

## kzlibrarypatterns **Basic Terminology**

### **Introduction**

In the automata theory, we have to deal with some mathematical preliminaries. As examples in finite automata and finite state machine the knowledge of set theory is necessary, in grammar and language section we need the basic knowledge of alphabet, string, and substring, and in the regular expression chapter we need the concept of prefix, suffix, etc. The knowledge of the basic operations on a set such as union, intersection, difference, Cartesian product, power set, and concatenation product are required throughout the syllabus of formal language and automata theory.

For this reason, in this chapter, we shall discuss some basic terminologies related to mathematics which are required for automata theory.

### **1.1 Basics of String**

A string has some features as follows:

- Symbol: A symbol is a user-defined entity.
- Alphabet: An alphabet is a finite set of symbols denoted by  $\Sigma$  in automata. An alphabet is a set of symbols used to construct a language. As an example,  $\{0, 1\}$  is a binary alphabet, and  $\{A, Z, a, \dots, z\}$  is an alphabet set for the English language.
- String: A string is defined as a sequence of symbols of finite length. A string is denoted by  $w$  in automata. As an example, 000111 is a binary string.

(The length of a string  $w$  is denoted by  $|w|$ . For the previous case,  $|w| = |000111| = 6$ .)

- Prefix: A prefix of a string is the string formed by taking any number of symbols of the string.

Example: Let us take a string  $w = 0111$ . For the particular string,  $\lambda$ , 0, 01, 011, and 0111 are prefixes of the string 0111. For a string of length  $n$ , there are  $n + 1$  number of prefixes.

- Proper prefix: For a string, any prefix of the string other than the string itself is called as the proper prefix of the string.

Example: For the string  $w = 0111$ , the proper prefixes are  $\lambda$ , 0, 01, and 011.

- Suffix: A suffix of a string is formed by taking any number of symbols from the end of the string.

Example: Let us take a string  $w = 0110$ . For the particular string,  $\lambda$ , 0, 10, 110, and 0110 are suffixes of the string 0110. For a string of length  $n$ , there are  $n + 1$  number of suffixes.

- Proper suffix: For a string, any suffix of the string other than the string itself is called as the proper suffix of the string.

Example: For the string  $w = 0110$ , the proper suffixes are  $\lambda$ , 0, 10, and 110

□ Substring: A substring of a string is defined as a string formed by taking any number of symbols of the string.

Example: For the string  $w = 012$ , the substrings are  $\lambda, 0, 1, 2, 01, 12$ , and  $012$ .

## 1.2 Basics of Set Theory

A set is a well-defined collection of objects. The objects used for constructing a set are called elements or the members of the set.

A set has some features as follows:

- A set is a collection of objects. This collection is regarded as a single entity.
- A set is comprised of distinct elements. If an element, say  $a$ , is in set  $S$ , then it is denoted as  $a \in S$ .
- A set has a well-defined boundary. If  $S$  is a set and  $a$  is any element, then depending on the properties of  $a$ , it can be said whether  $a \in S$  or  $a \notin S$ .
- A set is characterized by its property. In general, if  $p$  is the defined property for the elements of  $S$ , then  $S$  is denoted as  $S = \{a : a \text{ has the property } p\}$ .

Example:

- The set of all integers is denoted as  $S = \{a : a \text{ is an integer}\}$ .

Here,  $7 \in S$  but  $1/7 \notin S$ .

- The set of all odd numbers denoted as  $S = \{a : a \text{ is not divisible by } 2\}$ .

Here,  $7 \in S$  but  $8 \notin S$ .

- The set of prime numbers less than 100 is denoted as  $S = \{a : a \text{ is prime and less than } 100\}$ .

Here,  $23 \in S$  but  $98$  or  $101 \notin S$ .

### 1.2.1 Subset

Let there be two sets  $S$  and  $S_1$ .  $S_1$  is said to be a subset of  $S$  if every element  $S_1$  is an element of  $S$ .

Symbolically, it is denoted as  $S_1 \text{ subset } S$ .

The reverse of a subset is the superset. In the previous example,  $S$  is the superset of  $S_1$ .

Example:

- Let  $Z$  be the set of all integers.  $E$  is the set of all even numbers. All even numbers are natural numbers. So, it can be denoted as  $E \text{ subset } Z$ .

- Let  $S$  be the set of the numbers divisible by 6, where  $T$  is the set of numbers divisible by 2. Property says that if a number is divisible by 6, it must be divisible by 2 and 3. So, it can be denoted as  $S \text{ subset } T$ .

### 1.2.2 Finite and Infinite Set

A set is said to be finite if it contains no element or a finite number of elements. Otherwise, it is an infinite set.

Example:

- Let  $S$  be the set of one digit integers greater than 1.  $S$  is finite as its number of elements is 8.

□ Let  $P$  be the set of all prime numbers.  $T$  is infinite, as the number of prime numbers is infinite.

