

FORMAL LANGUAGES AND AUTOMATA THEORY

Module 1

Rijin IK

Assistant Professor
Department of Computer Science and Engineering
Vimal Jyothi Engineering College
Chemperi

May 12, 2024

- 1 Course Outcomes
- 2 Introduction to formal language theory
- 3 Finite Automata
- 4 Deterministic Finite Automata (DFA)
- 5 Nondeterministic finite automaton (NFA)
 - NFA with ϵ transitions (ϵ -NFA)
 - Equivalence of NFA with and without epsilon transitions
- 6 Regular Grammar

After the completion of the course the student will be able to

- ➊ Classify a given formal language into Regular, Context-Free, Context Sensitive, Recursive or Recursively Enumerable. [Cognitive knowledge level: Understand]
- ➋ Explain a formal representation of a given regular language as a finite state automaton, regular grammar, regular expression and Myhill-Nerode relation. [Cognitive knowledge level: Understand]
- ➌ Design a Pushdown Automaton and a Context-Free Grammar for a given context-free language. [Cognitive knowledge level : Apply]
- ➍ Design Turing machines as language acceptors or transducers. [Cognitive knowledge level: Apply]
- ➎ Explain the notion of decidability. [Cognitive knowledge level: Understand]

Symbols

- Symbols are indivisible objects or entity that cannot be defined
- Symbols are the atoms of the world of language
- A symbol is any single objects such as $\Delta, a, 0, 1, \#$ etc

Alphabets

An alphabet is a finite, nonempty set of symbols. It is denoted by Σ

Example:

- $\Sigma = \{0, 1\}$ is binary alphabet consisting of the symbols 0 and 1.
- $\Sigma = \{a, b, c \dots z\}$ is lowercase English alphabet.

String

A string (or word) is a finite sequence of symbols chosen from some alphabet.

Example:

- If $\Sigma = \{a, b, c\}$ then **abcbb** is a string formed from that alphabet.

Length of a string

- The length of a string w , denoted $|w|$, is the number of symbols composing the string.
- **Example:**
 - The string abcb has length 4.

Empty string (ϵ):

- The empty string denoted by ϵ , is the string consisting of zero symbols. Thus $|\epsilon| = 0$.

Operations on Strings

- 1 Concatenation of strings
- 2 String Reversal
- 3 Substring

Concatenation of strings

- The concatenation of two strings is the string formed by writing the first, followed by the second, with no intervening space.
- Concatenation of strings is denoted by \cdot (dot)
- That is, if w and x are strings, then wx is the concatenation of these two strings.
- **Example:**
 - The concatenation of dog and house is doghouse.
 - Let $x=0100101$ and $y=1111$ then $x \cdot y=01001011111$

String Reversal

- Reversing a string means writing the string backwards.
- It is denoted by w^R
- **Example:**
 - Reverse of the string **abcd** is **dcba**.

Note: If $w = w^R$, then that string is called palindrome.

Substring

- A substring is a part of a string.
- **Example:**
 - If $abcd$ is string then possible substrings are $\epsilon, a, b, c, d, ab, bc, cd, abc, bcd$ are proper substrings for the given string
- A **prefix** of a string is any number of leading symbols of that string.
- A **suffix** of a string is any number of trailing symbols.
- **Example:**
 - String abc has prefixes ϵ, a, ab , and abc ;
 - its suffixes are ϵ, c, bc , and abc .
- A prefix or suffix of a string, other than the string itself, is called a proper prefix or suffix.

Language

- A (formal) language is a set of strings of symbols from someone alphabet.
- It is denoted by L . We denote this language by Σ^* .
 - The empty set, ϕ , and the set consisting of the empty string $\{\epsilon\}$ are languages.

Example:

- If $\Sigma = \{a\}$, then $\Sigma^* = \{\epsilon, a, aa, aaa, \dots\}$.
- If $\Sigma = \{0, 1\}$, then $\Sigma^* = \{\epsilon, 0, 1, 00, 01, 10, 11, 000, \dots\}$.

Operations on languages

- 1 Union
- 2 Intersection
- 3 Complementation
- 4 Concatenation
- 5 Reversal
- 6 Kleene Closure
- 7 Positive Closure

Union

- If L_1 and L_2 are two languages over an alphabet Σ . Then the union of L_1 and L_2 is denoted by $L_1 \cup L_2$
- **Example**
 - $L_1 = \{0, 01, 011\}$ and $L_2 = \{001\}$, then $L_1 \cup L_2 = \{0, 01, 011, 001\}$

Intersection

- If L_1 and L_2 are two languages over an alphabet Σ . Then the intersection of L_1 and L_2 is denoted by $L_1 \cap L_2$
- **Example**
 - $L_1 = \{0, 01, 011\}$ and $L_2 = \{01\}$, then $L_1 \cap L_2 = \{01\}$

Complementation

- If L is a language over an alphabet Σ . then the complement of L denoted by L' is the language consisting of those strings that are not in L over the alphabet.
- **Example**
 - If $\Sigma = \{a, b\}$, and $L = \{a, b, aa\}$. then
 - $L' = \Sigma^* - L = \{\epsilon, , a, b, aa, bb, ab...\} - \{a, b, aa\} = \{\epsilon, bb, ab, ba...\}$.

Concatenation

- Concatenation of two languages L_1 and L_2 is the language $L_1 \cdot L_2$, each element of which is a string formed by combining one string of L_1 with another string of L_2 .
- **Example**
 - $L_1 = \{bc, bcc, cc\}$ and $L_2 = \{cc, ccc\}$, then
 - $L_1 \cdot L_2 = \{bccc, bccccc, bcccccc, cccc, cccccc\}$

Reversal

- If L is language, then L^R is obtained by reversing the corresponding string in L . This operation is similar to the reversal of a string.

$$L^R = \{w^R \mid w \in L\}$$

- **Example**

- If $L = \{0, 011, 0111\}$, then $L^R = \{0, 110, 1110\}$

Kleene Closure

- The Kleene closure (or just closure) of L , denoted L^* , is the set

$$L^* = \bigcup_{i=0}^{\infty} L^i$$

$$L^* = L^0 \cup L^1 \cup L^2 \cup \dots$$

- **Example**

- Let $L = \{10, 1\}$
- $L^* = L^0 \cup L^1 \cup L^2 \dots = \{\epsilon, 1, 10, 11, 111, 1111, \dots\}$

Positive Closure

- The positive closure of L , denoted L^+ , is the set

$$L^+ = \bigcup_{i=1}^{\infty} L^i$$

$$L^+ = L^1 \cup L^2 \cup L^3 \cup \dots$$

- **Example**

- Let $L = \{10, 1\}$
- $L^+ = L^1 \cup L^2 \cup L^3 \dots = \{1, 10, 11, 111, 1111, \dots\}$

Finite Automata or Finite Automaton

- Mathematical model of finite state system

Finite state system

- A system containing only finite number of states and transition among them is called finite state system(FSS)

State

- State of a system is an instantaneous description of that system.
- Which gives all relevant information necessary to determine how the system can evolve from that point on

Transition

- Change of states that occur spontaneously or in response to inputs to the states

Finite Automata

Finite Automata Model

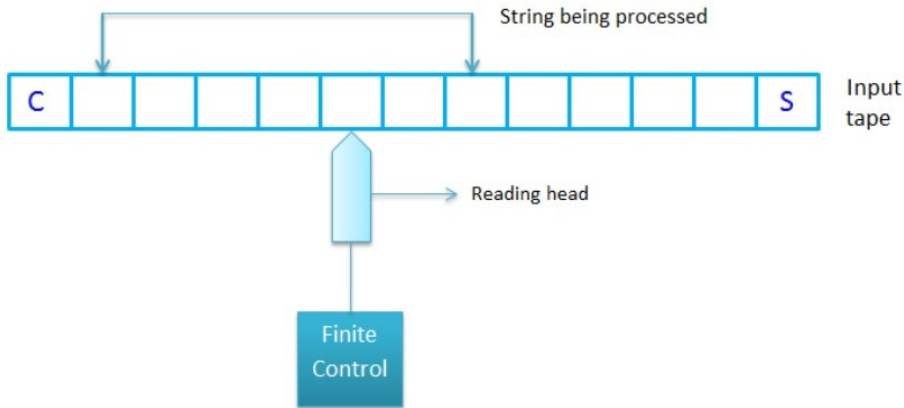


Figure: Block diagram of a finite automaton

Input tape:

- The input tape is divided into squares, each square containing a single symbol from the input alphabet Σ
- The end squares of the tape contain the endmarker
- The left-to-right sequence of symbols between the two endmarkers is the input string to be processed.

