THEORY OF COMPUTATION

Pumping Lemma

Dr. Faheem Masoodi
masoodifahim@uok.edu.in

Pumping Lemma.

An interesting question about languages is that whether there are some languages that cannot be recognized by a Finite Automata or for which, we cannot construct a regular expression and the answer to this all-important question can be found using a tool called Pumping lemma that was discovered in 1961 by Yeshoshua, Perles and Shamir. Pumping lemma for regular language is a theorem that assists us in establishing that a certain language is not regular but cannot help us to prove whether a language is regular. Pumping lemma proves that a particular language is not regular by using proof by contradiction. An intrinsic property of these non-regular languages is that they require some sort of memory for counting instances in a string.. For example a typical example of a non-regular language is $L = \{a^n b^n | \ n \ \geq 0\}$ in which some number of *a's* is followed by exactly the same number of *b's*. Now in order to guarantee the same number of b's, we need to count and store the information about number of a's in the string. Recall that Finite Automata have no auxiliary memory and are not capable of storing any information and therefore cannot be used to recognize non-regular languages. In this chapter, we will study the pumping lemma theorem and try to understand, how by way of contradiction, it can used to prove that a language is not regular.

Pumping Lemma Definition: For every regular L, $\exists\ p \geq 0$; s.t $\forall\ w \in L$, where $|w| \geq p$, $\exists$ a partition w=*abc*, satisfying the following three conditions:

- $|ab| \leq p$
- $|b| > 0$
- $\forall\ i \geq 0,\ ab^ic \in L$

If L is a regular language, then there exists a positive integer '*p'* known as pumping length, such that for all substrings $w \in L$, where $|w| \geq p$, there exists a partition of $w=abc$ such that $|ab| \leq p$ and $|b| > 0$ and for all $i \geq 0$, $ab^ic \in L$. Here *w* is partitioned into three parts *a, b and c* and length of *a and b* cannot exceed the pumping length, *a or z* may be null string but $b \neq \wedge$ and the middle part b^i allows *i* instances of b, that can be pumped into the string and $b^0 = \wedge$.

So for every regular language, the above statement holds true. In order to prove that a particular language is not regular, we will use the contrapositive pumping lemma.

Contrapositive of Pumping Lemma:

Since it is compulsory for every regular language to satisfy all conditions of pumping lemma, we try to form the contrapositive of pumping lemma by negating the conditions of actual theorem and thereby prove that the language is not regular. Let L be some language ($L \in \sum^*$). If for all $p \geq 0$, there exists any word $w \in L$, such that $|w| \geq p$ and for all partitions, of $w=abc$ such that $|ab| \leq p$ and $|b| > 0$ and there exists $i \geq 0$, such that $ab^ic \notin L$.

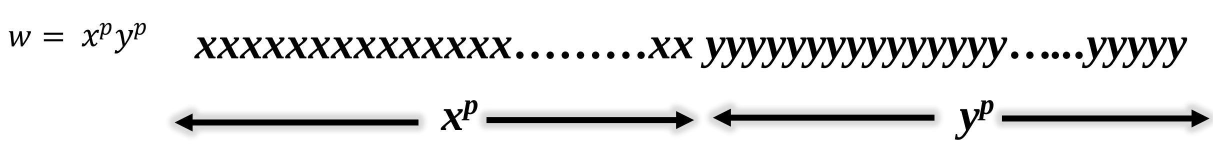
Lets analyze this contrapositive statement by comparing it to actual pumping lemma statement. The primary four differences have been summarized in that Table 1 below. While as in pumping lemma, we check if there exists any $p \geq 0$, in contrapositive, we check for all $p \geq 0$. In pumping lemma, we look for all $w \epsilon L$ strings, in contrapositive, we choose any one string. Additionally in pumping lemma, we check if there exists any partition $w=abc$, in contrapositive, we look for all partitions $w=abc$. Finally we conclude that for all i, ab^ic belongs to a regular language, while as in contapositve, we conclude that there exists some i, for which ab^ic does not belong to a regular language

Example 4.1

Let's consider the language $L = \{x^n y^n | \ n \ \geq 0\}$ discussed in the introduction of this chapter. This language contains strings that consist of sequences of *0's* followed by equal number of *1's*. Intuitively, a count of number of *0's* will have to be maintained so that number of *1's* can be verified. We will try to use the contrapositive statement of pumping lemma to prove that this language is not regular. We begin with the assumption that L is a regular language. Given a pumping length $p \geq 0$ by the pumping lemma, we choose any string $w \epsilon L$ s.t $|w| \geq p$, that will be partitioned as $w=abc$, where a,b *and* c can assume any value subject to the constraints that $|ab| \leq p$ *and* $|b| > 0$. Once the string is partitioned, we can choose any value of i such that, $ab^ic \notin L$. i.e. the language is not a regular.

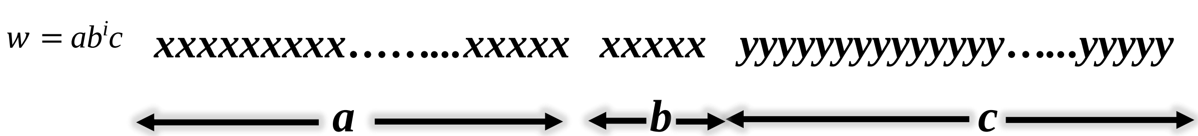
Here $\quad L = \{a^n b^n | \ n \ \geq 0\}$

Given $p \geq 0$, Choose a string $w = \ x^p y^p$



Next Partition $w=abc$, where a,b *and* c can assume any value subject to the constraints that

- $|ab| \leq p$ *and*
- $|b| > 0$



Now Choose a value for i, such that ab^ic is not in the given language. Lets consider a $i=0$, Then

$$ab^0c = ac \Rightarrow x^q y^p,$$

$$\text{where } q = p - |y| \neq p$$

Therefore, $ab^ic \notin L$ and we can conclude that language L is not regular.