ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Exam II, Theory of Computation 2018-2019

- Books, notes, communication, calculators, cell phones, computers, etc... are not allowed.
- Your explanations and proofs should be clear enough and in sufficient detail so that they are easy to understand and have no ambiguities.
- You are allowed to refer to material covered in the class including theorems without reproving them.
- **Do not touch until the start of the exam.**

Good luck!

Name: ______________________ **N° Sciper:** ______________

Problem 1 / 7 points	Problem 2 / 6 points	Problem 3 / 6 points	Problem 4 / 6 points

Total / 25

1 *(consisting of subproblems* ***a-b****, 7 pts)* **Basic questions.**

1a *(4 pts)* Consider two languages A and B. Which of the following statements are true?

1. A is regular $\implies$ A is decidable $\implies$ A is recognizable.
2. A and B are decidable $\implies$ $A \cup B$ is decidable.
3. A and B are unrecognizable $\implies$ $A \cup B$ is unrecognizable.
4. A is undecidable and recognizable $\implies$ $\overline{A}$ is unrecognizable.
5. $\overline{A}$ is decidable $\implies$ A is decidable.
6. $\{0^n 1^n : n \geq 0\}$ is decidable.
7. A is decidable and $A \leq_m B$ $\implies$ B decidable.
8. $\{\langle M, w\rangle \mid M$ is a TM and M doesn't halt on input w$\}$ is recognizable.

(A complete solution identifies ***all*** *true statements. A fully correct solution is worth* 4 *points. A solution with one mistake is worth* 3 *points. A solution with two mistakes is worth* 1 *point. Solutions with more mistakes are worth* 0 *points. A mistake is to either indicate falsely that a false statement is true or to not indicate that a true statement is true.)*

Solution:

Among the above statements, the following are true **1,2,4,5,6**

1b *(3 pts)* Let B be an undecidable language. Show that $B \leq_m \overline{B}$ implies that B and $\overline{B}$ are unrecognizable.

(In this problem you are asked to provide a formal proof of the statement. Recall that you are allowed to refer to material covered in the class including theorems without reproving them.)

Solution:

From the class, we know that: B is decidable iff B is recognizable and $\overline{B}$is recognizable, which is equivalent to B is undecidable iff B is unrecognizable or $\overline{B}$ is unrecognizable. **[1 point for stating the theorem and the equivalent formulation]**
We will now prove that B is unrecognizable implies $\overline{B}$ is unrecognizable, and that $\overline{B}$ is unrecognizable implies B is unrecognizable.

1. if B is unrecognizable then from $B \leq_m \overline{B}$ we get that $\overline{B}$ is also unrecognizable. **[1 point for considering the two cases, and proving the first one]**

2. if $\overline{B}$ is unrecognizable then from Exercise set VI we know that $B \leq_m \overline{B}$ implies $\overline{B} \leq_m B$, and thus, B is also unrecognizable. **[1 point, correct proofs of the fact $B \leq_m \overline{B} \iff \overline{B} \leq_m B$ have a positive impact on the grading of the whole exercise, but are not required to get full points]**

2 *(6 pts)* A DFA $D = (Q, \Sigma, \delta, q, F)$ is defined to be minimal if it has the minimum number of states among all DFAs that recognize the same language as D. More formally, D is minimal if

$$|Q| = \min_{\substack{D'=(Q',\Sigma,\delta',q',F'):\\ L(D)=L(D')}} |Q'|.$$

Show that the following language is decidable.

$$L_2 = \{\langle D \rangle \mid D \text{ is a minimal DFA}\}.$$

(In this problem you are asked to show that the language L_2 is decidable. Recall that you are allowed to refer to material covered in the class including theorems without reproving them.)

Solution: We will design an algorithm that given a description of a DFA D decides whether D is minimal.

The algorithm works as follows:

Algorithm 1 IsMinimal($D = (Q, \Sigma, \delta, q_s, F)$)

```
procedure IsMinimal(D = (Q, Σ, δ, q_s, F))
    for i ∈ [1, |Q| − 1] do
        for D' ∈ {(Q', Σ, δ', q', F') : |Q'| = i ∧ (Q', Σ, δ', q', F') is a description of a DFA} do
            if EQ(D, D') == TRUE then
                return FALSE
    return TRUE
```

where EQ is a Turing machine that for descriptions of two DFAs decides if they recognize the same language. We know that such a machine exists from the lectures.

We need to explain how can we implement second **for** in the algorithm. Notice that the number of DFAs on i states is bounded by $2^i \cdot i \cdot i^{i \cdot |\Sigma|}$, as the number of different δ' is at most $i^{i \cdot |\Sigma|}$ (for every state q and every letter $a \in \Sigma$ we have i choices for $\delta'(q, a)$), number of different q' is at most i and number of different F' is at most 2^i. In particular the number of DFAs on i states is finite so we can iterate over them in some arbitrary order (for instance we could iterate over binary strings of length $\lceil log(i) \rceil + i + (i \cdot |\Sigma|) \lceil log(i) \rceil$, where first $\lceil log(i) \rceil$ bits encode the starting state q', next i bits encode set F' and final $(i \cdot |\Sigma|) \lceil log(i) \rceil$ encode the transition function δ').