Reading head:

- The head examines only one square at a time and can move one square either to the left or to the right.
- For further analysis, we restrict the movement of the R-head only to the right side.
- Initially the head is placed at the leftmost square of the tape.

Finite control:

- The input to the finite control will usually be the symbol under the R-head, say a , and the present state of the machine, say q , to give the following outputs
 - A motion of the R-head along the tape to the next square
 - The next state of the finite state machine given by $\delta(q, a)$.

Acceptance of String by a Finite Automaton

- The FA accepts a string x if the sequence of transitions corresponding to the symbols of x leads from the start state to an accepting state(final state) and the entire string has to be consumed

Finite State Automata-Formal definition

- A finite automaton (FA) is a 5-tuple or quintuple

$$M = (Q, \Sigma, \delta, q_0, F)$$

- Q is a finite set of states
- Σ is a finite input alphabet
- δ is the transition function
- $q_0 \in Q$ is the initial state
- $F \subseteq Q$ is the set of final states

Finite Automata

Transition Diagram:

- A transition diagram is a directed graph associated with an FA in which the vertices of the graph correspond to the states of the FA.
- If there is a transition from state q to state p on input a , then there is an arc labelled a from state q to state p in the transition diagram.
 - State is denoted by \bigcirc
 - Transition is denoted by \rightarrow
 - Initial state is denoted by $\rightarrow \bigcirc$
 - Final state is denoted by $\bigcirc \bigcirc$

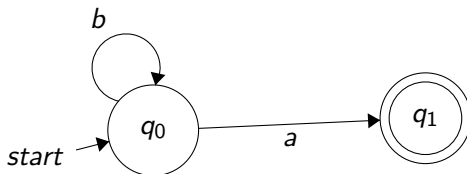


Figure: Transition diagram

Transition Table:

- A tabular representation in which rows correspond to states, columns correspond to inputs and entries correspond to next states.
- The start state is denoted by an arrow with no source.
- The accept state is denoted by a star.

State\input	a	b
$\rightarrow q0$	q1	q0
$*q1$	-	-

Table: Transition Table

There are two types of finite automata

- 1 Deterministic Finite Automata (DFA)
- 2 Non-Deterministic Finite Automata (NFA)

Deterministic Finite Automata (DFA)

Deterministic Finite Automata (DFA)

- It is a finite state machine that reads input symbols from an alphabet one by one and transitions between states based on a fixed set of rules.
- The transitions are deterministic, meaning that for a given input symbol and current state, there is only one next state.
- DFAs are capable of recognizing and accepting certain types of languages, known as regular languages.

Deterministic Finite Automata (DFA)

Deterministic Finite Automata (DFA)

- Formally, a deterministic finite automaton can be represented by a 5-tuple

$$M = (Q, \Sigma, \delta, q_0, F)$$

- Q is a finite set of states
- Σ is a finite input alphabet
- δ is the transition function mapping $Q \times \Sigma$ to Q i.e., $\delta(q, a)$ is a state for each state q and input symbol a .
- $q_0 \in Q$ is the initial state
- $F \subseteq Q$ is the set of final states. It is assumed here that there may be more than one final state.

Deterministic Finite Automata (DFA)

Steps to design a DFA

- 1 Understand the language for which the DFA has to be designed and write the language for the set of strings starting with minimum string that are accepted by FA.
- 2 Draw transition diagram for the minimum length string.
- 3 Obtain the possible transitions to be made for each state on each input symbol.
- 4 Draw the transition table.
- 5 Test DFA with few strings that are accepted and few strings that are rejected by the given language.
- 6 Represent DFA with tuples.

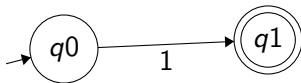
Deterministic Finite Automata (DFA)

Q:Design DFA that accepts all strings which starts with '1' over the alphabet $\{0,1\}$

- 1 Understand the language for which the DFA has to be designed and write the language for the set of strings starting with minimum string that are accepted by FA.

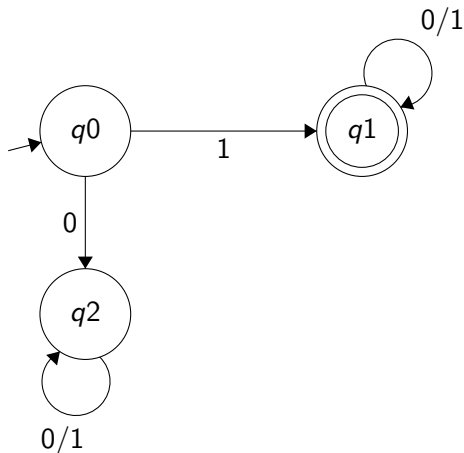
$$L = \{1, 10, 11, 100, 110, 101, 111, \dots\}$$

- 2 Draw transition diagram for the minimum length string.



Deterministic Finite Automata (DFA)

- 3 Obtain the possible transitions to be made for each state on each input symbol.



Deterministic Finite Automata (DFA)

4 Draw the transition table.

	0	1
$\rightarrow q_0$	q_2	q_1
$*q_1$	q_1	q_1
q_2	q_2	q_2

- Test DFA with few strings that are accepted and few strings that are rejected by the given language.
- Represent DFA with tuples.

Deterministic Finite Automata (DFA)

- 5 Test DFA with few strings that are accepted and few strings that are rejected by the given language.

- **Case i)** Let $w=1010 \in L$

$$\begin{aligned}\delta(q_0, 1010) &= \delta(q_1, 010) \\ &= \delta(q_1, 10) \\ &= \delta(q_1, 0) \\ &= q_1\end{aligned}$$

q_1 is final state and the entire string has been consumed i.e., given string is accepted by DFA.

- **Case ii)** Let $w=0001 \notin L$

$$\begin{aligned}\delta(q_0, 0001) &= \delta(q_2, 001) \\ &= \delta(q_2, 01) \\ &= \delta(q_2, 1) \\ &= q_2\end{aligned}$$

q_2 is not final state and the entire string has been consumed i.e., given string is rejected by DFA

- Represent DFA with tuples.

