CSC 320 - Lecture 08

#pumping-lemma #regular #languages

What Is The Pumping Lemma?

It states properties that hold for any regular language.

It Says. For every regular language L there is a positive integer p such that for any string in L that is of length at least p, there is a way to rewrite the string as xyz where:

1. $xy^iz \in L$ for all $i \geq 0$, that is: $xz \in L$, $xyz \in L$ (that is given), $xyyz \in L$, $xyyyz \in L$, ...
2. $|y| > 0$, i.e., $y \neq \varepsilon$
3. $|xy| \leq p$, that is substring xy cannot contain more than p symbols

Note. It does not tell us the value of p, we just know that such a natural number exists. (p is the pumping value).

Note. Pumping Lemma holds for *every* regular language.

What Is The Pumping Lemma Good For?

Use pumping lemma as a tool to prove that a certain language L is not regular.

To Prove. We argue by using contradiction.

1. We assume that L is regular
2. Pumping Lemma then guarantees for L that there is a $p \in \mathbb{N}$ such that for every string $s \in L$ of length at least p the three properties hold.

If we can find a string L in that...

- is of length at least p but that
- does **not** satisfy all three conditions listed in the pumping lemma, then we know that L cannot be regular.

Reason. Pumping Lemma holds for *every* regular language.

Note. It is crucial what string of length at least p to choose from L to derive a contradiction.

- Not every long string might lead to a contradiction.
- But just one string in L of length at least p that does not satisfies the conditions of the pumping lemma is sufficient to prove that a language is not regular.
- When choosing a string as counterexample to pumping lemma, think about what properties might make the language not regular.

Example

Let $L = \{0^n 1^n | n \geq 0\}$. Use pumping lemma to prove L is not regular. (The empty string is in the language).

Proof by contradiction. Assume that L is regular. Therefore, all properties of pumping lemma must hold for L, i.e.:

If $s \in L$ and $|s| \geq p$ then we can rewrite $s = xyz$ with...

1. for each $i \geq 0$: $xy^i z \in L$
2. $|y| > 0$ (ie, $y \neq \varepsilon$)
3. $|xy| \leq p$

We will choose $s = 0^p 1^p$. Where s is the string that we hope is a counterexample, that is we must reach a contradiction to the properties of the pumping lemma. For such a counterexample s, length of s must depend on $p : |s| \geq p$ required.

We confirm. $s \in L$ and $|s| \geq p$ (since $|s| = 2p$). Therefore (since L is regular is assumed) pumping lemma guarantees: s can be written as $s = xyz$ with $xyz \in L$ for any $i \geq 0$ and...

1. for each $i \geq 0$: $xy^i z \in L$
2. $|y| > 0$ (ie, $y \neq \varepsilon$)
3. $|xy| \leq p$

Question. Does there exist a rewriting into xyz for $s = 0^p 1^p = xyz$ such that 1., 2., and 3. hold? (Ex. 0...01...1; where 0..0 is of length p and 1...1 is of length p).

We want to rewrite it as xyz; where $y \neq \varepsilon$.

Our Goal. Show such rewriting does not exist.

Note. Just one rewriting of s that satisfies pumping lemma conditions is sufficient to show that string s **does not** serve as counterexample when trying to show that L is not regular. Therefore, we consider all cases to rewrite s:

- **Case 1**: String y consists of 0s only.

  0...01...1; where y is in 0...0. $xyyz \notin L$ (because we pump the number of y's which results in more 0s than 1s).

String xyz has more 0s than 1s. Therefore $xyz \notin L$, violating condition 1 of the pumping lemma. Contradiction.

- **Case 2**: String y consists of 1s only.

  0...01...1; where y is in 1...1. $xyyz \notin L$ (because we pump the number of y's which results in more 1s than 0s).

Just like above the string $xyyz$ has more 1s than 0s. Therefore $xyyz \notin L$, violating condition 1 of the pumping lemma. Contradiction.

- **Case 3**: String y consists of both 0s and 1s.

  0...01...1; where y is 01. $xyyz \notin L$ (because 0...001011...1 $\notin L$).

