Regular Language

Class Discussion

Can you draw a DFA that accepts the language $\{a^k b^k \mid k = 0,1,2,\ldots\}$ over the alphabet $\Sigma=\{a,b\}$?

Limitations of FA

Many languages are non-regular:

- $\{a^n b^n \mid n = 0,1,2,\ldots\}$
- $\{0^{i^2} \mid i \in 0,1,2,\ldots\}$
- $\{0^p \mid p \text{ is a prime number}\}$
- the set of all well-formed parentheses over the alphabet $\Sigma=\{(,\ )\}$
- the set of all palindromes over the alphabet $\Sigma=\{a, b\}$
- ……

Why Impossible?

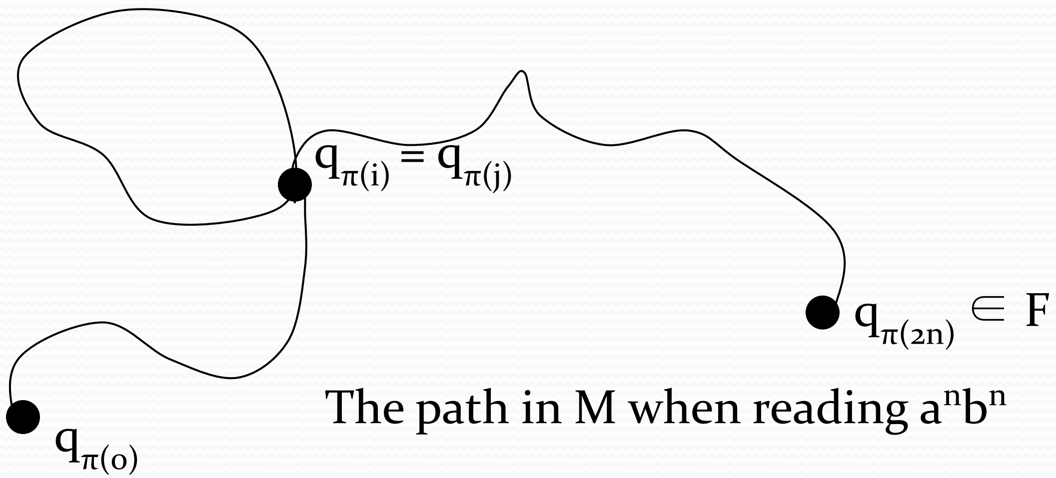
We want to prove that $L = \{a^i b^i \mid i = 0,1,2,\ldots\}$ is non-regular. Prove by contradiction:

- Assume that there is a DFA M that recognizes L. Let n be the total number of states in M.
- Consider the path followed by the input string $a^n b^n$ in M:

$$q_{\pi(0)} \xrightarrow{a} q_{\pi(1)} \xrightarrow{a} \ldots\ldots \xrightarrow{a} q_{\pi(n)} \xrightarrow{b} q_{\pi(n+1)} \xrightarrow{b} \ldots\ldots \xrightarrow{b} q_{\pi(2n)}$$

Why Impossible?

- Since M has only n states, there must be at least one state visited twice in the first n transitions. Let this state be visited at the i^{th} and the j^{th} steps, where $j > i$.



The path in M when reading $a^n b^n$

- By skipping the loop, $a^{n-(j-i)}b^n$ should also be accepted, but this is contradictory since $a^{n-(j-i)}b^n \notin L$

Another Example

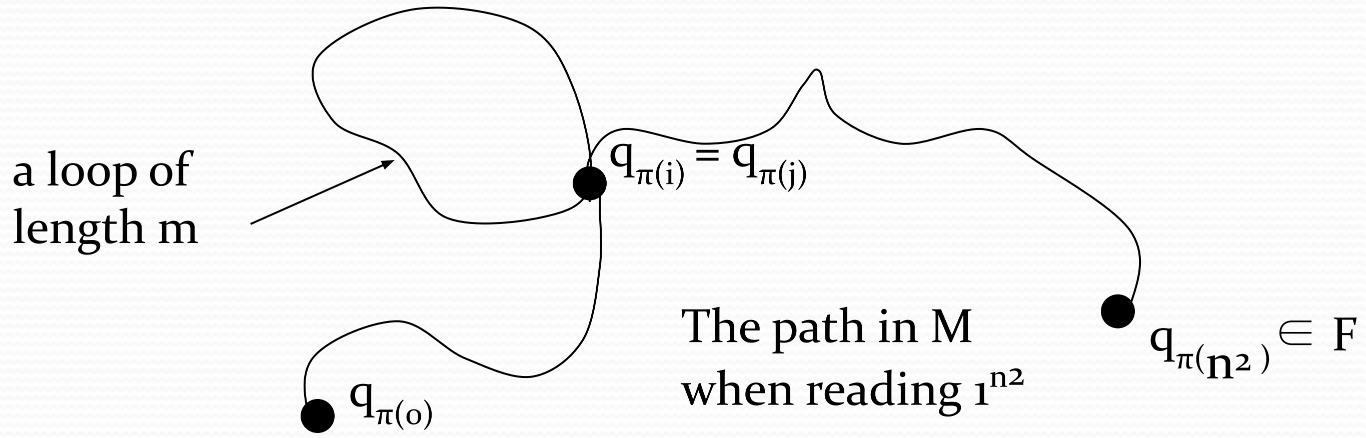
We want to prove that $L = \{1^{k2} \mid k > 0\}$ is non-regular. Prove by contradiction:

