# Lecture 4
# Properties of Regular Languages

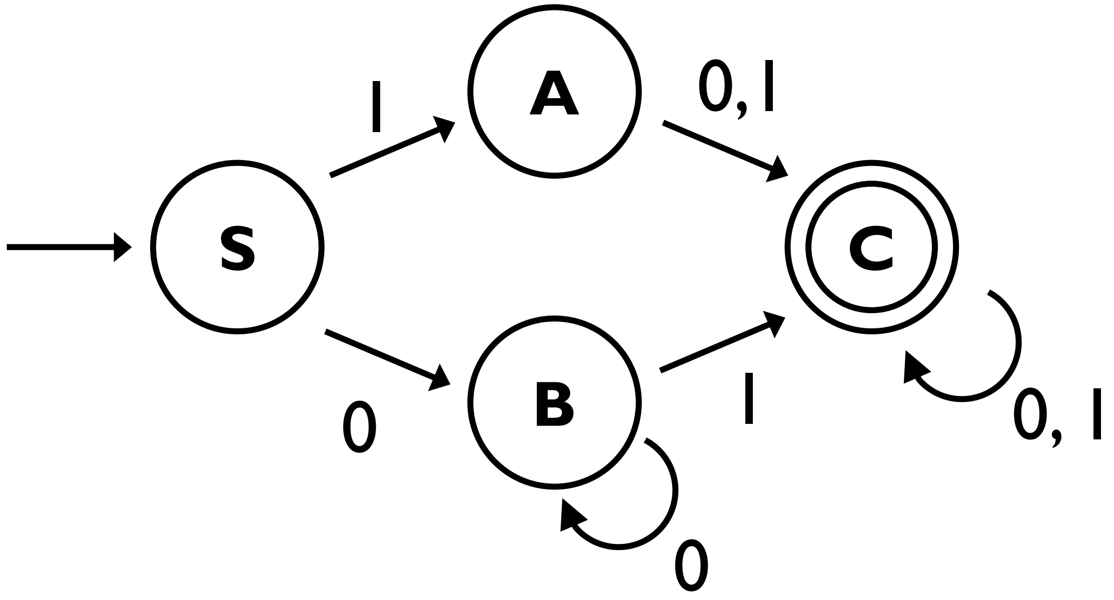
COSE215: Theory of Computation

Seunghoon Woo

Fall 2023

# Practice for Lecture 3

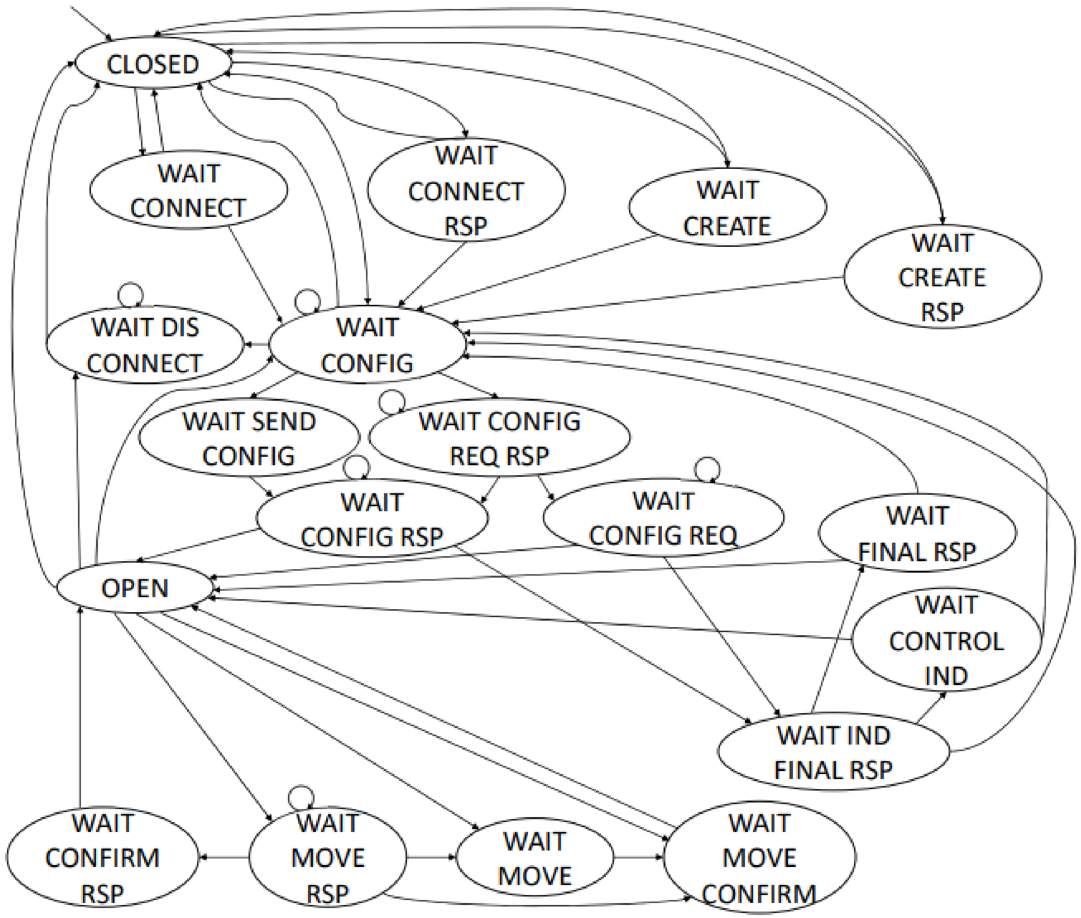
- **Find a <u>DFA</u> for the given regular expression, and then extract regular grammar**
  - $(1(0+1)+00^*1)(0+1)^*$




