So far we've been dealing with regular languages exclusively, but if all languages were regular, we wouldn't bother specifying that they're regular. They would just be languages. In fact, we've run into a language already which is irregular.

- \( \{ w \in \Sigma^* | \#(w, a) = \#(w, b)^2 \} \) (\( \Sigma = \{a, b\} \)).

Other examples are easy to construct.

- \( \{ 0^p \mid p \text{ prime} \} \)
- \( \{ w \in \Sigma^* | w \text{ is a palindrome} \} \)
- \( \{ 0^n 1^n \mid n \geq 2 \} \)
- \( \{ 0^n 1^n \mid n \geq 3 \} \)
- \( \{ 0^n 1^n \mid n \geq 13 \} \cup \{ 0^n 1^n \mid k \neq 13 \} \)

Intuitively, what makes a language irregular is that we need to remember more than a constant amount of information. There are a few ways to formalize this notion (some better than others).

**Pumping Lemma**

If \( A \) is regular, then \( \exists p \in \mathbb{N} \) such that \( A \) contains a string of length at least \( p \), \( w \) may be written as \( w = xyz \) satisfying

1. \( \forall i \geq 0, \ xy^iz \in A \)
2. \( |y| > 0 \) (i.e. \( y \neq \epsilon \))
3. \( |xy| \leq p \)

**Proof Sketch**

The idea is if we pick \( p \) to be the number of states, then any input string of length \( p \) visits \( p + 1 \) states and thus (by the pigeonhole principle) must visit some state more than once. This, however, creates a cycle. We can extend this accepting path by taking the cycle as many (or as few) times as we like.
Let \( D = (Q, \Sigma, \delta, q_0, F) \) be a DFA with \( L(D) = A \) and pick \( p = |A| \).

If \( n \in A \), let \( q_0 \) be such that \( n \geq p \).

Let \( w = \alpha_1 ... \alpha_n \in A \) be such that \( n \geq p \).

Let \( q_0, ... q_n \) be the sequence of states \( D \) visits on input \( w \).

Then since this sequence contains \( n+1 \geq p+1 = |A|+1 \) states, it follows from the pigeon hole principle that there exists \( 0 \leq i < j \leq p+1 \) such that \( q_i = q_j \). But \( q_i, q_{i+1}, ... q_j \) forms a cycle, so we may also form an accepting path by removing the cycle or repeating it as many times as we wish.

\[ y \]

\[ x \rightarrow y \rightarrow x \rightarrow y \rightarrow x \rightarrow y \]

Define \( x = \alpha_1 ... \alpha_{i-1}, y = \alpha_i ... \alpha_{j-1}, z = \alpha_j ... \alpha_n \).

Note that \( x, z \) may be \( \varepsilon \) by \( y \neq \varepsilon \), hence \( |y| > 0 \).

Moreover, \( w = xyz \) and \( xy^k z \in A \) for nonnegative integers \( k \).

Lastly, \( |xy| \leq p \) by construction since \( j \leq p+1 \).

This is all three conditions of the lemma, so we're done.

\[ \Box \]

**Ex:** Prove \( A = \{ w \in \Sigma^* | \#(w, a) = \#(w, b)^2 \} \) (\( \Sigma = \{a, b\} \))

**Pf:** Suppose \( A \) is regular. Then \( \exists p \) satisfying the pumping lemma.

We must find a way to break one of the three conditions of the pumping lemma to arrive at a contradiction.

We'll target the 3rd condition.

Pick \( w = a^pb^p \in A \). Since \( |xy| \leq p \), we have that \( xy = a^p \) however we divide \( x \) and \( y \). But then \( \#(xy^2z, a) > \#(xy^2z, b) \) and the PL insists \( xy^2z \in A \).

\( \Box \)

Thus \( A \) is not regular.
Ex) $B = \{0^p \mid p \text{ prime} \}$ is irregular.

\[ \text{PF} \]
Suppose $B$ is regular, and let $p'$ be such that $p' \geq p$. 

Pick $p_0$ prime such that $p_0 \geq p'$.

Then $0^{p_0} \in B$.

Then $y = 0^k$ for some $0 \leq k \leq p'$.

But $w_i = xy^2z$ for $i \geq 0$ has length $p_0 + (i-1)k$, and most $1wl$ are not prime, yet the PL insists $w_i \in B$.

\[ \therefore B \text{ is irregular.} \]

Use the pumping lemma to complete the remaining four exercises.

1) Pumping breaks the palindrome condition.

2) Pump $0^p1^p$. Since $|w_1| > 1$, we break the $0^n1^n$ condition.

3) Pump $0^p1^p$ down so that we end up with $0^k1^p$ and $k \neq p$.

4) The pumping lemma does not apply to this language!

The second set is regular, and the first set can be written as $x = \epsilon, y = 0$ to pump into the second set (either up or down).