# 3.14: The Pumping Lemma for Regular Languages

In this section we consider techniques for showing that particular languages are not regular.

## Introduction

 ${\sf Question} \colon {\sf Is}$ 

```
L = \{ 0^{n}1^{n} \mid n \in \mathbb{N} \} = \{\%, 01, 0011, 000111, \ldots \}
```

a regular language? Answer:

## Introduction

Question: Is

 $L = \{ 0^{n}1^{n} \mid n \in \mathbb{N} \} = \{\%, 01, 0011, 000111, \ldots \}$ 

a regular language?

Answer: No.

Intuitively, an automaton would have to have infinitely many states to accept this language. A finite automaton won't be able to keep track of how many 0's it has seen so far, and thus won't be able to insist that the correct number of 1's follow.

## Introduction

We could turn the preceding ideas into a direct proof that L is not regular.

Instead, we will first state a general result, called the Pumping Lemma for regular languages, for proving that languages are non-regular.

Next, we will show how the Pumping Lemma can be used to prove that L is non-regular.

Finally, we will prove the Pumping Lemma.

**Lemma 3.14.1 (Pumping Lemma for Regular Languages)** For all regular languages L,

**Lemma 3.14.1 (Pumping Lemma for Regular Languages)** For all regular languages *L*, there is a  $n \in \mathbb{N} - \{0\}$  such that,

**Lemma 3.14.1 (Pumping Lemma for Regular Languages)** For all regular languages *L*, there is a  $n \in \mathbb{N} - \{0\}$  such that, for all  $z \in \mathbf{Str}$ , if  $z \in L$  and  $|z| \ge n$ , then

**Lemma 3.14.1 (Pumping Lemma for Regular Languages)** For all regular languages *L*, there is a  $n \in \mathbb{N} - \{0\}$  such that, for all  $z \in$ **Str**, if  $z \in L$  and  $|z| \ge n$ , then there are  $u, v, w \in$ **Str** such that z = uvw and

**Lemma 3.14.1 (Pumping Lemma for Regular Languages)** For all regular languages L, there is a  $n \in \mathbb{N} - \{0\}$  such that, for all  $z \in \mathbf{Str}$ , if  $z \in L$  and  $|z| \ge n$ , then there are  $u, v, w \in \mathbf{Str}$  such that z = uvw and

(1)  $|uv| \leq n;$ 

**Lemma 3.14.1 (Pumping Lemma for Regular Languages)** For all regular languages L, there is a  $n \in \mathbb{N} - \{0\}$  such that, for all  $z \in \mathbf{Str}$ , if  $z \in L$  and  $|z| \ge n$ , then there are  $u, v, w \in \mathbf{Str}$  such that z = uvw and

(1)  $|uv| \le n$ ; (2)  $v \ne \%$ ; and

**Lemma 3.14.1 (Pumping Lemma for Regular Languages)** For all regular languages L, there is a  $n \in \mathbb{N} - \{0\}$  such that, for all  $z \in \mathbf{Str}$ , if  $z \in L$  and  $|z| \ge n$ , then there are  $u, v, w \in \mathbf{Str}$  such that z = uvw and

- (1)  $|uv| \le n;$ (2)  $v \ne \%;$  and
- (3)  $uv^i w \in L$ , for all  $i \in \mathbb{N}$ .

Before proving the Pumping Lemma, let's see how it can be used to prove that  $L = \{ 0^n 1^n \mid n \in \mathbb{N} \}$  is non-regular.

Before proving the Pumping Lemma, let's see how it can be used to prove that  $L = \{ 0^n 1^n \mid n \in \mathbb{N} \}$  is non-regular.

Suppose, toward a contradiction, that *L* is regular. Thus there is an  $n \in \mathbb{N} - \{0\}$  with the property of the Pumping Lemma.

Before proving the Pumping Lemma, let's see how it can be used to prove that  $L = \{ 0^n 1^n \mid n \in \mathbb{N} \}$  is non-regular.

Suppose, toward a contradiction, that *L* is regular. Thus there is an  $n \in \mathbb{N} - \{0\}$  with the property of the Pumping Lemma.

Suppose  $z = 0^{n}1^{n}$ . Since  $z \in L$  and  $|z| = 2n \ge n$ , it follows that there are  $u, v, w \in$ **Str** such that z = uvw and properties (1)–(3) of the lemma hold. Since  $0^{n}1^{n} = z = uvw$ , (1) tells us that

Before proving the Pumping Lemma, let's see how it can be used to prove that  $L = \{ 0^n 1^n \mid n \in \mathbb{N} \}$  is non-regular.

Suppose, toward a contradiction, that *L* is regular. Thus there is an  $n \in \mathbb{N} - \{0\}$  with the property of the Pumping Lemma.

