**Deerwalk Institute Of Technology**

**Compiler Design and Construction**

**Submitted By: Submitted To:**

**Name: Sagar Giri Satya Maharjan**

**Roll No. 205**

**Section: A**

**Date: August 21 2015**

## Alphabet, String and Language

* **Alphabet:**
  + - An alphabet is a finite set of objects called symbols.

Notation: Σ = {a, b, . . . , z}

* + - The symbols can be the letters such as {a, b, c}, bits {0, 1}, digits {0, 1, 2, 3... 9}, common characters like {$, #, ...} etc.
    - Example: {0,1} – Binary alphabets, {+, −, \*} – Special symbols
* **Strings:**
  + - Strings are denoted by lower case letters.
    - String is a finite sequence of symbols taken from some alphabet. E.g. 0110 is a string from binary alphabet, “compiler” is a string over alphabet {a, b, c ... z}.
    - Empty String is a string with zero occurrences of symbols. It is denoted by ‘ε’ (epsilon).
    - The length of a string w, denoted by |w|, is the number of positions for symbols in w. We have for every non empty strings, length (s) ≥ 0.
* **Languages**
  + - Languages are denoted by upper case letters.
    - A language over an alphabet Σ is a set of strings over Σ.

Notation: L, M, N, . . . for languages. |L| for the size (number of strings) of L.

* + - Σ∗ will denote a set of all strings over Σ. Then, a language **L** over Σ is just a subset of Σ∗. Hence, Σk denotes a set of all strings of length k over Σ.
    - Operations we can perform in language are:

→ Union, Intersection, Subtraction, Complementation, Concatenation