The **Limits** of Regular Languages
Friday Four Square!
4:15PM, Outside Gates
Announcements

- Problem Set 4 due right now.
- Problem Set 5 out, due next Friday at 2:15PM.
  - Play around with regular languages!
  - Explore nonregular languages.
- Midterm next Tuesday, 7PM – 10PM.
  - Alternate times will be sent out later tonight.
  - If you have contacted us prior to 3:30PM today about taking an alternate exam, you can show up to any of the alternate exam times.
The Limits of Regular Languages

- We have now seen three ways to think about regular languages:
  - DFAs
  - NFAs
  - Regular Expressions
- We have a huge array of closure properties.
- Is every language regular?
An Important Observation
An Important Observation
An Important Observation
An Important Observation
An Important Observation

Start: q₀

- q₀ to q₁: 0
- q₁ to q₂: 1
- q₁ to q₅: 0
- q₅ to q₂: 0
- q₂ to q₃: 1
- q₃ to q₄: 1
- q₄ to q₅: 0
- q₅ to q₅: 1
An Important Observation
An Important Observation
An Important Observation
An Important Observation
An Important Observation

The diagram illustrates a finite state machine with states $q_0$, $q_1$, $q_2$, $q_3$, $q_5$, and $q_4$. The transitions are labeled with inputs 0 and 1:

- From $q_0$, a 0 transition leads to $q_1$, a 1 transition leads to $q_2$, and a 0 transition loops back to $q_0$.
- From $q_1$, a 0 transition leads to $q_2$, a 1 transition leads to $q_3$, and a 1 transition loops back to $q_1$.
- From $q_2$, a 1 transition leads to $q_3$, and a transition labeled with 0, 1 leads to $q_5$.
- From $q_3$, a 1 transition leads to $q_4$, and a transition labeled with 0, 1 leads to $q_5$.
- From $q_5$, a transition labeled with 0, 1 leads to $q_2$.
- From $q_4$, a 1 transition loops back to $q_4$.

The states $q_2$, $q_3$, and $q_5$ are part of a loop, indicating that they can be revisited without consuming another symbol.
An Important Observation

Diagram:

- Start state: $q_0$
- States: $q_0, q_1, q_2, q_3, q_5, q_4$
- Transitions:
  - $q_0 \xrightarrow{0} q_1$
  - $q_1 \xrightarrow{1} q_2$
  - $q_2 \xrightarrow{0} q_5$
  - $q_2 \xrightarrow{1} q_3$
  - $q_3 \xrightarrow{1} q_4$
  - $q_5 \xrightarrow{0,1} q_5$

Input sequence: $0, 1, 1, 1, 0, 1, 1, 1, 1$
An Important Observation
An Important Observation
An Important Observation
An Important Observation

Diagram:

- Start state: $q_0$
- States: $q_0, q_1, q_2, q_3, q_4, q_5$
- Transitions:
  - $q_0$ to $q_1$ on 0
  - $q_0$ to $q_5$ on 0, 1
  - $q_1$ to $q_2$ on 1
  - $q_2$ to $q_3$ on 1
  - $q_3$ to $q_4$ on 1
  - $q_4$ to $q_2$ on 1
  - $q_5$ to $q_0$ on 0, 1

Input sequence: 0 1 1 1 0 1 1 1 1
An Important Observation
An Important Observation
An Important Observation
An Important Observation

![Diagram of a finite automaton with states and transitions.]

- **States**: $q_0, q_1, q_2, q_3, q_4, q_5$
- **Transitions**:
  - $q_0 \xrightarrow{0} q_1$
  - $q_1 \xrightarrow{0} q_2$
  - $q_2 \xrightarrow{1} q_3$
  - $q_3 \xrightarrow{1} q_4$
  - $q_5 \xrightarrow{0, 1} q_0$
  - $q_5 \xrightarrow{0} q_1$
  - $q_5 \xrightarrow{0, 1} q_2$
  - $q_5 \xrightarrow{0, 1} q_3$
  - $q_5 \xrightarrow{0, 1} q_4$

- **Start State**: $q_0$
- **Accepting States**: $q_4$
An Important Observation
An Important Observation

The image displays a nondeterministic finite automaton (NFA). The states are labeled with circles, and the transitions are indicated by arrows. The alphabet is labeled with 0 and 1. The transitions are as follows:

- Start state: $q_0$, transitions: $0 \rightarrow q_1, 1 \rightarrow q_2$.
- State $q_1$: transitions: $0 \rightarrow q_1, 1 \rightarrow q_3$.
- State $q_2$: transitions: $0 \rightarrow q_5, 1 \rightarrow q_3$.
- State $q_3$: transition: $1 \rightarrow q_4$.
- State $q_5$: loop transition: $0, 1 \rightarrow q_5$.

The accept state is $q_4$. The states $q_1$ and $q_3$ are labeled with yellow and gray, indicating specific transitions or states of interest.
An Important Observation
An Important Observation
An Important Observation
Visiting Multiple States

- Let $D$ be a DFA with $n$ states.
- Any string $w$ accepted by $D$ that has length at least $n$ must visit some state twice.
  - Number of states visited is equal to the length of the string plus one.
  - By the pigeonhole principle, some state is duplicated.
- The substring of $w$ between those revisited states can be removed, duplicated, tripled, etc. without changing the fact that $D$ accepts $w$. 

Informally

- Let $L$ be a regular language.
- If we have a string $w \in L$ that is “sufficiently long,” then we can split the string into three pieces and “pump” the middle.
- Write $w = xyz$.
- Then $xy^0z, xy^1z, xy^2z, \ldots, xy^n z, \ldots$ are all in $L$.
  - **Notation:** $y^n$ means “$n$ copies of $y$.”
The Weak Pumping Lemma

- The **Weak Pumping Lemma for Regular Languages** states that

  **For any** regular language L,

  There exists a positive natural number n such that

  **For any** w ∈ L with |w| ≥ n,

  There exists strings x, y, z such that

  **For any** natural number i,

  \[ w = xyz, \]

  \[ y \neq \varepsilon \]

  \[ xy^iz \in L \]
The Weak Pumping Lemma

The Weak Pumping Lemma for Regular Languages states that

∀ regular language L,

∃ a positive natural number n such that

∀ w ∈ L with |w| ≥ n,

∃ strings x, y, z such that

∀ natural number i,

w = xyz,

y ≠ ε

xy^i z ∈ L
The Weak Pumping Lemma

- The **Weak Pumping Lemma** for Regular Languages states that

  \[ \forall \text{ regular language } L, \exists \text{ a positive natural number } n \text{ such that } \forall w \in L \text{ with } |w| \geq n, \exists \text{ strings } x, y, z \text{ such that } \forall \text{ natural number } i, w = xyz, y \neq \varepsilon, xy^iz \in L \]
The Weak Pumping Lemma

The Weak Pumping Lemma for Regular Languages states that

For any regular language $L$,

There exists a positive natural number $n$ such that

For any $w \in L$ with $|w| \geq n$,

There exists strings $x, y, z$ such that

For any natural number $i$,

$$w = xyz,$$

$$y \neq \varepsilon$$

$$xy^iz \in L$$
The Weak Pumping Lemma

The Weak Pumping Lemma for Regular Languages states that there exists a positive natural number $n$ such that for any regular language $L$, and for any word $w \in L$ with $|w| \geq n$, there exist strings $x$, $y$, and $z$ such that:

For any natural number $i$,

$$w = xyz,$$

$y \neq \varepsilon$

$$xy^iz \in L$$

This number $n$ is sometimes called the pumping length.
The Weak Pumping Lemma

- The **Weak Pumping Lemma for Regular Languages** states that

  **For any** regular language $L$,

  There exists a positive natural number $n$ such that

  For any $w \in L$ with $|w| \geq n$,

  There exists strings $x, y, z$ such that

  For any natural number $i$,

  $w = xyz$,

  $y \neq \varepsilon$

  $xy^iz \in L$

Strings longer than the pumping length must have a special property.
The Weak Pumping Lemma

- The **Weak Pumping Lemma for Regular Languages** states that

  For any **regular language** \( L \),
  
  There exists a positive natural number \( n \) such that
  
  For any \( w \in L \) with \( |w| \geq n \),
  
  There exists strings \( x, y, z \) such that

  For any natural number \( i \),

  
  \[ w = xyz, \quad w \text{ can be broken into three pieces,} \]

  \[ y \neq \varepsilon \]

  \[ xy^iz \in L \]
The Weak Pumping Lemma

- The Weak Pumping Lemma for Regular Languages states that

For any regular language L,

There exists a positive natural number n such that

For any $w \in L$ with $|w| \geq n$,

There exists strings $x, y, z$ such that

For any natural number $i$, $w = xyz$, $w$ can be broken into three pieces,

$y \neq \varepsilon$ where the middle piece isn't empty,

$xy^iz \in L$
The Weak Pumping Lemma

- The **Weak Pumping Lemma for Regular Languages** states that

  For any regular language $L$, 
  
  There exists a positive natural number $n$ such that
  
  For any $w \in L$ with $|w| \geq n$, 
  
  There exists strings $x$, $y$, $z$ such that
  
  For any natural number $i$, 
  
  $w = xyz$, $w$ can be broken into three pieces,
  
  $y \neq \varepsilon$, where the middle piece isn't empty,
  
  $xy^iz \in L$, where the middle piece can be replicated zero or more times.
The Weak Pumping Lemma

- Let $\Sigma = \{0, 1\}$ and $L = \{ w \mid w$ contains 00 as a substring. $\}$
- Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”

```
1 0 0 1 0
```
The Weak Pumping Lemma