String $xyyz$ may have the same number of 0s and 1s, but they are out of order (some 1 s occur before 0s). Hence $xyyz \notin L$. Contradiction.

Since we **cannot rewrite** s satisfying the conditions of the pumping lemma, L is **not** regular.

We demonstrated. Pumping Lemma can serve as tool to show certain languages are not regular.

Missing. Proof of correctness for statement of Pumping Lemma.

Correctness of Pumping Lemma (PL)

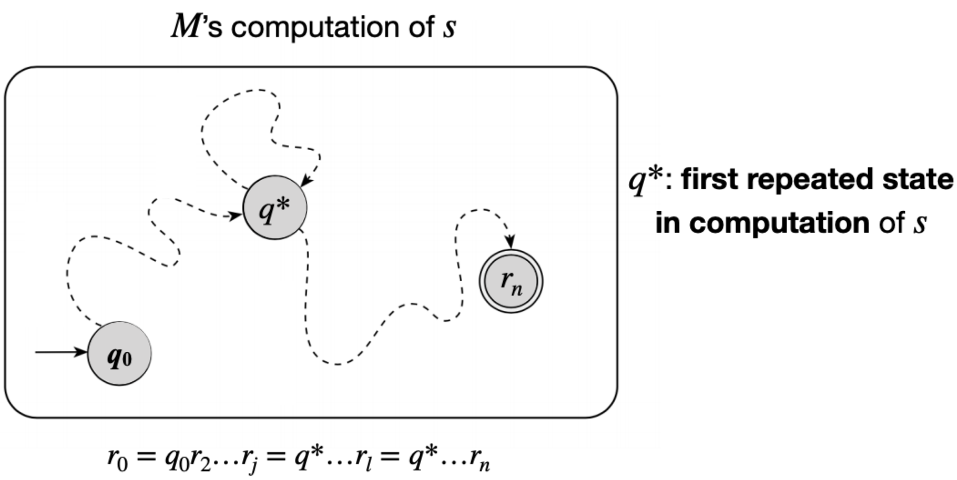
Plan for Proof. What does L look like? Either...

1. no strings in L are of length at least p, or
2. there exist strings in L of length at least p.

- **Case 1**: Three conditions hold for **all** strings of length at least p (therefore for **all strings in** L since no strings in L longer than $p-1$). (CHECKMARK)
- **Case 2**: Consider properties a DFA has for regular language L.

Let DFA M with $L(M) = L$ have p states. Let $s \in L$ and $n = |s| \geq p$. Consider computation sequence $r_0 r_1 \ldots r_n$ of states of M when processing input s.

- Start state, say $r_0 = q_0$, followed by sequence of states until reaching end of s in state, say $r_n = q_n$. Since $s \in L$: $r_n \in F$: $q_n \in F$.
- If $|s| = n$: sequence of states of M computing s has length $n + 1$ with $n + 1 > p = |Q|$.
- Therefore, sequence of states computing s must contain **at least one repeated state**.

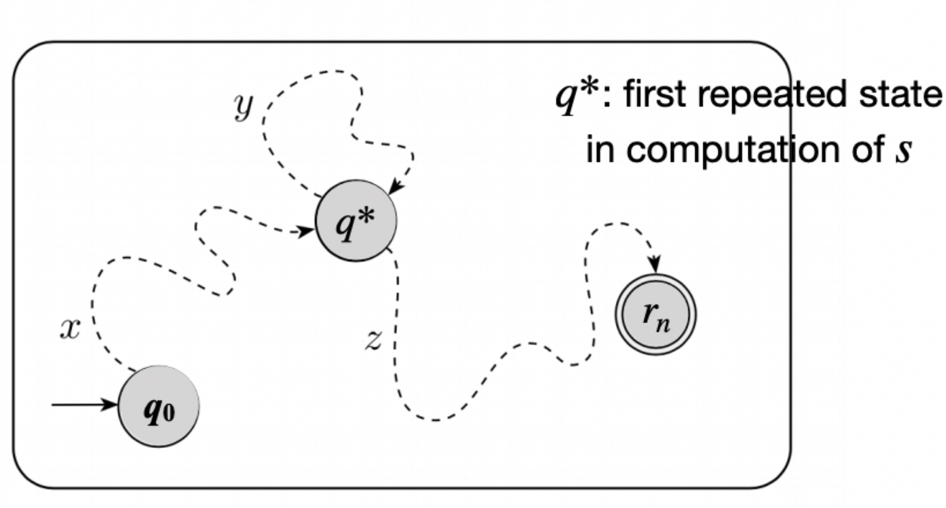


Proof of Pumping Lemma

Show: Can rewrite $s = xyz$

- x: substring of s appearing processed by M before reaching q^*
- y: substring between first two appearance of q^*
- z: suffix s