Deterministic Finite Automata (DFA)

6 Represent DFA with tuples.

$$DFA, M = (Q, \Sigma, \delta, q_0, F)$$

$$\text{where } Q = \{q_0, q_1, q_2\}$$

$$\Sigma = \{0, 1\}$$

$$\delta : \delta(q_0, 0) = q_2$$

$$\delta(q_0, 1) = q_1$$

$$\delta(q_1, 0) = q_1$$

$$\delta(q_1, 1) = q_1$$

$$\delta(q_2, 0) = q_2$$

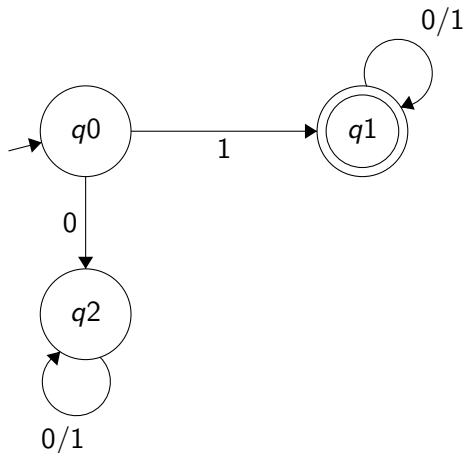
$$\delta(q_2, 1) = q_2$$

$$q_0\text{--initial state} = q_0$$

$$F\text{--finalstate} = \{q_1\}$$

Deterministic Finite Automata (DFA)

Dead State: A rejecting state that is essentially a dead end. Once the machine enters a dead state, there is no way for it to reach an accepting state, so we already know that the string is going to be rejected.



state q_2 is a dead state

Deterministic Finite Automata (DFA)

Language accepted by a DFA

- The language accepted by a DFA M is the set of all string accepted by M .
- It is denoted by $L(M)$.

$$L(M) = \{w \in \Sigma^* / M \text{ accepts } w\}$$

Deterministic Finite Automata (DFA)

Extending the transition function to string (Extended transition function - $\hat{\delta}$)

- It is a function that takes a state q and a string w and return a state p .
- **Example:**
 - consider a string $w = 101$
 - generally we write the transition function as

$$\delta(\delta(\delta(q_0, 1), 0), 1)$$

- Instead of that we can write.

$$\hat{\delta}(q_0, 101)$$

- generally $\hat{\delta}(q_0, w) = p$
- if $p \in F$ then w is accepted

Deterministic Finite Automata (DFA)

Properties of $\hat{\delta}$

① $\hat{\delta}(q_0, \epsilon) = q_0$

② $\hat{\delta}(q_0, wa) = \delta(\hat{\delta}(q_0, w), a)$

$$\delta : Q \times \Sigma \rightarrow Q$$

$$\hat{\delta} : Q \times \Sigma^* \rightarrow Q$$

③ $\hat{\delta} = \delta$

$$\begin{aligned}\hat{\delta}(q, a) &= \hat{\delta}(q, \epsilon a) \\ &= \delta(\hat{\delta}(q, \epsilon), a) \\ &= \delta(q, a) \\ \hat{\delta} &= \delta\end{aligned}$$

Deterministic Finite Automata (DFA)

Language accepted by a DFA

- Set of all string accepted by the finite automata.
- It is denoted by $L(M)$.

$$L(M) = \{w / \hat{\delta}(q_0, w) = p \text{ and } p \in f\}$$

Nondeterministic finite automaton (NFA)

Nondeterminism means a choice of moves for an automaton. Rather than prescribing a unique move in each situation

The transition function that a state & input symbol as arguments , but returns a set of zero , one or more states.

Nondeterministic finite automaton (NFA)

Nondeterministic finite automaton (NFA)

- Formally, a Nondeterministic finite automaton can be represented by a 5-tuple

$$M = (Q, \Sigma, \delta, q_0, f)$$

- Q is a finite set of states
- Σ is a finite input alphabet
- δ is the transition function mapping $Q \times \Sigma$ into 2^Q , (which is the power set of Q , the set of all subsets of Q ;))
- $q_0 \in Q$ is the initial state
- $F \subseteq Q$ is the set of final states. (It is assumed here that there may be more than one final state.)

Nondeterministic finite automaton (NFA)

Example:

NFA that accepts strings which contains either two consecutive 0's or two consecutive 1's.

$$L = \{00, 11, 100, 001, 110, 011, 111, 000, 0100, 1011, \dots\}$$

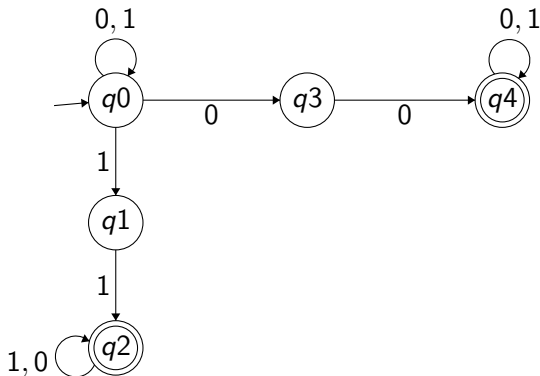


Figure: Transition Diagram

Nondeterministic finite automaton (NFA)

Example Cont..

State\input	0	1
$\rightarrow q_0$	$\{q_0, q_3\}$	$\{q_0, q_1\}$
q_1	ϕ	$\{q_2\}$
$*q_2$	$\{q_2\}$	$\{q_2\}$
q_3	$\{q_4\}$	ϕ
$*q_4$	$\{q_4\}$	$\{q_4\}$

Table: Transition Table

Nondeterministic finite automaton (NFA)

Example Cont..

Let the input, $w = 01001 \in L$

$$\begin{aligned}\delta(q_0, 0) &= \{q_0, q_3\} \\ \hat{\delta}(q_0, 01) &= \delta(\delta(q_0, 0), 1) \\ &= \delta(\{q_0, q_3\}, 1) \\ &= \delta(q_0, 1) \cup \delta(q_3, 1) \\ &= \{q_0, q_1\} \cup \phi \\ &= \{q_0, q_1\} \\ \hat{\delta}(q_0, 010) &= \delta(\hat{\delta}(q_0, 01), 0) \\ &= \delta(\{q_0, q_1\}, 0) \\ &= \delta(q_0, 0) \cup \delta(q_1, 0) \\ &= \{q_0, q_3\} \cup \phi \\ &= \{q_0, q_3\}\end{aligned}$$

Nondeterministic finite automaton (NFA)

Example Cont..

$$\begin{aligned}\hat{\delta}(q_0, 0100) &= \delta(\hat{\delta}(q_0, 010), 0) \\ &= \delta(\{q_0, q_3\}, 0) \\ &= \delta(q_0, 0) \cup \delta(q_3, 0) \\ &= \{q_0, q_3\} \cup \{q_4\} \\ &= \{q_0, q_3, q_4\} \\ \hat{\delta}(q_0, 01001) &= \delta(\hat{\delta}(q_0, 0100), 1) \\ &= \delta(\{q_0, q_3, q_4\}, 1) \\ &= \delta(q_0, 1) \cup \delta(q_3, 1) \cup \delta(q_4, 1) \\ &= \{q_0, q_1\} \cup \emptyset \cup \{q_4\} \\ &= \{q_0, q_1, q_4\}\end{aligned}$$

After the entire string is consumed, the FA is in the state q_4 . As q_4 is the final state, the string is accepted by FA

Nondeterministic finite automaton (NFA)

Example Cont..

$$M = (Q, \Sigma, \delta, q_0, F)$$

$$\begin{aligned} \text{where } Q &= \{q_0, q_1, q_2, q_3, q_4\} \\ \Sigma &= \{0, 1\} \\ \delta : \delta(q_0, 0) &= \{q_0, q_3\} \\ \delta(q_0, 1) &= \{q_0, q_1\} \\ \delta(q_1, 0) &= \phi \\ \delta(q_1, 1) &= \{q_2\} \\ \delta(q_2, 0) &= \{q_2\} \\ \delta(q_2, 1) &= \{q_2\} \\ \delta(q_3, 0) &= \{q_4\} \\ \delta(q_3, 1) &= \phi \\ \delta(q_4, 0) &= \{q_4\} \\ \delta(q_4, 1) &= \{q_4\} \\ q_0 &= q_0 \\ F &= \{q_2, q_4\} \end{aligned}$$

Nondeterministic finite automaton (NFA)