- Let $\Sigma = \{0, 1\}$ and $L = \{ w \mid w \text{ contains 00 as a substring.} \}$
- Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”
The Weak Pumping Lemma

• Let $\Sigma = \{0, 1\}$ and
  $L = \{ w \mid w \text{ contains } 00 \text{ as a substring.} \}$

• Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”

\[1 \ 0 \ 0 \ 0\]
The Weak Pumping Lemma

- Let \( \Sigma = \{0, 1\} \) and
  \( L = \{ w \mid w \text{ contains } 00 \text{ as a substring. } \} \)

- Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”
The Weak Pumping Lemma

- Let $\Sigma = \{0, 1\}$ and $L = \{ w \mid w$ contains 00 as a substring. $\}$
- Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”

```
1 0 0 1 1 0
```
The Weak Pumping Lemma

• Let $\Sigma = \{0, 1\}$ and $L = \{ w \mid w$ contains 00 as a substring. $\}$

• Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”
The Weak Pumping Lemma

• Let $\Sigma = \{0, 1\}$ and
  \[ L = \{ w \mid w \text{ contains } 00 \text{ as a substring. } \} \]

• Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”
The Weak Pumping Lemma

- Let \( \Sigma = \{0, 1\} \) and 
  \( L = \{ w \mid w \text{ contains 00 as a substring.} \} \)

- Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”

1 0 0
The Weak Pumping Lemma

- Let $\Sigma = \{0, 1\}$ and $L = \{ w \mid w \text{ contains } 00 \text{ as a substring.}\}$
- Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”

The first piece is just the empty string! This is perfectly fine.
The Weak Pumping Lemma

• Let $\Sigma = \{0, 1\}$ and
  $L = \{ w \mid w \text{ contains } 00 \text{ as a substring.} \}$

• Any string of length 3 or greater can be split into three pieces, the second of which can be "pumped."
The Weak Pumping Lemma

• Let $\Sigma = \{0, 1\}$ and
  $L = \{ w \mid w$ contains $00$ as a substring. $\}$

• Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”
The Weak Pumping Lemma

- Let $\Sigma = \{0, 1\}$ and $L = \{ w \mid w$ contains 00 as a substring. $\}$
- Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”
The Weak Pumping Lemma

- Let $\Sigma = \{0, 1\}$ and $L = \{ w \mid w$ contains 00 as a substring. $\}$
- Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”
The Weak Pumping Lemma

- Let $\Sigma = \{0, 1\}$ and $L = \{w | w$ contains 00 as a substring.$\}$
- Any string of length 3 or greater can be split into three pieces, the second of which can be "pumped."

```
1 1 1 0 0
```
The Weak Pumping Lemma

• Let $\Sigma = \{0, 1\}$ and $L = \{ w \mid w$ contains 00 as a substring. $\}$

• Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”
The Weak Pumping Lemma

- Let $\Sigma = \{0, 1\}$ and $L = \{ w \mid w \text{ contains 00 as a substring.} \}$
- Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”

| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
The Weak Pumping Lemma

• Let $\Sigma = \{0, 1\}$ and
  $L = \{ w \mid w \text{ contains } 00 \text{ as a substring.} \}$

• Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”

1 1 0 0 0 0 0 1
The Weak Pumping Lemma

• Let \( \Sigma = \{0, 1\} \) and
  \[ L = \{ \text{w} \mid \text{w contains 00 as a substring.} \} \]

• Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”

\[
\begin{array}{cccc}
1 & 1 & 0 & 0
\end{array}
\]
The Weak Pumping Lemma

- Let $\Sigma = \{0, 1\}$ and $L = \{ w \mid w$ contains 00 as a substring. $\}$
- Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”

1 1 0 0 0 0 0 1
The Weak Pumping Lemma

• Let $\Sigma = \{0, 1\}$ and $L = \{ w \mid w \text{ contains } 00 \text{ as a substring.} \}$

• Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”

0 0 1
0 0 1
0 0 1
1 0 0 1
The Weak Pumping Lemma

• Let $\Sigma = \{0, 1\}$ and $L = \{ w \mid w$ contains 00 as a substring. $\}$

• Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”

1 1 0 0 0 0 1 0 0 1 0 0 0 1
The Weak Pumping Lemma
The Weak Pumping Lemma

• Let $\Sigma = \{0, 1\}$ and
  $L = \{ \varepsilon, 0, 1, 00, 01, 10, 11 \}$
The Weak Pumping Lemma

- Let $\Sigma = \{0, 1\}$ and
  $L = \{ \varepsilon, 0, 1, 00, 01, 10, 11 \}$
- Any string of length 3 or greater can be split into three pieces, the second of which can be "pumped."
The Weak Pumping Lemma

- Let $\Sigma = \{0, 1\}$ and $L = \{\varepsilon, 0, 1, 00, 01, 10, 11\}$

- Any string of length 3 or greater can be split into three pieces, the second of which can be “pumped.”

The weak pumping lemma holds for finite languages because the pumping length can be longer than the longest string!
Testing Equality

- The **equality problem** is defined as follows:
  
  Given two strings $x$ and $y$, report whether $x = y$.

- Let $\Sigma = \{0, 1, ?\}$. We can encode the equality problem as a string of the form $x?y$.
  
  - “Is 001 equal to 110?” would be 001?110
  - “Is 11 equal to 11?” would be 11?11
  - “Is 110 equal to 110?” would be 110?110

- Let $EQUAL = \{ w?w \mid w \in \{0, 1\}^* \}$

- **Question**: Is $EQUAL$ a regular language?
The Weak Pumping Lemma

The **Weak Pumping Lemma for Regular Languages** states that

For any regular language \( L \),

**There exists** a positive natural number \( n \) such that

For any \( w \in L \) with \( |w| \geq n \),

**There exists** strings \( x, y, z \) such that

For any natural number \( i \),

\[
\begin{align*}
w &= xyz, & w \text{ can be broken into three pieces,} \\
y &\neq \varepsilon & \text{where the middle piece isn’t empty,} \\
xyz^i &\in L & \text{where the middle piece can be replicated zero or more times.}
\end{align*}
\]
Using the Weak Pumping Lemma

\[ EQUAL = \{ w?w \mid w \in \{0, 1\}^* \} \]
Using the Weak Pumping Lemma

\[\text{EQUAL} = \{ w?w \mid w \in \{0, 1\}^* \}\]
Using the Weak Pumping Lemma

\[ \text{EQUAL} = \{ \ w?w \mid w \in \{0, 1\}^* \ \} \]
Using the Weak Pumping Lemma

\[ \text{EQUAL} = \{ \text{w}x\text{w} \mid \text{w} \in \{0, 1\}^* \} \]
Using the Weak Pumping Lemma

\[ EQUAL = \{ w?w \mid w \in \{0, 1\}^* \} \]
Using the Weak Pumping Lemma

\[ \text{EQUAL} = \{ w?w \mid w \in \{0, 1\}^* \} \]
Using the Weak Pumping Lemma

\[ \text{EQUAL} = \{ ww^* | w \in \{0, 1\}^* \} \]
Using the Weak Pumping Lemma

\[ \text{EQUAL} = \{ w?w \mid w \in \{0, 1\}^* \} \]
Using the Weak Pumping Lemma

\[ \text{EQUAL} = \{ w?w \mid w \in \{0, 1\}^* \} \]
Using the Weak Pumping Lemma

\[ EQUAL = \{ \, w?w \mid w \in \{0, 1\}^* \, \} \]
Using the Weak Pumping Lemma

\[ EQUAL = \{ w?w \mid w \in \{0, 1\}^* \} \]
Using the Weak Pumping Lemma

\[ \text{EQUAL} = \{ \, w^2w \mid w \in \{0, 1\}^* \, \} \]
What's Going On?

- The weak pumping lemma says that for “sufficiently long” strings, we should be able to pump some part of the string.
- We can't pump any part containing the ?, because we can't duplicate or remove it.
- We can't pump just one part of the string, because then the strings on opposite sides of the ? wouldn't match.
- Can we formally show that $EQUAL$ is not regular?
Theorem: $\text{EQUAL}$ is not regular.

Proof: By contradiction; assume that $\text{EQUAL}$ is regular. Let $n$ be the pumping length guaranteed by the weak pumping lemma. Let $w = 0^n ? 0^n$. Then $w \in \text{EQUAL}$ and $|w| = 2n + 1 \geq n$. Thus by the weak pumping lemma, we can write $w = xyz$ such that $y \neq \varepsilon$ and for any natural number $i$, $xy^i z \in \text{EQUAL}$. Then $y$ cannot contain $?$, since otherwise if we let $i = 0$, $xy^0 z = xz$ does not contain $?$ and would not be in $\text{EQUAL}$. So $y$ is either completely to the left of the $?$ or completely to the right of the $?$. Let $|y| = k$, so $k > 0$. Since $y$ is completely to the left or right of the $?$, then $y = 0^k$. Now, we consider two cases:

Case 1: $y$ is to the left of the $?$.
Then $xy^2 z = 0^n 0^n \not\in \text{EQUAL}$, contradicting the weak pumping lemma.

Case 2: $y$ is to the right of the $?$.
Then $xy^2 z = 0^n \not\in \text{EQUAL}$, contradicting the weak pumping lemma.

In either case we reach a contradiction, so our assumption was wrong and $\text{EQUAL}$ is not regular. ■

For any regular language $L$,

There exists a positive natural number $n$ such that

For any $w \in L$ with $|w| \geq n$,

There exists strings $x, y, z$ such that

For any natural number $i$,

$w = xyz$,

$y \neq \varepsilon$

$xy^i z \in L$

Theorem: $\text{EQUAL}$ is not regular.
Theorem: \textit{EQUAL} is not regular.