Now we will prove that this Algorithm decides L_2. Previous discussion guarantees that Algorithm terminates. Moreover:

- If $D = (Q, \Sigma, \delta, q_s, F) \in L_2$ then no DFA with the number of states smaller than $|Q|$ will recognize the same language as D, so EQ will always return FALSE and in turn the Algorithm will return TRUE,
- If $D = (Q, \Sigma, \delta, q_s, F) \notin L_2$ then there exists DFA $D' = (Q', \Sigma, \delta', q', F')$ such that $|Q'| < |Q|$ and $L(D) = L(D')$. As the Algorithm checks all DFAs with smaller than D number of states it also, in particular, checks D' for which EQ will return TRUE and in turn the Algorithm returns FALSE.

Common mistake 1

Many solutions were based on ideas of the form:

1. remove unreachable states,
2. merge a set of states S if once reaching S you cannot escape S and all the states in S are either accepting or rejecting,
3. merge indistinguishable states.

These ideas make sense but it's really hard to make them work. For instance: 1) and 2) alone are not enough to rule out some non-minimal DFAs. The problem with 3) is that it's not trivial to come up with a correct definition for indistinguishable states and even with a correct definition an algorithm for removing them is not trivial.

Common mistake 2

Some solutions had the following idea "Iterate over all DFAs that recognize the same language as D and then check if the number of states is smaller". There are two problems with this solution: first is that the number of such DFAs is infinite (For instance all DFAs with $F = \emptyset$ recognize the same (empty) language) so it's not possible to iterate over them, second is that it's not clear how to even generate such a set of DFAs.

The overall grading scheme was the following:

- **[6 points]** Correct solution and proof.
- **[5 points]** Correct solution with mistakes or missing parts in the proof (for instance not saying why the number of DFAs is finite)
- **[1-2 points]** Various ideas that work for some cases (of the form of "Common mistake 1").

(This page is intentionally left blank.)

3 *(6 pts)* Consider the following language

$$L_3 = \{\langle M\rangle \mid \ M \text{ is a TM that halts on every input}\}\,.$$

Show that L_3 is unrecognizable by giving a mapping reduction from the unrecognizable language NR defined as follows:

$$NR = \{\langle M, w\rangle \mid \ M \text{ is a TM and M does not reject input } w\}\,.$$

In other words, show that $NR \leq_m L_3$.

*(In this problem you are asked to give a mapping reduction from the language NR to the language L_3 and to provide a formal proof of its correctness. A mapping reduction from another unrecognizable language, such as $\overline{A_{TM}}$, only gives partial points. You are **not** required to prove that L_3 and NR are unrecognizable. Recall that you are allowed to refer to material covered in the class including theorems without reproving them.)*

Solution: As a reduction consider the following computable function f that maps a pair (Turing machine, string) to a Turing machine:

$f(\langle M, w\rangle)$ = the following Turing machine:

«On input x:

1. Simulate M with w as input for $|x|$ steps.
2. If M rejects w (in $|x|$ steps or less) then go into an infinite loop.
3. Otherwise accept x. »

First if $\langle M, w\rangle \in NR$, then M either accepts w or loops with w as input. In any of these cases, for any input x the if condition in line 2 of $f(\langle M, w\rangle)$ will not be satisfied. Hence for any input x, $f(\langle M, w\rangle)$ will accept x hence $f(\langle M, w\rangle) \in L_3$.

Second if $\langle M, w\rangle \notin NR$ then M rejects w in k steps (for some $k > 0$). Hence, for all inputs x such that $|x| \geq k$, the if condition in line 2 will be satisfied therefore $f(\langle M, w\rangle)$ loops on all inputs x such that $|x| \geq k$. In particular $f(\langle M, w\rangle) \notin L_3$.

Notice that f is indeed computable (it is just the definition of a Turing machine) and we just proved that

$$\langle M, w\rangle \in NR \iff f(\langle M, w\rangle) \in L_3$$

which ends the proof.

(This page is intentionally left blank.)

4 *(6 pts)* Classify the following language into one of: decidable, undecidable but recognizable, unrecognizable.

$$L_4 = \{\langle M, D\rangle \mid \; M \text{ is a TM, } D \text{ is a DFA and } L(D) \cap L(M) \neq \phi\}\,.$$

Justify your answer with a formal proof.

(In this problem, you are asked to identify whether L_4 is (decidable and recognizable), (undecidable and recognizable), or (unrecognizable) and provide a formal correctness proof. Recall that you are allowed to refer to material covered in the class including theorems without reproving them.)

Solution:

Added 2025: See your Homework 2, Problem 1 :)