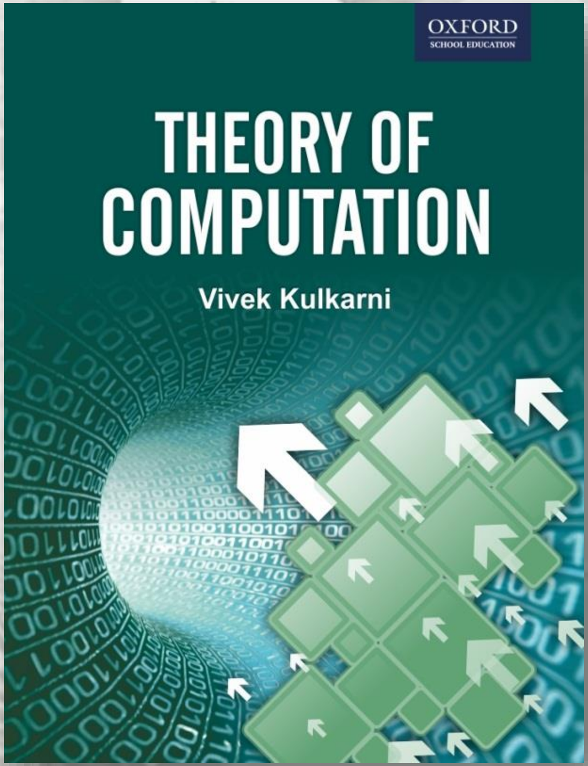
OXFORD
HIGHER EDUCATION

THEORY OF COMPUTATION

Vivek Kulkarni

Slides for Faculty Assistance



Chapter 1

Preliminaries

Author: Vivek Kulkarni

vivek_Kulkarni@yahoo.com

Outline

- Following topics are covered in the slides:
  - Basic concepts, namely, symbols, alphabets , strings, and language
  - Preliminaries related to set theory, relations, and set operations
  - Closure properties of relations
  - Basics of graph theory and graph properties
  - Formal verses natural languages
  - Introduction to mathematical induction

Symbol, Alphabet and Strings

- A **symbol** is an abstract or user-defined entity. For example, letters, digits, or any other characters that one wishes to consider as a part of the language that is being designed, are said to be symbols.

- An **alphabet** is a finite set of symbols. It is denoted by Σ. For example,
  - D = {0, 1, 2, …, 9}
  - X = {+, –, *, /, %}

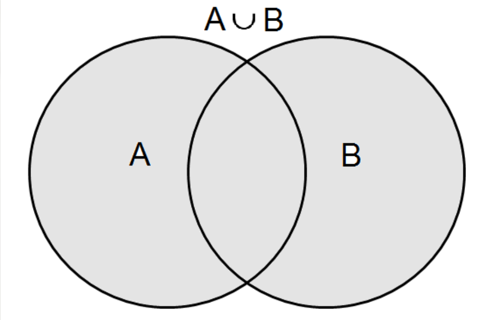
- **String** (or word) is defined as a finite sequence of symbols over a given alphabet. All the symbols of a string should come from the same alphabet set.

Sets

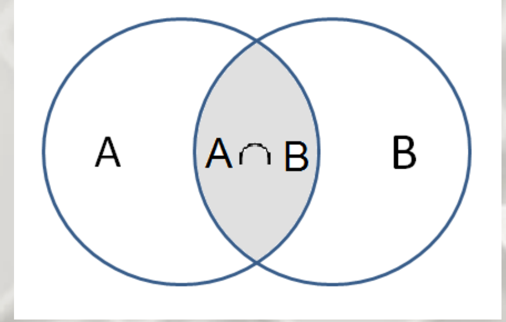
- A **set** is defined as a collection of well-defined and distinct objects. These objects, or entities, are called the *members* (or *elements*) of the set. For example, consider a set A such that, $A = \{1, 2, 3\}$. Here, 1, 2, and 3 are members of set A.
- Sets can be finite or infinite.
- Any two sets A and B, are considered equivalent if and only if they have precisely the same elements.
- **Cardinality** of a set is defined as the number of elements in the set. If A is any set, then its cardinality is denoted as '$| A |$'.