Proof: By contradiction; assume that \textit{EQUAL} is regular.

Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n \oplus 0^n \). Then \( w \in \text{EQUAL} \) and \( |w| = 2n + 1 \geq n \). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^i z \in \text{EQUAL} \). Then \( y \) cannot contain \( ? \), since otherwise if we let \( i = 0 \), \( xy^0 z = xz \) does not contain \( ? \) and would not be in \textit{EQUAL}. So \( y \) is either completely to the left of the \( ? \) or completely to the right of the \( ? \). Let \( |y| = k \), so \( k > 0 \). Since \( y \) is completely to the left or right of the \( ? \), then \( y = 0^k \).

Now, we consider two cases:

Case 1: \( y \) is to the left of the \( ? \). Then \( xy^2 z = 0^n \oplus 0^n \notin \text{EQUAL} \), contradicting the weak pumping lemma.

Case 2: \( y \) is to the right of the \( ? \). Then \( xy^2 z = 0^n \notin \text{EQUAL} \), contradicting the weak pumping lemma.

In either case we reach a contradiction, so our assumption was wrong and \textit{EQUAL} is not regular. \( \blacksquare \)
**Theorem:** \( \text{EQUAL} \) is not regular.

**Proof:** By contradiction; assume that \( \text{EQUAL} \) is regular.

Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n \oplus 0^n \). Then \( w \in \text{EQUAL} \) and \( |w| = 2n + 1 \geq n \). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in \text{EQUAL} \). Then \( y \) cannot contain \( \oplus \), since otherwise if we let \( i = 0 \), \( xy^0z \) does not contain \( \oplus \) and would not be in \( \text{EQUAL} \). So \( y \) is either completely to the left of the \( \oplus \) or completely to the right of the \( \oplus \). Let \( |y| = k \), so \( k > 0 \). Since \( y \) is completely to the left or right of the \( \oplus \), then \( y = 0^k \).

Now, we consider two cases:

**Case 1:** \( y \) is to the left of the \( \oplus \). Then \( xy^2z = 0^n \oplus 0^{n+k} \notin \text{EQUAL} \), contradicting the weak pumping lemma.

**Case 2:** \( y \) is to the right of the \( \oplus \). Then \( xy^2z = 0^n \oplus 0^n \notin \text{EQUAL} \), contradicting the weak pumping lemma.

In either case we reach a contradiction, so our assumption was wrong and \( \text{EQUAL} \) is not regular. ■

**For any** regular language \( L \),

**There exists** a positive natural number \( n \) such that

**For any** \( w \in L \) with \( |w| \geq n \),

**There exists** strings \( x, y, z \) such that

**For any** natural number \( i \),

\[
\begin{align*}
w &= xyz, \\
y &\neq \varepsilon \\
x y^i z &\in L
\end{align*}
\]
For any regular language $L$,
There exists a positive natural number $n$ such that
For any $w \in L$ with $|w| \geq n$,
There exists strings $x$, $y$, $z$ such that
For any natural number $i$,
\[ w = xyz, \]
\[ y \neq \varepsilon \]
\[ xy^i z \in L \]

Theorem: $EQUAL$ is not regular.

Proof: By contradiction; assume that $EQUAL$ is regular. Let $n$ be the pumping length guaranteed by the weak pumping lemma.
Theorem: \textit{EQUAL} is not regular.

Proof: By contradiction; assume that \textit{EQUAL} is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n \neq 0^n \). Then \( w \in \textit{EQUAL} \) and \( |w| = 2n + 1 \geq n \). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in \textit{EQUAL} \). Then \( y \) cannot contain \( ? \), since otherwise if we let \( i = 0 \), \( xy^0z = xz \) does not contain \( ? \) and would not be in \textit{EQUAL}. So \( y \) is either completely to the left of the \( ? \) or completely to the right of the \( ? \). Let \( |y| = k \), so \( k > 0 \). Since \( y \) is completely to the left or right of the \( ? \), then \( y = 0^k \). Now, we consider two cases:

Case 1: \( y \) is to the left of the \( ? \). Then \( xy^2z = 0^n ? 0^n \notin \textit{EQUAL} \), contradicting the weak pumping lemma.

Case 2: \( y \) is to the right of the \( ? \). Then \( xy^2z = 0^n ? 0^n+k \notin \textit{EQUAL} \), contradicting the weak pumping lemma.

In either case we reach a contradiction, so our assumption was wrong and \textit{EQUAL} is not regular. ■
Theorem: \( \text{EQUAL} \) is not regular.

Proof: By contradiction; assume that \( \text{EQUAL} \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma.

Let \( w = 0^n?0^n \). Then \( w \in \text{EQUAL} \) and \( |w| = 2n + 1 \geq n \). Thus, by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in \text{EQUAL} \).

Then \( y \) cannot contain \( ? \), since otherwise if we let \( i = 0 \), \( xy^0z \) does not contain \( ? \) and would not be in \( \text{EQUAL} \).

So \( y \) is either completely to the left of the \( ? \) or completely to the right of the \( ? \). Let \( |y| = k \), so \( k > 0 \). Since \( y \) is completely to the left or right of the \( ? \), then \( y = 0^k \). Now, we consider two cases:

Case 1: \( y \) is to the left of the \( ? \).

Then \( xy^2z = 0^n?0^n \notin \text{EQUAL} \), contradicting the weak pumping lemma.

Case 2: \( y \) is to the right of the \( ? \).

Then \( xy^2z = 0^n?0^n \notin \text{EQUAL} \), contradicting the weak pumping lemma.

In either case we reach a contradiction, so our assumption was wrong and \( \text{EQUAL} \) is not regular. ■
Theorem: $\text{EQUAL}$ is not regular.

Proof: By contradiction; assume that $\text{EQUAL}$ is regular. Let $n$ be the pumping length guaranteed by the weak pumping lemma. Let $w = 0^n?0^n$. Then $w \in \text{EQUAL}$ and $|w| = 2n + 1 \geq n$. Thus by the weak pumping lemma, we can write $w = xyz$ such that $y \neq \varepsilon$ and for any natural number $i$, $xy^iz \in \text{EQUAL}$. Then $y$ cannot contain $?$, since otherwise if we let $i = 0$, $xy^0z$ does not contain $?$ and would not be in $\text{EQUAL}$. So $y$ is either completely to the left of the $?$ or completely to the right of the $?$. Let $|y| = k$, so $k > 0$. Since $y$ is completely to the left or right of the $?$, then $y = 0^k$. Now, we consider two cases:

Case 1: $y$ is to the left of the $?$. Then $xy^2z = 0^n0^k?0^n \not\in \text{EQUAL}$, contradicting the weak pumping lemma.

Case 2: $y$ is to the right of the $?$. Then $xy^2z = 0^n?0^n0^k \not\in \text{EQUAL}$, contradicting the weak pumping lemma.

In either case we reach a contradiction, so our assumption was wrong and $\text{EQUAL}$ is not regular. ■

For any regular language $L$,

There exists a positive natural number $n$ such that

For any $w \in L$ with $|w| \geq n$,

There exists strings $x, y, z$ such that

For any natural number $i$,

$w = xyz$,

$y \neq \varepsilon$

$xy^iz \in L$

The hardest part of most proofs with the pumping lemma is choosing some string that we should be able to pump but cannot.
Theorem: \textit{EQUAL} is not regular.

\textit{Proof:} By contradiction; assume that \textit{EQUAL} is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n?0^n \).

For any regular language \( L \),
\textit{There exists} a positive natural number \( n \) such that

\textit{For any} \( w \in L \) with \(|w| \geq n\),
\textit{There exists} strings \( x, y, z \) such that
\textit{For any} natural number \( i \),
\( w = xyz \),
\( y \neq \epsilon \)
\( xy^iz \in L \)
Theorem: $EQUAL$ is not regular.

Proof: By contradiction; assume that $EQUAL$ is regular. Let $n$ be the pumping length guaranteed by the weak pumping lemma. Let $w = 0^n?0^n$. Then $w \in EQUAL$ and $|w| = 2n + 1 \geq n$. 

For any regular language $L$, there exists a positive natural number $n$ such that for any $w \in L$ with $|w| \geq n$, there exist strings $x, y, z$ such that for any natural number $i$, $w = xyz$, $y \neq \varepsilon$, $xy^iz \in L$. 

Theory: For any regular language L, there exists a positive natural number n such that for any w ∈ L with |w| ≥ n, there exists strings x, y, z such that for any natural number i, w = xyz, y ≠ ε, xy^iz ∈ L.
For any regular language $L$, there exists a positive natural number $n$ such that for any $w \in L$ with $|w| \geq n$, there exists strings $x, y, z$ such that for any natural number $i$,

$$w = xyz,$$

$y \neq \varepsilon$

$xy^iz \in L$

Theorem: $EQUAL$ is not regular.

Proof: By contradiction; assume that $EQUAL$ is regular. Let $n$ be the pumping length guaranteed by the weak pumping lemma. Let $w = 0^n?0^n$. Then $w \in EQUAL$ and $|w| = 2n + 1 \geq n$. 


Theorem: \textit{EQUAL} is not regular.

Proof: By contradiction; assume that \textit{EQUAL} is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n?0^n \). Then \( w \in \textit{EQUAL} \) and \( |w| = 2n + 1 \geq n \).
Theorem: \textit{EQUAL} is not regular.

Proof: By contradiction; assume that \textit{EQUAL} is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n \oplus 0^n \). Then \( w \in \textit{EQUAL} \) and \(|w| = 2n + 1 \geq n\). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in \textit{EQUAL} \).
**Theorem:** EQUAL is not regular.

