or two should suffice.)*

1. There are infinitely many Turing-recognizable languages.

   Solution. True. For example, every finite language is Turing-recognizable and there are infinitely many of these.

2. There exists a Turing-recognizable language L such that for every Turing machine M recognizing L, we have $|\langle M\rangle| > 10000$. (Assume encodings are in binary.)

   Solution. True. There are only finitely many (at most $2^{1}0000$) Turing machines with encoding of at most 10000 bits. Hence only finitely many languages satisfy the property above.

3. The following property of languages is undecidable.

$$P = \left\{ \langle M\rangle \;\middle|\; \begin{array}{l} M \text{ is a Turing machine and there exists a Turing machine } N \\ \text{that recognizes the same language where } |\langle N\rangle| \leq 10000. \end{array} \right\}$$

   Solution. True. The proof is by Rice's Theorem. We are told above that P is a property of languages. It is non-trivial, because any Turing machine with an encoding in fewer than 10000 bits has the property, and there is a Turing machine that lacks the property by part (2) above.

Problem 5. Let M be a 1-tape Turing machine with q states, and let w be a string of length n. Prove that if on input w the machine M does not move its head left in the first $n+q+1$ steps, then it *never* moves its head left on this input.

Solution.

Assume that M does not halt in the first $n+q+1$ steps. (Else it is obvious that it never moves its head left if it does not do so in the first $n+q+1$ steps.)

We will show that if the tape head does not move left in the first $|w|+q+1$ steps when Turing machine M is run on input w, then the tape head *never* moves left. In the first $|w|$ steps, the tape head moves right across the entire input. In each of the next $q+1$ steps, a blank is read from the tape, some symbol may be written, and then the tape head moves right again. In these $q+1$ steps, some state in the finite automaton must be visited twice, by the pigeonhole principle. Thus, M enters a loop in which the automaton iterates through the same sequence of states, blanks are always read from the tape, and the tape head *always* moves right.