Suppose  $z = 0^n 1^n$ . Since  $z \in L$  and  $|z| = 2n \ge n$ , it follows that there are  $u, v, w \in$ **Str** such that z = uvw and properties (1)–(3) of the lemma hold. Since  $0^n 1^n = z = uvw$ , (1) tells us that there are  $i, j, k \in \mathbb{N}$  such that

$$u = 0^i$$
,  $v = 0^j$ ,  $w = 0^k 1^n$ ,  $i + j + k = n$ .

By (2), we have that

Before proving the Pumping Lemma, let's see how it can be used to prove that  $L = \{ 0^n 1^n \mid n \in \mathbb{N} \}$  is non-regular.

Suppose, toward a contradiction, that *L* is regular. Thus there is an  $n \in \mathbb{N} - \{0\}$  with the property of the Pumping Lemma.

Suppose  $z = 0^n 1^n$ . Since  $z \in L$  and  $|z| = 2n \ge n$ , it follows that there are  $u, v, w \in$ **Str** such that z = uvw and properties (1)–(3) of the lemma hold. Since  $0^n 1^n = z = uvw$ , (1) tells us that there are  $i, j, k \in \mathbb{N}$  such that

$$u = 0^{i}, v = 0^{j}, w = 0^{k}1^{n}, i + j + k = n.$$

By (2), we have that  $j \ge 1$ , and thus that i + k = n - j < n. By (3), we have that

Before proving the Pumping Lemma, let's see how it can be used to prove that  $L = \{ 0^n 1^n \mid n \in \mathbb{N} \}$  is non-regular.

Suppose, toward a contradiction, that *L* is regular. Thus there is an  $n \in \mathbb{N} - \{0\}$  with the property of the Pumping Lemma.

Suppose  $z = 0^n 1^n$ . Since  $z \in L$  and  $|z| = 2n \ge n$ , it follows that there are  $u, v, w \in$ **Str** such that z = uvw and properties (1)–(3) of the lemma hold. Since  $0^n 1^n = z = uvw$ , (1) tells us that there are  $i, j, k \in \mathbb{N}$  such that

$$u = 0^{i}, v = 0^{j}, w = 0^{k}1^{n}, i + j + k = n.$$

By (2), we have that  $j \ge 1$ , and thus that i + k = n - j < n. By (3), we have that

$$0^{i+k}1^n = 0^i 0^k 1^n = uw = u\%w = uv^0 w \in L.$$

Thus i + k = n—contradiction. Thus L is not regular.

**Proof.** Suppose *L* is a regular language.

**Proof.** Suppose *L* is a regular language. Thus there is an NFA *M* such that L(M) = L. Let  $n = |Q_M|$ . Thus  $n \in \mathbb{N} - \{0\}$ .

**Proof.** Suppose *L* is a regular language. Thus there is an NFA *M* such that L(M) = L. Let  $n = |Q_M|$ . Thus  $n \in \mathbb{N} - \{0\}$ . Suppose  $z \in$ **Str**,  $z \in L$  and  $|z| \ge n$ . Let m = |z|. Thus  $1 \le n \le |z| = m$ . Since  $z \in L = L(M)$  and *M* is an NFA,

**Proof.** Suppose *L* is a regular language. Thus there is an NFA *M* such that L(M) = L. Let  $n = |Q_M|$ . Thus  $n \in \mathbb{N} - \{0\}$ . Suppose  $z \in$ **Str**,  $z \in L$  and  $|z| \ge n$ . Let m = |z|. Thus  $1 \le n \le |z| = m$ . Since  $z \in L = L(M)$  and *M* is an NFA, there is a valid labeled path for *M* 

$$q_1 \stackrel{a_1}{\Rightarrow} q_2 \stackrel{a_2}{\Rightarrow} \cdots q_m \stackrel{a_m}{\Rightarrow} q_{m+1},$$

that is labeled by z and where  $q_1 = s_M$ ,  $q_{m+1} \in A_M$  and  $a_i \in Sym$  for all  $1 \le i \le m$ .

**Proof.** Suppose *L* is a regular language. Thus there is an NFA *M* such that L(M) = L. Let  $n = |Q_M|$ . Thus  $n \in \mathbb{N} - \{0\}$ . Suppose  $z \in$ **Str**,  $z \in L$  and  $|z| \ge n$ . Let m = |z|. Thus  $1 \le n \le |z| = m$ . Since  $z \in L = L(M)$  and *M* is an NFA, there is a valid labeled path for *M* 

$$q_1 \stackrel{a_1}{\Rightarrow} q_2 \stackrel{a_2}{\Rightarrow} \cdots q_m \stackrel{a_m}{\Rightarrow} q_{m+1},$$

that is labeled by z and where  $q_1 = s_M$ ,  $q_{m+1} \in A_M$  and  $a_i \in Sym$ for all  $1 \le i \le m$ . Since  $|Q_M| = n$ , not all of the states  $q_1, \ldots, q_{n+1}$  are distinct. Thus, there are  $1 \le i < j \le n+1$  such that  $q_i = q_j$ .