**Proof:** By contradiction; assume that EQUAL is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n?0^n \). Then \( w \in \text{EQUAL} \) and \( |w| = 2n + 1 \geq n \). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in \text{EQUAL} \).
**Theorem:** \( \text{EQUAL} \) is not regular.

**Proof:** By contradiction; assume that \( \text{EQUAL} \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n?0^n \). Then \( w \in \text{EQUAL} \) and \( |w| = 2n + 1 \geq n \). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in \text{EQUAL} \).

At this point, we have some string that we should be able to split into pieces and pump. The rest of the proof shows that no matter what choice we make, this is impossible.
Theorem: \textit{EQUAL} is not regular.

Proof: By contradiction; assume that \textit{EQUAL} is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n?0^n \). Then \( w \in \textit{EQUAL} \) and \(|w| = 2n + 1 \geq n\). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in \textit{EQUAL} \). Then \( y \) cannot contain \(?\), since otherwise if we let \( i = 0 \), \( xy^iz = xz \) does not contain \(?\) and would not be in \textit{EQUAL}.

For any regular language \( L \),

\[ \text{There exists} \ \text{a positive natural number} \ n \ \text{such that} \]

For any \( w \in L \) with \(|w| \geq n\),

\[ \text{There exists} \ \text{strings} \ x, y, z \ \text{such that} \]

For any natural number \( i \),

\[ w = xyz, \]

\[ y \neq \varepsilon \]

\[ xy^iz \in L \]
Theorem: \textsc{Equal} is not regular.

Proof: By contradiction; assume that \textsc{Equal} is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n?0^n \). Then \( w \in \textsc{Equal} \) and \(|w| = 2n + 1 \geq n\). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \epsilon \) and for any natural number \( i \), \( xy^iz \in \textsc{Equal} \). Then \( y \) cannot contain \( ? \), since otherwise if we let \( i = 0 \), \( xy^iz = xz \) does not contain \( ? \) and would not be in \textsc{Equal}. So \( y \) is either completely to the left of the \( ? \) or completely to the right of the \( ? \).
Theorem: \( \text{EQUAL} \) is not regular.

Proof: By contradiction; assume that \( \text{EQUAL} \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n?0^n \). Then \( w \in \text{EQUAL} \) and \( |w| = 2n + 1 \geq n \). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in \text{EQUAL} \). Then \( y \) cannot contain \( ? \), since otherwise if we let \( i = 0 \), \( xy^iz = xz \) does not contain \( ? \) and would not be in \( \text{EQUAL} \). So \( y \) is either completely to the left of the \( ? \) or completely to the right of the \( ? \). Let \( |y| = k \), so \( k > 0 \).
Theorem: EQUAL is not regular.

Proof: By contradiction; assume that EQUAL is regular. Let n be the pumping length guaranteed by the weak pumping lemma. Let $w = 0^n?0^n$. Then $w \in EQUAL$ and $|w| = 2n + 1 \geq n$. Thus by the weak pumping lemma, we can write $w = xyz$ such that $y \neq \varepsilon$ and for any natural number $i$, $xy^iz \in EQUAL$. Then $y$ cannot contain $?$, since otherwise if we let $i = 0$, $xy^iz = xz$ does not contain $?$ and would not be in EQUAL. So $y$ is either completely to the left of the $?$ or completely to the right of the $?$. Let $|y| = k$, so $k > 0$. Since $y$ is completely to the left or right of the $?$, then $y = 0^k$.
For any regular language $L$, 
There exists a positive natural number $n$ such that 
For any $w \in L$ with $|w| \geq n$, 
There exists strings $x, y, z$ such that 
For any natural number $i$, 
\[ w = xyz, \]
\[ y \neq \varepsilon \]
\[ xy^i z \in L \]

**Theorem:** $EQUAL$ is not regular.

**Proof:** By contradiction; assume that $EQUAL$ is regular. Let $n$ be the pumping length guaranteed by the weak pumping lemma. Let $w = 0^n?0^n$. Then $w \in EQUAL$ and $|w| = 2n + 1 \geq n$. Thus by the weak pumping lemma, we can write $w = xyz$ such that $y \neq \varepsilon$ and for any natural number $i$, $xy^i z \in EQUAL$. Then $y$ cannot contain $?$, since otherwise if we let $i = 0$, $xy^i z = xz$ does not contain $?$ and would not be in $EQUAL$. So $y$ is either completely to the left of the $?$ or completely to the right of the $?$. Let $|y| = k$, so $k > 0$. Since $y$ is completely to the left or right of the $?$, then $y = 0^k$. Now, we consider two cases:
Theorem: \( \text{EQUAL} \) is not regular.

Proof: By contradiction; assume that \( \text{EQUAL} \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n?0^n \). Then \( w \in \text{EQUAL} \) and \( |w| = 2n + 1 \geq n \). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in \text{EQUAL} \). Then \( y \) cannot contain \( ? \), since otherwise if we let \( i = 0 \), \( xy^iz = xz \) does not contain \( ? \) and would not be in \( \text{EQUAL} \). So \( y \) is either completely to the left of the \( ? \) or completely to the right of the \( ? \). Let \( |y| = k \), so \( k > 0 \). Since \( y \) is completely to the left or right of the \( ? \), then \( y = 0^k \). Now, we consider two cases:

Case 1: \( y \) is to the left of the \( ? \).
For any regular language $L$, there exists a positive natural number $n$ such that for any $w \in L$ with $|w| \geq n$, there exist strings $x$, $y$, $z$ such that $w = xyz$, $y \neq \varepsilon$, $xy^i z \in L$.

**Theorem:** $\text{EQUAL}$ is not regular.

**Proof:** By contradiction; assume that $\text{EQUAL}$ is regular. Let $n$ be the pumping length guaranteed by the weak pumping lemma. Let $w = 0^n ? 0^n$. Then $w \in \text{EQUAL}$ and $|w| = 2n + 1 \geq n$. Thus by the weak pumping lemma, we can write $w = xyz$ such that $y \neq \varepsilon$ and for any natural number $i$, $xy^i z \in \text{EQUAL}$. Then $y$ cannot contain $?$, since otherwise if we let $i = 0$, $xy^0 z = xz$ does not contain $?$ and would not be in $\text{EQUAL}$. So $y$ is either completely to the left of the $?$ or completely to the right of the $?$. Let $|y| = k$, so $k > 0$. Since $y$ is completely to the left or right of the $?$, then $y = 0^k$. Now, we consider two cases:

**Case 1:** $y$ is to the left of the $?$. Then $xy^2 z = 0^{n+k} ? 0^n \notin \text{EQUAL}$, contradicting the weak pumping lemma.
Theorem: \textit{EQUAL} is not regular.

Proof: By contradiction; assume that \textit{EQUAL} is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n?0^n \). Then \( w \in \textit{EQUAL} \) and \( |w| = 2n + 1 \geq n \). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in \textit{EQUAL} \). Then \( y \) cannot contain \( ? \), since otherwise if we let \( i = 0 \), \( xy^iz = xz \) does not contain \( ? \) and would not be in \textit{EQUAL}. So \( y \) is either completely to the left of the \( ? \) or completely to the right of the \( ? \). Let \( |y| = k \), so \( k > 0 \). Since \( y \) is completely to the left or right of the \( ? \), then \( y = 0^k \). Now, we consider two cases:

Case 1: \( y \) is to the left of the \( ? \). Then \( xy^2z = 0^{n+k}?0^n \notin \textit{EQUAL} \), contradicting the weak pumping lemma.

Case 2: \( y \) is to the right of the \( ? \).
Theorem: $EQUAL$ is not regular.

Proof: By contradiction; assume that $EQUAL$ is regular. Let $n$ be the pumping length guaranteed by the weak pumping lemma. Let $w = 0^n?0^n$. Then $w \in EQUAL$ and $|w| = 2n + 1 \geq n$. Thus by the weak pumping lemma, we can write $w = xyz$ such that $y \neq \epsilon$ and for any natural number $i$, $xy^iz \in EQUAL$. Then $y$ cannot contain $?$, since otherwise if we let $i = 0$, $xy^iz = xz$ does not contain $?$ and would not be in $EQUAL$. So $y$ is either completely to the left of the $?$ or completely to the right of the $?$. Let $|y| = k$, so $k > 0$. Since $y$ is completely to the left or right of the $?$, then $y = 0^k$. Now, we consider two cases:

Case 1: $y$ is to the left of the $?$. Then $xy^2z = 0^{n+k}0^n \notin EQUAL$, contradicting the weak pumping lemma.

Case 2: $y$ is to the right of the $?$. Then $xy^2z = 0^n0^{n+k} \notin EQUAL$, contradicting the weak pumping lemma.
Theorem: EQUAL is not regular.

Proof: By contradiction; assume that EQUAL is regular. Let n be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n?0^n \). Then \( w \in \text{EQUAL} \) and \( |w| = 2n + 1 \geq n \). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in \text{EQUAL} \). Then \( y \) cannot contain \( ? \), since otherwise if we let \( i = 0 \), \( xy^iz = xz \) does not contain \( ? \) and would not be in EQUAL. So \( y \) is either completely to the left of the ? or completely to the right of the ?. Let \( |y| = k \), so \( k > 0 \). Since \( y \) is completely to the left or right of the ?, then \( y = 0^k \). Now, we consider two cases:

Case 1: \( y \) is to the left of the ?. Then \( xy^2z = 0^{n+k}0^n \notin \text{EQUAL} \), contradicting the weak pumping lemma.

Case 2: \( y \) is to the right of the ?. Then \( xy^2z = 0^n0^{n+k} \notin \text{EQUAL} \), contradicting the weak pumping lemma.

