Final

Note

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

Problems

1. For two given languages A and B, define $A \diamond B = \{xy \mid x \in A \text{ and } y \in B \text{ and } |x| = |y|\}$. Prove that, if A and B are regular, then $A \diamond B$ is context-free. (Hint: construct a PDA where the stack is used to ensure that x and y are of equal length.)

2. Give a formal description of a (single-tape deterministic) Turing machine that decides the language $\{1^k\#1^{k^2} \mid k \geq 1\}$.

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

4. A *useless state* in a pushdown automaton is a state that is never entered on any input. Show the decidability of the problem of determining whether a given pushdown automaton has a useless state.

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

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

7. 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?

8. Discuss the applicability of Rice's Theorem for each of the following problems (languages). Please give the reasons why or why not.
   (a) $REGULAR_{\text{TM}} = \{\langle M\rangle \mid M \text{ is a TM and } L(M) \text{ is regular}\}$.
   (b) $UNCOUNTABLE_{\text{TM}} = \{\langle M\rangle \mid M \text{ is a TM and } L(M) \text{ is uncountable}\}$.

9. In the proof that the *3SAT* problem is polynomially reducible to the *VERTEX_COVER* problem, we convert an arbitrary boolean expression in 3CNF (input of the *3SAT* problem) to an input graph of the *VERTEX_COVER* problem.
   (a) Please illustrate the conversion by drawing the graph that will be obtained from the following boolean expression:

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

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

10. 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:

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

    (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.

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.

- $ALL_{\text{CFG}} = \{\langle G\rangle \mid G \text{ is a CFG and } L(G) = \Sigma^*\}$. ALL_{CFG} is undecidable.

- (**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 $\overline{A} \leq_m \overline{B}$.
- $3SAT = \{\langle \phi \rangle \mid \phi$ 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.