CS 4510 Automata and Complexity 1/23/2023

Lecture 4: The Pumping Lemma

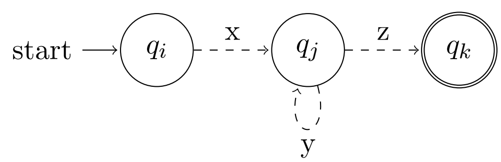
Lecturer: Abrahim Ladha *Scribe(s): Michael Wechsler*

1 Background

We previously mentioned that we have some intuition on the limitations of DFAs.

- ex) $\{a^n b^n \mid n \in \mathbb{N}\}$ should not be possible with finite memory since it needs a counter

Suppose we have a DFA, D, made up of p states. Consider a word, w, such that $|w| = p+1$. When we simulate $D(w)$, consider the sequence of states visited when deciding if $w \in L(D)$



By the Pigeonhole Principle, some state (q_j above) appears twice in this sequence, and our DFA must contain a loop. The details of the loop are not important, we just want to know that it exists.

⋆ Pigeonhole Principle: if you have $p+1$ pigeons and p pigeonholes, there must be at least 1 pigeonhole with greater than 1 pigeons in it

Claim: If $xyz \in L, \forall i, xy^i z \in L$ due to the DFA pictured above

Important Fact: If a language is regular, then it can be pumped

- Contrapositive (Equivalent): If a language cannot be pumped, then it is not regular
- Converse (**NOT** Equivalent): If a language can be pumped, then it is regular
  - This is not true and will be seen at the end of note section 2.3.2

2 Formula

The pumping lemma has many moving pieces and can be tricky to apply. I suggest you use this formula exactly. L is the language we want to prove is not regular.

1. Assume to the contrary, L is regular with pumping length p (pumping length is the number of states in L's DFA)
2. Choose some $s \in L$ such that $|s| \geq p$

3. For all cases $s = xyz$ such that $|xy| \leq p$ and $|y| > 0$
4. Choose $i \neq 1$ such that $xy^iz \notin L$
5. Conclude that L cannot be pumped, which means L is not regular

We require that $|s| \geq p$ so we have the pigeonhole condition, and we can pump the string. We require that $|y| > 0$ so we pump something non-trivial. For every language, you can pump $y = \varepsilon$ since its true that $\forall i \ \varepsilon^i = \varepsilon$. We require that $|xy| \leq p$ so that our pigeonhole condition occurs somewhere in what we have denoted as xy.

3 Examples

3.1 $L_1 = \{0^n1^n \mid n \in \mathbb{N}\}$

1. Assume to the contrary, L_1 is regular and $\exists$ a DFA, D, with a pumping length p

2. Let $s = 0^p1^p$
   $s \in L_1$ and $|s| = 2p \geq p$

3. There is only 1 case since the first p characters in the string are all 0s
   $x = 0^a, \ y = 0^b, \ z = 0^{p-a-b}1^p$
   Subject to $|xy| = a + b \leq p$ and $|y| = b > 0$

4. Look at $i = 2$
   $xy^2z = xyyz = 0^a0^b0^b0^{p-a-b}1^p = 0^{p+b}1^p$

5. We know that $b > 0$, so the number of 0s $\neq$ the number of 1s since $p + b > p$. Thus, L_1 cannot be pumped, and as a result, is not regular.

3.1.1 $L_2 = \{ww^R \mid w \in \Sigma^*\}$ (Even-length palindromes)

⋆ Note: a string "raised" to R is the reverse of the string

1. Assume to the contrary, L_2 is regular and $\exists$ a DFA, D, with a pumping length p

2. Let $s = 0^{p-1}110^{p-1}$ (We are choosing a poor s on purpose)
   $s \in L_2$ and $|s| = 2p \geq p$

3. The first p characters in the string are different, meaning there are several cases (2)

(a) $x = 0^a,\ y = 0^b,\ z = 0^{p-1-a-b}110^{p-1}$
Subject to $|xy| = a + b \leq p$ and $|y| = b > 0$

(b) $x = 0^a,\ y = 0^{p-1-a}1,\ z = 10^{p-1}$
Subject to $|xy| = a + p - 1 - a + 1 = p \geq p$ and $|y| = p - 1 - a + 1 > 0$