In either case we reach a contradiction, so our assumption was wrong and \( \text{EQUAL} \) is not regular.
**Theorem:** \( \textit{EQUAL} \) is not regular.

**Proof:** By contradiction; assume that \( \textit{EQUAL} \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Let \( w = 0^n\epsilon 0^n \). Then \( w \in \textit{EQUAL} \) and \( |w| = 2n + 1 \geq n \). Thus by the weak pumping lemma, we can write \( w = xyz \) such that \( y \neq \epsilon \) and for any natural number \( i \), \( xy^iz \in \textit{EQUAL} \). Then \( y \) cannot contain \( \epsilon \), since otherwise if we let \( i = 0 \), \( xy^0z = xz \) does not contain \( \epsilon \) and would not be in \( \textit{EQUAL} \). So \( y \) is either completely to the left of the \( \epsilon \) or completely to the right of the \( \epsilon \). Let \( |y| = k \), so \( k > 0 \). Since \( y \) is completely to the left or right of the \( \epsilon \), then \( y = 0^k \). Now, we consider two cases:

**Case 1:** \( y \) is to the left of the \( \epsilon \). Then \( xy^2z = 0^{n+k}0^n \notin \textit{EQUAL} \), contradicting the weak pumping lemma.

**Case 2:** \( y \) is to the right of the \( \epsilon \). Then \( xy^2z = 0^n0^{n+k} \notin \textit{EQUAL} \), contradicting the weak pumping lemma.

In either case we reach a contradiction, so our assumption was wrong and \( \textit{EQUAL} \) is not regular. ■
Nonregular Languages

- The weak pumping lemma describes a property common to all regular languages.
- Any language $L$ which does not have this property cannot be regular.
- What other languages can we find that are not regular?
A Canonical Nonregular Language

• Consider the language $L = \{ 0^n1^n \mid n \in \mathbb{N} \}$.  
  $$L = \{ \varepsilon, 01, 0011, 000111, 00001111, \ldots \}$$

• L is a **classic example** of a nonregular language.

• Intuitively: If you have only finitely many states in a DFA, you can't "remember" an arbitrary number of 0s.

• How would we prove that L is nonregular?
The Pumping Lemma as a Game

- The weak pumping lemma can be thought of as a game between you and an adversary.

- **You win** if you can prove that the pumping lemma fails.

- **The adversary wins** if the adversary can make a choice for which the pumping lemma succeeds.

- The game goes as follows:
  - **The adversary** chooses a pumping length $n$.
  - **You** choose a string $w$ with $|w| \geq n$ and $w \in L$.
  - **The adversary** breaks it into $x$, $y$, and $z$.
  - **You** choose an $i$ such that $xy^iz \notin L$ (if you can't, you lose!)
The Pumping Lemma Game

ADVERSARY

YOU
The Pumping Lemma Game

ADVERSARY

Maliciously choose pumping length n.

YOU
The Pumping Lemma Game

**ADVERSARY**

Maliciously choose pumping length $n$.

**YOU**

Cleverly choose a string $w \in L, |w| \geq n$.
The Pumping Lemma Game

ADVERSARY

Maliciously choose pumping length $n$.

Maliciously split $w = xyz$, $y \neq \varepsilon$

YOU

Cleverly choose a string $w \in L$, $|w| \geq n$
### The Pumping Lemma Game

<table>
<thead>
<tr>
<th><strong>ADVERSARY</strong></th>
<th><strong>YOU</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Maliciously choose pumping length n.</td>
<td></td>
</tr>
<tr>
<td>Maliciously split</td>
<td></td>
</tr>
<tr>
<td>$w = xyz, y \neq \varepsilon$</td>
<td></td>
</tr>
<tr>
<td>Cleverly choose a string</td>
<td></td>
</tr>
<tr>
<td>$w \in L,</td>
<td>w</td>
</tr>
<tr>
<td>Cleverly choose $i$ such that $xy^iz \notin L$</td>
<td></td>
</tr>
</tbody>
</table>
The Pumping Lemma Game

<table>
<thead>
<tr>
<th>ADVERSARY</th>
<th>YOU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maliciously choose pumping length $n$.</td>
<td>Cleverly choose a string $w \in L$, $</td>
</tr>
<tr>
<td>Maliciously split $w = xyz$, $y \neq \varepsilon$</td>
<td>Cleverly choose $i$ such that $xy^iz \notin L$</td>
</tr>
<tr>
<td>Grrr! Aaaargh!</td>
<td></td>
</tr>
</tbody>
</table>
The Pumping Lemma Game

\[ L = \{ 0^n1^n \mid n \in \mathbb{N} \} \]

<table>
<thead>
<tr>
<th>ADVERSARY</th>
<th>YOU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maliciously choose pumping length ( n ).</td>
<td>Cleverly choose a string ( w \in L,</td>
</tr>
<tr>
<td>Maliciously split ( w = xyz, y \neq \varepsilon )</td>
<td>Cleverly choose ( i ) such that ( xy^iz \notin L )</td>
</tr>
<tr>
<td>Grrr! Aaaargh!</td>
<td></td>
</tr>
</tbody>
</table>
The Pumping Lemma Game

\[ L = \{ 0^n1^n | n \in \mathbb{N} \} \]

**ADVERSARY**

Maliciously choose pumping length \( n \).

Maliciously split \( w = xyz, y \neq \varepsilon \)

Grrr! Aaaargh!

**YOU**

Cleverly choose a string \( w \in L, |w| \geq n \)

Cleverly choose \( i \) such that \( xy^i z \notin L \)

\[ 0^n1^n \]
Theorem: $L = \{ 0^n1^n \mid n \in \mathbb{N} \}$ is not regular.
**Theorem:** $L = \{ 0^n1^n \mid n \in \mathbb{N} \}$ is not regular.

**Proof:** By contradiction; assume $L$ is regular.

---

Theorem: $L = \{ 0^n1^n \mid n \in \mathbb{N} \}$ is not regular.

Proof: By contradiction; assume $L$ is regular.
Theorem: $L = \{ 0^n1^n \mid n \in \mathbb{N} \}$ is not regular.

Proof: By contradiction; assume $L$ is regular. Let $n$ be the pumping length guaranteed by the weak pumping lemma.

Consider the string $w = 0^n1^n$. Then $|w| = 2n \geq n$ and $w \notin L$, so we can write $w = xyz$ such that $y \neq \varepsilon$, and for any natural number $i$, $xy^iz \in L$. We consider three cases:

Case 1: $y$ consists solely of 0s. Then $xy^0z = xz = 0^{n-|y|}1^n$, and since $|y| > 0$, $y \notin L$.

Case 2: $y$ consists solely of 1s. Then $xy^0z = xz = 0^n1^{n-|y|}$, and since $|y| > 0$, $y \notin L$.

Case 3: $y$ consists of $k$ 0s followed by $m$ 1s. Then $xy^2z$ has the form $0^{n+k}1^{m-|y|}$, so $xy^2z \notin L$.

In all three cases we reach a contradiction, so our assumption was wrong and $L$ is not regular. ■
**Theorem:** $L = \{ 0^n1^n | n \in \mathbb{N} \}$ is not regular.

**Proof:** By contradiction; assume $L$ is regular. Let $n$ be the pumping length guaranteed by the weak pumping lemma. Consider the string $w = 0^n1^n$. Then $|w| = 2n \geq n$ and $w \not\in L$, so we can write $w = xyz$ such that $y \neq \varepsilon$, and for any natural number $i$, $xy^iz \not\in L$. We consider three cases:

**Case 1:** $y$ consists solely of 0s. Then $xy^0z = xz = 0^{n-|y|}1^n$, and since $|y| > 0$, $y \not\in L$.

**Case 2:** $y$ consists solely of 1s. Then $xy^0z = xz = 0^n1^{n-|y|}$, and since $|y| > 0$, $y \not\in L$.

**Case 3:** $y$ consists of $k$ 0s followed by $m$ 1s. Then $xy^2z$ has the form $0^{n+k}1^m0^k1^{n+m}$, so $xy^2z \not\in L$.

In all three cases we reach a contradiction, so our assumption was wrong and $L$ is not regular. $\blacksquare$
Theorem: \( L = \{ 0^n1^n \mid n \in \mathbb{N} \} \) is not regular.

Proof: By contradiction; assume \( L \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Consider the string \( w = 0^n1^n \). Then \( |w| = 2n \geq n \) and \( w \in L \), so we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in L \).
Theorem: \( L = \{ 0^n1^n \mid n \in \mathbb{N} \} \) is not regular.

Proof: By contradiction; assume \( L \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Consider the string \( w = 0^n1^n \). Then \( |w| = 2n \geq n \) and \( w \in L \), so we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in L \). We consider three cases:

Case 1: \( y \) consists solely of 0s. Then \( xy^0z = xz = 0^{n-|y|}1^n \), and since \( |y| > 0 \), \( y \notin L \).

Case 2: \( y \) consists solely of 1s. Then \( xy^0z = xz = 0^n1^{n-|y|} \), and since \( |y| > 0 \), \( y \notin L \).

Case 3: \( y \) consists of \( k \) 0s followed by \( m \) 1s. Then \( xy^2z \) has the form \( 0^{n+k}1^m0^k1^n+m \), so \( xy^2z \notin L \).

In all three cases we reach a contradiction, so our assumption was wrong and \( L \) is not regular. ■
Theorem: $L = \{ 0^n1^n \mid n \in \mathbb{N} \}$ is not regular.