The extended transition function

- ① $\hat{\delta}(q, \epsilon) = q$
 - NFA cannot change its state without accepting an input symbol
- ② $\hat{\delta}(q, wa) = \{p / \text{for some state } r \text{ in } \hat{\delta}(q, w), \text{ for some state } p \text{ in } \delta(r, a)\}$
 - $\hat{\delta}(q, w) = \{r_1, r_2, r_3, r_4, \dots\}$
 - $\delta(\hat{\delta}(q, w), a) = \delta(\{r_1, r_2, r_3, r_4, \dots\}, a) = \{p_1, p_2, p_3, p_4, \dots\}$
- ③ $\hat{\delta}(q, a) = \delta(q, a)$
 - $\hat{\delta}(q, a) = \hat{\delta}(q, \epsilon a) = \delta(\hat{\delta}(q, \epsilon), a) = \delta(q, a)$
 - $\hat{\delta} = \delta$
- ④ $\delta(p, w) = \bigcup_{q \text{ in } p} \delta(q, w)$
 - Suppose $p = \{q_1, q_2, q_3, \dots, q_n\}$

$$\begin{aligned}\delta(\{q_1, q_2, q_3, \dots, q_n\}, w) &= \delta(q_1, w) \cup \delta(q_2, w) \cup \dots \cup \delta(q_n, w) \\ &= \bigcup_{i=1}^n \delta(q_i, w)\end{aligned}$$

Nondeterministic finite automaton (NFA)

The Language accepted by an NFA

- Set of all string accepted by an NFA is called language accepted by an NFA

$$L(M) = \{x / \hat{\delta}(q, x) \text{ is in } f\}$$

or

$$L(M) = \{w / \hat{\delta}(q_0, w) \cap f \neq \phi\}$$

$\hat{\delta}(q, w)$ contain a final state then it is accepted

All DFA's are NFA's

Differences between NFA and DFA

NFA	DFA
A Nondeterministic finite automaton is a 5-tuple $M = (Q, \Sigma, \delta, q_0, f)$ where $\delta : Q \times \Sigma \text{ into } 2^Q$	A Deterministic finite automaton is a 5-tuple $M = (Q, \Sigma, \delta, q_0, f)$ where $\delta : Q \times \Sigma \text{ into } Q$
NFA is the one in which there exists many paths for a specific input from current state to next state.	DFA is a FA in which there is only one path for a specific input from current state to next state.
Next move cannot be determined by the present state and current input symbol.	Next move can be determined by the present state and current input symbol.
NFA is easier to construct.	DFA is more difficult to construct.
NFA requires less space.	dfa requires more space.
Time required for executing an input string is more.	Time required for executing an input string is less

Table: Differences between NFA and DFA

Applications of FA:

- Used in Lexical analysis phase of a compiler to recognize tokens.
- Used in text editors for string matching.
- Software for designing and checking the behavior of digital circuits.

Limitations of FA:

- FA's will have finite amount of memory.
- The class of languages recognized by FA s is strictly the regular set. There are certain languages which are non regular i.e. cannot be recognized by any FA.

NFA to DFA Conversion

Let, $M = (Q, \Sigma, \delta, q_0, F)$ is an NFA which accepts the language $L(M)$. There should be equivalent DFA denoted by $M' = (Q', \Sigma', q_0', \delta', F')$ such that $L(M) = L(M')$.

- 1 Initially $Q' = \phi$
- 2 Add q_0 of NFA to Q' . Then find the transitions from this start state.
- 3 In Q' , find the possible set of states for each input symbol. If this set of states is not in Q' , then add it to Q' .
- 4 In DFA, the final state will be all the states which contain F (final states of NFA)

Finite Automata

Q: Convert the given NFA to DFA.

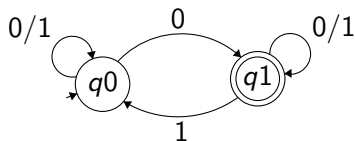


Figure: NFA

Solution: For the given transition diagram we will first construct the transition table.

δ	0	1
$\rightarrow q0$	$\{q0, q1\}$	$\{q0\}$
$*q1$	$\{q1\}$	$\{q0, q1\}$

Convert the given NFA to DFA

Now we will obtain δ' transition for state q_0 .

$$\begin{aligned}\delta'([q_0], 0) &= \{q_0, q_1\} \\ &= [q_0, q_1] \text{ (new state generated)} \\ \delta'([q_0], 1) &= \{q_1\} = [q_1]\end{aligned}$$

The δ' transition for state q_1 is obtained as:

$$\begin{aligned}\delta'([q_1], 0) &= \{q_1\} \\ \delta'([q_1], 1) &= \{q_0, q_1\}\end{aligned}$$

Now we will obtain δ' transition on $[q_0, q_1]$.

$$\begin{aligned}\delta'([q_0, q_1], 0) &= \delta(q_0, 0) \cup \delta(q_1, 0) \\ &= \{q_0, q_1\} \cup \{q_1\} \\ &= \{q_0, q_1\} \\ &= [q_0, q_1]\end{aligned}$$

Convert the given NFA to DFA

Similarly

$$\begin{aligned}\delta'([q0, q1], 1) &= \delta(q0, 1) \cup \delta(q1, 1) \\ &= \{q0\} \cup \{q0, q1\} \\ &= \{q0, q1\} \\ &= [q0, q1]\end{aligned}$$

As in the given NFA, $q1$ is a final state, then in DFA wherever, $q1$ exists that state becomes a final state. Hence in the DFA, final states are $[q1]$ and $[q0, q1]$.

The transition table for the constructed DFA will be:

δ	0	1
$\rightarrow[q0]$	$[q0, q1]$	$[q0]$
$*[q1]$	$[q1]$	$[q0, q1]$
$*[q0, q1]$	$[q0, q1]$	$[q0, q1]$

Table: Transition table

Convert the given NFA to DFA

The Transition diagram will be:

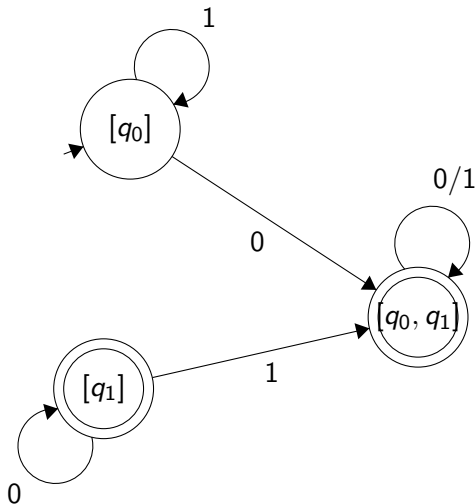


Figure: Transition Digram - DFA

Convert the given NFA to DFA

Even we can change the name of the states of DFA.

$$A = [q_0]$$

$$B = [q_1]$$

$$C = [q_0, q_1]$$

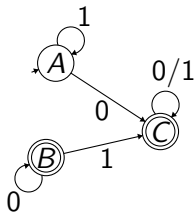


Figure: Transition diagram -DFA

Upon careful inspection, it appears that state B is unreachable from the initial state A . Thus, for the purpose of minimizing the DFA, it is advisable to remove state B .

Equivalence of NFA and DFA

- If L is a Language accepted by an NFA then there is a DFA that accept L
 - i.e for every NFA there exist an equivalent DFA
 - The transition of NFA to DFA is normally called a subset construction

Equivalence of NFA and DFA

Theorem 1.0

Let language $L \subseteq \Sigma^*$, and suppose L is accepted by NFA $N = (\Sigma, Q, q_0, F, \delta)$. There exists a DFA $D = (\Sigma, Q', q_0', F', \delta')$ that also accepts L . ($L(N) = L(D)$).

By allowing each state in the DFA D to represent a set of states in the NFA N , we are able to prove through induction that D is equivalent to N . Before we begin the proof, let's define the parameters of D :

- Q' is equal to the powerset of Q , $Q' = 2^Q$
- $q_0' = \{q_0\}$
- F' is the set of states in Q' that contain any element of F ,
 $F' = \{q \in Q' / q \cap F \neq \phi\}$
- δ' is the transition function for D .
 $\delta'(q, a) = \bigcup_{p \in q} \delta(p, a)$ for $q \in Q'$ and $a \in \Sigma$.

