

Exercise III, Theory of Computation 2025

These exercises are for your own benefit. Feel free to collaborate and share your answers with other students. Solve as many problems as you can and ask for help if you get stuck for too long. Problems marked * are more difficult but also more fun :).

These problems are taken from various sources at EPFL and on the Internet, too numerous to cite individually.

- 1 Prove that the language $L = \{0^{n^2}1^n \mid n \geq 0\}$ is not regular.

Solution: The only technique to prove that a given language is not regular, that we know of, is using the pumping lemma: Assume for the sake of contradiction that L is regular. Then by the pumping lemma, L has a pumping length p . Now consider $w = 0^{p^2}1^p \in L$. The pumping lemma now tells us that we can decompose w as

$$w = xyz, \text{ where } |xy| \leq p, |y| \geq 1 \text{ and } xy^iz \in L \text{ for all } i \geq 0.$$

Since w starts with $p^2 \geq p$ zeros, xy and therefore also y consist entirely of only zeros. Since $|y| \geq 1$, we have $y = 0^k$ for some integer $k \geq 1$. But now if we just choose $i = 2$ for example, we deduce that

$$xy^2z = 0^{p^2+k}1^p \in L,$$

a contradiction since $p^2 + k > p^2$. We conclude that L is not regular. In fact, any choice of $i \neq 1$ would have done the job as well.

- 2 Prove that the language $L = \{0^n1^m \mid 0 \leq n \leq m \text{ OR } 0 \leq 2m \leq n\}$ is not regular.

Solution: We again use the pumping lemma. Assume for the sake of contradiction that L is a regular language, let p be its pumping length, and consider

$$w = 0^{p+1}1^{p+1} \in L.$$

By the pumping lemma, w can be decomposed as

$$w = xyz, \text{ where } |xy| \leq p, |y| \geq 1 \text{ and } xy^iz \in L \text{ for all } i \geq 0.$$

Since w starts with $p+1$ zeros, $y = 0^k$ for some integer k with $1 \leq k \leq p$. But now, choosing $i = 2$, we find that

$$xy^2z = 0^{p+1+k}1^{p+1} \in L.$$

However, $p+1 < p+1+k < 2(p+1)$ for any $1 \leq k \leq p$, which contradicts the definition of L . Hence, L is not regular. Unlike in the previous exercise, here the choice of $i = 2$ was crucial.

- 3 Let L be language over the singleton alphabet $\{1\}$ consisting of all the strings whose lengths are prime numbers. Thus, $L = \{11, 111, 11111, 1111111, \dots\}$. Prove that L is not regular.

Solution: Once again, we will make use of the pumping lemma. Suppose that L is a regular language with pumping length p . Let q be any prime number strictly greater than p , which exists since the set of prime numbers is infinite. Now $w = 1^q \in L$ can be decomposed as

$$w = xyz, \text{ where } |xy| \leq p, |y| \geq 1 \text{ and } xy^iz \in L \text{ for all } i \geq 0.$$

Let us write $y = 1^k$ for some integer $k \geq 1$. But now, for any $i \geq 0$,

$$xy^iz = 1^{q+(i-1)k} \in L.$$

But if we choose $i = q+1$ then $q+(i-1)k = q+qk = q(k+1)$ is clearly composite, a contradiction. We conclude that L is not regular.

Note that in this exercise, the choice of i in the pumping lemma was more subtle than in the previous two. In particular, it depended on the initial string $w \in L$ we chose.

- 4 A *palindrome* is a string that reads the same forwards and backwards.

Let L be the language over the alphabet $\{0, 1\}$ consisting of all the palindromes. For example, we have $00 \in L$, $10101 \in L$ and $\varepsilon \in L$. Prove that L is not regular.