Proof: By contradiction; assume $L$ is regular. Let $n$ be the pumping length guaranteed by the weak pumping lemma. Consider the string $w = 0^n1^n$. Then $|w| = 2n \geq n$ and $w \in L$, so we can write $w = xyz$ such that $y \neq \varepsilon$ and for any natural number $i$, $xy^iz \in L$. We consider three cases:

Case 1: $y$ consists solely of 0s.
Theorem: \( L = \{ 0^n1^n \mid n \in \mathbb{N} \} \) is not regular.

Proof: By contradiction; assume \( L \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Consider the string \( w = 0^n1^n \). Then \( |w| = 2n \geq n \) and \( w \in L \), so we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in L \). We consider three cases:

Case 1: \( y \) consists solely of 0s. Then \( xy^0z = xz = 0^{n-|y|}1^n \), and since \( |y| > 0 \), \( xz \notin L \).
Theorem: \( L = \{ 0^n1^n \mid n \in \mathbb{N} \} \) is not regular.

Proof: By contradiction; assume \( L \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Consider the string \( w = 0^n1^n \). Then \( |w| = 2n \geq n \) and \( w \in L \), so we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in L \). We consider three cases:

Case 1: \( y \) consists solely of 0s. Then \( xy^0z = xz = 0^{n-|y|}1^n \), and since \( |y| > 0 \), \( xz \notin L \).

Case 2: \( y \) consists solely of 1s.

Case 3: \( y \) consists of \( k \) 0s followed by \( m \) 1s. Then \( xy^2z \) has the form \( 0^{n+k}1^m0^k1^{n+m} \), so \( xy^2z \notin L \). In all three cases we reach a contradiction, so our assumption was wrong and \( L \) is not regular. ■
Theorem: L = \{ 0^n1^n | n \in \mathbb{N} \} is not regular.

Proof: By contradiction; assume L is regular. Let n be the pumping length guaranteed by the weak pumping lemma. Consider the string \( w = 0^n1^n \). Then \( |w| = 2n \geq n \) and \( w \in L \), so we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in L \). We consider three cases:

Case 1: \( y \) consists solely of 0s. Then \( xy^0z = xz = 0^{n-|y|}1^n \), and since \( |y| > 0 \), \( xz \notin L \).

Case 2: \( y \) consists solely of 1s. Then \( xy^0z = xz = 0^n1^{n-|y|} \), and since \( |y| > 0 \), \( xz \notin L \).
Theorem: \( L = \{ 0^n1^n \mid n \in \mathbb{N} \} \) is not regular.

Proof: By contradiction; assume \( L \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Consider the string \( w = 0^n1^n \). Then \( |w| = 2n \geq n \) and \( w \in L \), so we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in L \). We consider three cases:

Case 1: \( y \) consists solely of 0s. Then \( xy^0z = xz = 0^{n-|y|}1^n \), and since \( |y| > 0 \), \( xz \notin L \).

Case 2: \( y \) consists solely of 1s. Then \( xy^0z = xz = 0^n1^{n-|y|} \), and since \( |y| > 0 \), \( xz \notin L \).

Case 3: \( y \) consists of \( k \) 0s followed by \( m \) 1s.

In all three cases we reach a contradiction, so our assumption was wrong and \( L \) is not regular. ■
**Theorem:** \( L = \{ 0^n1^n \mid n \in \mathbb{N} \} \) is not regular.

**Proof:** By contradiction; assume \( L \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Consider the string \( w = 0^n1^n \). Then \(|w| = 2n \geq n\) and \( w \in L \), so we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in L \). We consider three cases:

**Case 1:** \( y \) consists solely of 0s. Then \( xy^0z = xz = 0^{n-|y|}1^n \), and since \(|y| > 0\), \( xz \notin L \).

**Case 2:** \( y \) consists solely of 1s. Then \( xy^0z = xz = 0^n1^{n-|y|} \), and since \(|y| > 0\), \( xz \notin L \).

**Case 3:** \( y \) consists of \( k \) 0s followed by \( m \) 1s.

\[
0 0 0 0 0 0 1 1 1 1 1 1
\]
**Theorem:** \( L = \{ 0^n1^n \mid n \in \mathbb{N} \} \) is not regular.

**Proof:** By contradiction; assume \( L \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Consider the string \( w = 0^n1^n \). Then \( |w| = 2n \geq n \) and \( w \in L \), so we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in L \). We consider three cases:

**Case 1:** \( y \) consists solely of 0s. Then \( xy^0z = xz = 0^n1^n \), and since \( |y| > 0 \), \( xz \notin L \).

**Case 2:** \( y \) consists solely of 1s. Then \( xy^0z = xz = 0^n1^n \), and since \( |y| > 0 \), \( xz \notin L \).

**Case 3:** \( y \) consists of \( k \) 0s followed by \( m \) 1s. Then \( xy^2z \) has the form \( 0^{n+k}1^m0^k1^{n+m} \), so \( xy^2z \notin L \).

In all three cases we reach a contradiction, so our assumption was wrong and \( L \) is not regular. ■
Theorem: $L = \{0^n1^n \mid n \in \mathbb{N}\}$ is not regular.

Proof: By contradiction; assume $L$ is regular. Let $n$ be the pumping length guaranteed by the weak pumping lemma. Consider the string $w = 0^n1^n$. Then $|w| = 2n \geq n$ and $w \in L$, so we can write $w = xyz$ such that $y \neq \varepsilon$ and for any natural number $i$, $xy^iz \in L$. We consider three cases:

Case 1: $y$ consists solely of 0s. Then $xy^0z = xz = 0^{n-|y|}1^n$, and since $|y| > 0$, $xz \notin L$.

Case 2: $y$ consists solely of 1s. Then $xy^0z = xz = 0^n1^{n-|y|}$, and since $|y| > 0$, $xz \notin L$.

Case 3: $y$ consists of $k$ 0s followed by $m$ 1s. Then $xy^2z$ has the form $0^{n+k}1^m$, so $xy^2z \notin L$.

In all three cases we reach a contradiction, so our assumption was wrong and $L$ is not regular. ■
Theorem: \( L = \{ 0^n 1^n \mid n \in \mathbb{N} \} \) is not regular.

Proof: By contradiction; assume \( L \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Consider the string \( w = 0^n 1^n \). Then \( |w| = 2n \geq n \) and \( w \in L \), so we can write \( w = xyz \) such that \( y \neq \epsilon \) and for any natural number \( i \), \( xy^i z \in L \). We consider three cases:

**Case 1:** \( y \) consists solely of 0s. Then \( xy^0 z = xz = 0^{n-|y|} 1^n \), and since \( |y| > 0 \), \( xz \notin L \).

**Case 2:** \( y \) consists solely of 1s. Then \( xy^0 z = xz = 0^n 1^{n-|y|} \), and since \( |y| > 0 \), \( xz \notin L \).

**Case 3:** \( y \) consists of \( k \) 0s followed by \( m \) 1s. 

\[
0000000111110011111111
\]
**Theorem:** \( L = \{ 0^n1^n \mid n \in \mathbb{N} \} \) is not regular.

**Proof:** By contradiction; assume \( L \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Consider the string \( w = 0^n1^n \). Then \( |w| = 2n \geq n \) and \( w \in L \), so we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in L \). We consider three cases:

**Case 1:** \( y \) consists solely of 0s. Then \( xy^0z = xz = 0^{n-|y|}1^n \), and since \( |y| > 0 \), \( xz \notin L \).

**Case 2:** \( y \) consists solely of 1s. Then \( xy^0z = xz = 0^n1^{n-|y|} \), and since \( |y| > 0 \), \( xz \notin L \).

**Case 3:** \( y \) consists of \( k \) 0s followed by \( m \) 1s. Then \( xy^2z \) has the form \( 0^n1^m0^k1^n \), so \( xy^2z \notin L \).
Theorem: L = \{ 0^n1^n \mid n \in \mathbb{N} \} is not regular.

Proof: By contradiction; assume L is regular. Let n be the pumping length guaranteed by the weak pumping lemma. Consider the string \( w = 0^n1^n \). Then \( |w| = 2n \geq n \) and \( w \in L \), so we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in L \). We consider three cases:

Case 1: \( y \) consists solely of 0s. Then \( xy^0z = xz = 0^{n-|y|}1^n \), and since \( |y| > 0 \), \( xz \notin L \).

Case 2: \( y \) consists solely of 1s. Then \( xy^0z = xz = 0^n1^{n-|y|} \), and since \( |y| > 0 \), \( xz \notin L \).

Case 3: \( y \) consists of \( k \) 0s followed by \( m \) 1s. Then \( xy^2z \) has the form \( 0^n1^m0^k1^n \), so \( xy^2z \notin L \).

In all three cases we reach a contradiction, so our assumption was wrong and L is not regular.
**Theorem:** \( L = \{ 0^n1^n \mid n \in \mathbb{N} \} \) is not regular.

**Proof:** By contradiction; assume \( L \) is regular. Let \( n \) be the pumping length guaranteed by the weak pumping lemma. Consider the string \( w = 0^n1^n \). Then \( |w| = 2n \geq n \) and \( w \in L \), so we can write \( w = xyz \) such that \( y \neq \varepsilon \) and for any natural number \( i \), \( xy^iz \in L \). We consider three cases:

**Case 1:** \( y \) consists solely of 0s. Then \( xy^0z = xz = 0^{n-|y|}1^n \), and since \( |y| > 0 \), \( xz \not\in L \).

**Case 2:** \( y \) consists solely of 1s. Then \( xy^0z = xz = 0^n1^{n-|y|} \), and since \( |y| > 0 \), \( xz \not\in L \).

**Case 3:** \( y \) consists of \( k \) 0s followed by \( m \) 1s. Then \( xy^2z \) has the form \( 0^n1^m0^k1^n \), so \( xy^2z \not\in L \).

