
بسم الله الرحمن الرحيم

Introduction

Dr. Nasir Jalal PhD (Computer Science)

Lecturer

CS & IT Department, CUVAS Bahawalpur

Course Detail

BS CS 5th Semester Fall 2022-2026

Course: Theory of Programming Languages

Course Code: CS IT 507

Lecture: 4 (18-09-2024)

Outline Lecture 4

- Revision of Lecture 3
- Prerequisite of GFG
- What is Language
- Difference between Spoken language and Computer Language.
- What is CFG

Symbols

Symbols are an entity or individual objects, which can be any letter, alphabet or any picture.

- Example:

1, a, b, #

Alphabets

Definition:

A finite non-empty set of symbols (letters), is called an alphabet. It is denoted by Σ (Greek letter sigma).

Example:

Σ={a,b}

Σ={0,1} //important as this is the language
//which the computer understands.

Σ={i,j,k}

- Examples:

∑ = {a, b}

∑ = {A, B, C, D}

∑ = {0, 1, 2}

∑ = {0, 1, ..., 5]

∑ = {#, β, Δ}

String

It is a finite collection of symbols from the alphabet. The string is denoted by w.

- **Example 1:**

If ∑ = {a, b}, various string that can be generated from ∑ are {ab, aa, aaa, bb, bbb, ba, aba.....}.

- **A string with zero occurrences of symbols is known as an empty string. It is represented by ε.**
- **The number of symbols in a string w is called the length of a string. It is denoted by |w|.**
- **Example 2:**

w = 010

|w| = 3

Words

- Definition:

  Words are strings belonging to some language.

Example:

If Σ= {f} then a language L can be defined as

L={f^n : n=1,2,3,…..} or L={f,ff,fff,….}

Here f,ff,fff,… are the words of L

Valid/In-valid alphabets

- While defining an alphabet, an alphabet may contain letters consisting of group of symbols for example Σ_1= {B, aB, bab, d}.

- Now consider an alphabet
  Σ_2= {B, Ba, bab, d} and a string BababB.

Scan “BababB”

This string can be **tokenized** in two different ways

- (Ba), (bab), (B)
- (B), (abab), (B)

Which shows that the second group cannot be identified as a string, defined over Σ

NOTE

- All words are strings, but not all strings are words.

Language

A language is a collection of appropriate string. A language which is formed over Σ can be **Finite** or **Infinite**.

- **Example: 1**

L1 = {Set of string of length 2} = {aa, bb, ba, bb} **Finite Language**

- **Example: 2**

L2 = {Set of all strings starts with 'a'} = {a, aa, aaa, abb, abbb, ababb} **Infinite Language**

Remarks

- While defining an alphabet of letters consisting of more than one symbols, no letter should be started with the letter of the same alphabet *i.e.* one letter should not be the prefix of another. However, a letter may be ended in the letter of same alphabet *i.e.* one letter may be the suffix of another.
- Σ_1= {B, aB, bab, d}
- Σ_2= {B, Ba, bab, d}

Σ_1 is a valid alphabet while Σ_2 is an in-valid alphabet.

Length of Strings

- Definition:

  The length of string s, denoted by |s|, is the number of letters in the string.

- Example:

  Σ={a,b}

  s=ababa

  |s|=5

Example:2 String Length

Σ= {B, aB, bab, d}

s=BaBbabBd

Tokenizing=(B), (aB), (bab), (d)

|s|=4

Reverse of a String

- **Definition:**

  The reverse of a string s denoted by Rev(s) or s^r, is obtained by writing the letters of s in reverse order.

- **Example:**

  If s=abc is a string defined over Σ={a,b,c}

  then Rev(s) or s^r = cba

Reverse of a String

- Example 2:

  Σ= {B, aB, bab, d}

  s=BaBbabBd

  Rev(s)=dBbabaBB

CFG Overview

- A context-free grammar (CFG) consisting of a finite set of grammar rules is a quadruple

Terminals Represented using small case letter like a, b, c etc

Non-Terminals: The symbols that must be replaced by other things are called non-terminals. These symbols are represented using a capital letter like A, B, C, etc

Productions: The grammatical rules are often called productions or Principals.

Start Symbol:

CFG Overview

A right-regular grammar is a **context-free grammar** in which the right-hand side of every production rule has one of the following forms: the empty string; a string consisting of a single non-terminal symbol; or a string consisting of a single terminal symbol followed by a single non-terminal symbol.

CFG Overview

CFG Example

$\sum = \{a,b\}$

productions:

1. $S \rightarrow XaaX$
2. $X \rightarrow aX$
3. $X \rightarrow bX$
4. $X \rightarrow \varepsilon$

This grammar defines the language expressed by Regular Expression $(a+b)^*aa(a+b)^*$.

CFG Example

Some of the strings generated by CFG are

aa

aabaababaabb

aaaa

aabaabbbaaabbb

aaabbbbbaaabaaa

aaabbbaaba

aaaab

aaaaaabaab

aaaababbaabb

aabbbbaaaaaabaaa

Process to generate string aaabbbaaba in explain in table

Rule	Application	Result
Start → S	Start	S
S → XaaX	S	XaaX
X → aX	XaaX	aXaaX
X → aX	aXaaX	aaXaaX
X → aX	aaXaaX	aaaXaaX
X → bX	aaaXaaX	aaabXaaX
X → bX	aaabXaaX	aaabbXaaX
X → bX	aaabbXaaX	aaabbbXaaX
X → ε	aaabbbXaaX	aaabbbaaX
X → bX	aaabbbaaX	aaabbbaabX
X → aX	aaabbbaabX	aaabbbaabaX
X → ε	aaabbbaabaX	aaabbbaaba

THANKS