Solution: To no one's surprise, we again use the pumping lemma. Suppose for the sake of contradiction that L is regular and let p be the pumping length of L . Consider the string $w = 0^p 1 0^p \in L$. By the pumping lemma, w can be decomposed as

$$w = xyz, \text{ where } |xy| \leq p, |y| \geq 1 \text{ and } xy^iz \in L \text{ for all } i \geq 0.$$

Since $|xy| \leq p$, we have $y = 0^k$ for some $k \geq 1$ and choosing $i = 0$, we obtain

$$xz = 0^{p-k} 1 0^p \in L,$$

which is a contradiction since the single 1 is not in the middle of the string. Hence, we can conclude that L is not regular.

Note that any choice of $i \neq 1$ would have done the job as well, and there are many different initial strings $w \in L$ one could have considered.

5 For any string w over the alphabet $\{0, 1\}$, define its *balance* by

$$\text{bal}(w) = (\text{number of 1s in } w) - (\text{number of 0s in } w).$$

5a For some language L over the alphabet $\{0, 1\}$, define

$$B_L = \{w \in \{0, 1\}^* \mid \text{bal}(w) = \text{bal}(u) \text{ for some } u \in L\}.$$

If L is regular, must B_L be regular as well? What about the other direction?

5b* Is the following language regular?

$$L = \{w \in \{0, 1\}^* \mid -5 \leq \text{bal}(s) \leq 5 \text{ for every substring } s \text{ of } w\}$$

Hint: Keep track of the extreme values of bal for substrings ending at the current position.

Solution:

5a The answer to the first question is no: There exists a regular language L such that B_L is not regular. To prove this, consider the regular language $L = \{\varepsilon\}$. In this case, B_L is the set of strings that contain an equal number of zeros and ones. We can now use the pumping lemma to show that B_L is not regular: Assume for the sake of contradiction that B_L is regular and let p be the pumping length of B_L . Consider $w = 0^p 1^p \in B_L$ and note that the pumping lemma guarantees that $0^{p+k} 1^p \in B_L$ for some $k \neq 0$, a contradiction.

The answer to the second question is no as well: There exists a non-regular language L for which B_L is regular. Consider the language

$$L = \{0^n 1^m \mid 0 \leq n \leq m \text{ OR } 0 \leq 2m \leq n\}$$

from exercise 2, where we have seen that L is not regular. But now note that for every $k \in \mathbb{Z}$, there is some $u \in L$ with $\text{bal}(u) = k$, given by

$$u = \begin{cases} 1^k \in L & \text{if } k \geq 0 \\ 0^{2|k|} 1^{|k|} \in L & \text{if } k < 0. \end{cases}$$

Thus, we have $B_L = \{0, 1\}^*$ which is clearly a regular language.

5b Yes, this language is regular. The key insight is that in order to check membership in L , all we need to keep in memory while traversing the string is the smallest and largest possible balance of a substring ending at the current position. If these values never leave the interval $[-5, 5]$ until we reach the end of the string, then we accept our input, otherwise, we reject it. For a given position in some string $w \in L$, let x and y denote the smallest and largest balance of all substrings of w ending in that position. Note that $x \leq 0$ and $y \geq 0$ since we can always take into consideration the empty string which has balance 0. Also note that $|x - y| \leq 5$ by definition of L , since otherwise the set-difference of the two witnessing substrings is a substring with disallowed balance.