- $S \rightarrow 1A \mid 0B$
- $A \rightarrow 0C \mid 1C \mid 0 \mid 1$
- $B \rightarrow 0B \mid 1C \mid 1$
- $C \rightarrow 0C \mid 1C \mid 0 \mid 1$

# Example State Machine

- **Bluetooth L2CAP**



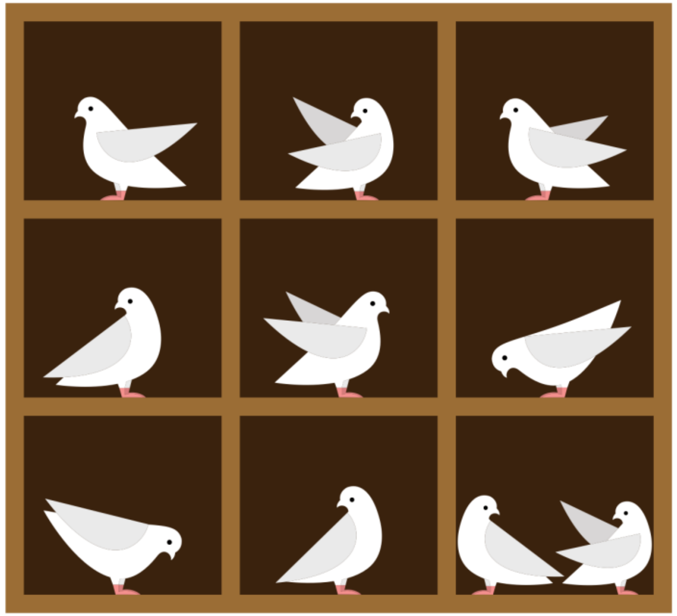
# Contents

# Identifying non-regular languages

- **Non-regular languages**
  - Non-regular languages cannot be recognized by finite automata
  - Claim: $L = \{0^n 1^n \mid n \geq 0\}$ is not regular
    - L is a regular language if we can construct a DFA for L
    - However, DFA has limited temporary storage, so it cannot remember $n$

# Identifying non-regular languages

- **Pigeonhole principle**
  - If $n$ pigeons are placed in $m$ pigeon holes,
    - Then one hole will contain at least $n/m$ pigeons
  - If we put $n$ objects into $m$ boxes and $n > m$, then at least one box must have more than one object in it

