Nonregular language I

We are interested in the limit of finite automata

- Some languages not recognized

\[ \{0^n1^n \mid n \geq 0 \} \]

- We might remember \#0 first
- But \# of possible \( n \)’s is \( \infty \)
- Thus we cannot recognize it by finite automata
- However, this is not a formal proof
- It may be difficult to quickly tell if a language is regular or not
Consider two languages

\[ C = \{ w \mid \#0 = \#1 \} \]
\[ D = \{ w \mid \#01 = \#10 \} \]

It seems that both are not regular
Indeed, \( C \) is not regular but \( D \) is
This is an exercise in the book, so we don’t give details

To formally prove a language is not regular, we will introduce the pumping lemma
Pumping lemma I

Strategy: by contradiction
We prove

\[ \text{regular} \Rightarrow \text{some properties} \]

If “some properties” cannot hold, then the language is not regular
If $A$ regular $\Rightarrow \exists p$ (pumping length) such that 
$\forall s \in A, |s| \geq p, \exists x, y, z$ such that $s = xyz$ and

1. $\forall i \geq 0, xy^iz \in A$
2. $|y| > 0$
3. $|xy| \leq p$

Note that for $y^i$, we have $y^0 = \epsilon$
Proof of pumping lemma I

- Because $A$ is regular, $\exists$ a DFA to recognize $A$
- Let $p = \# \text{ states of this DFA}$
- If no string $s$ such that $|s| \geq p$, then the theorem statement is satisfied
- Now consider any $s$ with $|s| \geq p$

\[ s = s_1 \cdots s_n \]
Proof of pumping lemma II

- To process this string, assume the state sequence is

\[ q_1 \cdots q_{n+1} \]

- Because \(|s| \geq p\), we have

\[ n + 1 > p \]

- In \(1 \ldots p + 1\) two states must be the same (pigeonhole principle)

Fig 1.72
Proof of pumping lemma III
Proof of pumping lemma IV

- Assume $q_j$ and $q_l$ with $j \leq p + 1$, $l \leq p + 1$

are two same states. Then

$$x = s_1 \cdots s_{j-1}, \quad y = s_j, \cdots s_{l-1}, \quad z = s_l \cdots s_n$$

Because $j \neq l$,

$$|y| > 0$$

From $l \leq p + 1$, we have

$$|xy| \leq p$$

Thus all conditions are satisfied
Let's apply pumping lemma to prove that

\[ B = \{0^n1^n \mid n \geq 0\} \]

is not regular

Assume \( B \) is regular. From the lemma there is \( p \) such that

\[ s = 0^p1^p = xyz \]

By the lemma, \(|y| > 0\) and \(xy^iz \in B, \forall i \geq 0\)
Example 1.73 II

1. If

\[ y = 0 \cdots 0 \]

then

\[ xy = 0 \cdots 0 \text{ and } z = 0 \cdots 01 \cdots 1 \]

Thus

\[ xyyz : \#0 > \#1 \]

2. Then \( xy^2z \notin B \), a contradiction
Example 1.73 III

3. If

\[ y = 1 \cdots 1, \]

similarly

\[ xy^2z \notin B \text{ as } \#0 < \#1 \]

4. If

\[ y = 0 \cdots 01 \cdots 1 \]

then

\[ xyyz \notin B \text{ as it is not in the form of } 0^?1^? \]
Example 1.73 IV

We see that the condition

\[ |xy| \leq p \]

is not used
Example 1.39 |

- $C = \{ w \mid \#0 = \#1 \}$
- We follow the previous example to have

\[ s = 0^p 1^p = xyz \]

- However, we cannot get the needed contradiction for the case of

\[ y = 0 \cdots 01 \cdots 1 \]
Example 1.39 II

- Earlier we said
  
  \[ xyyz \text{ not in the form of } 0^?1^? \]

  but now we only require

  \[ \#0 = \#1 \]

- It is possible that

  \[ x = \epsilon, \ z = \epsilon, \ y = 0^p1^p \]

  and then

  \[ |y| > 0 \text{ and } xy^iz \in C \]
Example 1.39 III

- The 3rd condition should be applied

\[ |xy| \leq p \Rightarrow y = 0 \cdots 0 \text{ in } s = 0^p1^p \]

Then

\[ xyyz \not\in C \]

- Question: the pumping says

\[ \forall s \in A, \cdots \]

but why in the examples we analyzed a particular \( s \)?
And it seems that the selection of $s$ is important. Why?

Below we will explain our use of the lemma