- Assume that there is a DFA M that recognizes L. Let n be the total number of states in M.
- M should also accept the string 1^{n2}

$$q_{\pi(0)} \xrightarrow{1} q_{\pi(1)} \xrightarrow{1} q_{\pi(2)} \xrightarrow{1} \ldots\ldots \xrightarrow{1} q_{\pi(n2)}$$

Another Example

- Since $n^2 \geq n$ and M has only n states, there must be at least two equal states from $q_{\pi(0)}$ to $q_{\pi(n^2)}$. Let $q_{\pi(i)} = q_{\pi(j)}$ be the first repeated state where $j-i = m \leq n$.



- By repeating the loop one more time, $1^{(n_2+m)}$ is also accepted by M, which is a contradiction, since (n^2+m) cannot be a square (the next square after n^2 is $(n+1)^2$ but $n^2+m < (n+1)^2$).

Pumping Lemma

- L: regular language
- There exists a constant n s.t. every string $z \in L$ with $|z| \geq n$
- We can find $z = uvw$
  - $v \neq \varepsilon$ i.e. $|v| \geq 1$
  - $|uv| \leq n$
  - For all $i \geq 0$, string uv^iw also in L

Proof of Pumping Lemma

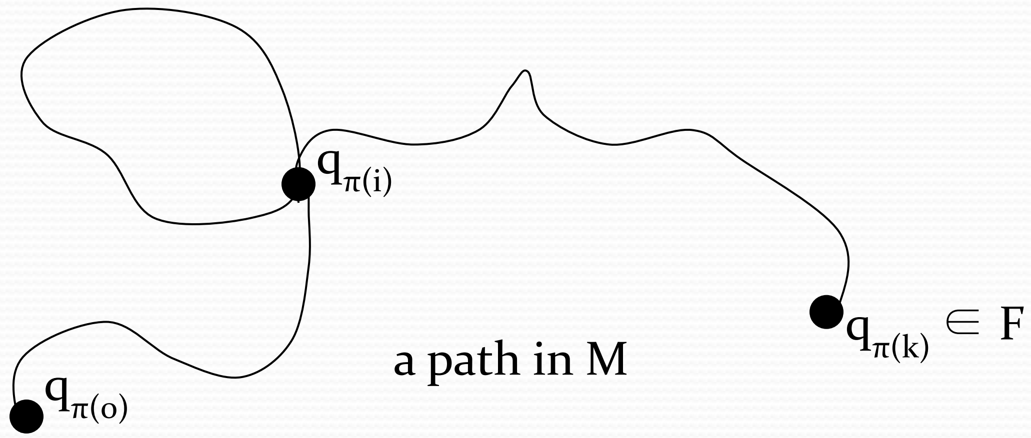
- If L is a regular language, there are DFAs that recognize L. Let n be the number of states in one such DFA (called M) recognizing L.
- If z is a word in L with $|z| = k \geq n$:

$$q_{\pi(0)} \xrightarrow{z(1)} q_{\pi(1)} \xrightarrow{z(2)} q_{\pi(2)} \xrightarrow{z(3)} \ldots\ldots \xrightarrow{z(k)} q_{\pi(k)}$$

where $z(i)$ is the i^{th} symbol of the string z.

- Since $k \geq n$ and M has only n states, there must be at least one repeated states from $q_{\pi(0)}$ to $q_{\pi(k)}$. Let $q_{\pi(i)}$ be the first such repeated state.

Proof of Pumping Lemma



Let u be the string obtained by traversing from $q_{\pi(o)}$ to $q_{\pi(i)}$, and v be the string obtained by traversing the loop once ($|v|\geq 1$). In the traversal from $q_{\pi(o)}$ to $q_{\pi(i)}$ and then through the loop once back to $q_{\pi(i)}$, nothing except $q_{\pi(i)}$ repeats, so $|uv|\leq n$. By traversing the loop 0 or more times, we obtain uv^iw where $i \geq 0$. These strings should all be accepted by M since they can all reach $q_{\pi(k)}$ which is a final state in M.

