Regular languages

Definition: a language is called a regular language if some finite automaton recognizes it

Regular languages use finite memory (finite states)
Non-regular languages require infinite memory

Are the following regular?

L1 = \{w \mid w \text{ has at least 100 1's}\}
Yes: Start at q₀, For each 1 qₖ \rightarrow qₖ₊₁. F=\{q₁₀₀\}

L2 = \{w \mid w \text{ has same number of 0's and 1's}\}
No: unknown number of states

L3 = \{w \mid w \text{ is of the form } 0^n1^n, n>0\}
No: unknown number of states

What about this class of languages

\(\Sigma = \{a, b\}\)

\(L_n = \{w \mid w \text{ contains } n \text{ b's in a row} \}\)
- \(L_1 = \{abbba, aabbbba, ababbbba, \ldots\}\)
- \(L_2 = \{babbba, bbbbb, aabbbbba, \ldots\}\)

\(L_n\) is regular for each value of \(n\)
Regular languages can be infinite
• E.g., a(ba)*b

For FSA to generate an infinite set of strings, there must be a loop between some states

Pumping lemma

Every string in regular language L with length greater than or equal to the pumping length p can be “pumped”

Every string s ∈ L (|s|≥p) can be written as xyz where
1. For each i ≥ 0, xy^i z ∈ L
2. |y| > 0
3. |xy| ≤ p

If L violates pumping lemma, then it is not regular

Pumping lemma, continued
1. For each i ≥ 0, xy^i z ∈ L
   There is a loop
2. |y| > 0
   There is a loop of letters (not of ε, which would effectively not be a loop)
3. |xy| ≤ p
   Not allowed more states than pumping length (keep memory finite!)

Proof idea
If |s| ≤ p, trivially true

If |s|>p, consider the states the FSA goes through
• Since there are only p states, |s|>p, one state must be repeated
• Pigeonhole principle: There must be a cycle
Example: How Regular Language 
$L = \{a^n b^m c^n a, n \geq 1, m \geq 1\}$ demonstrates pumping lemma 
Pumping length: $p = 6$ – all words $w$ where $|w| \geq 6$ are pumpable 
Example word: $w = \text{abbaca}$  \hspace{1cm} |w| = 6  
What if $x = a$  \hspace{0.5em} $y = \text{bb}$  \hspace{0.5em} $z = \text{aca}$?
\hspace{1cm} $xy^i z = a(bb)^i aca = \text{abbbbb} \in L$ --- good!  
\hspace{1cm} $xy^0 z = a(bb)^0 aca = \text{aca} \notin L$ --- no good!  
For every word $w$ where $|w| \geq 6$, there must be some $xyz$ division where $xy^i z \in L$ for all $i \geq 0$

Example continued: 
$L = \{a^n b^m c^n a, n \geq 1, m \geq 1\}$ 
Pumping length: $p = 6$ 
Example word: $w = \text{abbaca}$  \hspace{1cm} |w| = 6  
What if $x = \text{ab}$  \hspace{0.5em} $y = \text{b}$  \hspace{0.5em} $z = \text{aca}$?
\hspace{1cm} $xy^i z = \text{ab(b)}^i aca = \text{abbbbb} \in L$ --- good!  
\hspace{1cm} $xy^0 z = \text{ab(b)}^0 aca = \text{aca} \notin L$ --- good! 
So, there DOES exist a division $xyz$ where $xy^i z \in L$ for all $i \geq 0$

Example continued: 
$L = \{a^n b^m c^n a, n \geq 1, m \geq 1\}$ 
Pumping length: $p = 6$ 
One more example word: $w = \text{abacca}$  \hspace{1cm} |w| = 6  
What if $x = \text{aba}$  \hspace{0.5em} $y = \text{c}$  \hspace{0.5em} $z = \text{ca}$?
\hspace{1cm} $xy^i z = \text{abac}^i \text{ca} = \text{abaccca} \in L$ --- good!  
\hspace{1cm} $xy^0 z = \text{abac}^0 \text{ca} = \text{abaca} \in L$ --- good! 
So, there DOES exist a division $xyz$ where $xy^i z \in L$ for all $i \geq 0$
This will work for all $w \in L$, $|w| \geq 6$

Common pumping proof-by-contradiction 

Define a simple word $w$ that is guaranteed to have more than $p$ symbols, and you know the first $p$ symbols 
Show repetition of intermediate $y$ string violates language rules
Prove $B = \{0^n 1^n \}$ is not regular

Proof by contradiction: assume $B$ is regular
thus, any $w \in B$ can be “pumped” if $|w| > p$

First suggestion: $w = 0011$, $x = 0$, $y = 01$, $z = 1$ – counterexample
$xy^2z = 001011 \notin B$

Close! But maybe $|0011| \leq p$, how do we know this will be problem when $|w| > p$

Our solution: Let $w = 0^p 1^p |w| > p$, so must be “pump”-able
$|xy| \leq p$ so, $x = 0^f \ y = 0^g$, $f + g \leq p$ and $g > 0$
When we pump $w$: $xy^2z$, we get $p + g$ 0's followed by $p$ 1's. $xy^2z \notin B$

Contradiction, pumped $w \notin B$

Prove $F = \{ww \mid w = 0^* 1^* \}$ is not regular

Proof by contradiction: assume $F$ is regular
thus, any $v \in F$ can be “pumped” if $|v| > p$

• Our solution: Let $w = 0^p \ 1^p \ |w| > p$ so must be “pump”-able
$|xy| \leq p$ so, $x = 0^f \ y = 0^g$, $f + g \leq p$ and $g > 0$
When we pump $w$: $xy^2z$, we get $p + g$ 0's followed by $10^p 1$. $xy^2z \notin B$

Contradiction, pumped $w \notin F$

$F = \{11, 00, 0101, 1010, 11011101, \ldots\}$

Prove $E = \{1^n \}$ is not regular

Proof by contradiction: assume $E$ is regular
thus, any $w \in E$ can be “pumped” if $|w| > p$

Our solution: Let $w = 1^p \ \ |w| > p$, so must be “pump”-able
$|xy| \leq p$ so $|y| \leq p$
$|xy^2z| \leq p^2 + p$

What’s the length of the next-biggest string after $|w| = p^2$
$|w_{\text{next-biggest}}| = (p + 1)^2 = p^2 + 2p + 1$
Pumping $w$ once gives length at most $p^2 + p < p^2 + 2p + 1$
Thus, $xy^2z \notin E$

Contradiction, pumped $w \notin E$

$F = \{11, 00, 0101, 1010, 11011101, \ldots\}$
Prove $A = \{0^i1^j \mid i > j > 0\}$ is not regular

Proof by contradiction: assume $A$ is regular

thus, any $w \in A$ can be "pumped" if $|w| > p$

Our solution:

Let $w = 0^{p+1}1^p$ \(|w| > p\), so must be "pump"-able

$|xy| \leq p$ so, $x = 0^f y = 0^g$, $f + g \leq p$ and $g > 0$

Let's say $xy = 0^p$. So $z = 01^p$

When we pump $w$: $xy^2z$, we get $0^p01^p \rightarrow 0^p0^g1^p \in A$

Let's try pumping down: $xy^2z$

we get $xz -> 001^p$

Number of 0s: $f + 1$

Number of 1s: $p + g \geq f + 1$

$f + 1 \leq p$

number of 0s < number of 1s

$xy^2z \notin A$

Contradiction, pumped $w \notin A$