Final

Note

This is a closed-book exam. Each problem accounts for 10 points, unless otherwise marked.

Problems

1. A *synchronizing sequence* for a DFA $M = (Q, \Sigma, \delta, q_0, F)$ and some "home" state $h \in Q$ is a string $s \in \Sigma^*$ such that, for every $q \in Q$, $\delta(q, s) = h$. A DFA is said to be *synchronizable* if it has a synchronizing sequence for some state. Prove that, if M is a k-state synchronizable DFA, then it has a synchronizing sequence of length at most k^3. (Note: $\delta(q, s)$ equals the state where M ends up when M starts from state q and reads input s.)

2. Show that single-tape TMs that cannot write on the portion of the tape containing the input string recognize only regular languages.

3. Let $CONTAIN_{\text{PDA\_DFA}} = \{\langle M_1, M_2\rangle \mid M_1$ is a PDA and M_2 is a DFA such that $L(M_1) \subseteq (M_2)\}$. Show that $CONTAIN_{\text{PDA\_DFA}}$ is decidable.

4. Show that if A is Turing-recognizable and $A \leq_m \overline{A}$, then A is decidable.

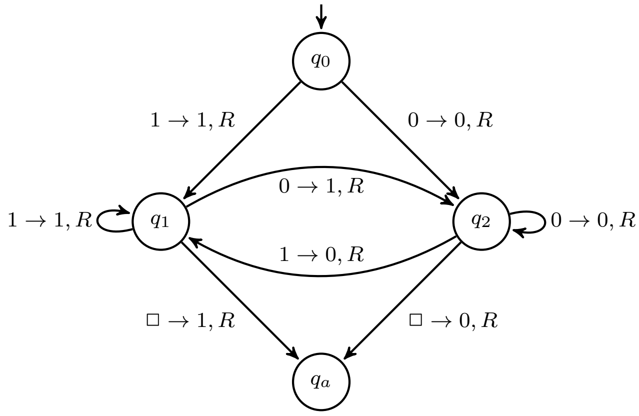
5. If A is reducible to B and B is a regular language, does that imply that A is a regular language? Why or why not?

6. Discuss the applicability of Rice's Theorem for each of the following problems (languages). Please give the reasons why or why not.

   (a) $FINITE_{\text{TM}} = \{\langle M\rangle \mid M$ is a TM and $L(M)$ is finite$\}$.

   (b) $W\_USELESS_{\text{TM}} = \{\langle M\rangle \mid M$ is a TM with useless states$\}$. (A *useless state* in a Turing machine is one that is never entered on any input string.)

7. In the proof of the Cook-Levin theorem, which states that SAT is NP-complete, we used 2×3 windows of cells to formulate the constraint that the configuration of each row (except the first one) in the $n^k \times n^k$ tableau legally follows the configuration of the preceding row. Suppose the Turing machine for the NP language being reduced is as follows. (Here, $\square$ is used in place of $\sqcup$, because of a typesetting problem.)



Which of the following 2×3 windows of cells are illegal? Why?

1	0	1
q_2	0	1

1	1	0
1	1	0

1	q_1	0
1	0	q_2

q_0	1	0
1	q_1	0

1	0	q_a
1	0	$\square$

1	0	q_2
1	0	1

8. Following the previous problem, why couldn't we simply use two entire rows of cells to formulate the constraint?

9. In the proof that the *3SAT* problem is polynomially reducible to the *SUBSET_SUM* problem, we convert an arbitrary boolean expression in 3CNF (input of the *3SAT* problem) to a set of numbers and a target number as input of the *SUBSET_SUM* problem.

   (a) Please illustrate the conversion by giving the set of numbers and the target number that will be obtained from the following boolean expression:

   $$(x + y + \overline{z}) \cdot (\overline{x} + y + z) \cdot (x + \overline{y} + z).$$

   (b) The original boolean expression is satisfiable. As a demonstration of why the reduction is correct, please use the resulting set of numbers and the target number to argue that it is indeed the case.

10. Let $HITTING\_SET = \{\langle S, C, k\rangle \mid S$ is a set and C is a collection of subsets of S such that S contains a hitting set of size k for $C\}$. A hitting set in S for C is a subset $S' \subseteq S$ such that S' contains at least one element from each subset in C. Prove that *HITTING_SET* is NP-complete. (Hint: reduction from *VERTEX_COVER*.)

Appendix

- Given a language $L \subseteq \Sigma^*$, define a binary relation R_L over Σ^* as follows:

$$xR_Ly \text{ iff } \forall z \in \Sigma^*(xz \in L \leftrightarrow yz \in L)$$

R_L can be shown to be an equivalence relation.

(**Myhill-Nerode Theorem**) With R_L defined as above, the following are equivalent:

1. L is regular.
2. R_L is of finite index.

Moreover, the index of R_L equals the number of states in the smallest DFA that recognizes L.

Note: the *index* of an equivalence relation is the number of equivalence classes it induces.

- (**Rice's Theorem**) Any problem P about Turing machines satisfying the following two conditions is undecidable:

  1. For any TMs M_1 and M_2, where $L(M_1) = L(M_2)$, we have $\langle M_1 \rangle \in P$ iff $\langle M_2 \rangle \in P$.
  2. P is nontrivial, i.e., there exist TMs M_1 and M_2 such that $\langle M_1 \rangle \in P$ and $\langle M_2 \rangle \notin P$.

- Language A is **mapping reducible** (many-one reducible) to language B, written $A \leq_m B$, if there is a computable function $f : \Sigma^* \longrightarrow \Sigma^*$, where for every w,

$$w \in A \Longleftrightarrow f(w) \in B.$$

- $A \leq_m B$ is equivalent to $\overline{A} \leq_m \overline{B}$.

- $3SAT = \{\langle \phi \rangle \mid \phi \text{ is a satisfiable 3CNF-formula}\}$. (A 3CNF-formula is a CNF-formula where all the clauses have three literals.) $3SAT$ is NP-complete.

- $VERTEX\_COVER = \{\langle G, k \rangle \mid G$ is an undirected graph that has a k-node vertex cover$\}$. (A *vertex cover* of an undirected graph G is a subset of the nodes where every edge of G touches one of those nodes.) $VERTEX\_COVER$ is NP-complete.

- $SUBSET\_SUM = \{\langle S, t \rangle \mid S = \{x_1, \cdots, x_k\}$ and for some $\{y_1, \cdots, y_l\} \subseteq S$, we have $\sum y_i = t\}$. $SUBSET\_SUM$ is NP-complete.