Nonregular Languages

How can we show that a language is not regular?
The Pumping Lemma

Theorem. (Pumping Lemma) Let $L$ be an arbitrary regular language. Then there exists a positive integer $p$ with the following property:

Given an arbitrary member $w$ of $L$ having length at least $p$ (i.e., $|w| \geq p$), $w$ can be divided into three parts, $w = xyz$, such that

- $|y| \geq 1$ (the middle part is nonempty),
- $|xy| \leq p$ (the first two parts together have length at most $p$), and
- for each $i \geq 0$, $xy^i z \in L$ (removing or repeating the middle part produces members of $L$).
Proof of the Pumping Lemma

Let $L$ be an arbitrary regular language. Then there is an FA, say $M$, that decides $L$. Let $p$ be the number of states of $M$.

Let $w$ be an arbitrary member of $L$ having length $n$ with $n \geq p$.

Let $q_0, q_1, \ldots, q_n$ be the states that $M$ on input $w$. That is, for each $i$, after reading the first $i$ symbols of $w$, $M$ is at $q_i$.

Clearly, $q_0$ is the initial state of $M$. Also, because $w \in L$, $q_n$ is a final state of $M$. 
The Pigeonhole Principle

We are placing a number of pigeons in a number of holes.

If there are more pigeons than there are holes, at least one hole should host more than one pigeon.
The Pigeonhole Principle

We are placing a number of pigeons in a number of holes.

If there are more pigeons than there are holes, at least one hole should host more than one pigeon.

Consider $q_0, \ldots, q_p$ (the first $p + 1$ states that $M$ goes through on input $w$). By the pigeonhole principle, there exist $c$ and $d$, $0 \leq c < d \leq p$, such that $q_c = q_d$.

Pick an arbitrary such pair $(c, d)$. 
Proof of the Pumping Lemma (cont’d)

Let $x = w_1 \ldots w_c$, $y = w_{c+1} \ldots w_d$, and $z = w_{d+1} \ldots w_n$. Then

- $|y| \geq 1$,
- $|xy| \leq p$,
- $M$ transitions from $q_0$ to $q_c$ on $x$,
- $M$ transitions from $q_c$ to $q_c$ on $y$,
- $M$ transitions from $q_c$ to $q_n$ on $z$.

Thus, for every $i \geq 0$, $M$ transitions from $q_0$ to $q_n$ on $xy^i z$, and so $xy^i z$ is a member of $L$. 

CSC527, Chapter 1, Part 4 © 2012 Mitsunori Ogihara
Application of the Pumping Lemma

Example 1: \( B = \{0^n1^n \mid n \geq 0\} \) is not regular.
Application of the Pumping Lemma

Example 1: $B = \{0^n1^n \mid n \geq 0\}$ is not regular.

Proof Assume that $B$ is regular. Let $p$ be a constant from the pumping lemma for $B$. Let $w = 0^p1^p$. 
**Application of the Pumping Lemma**

**Example 1:** \( B = \{0^n1^n \mid n \geq 0\} \) is not regular.

**Proof** Assume that \( B \) is regular. Let \( p \) be a constant from the pumping lemma for \( B \). Let \( w = 0^p1^p \).

Then \( w \) is in \( B \) so it can be divided into \( w = xyz \) such that

- \(|y| \geq 1, \)
- \(|xy| \leq p, \) and
- for each \( i \geq 0 \), \( xy^iz \in B \).
Application of the Pumping Lemma

Example 1: \( B = \{0^n1^n \mid n \geq 0\} \) is not regular.

Proof Assume that \( B \) is regular. Let \( p \) be a constant from the pumping lemma for \( B \). Let \( w = 0^p1^p \).

Then \( w \) is in \( B \) so it can be divided into \( w = xyz \) such that

- \( |y| \geq 1 \),
- \( |xy| \leq p \), and
- for each \( i \geq 0 \), \( xy^iz \in B \).

Since \( |xy| \leq p \), both \( x \) and \( y \) consist solely of 0s. The word \( xyyz \) has more 0s than 1s, and thus, not in \( B \). However, by the pumping lemma, \( xyyz \in B \), a contradiction. Hence, \( B \) is not regular.
Illustrating Conversation

I think "0^n1^n" is regular...
I think "$0^n1^n$" is regular...

You're wrong, and here's why. Assuming it's regular, how many states do you need to build an FA for it?
Illustrating Conversation

Well, it must be large, but I think it should be less than a trillion...
Illustrating Conversation

Well, it must be large, but I think it should be less than a trillion...

Let $p$ be the number and $w=0^p1^p$. The Pumping Lemma divides this into $xyz$. What is $y$?
Well, $y = 0^k$ for some positive $k...$
Illustrating Conversation

Well, $y = 0^k$ for some positive $k$...

Then $xz$ must be a member, but it has fewer 0s than 1s, so it can't be. We thus have a contradiction.
Illustrating Conversation

Oh, I was so naive. I should have taken CSC527...

It's not too late.
Example 2

\[ C = \{ w \mid w \in \{0, 1\}^* \text{ and has an equal number of 0s and 1s } \} \text{ is not regular.} \]
Example 2

\[ C = \{ w \mid w \in \{0, 1\}^* \text{ and has an equal number of 0s and 1s } \} \] is not regular.

**Proof**  Assume, to the contrary, that \( C \) is regular. Let \( p \) be the constant from the pumping lemma for \( C \).
Example 2

\( C = \{ w \mid w \in \{0, 1\}^* \text{ and has an equal number of } 0\text{s and } 1\text{s } \} \) is not regular.

**Proof** Assume, to the contrary, that \( C \) is regular. Let \( p \) be the constant from the pumping lemma for \( C \).