Using Pumping Lemma

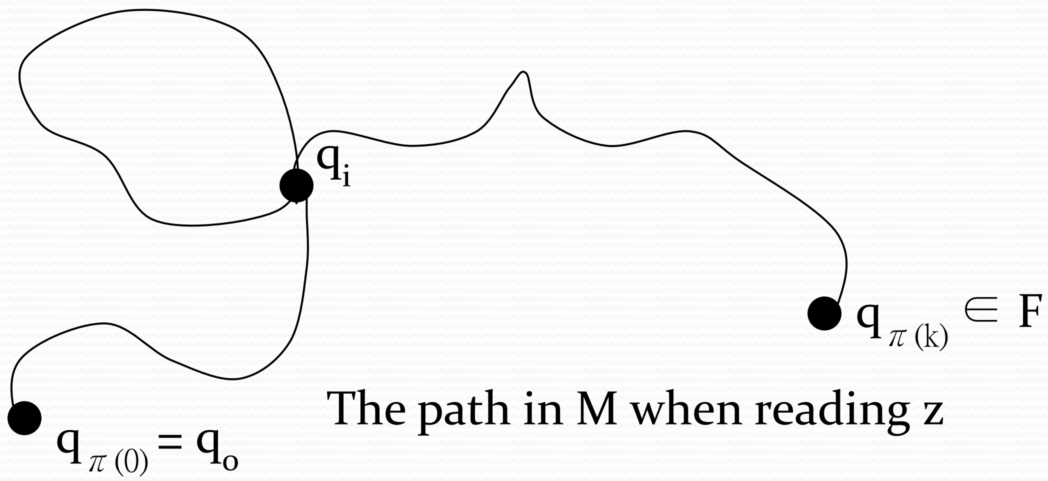
Use the proof of pumping lemma

- Suppose L is regular, then we can construct a DFA M that accepts L
- Suppose M has n states
- Carefully choose a string z where $|z| = k \geq n$

Using Pumping Lemma

Claim that

- when accepting z, there must be at least one state visited twice in the first n transitions (by pigeonhole principle)
- Let the state be q_i ,$0 \le i \le n$



The path in M when reading z

Using Pumping Lemma

Claim that

- By skipping the loop or repeating the loop for certain times to generate new string z’
- M does not accept z’
- So contradiction occurs

Example of using pumping lemma

- Consider a language L:

  The set of all strings over {0,1} with number of '1's is three times number of '0's
- Show that L is non-regular

Example

- Assume L is regular
- Let n be the constant in the lemma
- Let $z = 0^n 1^{3n} \in L$
- $uvw = z = 0^p 0^q 0^{n-p-q} 1^{3n}$, where $|uv| = p+q \leq n$, $|v|=q \geq 1$

Example

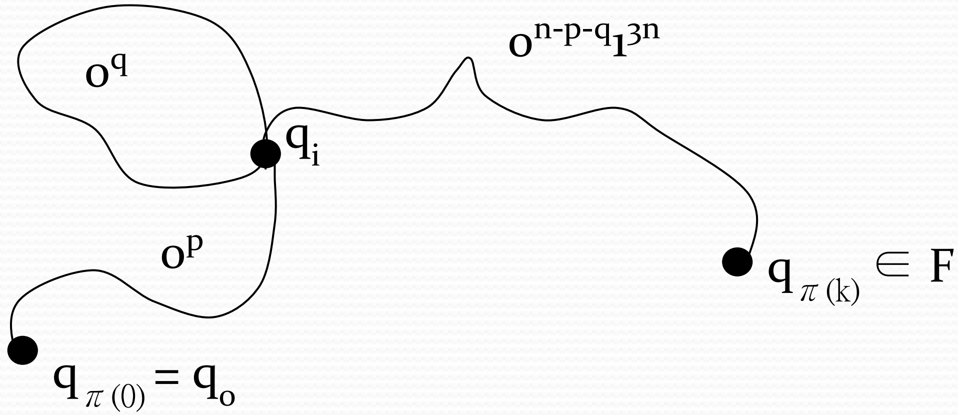
- By Pumping Lemma, uv^iw is also in L for all i
- $0^p0^{qi}0^{n-p-q}1^{3n} \in$ L for all $i \geq 0$
- For $i=0$, $uv^iw = 0^{n-q}1^{3n} \notin L$
- There is a contradiction

Example

- Suppose L is regular, then we can construct a DFA M that accepts L
- Suppose M has n states
- Let $z = 0^n 1^{3n} \in L$
- When accepting z, there must be at least one state visited twice in the first n transitions
- Let the state be q_i ,where $0 \leq i \leq n$

Example

- The path when reading z can be illustrated by the following diagram:



- By skipping the loop, we have $z' = 0^{n-q}1^{3n} \notin L$
- By the property of M, z' should be accepted by M
- So by contradiction, there is no such M

Example

We want to prove that $L = \{1^y \mid y \text{ is a prime}\}$ is non-regular. We can make use of the Pumping Lemma:

- If L is a regular language, it follows the Pumping Lemma. Let n be the constant in the lemma.
- Consider $z=1^p$ where p is a prime and $p \geq n$. (The number of primes are infinite, so such a p exists.)
- According to the lemma, we can write z into uvw where $|uv| \leq n$ and $|v| \geq 1$ and $uv^iw \in L$ for all $i \geq 0$.