Equivalence of NFA and DFA

Now we will prove that $\hat{\delta}'(q_0, x) = \hat{\delta}(q_0, x)$ for every x . ie, $L(D) = L(N)$

Basis Step:

- Let x be the empty string ϵ .
- By the definition of extended transition function, NFA & DFA cannot change its state without accepting an input symbol

$$\hat{\delta}'(q_0, \epsilon) = \{q_0\}$$

$$\hat{\delta}(q_0, \epsilon) = \{q_0\}$$

Therefor $\delta'(q_0, \epsilon) = \delta(q_0, \epsilon)$ is true for $x = \epsilon$.

Inductive Step:

- Assume that for any y with $|y| \geq 0$, $\hat{\delta}'(q_0, y) = \hat{\delta}(q_0, y)$.
 - Let both state be $p = \{p_1, p_2, p_3 \dots\}$
- Let $n = |y|$, then we need to prove that for a string z with $|z| = n + 1$, $\hat{\delta}'(q_0, z) = \hat{\delta}(q_0, z)$.

Equivalence of NFA and DFA

- We can represent the string z as a concatenation of string y ($|y| = n$) and symbol a from the alphabet Σ . ($a \in \Sigma$). So, $z = ya$.

$$\begin{aligned}\hat{\delta}'(q_0, z) &= \hat{\delta}'(q_0, ya) \\ &= \delta'(\hat{\delta}'(q_0, y), a) \\ &= \delta'(\hat{\delta}(q_0, y), a) \quad (\text{by assumption}) \\ &= \bigcup_{p \in \hat{\delta}(q_0, y)} \delta'(p, a) \quad (\text{by definition of } \hat{\delta}) \\ &= \bigcup_{p \in \hat{\delta}(q_0, y)} \delta(p, a) \quad (\text{by assumption}) \\ &= \hat{\delta}(q_0, ya) \\ &= \hat{\delta}(q_0, z)\end{aligned}$$

- DFA D accepts a string x iff $\hat{\delta}'(q_0, x) \in F'$. From the above it follows that D accepts x iff $\hat{\delta}(q_0, x) \cap F \neq \phi$.
- So a string is accepted by DFA D if, and only if, it is accepted by NFA N .

NFA with ϵ transitions(ϵ -NFA)

In the case of ϵ -NFA without applying an input symbol the state will change

Formal Definition of NFA with ϵ transitions

$$M = (Q, \Sigma, \delta, q_0, F)$$

where

- Q is a set of states,
- Σ is the alphabet,
- δ is the transition function $\delta : Q \times \{\Sigma \cup \epsilon\}$ into 2^Q ,
- q_0 is the initial state,
- $F \subset Q$ is the set of final (or accepting) states.

$\delta(q, a)$ will consists of all states p such that there is a transition labelled a from q to p . Where a is either ϵ or a symbol in Σ .

NFA with ϵ transitions(ϵ -NFA)

Example:

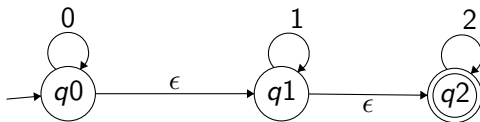


Figure: Transition diagram

δ	0	1	2	ϵ
$\rightarrow q0$	$\{q0\}$	ϕ	ϕ	$\{q1\}$
$q1$	ϕ	$\{q1\}$	ϕ	$\{q2\}$
$*q2$	ϕ	ϕ	$\{q2\}$	ϕ

Table: Transition table

ϵ -closure(q)

- It is a set of states that can be reached from q along the path labelled by ϵ only

$$\epsilon - \text{closure}(p_i) = \{q_j / q_j \text{ is reachable from } p_i \text{ using zero or more } \epsilon - \text{transition}\}$$

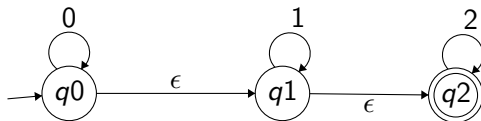


Figure: Transition diagram

$$\epsilon - \text{closure}(q_0) = \{q_0, q_1, q_2\}$$

$$\epsilon - \text{closure}(q_1) = \{q_1, q_2\}$$

$$\epsilon - \text{closure}(q_2) = \{q_2\}$$

NFA with ϵ transitions(ϵ -NFA)

Extended transition function($\hat{\delta}$) in ϵ -NFA

① $\hat{\delta}(q, \epsilon) = \epsilon - \text{closure}(q)$

② for w in Σ^* and a in Σ

$$\begin{aligned}\hat{\delta}(q, wa) &= \epsilon - \text{closure}(p) \\ &= \{p / \text{for some } r \text{ in } \hat{\delta}(q, w), p \text{ is in } \delta(r, a)\}\end{aligned}$$

③ $\delta(R, a) = \bigcup_{q \text{ in } R} \delta(q, a)$

④ $\hat{\delta}(R, w) = \bigcup_{q \text{ in } R} \hat{\delta}(q, w)$

⑤ $\hat{\delta}(q, a) = \epsilon - \text{closure}(\delta(\hat{\delta}(q, \epsilon), a))$

Conversion of Epsilon-NFA to NFA

- 1 Find out all the ϵ -transitions from each state from Q . That will be called as $\epsilon - closure(q_i)$ where, $q_i \in Q$.
- 2 Then, δ_1 transitions can be obtained. The δ_1 transitions means an ϵ -closure on δ moves.
- 3 Step 2 is repeated for each input symbol and for each state of given NFA.
- 4 By using the resultant status, the transition table for equivalent NFA without ϵ can be built.

Conversion of Epsilon-NFA to NFA

- 1 Find out all the ϵ -transitions from each state from Q . That will be called as $\epsilon - closure(q_i)$ where, $q_i \in Q$.
- 2 Then, δ_1 transitions can be obtained. The δ_1 transitions means an ϵ -closure on δ moves.
- 3 Step 2 is repeated for each input symbol and for each state of given NFA.
- 4 By using the resultant status, the transition table for equivalent NFA without ϵ can be built.

Equivalence of NFA with and without epsilon transitions

Example: Convert the given NFA with epsilon to NFA without epsilon.

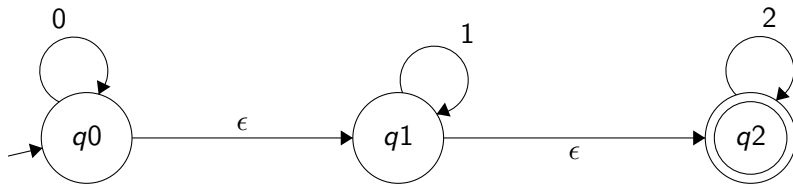


Figure: NFA-With ϵ transition

Equivalence of NFA with and without epsilon transitions

Solution: We will first obtain ϵ - *closure* of each state

$$\epsilon - \text{closure}(q_0) = \{q_0, q_1, q_2\}$$

$$\epsilon - \text{closure}(q_1) = \{q_1, q_2\}$$

$$\epsilon - \text{closure}(q_2) = \{q_2\}$$

Now we will obtain δ^1 transitions for each state on each input symbol

$$\begin{aligned}\delta'(q_0, 0) &= \epsilon - \text{closure}(\delta(\hat{\delta}(q_0, \epsilon), 0)) \\ &= \epsilon - \text{closure}(\delta(\epsilon - \text{closure}(q_0), 0)) \\ &= \epsilon - \text{closure}(\delta(q_0, q_1, q_2), 0) \\ &= \epsilon - \text{closure}(\delta(q_0, 0) \cup \delta(q_1, 0) \cup \delta(q_2, 0)) \\ &= \epsilon - \text{closure}(q_0 \cup \phi \cup \phi) \\ &= \epsilon - \text{closure}(q_0) \\ &= \{q_0, q_1, q_2\}\end{aligned}$$

