Syntax and Semantics
Exercise Session 5

Before presenting the solutions of the exercises, we recall the statement of the pumping lemma and we describe how it can be used to prove that a language is not regular.

Pumping Lemma. *If $L \subseteq \Sigma^*$ is a regular language, then there exists a number $p \in \mathbb{N}$, called* pumping length *for L, such that for any word $w \in L$ with $|w| \geq p$, there exist $x, y, z \in \Sigma^*$ such that $w = xyz$ and the following hold*

(i) $xy^iz \in L$ *for any* $i \in \mathbb{N}$.

(ii) $|y| > 0$

(iii) $|xy| \leq p$

How to use the Pumping Lemma: The pumping lemma is often used to prove that a language L is not regular. This is done by contradiction (i.e., showing that the assumption that L is regular leads to a contradiction) following the steps listed below:

1. Assume that L is regular and that $p \in \mathbb{N}$ is a pumping length for L;
2. Choose a word $w \in L$ s.t. $|w| \geq p$ (notice that w depends on p);
3. Show that there is no way to chose $x, y, z \in \Sigma^*$ and satisfy $w = xyz$, (i), (ii) and (iii) at the same time.

In the last step it is often the case that one shows that in order to fulfill $w = xyz$, (ii) and (iii), the substrings x and y get a shape that allows one to find an index $j \in \mathbb{N}$ such that $xy^jz \notin L$, making (i) impossible to hold.

Remark. *We recall that regular languages are closed under complement, intersection, union, concatenation, and Kleene star. Sometimes it can be helpful to exploit such closure properties in order to either make easier the application of the pumping lemma or reduce the problem to a well known one.*

Take for instance the language $L \subseteq \{a, b, c\}^*$ defined as

$$L = \left\{ w \in \{a, b, c\}^* \,\middle|\, \begin{array}{l} \textit{if } w \textit{ has an odd number of } a\textit{'s} \\ \textit{then its has as many } c\textit{'s as } b\textit{'s} \end{array} \right\}$$

We proceed by contradiction. Assume L to be regular. Since regular languages are closed by intersection we have that $L' = L \cap \mathcal{L}(c^* ab^*)$ is regular. Notice that

$$L' = \{c^n ab^n \mid n \geq 0\} .$$

Following the steps listed above one can easily prove that L' is not regular[1]*. This leads to a contradiction thus L cannot be regular* □

Exercise 1.
Prove that if a regular language $L \subseteq \Sigma^*$ has a pumping length p, then any $q \geq p$ is also a pumping length for L.

Solution 1.
Consider $w \in L$ such that $|w| \geq q$. By hypothesis $p \leq q$, therefore $|w| \geq p$. By hypothesis p is a pumping length of L, hence there exist $x, y, z \in \Sigma^*$ such that $|xy| \leq p$, $|y| > 0$ and $xy^i z \in L$ for all $i \in \mathbb{N}$. By $p \leq q$, we have also that $|xy| \leq q$. This proves that q is a pumping length for L. □

Exercise 2.
Prove that the following languages are not regular

$$\begin{aligned} L_1 &= \{a^i b^j c^k \mid i, j, k \geq 0 \text{ and if } i = 1 \text{ then } j = k\} \\ L_2 &= \{0^n 1^m 0^n \mid m, n \geq 0\} \\ L_3 &= \{0^m 1^n \mid m \neq n\} . \end{aligned}$$

[1] Left as exercise.

Solution 2.

(L_1) Assume that L_1 is regular. Regular languages are closed by intersection thus $L = L_1 \cap \mathcal{L}(a(b \cup c)^*)$ is regular. Observe that

$$L = \{ab^n c^n \mid n \in \mathbb{N}\} .$$

Let $p \geq 1$ be a pumping length for L and consider the word $w = ab^p c^p$. By construction $w \in L$ and $|w| \geq p$. Take $x, y, z \in \{a, b, c\}^*$ such that $w = xyz$, $|xy| \leq p$ and, $|y| > 0$. By construction we have that xy is of the form ab^q for some $q \leq p$. We distinguish two cases:

$x = \varepsilon$) we have that $y = ab^q$, therefore $xy^0z = z$ does not begin with the symbol a. Thus $xy^0z \notin L$.

$x \neq \varepsilon$) we have that $y = b^r$ for some $0 < r \leq q$. Therefore $xy^0z = xz$ is of the form $ab^{p-r}c^p$ where $p - r < p$. Thus $xy^0z \notin L$.

By the pumping lemma this contradicts the fact that L is regular, from which it follows that L_1 isn't regular either.

(L_2) Assume that L_2 is regular and that $p \geq 1$ is a pumping length for it. Consider the word $w = 0^p10^p$. By construction $w \in L_2$ and $|w| \geq p$. Take $x, y, z \in \{0, 1\}^*$ such that $w = xyz$, $|xy| \leq p$ and, $|y| > 0$. By $|xy| \leq p$ and $|y| > 0$, we have that y is of the form 0^q for some $0 < q \leq p$. For this reason, the word $xy^0z = xz$ is of the form $0^{p-q}10^p$ where $p - q < p$. Therefore $xy^0z \notin L_2$. By the pumping lemma, this contradicts the assumption that L_2 is regular.