Set Operations

- The **union** of two sets is defined as: $A \cup B = \{x \mid x \in A$, or $x \in B)$
  - For example: If $A = \{1, 2, 3\}$, and
    B = {1, 3, 4, 6}, then,
    $A \cup B = \{1, 2, 3, 4, 6\}$.



- The **intersection** of two sets is defined as: $A \cap B = \{x \mid x \in A$ and $x \in B)$
  - For example: If $A = \{1, 2, 3\}$, and
    $B = \{1, 3, 4, 6\}$, then,
    $A \cap B = \{1, 3\}$.



Set Operations continued …

- The **difference** of two sets is defined as: $\underline{A} - \underline{B} = \{\underline{x} \mid \underline{x} \in \underline{A} \text{ and } \underline{x} \notin \underline{B}\}$,

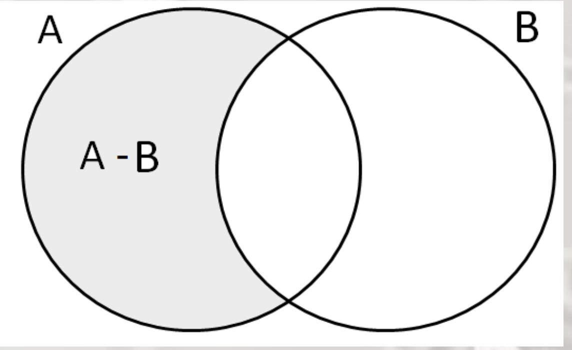
  or, $\underline{A} - \underline{B} = \underline{A} - (\underline{A} \cap \underline{B})$

  - For example:

    If $A = \{1, 2, 3, 7, 9\}$, and

    $B = \{1, 3, 4, 6\}$, then,

    $A - B = \{2, 7, 9\}$.



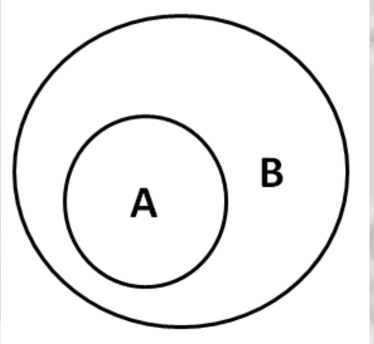
- The **Cartesian product** of two sets is defined as:

  $A \times B = \{(a, b) \mid a \in A \text{ and } b \in B, \forall a \,\&\, \forall b\}$

  - For example: $\{a, b\} \times \{a, b\} = \{(a, a), (a, b), (b, a), (b, b)\}$

Set Operations continued …

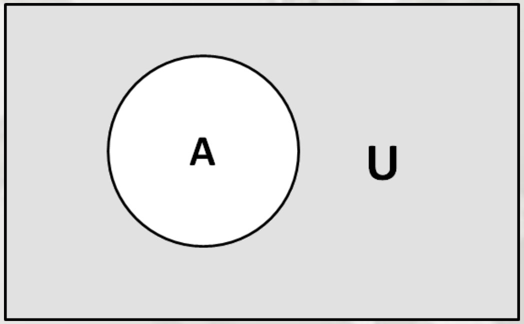
- If every member of set A is a member of set B, then set A is said to be a ***subset*** of set B. We write this as: $A \subseteq B$. Here, set B is said to be the ***superset*** of set A.
  - For example: $\{1, 4\} \subseteq \{1, 2, 3, 4, 5\}$
  - The empty set ϕ is a subset of every set.
  - Every set is a subset of itself.