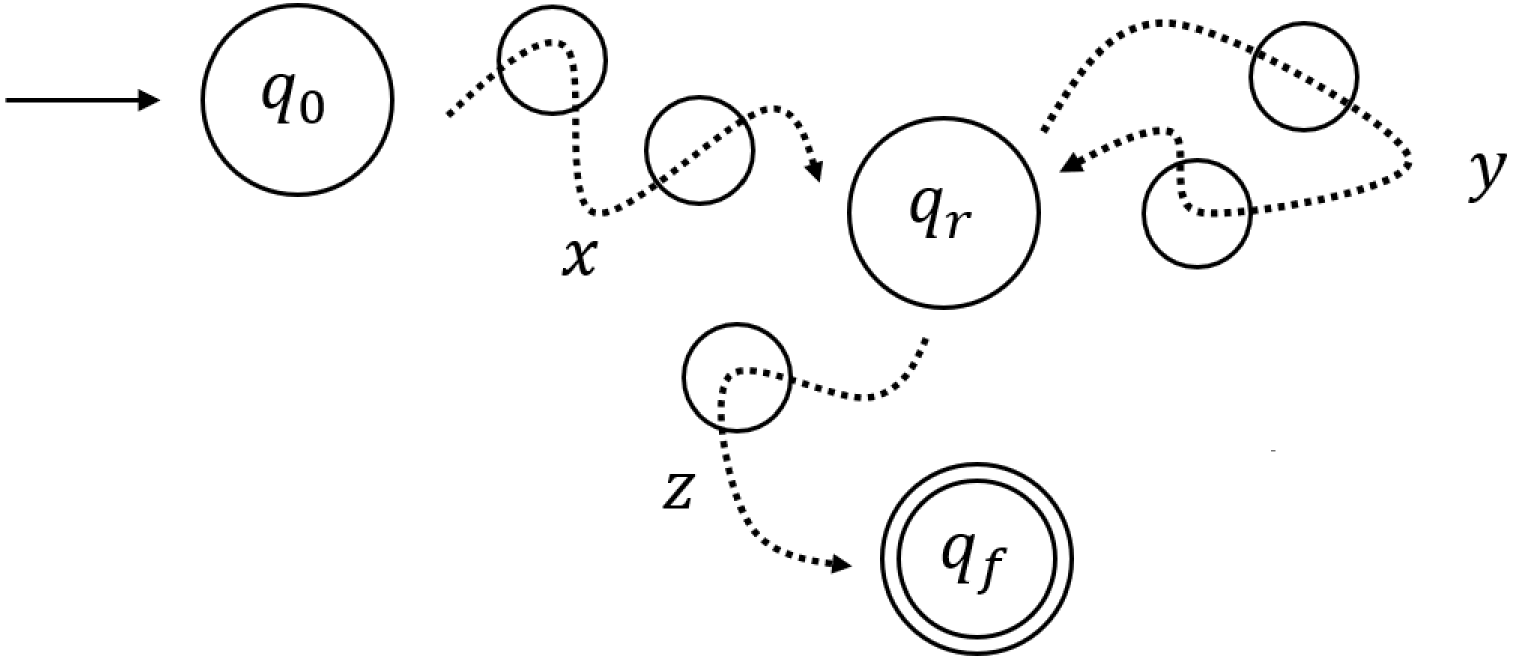


# Identifying non-regular languages

- **Basic idea to identify non-regular languages**
  - Consider a finite automaton with $n$ state
  - Given an input string with length $m$ where $m > n$
  - Then, one or more states will inevitably be visited multiple times

