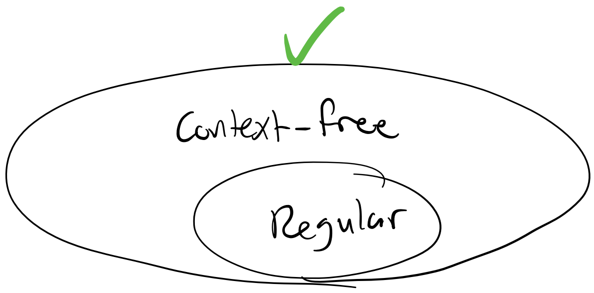
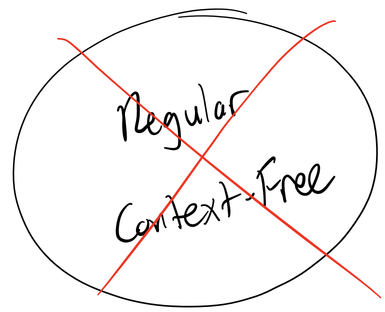
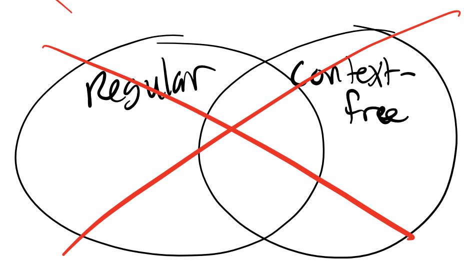
CSCI 301 - Lecture 28: The Pumping Lemma

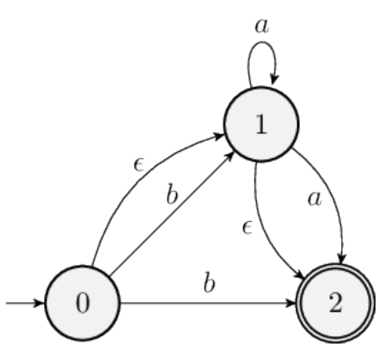
Possible realities:







NFA M:



Suppose there is a string $w \in L(M)$ and $|w| > 2$. What do we know about w?

- at most 1 b at beginning
- any amt of a's ≥ 2
- never went $0 \to 2$
- a's are at the end

The Pumping Lemma for Regular Languages

Let A be a regular language. There exist an integer p, called the pumping length, such that for every string $s \in A$ with $|s| \geq p$, s can be written as $s = xyz$, where:

- $y \neq \varepsilon$, or $|y| > 0$
- $|xy| \leq p$
- for all $i \geq 0$, $xy^iz \in A$

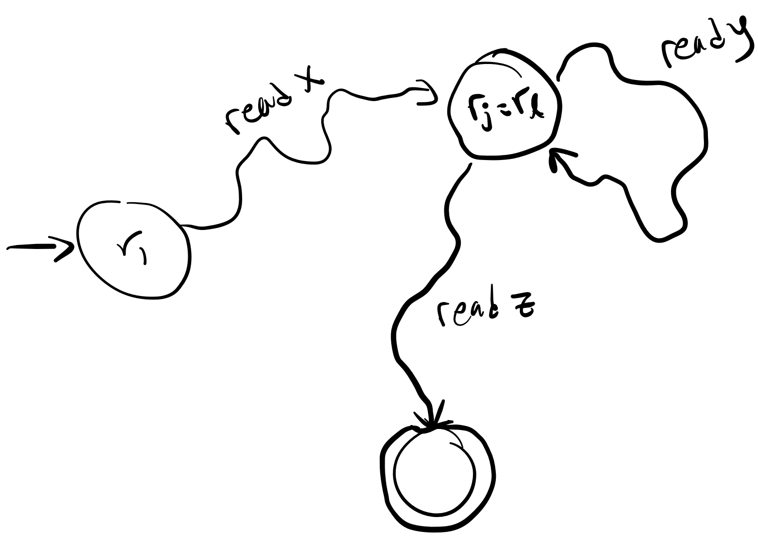
Proof sketch/intuition:

String symbols:

$S = S_1 S_2 S_3 \ldots S_n$

Machine states:

$r = r_1, r_2, r_3, \ldots, r_p, \ldots, r_k, \ldots, r_{n+1}$



Suppose A is regular. Let:

- M be a DFA where $L(M) = A$
- $p = |Q|$, the number of states in M.
- $s \in A$ where $|s| = n \geq p$

Observe:

- M visits $n+1 > p$ states while processing s.
- So there is some state visited twice by the time $p+1$ symbols are read.

$$r = \underbrace{r_1, r_2, \dots r_j, \dots r_\ell}_{\leq p+1}, \dots, r_{n+1}$$

Let:

- $r_j = r_\ell$ be the repeated state (with $j < \ell$, WLOG).

Notice $j < \ell \leq p+1$, because the state must have been repeated before we ran out of states.

- $x = s_1 s_2 \cdots s_{j-1}$
- $y = s_j s_{j+1} \cdots s_{\ell-1}$
- $z = s_\ell \cdots s_n$

1. $j < \ell$, so $y \neq \varepsilon$
2. $|xy| \leq p$ because $\ell \leq p+1$ so $\ell - 1 \leq p$
3. $xy^iz \in A$

Prove that $A = \{0^n 1^n : n \geq 0\}$ is not regular.

The pumping lemma (concisely):

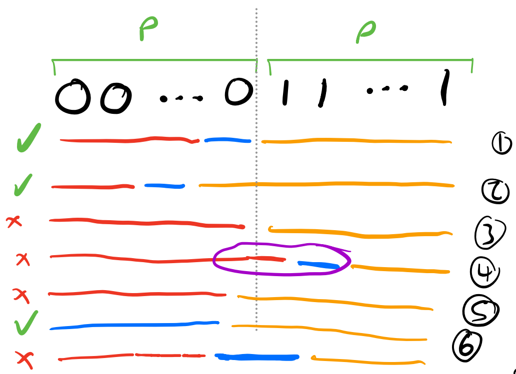
If A is regular, then $\exists p$ s.t. $\forall s \in A$, if $|s| \geq p$, then $s = xyz$, with
- $y \neq \varepsilon$
- $|xy| \leq p$
- $xy^iz \in A$ $\forall i \in \mathbb{N}$

Proof: By contradiction. Suppose A is regular, and let p be the pumping length. Consider the string: $0^p 1^p$

(P.L. conditions):

- $s \in A$? ✓
- $|s| \geq p$? ✓ $|s| = 2p$

So by the pumping lemma, s can be written as $s = xyz$



Which of these are possible?

Because $|xy| < p$, y contains only 0's. Because $y \neq \varepsilon$, it contains at least 1 zero.

$$xy^2z \notin A$$

$\{1^{n^2}\}$

Let $s = 1^{p^2}$

$|s| = p^2$

$s = xyz$

$s' = xy^2z \quad |s'| = p^2 + |y|$

$|y| \leq p$

$|s'| \leq p^2 + p$

$|s'| = q^2, q > p$

$(p+1)^2 = p^2 + 2p + 1 > p^2 + p \geq |s'|$

-

Lemma (The Pumping Lemma for Context-Free Languages): Let L be a context-free language. Then there exists an integer $p \geq 1$, called the pumping length, such that every string s in L with $|s| \geq p$ can be written as $s = uvxyz$, where

- $|vy| \geq 1$ (i.e., v and y are not both empty)
- $|vxy| \leq p$, and
- $uv^ixy^iz \in L$ for all $i \geq 0$.