- The **power set** of a set A is the set of all subsets of A, including itself, and the empty set, ϕ. It is denoted by 2^A.
  - For example, if $A = \{0, 1, 2\}$, then, $2^A = \{\phi, \{0\}, \{1\}, \{2\}, \{0, 1\}, \{0, 2\}, \{1, 2\}, \{0, 1, 2\}\}$

Set Operations continued …

- A set that encompasses all possible sets that can exist is called a ***Universal Set***, and is denoted by U. The **complement** of any set A is defined as: $A' = U - A$



- **Concatenation** of two sets A and B, is defined as: $A \cdot B = AB = \{x \mid x = ab, \forall a \in A \text{ and } \forall b \in B\}$. This means that every string from set A is concatenated with each string in set B.
  - For example, if $A = \{000, 111\}$, and $B = \{101, 010\}$, then, $AB = \{000101, 000010, 111101, 111010\}$
- **Closure** of a set is defined as: $S^* = S^0 \cup S^1 \cup S^2 \ldots$, where, $S^0 = \{\in\}$, and, $S^i = S^{i-1} \cdot S$; for $i > 0$. Closure of a set is thus a repetitive concatenation of the set to itself.
  - For example: If $S = \{01, 11\}$, then, $S^* = S^0 \cup S^1 \cup S^2 \cup S^3 \cup \ldots = \{\in, 01, 11, 0101, 0111, 1101, 1111, 010101, 010111, \ldots\}$

Countable and Uncountable Sets

- Countability is the property which signifies the existence of a successor. For instance, given any integer i, one can always find its successor '$i + 1$'.

- Finite sets are always countable. Likewise, infinite sets that can be placed in one-to-one correspondence with the set of natural numbers , $N = \{1, 2, 3, 4, 5, \ldots\}$, are said to be *countably infinite*, or just *countable*, or *enumerable*.

- Some infinite sets are *uncountable*. For example, let us consider the set of real numbers R: One cannot find the successor for any given real number. This is because, between any two real numbers there are infinite number of other real numbers. Hence the set 'R' is an infinite set, which is uncountable.

Relations

- A **relation** is a set of ordered pairs (or tuples), where the first component of the pair is from the set called the *domain*, and the second component is from the set called the *range* (or *co-domain*).

- A ***binary relation*** can be defined as follows:

  ${}_AR_B = \{(a, b) \mid a \in A \text{ and } b \in B\}$, where, set A is the domain set, and set B is the range set.
- Every relation is a subset of the Cartesian product of domain and range sets: ${}_AR_B \subseteq (A \times B)$

- A relation may encompass four types of associations: 'one-to-one', 'one-to-many', 'many-to-one', and 'many-to-many'.

Properties of Relations

- If R is a relation on set S (domain and range is the same set S), then it is said to be:
  - ***Reflexive***, if ${}_aR_a$ exists for all a in S.
  - ***Transitive***, if ${}_aR_b$ and ${}_bR_c$ imply ${}_aR_c$, for all a, b, and c in S.
  - ***Symmetric***, if ${}_aR_b$ implies ${}_bR_a$, for all a and b in S.
  - ***Anti-symmetric***, if ${}_aR_b$ does not imply ${}_bR_a$, for all a and b in S.

- If a relation is reflexive, transitive, as well as symmetric, then it is said to be an ***equivalence** relation*. If a relation is reflexive, transitive, and anti-symmetric, then it is said to be a ***partial ordering** relation*.

Closure Properties of Relations

- The **transitive closure** of a relation R, which is denoted by R^+, is defined as follows:
  - If $(a, b) \in R$, then (a, b) is in R^+
  - If $(a, b) \in R^+$ and $(b, c) \in R^+$, then (a, c) is in R^+
  - For example, let $S = \{1, 2, 3\}$, and R is a relation on S, such that $R = \{(1, 2), (2, 2), (2, 3)\}$; then, $R^+ = \{(1, 2), (2, 2), (2, 3), (1, 3)\}$

- **Reflexive and transitive closure** of a relation R, which is denoted by R^*, is defined as, $R^* = R^+ \cup \{(a, a) \mid \forall a \in S\}$, where R is a relation defined over set S.
  - For example, for the above set S and relation R:

    $R^* = \{(1, 1), (1, 2), (1, 3), (2, 2), (2, 3), (3, 3)\}$

- **Symmetric closure** of a relation R is defined as, If $(a, b) \in R$ then (a, b) and (b, a) are in the symmetric closure of R. Thus, symmetric closure of $R = \underline{R} \cup \{(b, a) \mid (a, b) \in R\}$. In other words, symmetric closure of R is the union of R with its inverse relation, R^{-1}.
  - For example, let us consider relation, $R = \{(1, 2), (2, 2), (2, 3)\}$ over set $S = \{1, 2, 3\}$, then, symmetric closure of $R = \{(1, 2), (2, 2), (2, 3), (2, 1), (3, 2)\}$.