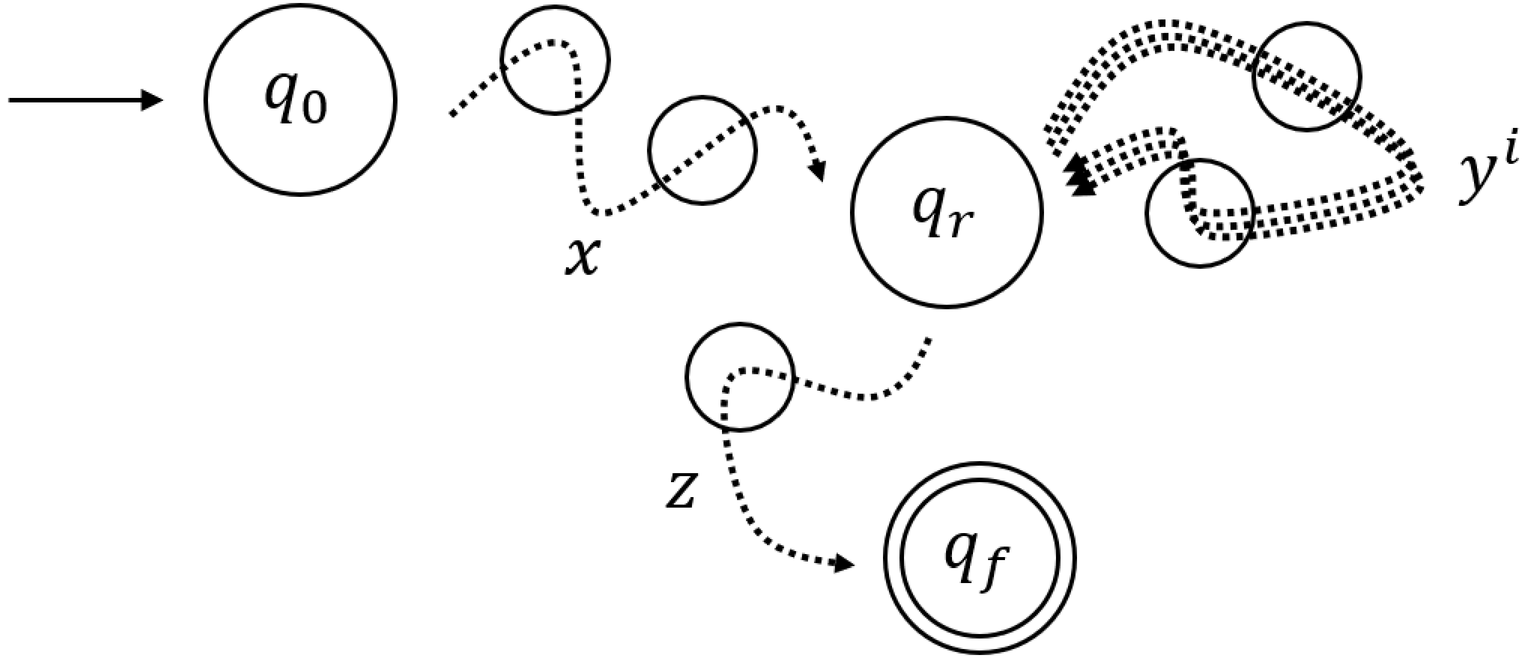
# Identifying non-regular languages

- **Pumping lemma**



# Identifying non-regular languages

- **Pumping lemma**



# Identifying non-regular languages

- **Pumping lemma**
  - Let $L$ be a regular language
  - There exists a positive integer $m$ such that any $w \in L$, if $|w| \geq m$, there exists $w = xyz$ such that
    - ❖ $|xy| \leq m$
    - ❖ $|y| \geq 1$
    - ❖ For all $i \geq 0,\ xy^i z \in L$

# Identifying non-regular languages

- **Pumping lemma**
  - Let $L$ be a regular language
  - There exists a positive integer $m$ such that any $w \in L$, if $|w| \geq m$, there exists $w = xyz$ such that
    - ❖ $|xy| \leq m$
    - ❖ $|y| \geq 1$
    - ❖ For all $i \geq 0,\ xy^i z \in L$
  - A string of sufficiently large length ($|w| \geq m$) can be represented in the form of $xyz$, and $xy^i z$, which is pumped "$y$" $i$ times, can also always belong to this language

# Identifying non-regular languages

- **Proof of the pumping lemma**
  - If $L$ is regular, there exists a DFA $M = (Q, \Sigma, \delta, q_0, F)$ that recognizes it

# Identifying non-regular languages

- **Proof of the pumping lemma**
  - If $L$ is regular, there exists a DFA $M = (Q, \Sigma, \delta, q_0, F)$ that recognizes it
  - Let $|Q| = n + 1$ ($M$ has states labeled $q_0, q_1, \ldots, q_n$)