In all three cases we reach a contradiction, so our assumption was wrong and \( L \) is not regular. ■
Counting Symbols

- Consider the alphabet $\Sigma = \{0, 1\}$ and the language $\text{BALANCE} = \{w \mid w \text{ contains an equal number of 0s and 1s.}\}$

- For example:
  - $01 \in \text{BALANCE}$
  - $110010 \in \text{BALANCE}$
  - $11011 \notin \text{BALANCE}$

- **Question:** Is $\text{BALANCE}$ a regular language?
BALANCE and the Weak Pumping Lemma

\[ BALANCE = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \} \]
**BALANCE** and the Weak Pumping Lemma

\[ \text{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \} \]
BALANCE and the Weak Pumping Lemma

$BALANCE = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \}$
\textit{BALANCE} and the Weak Pumping Lemma

\textit{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \}
**BALANCE** and the Weak Pumping Lemma

\[ \text{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s. } \} \]
BALANCE and the Weak Pumping Lemma

\[
\text{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \}
\]
\textit{BALANCE} and the Weak Pumping Lemma

\[ \textit{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \} \]
BALANCE and the Weak Pumping Lemma

BALANCE = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \}
$BALANCE$ and the Weak Pumping Lemma

$BALANCE = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \}$
**BALANCE** and the Weak Pumping Lemma

**BALANCE** = \( \{ w \mid w \text{ contains an equal number of 0s and 1s.} \} \)

1 0 0 1
BALANCE and the Weak Pumping Lemma

\[ \text{BALANCE} = \{ w | w \text{ contains an equal number of 0s and 1s.} \} \]
BALANCE and the Weak Pumping Lemma

\[ \text{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \} \]
\textit{BALANCE} and the Weak Pumping Lemma

\textit{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \}
BALANCE and the Weak Pumping Lemma

\[ \text{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s. } \} \]
An Incorrect Proof

Theorem: BALANCE is regular.

Proof: We show that BALANCE satisfies the condition of the pumping lemma. Let \( n = 2 \) and consider any string \( w \in BALANCE \) such that \( |w| \geq 2 \). Then we can write \( w = xyz \) such that \( x = z = \varepsilon \) and \( y = w \), so \( y \neq \varepsilon \). Then for any natural number \( i \), \( xy^i z = w^i \), which has the same number of 0s and 1s. Since BALANCE passes the conditions of the weak pumping lemma, BALANCE is regular. ■
An Incorrect Proof

Theorem: \textsc{Balance} is regular.

Proof: We show that \textsc{Balance} satisfies the condition of the pumping lemma. Let \( n = 2 \) and consider any string \( w \in \textsc{Balance} \) such that \(|w| \geq 2\). Then we can write \( w = xyz \) such that \( x = z = \varepsilon \) and \( y \neq \varepsilon \). Then for any natural number \( i \), \( xy^iz = w^i \), which has the same number of 0s and 1s. Since \textsc{Balance} passes the conditions of the weak pumping lemma, \textsc{Balance} is regular. ■
The Weak Pumping Lemma

- The **Weak Pumping Lemma for Regular Languages** states that

  **For any** regular language $L$,

  **There exists** a positive natural number $n$ such that

  **For any** $w \in L$ with $|w| \geq n$,

  **There exists** strings $x, y, z$ such that

  **For any** natural number $i$,

  $$w = xyz, \quad w \text{ can be broken into three pieces,}$$

  $$y \neq \varepsilon, \quad \text{where the middle piece isn't empty,}$$

  $$xy^iz \in L, \quad \text{where the middle piece can be replicated zero or more times.}$$
The Weak Pumping Lemma

- The **Weak Pumping Lemma for Regular Languages** states that

  For any regular language \( L \), there exists a positive natural number \( n \) such that

  For any \( w \in L \) with \( |w| \geq n \),

  There exists strings \( x, y, z \) such that

  For any natural number \( i \),

  \[ w = xyz, \quad w \text{ can be broken into three pieces}, \]

  \[ y \neq \varepsilon \quad \text{where the middle piece isn't empty}, \]

  \[ xy^iz \in L \quad \text{where the middle piece can be replicated zero or more times}. \]
Caution with the Pumping Lemma

- The weak (and full) pumping lemma describe a **necessary** condition of regular languages.
  - $L$ is regular $\rightarrow$ $L$ passes the pumping lemma
- The weak (and full) pumping lemma is not a **sufficient** condition of regular languages.
  - “$L$ passes the pumping lemma $\rightarrow$ $L$ is regular” is **not true**.
- If a language fails the pumping lemma, it is definitely **not** regular.
- If a language passes the pumping lemma, we learn nothing about whether it is regular or not.
**BALANCE is Not Regular**

- The language *BALANCE* can be proven not to be regular using a stronger version of the pumping lemma.
- To see the full pumping lemma, we need to revisit our original insight.
An Important Observation
An Important Observation
An Important Observation

The diagram above illustrates a finite automaton. The states are labeled as follows: $q_0$, $q_1$, $q_2$, $q_3$, $q_4$, and $q_5$. The transitions are labeled with inputs 0, 1, and 0, 1. The automaton starts at $q_0$ with the input string 0 1 1 0 1 1 1.
An Important Observation

Diagram:

- Start state: $q_0$
- Transitions:
  - $q_0$: 0 to $q_1$ and 1 to $q_2$
  - $q_1$: 1 to $q_2$ and 1 to $q_3$
  - $q_2$: 0, 1 to $q_5$
  - $q_3$: 1 to $q_4$
  - $q_4$: 1 to $q_4$

States:
- $q_0$, $q_1$, $q_2$, $q_3$, $q_4$, $q_5$
Weak Pumping Lemma Intuition

- Let D be a DFA with n states.
- Any string \( w \) accepted by D that has length at least n must visit some state twice.

  - Number of states visited is equal to \( |w| + 1 \).
  - By the pigeonhole principle, some state is duplicated.
- The substring of \( w \) in-between those revisited states can be removed, duplicated, tripled, quadrupled, etc. without changing the fact that \( w \) is accepted by D.
Pumping Lemma Intuition

- Let D be a DFA with n states.
- Any string $w$ accepted by D that has length at least $n$ must visit some state twice within its first $n$ characters.
  - Number of states visited is equal $n + 1$.
  - By the pigeonhole principle, some state is duplicated.
- The substring of $w$ in-between those revisited states can be removed, duplicated, tripled, quadrupled, etc. without changing the fact that $w$ is accepted by D.
The Weak Pumping Lemma

For any regular language $L$,

There exists a positive natural number $n$ such that

For any $w \in L$ with $|w| \geq n$,

There exists strings $x, y, z$ such that

For any natural number $i$,

$$w = xyz, \quad w \text{ can be broken into three pieces,}$$

$$y \neq \varepsilon \quad \text{where the middle piece isn't empty,}$$

$$xy^iz \in L \quad \text{where the middle piece can be replicated zero or more times.}$$
The Pumping Lemma

For any regular language $L$, there exists a positive natural number $n$ such that for any $w \in L$ with $|w| \geq n$, there exists strings $x, y, z$ such that for any natural number $i$, $w = xyz$, $|xy| \leq n$, $y \neq \varepsilon$, $xy^i z \in L$ where the first two pieces occur at the start of the string, where the middle piece isn't empty, where the middle piece can be replicated zero or more times.
Why This Change Matters

- The restriction $|xy| \leq n$ means that we can limit where the string to pump must be.
- If we specifically craft the first $n$ characters of the string to pump, we can force $y$ to have a specific property.
- We can then show that $y$ cannot be pumped arbitrarily many times.
BALANCE and the Pumping Lemma

\[ \text{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \} \]
**BALANCE** and the Pumping Lemma

**BALANCE** = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \}

Suppose the pumping length is 4.
$BALANCE$ and the Pumping Lemma

$BALANCE = \{ w \mid w \text{ contains an equal number of 0s and 1s. } \}$

Suppose the pumping length is 4.

\[ \begin{array}{cccccc}
0 & 0 & 0 & 0 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 \\
\end{array} \]
$\textit{BALANCE}$ and the Pumping Lemma

$\textit{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \}$

Suppose the pumping length is 4.

Since $|xy| \leq 4$, the string to pump must be somewhere in here.
**BALANCE** and the Pumping Lemma

**BALANCE** = \{ w | w contains an equal number of 0s and 1s. \}

Suppose the pumping length is 4.

![Example String](image)
BALANCE and the Pumping Lemma

\[ \text{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \} \]

Suppose the pumping length is 4.

\[
\begin{array}{ccccccc}
0 & 0 & 1 & 1 & 1 & 1 & 1 \\
\end{array}
\]
\textit{BALANCE} and the Pumping Lemma

\textit{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s. } \}

Suppose the pumping length is 4.
**BALANCE** and the Pumping Lemma

\[ BALANCE = \{ w \mid w \text{ contains an equal number of 0s and 1s.} \} \]

Suppose the pumping length is 4.
BALANCE and the Pumping Lemma

$\textit{BALANCE} = \{ w \mid w \text{ contains an equal number of 0s and 1s. } \}$

Suppose the pumping length is 4.

1 1 1 1
BALANCE and the Pumping Lemma

BALANCE = \{ w | w \text{ contains an equal number of 0s and 1s.} \}

Suppose the pumping length is 4.

\[ \underbrace{1 \ 1 \ 1 \ 1} \]
**BALANCE** and the Pumping Lemma

**BALANCE** = \{ w \mid w \text{ contains an equal number of } 0s \text{ and } 1s. \}

Suppose the pumping length is 4.

```
0 0 0 0 0 1 1 1 1 1
```
**BALANCE** and the Pumping Lemma

**BALANCE** = \{ w | w contains an equal number of 0s and 1s. \}

Suppose the pumping length is 4.

```
  0  0  0  1  1  1  1  1
```
**BALANCE** and the Pumping Lemma

**BALANCE** = \{ w | w contains an equal number of 0s and 1s. \}

Suppose the pumping length is 4.

