Introduction to Computability Theory

Lecture 4: Non Regular Languages
Prof. Amos Israeli

Lecture Outline
1. Motivate the Pumping Lemma.
2. Present and demonstrate the pumping concept.
3. Present and prove the Pumping Lemma.
4. Use the pumping lemma to prove that some languages are not regular.

Introduction and Motivation
In this lecture we ask: Are all languages regular? The answer is negative. The simplest example is the language $B = \{a^n b^n | n \geq 0\}$.

Try to think about this language.

If we try to find a DFA that recognizes the language $B = \{a^n b^n | n \geq 0\}$, it seems that we need an infinite number of states, to “remember” how many $a$-s we saw so far.

Note: This is not a proof!
Perhaps a DFA recognizing $B$ exists, but we are not clever enough to find it?
The Pumping Lemma is the formal tool we use to prove that the language $B$ (as well as many other languages) is not regular.

Consider the following NFA, denoted by $N$:

It accepts all words of the form $(0 \cup 1)(01)^*$.

What is Pumping?

Consider now the word $110 \in L(N)$.

Pumping means that the word 110 can be divided into two parts: 1 and 10, such that for any $i \geq 0$, the word $1(10)^i \in L(n)$.

We say that the word 110 can be pumped. For $i = 0$ this is called down pumping. For $i > 1$ this is called up pumping.

A more general description would be:

A word $w \in L$, can be pumped if $w = xy$ and for each $i \geq 0$, it holds that $xy^i \in L$.

Note: the formal definition is a little more complex than this one.
The Pumping Lemma

Let $A$ be a regular language. There exists a number $p$ such that for every $w \in A$, if $|w| \geq p$ then $w$ may be divided into three parts $w = xyz$, satisfying:

1. for each $i \geq 0$, it holds that $xy^iz \in A$.
2. $|y| > 0$.
3. $|xy| \leq p$.

Note: Without req. 2 the Theorem is trivial.

Demonstration Continuation

In terms of the previous demonstration we have:

1. $p = 3$
2. For $w = 110$, we get:
   \[
   x = 1, \quad y = 10, \quad z = \epsilon.
   \]

Proof of the Pumping Lemma

Let $D$ be a DFA recognizing $A$ and let $p$ be the number of states of $D$. If $A$ has no words whose length is at least $p$, the theorem holds vacuously. Let $w \in A$ be an arbitrary word such that $|w| \geq p$. Denote the symbols of $w$ by $w = w_0, w_2, ..., w_m$ where $m = |w| \geq p$.

Proof of the Pumping Lemma

Assume that $q_0, q_1, ..., q_p, ..., q_m$ is the sequence of states that $D$ goes through while computing with input $w$. For each $k$, $0 \leq k < m$,\[ \delta(q_k, w_k) = q_{k+1}. \]Since $w \in A$, $q_m \in F_D$.

Since the sequence $q_0, q_1, ..., q_p$ contains $p + 1$ states and since the number of states of $D$ is $p$, that there exist two indices $0 \leq i < j \leq p$ such that $q_i = q_j$. 
Proof of the Pumping Lemma

Denote \( x = w_1 w_2 ... w_{i-1} \), \( y = w_i w_{i+1} ... w_{j-1} \) and 
\( z = w_j w_{j+1} ... w_m \).

**Note:** Under this definition \(|y| > 0\) and \(|xy| \leq p\).

By this definition, the computation of \( D \) on 
\( x = w_1 w_2 ... w_{i-1} \) starting from \( q_i \), ends at \( q_i \).

By this definition, the computation of \( D \) on 
\( z = w_j w_{j+1} ... w_m \), starting from \( q_j \), ends at \( q_m \)
which is an accepting state.

---

Example: \( L = \{a^n b^n \} \).  

**Lemma:** The language \( L = \{a^n b^n \} \) is not regular.

**Proof:** Assume towards a contradiction that \( L \) is
regular and let \( p \) be the **pumping length** of \( L \).

Let \( w = a^p b^p \). By the Pumping Lemma there
exists a division of \( w \), \( w = xyz \), such that
\(|xy| \leq p\), and \( w \) can be pumped.

This means that \( xy = a^k \), where \( k \leq p \).

Since \(|y| > 0\) we conclude \( y = a^l \)

---

Proof of the Pumping Lemma

The computation of \( D \) on \( y = w_i w_{i+1} ... w_{j-1} \)
starting from \( q_i \), ends at \( q_j \). Since \( q_i = q_j \), this
computation starts and ends at the same state.

Since it is a circular computation, it can repeat itself \( k \) times for any \( k \geq 0 \).

In other words: for each \( i \geq 0 \), \( xy^i z \in A \).
Q.E.D.

---

Example: \( L = \{a^n b^n \} \) (Cont.)

This however implies that in \( xy^0 z \), the number
of \( a \)-s is smaller then the number of \( b \)-s.

It also means that for every \( i > 1 \), the number of
\( a \)-s in \( xy^i z \) is larger then the number of
\( b \)-s. Both cases constitute a contradiction.

**Note:** Each one of these cases is separately sufficient for the proof.
**Discussion**

This is what we got so far:

$$\text{RL-s Ex: } \{a^n \mid n \geq 0\}$$

$$\{a^n b^n \mid n \geq 0\}$$

**Lecture Recap**

1. Motivated the Pumping Lemma.
2. Presented and demonstrated the **pumping** concept.
3. Presented and proved the **Pumping Lemma**.
4. Used the pumping lemma to prove that some languages are not regular.