**Proof (cont.).** Hence, our path looks like:

 $q_1 \stackrel{a_1}{\Rightarrow} \cdots q_{i-1} \stackrel{a_{i-1}}{\Rightarrow} q_i \stackrel{a_i}{\Rightarrow} \cdots q_{j-1} \stackrel{a_{j-1}}{\Rightarrow} q_j \stackrel{a_j}{\Rightarrow} \cdots q_m \stackrel{a_m}{\Rightarrow} q_{m+1}.$ 

**Proof (cont.).** Hence, our path looks like:

$$q_1 \stackrel{a_1}{\Rightarrow} \cdots q_{i-1} \stackrel{a_{i-1}}{\Rightarrow} q_i \stackrel{a_i}{\Rightarrow} \cdots q_{j-1} \stackrel{a_{j-1}}{\Rightarrow} q_j \stackrel{a_j}{\Rightarrow} \cdots q_m \stackrel{a_m}{\Rightarrow} q_{m+1}.$$
Let

$$u = a_1 \cdots a_{i-1}, \quad v = a_i \cdots a_{j-1}, \quad w = a_j \cdots a_m.$$

Then z = uvw. Since |uv| = j - 1 and  $j \le n + 1$ , we have that  $|uv| \le n$ . Since i < j, we have that  $i \le j - 1$ , and thus that  $v \ne \%$ .

Proof (cont.). Finally, since

 $q_i \in \Delta(\{q_1\}, u), \quad q_j \in \Delta(\{q_i\}, v), \quad q_{m+1} \in \Delta(\{q_j\}, w)$ 

and  $q_i = q_j$ , we have that

 $q_j \in \Delta(\{q_1\}, u), \quad q_j \in \Delta(\{q_j\}, v), \quad q_{m+1} \in \Delta(\{q_j\}, w).$ 

Thus, we have that  $q_{m+1} \in \Delta(\{q_1\}, uv^i w)$  for all  $i \in \mathbb{N}$ . But  $q_1 = s_M$  and  $q_{m+1} \in A_M$ , and thus  $uv^i w \in L(M) = L$  for all  $i \in \mathbb{N}$ .  $\Box$ 

We could show that L' is non-regular using the Pumping Lemma.

But we can also prove this result by using some of the closure properties of Section 3.12 plus the fact that  $L = \{ 0^n 1^n \mid n \in \mathbb{N} \}$  is non-regular.

We could show that L' is non-regular using the Pumping Lemma.

But we can also prove this result by using some of the closure properties of Section 3.12 plus the fact that  $L = \{ 0^n 1^n \mid n \in \mathbb{N} \}$  is non-regular.

Suppose, toward a contradiction, that L' is regular. It is easy to see that  $\{0\}$  and  $\{1\}$  are regular (e.g., they are generated by the regular expressions 0 and 1). Thus, by Theorem 3.12.29, we have that  $\{0\}^*\{1\}^*$  is regular. Hence, by Theorem 3.12.29 again, it follows that —contradiction. Thus L'is non-regular.

We could show that L' is non-regular using the Pumping Lemma.

But we can also prove this result by using some of the closure properties of Section 3.12 plus the fact that  $L = \{ 0^n 1^n \mid n \in \mathbb{N} \}$  is non-regular.

Suppose, toward a contradiction, that L' is regular. It is easy to see that  $\{0\}$  and  $\{1\}$  are regular (e.g., they are generated by the regular expressions 0 and 1). Thus, by Theorem 3.12.29, we have that  $\{0\}^*\{1\}^*$  is regular. Hence, by Theorem 3.12.29 again, it follows that  $= L' \cap \{0\}^*\{1\}^*$  is regular—contradiction. Thus L' is non-regular.

We could show that L' is non-regular using the Pumping Lemma.

But we can also prove this result by using some of the closure properties of Section 3.12 plus the fact that  $L = \{ 0^n 1^n \mid n \in \mathbb{N} \}$  is non-regular.

Suppose, toward a contradiction, that L' is regular. It is easy to see that  $\{0\}$  and  $\{1\}$  are regular (e.g., they are generated by the regular expressions 0 and 1). Thus, by Theorem 3.12.29, we have that  $\{0\}^*\{1\}^*$  is regular. Hence, by Theorem 3.12.29 again, it follows that  $L = L' \cap \{0\}^*\{1\}^*$  is regular—contradiction. Thus L' is non-regular.

# Forlan Implementation of Pumping Lemma

The textbook describes the implementation in Forlan of the idea behind the Pumping Lemma.