Equivalence of NFA with and without epsilon transitions

Solution cont..

$$\begin{aligned}\delta'(q_0, 1) &= \epsilon - \text{closure}(\delta(\hat{\delta}(q_0, \epsilon), 1)) \\ &= \epsilon - \text{closure}(\delta(q_0, q_1, q_2), 1)) \\ &= \epsilon - \text{closure}(\delta(q_0, 1) \cup \delta(q_1, 1) \cup \delta(q_2, 1)) \\ &= \epsilon - \text{closure}(\phi \cup q_1 \cup \phi) \\ &= \epsilon - \text{closure}(q_1) \\ &= \{q_1, q_2\} \\ \delta'(q_0, 2) &= \epsilon - \text{closure}(\delta(\hat{\delta}(q_0, \epsilon), 2)) \\ &= \epsilon - \text{closure}(\delta(q_0, q_1, q_2), 2)) \\ &= \epsilon - \text{closure}(\delta(q_0, 2) \cup \delta(q_1, 2) \cup \delta(q_2, 2)) \\ &= \epsilon - \text{closure}(\phi \cup \phi \cup q_2) \\ &= \epsilon - \text{closure}(q_2) \\ &= \{q_2\}\end{aligned}$$

Equivalence of NFA with and without epsilon transitions

Solution cont..

$$\begin{aligned}\delta'(q1, 0) &= \epsilon - \text{closure}(\delta(\hat{\delta}(q1, \epsilon), 0)) \\ &= \epsilon - \text{closure}(\delta(q1, q2), 0)) \\ &= \epsilon - \text{closure}(\delta(q1, 0) \cup \delta(q2, 0)) \\ &= \epsilon - \text{closure}(\phi \cup \phi) \\ &= \epsilon - \text{closure}(\phi) \\ &= \phi \\ \delta'(q1, 1) &= \epsilon - \text{closure}(\delta(\hat{\delta}(q1, \epsilon), 1)) \\ &= \epsilon - \text{closure}(\delta(q1, q2), 1)) \\ &= \epsilon - \text{closure}(\delta(q1, 1) \cup \delta(q2, 1)) \\ &= \epsilon - \text{closure}(q1 \cup \phi) \\ &= \epsilon - \text{closure}(q1) \\ &= \{q1, q2\}\end{aligned}$$

Equivalence of NFA with and without epsilon transitions

Solution cont..

$$\begin{aligned}\delta'(q1, 2) &= \epsilon - \text{closure}(\delta(\hat{\delta}(q1, \epsilon), 2)) \\ &= \epsilon - \text{closure}(\delta(q1, q2), 2)) \\ &= \epsilon - \text{closure}(\delta(q1, 2) \cup \delta(q2, 2)) \\ &= \epsilon - \text{closure}(\phi \cup q2) \\ &= \epsilon - \text{closure}(q2) \\ &= \{q2\} \\ \delta'(q2, 0) &= \epsilon - \text{closure}(\delta(\hat{\delta}(q2, \epsilon), 0)) \\ &= \epsilon - \text{closure}(\delta(q2), 0)) \\ &= \epsilon - \text{closure}(\delta(q2, 0)) \\ &= \epsilon - \text{closure}(\phi) \\ &= \phi\end{aligned}$$

Equivalence of NFA with and without epsilon transitions

Solution cont..

$$\begin{aligned}\delta'(q_2, 1) &= \epsilon - \text{closure}(\delta(\hat{\delta}(q_2, \epsilon), 1)) \\ &= \epsilon - \text{closure}(\delta(q_2), 1) \\ &= \epsilon - \text{closure}(\delta(q_2, 1)) \\ &= \epsilon - \text{closure}(\phi) \\ &= \phi \\ \delta'(q_2, 2) &= \epsilon - \text{closure}(\delta(\hat{\delta}(q_2, \epsilon),)) \\ &= \epsilon - \text{closure}(\delta(q_2), 2)) \\ &= \epsilon - \text{closure}(\delta(q_2, 2)) \\ &= \epsilon - \text{closure}(q_2) \\ &= \{q_2\}\end{aligned}$$

Equivalence of NFA with and without epsilon transitions

Solution cont..

Now, we will summarize all the computed δ' transitions as given below

$$\delta'(q_0, 0) = \{q_0, q_1, q_2\}$$

$$\delta'(q_0, 1) = \{q_1, q_2\}$$

$$\delta'(q_0, 2) = \{q_2\}$$

$$\delta'(q_1, 0) = \{\phi\}$$

$$\delta'(q_1, 1) = \{q_1, q_2\}$$

$$\delta'(q_1, 2) = \{q_2\}$$

$$\delta'(q_2, 0) = \{\phi\}$$

$$\delta'(q_2, 1) = \{\phi\}$$

$$\delta'(q_2, 2) = \{q_2\}$$

Equivalence of NFA with and without epsilon transitions

Solution cont.. The transition table is given below

δ	0	1	2
\rightarrow^*q0	$\{q0, q1, q2\}$	$\{q1, q2\}$	$\{q2\}$
$q1$	ϕ	$\{q1, q2\}$	$\{q2\}$
*q2	ϕ	ϕ	$\{q2\}$

Table: Transition table NFA without ϵ transition

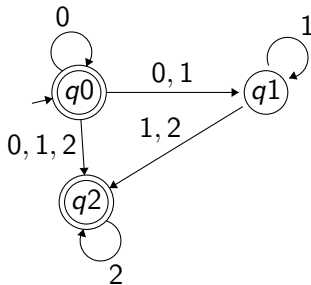


Figure: Transition diagram NFA without ϵ transition

Conversion NFA with epsilon to DFA

The method for converting the NFA with ϵ to DFA is explained below

- Consider $M = (Q, \Sigma, \delta, q_0, F)$ is NFA with ϵ .
- We have to convert this NFA with ϵ to equivalent DFA denoted by $M_D = (Q_D, \Sigma, \delta_D, q_0, F_D)$
- Then obtain, $\epsilon - \text{closure}(q_0) = \{p_1, p_2, p_3, \dots, p_n\}$
- then $[p_1, p_2, p_3, \dots, p_n]$ becomes a start state of DFA
- now $[p_1, p_2, p_3, \dots, p_n] \in Q_D$
- We will obtain δ transition on $[p_1, p_2, p_3, \dots, p_n]$ for each input.
 - $\delta_D([p_1, p_2, p_3, \dots, p_n], a) = \epsilon - \text{closure}(\delta(p_1, a) \cup \delta(p_2, a) \cup \dots \cup \delta(p_n, a))$
 - $= \bigcup_{i=1}^n \epsilon - \text{closure}(p_i, a)$
Where a is input $\in \Sigma$
- The state obtained $[p_1, p_2, p_3, \dots, p_n] \in Q_D$.
- The states containing final state in p_i is a final state in DFA

Conversion NFA with epsilon to DFA

Example: Convert the following NFA with epsilon to equivalent DFA

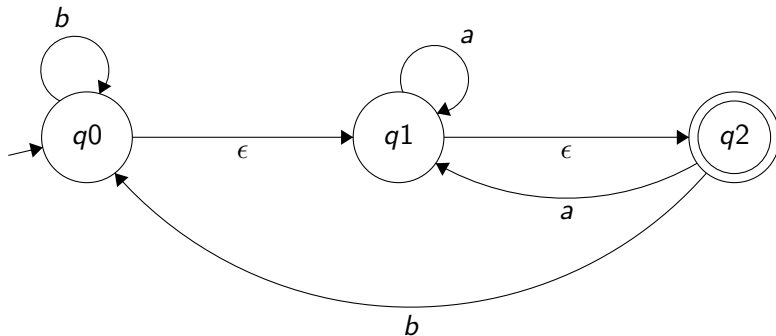


Figure: NFA with ϵ transition

Conversion NFA with epsilon to DFA

Solution:

To convert this NFA with epsilon, we will first find the ϵ -closures,

$$\epsilon - \text{closure}(q_0) = \{q_0, q_1, q_2\}$$

$$\epsilon - \text{closure}(q_1) = \{q_1, q_2\}$$

$$\epsilon - \text{closure}(q_2) = \{q_2\}$$

When, $\epsilon - \text{closure}(q_0) = \{q_0, q_1, q_2\}$, we will call this state as A.