# Identifying non-regular languages

- **Proof of the pumping lemma**
  - If $L$ is regular, there exists a DFA $M = (Q, \Sigma, \delta, q_0, F)$ that recognizes it
  - Let $|Q| = n + 1$ ($M$ has states labeled $q_0, q_1, \dots, q_n$)
  - Take a string $w \in L$ such that $|w| \geq m = n + 1$

# Identifying non-regular languages

- **Proof of the pumping lemma**
  - If $L$ is regular, there exists a DFA $M = (Q, \Sigma, \delta, q_0, F)$ that recognizes it
  - Let $|Q| = n + 1$ ($M$ has states labeled $q_0, q_1, \ldots, q_n$)
  - Take a string $w \in L$ such that $|w| \geq m = n + 1$
  - Consider the set of states $M$ goes through as it processes $w$: $q_0, q_i, q_j, \ldots, q_f$

# Identifying non-regular languages

- **Proof of the pumping lemma**
  - If $L$ is regular, there exists a DFA $M = (Q, \Sigma, \delta, q_0, F)$ that recognizes it
  - Let $|Q| = n + 1$ ($M$ has states labeled $q_0, q_1, \dots, q_n$)
  - Take a string $w \in L$ such that $|w| \geq m = n + 1$
  - Consider the set of states $M$ goes through as it processes $w$: $q_0, q_i, q_j, \dots, q_f$
  - Since this sequence has $|w| + 1$ states, at least one state must be repeated
    - ❖ Such a repetition must start no later than the $n^{th}$ move
    - ❖ E.g., $q_0, q_i, q_j, \dots, q_r, \dots, q_r, \dots, q_f$

# Identifying non-regular languages

- **Proof of the pumping lemma**
  - If $L$ is regular, there exists a DFA $M = (Q, \Sigma, \delta, q_0, F)$ that recognizes it
  - Let $|Q| = n + 1$ ($M$ has states labeled $q_0, q_1, \ldots, q_n$)
  - Take a string $w \in L$ such that $|w| \geq m = n + 1$
  - Consider the set of states $M$ goes through as it processes $w$: $q_0, q_i, q_j, \ldots, q_f$
  - Since this sequence has $|w| + 1$ states, at least one state must be repeated
    - ❖ Such a repetition must start no later than the $n^{th}$ move
    - ❖ E.g., $q_0, q_i, q_j, \ldots, q_r, \ldots, q_r, \ldots, q_f$
  - This indicates that there must be substrings $x, y, z$ of $w$ such that
    - ❖ $\delta^*(q_0, x) = q_r, \quad \delta^*(q_r, y) = q_r, \quad \delta^*(q_r, z) = q_f$ (with $|xy| \leq n + 1$ and $|y| \geq 1$)

# Identifying non-regular languages

- **Proof of the pumping lemma**
  - If $L$ is regular, there exists a DFA $M = (Q, \Sigma, \delta, q_0, F)$ that recognizes it
  - Let $|Q| = n + 1$ ($M$ has states labeled $q_0, q_1, \dots, q_n$)
  - Take a string $w \in L$ such that $|w| \geq m = n + 1$
  - Consider the set of states $M$ goes through as it processes $w$: $q_0, q_i, q_j, \dots, q_f$
  - Since this sequence has $|w| + 1$ states, at least one state must be repeated
    - ❖ Such a repetition must start no later than the $n^{th}$ move
    - ❖ E.g., $q_0, q_i, q_j, \dots, q_r, \dots, q_r, \dots, q_f$
  - This indicates that there must be substrings $x, y, z$ of $w$ such that
    - ❖ $\delta^*(q_0, x) = q_r$, $\delta^*(q_r, y) = q_r$, $\delta^*(q_r, z) = q_f$ (with $|xy| \leq n + 1$ and $|y| \geq 1$)
  - From this, the followings can be satisfied
    - ❖ $\delta^*(q_0, xz) = q_f$, $\delta^*(q_0, xy^2z) = q_f$, $\delta^*(q_0, xy^3z) = q_f$, $\boldsymbol{\delta^*(q_0, xy^iz) = q_f}$

