Parul® University

NAAC A++

Introduction to Alphabet, languages and grammars

Chapter - 1: Introduction

Prof. Riddhi Atulkumar Mehta
Assistant Professor
Department of Computer Science and Engineering

Parul® University

NAAC A++

INDEX

Content

What is Theory of Computation?

Theory of Computation is the branch of computer science that deals with how problems can be solved using algorithms and how efficiently they can be solved.

Main Areas:

- Automata Theory: Study of abstract machines (finite automata, pushdown automata, etc.)
- Formal Languages: Study of syntax and structure of languages
- Computability Theory: What problems can be solved
- Complexity Theory: How efficiently problems can be solved

Alphabet (Σ)

Definition:

An alphabet is a finite, non-empty set of symbols.

Notation:

Σ (Greek letter sigma)

Examples:

- Σ = {0, 1} → Binary alphabet
- Σ = {a, b, c, ..., z} → English alphabet
- Σ = {+, *, (,), a, b}

Strings and Their Properties

String:

A finite sequence of symbols taken from a given alphabet.

Examples:

If Σ = {a, b}, then "ab", "aab", and "bbaaa" are strings over Σ.

Key Concepts:

Length (|w|): Number of symbols in string w

|"abc"| = 3

Empty String (ε): A string with zero symbols

|ε| = 0

Language

Definition:

A language is a set of strings formed using the symbols of a given alphabet.

Examples:

- L = {“a”, “ab”, “abc”}
- Over Σ = {0, 1}, L = {ε, 0, 11, 101} is a binary language

Types of Languages:

- Finite Language: Has a limited number of strings
- Infinite Language: Has unlimited strings (e.g., $L = \{a^n \mid n \geq 0\}$)

Language Construction Methods

Important Operations on Languages:

1.Union ($L_1 \cup L_2$):

Set of strings that belong to L_1 or L_2 or both

2.Concatenation ($L_1 \cdot L_2$):

All strings formed by taking a string from L_1 followed by a string from L_2

3.Kleene Star (L*):

Set of all strings including ε, formed by zero or more concatenations of strings from L

Example:

If L = {a}, then L* = {ε, a, aa, aaa, ...}

Grammar

Definition:
A grammar is a formal set of rules used to generate strings in a language.

A grammar G is a 4-tuple:

$G = (V, \Sigma, P, S)$

Where:

- V: Set of variables (non-terminals)
- Σ: Set of terminals (alphabet symbols)
- P: Set of production rules
- S: Start symbol ($S \in V$)

Parul® University | NAAC GRADE A++

https://paruluniversity.ac.in/