Formally, construct the DFA $M = (Q, \{0, 1\}, \delta, q_0, F)$ with

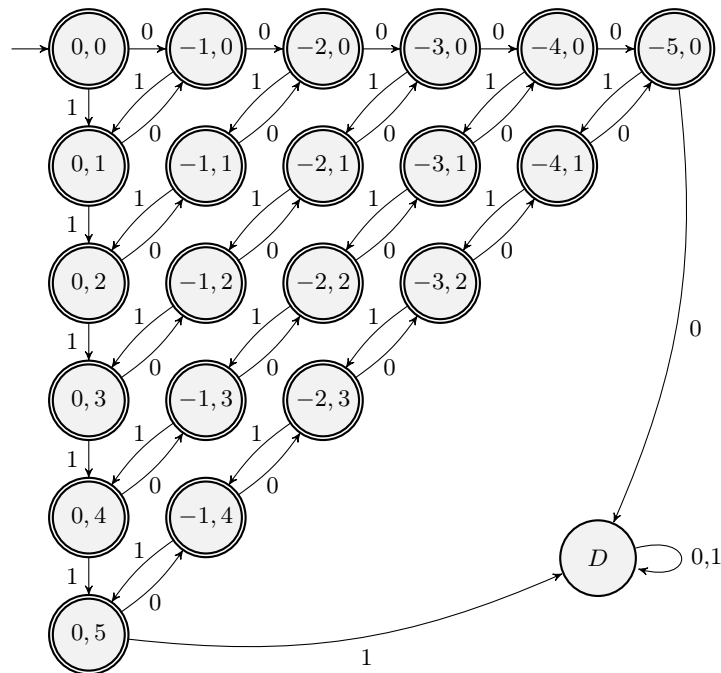
- states $Q = \{(x, y) \in \mathbb{Z}^2 \mid x \leq 0, y \geq 0, |x - y| \leq 5\} \cup \{D\}$,
- transition function $\delta : Q \times \Sigma \rightarrow Q$ given by

$$\delta((x, y), a) = \begin{cases} (x - 1, \max\{y - 1, 0\}) & \text{if } a = 0, x > -5 \\ D & \text{if } a = 0, x = -5 \\ (\min\{0, x + 1\}, y + 1) & \text{if } a = 1, y < 5 \\ D & \text{if } a = 1, y = 5 \end{cases} \quad \text{and } \delta(D, a) = D$$

for all $(x, y) \in [-5, 0] \times [0, 5]$ and both $a \in \{0, 1\}$,

- start state $q_0 = (0, 0) \in Q$, and
- accepting states $F = Q \setminus \{D\}$.

The automaton is illustrated below.



If $x \in L$, then M never goes to D and thus accepts. Conversely, if $x \notin L$ then this is witnessed by some substring s of w with $\text{bal}(s) \notin [-5, 5]$ and at the latest after reading the last symbol of s , M will go to state D and thus reject. We conclude that $L(M) = L$ and therefore that L is regular.

6* Let L be a language that satisfies the condition of the pumping lemma. Must L be regular?

Hint: *Modify the pumping lemma to look at the back of the string instead of the front.*

Solution: The answer is no, we can construct a language that satisfies the condition of the pumping lemma but is not regular. But how are we going to show that our constructed language is not regular if we cannot use the pumping lemma to get a contradiction? Well, since the pumping lemma is the only tool we know for proving a language not regular, the only option is to reuse the ideas behind the pumping lemma but apply them slightly differently.

Consider the language

$$L = \{u01^p \mid u \in \{0, 1\}^*, p \text{ is a prime number}\} \cup \{1\}^*.$$

It is not hard to check that L satisfies the condition of the pumping lemma with pumping length $p = 1$, by showing that any word in L stays in L after pumping up (or removing) the first symbol.

Now note that in the proof of the pumping lemma we made the somewhat arbitrary choice to look for a repeating state at the beginning of the DFA computation. More generally, if p is the pumping length of L , there must be a repeating state among any p consecutive steps of the computation. In particular, we can get another version of the pumping lemma that finds a substring to pump towards the **end** of the string:

Backwards Pumping Lemma. For any regular language A , there is an integer $p \geq 1$ such that every $w \in A$ can be written $w = xyz$ with $|yz| \leq p$, $|y| \geq 1$ and $xy^iz \in A$ for all $i \geq 0$.

But this version of the pumping lemma now shows that L is not regular: If it were and had pumping length p , then for some prime $q \geq p$, we could pump up the end of $01^q \in L$ to obtain $01^k \in L$ for some composite k , just like in exercise 3. Thus, we conclude that while satisfying the condition of the original pumping lemma, the language L is not regular.