Example

- Let $|v| = m$.
- Consider the case when $i = p+1$. Then $|uv^iw| = p + pm = p\ (m+1)$, which is not a prime.
- So it is not true that uv^iw is in L for all $i \geq 0$.
- Therefore L is not a regular language.

Example

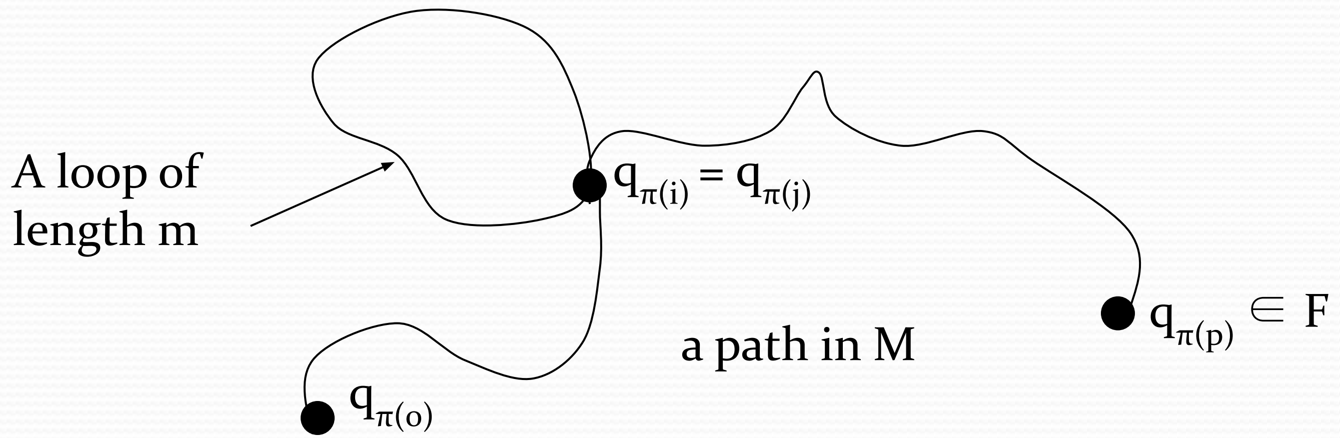
We can also prove that $L = \{1^y \mid y \text{ is a prime}\}$ is non-regular by following the argument directly:

- Assume that there is a DFA M that recognizes L. Let n be the total number of states in M.
- We know that the number of primes are infinite, so there exists a prime $p \geq n$.

$$q_{\pi(0)} \xrightarrow{1} q_{\pi(1)} \xrightarrow{1} \ldots\ldots \xrightarrow{1} q_{\pi(p-1)} \xrightarrow{1} q_{\pi(p)}$$

Example

- Since $p \geq n$ and M has only n states, there must be at least two equal states from $q_{\pi(0)}$ to $q_{\pi(p)}$. Let them be $q_{\pi(i)}$ and $q_{\pi(j)}$ where $j-i = m > 0$.



- By repeating the loop p+1 times, $1^{(p-m) + (p+1)m} = 1^{p(m+1)}$ is also accepted by M, which is a contradiction since $p(m+1)$ is not a prime.

Closure Properties

Regular sets are said to be *closed* under an operation *op* if the application of *op* to a regular set will result in a regular set.

For example, if the union of two regular sets will result in a regular set, regular sets are said to be closed under union.

Closure Properties

Are regular sets closed under:

- Union ?
- Concatenation ?
- Kleene Closure ?

Why?

Closure Properties

Are regular sets closed under complementation?

If A is a regular set over Σ, is A' = Σ^*-A regular? Why?

If A is regular, there exists a DFA M recognizing A. Given M, we can construct a DFA M' for A' by copying M to M' except that all final states in M are changed to non-final, and all non-final states to final.

Closure Properties

Are regular sets closed under intersection?

If A and B are regular sets, is $C = A \cap B$ regular? Why?

$$C = A \cap B = \overline{\overline{A} \cup \overline{B}}$$

Since regular sets are closed under union and complementation, they are also closed under intersection.

Closure Properties (Summary)

Regular sets are closed under:

- Union
- Concatenation
- Kleene Star
- Complementation
- Intersection
- Reversal

Thank you