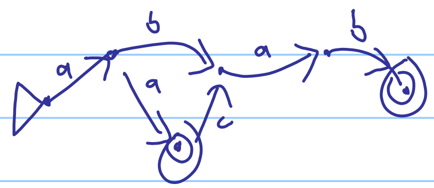
Some languages aren't regular

Note Title 3/13/2014

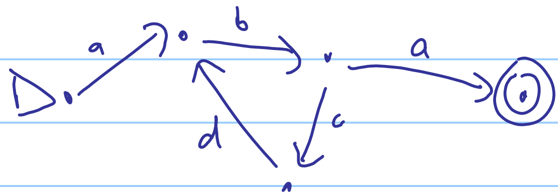
Claim 1: The transition graph of a dfa for any infinite regular language has cycles.

Proof: Each string in the language corresponds to a walk in the dfa's transition graph. If there are no cycles, there are only finitely many walks. $\square$

Examples:

- 

  No (directed) cycles, so finite language.

- 

  Has a cycle, so language is infinite.

Definition A string in a language can be pumped if we can repeat any number of copies of some substring while staying in the language.

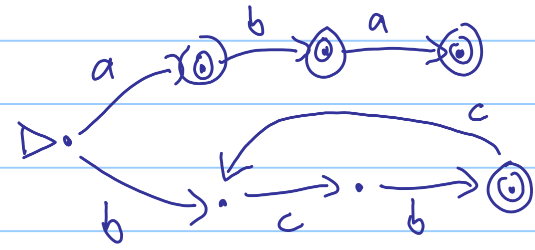
Examples:

- $L = \{ a^n (bc)^m d : n \geq 3, m \geq 0 \}$

  most strings in this language can be pumped.
  e.g. 'aaabcbcd' extends to
  'aaabcbcbc...bcd'

- $L = \{$ java programs $\}$
  every string in L containing whitespace can be pumped, by repeating the whitespace.

- L defined by



The strings a, ab, aba cannot be pumped.
The strings bcbccb, bcbccbccb can be pumped.

<u>Formal definition</u> String s in language L can be <u>pumped</u> if we can write $s = s_1 s_2 s_3$ with s_2 nonempty, such that

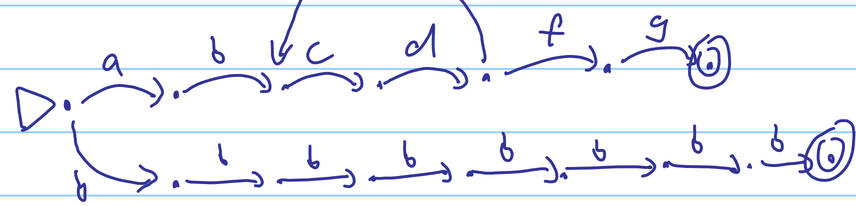
$$s_1 s_2^n s_3 \in L \quad \text{for all } n \geq 0.$$

example: In the last example above, s = bcbccb can be pumped with $s_1 = b$, $s_2 = cbc$, $s_3 = cb$

[informal version of "simplified pumping lemma"]
<u>Claim 2</u> Given an infinite, regular language L, every sufficiently long string in L can be pumped.

(proof is given later. Examples first.)

example: ① L defined by



All strings in L with length ≥ 9 can be pumped.
e.g. bbbbbbbb can't be pumped.
abcdecdfg can be pumped.

abcdecdecdfg can be pumped.

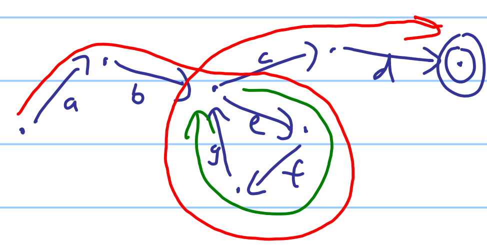
(2) L given by $\{ a^n b^m a^n : n \geq 0, \ 0 \leq m \leq 3 \}$
[NB L is not regular, as we will soon see].

The string $\underbrace{aaa\ldots a}_{\text{even number of a's}}$ can be pumped.
But aa...aabaa...a cannot be pumped.
More generally, there exist arbitrarily long strings in L that cannot be pumped. So L isn't regular, by claim 2.

"Simplified pumping lemma"
Formal statement of claim 2 : Let L be an infinite, regular language. Then there exists some $M > 0$ such that all strings in L, with length $\geq M$, can be pumped.

Proof: By claim 1, a transition graph for L contains a cycle. The graph is finite, so we must traverse a cycle after some maximum number of steps M. But this means any $s \in L$ with length $\geq M$ traverses a cycle. So we can pump s by adding any number of trips around the cycle.

Basic idea:



s is red

cycle is green

s is abefgcd

pumped version is $ab(efg)^n cd$ for $n \geq 0$. $\square$

We call Claim 2 the simplified Pumping Lemma.
We can use it to prove that some languages aren't regular.

Examples of applying the (simplified) Pumping Lemma:

<u>Claim</u> $a^n b^n$ isn't regular.

proof: No strings in this language can be pumped. This contradicts the (simplified) Pumping Lemma, which states that all sufficiently long strings in an infinite regular language can be pumped.

<u>Claim</u> $L = \{ a b^n c d e a^{2n} : n \geq 0 \}$ isn't regular.

proof: Suppose L is regular and aim for a contradiction. By the simplified Pumping Lemma, there exists some M such that all $s \in L$ with length $\geq M$ can be pumped. Consider the string

$$w = a b^M c d e a^{2M}$$

$w \in L$, and $|w| \geq M$, but w can't be pumped, contradicting the pumping lemma.

notation: $|s|$ is the length of s

<u>Claim 3</u> (Full pumping lemma) L infinite & regular. Then there exists M such that all $s \in L$ with $|s| \geq M$ can be pumped <u>before</u> the Mth symbol.

← technically, $|s_1 s_2| < M$

proof: The same proof works, since the cycle must be traversed in the first M transitions.

Example: $L = \{ a^n b^m c^n : n, m \geq 0 \}$ is not regular.

Proof: Note that the simplified Pumping Lemma isn't quite good enough, since we can always pump 'b's into a long string in L. But the Full Pumping Lemma gives us what we need:

Assume L regular. Then have M such that all $s \in L$ with $|s| \geq M$ can be pumped before M. Take $s = a^M c^M$. Now we can't pump before M and stay inside L, contradicting the full pumping lemma.

Example: Java is not regular.

proof: Assume it is. Restrict attention to the alphabet '{', '}' by ignoring all other symbols. Get M from pumping lemma. Consider the string $\{^M \}^M$. It can't be pumped before M, contradicting the pumping lemma. $\square$