(L_3) Assume that L_3 is regular. Since regular languages are closed by complement, we have that $\bar{L}_3$ is regular as well. Let then $p \geq 1$ be a pumping length for $\bar{L}_3$ and take the word $w = 0^p1^p$. By construction $w \in \bar{L}_3$ and $|w| \geq p$. Take $x, y, z \in \{0, 1\}^*$ such that $w = xyz$, $|xy| \leq p$ and, $|y| > 0$. By construction, $y = 0^q$ for some $0 < q \leq p$. Thus, the word xy^0z is of the form $0^{p-q}1^p$. Since $p > 0$ we have that $xy^0z = 0^{p-q}1^p \in L_3$, thus $xy^0z \notin \bar{L}_3$. By the pumping lemma, this contradicts the fact that $\bar{L}_3$ is regular, which is in contradiction with the assumption that L_3 is regular. $\square$

Exercise 3.
Consider the alphabet Σ and let $L \subseteq \Sigma^*$. We say that two strings $x, y \in \Sigma^*$ are indistinguishable by L, and we write $x \equiv_L y$, if for any string $z \in \Sigma^*$, $xz \in L$ if and only if $yz \in L$.

Prove that $\equiv_L$ is an equivalence relation.

Solution 3.
We have to show that $\equiv_L$ is reflexive, symmetric, and transitive.

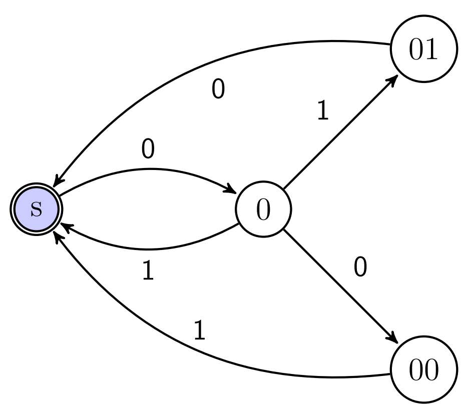
Reflexive Immediate by definition of $\equiv_L$.

Symmetric Immediate by definition of $\equiv_L$.

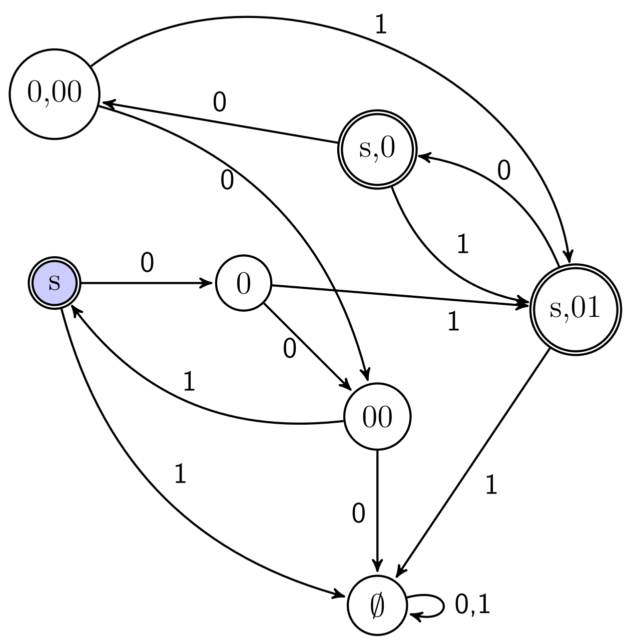
Transitive Consider $x, y, z \in \Sigma^*$ and assume that $x \equiv_L y$ and $y \equiv_L z$. By definition we have that for all $w, w' \in \Sigma^*$, $xw \in L$ iff $yw \in L$ and, $yw' \in L$ iff $zw' \in L$. This implies that, for all $w \in \Sigma^*$, $xw \in L$ iff $zw \in L$. Thus, by definition, $x \equiv_L z$. $\square$

Exercise 4.
Construct an NFA that recognizes the language $(01 \cup 001 \cup 010)^*$. Convert this NFA to an equivalent DFA.

Solution 4.
Here is an NFA that recognizes $(01 \cup 001 \cup 010)^*$:



Here is an equivalent DFA:



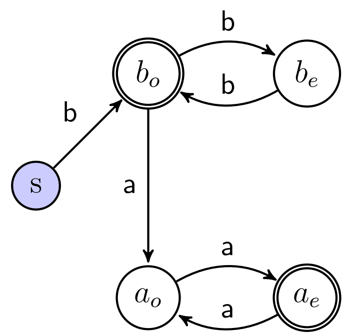
$\square$

Exercise 5.
Consider the language over the alphabet $\Sigma = \{a, b\}$.

$$L = \left\{ w \,\middle|\, \begin{array}{l} w \text{ has an even number of } a\text{'s and an odd number of } b\text{'s,} \\ \text{and } w \text{ does not contain the substring } ab \end{array} \right\}.$$

Prove that L is regular. Find a regular expression that generates L.

Solution 5.
Here is an NFA that recognizes L



From the above NFA one can see that $L = b(bb)^*(aa)^*$. $\square$