M's computation of $s = xyz$



Note. This uses a famous concept: **Pigeon Hole Principle**. Fancy name for simple fact that **if p pigeons are placed into fewer than p holes, some hole has to have more than one pigeon in it**.

1. For each $i \geq 0$: $xy^iz \in L$
    - Computing M on $xyyz$ implies $xyyz \in L$
    - Computing xz implies $xz \in L$
    - Computing xy^iz for $i > 2$ implies $xy^iz \in L$
    - Thus $xy^iz \in L$ for $i \leq 0$
2. In DFA computation no state can be repeated without processing at least one symbol (thus if there is a y its length must be greater than 0)
    - $|y| > 0$
3. Must be among first $p + 1$ states.
    - $r_0 = q_0 r_2 \ldots r_j = q^* \ldots r_1 = q^* \ldots r_n$
        - $r_0 = q_0 r_2 \ldots r_j = q^* \ldots r_1 = q^*$ is $l \leq p + 1$
    - $|xy| > p$? (No!)
    - $|xy| \leq p$

Example

Show that $L = \{w \in \{0,1\}^* | w \text{ has an equal number of 0s and 1s}\}$ is not regular. We assume L is regular. Let p be pumping length given by pumping lemma.

Next. Choose string s to achieve contradiction.

Note. **Choice of string -** $s \in L$ of length at least p - to derive a contradiction, **is crucial**. Not every long string might lead to contradiction.

But. **Finding just one string that does not satisfies conditions of pumping lemma proves language is not regular**.

We choose $s = 0^p 1^p$. Where $s \in L$; $|s| \geq p$.

PL guarantees that rewriting of s into xyz, satisfies the three conditions. Because of 3, **no matter the rewriting**: $|xy \leq p|$. (=> y) consists only of 0s (and of at least one 0 due to 2). But $xyz \notin L$. Contradiction to 1.

For 0..01...1 where 0...0 is of length p and 1...1 is of length p, we know that $xy = 0\ldots0$ and $y = 0\ldots0$. But if we do $xyyz$ then we add more 0s but we do not add more 1s.

Example

A possible unlucky choice of s when trying to show that $L = \{w \in \{0,1\}^* | w \text{ has an equal number of 0s and 1s}\}$ is not regular.

$s = (01)^p \in L$ but it cannot be used successfully to achieve a contradiction using the pumping lemma! **String choice is crucial!**

Example

Show that $L = \{ww | w \in \{0,1\}^*\}$ is not regular.

We assume that L is regular.

$\Rightarrow$ Pumping lemmas guarantees that there is a $p \in \mathbb{N}$ with: for all $s \in L$ with $|s| \geq p : s$ can be rewritten as $s = xyz$ with $xyz \in L$ and

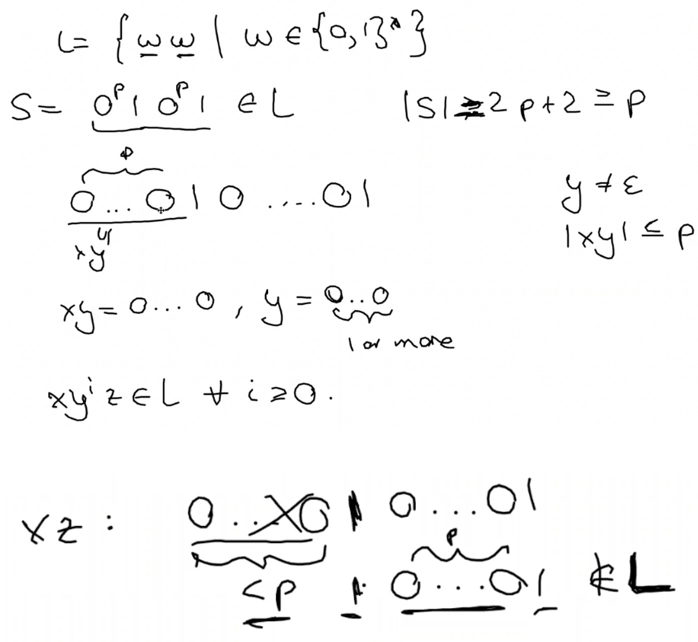
1. for each $i \geq 0$: $xy^iz \in L$
2. $|y| > 0$ (ie, $y \neq \varepsilon$)
3. $|xy| \leq p$