Let \( w = 0^p1^p \). Then \( w = xyz \) such that \(|xy| \leq p\), \(|y| \geq 1\), and for every \( i \geq 0\), \( xy^iz \in C \).
Example 2

\[ C = \{ w \mid w \in \{0, 1\}^* \text{ and has an equal number of 0s and 1s } \} \]

is not regular.

**Proof**  Assume, to the contrary, that \( C \) is regular. Let \( p \) be the constant from the pumping lemma for \( C \).

Let \( w = 0^p1^p \). Then \( w = xyz \) such that \( |xy| \leq p \), \( |y| \geq 1 \), and for every \( i \geq 0 \), \( xy^iz \in C \).

Let \( w' = xz \). Then \( w' \in C \) but \( w' \) has fewer 0s than 1s.
**Example 3**

The language $F = \{vv \mid v \in \{0, 1\}^*\}$ is not regular ($F$ is the language of all even length strings over $\{0, 1\}$ whose first half is identical to the second half).

**Proof** Assume, to the contrary, that $F$ is regular. Let $p$ be a constant for which the pumping lemma holds for $F$. 
Example 3

The language \( F = \{vv \mid v \in \{0, 1\}^*\} \) is not regular (\( F \) is the language of all even length strings over \( \{0, 1\} \) whose first half is identical to the second half).

**Proof** Assume, to the contrary, that \( F \) is regular. Let \( p \) be a constant for which the pumping lemma holds for \( F \).

Let \( w = 0^p1^p0^p1^p \). Then, \( w \) is divided into \( w = xyz \) such that \(|y| > 0, |xy| \leq p, \) and \((\forall i \geq 0)[xy^i z \in F] \). Here \( y \in 0^* \) since \( w \) begins with \( 0^p \).
Example 3

The language $F = \{ vv \mid v \in \{0, 1\}^* \}$ is not regular ($F$ is the language of all even length strings over $\{0, 1\}$ whose first half is identical to the second half).

Proof Assume, to the contrary, that $F$ is regular. Let $p$ be a constant for which the pumping lemma holds for $F$.

Let $w = 0^p1^p0^p1^p$. Then, $w$ is divided into $w = xyz$ such that $|y| > 0$, $|xy| \leq p$, and $(\forall i \geq 0)[xy^iz \in F]$. Here $y \in 0^*$ since $w$ begins with $0^p$.

Pick $i = 0$, we have $0^q1^p0^p1^p \in F$, where $q < p$. This word cannot be decomposed as $uu$. This is a contradiction.
Example 4

\[ D = \{1^n^2 \mid n \geq 0\} \] is not regular.
Example 4

\[ D = \{1^{n^2} \mid n \geq 0\} \text{ is not regular.} \]

Proof Assume, to the contrary, that \( D \) is regular. Let \( p \) be a constant for which the pumping lemma holds for \( D \).
Example 4

\[ D = \{1^n^2 \mid n \geq 0\} \text{ is not regular.} \]

**Proof** Assume, to the contrary, that \( D \) is regular. Let \( p \) be a constant for which the pumping lemma holds for \( D \).

Let \( w = 1^p^2 \). Then \( w = xyz \) for some \( x, y, z \) such that \( |y| \geq 1 \), \( |xy| \leq p \), and \( (\forall i \geq 0)[xy^i z \in D] \).
Example 4

\[ D = \{1^{n^2} \mid n \geq 0\} \text{ is not regular.} \]

**Proof** Assume, to the contrary, that \( D \) is regular. Let \( p \) be a constant for which the pumping lemma holds for \( D \).

Let \( w = 1^{p^2} \). Then \( w = xyz \) for some \( x, y, z \) such that \( |y| \geq 1 \), \( |xy| \leq p \), and \( (\forall i \geq 0) [xy^iz \in D] \).

Let \( l = |y| \). Then \( 1 \leq l \leq p \). By plugging in \( i = 2 \), we have \( 1^{p^2+l} \in D \), but \( p^2 + l \leq p^2 + p < (p + 1)^2 \), a contradiction. \( \square \)
Example 5

\[ E = \{0^i1^j \mid i > j\} \text{ is not regular.} \]
Example 5

\[ E = \{0^i1^j \mid i > j\} \] is not regular.

**Proof** Assume, to the contrary, that \( E \) is regular. Let \( p \) be a constant for which the pumping lemma holds for \( E \).
Example 5

\[ E = \{0^i1^j \mid i > j\} \] is not regular.

**Proof** Assume, to the contrary, that \( E \) is regular. Let \( p \) be a constant for which the pumping lemma holds for \( E \).

Let \( w = 0^p1^{p-1} \). Then \( w = xyz \) for some \( x, y, z \) such that \( |y| \geq 1 \), \( |xy| \leq p \), and \((\forall i \geq 0)[xy^i z \in E]\). Here \( y \in 0^* \) since the first \( p \) symbols of \( w \) are all 0.
Example 5

\[ E = \{0^i1^j \mid i > j\} \text{ is not regular.} \]

**Proof** Assume, to the contrary, that \( E \) is regular. Let \( p \) be a constant for which the pumping lemma holds for \( E \).

Let \( w = 0^p1^{p-1} \). Then \( w = xyz \) for some \( x, y, z \) such that \( |y| \geq 1, |xy| \leq p \), and \((\forall i \geq 0)[xy^iz \in E]\). Here \( y \in 0^* \) since the first \( p \) symbols of \( w \) are all 0.

With \( i = 0 \), we have \( 0^q1^{p-1} \in E \), where \( q \leq p-1 \), a contradiction.