Now, let us find transition on A with every input symbol

$$\begin{aligned}\delta'(A, a) &= \epsilon - \text{closure}(\delta(A, a)) \\ &= \epsilon - \text{closure}(\delta(q_0, q_1, q_2), a)) \\ &= \epsilon - \text{closure}(\delta(q_0, a) \cup \delta(q_1, a) \cup \delta(q_2, a)) \\ &= \epsilon - \text{closure}(\phi \cup q_1 \cup q_2) \\ &= \epsilon - \text{closure}(q_1 \cup q_2) \\ &= \{q_1, q_2\} \text{ let us call it as state B}\end{aligned}$$

Conversion NFA with epsilon to DFA

Solution cont..

$$\begin{aligned}\delta'(A, b) &= \epsilon - \text{closure}(\delta(A, b)) \\ &= \epsilon - \text{closure}(\delta(q_0, q_1, q_2), b)) \\ &= \epsilon - \text{closure}(\delta(q_0, b) \cup \delta(q_1, b) \cup \delta(q_2, b)) \\ &= \epsilon - \text{closure}(q_0 \cup \phi \cup q_0) \\ &= \epsilon - \text{closure}(q_0) \\ &= \{q_0, q_1, q_2\} \text{ its nothing but state } A \\ \delta'(B, a) &= \epsilon - \text{closure}(\delta(B, a)) \\ &= \epsilon - \text{closure}(\delta(q_1, q_2), a)) \\ &= \epsilon - \text{closure}(\delta(q_1, a) \cup \delta(q_2, a)) \\ &= \epsilon - \text{closure}(q_1 \cup q_2) \\ &= \epsilon - \text{closure}(q_1) \\ &= \{q_1, q_2\} \text{ its nothing but state } B\end{aligned}$$

Conversion NFA with epsilon to DFA

Solution cont..

$$\begin{aligned}\delta'(B, b) &= \epsilon - \text{closure}(\delta(B, b)) \\ &= \epsilon - \text{closure}(\delta(q1, q2), b)) \\ &= \epsilon - \text{closure}(\delta(q1, b) \cup \delta(q2, b)) \\ &= \epsilon - \text{closure}(\phi \cup q0) \\ &= \epsilon - \text{closure}(q0) \\ &= \{q0, q1, q2\} \text{ its nothing but state } A\end{aligned}$$

Conversion NFA with epsilon to DFA

Solution cont.. Hence, the transition table for the generated DFA is as follows

δ	a	b
\rightarrow^*A	B	A
$*B$	B	A

Table: DFA without ϵ

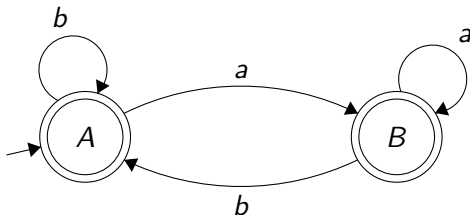


Figure: DFA without ϵ diagram

Equivalence of NFA with and without epsilon transitions

Theorem 1.1

If L is a language accepted by NFA with ϵ -transition, Then L is accepted by NFA without ϵ -transition.

Proof:

- Let $M = (Q, \Sigma, \delta, q_0, F)$ be a NFA with ϵ -transition, then we need to construct an NFA, $M' = (Q, \Sigma, \delta', q_0, F')$
- Where

$$F' = \begin{cases} F \cup \{q_0\} & \text{if } \epsilon\text{-closure}(q_0) \text{ contains a state from } F, \\ F & \text{Otherwise} \end{cases}$$

and

$$\delta'(q, a) = \delta(q, a) \text{ for } q \in Q \text{ and } a \in \Sigma$$

Equivalence of NFA with and without epsilon transitions

Proof Cont..

Basis

- $|x| = 1$ then x is a symbol a and

$$\delta'(q_0, a) = \delta(q_0, a) \text{ by definition of } \delta'$$

Induction

- Assume the hypothesis is true for the string of length m or less

$$\text{i.e. } \hat{\delta}'(q_0, x) = \hat{\delta}(q_0, x)$$

To prove this theorem we must show that

$$\hat{\delta}'(q_0, xa) = \hat{\delta}(q_0, xa) \text{ where } |xa| = m + 1$$

Equivalence of NFA with and without epsilon transitions

$$\begin{aligned}\hat{\delta}(q_0, xa) &= \delta'(\hat{\delta}(q_0, x), a) \\ &= \delta'(\hat{\delta}(q_0, x), a) \text{ by hypothesis}\end{aligned}$$

let $\hat{\delta}(q_0, x) = P = \{q_0, q_1, q_2, q_3, \dots, q_n\}$

$$\begin{aligned}\text{therefor } \hat{\delta}'(q_0, xa) &= \delta'(P, a) \\ &= \bigcup_{q \in P} \delta'(q, a) \\ &= \bigcup_{q \in P} \delta(q, a) \\ &= \delta(P, a) \\ &= \delta(\hat{\delta}(q_0, x), a) \\ &= \hat{\delta}(q_0, xa)\end{aligned}$$

Thus $\delta'(q_0, xa) = \hat{\delta}(q_0, xa)$

Regular Grammar

Chomsky Hierarchy in Theory of Computation

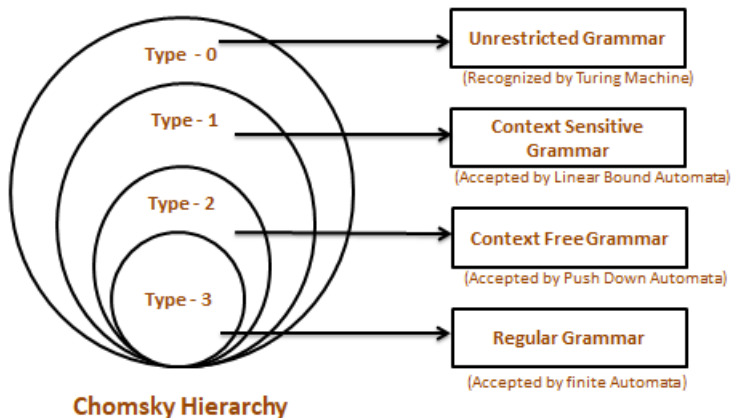


Figure: Chomsky Hierarchy in Theory of Computation

Regular Grammar(RG)-Type 3

- Regular grammar generates regular language.
- Regular Languages are recognized by using finite state automata.