4. Choose i for the 2 cases

   (a) Look at $i = 2$
   $xy^2z = xyyz = 0^a0^b0^b0^{p-1-a-b}110^{p-1} = 0^{p-1+b}110^{p-1}$
   Since $b > 0$, we know that $p - 1 + b \neq p - 1$. Therefore, the two sections of 0s are unequal and the string is not a palindrome

   (b) Look at $i = 0$
   $xy^0z = xz = 0^a10^{p-1}$
   Since there is only one 1, this is not an even-length palindrome

5. For both cases, the language could not be pumped. Therefore, L_2 is not regular.

⋆ Note: Letting $s = 0^p0^p$ would result in a successful pump, but this does not mean the language is regular. Try to choose s carefully to avoid this situation.

3.2 $L_3 = \{ww \mid w \in \Sigma^*\}$, assume $\Sigma = \{0, 1\}$

1. Assume to the contrary, L_3 is regular and $\exists$ a DFA, D, with a pumping length p

2. Let $s = 0^p10^p1$
   $s \in L_3$ and $|s| = 2p + 2 \geq p$

3. There is only 1 case since the first p characters in the string are all 0s
   $x = 0^a,\ y = 0^b,\ z = 0^{p-a-b}10^p1$
   Subject to $|xy| = a + b \leq p$ and $|y| = b > 0$

4. Look at $i = 2$
   $xy^2z = xyyz = 0^a0^b0^b0^{p-a-b}10^p1 = 0^{p+b}10^p1$

   Take the right-most $p + 2$ characters in xy^2z. This string, which we'll call $w_2 = 10^p1$. Now, there are two cases for the leftmost string, $w_1 = 0^{p+b}$.

   (a) If $b = 1$, xy^2z is not even length, and therefore not in L_3
   (b) If $b > 1$, the midpoint of $xy^2z = w_1w_2$ is in the first block of 0s. We can tell that $w_1 \neq w_2$, and therefore, xy^2z is not in L_3

5. Both cases end with the pumped string not being in L_3. Thus, L_3 cannot be pumped and is not regular.

3.3 $L_4 = \{0^n 1^n 2^n \mid n \in \mathbb{N}\}$ - A note on choosing a BAD s

- Assume to the contrary, L_4 is regular and $\exists$ a DFA, D, with a pumping length p
- Let $s = 0^{\lfloor p/3 \rfloor + 1} 1^{\lfloor p/3 \rfloor + 1} 2^{\lfloor p/3 \rfloor + 1}$
  $s \in L_4$ and $|s| > p + 1 \geq p$
- There are 6 cases for this string

```
0............01............12............2
| x | y |              z                  |
| x |       y      |         z            |
| x |            y             |    z     |
|      x      | y |          z            |
|      x      |        y       |    z     |
|              x               | y | z    |
```

There are too many cases. Choose a better s instead of continuing this proof. For each of the six cases, you would be required to construct x, y, z choose i, and show $xy^i z \notin L$. It can be easy to miss a case. A better choice is $s = 0^p 1^p 2^p$ which has a similar proof as shown previously.

3.4 $L_5 = \{1^{n^2} \mid n \in \mathbb{N}\}$

1. Assume to the contrary, L_5 is regular and $\exists$ a DFA, D, with a pumping length p
2. Let $s = 1^{p^2}$
   $s \in L_5$ and $|s| = p^2 \geq p$
3. There is only 1 case since the first p characters in the string are all 1s
   $x = 1^a,\ y = 1^b,\ z = 1^{p^2 - a - b}$
   Subject to $|xy| = a + b \leq p$ and $|y| = b > 0$
4. Look at $i = 2$
   $xy^2 z = xyyz = 1^a 1^b 1^b 1^{p^2 - a - b} = 1^{p^2 + b}$

   Since $b > 0$, $p^2 + b > p^2$
   Since $a + b \leq p$, $b \leq p$
   Thus $p^2 + b \leq p^2 + p < p^2 + p + (p + 1) = p^2 + 2p + 1 = (p + 1)^2$

   By the first and third lines, we know $|1^{p^2}| < |1^{p^2 + b}| < |1^{(p+1)^2}|$
5. By the last line, we can see that $xy^2 z$ falls between two adjacent strings in L_5. Therefore, it is not in L_5. Thus, L_5 cannot be pumped and is not regular.