# Identifying non-regular languages

- **Pumping lemma**
  - Show that $L = \{a^n b^n : n \geq 0\}$ is not regular

# Identifying non-regular languages

- **Pumping lemma**
  - Show that $L = \{a^n b^n : n \geq 0\}$ is not regular
    - ❖ Assume that $L$ is regular, so that the pumping lemma must hold
    - ❖ Let $m = n$
    - ❖ Because $|xy| \leq m$, the substring $y$ must consist entirely of $a$'s (suppose $|y| = k$)
    - ❖ When $i = 0$, then $w_0 = a^{m-k} b^m$
    - ❖ $a^{m-k} b^m$ clearly not in $L$ => ***L* is not regular**

# Identifying non-regular languages

- **Pumping lemma**
  - Let $\Sigma = \{a, b\}$. Show that $L = \{ww^R : w \in \Sigma^*\}$ is not regular
    - $w^R$ = string reverse
      - E.g., $w = abb$, then $w^R = bba$

# Identifying non-regular languages

- **Pumping lemma**
  - Let $\Sigma = \{a, b\}$. Show that $L = \{ww^R : w \in \Sigma^*\}$ is not regular
    - ❖ Assume that $L$ is regular, so that the pumping lemma must hold
    - ❖ Consider a positive integer $m$ and let $w' = ww^R$ be $a^m b^m b^m a^m$
    - ❖ Because $|xy| \leq m$, the substring $y$ must consist entirely of $a$'s (suppose $|y| = k$)
    - ❖ When $i = 0$, then $w' = a^{m-k} b^m b^m a^m$ $(\notin L)$
    - ❖ ***L* is not regular**

$$\overbrace{aaa \ldots aaa}^{a^m}\overbrace{bbb \ldots bbb}^{b^m}\overbrace{bbb \ldots bbb}^{b^m}\overbrace{aaa \ldots aaa}^{a^m}$$

# Identifying non-regular languages

- **Pumping lemma**
  - Pumping Lemma is violated => not a regular language
  - Pumping Lemma is not violated => do not know if it is regular language or not

# Identifying non-regular languages

- **Pumping lemma**
  - Let $\Sigma = \{a, b\}$. Show that $L = \{w \in \Sigma^*: n_a(w) < n_b(w)\}$ is not regular

# Identifying non-regular languages

- **Pumping lemma**
  - Let $\Sigma = \{a, b\}$. Show that $L = \{w \in \Sigma^*: n_a(w) < n_b(w)\}$ is not regular
    - ❖ Assume that $L$ is regular, so that the pumping lemma must hold
    - ❖ Consider a positive integer $m$ and $w = a^m b^{m+1}$
    - ❖ Because $|xy| \leq m$, the substring $y$ must consist entirely of $a$'s (suppose $|y| = k \geq 1$)
    - ❖ When $i = 2$, then $w = a^{m+k} b^{m+1}$ $(\notin L)$
    - ❖ **$L$ is not regular**

# Identifying non-regular languages

- **Pumping lemma**
  - Let $\Sigma = \{a, b\}$. Show that $L = \{ww | w \in \Sigma^*\}$ is not regular

# Identifying non-regular languages

- **Pumping lemma**
  - Let $\Sigma = \{a, b\}$. Show that $L = \{ww | w \in \Sigma^*\}$ is not regular
    - ❖ Assume that $L$ is regular, so that the pumping lemma must hold
    - ❖ Consider a positive integer $m$ and $w = a^m b a^m b$
    - ❖ Because $|xy| \leq m$, the substring $y$ must consist entirely of $a$'s (suppose $|y| = k \geq 1$)
    - ❖ When $i = 0$, then $w = a^{m-k} b a^m b$ $(\notin L)$
    - ❖ **$L$ is not regular**

# Identifying non-regular languages

- **Pumping lemma: practice**
  - Show that $L = \{a^i b^j c^k \mid i + j \leq k\}$ is not regular

# Identifying non-regular languages

- **Pumping lemma: practice**
  - Show that $L = \{0^{n^2} \mid n \geq 0\}$ is not regular

# Next Lecture

- **Context-Free Languages**