Formal definition of Regular Grammar(RG)-Type 3

- A Regular Grammar 'G' can be formally described using four tuples as $G = (V, T, P, S)$
- where,
 - V = Set of Variables or Non-Terminal Symbols(Capital letters)
 - T = Set of Terminal Symbols(Small letters, Numbers etc)
 - S = Start Symbol $S \in V$
 - P = Production Rules for Terminals and Non-Terminals

All productions are of the form

$$A \rightarrow xB$$

$$A \rightarrow Cx$$

$$A \rightarrow x$$

where $A, B, C \in V, x \in T^*$

Types of regular grammar:

- 1 Right linear grammar(RLG)
- 2 Left Linear grammar(LLG)

Regular Grammar

Right-linear grammar

- A grammar is said to be Right-linear if all productions are of the form:

$$A \rightarrow xB$$

$$A \rightarrow xC$$

$$A \rightarrow y$$

where $A, B, C \in V, x, y \in T^*$

Example

$$G = (\{S, B\}, \{a, b\}, P, S),$$

$$P =$$

$$S \rightarrow aB \mid bS \mid \epsilon$$

$$B \rightarrow aS \mid bB$$

Regular Grammar

Left Linear grammar

- A grammar is said to be Left Linear if all productions are of the form:

$$A \rightarrow Bx$$

$$A \rightarrow Cy$$

$$A \rightarrow x$$

where $A, B, C \in V, x, y \in T^*$

Example

$$G = (\{S, B\}, \{a, b\}, P, S),$$

$$P =$$

$$S \rightarrow Ba|Sb|\epsilon$$

$$B \rightarrow Sa|Bb$$

Regular Grammar

Derivation of sentences Grammar

$$A \rightarrow aB$$

$$B \rightarrow aB$$

$$B \rightarrow bB$$

$$B \rightarrow \epsilon$$

Derivation from the above grammar

$$A \Rightarrow aB \Rightarrow a$$

$$A \Rightarrow aB \Rightarrow aaB \Rightarrow aa$$

$$A \Rightarrow aB \Rightarrow abB \Rightarrow ab$$

$$A \Rightarrow aB \Rightarrow aaB \Rightarrow aaaB \Rightarrow aaa$$

Language generated by the grammar is $L(G) = \{ax \mid x \in \{a, b\}^*\}$

Regular Grammar

- Language generated by the grammar is $L(G) = \{ax \mid x \in \{a, b\}^*\}$
- i.e $a(a + b)^*$ (Called Regular Expression)

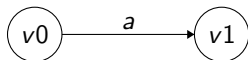
Language generated by Regular Grammar

- The class of languages generated by the family of regular grammar is precisely the set of Regular Language
 - The language generated by a Regular grammar is Regular
 - Any Regular Language can be generated by a Regular Grammar

Converting Regular Grammar to DFA

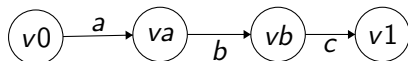
Assume that a regular grammar is given in its right-linear form, this grammar may be easily converted to a DFA. A right-linear grammar, defined by $G = (V, T, S, P)$, may be converted to a DFA, defined by $M = (Q, \Sigma, \delta, q_0, F)$ by:

- 1 Create a state for each variable.
- 2 Convert each production rule into a transition.
 - a If the production rule is of the form $V_i \rightarrow aV_j$, where $a \in T$, add the transition $\delta(V_i, a) = V_j$ to M . For example, $V_0 \rightarrow aV_1$ becomes:

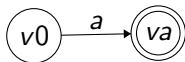


Regular Grammar

- b If the production rule is of the form $V_i \rightarrow wV_j$, where $w \in T^*$, create a series of states which derive w and end in V_j . Add the states in between to Q . For example, $V_0 \rightarrow abcV_1$ becomes



- c If the production rule is of the form $V_i \rightarrow w$, where $w \in T^*$, create a series of states which derive w and end in a final state. For example, $V_0 \rightarrow a$ becomes:



Regular Grammar

Theorem

Let $G=(V,T,P,S)$ be a right linear grammar. then $L(G)$ is regular language

Proof

We assume that $V = \{V_0, V_1, \dots\}$, that $S = V_0$, and that we have productions of the form $V_0 \rightarrow v_1 V_i, V_i \rightarrow v_2 V_j, \dots$ or $V_n \rightarrow v_l, \dots$. If w is a string in $L(G)$, then because of the form of the productions

$$\begin{aligned} V_0 &\implies v_1 V_i \\ &\implies v_1 v_2 V_j \\ &\quad \cdot \\ &\quad \cdot \\ &\quad \cdot \\ &\implies v_1 v_2 \dots v_k V_n \\ &\implies v_1 v_2 \dots v_k v_l = w. \end{aligned}$$

Regular Grammar

The automaton to be constructed will reproduce the derivation by consuming each of these v 's in turn. The initial state of the automaton will be labeled V_0 , and for each variable V_i there will be a nonfinal state labeled V_i . For each production

$$V_i \Longrightarrow a_1 a_2 \dots a_m V_j,$$

the automaton will have transitions to connect V_i and V_j that is, δ will be defined so that

$$\delta^*(V_i, a_1 a_2 \dots a_m) = V_j$$

For each production

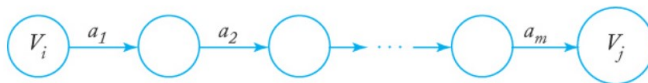
$$V_i \Longrightarrow a_1 a_2 \dots a_m,$$

the corresponding transition of the automaton will be

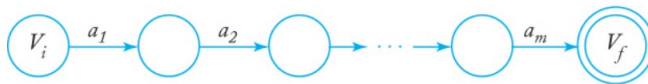
$$\delta^*(V_i, a_1 a_2 \dots a_m) = V_f,$$

Regular Grammar

where V_f is a final state. The intermediate states that are needed to do this are of no concern and can be given arbitrary labels.



Represents $V_i \rightarrow a_1 a_2 \dots a_m V_j$



Represents $V_i \rightarrow a_1 a_2 \dots a_m$

Regular Grammar

Suppose now that $w \in L(G)$ so that the above equation is satisfied. In the nfa there is, by construction, a path from V_0 to V_i labeled v_1 , a path from V_i to V_j labeled v_2 , and so on, so that clearly

$$V_f \in \delta^*(V_0, w),$$

and w is accepted by M .

Conversely, assume that w is accepted by M . Because of the way in which M was constructed, to accept w the automaton has to pass through a sequence of states V_0, V_i, \dots to V_f , using paths labeled v_1, v_2, \dots . Therefore, w must have the form

$$w = v_1 v_2 v_k v_l$$

and the derivation

$$V_0 \Rightarrow v_1 V_i \Rightarrow v_1 v_2 V_j \xRightarrow{*} v_1 v_2 \dots v_k V_k \Rightarrow v_1 v_2 \dots v_k v_l$$

is possible. Hence w is in $L(G)$, and the theorem is proved.

Regular Grammar

Theorem

If L is a regular language on the alphabet Σ , then there exists a right-linear grammar $G = (V, \Sigma, S, P)$ such that $L = L(G)$.

Proof

Let $M = (Q, \Sigma, \delta, q_0, F)$ be a dfa that accepts L . We assume that $Q = \{q_0, q_1, \dots, q_n\}$ and $\Sigma = \{a_1, a_2, \dots, a_m\}$. Construct the right-linear grammar $G = (V, \Sigma, S, P)$ with $V = \{q_0, q_1, \dots, q_n\}$ and $S = q_0$. For each transition $\delta(q_i, a_j) = q_k$ of M , we put in P the production

$$q_i \rightarrow a_j q_k.$$

In addition, if q_k is in F , we add to P the production

$$q_k \rightarrow \lambda.$$

Regular Grammar

Proofcont..

We first show that G defined in this way can generate every string in L . Consider $w \in L$ of the form $w = a_i a_j \dots a_k a_l$. For M to accept this string it must make moves via

$$\delta(q_0, a_i) = q_p,$$

$$\delta(q_p, a_j) = q_r,$$

.

.

$$\delta(q_s, a_k) = q_t,$$

$$\delta(q_t, a_l) = q_f \in F.$$

Regular Grammar

Proofcont..

By construction, the grammar will have one production for each of these δ 's. Therefore, we can make the derivation

$$\begin{aligned} q_0 &\Longrightarrow a_i q_p \Longrightarrow a_i a_j q_r \stackrel{*}{\Longrightarrow} a_i a_j \dots a_k q_t \\ &\Longrightarrow a_i a_j \dots a_k a_l q_f \Longrightarrow a_i a_j \dots a_k a_l, \end{aligned}$$

with the grammar G , and $w \in L(G)$.

Conversely, if $w \in L(G)$, then its derivation must have the form the above equation. But this implies that $\delta^*(q_0, a_i a_j \dots a_k a_l) = q_f$, completing the proof.