0 0 0 1 1
Theorem: BALANCE is not regular.

Proof: By contradiction; assume that BALANCE is regular. Let n be the length guaranteed by the pumping lemma. Consider the string w = 0^n1^n. Then |w| = 2n ≥ n and w ∈ BALANCE. Therefore, there exist strings x, y, and z such that w = xyz, |xy| ≤ n, y ≠ ε, and for any natural number i, xy^iz ∈ BALANCE. Since |xy| ≤ n, y must consist solely of 0s. But then xy^2z = 0^{n+|y|}1^n, and since |y| > 0, xy^2z ∉ BALANCE. We have reached a contradiction, so our assumption was wrong and BALANCE is not regular. ■
Theorem: BALANCE is not regular.

Proof: By contradiction; assume that BALANCE is regular.

Let $n$ be the length guaranteed by the pumping lemma. Consider the string $w = 0^n1^n$. Then $|w| = 2n \geq n$ and $w \in \text{BALANCE}$. Therefore, there exist strings $x$, $y$, and $z$ such that $w = xyz$, $|xy| \leq n$, $y \neq \varepsilon$, and for any natural number $i$, $xy^i z \in \text{BALANCE}$. Since $|xy| \leq n$, $y$ must consist solely of 0s. But then $xy^2 z = 0^{n+|y|}1^n$, and since $|y| > 0$, $xy^2 z \notin \text{BALANCE}$. We have reached a contradiction, so our assumption was wrong and BALANCE is not regular. ■
Theorem: BALANCE is not regular.

Proof: By contradiction; assume that BALANCE is regular. Let n be the length guaranteed by the pumping lemma.
Theorem: BALANCE is not regular.

Proof: By contradiction; assume that BALANCE is regular. Let $n$ be the length guaranteed by the pumping lemma. Consider the string $w = 0^n1^n$. Then $|w| = 2n \geq n$ and $w \in \text{BALANCE}$. Therefore, there exist strings $x, y, z$ such that $w = xyz$, $|xy| \leq n$, $y \neq \varepsilon$, and for any natural number $i$, $xy^iz \in \text{BALANCE}$. Since $|xy| \leq n$, $y$ must consist solely of 0s. But then $xy^2z = 0^{n+|y|}1^n$, and since $|y| > 0$, $xy^2z \not\in \text{BALANCE}$. We have reached a contradiction, so our assumption was wrong and BALANCE is not regular. ■
**Theorem:** BALANCE is not regular.

**Proof:** By contradiction; assume that BALANCE is regular. Let $n$ be the length guaranteed by the pumping lemma. Consider the string $w = 0^n 1^n$. Then $|w| = 2n \geq n$ and $w \in BALANCE$. Since $|w| = 2n$, $y$ must consist solely of 0s. But then $xy^2z = 0^{n+|y|}1^n$, and since $|y| > 0$, $xy^2z \not\in BALANCE$. We have reached a contradiction, so our assumption was wrong and BALANCE is not regular. ■
Theorem: BALANCE is not regular.

Proof: By contradiction; assume that BALANCE is regular. Let $n$ be the length guaranteed by the pumping lemma. Consider the string $w = 0^n1^n$. Then $|w| = 2n \geq n$ and $w \in BALANCE$. Therefore, there exist strings $x, y,$ and $z$ such that $w = xyz$, $|xy| \leq n$, $y \neq \varepsilon$, and for any natural number $i$, $xy^iz \in BALANCE$. 


Theorem: BALANCE is not regular.

Proof: By contradiction; assume that BALANCE is regular. Let n be the length guaranteed by the pumping lemma. Consider the string \( w = 0^n1^n \). Then \( |w| = 2n \geq n \) and \( w \in \text{BALANCE} \). Therefore, there exist strings \( x, y, \) and \( z \) such that \( w = xyz \), \( |xy| \leq n \), \( y \neq \varepsilon \), and for any natural number \( i \), \( xy^iz \in \text{BALANCE} \). Since \( |xy| \leq n \), \( y \) must consist solely of 0s.
Theorem: BALANCE is not regular.

Proof: By contradiction; assume that BALANCE is regular. Let n be the length guaranteed by the pumping lemma. Consider the string $w = 0^n1^n$. Then $|w| = 2n \geq n$ and $w \in BALANCE$. Therefore, there exist strings $x, y, z$ such that $w = xyz$, $|xy| \leq n$, $y \neq \varepsilon$, and for any natural number $i$, $xy^iz \in BALANCE$. Since $|xy| \leq n$, $y$ must consist solely of 0s.

This is why the pumping lemma is more powerful than the weak pumping lemma. We can force $y$ to be made purely of 0s, rather than some combination of 0s and 1s.
Theorem: BALANCE is not regular.

Proof: By contradiction; assume that BALANCE is regular. Let n be the length guaranteed by the pumping lemma. Consider the string \( w = 0^n1^n \). Then \( |w| = 2n \geq n \) and \( w \in BALANCE \). Therefore, there exist strings x, y, and z such that \( w = xyz \), \( |xy| \leq n \), \( y \neq \varepsilon \), and for any natural number i, \( xy^iz \in BALANCE \). Since \( |xy| \leq n \), y must consist solely of 0s. But then \( xy^2z = 0^{n+|y|}1^n \), and since \( |y| > 0 \), \( xy^2z \notin BALANCE \).
Theorem: BALANCE is not regular.

Proof: By contradiction; assume that BALANCE is regular. Let $n$ be the length guaranteed by the pumping lemma. Consider the string $w = 0^n1^n$. Then $|w| = 2n \geq n$ and $w \in BALANCE$. Therefore, there exist strings $x$, $y$, and $z$ such that $w = xyz$, $|xy| \leq n$, $y \neq \varepsilon$, and for any natural number $i$, $xy^iz \in BALANCE$. Since $|xy| \leq n$, $y$ must consist solely of 0s. But then $xy^2z = 0^{n+|y|}1^n$, and since $|y| > 0$, $xy^2z \notin BALANCE$. We have reached a contradiction, so our assumption was wrong and BALANCE is not regular.
**Theorem:** BALANCE is not regular.

**Proof:** By contradiction; assume that BALANCE is regular. Let $n$ be the length guaranteed by the pumping lemma. Consider the string $w = 0^n1^n$. Then $|w| = 2n \geq n$ and $w \in \textit{BALANCE}$. Therefore, there exist strings $x$, $y$, and $z$ such that $w = xyz$, $|xy| \leq n$, $y \neq \varepsilon$, and for any natural number $i$, $xy^iz \in \textit{BALANCE}$. Since $|xy| \leq n$, $y$ must consist solely of 0s. But then $xy^2z = 0^{n+|y|}1^n$, and since $|y| > 0$, $xy^2z \not\in \textit{BALANCE}$. We have reached a contradiction, so our assumption was wrong and BALANCE is not regular. ■
Summary of the Pumping Lemma

• Using the pigeonhole principle, we can prove the weak pumping lemma and pumping lemma.

• These lemmas describe essential properties of the regular languages.

• Any language that fails to have these properties cannot be regular.
Non-Closure Properties

- Over the past three lectures we've developed an array of useful closure properties of regular languages.
- Now that we know the existence of nonregular languages, we can show that some transformations don't preserve regularity.
Finite Unions

- If $L_1$ and $L_2$ are regular, then $L_1 \cup L_2$ are regular.
- By induction, we can prove the following:
  - Let $L_1, L_2, \ldots, L_n$ be regular languages.
  - Then
    $$ \bigcup_{i=1}^{n} L_i $$
    is a regular language.
- This is a **finite union** of regular languages.
Finite Unions, Intuitively
Finite Unions, Intuitively
Finite Unions, Intuitively

\[ \text{start} \quad \longrightarrow \quad L_1 \]

\[ \text{start} \quad \longrightarrow \quad L_2 \]

\[ \text{start} \quad \longrightarrow \quad L_n \]

\[ \ldots \]
Finite Unions, Intuitively

\[ L_1 \]

\[ L_2 \]

\[ L_n \]
Infinite Unions

- Let $L_1, L_2, L_3, \ldots$ be an infinite number of regular languages.
- Is
  \[
  \bigcup_{i=1}^{\infty} L_i
  \]
  is a regular language?
- This is an **infinite union** of regular languages.
Infinite Unions, Intuitively

\[ L_1 \]

\[ L_2 \]

\[ L_3 \]

...
There are infinite many states here! We can't necessarily just combine them into one big machine.
Infinite Unions

• The infinite union of certain infinite sequences of regular languages is regular.
  • For example, Kleene star.
• Intuitively, though, it doesn't seem like we can always do this.
• Can we find a counterexample?
The Key Idea

- Find a nonregular language $N$ and a sequence of regular languages $L_1, L_2, L_3, \ldots$ such that

$$N = \bigcup_{i=1}^{\infty} L_i$$

- Conclude that the regular languages are not closed under infinite union.
Disproving Closure

- A language containing just one string is called a **singleton language**.
- Any singleton language is regular. (Why?)
- For example, these languages are all regular: 
  \{ \varepsilon \}, \{ 01 \}, \{ 0011 \}, \{ 000111 \}, \{ 00001111 \}, \ldots
- What is the infinite union of all these languages?
  \{ 0^n1^n \mid n \in \mathbb{N} \}, which is not regular!
Next Time

• **Beyond Regular Languages**
  • Context-free languages.
  • Context-free grammars.