

CSC B36 Additional Notes
proving languages **not** regular using Pumping Lemma

© Nick Cheng

★ **Introduction**

The Pumping Lemma is used for proving that a language is **not** regular. Here is the Pumping Lemma.

If L is a regular language, then there is an integer $n > 0$ with the property that:

(*) for any string $x \in L$ where $|x| \geq n$, there are strings u, v, w such that

- (i) $x = uvw$,
- (ii) $v \neq \epsilon$,
- (iii) $|uv| \leq n$,
- (iv) $uv^k w \in L$ for all $k \in \mathbb{N}$.

To prove that a language L is **not** regular, we use proof by contradiction. Here are the steps.

1. Suppose that L is regular.
2. Since L is regular, we apply the Pumping Lemma and assert the existence of a number $n > 0$ that satisfies the property (*).
3. Give a particular string x such that
 - (a) $x \in L$,
 - (b) $|x| \geq n$.

This the trickiest part. A wrong choice here will make step 4 impossible.

4. By Pumping Lemma, there are strings u, v, w such that (i)-(iv) hold. Pick a particular number $k \in \mathbb{N}$ and argue that $uv^k w \notin L$, thus yielding our desired contradiction.

What follows are two example proofs using Pumping Lemma.

★ A (relatively) easy example

Let $L = \{0^k 1^k : k \in \mathbb{N}\}$. We prove that L is not regular.

[step 1]

By way of contradiction, suppose L is regular.

[step 2]

Let n be as in the Pumping Lemma.

[step 3]

Let $x = 0^n 1^n$.

Then $x \in L$ [definition of L]

and $|x| = 2n \geq n$.

[step 4]

By Pumping Lemma, there are strings u, v, w such that

- (i) $x = uvw$,
- (ii) $v \neq \epsilon$,
- (iii) $|uv| \leq n$,
- (iv) $uv^k w \in L$ for all $k \in \mathbb{N}$.

Let y be the prefix of x with length n . I.e., y is the first n symbols of x .

By our choice of x , $y = 0^n$.

By (i) and (iii), $uv = 0^j$ for some $j \in \mathbb{N}$ with $0 \leq j \leq n$.

Combining with (ii), $v = 0^j$ for some $j \in \mathbb{N}$ with $0 < j \leq n$.

By (iv), $uv^2 w \in L$. (#)

Aside: We are picking $k = 2$. Indeed, any $k \neq 1$ will do here.

However, $uv^2 w = uvvw$

$$= 0^{n+j} 1^n$$

$\notin L$, [definition of L ; since $j > 0$, $n + j \neq n$]

which contradicts (#).

Therefore L is not regular. $\square$

★ A harder example

Let $L = \{(10)^p 1^q : p, q \in \mathbb{N}, p \geq q\}$. We prove that L is not regular.

[step 1]

By way of contradiction, suppose L is regular.

[step 2]

Let n be as in the Pumping Lemma.

[step 3]

Let $x = (10)^n 1^n$.

Then $x \in L$ [definition of L]

and $|x| = 3n \geq n$.

[step 4]

By Pumping Lemma, there are strings u, v, w such that

- (i) $x = uvw$,
- (ii) $v \neq \epsilon$,
- (iii) $|uv| \leq n$,
- (iv) $uv^k w \in L$ for all $k \in \mathbb{N}$.

Let y be the prefix of x with length n .

By our choice of x , $y = (10)^{\frac{n}{2}}$ if n is even, and $y = (10)^{\frac{n-1}{2}} 1$ if n is odd.

By (i) and (iii), uv is a prefix of y , and

- $uv = (10)^j$ for some $j \in \mathbb{N}$ with $0 \leq j \leq \frac{n}{2}$, or
- $uv = (10)^j 1$ for some $j \in \mathbb{N}$ with $0 \leq j < \frac{n}{2}$.

Combining with (ii) — depending on whether $|uv|$ is even or odd,

- v is some nonempty substring of $(10)^j$ for some j where $0 \leq j \leq \frac{n}{2}$, or
- v is some nonempty substring of $(10)^j 1$ for some j where $0 \leq j < \frac{n}{2}$.

There are 3 cases to consider:

- (a) v starts with 0 and ends with 0.
- (b) v starts with 1 and ends with 1.
- (c) v starts and ends with different symbols.

For case (a), $uv^0w = uw$ contains 110 as a substring.

Thus $uv^0w \notin L$, [110 is not a substring of any string in L] which contradicts (iv).

For case (b), there are two possibilities: uv^0w contains 00 as a substring or uv^0w begins with 0.

Either one contradicts (iv). [details left to reader]

For case (c), $v = (10)^i$ or $v = (01)^i$, where $0 < i$.

So $|v| = 2i$.

Thus $uv^0w = uw = (10)^{n-i} 1^n \notin L$, [definition of L ; $n - i < n$]

which contradicts (iv).

We reach a contradiction in all cases.

Therefore L is not regular. $\square$