Suggestion for counterexample:

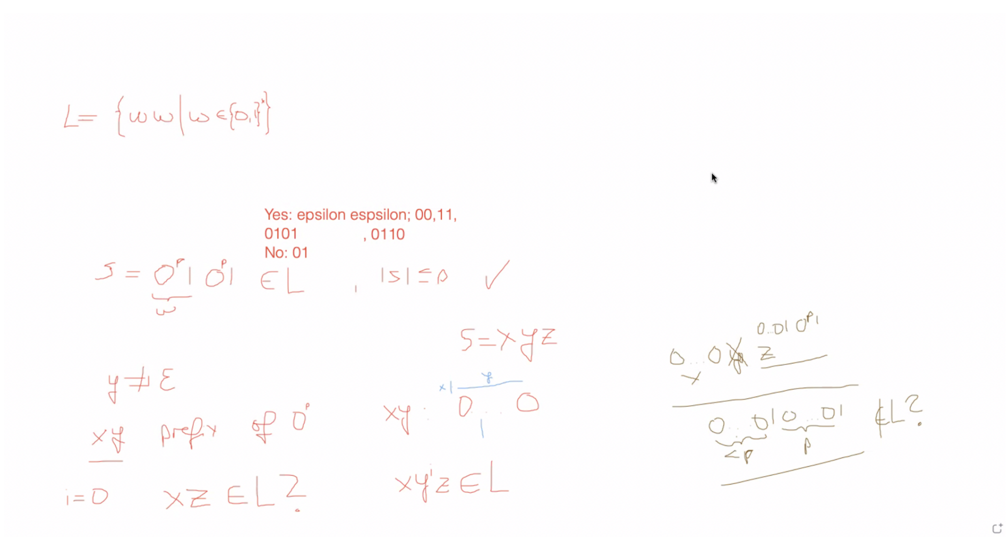
We verify: $s = 0^p 1 0^p 1 = ww$ for $s = 0^p 1 0^p 1 w = 0^p 1$, thus $s \in L$ and $|s| = 2p + 2p \geq p$.

Thus, we show for $s = 0^p 1 0^p 1$ that there is no rewriting for $s = xyz$ such that the three properties hold.

Since $|xy| \leq p$ and $y \neq \varepsilon$ we know that xy is a prefix of 0^p, that is y consists of 0s only, with $1 \leq |y| \leq p$. Consider "pumping up" s to $xyyz$. $xyyz = 0^k 1 0^p 1$ with $k > p$. But $xyyz \notin L$, in contradiction to 1.



Office Hours



Note. 0110 is not in the language.

Example

Show that $L = \{0^i 1^j | i > j\}$ is not regular.

Suggestion. $S = 0^p 1^{p-1}$. However, we will use $S = 0^{P+1} 1^p$. We know $y \neq \varepsilon$ and y is 1 or more 0s. $|xy| \leq p \Rightarrow xy$ consists of 1 or more 0s but no more than p. y consists of no more than p 0s. Let us try...

- $xyyz = 0^k 1^p$, $k > p$ will not work.
- $xyyyz$ will not work either.
- $xz = 0^{p+1-l} 1^p$ (y has l 0s), $1 \leq l \leq p$. So $0^k 1^p$, $k \leq p$. And thus, $xz \notin L$.

Previous Lecture

[Lecture07]

Next Lecture

[Lecture09]