Graph

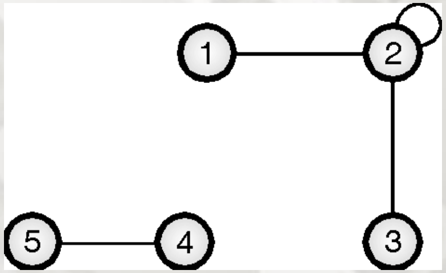
- A **graph** is formally defined by a tuple:

  $G = (V, E)$ where,

  V = Finite set of vertices or nodes, and

  $E = \{(v_1, v_2) \mid v_1, v_2 \in V\}$, i.e.,

  finite set of edges connecting the vertices.

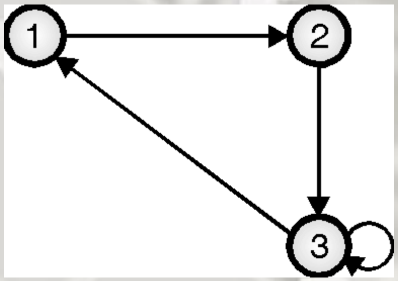


- Two vertices are said to be adjacent if they are connected by an edge.
- A **digraph** (or, directed graph) is denoted by:

  $G = (V, E)$, where,

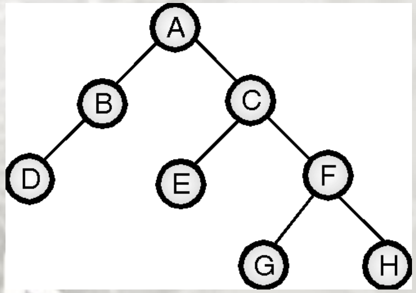
  V: Finite set of vertices, and

  E: Finite set of *ordered pairs* of vertices called *arcs*.



Graph

- A **Tree** is a digraph with the following properties:
  - There exists one vertex called the *root* vertex that does not have a predecessor and from which, there is a path to every other vertex in the graph.
  - Each vertex other than the root has exactly one predecessor; the immediate predecessor of a node is called the *parent node*.
  - The successors of each vertex are ordered from the left; the immediate successor of a node is called the *child node*.



- A graph can be defined as a relation over a set of vertices. It is not merely a diagram, but a visualization of the underlying relation.

Language

- A language is defined as a set of strings comprising symbols from one alphabet.
- Note that the null set ϕ and the set consisting of empty string, i.e., $\{\in\}$, are also considered as languages.
- The set of all strings over a fixed alphabet Σ is a language, and is denoted by Σ^*.
- For example, let $\Sigma = \{a\}$; then, $\Sigma^* = \{\in, a, aa, aaa, \ldots\}$. Similarly, let $\Sigma = \{0, 1\}$; then, $\Sigma^* = \{\in, 0, 1, 00, 01, 10, 11, 000, \ldots\}$
- Note that $\Sigma^* = \Sigma^{**}$
- All programming languages are **formal languages**. The term *formal* used here emphasises that the *form* of the strings of symbols has more importance than anything else.

Mathematical Induction

- **Principle of mathematical induction:**

  Let $S(n)$ denote the statement to be proved involving variable n, and let us suppose:
  - $S(1)$ is true;
  - If $S(k)$ is true for $n = k$, and ' $S(k + 1)$' is also true,

  then, $S(n)$ is true for all values of n.

- The following steps are involved in inductive proof;
  - **Induction basis**: This step tests if the statement, $S(n)$ holds true when n is equal to its lowest possible value. Usually, $n = 0$, or $n = 1$.
  - **Induction hypothesis** (or inductive hypothesis): In this step, it is assumed that $S(n)$ is true for some value of n, i.e., for $n = k$.
  - **Inductive step:** This step tests if the statement also holds when $n = k + 1$. If the step is true for $n = k + 1$, then it is true for all values of n.