

EECS 2001A : Introduction to the Theory of Computation

Suprakash Datta

Course page: <http://www.eecs.yorku.ca/course/2001>
Also on Moodle

The Big Question

- Are there languages that are not regular?

- How can we show a language is NOT regular?

An Example Problem

$$L = \{0^n 1^n \mid n \in \mathbb{N}\}$$

- ‘Playing around’ with FA convinces you that the ‘finiteness’ of FA is problematic for “all $n \in \mathbb{N}$ ”
- The problem occurs between the 0^n and the 1^n
- Informal: the memory of a FA is limited by the the number of states $|Q|$

Proving Non-regularity

- Prove a general statement – **no** DFA exists for a given problem
- Cannot assume an automaton structure or a specific strategy
- Need an argument that holds for **all** DFA's

Repeating DFA Paths

- Consider an accepting DFA M with size $|Q|$
- On a string of length p , $p + 1$ states get visited
- For an accepting path $p \geq |Q|$, there must be a j such that the computational path looks like: $q_1, \dots, q_j, \dots, q_j, \dots, q_k$
- The action of the DFA in q_j, \dots, q_j is always the same. If we repeat (or ignore) the q_j, \dots, q_j part, the new path will again be an accepting path

Line of Reasoning

Proof by contradiction:

- Assume that L is regular
- Hence, there is a DFA M that recognizes L
- For strings of length $> |Q|$ the DFA M has to 'repeat itself'
- Show that M will accept strings outside L
- Conclude that the assumption was wrong, so L is not regular

Note that we use the simple DFA, not the more elaborate (but equivalent) NFA or GNFA

Pumping Lemma

For every regular language L , there is a **pumping length** p , such that for any string $s \in L$ and $|s| \geq p$, we can write $s = xyz$ with

- 1 $xy^iz \in L$ for every $i \in \{0, 1, 2, \dots\}$

- 2 $|y| \geq 1$

- 3 $|xy| \leq p$

Note that:

(1) implies that $xz \in L$.

(2) says that y cannot be the empty string ϵ .

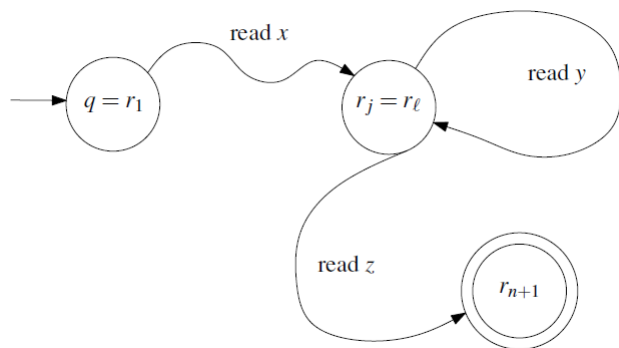
(3) is not always used

Formal Proof of the Pumping Lemma

- Let $M = (Q, \Sigma, \delta, q_1, F)$ with $Q = \{q_1, \dots, q_p\}$
- Let $s = s_1 \dots s_n \in L(M)$ with $|s| = n \geq p$
- Computational path of M on s is the sequence r_1, \dots, r_{n+1} , with each $r_i \in Q$, and with $r_1 = q_1$, $r_{n+1} \in F$ and $r_{t+1} = \delta(r_t, s_t)$ for $1 \leq t \leq n$
- When DFA M reads s , it visits the states r_1, \dots, r_{n+1} , and since $s \in L(M)$, $r_{n+1} \in F$
- Because $n + 1 \geq p + 1 > p$, there exist j, k such that $r_j = r_k$ (with $j < k$ and $k \leq p + 1$)

Formal Proof of the Pumping Lemma

- Let $x = s_1..s_{j-1}$, $y = s_j..s_{k-1}$, and $z = s_k...s_{n+1}$.
 x takes M from $q_1 = r_1$ to r_j , y takes M from r_j to r_j , and z takes M from r_j to $r_{n+1} \in F$
 As a result: xy^iz takes M from q_1 to $r_{n+1} \in F$ ($i \geq 0$)



Pumping Lemma-based Proof: Example 1

$$B = \{0^n 1^n \mid n \in \{0, 1, \dots\}\}$$

- Assume that B defined above is regular
- Let p be the pumping length, and $s = 0^p 1^p \in B$
- P.L.: There exists $s = xyz = 0^p 1^p$, with $xy^i z \in B$ for all $i \geq 0$
- Three options for y :
 - $y = 0^k$, hence $xyyz = 0^{p+k} 1^p \notin B$
 - $y = 1^k$ hence $xyyz = 0^p 1^{k+p} \notin B$
 - $y = 0^k 1^m$, hence $xyyz = 0^p 1^m 0^k 1^p \notin B$
- Conclusion: The pumping lemma does not hold, the language B is not regular.

Pumping Lemma-based Proof: Example 2

$$C = \{ww \mid w \in \{0, 1\}^*\}$$

- Let p be the pumping length, and take $s = 0^p 1 0^p 1 \in B$
- and take $s = 0^p 1 0^p 1$
- Let $s = xyz = 0^p 1 0^p 1$ with condition 3) $|xy| \leq p$
- Only one option: $y = 0^k$, so $xyyz = 0^{p+k} 1 0^p 1 \notin B$
- Without (3) this would have been a pain.

Intersecting Regular Languages

$$C = \{w \mid w \in \{0, 1\}^*, w \text{ has equal numbers of } 0, 1\}$$

- Using the Pumping Lemma is not easy
- Idea: If C is regular and F is regular, then the intersection $C \cap F$ has to be regular as well
- Assume that C is regular
- Take the regular language $F = \{0^n 1^m \mid n, m \in \mathbb{N}\}$, then the intersection is $C \cap F = \{0^n 1^n \mid n \in \mathbb{N}\}$
- But we know that $C \cap F$ is not regular
- Conclusion: C is not regular

Example of Pumping Down

$$E = \{0^i 1^j \mid i, j \in \mathbb{N}, i \geq j\}$$

- Problem: 'pumping up' $s = 0^p 1^p$ with $y = 0^k$ gives $xyyz = 0^{p+k} 1^p$, $xy^3z = 0^{p+2k} 1^p$, which are all in E (hence do not give contradictions)
- Solution: pump down to $xz = 0^{p-k} 1^p$
- Overall for $s = xyz = 0^p 1^p$ (recall $|xy| \leq p$): $y = 0^k$, hence $xz = 0^{p-k} 1^p \notin E$
- Contradiction: E is not regular

Pumping Lemma Usage: Summary

- You are given a pumping number
- You choose a string s
- There exist x, y, z (satisfying some criteria)
- You choose i in xy^iz , and show it violates criterion of set for that i
- Now you can conclude that the language is not regular

Note: There are other, less general, less popular techniques for showing a language is not regular