# Pumping Lemma for Regular Languages

**Finite memory • repeated states • nonregularity**

How to prove that a language cannot be recognized by any finite automaton.

# Learning outcomes

### Understand

Explain why sufficiently long accepted strings repeat a DFA state.

### State

Write the pumping lemma with the correct quantifiers and conditions.

### Apply

Prove nonregularity using a carefully selected witness string.

### Evaluate

Identify invalid proofs and the limitations of the lemma.

# Why must a DFA repeat a state?

### Finite memory

A DFA with p states can remember only one of p possible current states.

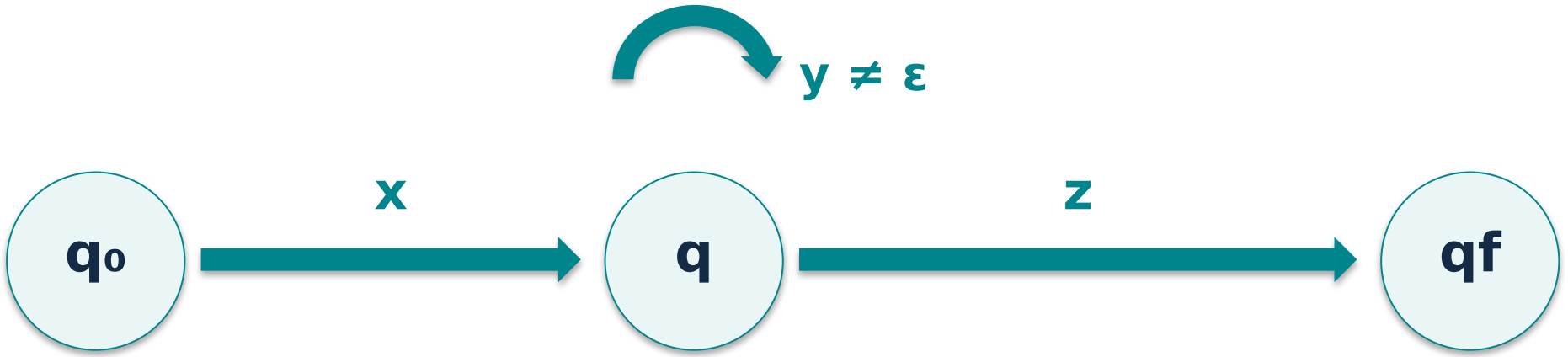
### Pigeonhole principle

Reading p symbols visits p + 1 states, including the initial state.

### Repeated state

At least two of those visits are to the same state; the intervening input forms a loop.

# The loop behind the lemma




x reaches the repeated state; y returns to it; z leads to an accepting state. Repeating or skipping y preserves acceptance.

# Formal statement of the pumping lemma

### Assumption

If L is regular, there exists an integer $p \geq 1$ such that every $w \in L$ with $|w| \geq p$ can be written $w = xyz$.

### Conditions

$|xy| \leq p$, $|y| \geq 1$, and $xy^iz \in L$ for every integer $i \geq 0$.

### Meaning

The nonempty substring y can be repeated any number of times, including zero.

# Quantifiers: who chooses what?

### 1. Pumping length

Assume L is regular. The lemma supplies p; it is not chosen by the proof writer.

### 2. Witness

Choose $w \in L$ with $|w| \geq p$, typically in terms of p.

### 3. Decomposition

Consider every valid split xyz; then choose an $i \geq 0$ that forces $xy^iz$ outside L.

**To refute regularity, defeat every valid split of your chosen witness.**

# Proof of the lemma

### Let M be a DFA

Take $p = |Q|$. For an accepted word w of length at least p, consider the first p transitions.

### Find a repeated state

Among visits $q_0,\ldots,q_p$, there exist indices $0 \le j < k \le p$ with $q_j = q_k$.

### Split the input

x is the first j symbols, y the next k−j, and z the remainder. The loop y can be traversed i times for every $i \ge 0$.

# Proof strategy for nonregularity

### Assume regularity

Let p be the pumping length supplied by the lemma.

### Choose a witness

Select $w \in L$ with $|w| \geq p$; use $|xy| \leq p$ to constrain y.

### Pump and contradict

For every valid decomposition, show some i gives $xy^iz \notin L$. Conclude L is not regular.

# Example 1: equal numbers in blocks

## Language

$L = \{0^n 1^n : n \geq 0\}$.

## Assume and choose

Suppose L is regular with pumping length p. Choose $w = 0^p 1^p$; $w \in L$ and $|w| = 2p \geq p$.

## Constrain the split

Since $|xy| \leq p$ and $|y| \geq 1$, y contains only zeros: $y = 0^k$ with $1 \leq k \leq p$.

# Example 1: pump down

### Choose i = 0

$xy^0z = xz = 0^{p-k}1^p$.

### Compare counts

There are $p-k$ zeros and $p$ ones. Since $k \geq 1$, the counts differ.

### Contradiction

$xz \notin L$, contrary to the pumping lemma. Therefore $L$ is not regular.

**The same argument works for every permitted decomposition.**

# Example 2: equal counts in any order

### Language

$L = \{w \in \{a,b\}^* : \#a(w) = \#b(w)\}$.

### Witness

Assume regularity, let p be the pumping length, and choose $w = a^p b^p$.

### Pump up

Every valid y is $a^k$, $k \geq 1$. With $i = 2$, $xy^2z = a^{p+k}b^p$, which has more a symbols than b symbols.

**Contradiction: L is not regular.**

# Example 3: unary square lengths

### Language

$L = \{a^{n^2} : n \geq 0\}$. Assume regularity with pumping length $p$.

### Witness and split

Choose $w = a\hat{}(p^2)$. Every valid $y = a^k$ with $1 \leq k \leq p$.

### Pump up

For $i = 2$, length $= p^2 + k$. Since $p^2 < p^2 + k < (p+1)^2$, it is not a square.

**Thus $xy^2z \notin L$ and L is not regular.**

# A regular language that pumps

**Language**

L = (ab)*. A pumping length p = 2 suffices.

**Any sufficiently long word**

For $w = (ab)^n$, $n \geq 1$, choose $x = \varepsilon$, $y = ab$, $z = (ab)^{n-1}$.

**Check**

$|xy| = 2$, $|y| = 2$, and $xy^iz = (ab)^{i+n-1} \in L$ for every $i \geq 0$.

# Common mistakes and limitations

### Wrong quantifier

Choosing one convenient decomposition is not enough for a nonregularity proof.

### Wrong witness

The witness must belong to L and have length at least p.

### Wrong conclusion

The pumping property is necessary, not sufficient, for regularity. Use a DFA or regular expression to prove regularity.

# Classroom exercises

### Exercise 1

Prove that $L = \{a^n b^{2n} : n \geq 0\}$ is not regular.

### Exercise 2

Prove that $L = \{ww^R : w \in \{0,1\}^*\}$ is not regular (even-length palindromes).

### Exercise 3

Give a pumping length and valid decomposition for every sufficiently long word in $0^*1^*$.

# Solutions and recap

### Exercise 1

$w = a^p b^{2p}$; $y = a^k$. Pump $i = 0$: $a^{p-k} b^{2p}$ violates the 1:2 ratio.

### Exercise 2

$w = 0^p 110^p$; $y = 0^k$. Pump $i = 0$: unequal outer zero blocks, so not a palindrome.

### Exercise 3

$p = 1$. Choose $x = \varepsilon$, $y$ = first symbol, $z$ = remaining suffix. Repeating or deleting $y$ preserves $0^*1^*$.

**Remember: choose the witness, constrain every split, pump, contradict.**