

Lecture #9: Nonregular Languages, Part One

Key Concepts

Pumping Lemma: Let Σ be an alphabet and let $A \subseteq \Sigma^*$.

If A is a regular language, then there is a number $p \geq 1$ (called the **pumping length** for A) — which only depends on A — such that if s is any string in A with length at least p , then s can be divided into three pieces $s = xyz$ (for $x, y, z \in \Sigma^*$), satisfying the following three conditions.

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

Note: y^i is the concatenation of i copies of the string y .

A Process to Prove that a Language $L \subseteq \Sigma^*$ is not Regular:

1. Assume — to obtain a contradiction — that L is a regular language.
2. Observe that the conditions in the Pumping Lemma (with L used as the language called “ A ”) are satisfied.
3. Introduce the “pumping length” — which is generally named p . Note that **you do not get to choose this value** — you can only assume that it exists (and is a positive integer).
4. Introduce a string in Σ — which is usually named s , because that is what is called in the statement of the Pumping Lemma.

You do get to choose this string — and you will need to think about the language, L , that you are working with — and its definition will almost always depend on the (unknown) pumping length p .

5. Show that $s \in L$ and $|s| \geq p$.

6. Conclude that there exist strings $x, y, z \in \Sigma^*$ such that $s = xyz$ and properties #1, #2 and #3, given in the statement of the Pumping Lemma (again, with L used in place of A) all hold.

Note that **you do not get to choose** the strings x , y , and z — you may only assume that these strings exist.

7. Use properties #1, #2 and #3 to get a contradiction **without making any other assumptions**.
8. Conclude that — since only one assumption was made — this assumption must be false. In other words, L is *not* a regular language.

More About This Process

- When carrying out this process, you *do* get to choose the value for the integer i , mentioned in property (a). It is frequently — but not always — possible to choose i to be either 0 or 2, when you are using this process. However, it is sometimes necessary to think more about the language L that is being considered, and to choose i to be something else.
- Note how the details in this process (including the order in which things are introduced and which values you get to choose, along with which value you *do not* get to choose) reflect the structure of the Pumping Lemma.

In the lecture notes, this process was used to show that if $\Sigma = \{a, b\}$ and

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

then L is *not* a regular language.