### 1.2.3 Equal

Two sets  $S$  and  $S_1$  are said to be equal if  $S$  is a subset of  $S_1$  and  $S_1$  is a subset of  $S$ .

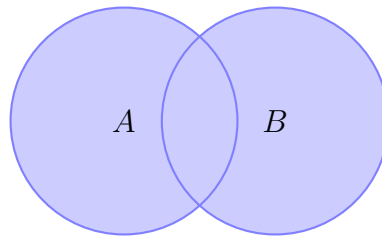
### 1.2.4 Algebraic Operations on Sets

□ Union: If there are two sets  $A$  and  $B$ , then their union is denoted by  $A \cup B$ .

Let  $A = \{2, 3, 4\}$  and  $B = \{3, 5, 6\}$ . Then,  $A \cup B = \{2, 3, 4, 5, 6\}$ . In general,  $A \cup B = \{x | x \in A \text{ or } x \in B\}$ .

Diagrammatically, the union operation on two sets can be represented as shown in Fig. 1.1. This diagrammatic representation of sets is called the Venn diagram.

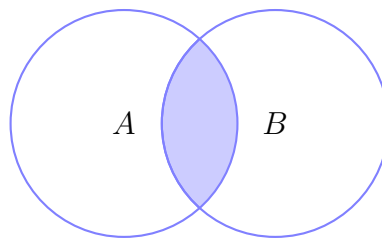
Fig.1.1.



□ Intersection: If there are two sets  $A$  and  $B$ , then their intersection is denoted by  $A \cap B$ . Let  $A = \{2, 3, 4\}$  and  $B = \{3, 4, 5, 6\}$ . Then,  $A \cap B = \{3, 4\}$ . In general,  $A \cap B = \{x | x \in A \text{ and } x \in B\}$ .

The Venn representation of the intersection operation on two sets can be represented as shown in Fig. 1.2.

Fig.1.2.

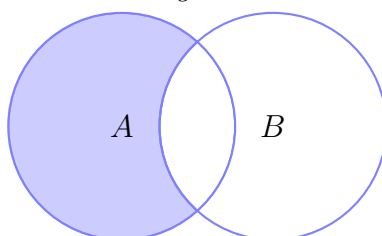


□ Difference: If there are two sets  $A$  and  $B$ , then their difference is denoted by  $AB$ . Let  $A = \{2, 3, 4, 5\}$  and  $B = \{3, 4\}$ . Then,  $AB = \{2, 5\}$ . In general,  $AB = \{x | x \in A \text{ and } x \notin B\}$ .

The Venn representation of the difference operation on two sets can be represented

as shown in Fig. 1.3.

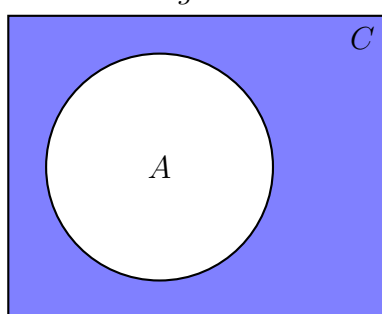
Fig.1.3.



□ Complementation: The complement of a set  $A$ , which is a subset of a large set  $U$  is denoted by  $A^c$  or  $A'$ , defined by  $A' = \{x \in U : x \notin A\}$ .

The Venn representation of the complement operation is given in Fig. 1.4.

Fig.1.4.



□ Cartesian product: If there are two sets  $A$  and  $B$ , then their Cartesian product is denoted by  $A \times B$ . Let  $A = \{2, 3, 4, 5\}$  and  $B = \{3, 4\}$ . Then,  $A \times B = \{(2, 3), (2, 4), (3, 3), (3, 4), (4, 3), (4, 4), (5, 3), (5, 4)\}$ . In general,  $A \times B = \{(a, b) | a \in A \text{ and } b \in B\}$ .

□ Power set: The power set of a set  $A$  is the set of all possible subsets of  $A$ . Let  $A = \{a, b\}$ . Then, the power set of  $A$  is  $\{(\emptyset), (a), (b), (a, b)\}$ . For a set of elements  $n$ , the number of elements of the power set of  $A$  is  $2^n$ .

□ Concatenation: If there are two sets  $A$  and  $B$ , then their concatenation is denoted by  $A.B$ . Let  $A = \{a, b\}$  and  $B = \{c, d\}$ . Then,  $A.B = \{ac, ad, bc, bd\}$ . In general,  $A.B = \{ab | a \text{ is in } A \text{ and } b \text{ is in } B\}$ .

### 1.2.5 Properties Related to Basic Operation

Some properties related to basic operations on set are as follows:

□  $A \cup \emptyset = A$ ,  $A \cap \emptyset = \emptyset$  ( $\emptyset$  is called null set)

□  $A \cup U = U$ ,  $A \cap U = A$  (where  $A \subset U$ )

□  $A \cup A = A$ ,  $A \cap A = A$  (idempotent law)