CPSC 351 — Tutorial Exercise #5
Nonregular Languages

About This Exercise

The goal of this exercise is to help you to understand, and give you practice in using, various techniques for proving that languages are not regular.

Getting Started

These initial problems will probably not be discussed during the tutorial. Please discuss them during office hours with the instructor, if you can, if you have trouble solving them.

1. Let $\Sigma = \{\text{a}\}$ and let $L = \{\omega \in \Sigma^\star \mid \text{the length of } \omega \text{ is even}\}$. Please identify the ***error*** that has been made, in the following incorrect proof, as precisely as you can.

   Claim: The above language, L, is not regular.

   Proof. Suppose, to obtain a contradiction, that L is a regular language.
   Then it follows, by the Pumping Lemma, that there is a number $p \geq 1$ such that if s is any string in L with length at least p, then s can be divided into three pieces, $s = xyz$ (for $x, y, z \in \Sigma^\star$) satisfying the following three conditions.

   1. $xy^iz \in L$ for every integer i such that $i \geq 0$.
   2. $|y| > 0$ (so that $y \neq \lambda$).
   3. $|xy| \leq p$.

   Consider the string $s = a^{2p}$ where $p \geq 1$ is the number whose existence follows by the above claim.

   - Since $2p$ is an even number, $s \in L$.
   - Since $p \geq 0$, $|s| = 2p \geq p$.

It now follows, by the above, that there exist strings $x, y, z \in \Sigma^\star$ such that $s = xyz$ and properties #1, #2 and #3, above, are satisfied.

With that noted, let $x = \lambda$, let $y = \mathtt{a}$, and let $z = \mathtt{a}^{2p-1}$. Then $|y| = 1 > 0$, so that property #2 is satisfied, and $|xy| = 1 \leq p$, so that property #3 is satisfied too. However, if $i = 0$ then i is an integer such that $i \geq 0$ and

$$xy^iz = xy^0z = \lambda \cdot \lambda\lambda\mathtt{a}^{2p-1} = \mathtt{a}^{2p-1} \notin L,$$

since $|xy^iz| = 2p - 1$ is not an even number in this case. Thus property #1 is not satisfied and a ***contradiction*** has been obtained.

It follows that the original assumption must be false — and the above language L is not regular, as claimed. □

2. Say whether each of the following statements is ***true*** or ***false***. Then write it using mathematical notation (so that the symbols "$\exists$" and "$\forall$" will be used); you may assume that it is "understood" that the types of values are integers here — you do not need to include this in the expressions that you write.

   (a) "For every integer x, there exists an integer y such that x is strictly less than y."

   (b) "There exists an integer y such that, for every integer x, x is strictly less than y".

   Note: The main difference between the sentences given in parts (a) and (b) is the order in which quantified variables are introduced. What (if anything) does suggest that something that you need to be careful about, when using results like "The Pumping Lemma", to prove things?

3. Let $\Sigma = \{\mathtt{a}, \mathtt{b}\}$ and let $L = \emptyset$. Please identify the ***error*** that has been made, in the following incorrect proof, as precisely as you can — assuming that you have already proved that the languages L_1 and L_2 are not regular. (This error is similar to one that students have made, when writing proofs that languages are not regular, in the past.)

   Claim: The above language, L, is not regular.

   Proof. Recall that the languages

   $$L_1 = \{\mathtt{a}^n\mathtt{b}^n \mid n \in \mathbb{Z} \text{ and } n \geq 0\}$$

   and

   $$L_2 = \{\mathtt{a}^n\mathtt{b}^m \mid n \in \mathbb{Z}, m \in \mathbb{Z}, n, m \geq 0 \text{ and } n \neq m\}$$

   are languages such that $L_1 \subseteq \Sigma^\star$, $L_2 \subseteq \Sigma^\star$, L_1 is not regular, and L_2 is not regular. Since $L_1 \cap L_2 = \emptyset = L$ it follows that the language $L = \emptyset$ is not regular, as well. □

Problems To Be Discussed

As time permits, the following problems will be discussed during the tutorial.

4. Let $\Sigma = \{\mathtt{a}, \mathtt{b}\}$. Using the "Pumping Lemma", prove that the language

$$L = \{\mathtt{a}^n\mathtt{b}^m \mid m < n\}$$

is not a regular language (where "$<$" represents the relation "strictly less than" so that, for example, it is not true that $3 < 3$).

5. Let $\Sigma = \{\mathtt{a}\}$ and and suppose that the language

$$L_p = \{\mathtt{a}^n \mid n \text{ is a prime number}\} \subseteq \Sigma^\star$$

is *not* a regular language. Recall, as well, that a positive integer n is ***composite*** if $n \geq 2$ and there exist integers k and ℓ such that $2 \leq k, \ell \leq n-1$ and $k \times \ell = n$. Every positive integer n, such that $n \geq 2$, is either prime or composite — but not both. On the other hand, the integers 0 and 1 are neither prime nor composite.

Let $L_c \subseteq \Sigma^\star$ be the language

$$L_c = \{\mathtt{a}^n \mid n \text{ is composite}\}.$$

Use this information, and one or more of the ***closure properties*** for regular languages that have now been introduced in this course, to prove that L_c is *not* a regular language.

The next problem is a bit more challenging — and requires you think, more carefully, about how to choose the number "i" that is used in a proof that uses the Pumping Lemma. You may use the fact that there are infinitely many primes — so that there is no integer is larger than all of them — without proving it.

6. Let $\Sigma = \{\mathtt{a}\}$. Prove that the language

$$L = \{\mathtt{a}^n \mid n \text{ is a prime number}\} \subseteq \Sigma^\star$$

is not a regular language.

Note: There are infinitely many prime numbers — so there is no integer that is larger than all of them. You may use this fact without proving it.